HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS

by

Foo Say Wei
March 1983

A thesis submitted for the degree of Doctor of Philosophy
in the Faculty of Engineering, University of London

Department of Electrical Engineering
Imperial College of Science and Technology
University of London

To my wife and my parents,
for their love and confidence in me.

ABSTRACT

A new method of coding speech signals is introduced and its application to the low-bit-rate transmission of speech is examined. The new method, which is referred to as hybrid frequency-domain coding (HFDC), is based on a combination of sub-band coding and transform coding techniques and thus involves block coding.

In one version of the HFDC scheme, a fixed data rate is used. In this version, blocks of speech signal are coded using a fixed number of bits. The technique is of relatively low complexity and results of subjective tests indicate that the quality of speech produced by the technique at 7.2 kbit/s is comparable with that produced by 19 kbit/s continuously variable slope delta modulation (CVSD); and that at 9.6 kbit/s, it is comparable with 23 kbit/s CVSD.

In a further application of the HFDC scheme, a variable bit rate technique is used which takes account of the nonstationary nature of speech. Block coding is again used but the number of bits allocated is made to depend on the characteristics of the block. Subjective tests have shown that improved speech quality can be obtained at the expense of slightly increased end-to-end delay.

Quadrature mirror filters are important in sub-band coding and a new method of designing these filters is described. The effects of finite-word-length on the performance of systems using quadrature mirror filters are considered and formulae are derived relating to the bounds and variances of the errors due to analog-to-digital conversion; quantization of the filter coefficients; roundoff in arithmetic computations and the quantization associated with compression encoding. The results obtained from extensive simulations are shown to be in excellent agreement with the theoretically derived values.

Various other issues relating to the design of the HFDC scheme are considered. These include the analysis of errors resulting from the use of the block-floating-point fast Walsh-Hadamard transform, and the automatic classification of speech into voiced, unvoiced and silent intervals.

ACKNOWLEDGEMENTS

I wish first and foremost to express my gratitude to Professor Laurence F.Turner whom I am privileged to have had as supervisor. I am most impressed by his patience, the encouragement he gave me during the initial period of my research and the advice and critical discussions he made in the course of the preparation of the various papers and the thesis.

I am much obliged to Dr.Lynda V.White of the Department of Mathematics for her valuable advice on the design and analysis of experiments for subjective tests of speech coding schemes.

The various forms of assistance extended to me in connection with the research conducted in this thesis are too many to be exhaustively listed. In particular, I wish to thank Dr.Lim Yong Ching, Dr.Peter Kwan and Mr.Zhong Yi Xing for their friendships; Dr.Eleftherios Frangoulis and Dr.Abrahaam Alcaim for the stimulating discussions ; Dr.David Horn of Bell Laboratory and Dr.John N.Holmes of Joint Speech Research Unit for providing the recorded speech materials; Prof. Peter Noll of Technische Universitat Berlin, Germany, for providing information on bit allocation and optimization; Messrs.Ian J.Dedic, Farzin Deravi and Katsumi Yoshida for their assistance in the subjective tests and in general, all other members of the department for discussions and advice in one way or another.

Last but not least, I would like to express my thanks to Dr.Tay Eng Soon for his support and encouragement; and to the Ministry of Defence for granting the DSO scholarship and the paid study leave.

TABLE OF CONTENTS

TITLE PAGE 1

DEDICATION 2

ABSTRACT 3

ACKNOWLEDGEMENTS 4

TABLE OF CONTENTS 5

LIST OF FIGURES 10

LIST OF TABLES 17

CHAPTER 1: INTRODUCTION 19

1.1 Preliminary 20

1.2 Production and Perception of Speech and Characteristics of Speech Signals 21

1.2.1 Production of Speech 21

1.2.2 Perception of Speech 23

1.2.3 Characteristics of Speech Signals 25

1.3 Speech Coders and Limits of Speech Coding 28

1.3.1 Ultimate Limits of Bit-Rate Reduction for Speech Information 29

1.3.2 Waveform Coding Systems 31

1.3.3 Source Modelling Systems 32

1.3.4 Intermediate Systems 35

1.4 Organisation of the Thesis 37

CHAPTER 2: REVIEW OF RELATED WORKS 47

2.1 Introduction 48

2.2 Sub-band Coding of Speech Signals 48

2.3 Orthogonal Transforms 52

2.3.1 Walsh-Hadamard Transform (WHT) 53

2.3.2 Discrete Fourier Transform (DFT) 59

2.3.3 Discrete Cosine Transform (DCT) 60

2.4 Transform Coding 62

2.5 Variable-Bit-Rate Coding of Speech Signals 70

2.5.1 Extremal Coding 71

2.5.2 Time Encoding 72

2.5.3 Adaptive Aperture Coding 74

2.5.4 Variable-Rate ADPCM Coding 76

2.6 Motivation for Hybrid Frequency-Domain Coding 77

CHAPTER 3: FIXED-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS 91

3.1 Introduction 92

3.2 The Partitioning of the Spectrum and the Design of the Filter Banks 93

3.3 The Allocation of Bits to Sub-bands 95

3.4 The Selection of Transform 98

3.5 The Selection of Block Size 102

3.6 The Methods of Signal Quantization 102

3.6.1 Quantization of Signals in Bands B1 and B2 102

3.6.2 Quantization of Signals in Bands B3 and B4 104

3.6.3 Coding of Side-Information 106

3.6.4 Quantizer for Transform Coefficients 108

3.7 System Delay Consideration 108

3.8 Results of Subjective Tests 112

CHAPTER 4: VARIABLE-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS 134

4.1 Introduction 135

4.2 Theoretical Basis of Bit Allocation to Segments of Speech Signals 137

4.3 Methods of Variable-Rate Coding Based on Signal Segmentation 139

4.3.1 A Super Block Processing Approach 139

4.3.2 A 'Limited-Type' Approach 140

4.4 A New Technique of Variable-Bit-Rate Coding 141

4.4.1 General Comments 141

4.4.2 A Control Strategy for Variable-Bit-Rate Coding 143

4.4.3 Buffer Requirements 145

4.4.4 Selection of the Parameters 146

4.5 Application of the Variable-Bit-Rate Technique to Hybrid Frequency-Domain Coding 148

4.6 Simulation of a Variable-Bit-Rate Hybrid Frequency-Domain Coder and Results of Subjective Tests 149

4.6.1 Bit Allocation 150

4.6.2 Method of Quantization 151

4.6.3 Results of Subjective Tests 152

4.7 Conclusions 153

CHAPTER 5: DESIGN OF NON-RECURSIVE QUADRATURE MIRROR FILTERS 161

5.1 Introduction 162

5.2 Theory of Quadrature Mirror Filters 163

5.3 Problem Formulation 168

5.4 Quadrature Mirror Filters of Two and Four Taps 172

5.5 Quadrature Mirror Filters of Larger Number of Taps 174

5.6 Conclusions 178

CHAPTER 6: EFFECTS OF FINITE-WORD-LENGTH ON THE PERFORMANCE OF SPLIT-BAND SYSTEMS USING QUADRATURE MIRROR FILTERS 209

6.1 Introduction 210

6.2 Number Representation and Noise Model Used 211

6.3 Structure and Coefficient Scaling of the Filters 213

6.4 Effects of Quantization and Finite-Word-Length in Systems Using Quadrature Mirror Filters 215

6.4.1 Effects of Quantization of the Input Signal 215

6.4.2 Effects of Quantization of Filter Coefficients 217

6.4.3 Error Bounds and Statistics Associated with Finite-Word-Length Coefficients 219

6.4.4 Effects of Arithmetic Roundoff 223

6.4.5 Analysis of Errors Arising from Arithmetic Roundoff 224

6.4.6 Effects of Combined Coefficient Quantization and Arithmetic Roundoff 226

6.4.7 Analysis of Errors Associated with Coefficient Quantization and Arithmetic Roundoff 227

6.4.8 Effects of Quantization for Transmission Bit Rate Reduction 229

6.4.9 Analysis of Errors Associated with Quantization for Transmission Bit Rate Reduction 230

6.5 Simulation Results 231

6.6 Conclusions 234

CHAPTER 7: ERROR ANALYSIS OF BLOCK-FLOATING-POINT FAST WALSH-HADAMARD TRANSFORMS 245

7.1 Introduction 246

7.2 Error Analysis 248

7.3 Simulation Results 253

7.4 Conclusions 255

CHAPTER 8: APPLICATION OF SUB-BAND ENERGY RATIO TO VOICED-UNVOICED-SILENCE CLASSIFICATION OF SPEECH SIGNALS 262

8.1 Introduction 263

8.2 Description of the Algorithm 265

8.3 Determination of Thresholds 266

8.4 Experiments and Results 267

8.5 Computational Load 270

8.6 Conclusions 271

CHAPTER 9: CONCLUSIONS 278

9.1 Conclusions 279

9.2 Suggestions for Further Research 283

REFERENCES 294

APPENDIX A PARAMETERS USED IN THE CVSD SCHEME 307

APPENDIX B SUBJECTIVE TESTS OF FIXED-RATE HFDC 311

APPENDIX C SUBJECTIVE TESTS OF VARIABLE-RATE HFDC 328

APPENDIX D NON-RECURSIVE FILTERS OF EVEN ORDER 341

APPENDIX E THE ANALYSIS OF THE ERROR $\mathcal{E}_{cr}$ 343

APPENDIX F EXPRESSION OF $W^2(\omega)$ IN CLOSED FORM 348

APPENDIX G CHARACTERISTICS OF QUADRATURE MIRROR FILTERS USED IN THE SIMULATION 350

LIST OF FIGURES

Fig.1.1	Cross-sectional view of the human vocal mechanism.	40
Fig.1.2	Schematic diagram of the human auditory mechanism.	41
Fig.1.3	Amplitude density of speech signals and theoretical Gamma and Laplace probability densities. (After Paez and Glisson [2].)	42
Fig.1.4	Autocorrelation functions of speech signals. (After Noll [3].)	43
Fig.1.5	Long-time power density spectrum for continuous speech. (After Dunn and White [4].)	43
Fig.1.6	Waveform and spectrum of a typical unvoiced segment of speech signal.	44
Fig.1.7	Waveform and spectrum of a typical voiced segment of speech signal.	45
Fig.1.8	Block diagram of a simple speech production model.	46
Fig.2.1	Block diagram of a general sub-band coding system.	80
Fig.2.2	Partition of speech spectrum into bands.	81
Fig.2.3	Illustration of integer band sampling technique for one sub-band; (a) Block diagram; (b) Amplitude spectra.	82
Fig.2.4	Frequency domain illustration of the movable upper sub-bands.	83
Fig.2.5	Signal flow graph of fast Hadamard-ordered Walsh-Hadamard transform.	84
Fig.2.6	Illustration of block reversal.	85
Fig.2.7	Signal flow graph of fast Walsh-ordered Walsh-Hadamard transform.	86
Fig.2.8	Illustration of coding and estimation of spectral envelope.	87
Fig.2.9	Variation of noise spectrum with parameter γ.	88

Fig.2.10 Illustration of extremal coding; (a) Original waveform; (b) Interpolated waveform. 89

Fig.2.11 Illustration of aperture coding. 90

Fig.2.12 Aperture with (a) first-order prediction: $a_1=1$; (b) more general prediction. 90

Fig.3.1(a) Block diagram of the encoder for hybrid frequency-domain coding scheme. 116

Fig.3.1(b) Block diagram of the decoder for hybrid frequency-domain coding scheme. 117

Fig.3.2 Frequency domain illustration of partition of speech spectrum. 118

Fig.3.3(a) Band-splitting at the transmitter. 119

Fig.3.3(b) Band-combining at the receiver. 120

Fig.3.4 Effects of different transforms on SSNR of Band 3 signals. (a) Side-information unquantized; (b) Side-information quantized. 121

Fig.3.5 Effects of different transforms on SSNR of Band 4 signals. (a) Side-information unquantized; (b) Side-information quantized. 122

Fig.3.6(a) Variation of r.m.s. energy within segments of a sub-band (B3) signal of speech. 123

Fig.3.6 (b-d) Comparisons of $SSNR_m$'s of the sub-band signal coded using the DCT and those of the signal coded using the WHT. Block length=40 ms. Number of samples per block=64. Total number of bits per block allocated for quantization of the transform coefficients=80. (b) using exact standard deviation; (c) using unquantized side-information with 5 group values; (d) using qunatized side-information with 5 group values. 124

Fig.3.7(a) Variation of r.m.s. energy within segments of a sub-band (B4) signal of speech. 125

Fig.3.7 (b-d) Comparisons of $SSNR_m$'s of the sub-band signal coded using the DCT and those of the signal coded using the WHT. Block length=40 ms. Number of samples per block=128. Total number of bits per block allocated for quantization of the transform coefficients=160. (b) using exact standard deviation; (c) using unquantized side-information with 5 group values; (d) using qunatized side-information with 5 group values. 126

Fig.3.8 Effects of different coding methods on SSNR of full-band signals. (a) Side-information unquantized; (b) Side-information quantized. 127

Fig.3.9 Illustration of the effects of transform block size on SSNR; (a) Band 3; (b) Band 4. 128

Fig.3.10 Illustration of coding and estimation of spectral envelope; (a) Original spectra; (b) Averages of groups; (c) Estimation by interpolation. 129

Fig.3.11 Distribution of ratios of group standard deviation to average block standard deviation; (a) Band 3; (b) Band 4. 130

Fig.3.12 Probability density functions for normalized transform values of signals; (a) Band 3; (b) Band 4. 131

Fig.3.13 Effects of types of quantizer on SSNR; (a) Band 3; (b) Band 4. 132

Fig.3.14 Relative subjective preference of quality of speech produced by HFDC and that produced by CVSD at different bit rates; (a) Experiment 1 (HFDC at 7.2 kbit/s); (b) Experiment 2 (HFDC at 9.6 kbit/s). 133

Fig.4.1 Distributions of energy within segments of voiced speech, unvoiced speech and silence. Segment length=10 ms. 155

Fig.4.2 Variation of r.m.s. values of segments of the telephone conversation. 156

Fig.4.3 Variation of r.m.s. values of segments of the radio announcement. 157

Fig.4.4 Distribution of energy within segments of the telephone conversation. Segment length=10 ms. 158

Fig.4.5 Distribution of energy within segments of the radio announcement. Segment length=10 ms. 159

Fig.4.6 Schematic diagram of the variable-rate hybrid frequency-domain coder; (a) Encoder; (b) Decoder. 160

Fig.5.1 Illustration of two-band splitting using quadrature mirror filters. 180

Fig.5.2 Frequency response of 4-point FIR lowpass filter. 181

Fig.5.3 Augmented impulse response. 182

Fig.5.4 Maximum peak-to-peak deviation of the overall response as a function of cutoff frequency expressed in percentage of transition width from $f_s/4$. (solution B, N=16, K=1). 183

Fig.5.5 Stopband attenuation as a function of cutoff frequency expressed in percentage of transition width from $f_s/4$. (solution B, N=16, K=1). 184

Fig.5.6 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution A, N=16, K=1). 185

Fig.5.7 Stopband attenuation for various transition widths. (solution A, N=16, K=1). 186

Fig.5.8 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N=16, TW=5/32). 187

Fig.5.9 Stopband attenuation for various values of K. (solution B, N=16, TW=5/32). 188

Fig.5.10 Maximum peak-to-peak deviation of the overall response for various values of K. (solution A, N=16, TW=5/32). 189

Fig.5.11 Stopband attenuation for various values of K. (solution A, N=16, TW=5/32). 190

Fig.5.12 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=16, K=1, TW=5/32). 191

Fig.5.13 Comparison of stopband attenuation for solution A and solution B. (N=16, K=1, TW=5/32). 192

Fig.5.14 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution B, N=8, K=1). 193

Fig.5.15 Stopband attenuation for various transition widths. (solution B, N=8, K=1). 194

Fig.5.16 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N=8, TW=5/16). 195

Fig.5.17 Stopband attenuation for various values of K. (solution B, N=8, TW=5/16). 196

Fig.5.18 Maximum peak-to-peak deviation of the overall response for various values of K. (solution A, N=8, TW=5/16). 197

Fig.5.19 Stopband attenuation for various values of K. (solution A, N=8, TW=5/16). 198

Fig.5.20 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=8, K=1, TW=5/16). 199

Fig.5.21 Comparison of stopband attenuation for solution A and solution B. (N=8, K=1, TW=5/16). 200

Fig.5.22 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution B, N=32, K=1). 201

Fig.5.23 Stopband attenuation for various transition widths. (solution B, N=32, K=1). 202

Fig.5.24 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N=32, TW=5/64). 203

Fig.5.25 Stopband attenuation for various values of K. (solution B, N=32, TW=5/64). 204

Fig.5.26 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=32, K=1, TW=5/64). 205

Fig.5.27 Comparison of stopband attenuation for solution A and solution B. (N=32, K=1, TW=5/64). 206

Fig.5.28 Frequency response of a 32-point FIR lowpass filter. 207

Fig.5.29 Frequency response of the overall system based on filters with lowpass frequency response shown in Fig.5.28. 208

Fig.6.1 Block diagram of a shared multiplier structure. 235

Fig.6.2 Variation of the weighting function with angular frequency; (a) N=4; (b) N=32. 236

Fig.6.3 Error variance $\sigma^2_{cr}(\omega)$ of the frequency response of systems due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32. 237

Fig.6.4 Illustration of deviation of peak-to-peak ripple from nominal value due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32. 239

Fig.6.5 Error variances of output due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32. 240

Fig.6.6 Error variances of output due to arithmetic roundoff; (a) N=8; (b) N=16; (c) N=32. 242

Fig.6.7 Error variances of output due to coefficient quantization and arithmetic roundoff; (a) N=8; (b) N=16; (c) N=32. 243

Fig.6.8 Error variances of output due to quantization for bit rate reduction. 244

Fig.7.1 Experimental error results : random numbers from a uniform distribution; -1<random numbers<1; word-length = 16 bits. 256

Fig.7.2 Experimental error results: random numbers from a uniform distribution; 0≤random numbers<1; word-length = 16 bits. 257

Fig.7.3 Experimental error results: random numbers from a Gaussian distribution N(0.,0.25); -1<random numbers<1; word-length = 16 bits. 258

Fig.7.4 Normalized errors of block-floating-point WHT as a function of block size using data from a segment of speech signals. Word-length = 10 bits. 259

Fig.7.5 Comparisons of errors resulting from block-floating-point WHT and those from block-floating-point FFT using random numbers from uniform distribution as input. 260

Fig.7.6 Comparisons of errors resulting from block-floating-point WHT and those from block-floating-point FFT using random numbers from Gaussian distribution as input. 261

Fig.8.1 Waveform of a segment of speech signal. 272

Fig.8.2 Variation of the energy content within segments of sub-band speech signals. Segment length = 5 ms. 274

Fig.8.3	Distribution of logarithm of energy ratio. Segment length = 10 ms. (a) Test signal 1; (b) Test signal 2.	275
Fig.8.4	Distribution of logarithm of energy within segments of speech signals. Segment length = 10 ms. (a) Test signal 1; (b) Test signal 2.	276
Fig.9.1	Waveform of a voiced segment of speech signal.	288
Fig.9.2	Absolute values of the Walsh-Hadamard transform coefficients.	289
Fig.9.3	Absolute values of the cal components of the Walsh-Hadamard transform.	290
Fig.9.4	Absolute values of the sal components of the Walsh-Hadamard transform.	291
Fig.9.5	Power spectrum of the Walsh-ordered WHT of the voiced segment shown in Fig.9.1.	292
Fig.9.6	Power spectrum of the discrete Fourier transform of the voiced segment shown in Fig.9.1.	293
Fig.A.1	Block diagram of CVSD coder.	310
Fig.B.1	Scatter plots for residuals; (a) Experiment 1; (b) Experiment 2.	325
Fig.B.2	Normal plots for residuals; (a) Experiment 1; (b) Experiment 2.	326
Fig.B.3	Subjective preference of quality of speech produced by CVSD against that produced by HFDC with the 90% confidence limits imposed; (a) Experiment 1; (b) Experiment 2.	327
Fig.C.1	Scatter plots for residuals; (a) Experiment 1; (b) Experiment 2.	339
Fig.C.2	Normal plots for residuals; (a) Experiment 1; (b) Experiment 2.	340

LIST OF TABLES

Table 1.1	Summary of Speech Coding Techniques	39
Table 3.1	Partition of Speech Band	95
Table 3.2	Multiplication Factor for Step-Size Adaptation	104
Table 3.3	Bit Allocation for HFDC (bit/s)	113
Table 3.4	Comparison of HFDC at 7.2 kbit/s and CVSD	114
Table 3.5	Comparison of HFDC at 9.6 kbit/s and CVSD	115
Table 4.1	Percentage of Segments Coded as Type-3	148
Table 4.2	Allocation of Bits for Average Rate of 7.2 kbit/s	150
Table 4.3	Allocation of Bits for Average Rate of 9.6 kbit/s	150
Table 4.4	Percentage Preference Score for Experiment 1	153
Table 4.5	Percentage Preference Score for Experiment 2	153
Table 8.1	Confusion Matrices for Test Signal 1 (SNR=40 dB)	269
Table 8.2	Confusion Matrices for Test Signal 2 (SNR=25 dB)	269
Table B.1	Arrangements of Sentences on Tapes	313
Table B.2	Results of Experiment 1	314
Table B.3	Results of Experiment 2	315
Table B.4	Transformed Data of Experiment 1	316
Table B.5	Transformed Data of Experiment 2	317
Table B.6	Analysis of Variance Table	317
Table B.7	ANOVA Table for Experiment 1	318
Table B.8	ANOVA Table for Experiment 2	319
Table B.9	Residuals for Experiment 1	321
Table B.10	Residuals for Experiment 2	321

Table B.11 Percentage Preference Score for CVSD (Experiment 1: HFDC at 7.2 kbit/s) 324

Table B.12 Percentage Preference Score for CVSD (Experiment 2: HFDC at 9.6 kbit/s) 324

Table C.1 Arrangements of Sentences on Tapes 330

Table C.2 Results of Experiment 1 332

Table C.3 Results of Experiment 2 332

Table C.4 Transformed Data for Experiment 1 333

Table C.5 Transformed Data for Experiment 2 333

Table C.6 ANOVA Table for Experiment 1 334

Table C.7 ANOVA Table for Experiment 2 334

Table C.8 Residuals for Experiment 1 335

Table C.9 Residuals for Experiment 2 336

Table C.10 Percentage Preference Score for Experiment 1 337

Table C.11 Percentage Preference Score for Experiment 2 338

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 PRELIMINARY

There are many advantages in representing speech signals in digital rather than analog form. Firstly, whereas an analog waveform becomes increasingly distorted when transmitted over long distance, a digital representation remains undistorted, provided that binary decision can be made accurately. Secondly, the advent of very large scale integration (VLSI) promises low-cost implementation of even complex systems based on digital techniques. Thirdly, the coding of speech signals in digital format allows the integration of telegraphy and other digital data within one network. Fourthly, the storage of speech in digital form permits fast and easy access with modern day computing systems. Fifthly, the encryption of digitally coded speech is easier and offers a higher level of security than can be obtained with analog methods. Furthermore, the performance of analog equipment tends to be much more temperature and humidity dependent than digital equipment.

Ideally, it is desirable to digitize speech signals using as few bits as are necessary to preserve the desired quality. There are many incentives for encoding speech signals at low bit rates. One advantage of using a reduced number of bits is that data transmission rates can

be reduced, with the effect that either speech signals can be transmitted using a reduced channel bandwidth, or an increased number of speech signals can be multiplexed and transmitted over a channel of large capacity. A second advantage is that for applications such as voice storage and response systems, the storage requirement can be reduced. Besides, if only very few bits are required for the representation of salient features of a speech utterance, then speech recognition techniques using template matching may be simplified.

In the search for new and better ways of encoding speech signals, an understanding of the relevant features of the signals and the modes of production are essential. Thus, in Section 1.2, certain aspects of human speech production and perception as well as the characteristics of speech signals that are most relevant to speech coding are examined. A panoramic view of the present state of the art of speech coding and some estimates of the limits of bit-rate reduction achievable are given in Section 1.3. This is followed in Section 1.4 by a description of the organization of the thesis.

1.2 PRODUCTION AND PERCEPTION OF SPEECH AND CHARACTERISTICS OF SPEECH SIGNALS

1.2.1 Production of Speech

A cross-sectional view of the human vocal organs is given in Fig.1.1. Speech is formed from a succession of speech sounds called phonemes according to a set of linguistic rules and speech sounds are produced when air is forced from the lungs through the bronchi, trachea, the vocal cords, the vocal tract and the nostril/mouth. The

basic mechanisms of excitation depend on the speech sounds to be generated.

During the generation of voiced sounds such as /a/, /e/, /i/, /o/, /u/, air is pushed through the vocal cords, forcing the vocal cords into vibration. The rate of vibration depends on the length, thickness and tension of the vocal cords. As a result, quasi-periodic pulses of air are produced which are modulated by resonances in the vocal tract. Different sounds are produced for different configurations of the vocal tract. The shape of the vocal tract varies as a function of time through the movements of the articulators: the lips, the jaw, the tongue and the velum. Voiced nasal sounds such as /n/ and /m/ in 'run' and 'rum' respectively are generated when the velum opens so that the nasal cavity forms part of the acoustic tube. Thus the signals of voiced sounds are quasi-periodic, and they contain the pitch frequency together with its many harmonics.

Fricatives, or unvoiced sounds, such as /f/ in 'fish' and /sh/ in 'she', are generated when the air is forced at a high velocity through the constriction created by the articulators somewhere along the vocal tract. This generates a broad-spectrum noise source which is used to excite the vocal tract. Thus the signals of unvoiced sounds are generally noise-like.

Voiced fricatives such as /v/ in 'veal' are produced with both vocal cord vibration and constriction being present at the same time.

Plosive sounds such as /p/ in 'pop' are produced by building up pressure in the mouth and abruptly releasing the air.

An understanding of the mechanism of speech production enables the modelling of such a system and facilitates the estimation of essential parameters of the model from the received sounds. Knowledge of auditory perception, on the other hand, provides a guide to those attributes that are important in human speech perception.

1.2.2 Perception of Speech

Generally, the receiving end of a speech communication system is a human being. Human perception results from the processing of the sound pressure wave impinging on the eardrum. A 'cut-out' view of the human auditory organs is shown in Fig.1.2. It is generally established that the external ear serves to focus sound onto the eardrum; the middle ear mainly carries out the impedance transformation between the air medium of the outer ear and the liquid medium of the inner ear; and the inner ear processes the sound transmitted through the middle ear. However, little is known about how the sounds are perceived and processed by the brain.

Although in general the listeners are more interested in extracting the message (words or sentences) from the speech sounds, the emotional contents in the forms of pitch, pitch variation, loudness, stress pattern and duration of phonemes can also affect the semantics and pragmatics of the utterance.

Numerous experiments have been carried out in the area of human perception of speech sounds and some inferences from these experiments are described in the following paragraphs.

It has been established [1] that the selectivity of the ears analysis varies with frequency, and that the spectral structure of complex sounds such as speech cannot be resolved in a bandwidth less than certain critical bands. Experiments have shown [1] that the widths of the critical bands up to about 1 kHz are approximately 100 Hz and above 1 kHz, they are about one sixth of the center frequencies of the respective bands.

A model has been put forward that the inner ear be regarded as an approximation to a very large number of extensively overlapping broad bandpass filters of bandwidths equal to the critical bands. A 'hair-cell', which has a highly non-linear response, is connected to the output of each filter. The hair cells code the filter outputs in the form of nerve pulses which are transmitted to the higher auditory centres for further processing.

Experience has shown that group-delay distortion with a speech signal is difficult to perceive if it varies by a small amount within one critical bandwidth. Human auditory systems are accustomed to perceiving speech in the presence of time-dispersion as a result of reverberation in the normal speaking environment. It is thus not surprising that much larger random group delay distortion is not detected as being unnatural.

As a result of the ears' filtering action, low-level sounds in one region of the spectrum are masked by higher-level sounds in other regions. Considerable masking of high frequency sounds by low frequency sounds also occurs outside of the critical bandwidth. This

explains the insensitivity of listeners to low-level background noise, when the wanted signal is of higher level.

Experiments also indicate that the noise in the formant regions is masked by the speech signal itself. Thus, efficient coding schemes may exploit this feature by shaping the noise spectrum in such a way that spectral components of the noise in the formant regions are allowed to have higher power than those components in the interformant regions in addition to allowing larger spectral noise power for the high frequency end of the speech spectrum.

1.2.3 Characteristics of Speech Signals

For efficient signal processing, certain specific features of speech signals have to be exploited. Three aspects of speech signals that are particularly important to speech coding will now be described. These are: the amplitude, the spectrum and the variation of the signal as a function of time.

AMPLITUDE

(a) Distributions

By analysing the amplitudes of speech signals over a long duration of time, Paez and Glisson [2] have shown that the probability density function of speech amplitudes can be very well described as a Gamma distribution of the form

$$p(x) = \left[\frac{\sqrt{3}}{8\pi\sigma_x|x|}\right]^{1/2} \exp\left(\frac{-\sqrt{3}|x|}{2\sigma_x}\right) \qquad (1.1)$$

A somewhat simpler approximation is the Laplace density function

$$p(x) = \frac{1}{\sqrt{2}\sigma_x}\exp\left(\frac{-\sqrt{2}|x|}{\sigma_x}\right) \qquad (1.2)$$

where $|x|$ denotes the absolute value of the signal and σ_x^2 its variance.

The probability distribution of the amplitudes of speech signals as obtained by Paez and Glisson [2] is shown in Fig.1.3 together with the Gamma and Laplace probability densities. From this figure, it can be seen that there is a very high proportion of zero and near zero amplitudes and a significant proportion of very high amplitudes. It is also clear that both the Gamma density and the Laplace density functions are reasonably close to the actual distribution, although the Gamma density gives a closer approximation than the Laplace density.

(b) Autocorrelation function

An estimate of the long-time average autocorrelation function of samples of speech signals given by Noll [3] is shown in Fig.1.4. From this figure, it can be seen that there exists a high degree of correlation between adjacent samples. However, the correlation decreases rapidly for greater spacing. Also, for the same spacing, the correlation is higher for lowpass filtered speech than for highpass filtered speech.

Unlike the long-time autocorrelation where the autocorrelation function decreases monotonically with spacing, the short-time autocorrelation function shows greater variation of values with

spacing [4]. In addition, the average short-time autocorrelation for voiced speech is very much higher than that for unvoiced speech [4]. As voiced sounds are quasi-periodic, the autocorrelation function computed from short-time voiced segments is also quasi-periodic.

SPECTRUM

The long-term power density spectrum of speech illustrated in Fig.1.5 was obtained by Dunn and White [5] by averaging the power output of a set of bandpass filters using about one minute of continuous speech. As can be seen from the figure, the peak is around 250-500 Hz and the spectrum rolls off at approximately 8-10 dB per octave beyond 500 Hz. Although the long-term statistics show that the high-frequency components contribute very little to the total speech energy, the short-time spectra of certain unvoiced segments of speech signal, an example of which is shown in Fig.1.6, indicate that the spectral density of the high-frequency components is in fact higher than that of the low-frequency components.

Examination of the spectral envelopes of short-time spectra of voiced segments, of which Fig.1.7 is a typical example, show not only that the power of low-frequency components are larger but also that local peaks exist. These peaks are due to resonances in the vocal tract and correspond essentially to the formants. Also, in general, the width of the formant at the low end of the speech spectrum is narrower than that at the high end.

For voiced speech which is quasi-periodic, its frequency spectrum also shows that most of the energy is contained in narrow

lobes spaced at the harmonics of the pitch frequency. These lobes are often referred to [6] as 'spectral fine structure'. The spectral fine structure is dependent mainly on the pitch period and the degree of voiced periodicity. Various techniques to exploit this quasi-periodic feature of voiced speech for speech coding have been investigated, a notable example being the Harmonic compression technique [7-10].

VARIATION IN TIME

Continuous speech consists of concatenation of voiced sounds, unvoiced sounds and pauses. Statistics have shown that the energy contents of voiced sounds are very much higher than those of unvoiced sounds. In a noise free environment, the energy contents of pauses or silent intervals are zero. Thus the power of speech signals varies considerably over time. In general, voiced speech statistics are relatively stationary over pitch periods. For longer duration, speech signals are nonstationary. These properties can be profitably exploited for efficient speech coding and is in fact investigated as part of the work reported on in this thesis.

1.3 SPEECH CODERS AND LIMITS OF SPEECH CODING

The rapid development in the technology of large scale integration in the recent years has given tremenduous impetus to research in digital speech coding. Developments in integrated circuits have meant that increasingly sophisticated coding schemes have become possible and economically viable. However, in any field of research, it is desirable to know whether the work being carried out is leading in the right direction and whether the aims are attainable or are

beyond naturally imposed limits. Thus, a discussion of the theoretical limits imposed by natural law on the digital coding of speech signals is given in the following section, followed by a brief summary of the various coding techniques.

1.3.1 Ultimate Limits of Bit-Rate Reduction for Speech Information

In the case of speech coding, the starting point in an investigation of limits of bit-rate reduction is the structure of the language. Most languages have 30 to 50 phonemes and the English language actually contains 42 phonemes. If silence or pauses are also counted, there are 43 symbols to be coded for the English language. With equal code length for each symbol, 6 bits per symbol is required. On account of the physical limitations of the human vocal mechanism, natural speech produced by human contains on the average about 10 phonemes per second. Thus, 60 bit/s are required to code the phonemic content of speech.

Dewey has carried out statistical studies relating to the relative frequency of the occurrence of the 42 phonemes [11,12], and using Shannon's mathematical theory of information [13], the entropy of speech has been shown to be 4.9 bit/phoneme. Thus it is theoretically possible in terms of phonemes, to transmit speech in the English language at a rate of 49 bit/s. However, the above calculations are based entirely on the phonemic content of speech. Information relating to the emotional content of speech and information about other charateristics of speech and the speaker, which are manifested through aspects such as pitch, harmonic content, pitch variation, loudness, stress variation etc., have not been taken

into consideration. The coding and transmission of this information requires additional bits. If it is assumed that 1000 combinations for all phonemes are sufficient for this purpose, then an additional 10log2(1000) = 99.6 bits are required. Thus, as a rough estimate, essential information on speech sounds can be compressed into about 150 bit/s [14].

Sophisticated coders which analyse phonemes, pitch, stress etc. are very complicated and generally resort to artificial intelligence techniques such as pattern recognition and synthesis by rule. In situations where relatively high bit rates can be used and where simple, good quality, robust and cheap coders are required, other techniques are desirable. These other techniques can be broadly classified under the headings of waveform coding, source modelling (vocoding) and intermediate (hybrid) methods. A summary of the quality and complexity of a range of coders is given in Table 1.1. This table is compiled based on the review by Flanagan et al [15] and the survey by Holmes [16]. In general, a waveform preservation approach is the least complicated and the reproduced speech is of high quality. However, bit rates in excess of 16 kbit/s are required to achieve such result. On the other hand, vocoding, which is based on an estimation of the parameters of the speech signals in accordance with a speech production model, is more complicated than waveform coding methods. Besides, the reproduced speech has a synthetic quality. These schemes do, however, have the advantage that the speech signals can be coded at bit rates on the order of 2 kbit/s. Hybrid methods are based on a combination of the vocoding and waveform coding techniques. By combining the advantages of vocoding and waveform coding, it is

possible to achieve intermediate performance at bit rates in the range of 4 to 16 kbit/s.

1.3.2 Waveform Coding Systems

Waveform coding methods make use of the signals in sampled-data form. If all the samples are to be coded and if frequency aliasing is to be avoided, then the minimum sampling rate required is determined by the well-known sampling theorem [17] which states that if a function f(t) contains no frequencies higher than B Hz, it is completely determined by giving its ordinates at a series of points spaced 1/2B seconds apart.

Linear pulse-code-modulation (PCM) [18] is the most basic waveform coding technique. In this scheme, samples are simply linearly quantized with no attempt being made to exploit the correlations among samples. For speech signals sampled at a rate equal to or greater than the Shannon sampling rate, adjacent samples are highly correlated. By encoding the prediction error signal d(n) which is the difference between the unquantized input sample x(n) and the estimate, $\hat{x}(n)$, of the input sample and which is in general smaller than x(n), a gain equal to σ_x^2/σ_d^2 in signal-to-noise ratio (SNR) can be obtained over that obtained by simply coding the amplitudes of x(n). In the above expression, σ_x^2 and σ_d^2 are the variances of x(n) and d(n) respectively. This form of speech coding is often referred to [19,20] as differential pulse-code-modulation (DPCM).

By using additional preceeding samples to predict the amplitude of the sample, the difference between the actual value and the

predicted value can be made smaller and hence an increased gain in SNR can be achieved [21]. The prediction can be based on the spectral envelope using a linear prediction model together with the spectral fine structure using a pitch model. As speech is a nonstationary process, the prediction model should be made adaptive with time if more efficient coding is to be achieved. This form of coder is referred to [22] as adaptive predictive coder. With increasing order of prediction, the difference between the actual and predicted samples becomes smaller and the SNR gain also increases [23].

Delta modulation [24] is a special case of DPCM where only one bit is used to indicate whether the difference is positive or negative. A sampling rate in excess of the Shannon rate is usually chosen so that the adjacent samples are highly correlated. By adaptively modifying the step-size, significant improvement can be achieved. Many forms of adaptive delta-modulation (ADM) are reported in the literatures [25-29], of which continuously variable slope delta modulation (CVSD) is perhaps one of the most popular.

1.3.3 Source Modelling Systems

Of the speech coders which are based on the source modelling technique, three principal types in particular have attracted considerable attention. They are known as channel vocoders, linear prediction coders and formant vocoders. The difference among them lie essentially in the way they model and extract the spectral envelope of speech sounds.

For each of these coding systems, a model of the voice production system is established, and the various parameters of the model are estimated from the speech signals. The set of important parameters are then coded in an efficient way. The speech is regenerated with the voiced production model using these sets of parameters. In this way, a very substantial reduction in bit rate can be achieved.

With most of the vocoders, linear separability of the sound source model and the vocal tract model is assumed and a distinction is made between the voiced and unvoiced sounds. Voiced sounds are modelled as being generated by a quasi-periodic train of pulses at pitch intervals which are used to drive a transfer function modelling the shape of the vocal tract; whereas for unvoiced sounds, the sound source is modelled as a pseudorandom noise sequence.

The essential information extracted by the vocoders from a short segment of speech signals is the voiced/unvoiced indication, together with the parameters describing the spectral envelope, and the pitch frequency when the segment is classified as voiced speech. The block diagram of a simple speech production model is shown in Fig.1.8.

Clearly, the quality of the reproduced speech depends very much on the goodness of fit of the model and the precision with which the various parameters such as the average power, the model parameters and the pitch are estimated and coded. The ability to classify sounds accurately as voiced or unvoiced also has a bearing on the quality of the reconstructed speech.

In linear prediction coders [30], the spectrum shaping is achieved by a recursive sampled-data filter which provides an all-pole transfer function. The filter coefficients are derived based on minimization of the mean-square error. The order of the filter is typically between 8 and 12, resulting in a corresponding 8 to 12 poles, either as complex conjugate pole pairs or real poles.

For channel vocoders [31-33], the speech signal is passed through a bank of typically 14 to 25 bandpass filters. The average short-time spectral power of each of the 14 to 25 contiguous spectral bands are sampled once every 20 ms and transmitted together with the information about pitch and the voiced/unvoiced decision. The spectral bands are determined based on auditory perception, and are not of equal bandwidth.

Formant vocoders [34-40] extract only the frequencies and amplitudes of a few, typically three, resonance peaks called formants. The bandwidths of the resonances are estimated based on some priori knowledge. A highly complicated analysis is required to estimate accurately the formant frequencies.

Numerous other vocoding techniques can be found in the literature [41-45].

A further reduction of the bit rate can be achieved by exploiting the residual correlation among the coefficients of the transfer function used for modelling the vocal tract [46,47]. Other methods that have been tried for achieving a further reduction in bit rate include: frame interpolation [48], variable-frame-rate coding [49], pattern matching [50,51] and vector quantization [52,53].

The quality of speech produced by vocoders designed so far tends generally to be synthetic and unnatural. Futhermore, the quality of the speech produced is speaker dependent and improvement in quality is not significant when higher bit rates are employed. Thus, even though vocoders can encode and produce highly intelligible speech at very low bit rates, the quality of speech reproduced is not considered to be suitable for commercial telephone communication.

1.3.4 Intermediate Systems

There is a very wide range of intermediate systems which attempt to fill the performance gap between the vocoding and waveform coding systems.

One such system consists of the voice excited vocoders [54,55]. In this system, the 'baseband' (typically 0 to 800 Hz) signal which covers the whole range of the fundamental frequency of speech, is coded using waveform coding. The upper band (800 to 4000 Hz) is coded using vocoder techniques. The excitation signal for the vocoder section can be obtained by an appropriate combination of nonlinear operation and spectral flattening on the decoded baseband signal.

As described in Section 1.2, the distribution of the average power spectral density of speech signals is non-uniform. Further, the bandwidth of the formant at the low frequency end of the spectrum is typically much narrower than that at the high frequency end. Thus, an efficient method of speech coding is to divide the speech signals into frequency components which are then separately coded: allocating more bits for those components that contribute more to the perceptual

quality of speech sounds. This form of coding is referred to [6] as frequency-domain coding.

Sub-band coding and transform coding are two of the frequency-domain coding techniques. In sub-band coding, the division is carried out by a bank of filters, usually 4 to 8, and the sub-band signals are coded using waveform coding techniques. The bandwidths of the sub-bands are wide relative to the pitch striations in the voiced speech spectrum.

In transform coding, on the other hand, the division into frequency components is effected by a suitable transform on blocks of speech signals. In order to capitalize on the pitch structure of the speech signals, as well as the formant structure, the number of frequency channels used in transform coding is much larger than in sub-band coding, and block sizes of 64 to 512 samples are used. Further details of sub-band coding and transform coding are given in Chapter 2.

Another intermediate system makes use of the harmonic scaling technique. In this technique, the local stationarity and the pitch structure of voiced speech signals are exploited. Through harmonic compression, the bandwidth and sampling rate of the speech signals can be reduced by a small integer number. At the receiver, the bandwidth and sampling rate are restored through harmonic expansion. It has been reported [7] that speech compressed and re-expanded using a ratio of 2:1 remains fully intelligible.

The harmonic compression (or expansion) can be realized either in the time domain by selectively eliminating (or inserting) pitch

periods [8] or in the frequency domain by reducing (or increasing) the spectral gap between the pitch harmonics [9,10].

As sub-band coding and transform coding do not specifically make use of the harmonic structure of speech signal, these techniques can be combined with harmonic scaling to achieve further bit-rate reduction [56].

1.4 ORGANISATION OF THE THESIS

The aim of the work reported on in this thesis was to investigate simple coding techniques capable of producing good quality speech in the range 7.2 kbit/s to 9.6 kbit/s.

Specifically, investigations into the combined use of sub-band coding and transform coding were carried out. These two techniques are known to be very good in their own rights and a discussion of sub-band coding and transform coding techniques, together with other work related to the thesis is presented in Chapter 2. The new method of coding involving the combined use of sub-band coding and transform coding is referred to as hybrid frequency-domain coding. A fixed-rate hybrid frequency-domain coding scheme is described in Chapter 3. Subjective tests were carried out on speech obtained with fixed-rate hybrid frequency-domain coding and the results are also contained in Chapter 3.

The nonstationary nature of speech is used to advantage in a variable-bit-rate technique based on segmentation of speech signals into three types. The technique and the results of its application to hybrid frequency-domain coding are discussed in Chapter 4.

Several issues relating to the implementation and performance of the schemes suggested are also examined in detail. One of the main components in the implementation of the hybrid frequency-domain coder, the quadrature mirror filters, is dealt with in depth in Chapters 5 and 6. Specifically, a new method of designing the filters is given in Chapter 5, and in Chapter 6, theoretical and experimental studies relating to the effects of finite-word-length on the performance of split-band systems using quadrature mirror filters are considered. In Chapter 7, the errors that arise when the fast Walsh-Hadamard transform is performed using a block-floating-point data representation are studied and a comparison is made between these errors and those arising from the fast Fourier transform. A new method for the machine classification of segments of speech signals into silence, voiced and unvoiced sounds is introduced in Chapter 8. Finally, the main results of the research are summarised in Chapter 9. Some suggestions for further work are also included in this chapter.

Table 1.1 SUMMARY OF SPEECH CODING TECHNIQUES

APPROACH	QUALITY	EXAMPLES	BIT RATE (kbit/s)	COMPLEXITY
Waveform Preservation	Toll	Log PCM	56	low
	Toll	Adaptive Delta Modulation	40	low
	Toll	Adaptive Differential PCM	32	low
	Toll	Pitch Predictive ADPCM	24	low
	Toll	Adaptive Predictive Coder	16	high
	Communications	Log PCM	36	low
	Communications	Adaptive Delta Modulation	24	low
	Communications	Adaptive Differential PCM	16	low
	Communications	Adaptive Predictive Coder	7.2	high
Modelling and Waveform Coding	Toll	Sub-band Coder	24	medium
	Toll	Adaptive Transform Coder	16	high
	Toll	Phase Vocoder	16	high
	Toll	Voice Excited Vocoder	16	high
	Communications	Sub-band Coder	9.6	medium
	Communications	Adaptive Transform Coder	7.2	high
	Communications	Phase Vocoder	7.2	high
	Communications	Voice Excited Vocoder	7.2	high
Modelling	Synthetic (Highly intelligible but unnatural)	Channel Vocoder	2.4	high
		Linear Prediction Coder	2.4	high
		LPC with Orthogonalized Coefficients	1.2	high
		Formant Vocoder	0.5	very high
Artificial Intelligence	Intelligible but highly unnatural	Phonetic Vocoder	0.1-0.2	very high

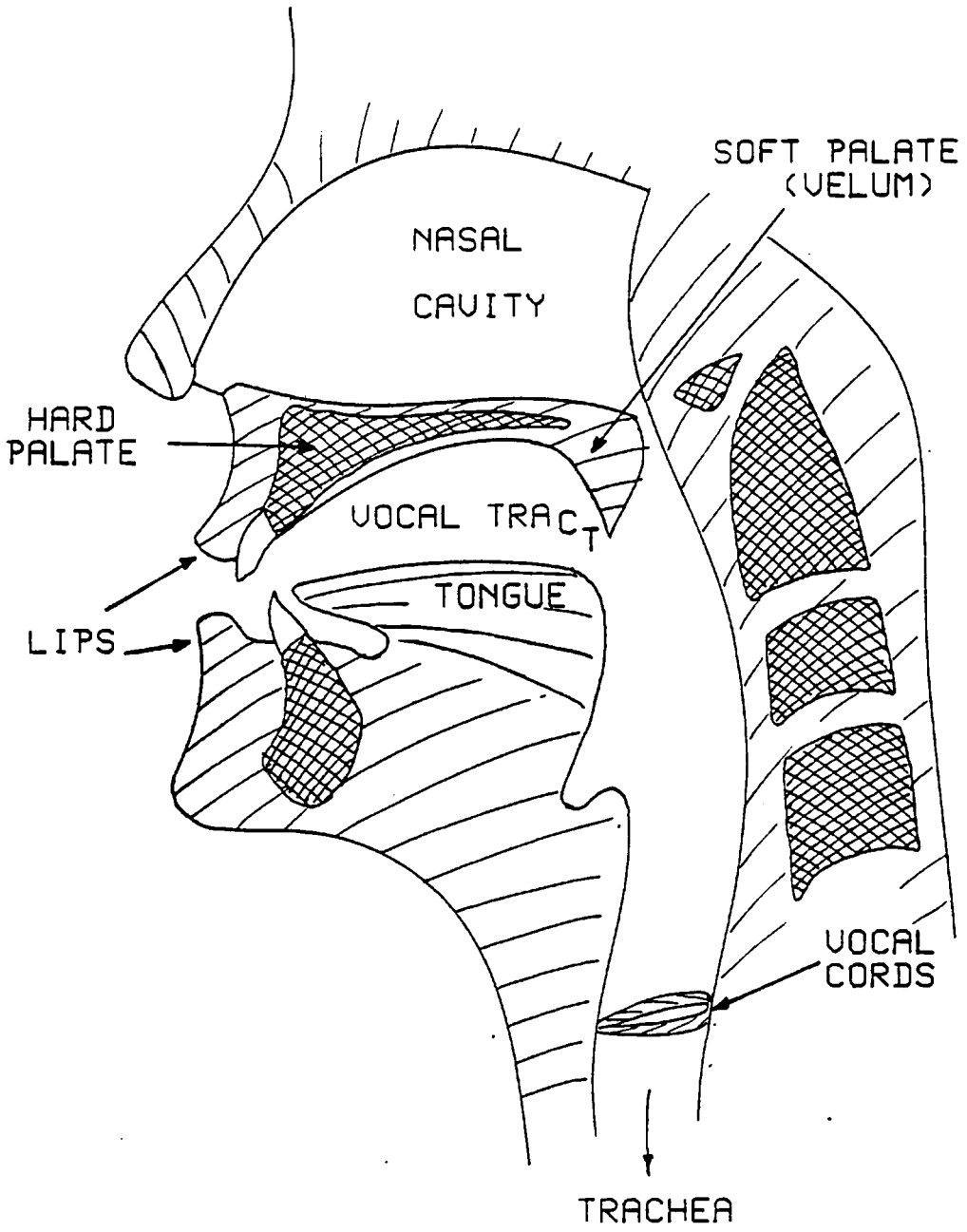


Fig.1.1 Cross-sectional view of the human vocal mechanism.

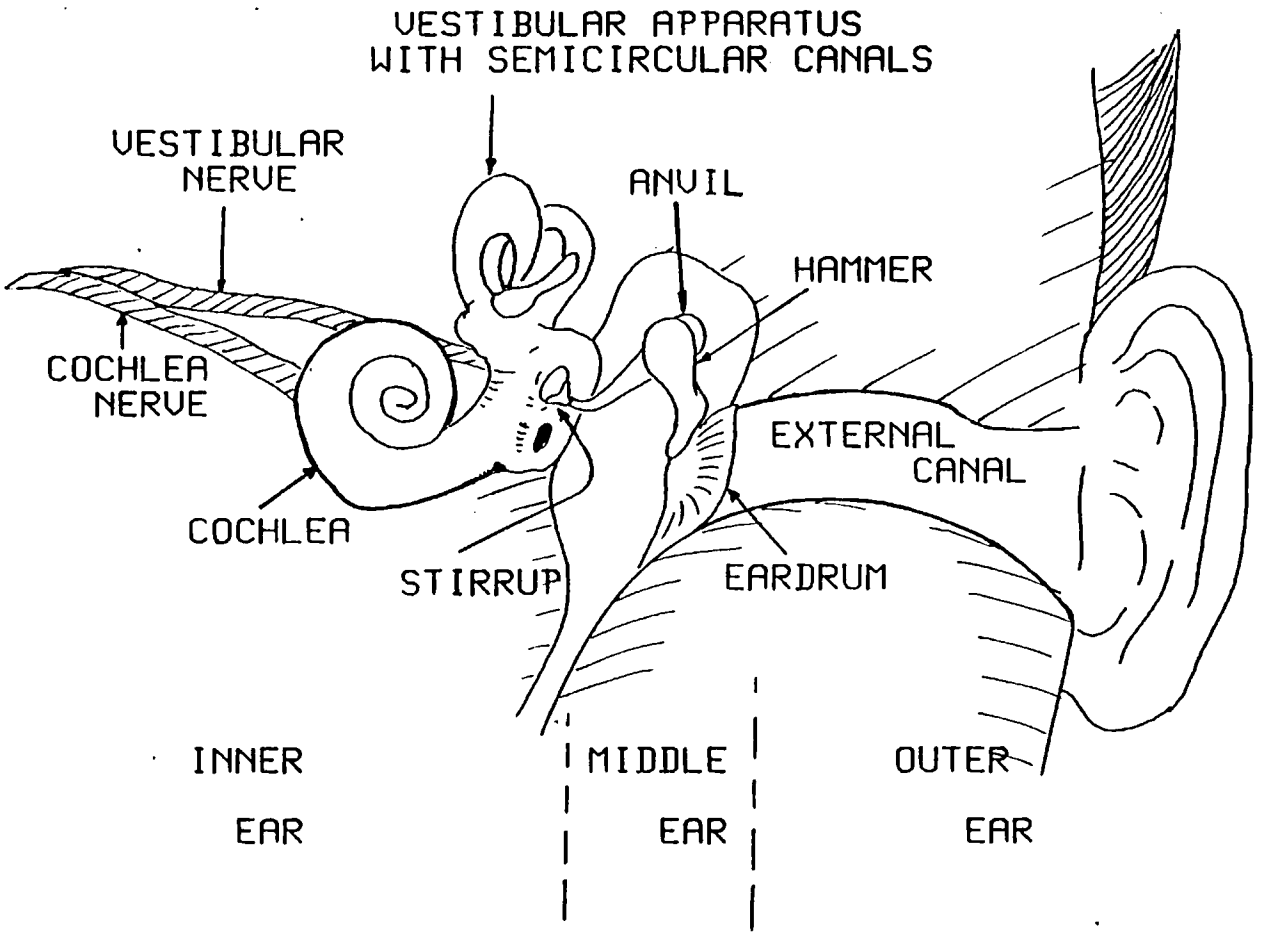


Fig.1.2 Schematic diagram of the human auditory mechanism.

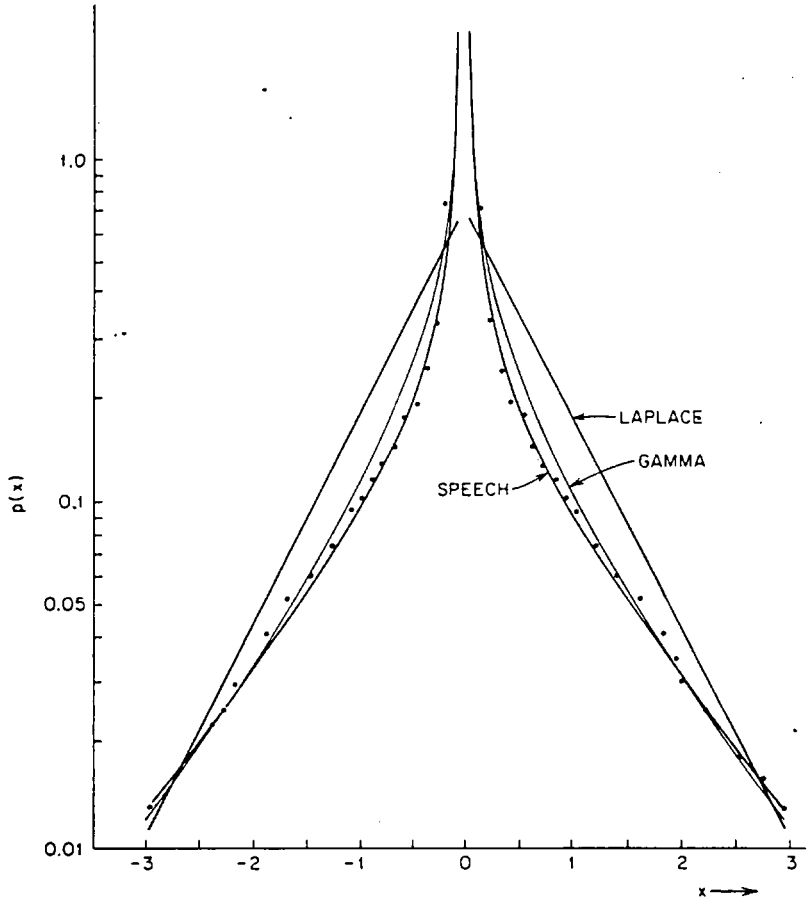


Fig.1.3 Amplitude density of speech signals and theoretical Gamma and Laplace probability densities. (After Paez and Glisson [2].)

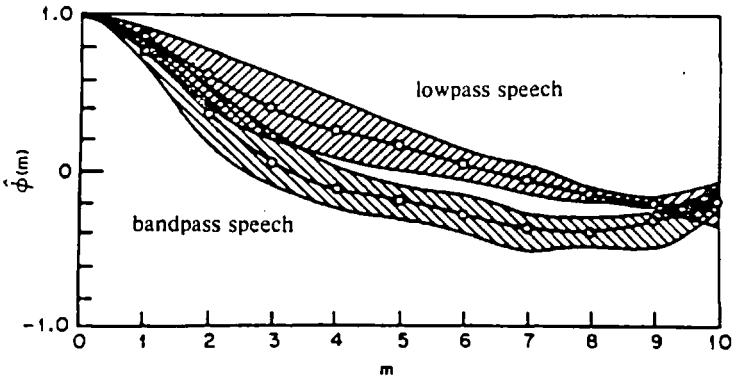


Fig.1.4 Autocorrelation functions of speech signals. (After Noll [3].)

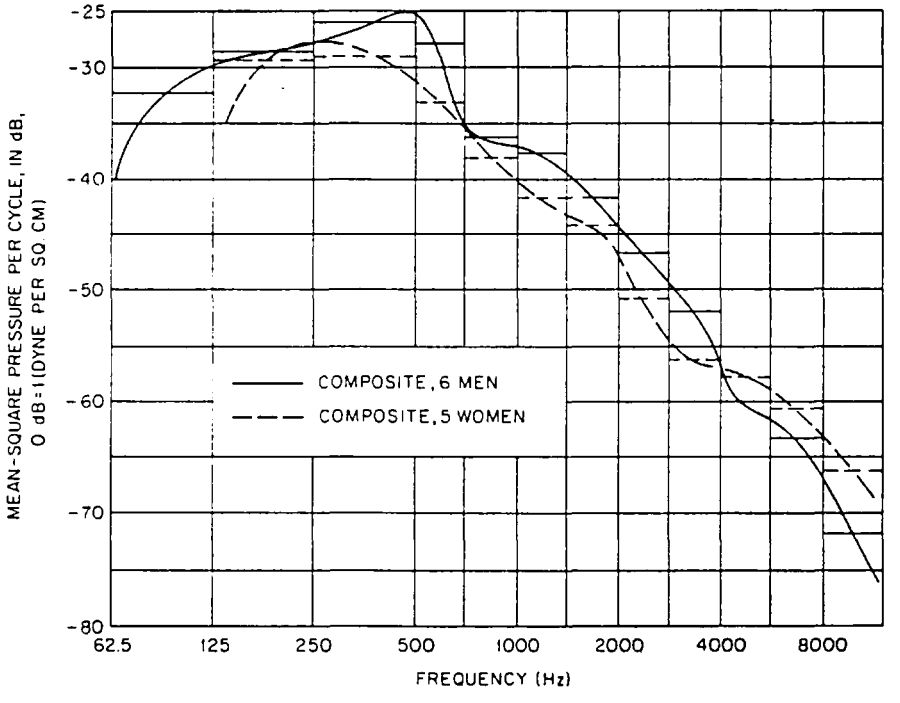


Fig.1.5 Long-time power density spectrum for continuous speech. (After Dunn and White [4].)

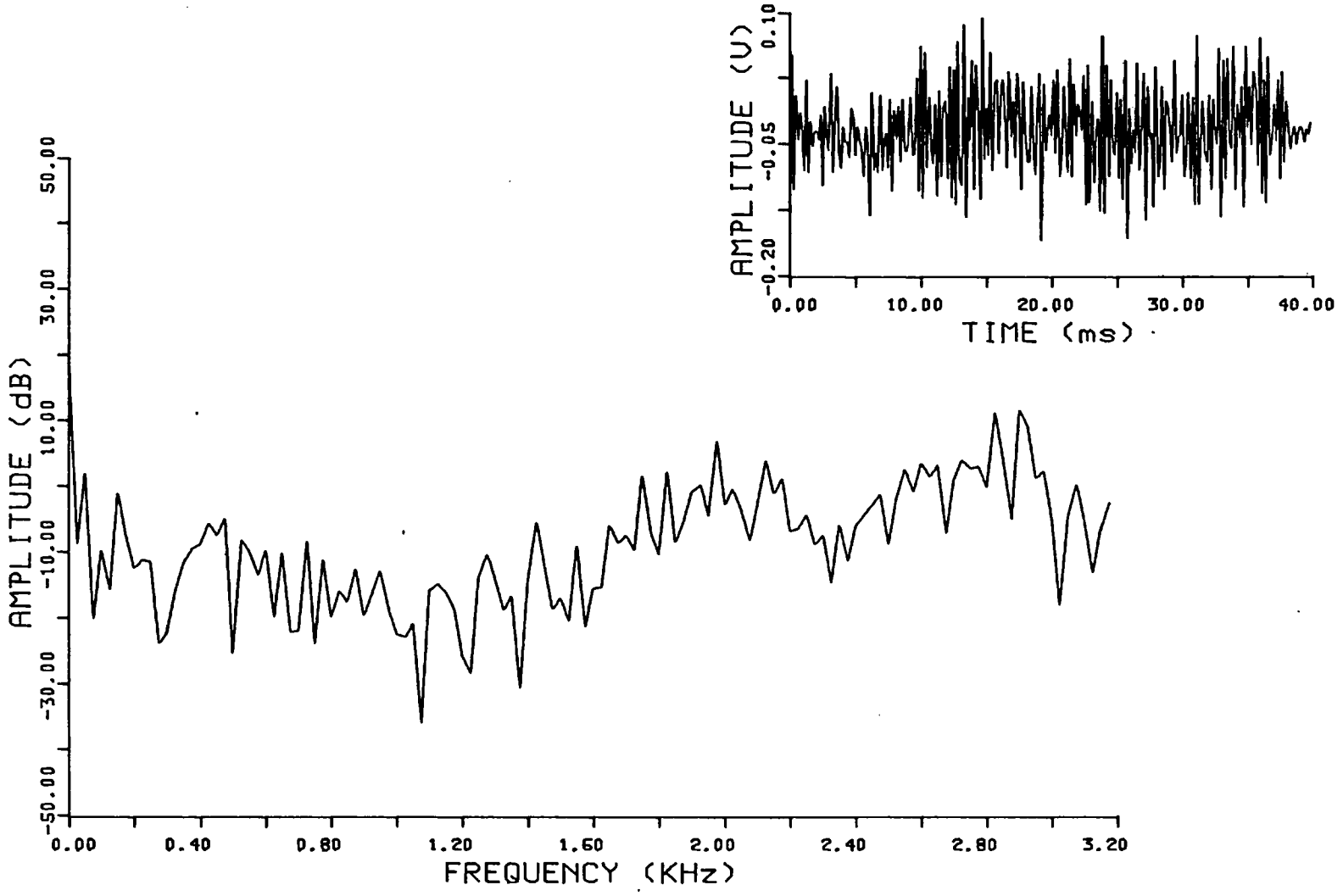


Fig.1.6 Waveform and spectrum of a typical unvoiced segment of speech signal.

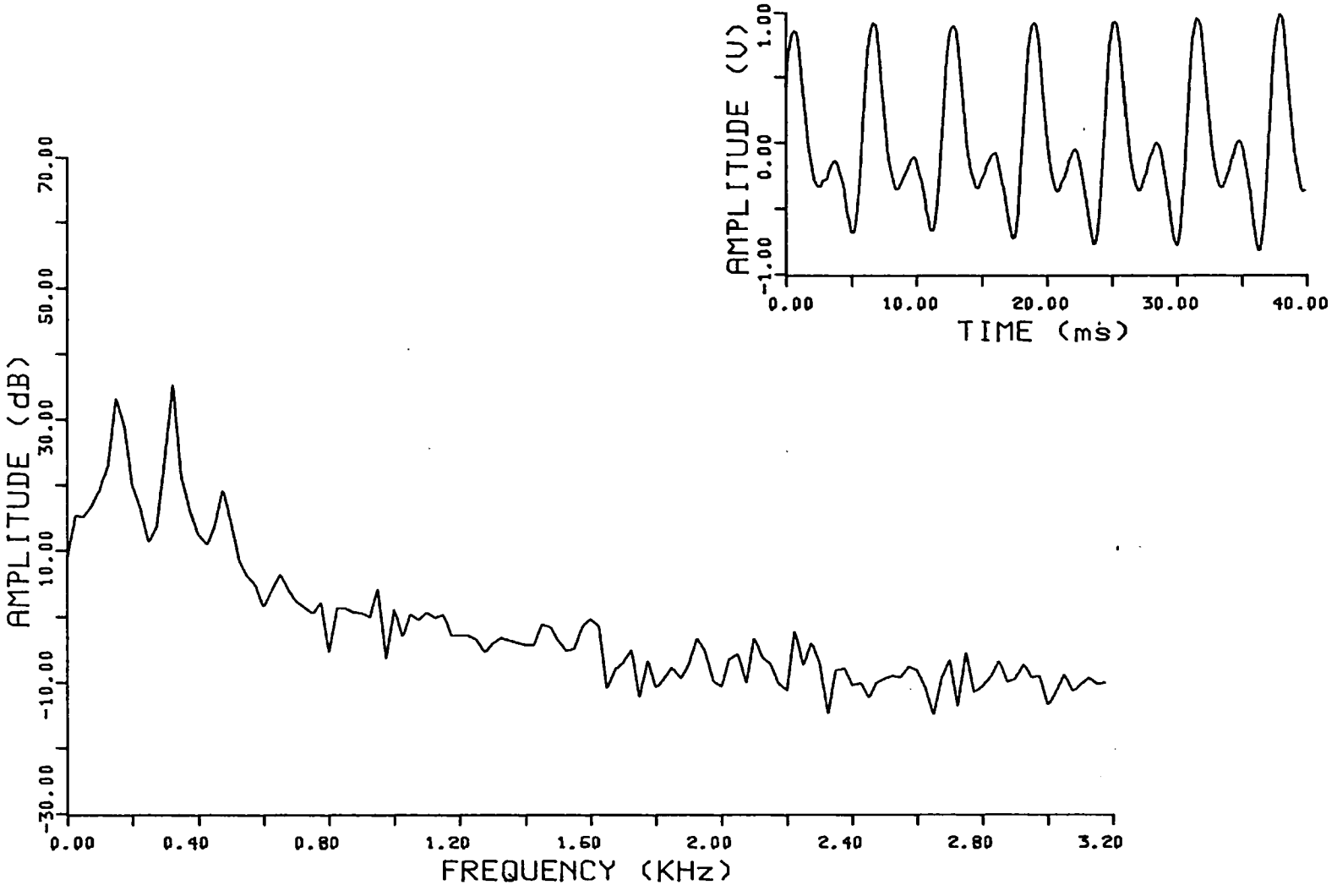


Fig.1.7 Waveform and spectrum of a typical voiced segment of speech signal.

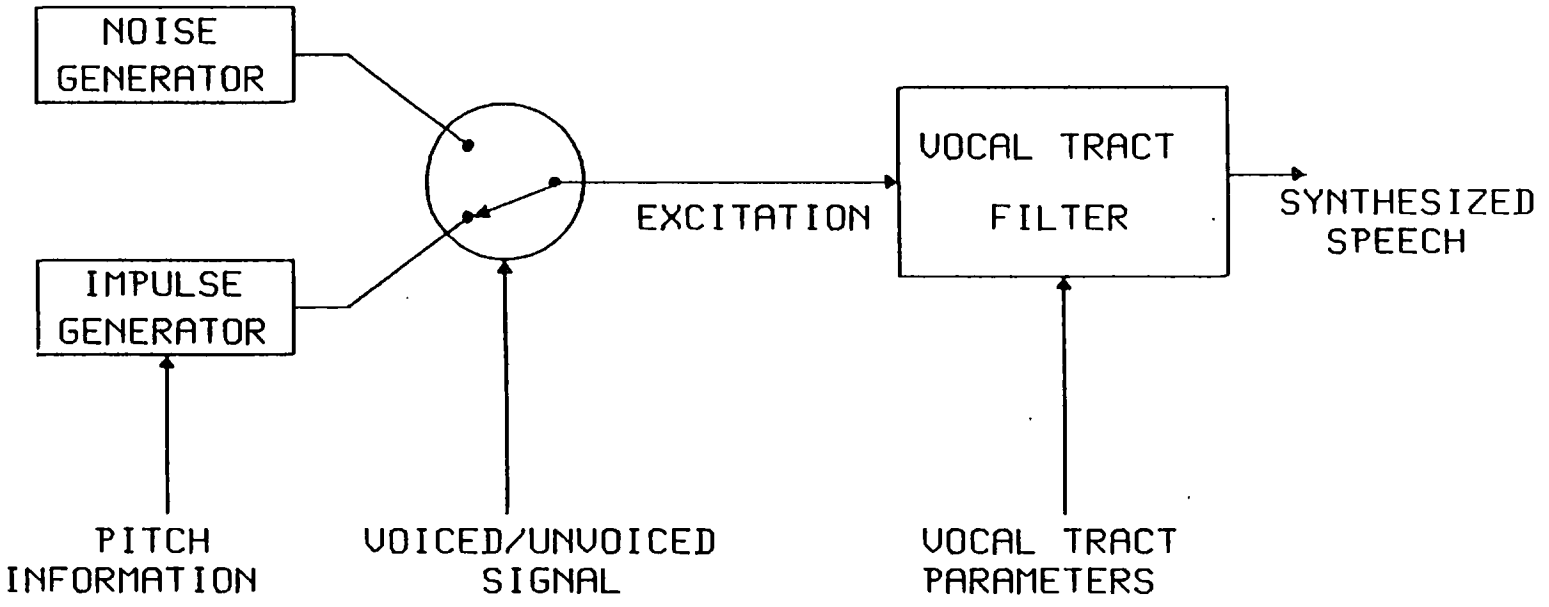


Fig.1.8 Block diagram of a simple speech production model.

CHAPTER 2

REVIEW OF RELATED WORKS

CHAPTER 2

REVIEW OF RELATED WORKS

2.1 INTRODUCTION

In this chapter, a review of some of the topics that are related to the work reported on in this thesis is presented. The two frequency-domain coding techniques, namely, sub-band coding and transform coding, from which the hybrid frequency domain coding technique is derived, are described in Sections 2.2 and 2.4. The orthogonal transforms used in the work reported in the main body of the thesis are summarized in Section 2.3. Some variable-rate source coding techniques are examined in Section 2.5 with attention being directed to speech coding. This is followed in Section 2.6 by a discussion of the advantages and disadvantages of the two frequency-domain techniques and the motivation for using the hybrid frequency-domain coding technique.

2.2 SUB-BAND CODING OF SPEECH SIGNALS

The sub-band coding of speech has been studied extensively by Crochiere et al [57-63]. A block diagram of a general sub-band coding system is shown in Fig.2.1. In this form of coding, the speech signal is bandpass filtered into M sub-bands. A frequency-domain illustration of 4-band partitioning is given in Fig.2.2. The sub-band signals are

then lowpass translated, sampled at twice the respective bandwidths, encoded and multiplexed for transmission. The number of bits to be used for encoding the bands is selected based on the relative importance of the signals. In general, the important formant and pitch information are contained in the lower bands, whereas the upper bands contain the perceptually less critical fricatives and noise-like sounds; so for a fixed allocation scheme, more bits per sample are allocated in encoding signals in the lower bands.

As the correlations among samples of the lowpass translated and decimated sub-band signals are low, adaptive PCM [64-66] rather than differential PCM quantization method is used to encode the sub-band signals.

At the receiver, the reverse processes take place. The sub-band signals are demultiplexed, decoded, demodulated and bandpass translated to their original spectral positions. They are then combined to give a replica of the input speech.

The rationale for encoding speech in sub-bands is as follows:

i. Since quantization noise is not equally perceptible at all frequencies, the distribution of the noise can then be controlled by encoding the different sub-band signals using different numbers of quantization levels.

ii.Since quantization noise is confined to the sub-bands, the quantization noise in one sub-band is prevented from masking signals in other sub-bands.

In a simple design [57], fixed bit allocation is adopted based on the long-term spectral characteristics of the speech signal. A

dynamic bit allocation technique was studied by Grauel [63] in which the numbers of bits allocated to the different bands are dynamically adjusted at regular intervals according to the spectral power in the bands. Grauel has reported that by using dynamic bit allocation, the bit rate can be reduced by more than 2.5 kbit/s without degrading the speech quality. Similar studies were made by Ramstad [67] using 13 uniform bands each of bandwidth 250 Hz.

There are several ways to carry out the filtering, translation and downsampling processes. Complicated modulation, demodulation and translation techniques can be avoided by using the integer band sampling technique [57]. The application of this technique for one sub-band is illustrated in Fig.2.3. In this form of band-partitioning, the M sub-band bandwidths are chosen such that the band edges are $m_i f_i$ and $(m_i+1)f_i$ for $i=1,\ldots,M$ where m_i is an integer which may be different for different bands and f_i is the sampling rate of the signals in the i^{th} sub-band. This constraint is necessary to avoid frequency aliasing in the sampling process. In order to simplify the process of decimation and interpolation, it is desirable that f_i be chosen such that f_s/f_i is an integer, where f_s is the sampling rate of the input signal.

Estaban and Galand introduced the use of quadrature mirror filters in sub-band coding [62]. Through the use of quadrature mirror filters, interband frequency aliasing can be cancelled and a flat passband can be obtained with lower order filters than with conventional finite-impulse-response (FIR) filters. The studies of the design of quadrature mirror filters and the effects of finite-word-length on the performance of split-band systems using

quadrature mirror filters are reported in Chapter 5 and Chapter 6 respectively.

If frequency aliasing is to be avoided, the sampling rate of the signals in each sub-band cannot be lower than the lower bound stipulated by the sampling theorem. As each sample must be coded with at least one bit, the use of sub-band coding technique at very low bit rates thus has to depend on a selective discarding of bands of less significant information. For example, to encode speech at 4.8 kbit/s using sub-band coding, Crochiere and Sambur [59] suggested the use of two fixed sub-bands at the lower end of the speech spectrum together with two movable upper sub-bands as shown in Fig.2.4. The center frequencies of the two upper filters are varied according to the dynamic movement of the vocal tract resonances, thus capturing those spectral components of the signals that are perceptually most significant. To achieve low transmission rate, bandwidths of the filters have to be narrow. However, this leaves spectral gaps in the speech spectrum and the regenerated speech has been found to have a reverberant quality.

It has been reported [7] that the time domain samples can be reduced by a factor of up to 2 without serious degradation in the perceptual quality of speech using the simple and efficient time domain harmonic scaling (TDHS) algorithms. The combination of sub-band coding with TDHS was investigated by Malah et al [56]. The method requires an estimation of the pitch and involves an additional delay up to 100 ms for the pitch extraction and harmonic scaling operations.

Furthermore, although at low bit rates, the speech produced is less noisy and less reverberant compared with the speech produced by simple sub-band coding at the same bit rate, it is distorted.

2.3 ORTHOGONAL TRANSFORMS

As the main interest is in the digital processing of signals, most of the discussion that follows is focused on discrete transforms.

A set of sampled functions $\{u_k(n)\} = \{u_0(n), u_1(n), \ldots\}$ is said to be orthogonal over the discrete time interval n=0 to (N-1) if

$$u_k(n)u_m(n) = c, \text{ (c being a constant) if } k = m$$
$$= 0 \qquad \text{if } k \neq m \qquad (2.1)$$

When c = 1, the set is called an orthonormal set.

Any set of discrete values {x(n)} of length N can be represented in terms of the orthogonal functions as

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)u_k(n) \qquad \text{for} \qquad n = 0,1,\ldots,N-1 \qquad (2.2)$$

where {X(k)} are the transform coefficients which in turn are related to {x(n)} and the orthogonal functions by

$$X(k) = \sum_{n=0}^{N-1} x(n)u_n(k) \qquad \text{for} \qquad k = 0,1,\ldots,N-1 \qquad (2.3)$$

In matrix notation,

$$[X] = [T][x] \qquad (2.4)$$

where the transform matrix [T] consists of N basis vectors $[u_n]$.

Similarly, Equation (2.2) can be expressed in matrix form as,

$$[x] = [T]^{-1}[X] \tag{2.5}$$

where $[T]^{-1}$ is the inverse transform matrix.

For transform [T] consisting of an orthonormal set of functions, it can be shown that the inverse of [T] is equal to the transpose of [T]. Mathematically,

$$[T]^{-1} = [T]^{t} \tag{2.6}$$

2.3.1 Walsh-Hadamard Transform (WHT)

The Walsh-Hadamard transform is derived from a set of Walsh functions [68,69] which are comprised of only two values, +1 and -1. The transform derived from Walsh functions arranged in natural order (known as Hadamard functions) is referred to as the Hadamard-ordered Walsh-Hadamard transform, and that obtained from Walsh-functions ordered in increasing sequency order is called the sequency-ordered or the Walsh-ordered Walsh-Hadamard transform.

The one dimensional discrete Hadamard-ordered WHT of size N of a sequence $\{x(n);0 \leqslant n \leqslant N-1\}$ may be defined as

$$B(l) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)\mathrm{Had}(l,n) \qquad l = 0,1,\ldots,N-1 \tag{2.7}$$

where $B(l)$ is the l^{th} Hadamard coefficient and $\mathrm{Had}(l,n)$ is the discrete Hadamard function.

The discrete Hadamard functions Had(l,n), l,n=0,1,...,N-1, can be represented by

$$\begin{aligned} \text{Had}(l,n) &= +1 \quad \text{if} \quad l_i n_i \text{ is 0 or even} \\ &= -1 \quad \text{if} \quad l_i n_i \text{ is odd} \end{aligned} \tag{2.8}$$

where l_i, n_i denote the i^{th} bit in the binary representation of l and n, that is, $l=(l_{m-1}\cdots l_1 l_0)_{\text{binary}}$ and $n=(n_{m-1}\cdots n_1 n_0)_{\text{binary}}$ in which $m = \log_2 N$.

The inverse Hadamard-ordered WHT is given by

$$x(n) = \sum_{l=0}^{N-1} B(l)\text{Had}(n,l) \qquad n = 0,1,\ldots,N-1 \tag{2.9}$$

In matrix representation, the Hadamard matrix of order N, $[H]_N$, can be written in terms of the Hadamard functions as

$$\begin{bmatrix} \text{Had}(0,0) & \cdots & \text{Had}(0,N-1) \\ \text{Had}(1,0) & \cdots & \text{Had}(1,N-1) \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ \text{Had}(N-2,0) & \cdots & \text{Had}(N-2,N-1) \\ \text{Had}(N-1,0) & \cdots & \text{Had}(N-1,N-1) \end{bmatrix} \tag{2.10}$$

Alternatively, the Hadamard matrices of order $2^{(m+1)}$, where m is an integer, can be obtained by using the simple recursive relationship:

$$[H]_{2^{(m+1)}} = \begin{bmatrix} [H]_{2^m} & [H]_{2^m} \\ [H]_{2^m} & -[H]_{2^m} \end{bmatrix} \tag{2.11}$$

and $$[H]_{2^0} = \text{Had}(0,0) = 1 \tag{2.12}$$

The transform coefficients of the Hadamard-ordered WHT are not arranged in increasing sequency order. In applications where it is desirable to have the sequencies associated with the transform coefficients in normal increasing order, Walsh-ordered WHT should be used. The one dimensional discrete Walsh-ordered WHT of size N of a sequence $\{x(n);0 \leqslant n \leqslant N-1\}$ is related to the discrete Walsh function Wal(k,n) by

$$W(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)\text{Wal}(k,n) \qquad k = 0,1,\ldots,N-1 \qquad (2.13)$$

The inverse transform is given by

$$x(n) = \sum_{k=0}^{N-1} W(k)\text{Wal}(n,k) \qquad n = 0,1,\ldots,N-1 \qquad (2.14)$$

It should be noted that, for both the Hadamard-ordered and the Walsh-ordered transforms, apart from the multiplier (1/N), the inverse transform matrix is identical to the transform matrix.

The discrete Walsh functions Wal(k,n), n,k, = 0,1,...,N-1 can be obtained by a different ordering of the discrete Hadamard functions. Mathematically, for k,n = 0,1,...,N-1,

$$\text{Wal}(k,n) = \text{Had}(l,n) \qquad (2.15)$$

where l is obtained from k by bit-reversal of the Gray code equivalent, g, of the binary code k. For example, if

$$k = (k_{m-1}k_{m-2}\cdots k_1k_0)_{\text{binary}}$$

and $$g = (g_{m-1}g_{m-2}\cdots g_1g_0)_{\text{binary}}$$

then $$l = (g_0g_1\cdots g_{m-2}g_{m-1})_{\text{binary}}$$

where $g_i = k_i \oplus k_{i+1}$ in which $\oplus$ denotes modulo 2 addition. Clearly, k can be similarly obtained from l.

The Walsh coefficients W(k), k=0,1,...,N-1 are also related to the Hadamard coefficients B(*l*), *l* =0,1,...,N-1 by

$$W(k) = B(l) \tag{2.16}$$

where *l* is obtained from k by bit-reversal of the binary-to-Gray code conversion of k as detailed above.

An algorithm for fast computation of the Walsh-Hadamard transform can be derived using a matrix partitioning technique. As the Hadamard-ordered WHT and the Walsh-ordered WHT are different in the ordering of the transform vectors, they are described separately.

The fast Hadamard-ordered WHT is best illustrated by an example. With N=8,

$$[B] = \frac{1}{8}[H]_8[X] \tag{2.17}$$

The matrix $[H]_8$ can be partitioned and expressed in terms of $[H]_4$. Thus,

$$[B] = \frac{1}{8}\begin{bmatrix} [H]_4 & [H]_4 \\ [H]_4 & -[H]_4 \end{bmatrix}\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \\ X(4) \\ X(5) \\ X(6) \\ X(7) \end{bmatrix}$$

which can be rewritten as

$$\begin{bmatrix} B(0) \\ B(1) \\ B(2) \\ B(3) \end{bmatrix} = \frac{1}{8}[H]_4\begin{bmatrix} X_1(0) \\ X_1(1) \\ X_1(2) \\ X_1(3) \end{bmatrix}$$

and $$\begin{bmatrix} B(4) \\ B(5) \\ B(6) \\ B(7) \end{bmatrix} = \frac{1}{8} [H]_4 \begin{bmatrix} X_1(4) \\ X_1(5) \\ X_1(6) \\ X_1(7) \end{bmatrix}$$

where $$X_1(i) = X(i) + X(i+4) \qquad i=0,1,2,3$$

and $$X_1(i) = X(i-4) - X(i) \qquad i=4,5,6,7 \qquad (2.18)$$

By partitioning $[H]_4$ again, the following equations are obtained

$$\begin{bmatrix} B(0) \\ B(1) \end{bmatrix} = \frac{1}{8} [H]_2 \begin{bmatrix} X_2(0) \\ X_2(1) \end{bmatrix}$$

$$\begin{bmatrix} B(2) \\ B(3) \end{bmatrix} = \frac{1}{8} [H]_2 \begin{bmatrix} X_2(2) \\ X_2(3) \end{bmatrix}$$

$$\begin{bmatrix} B(4) \\ B(5) \end{bmatrix} = \frac{1}{8} [H]_2 \begin{bmatrix} X_2(4) \\ X_2(5) \end{bmatrix}$$

$$\begin{bmatrix} B(6) \\ B(7) \end{bmatrix} = \frac{1}{8} [H]_2 \begin{bmatrix} X_2(6) \\ X_2(7) \end{bmatrix}$$

where $$X_2(i) = X_1(i) + X_1(i+2) \quad i=0,1,4,5$$

and $$X_2(i) = X_1(i-2) - X_1(i) \quad i=2,3,6,7 \qquad (2.19)$$

Since $$[H]_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad (2.20)$$

Equation (2.19) can be expressed as

$$8B(i) = X_3(i) \qquad i=1,\ldots,8$$

where $$X_3(i) = X_2(i) + X_2(i+1) \qquad i=0,2,4,6$$

and $$X_3(i) = X_2(i-1) - X_2(i) \qquad i=1,3,5,7 \qquad (2.21)$$

Thus, apart from the (1/8) multiplier, the whole sequence of operations involves only additions and subtractions. These operations are represented in the form of signal flow graph in Fig.2.5. It should be noted that, in general, the number of additions and subtractions required for a transform of size N, where N is a power of 2, is $N\log_2 N$.

Although Walsh-ordered WHT coefficients can be readily obtained from the Hadamard-ordered WHT coefficients using Equation (2.16), the method is not very efficient since it involves a binary-to-Gray code conversion in addition to bit reversal. A better method proposed by Manz [70] makes use of a simple modification of the Hadamard-ordered WHT. With this method, the first step is to 'bit-reverse' the input and order it in ascending index order. Let the new sequence be denoted by $\{\hat{X}(m), m=0,\ldots,N-1\}$. This is followed by a procedure of locating the blocks where reversals occur. A block is defined as a group of additions/subtractions which is disjoint from its neighbours. In general, the k^{th} iteration has 2^{k-1} blocks. The term 'reversal' is used to mean that subtractions are performed in place of additions, and vice versa.

The positions of blocks and the block with reversal is illustrated in Fig.2.6 for a transform size N=8. The final values obtained using this modified procedure are then the Walsh-ordered WHT coefficients. The signal flow graph for the fast Walsh-ordered WHT is illustrated in Fig.2.7 for a transform size N=8.

2.3.2 Discrete Fourier Transform (DFT)

The discrete Fourier transform (DFT) of a finite duration sequence $\{x(n), 0 \leqslant n \leqslant N-1\}$ is defined as

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j(2\pi nk/N)} \qquad k = 0,1,\ldots,N-1 \qquad (2.22)$$

and the inverse transform is defined as

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{j(2\pi nk/N)} \qquad n = 0,1,\ldots,N-1 \qquad (2.23)$$

Several fast algorithms [71,73] are available for the computation of DFT. For most practical cases where N is equal to a power of 2, the N-point DFT of $\{x(n)\}$ can be separated into two parts as

$$\begin{aligned} X(k) &= \sum_{\substack{n=0 \\ n \text{ even}}}^{N-1} x(n)e^{j(2\pi nk/N)} + \sum_{\substack{n=0 \\ n \text{ odd}}}^{N-1} x(n)e^{j(2\pi nk/N)} \\ &= \sum_{n=0}^{N/2-1} x(2n)e^{j(4\pi nk/N)} + \sum_{n=0}^{N/2-1} x(2n+1)e^{j[2\pi(2n+1)k/N]} \end{aligned} \qquad (2.24)$$

Define two (N/2)-point sequences $x_1(n)$ and $x_2(n)$ such that

$$x_1(n) = x(2n) \qquad n = 0,1,\ldots,(N/2)-1 \qquad (2.25)$$

$$x_2(n) = x(2n+1) \qquad n = 0,1,\ldots,(N/2)-1 \qquad (2.26)$$

and let $X_1(k)$ and $X_2(k)$ be the (N/2)-point DFT's of $x_1(n)$ and $x_2(n)$ for $0 \leqslant k \leqslant (N/2)-1$. Equation (2.24) can be rewritten as

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/2-1} x_1(n)e^{j[2\pi nk/(N/2)]} + e^{j(2\pi k/N)} \sum_{n=0}^{N/2-1} x_2(n)e^{j[2\pi nk/(N/2)]} \\ &= X_1(k) + e^{j(2\pi k/N)}X_2(k) \qquad \text{for} \qquad 0 \leqslant k \leqslant (N/2)-1 \end{aligned} \qquad (2.27)$$

Similarly, it can be shown that for $(N/2) \leqslant k \leqslant N-1$,

$$X(k) = X_1[k-(N/2)] + e^{j(2\pi k/N)}X_2[k-(N/2)] \qquad (2.28)$$

Therefore, it can be seen that an N-point DFT can be obtained from two (N/2)-point DFT's.

The scheme above can be iterated to evaluate the (N/2)-point DFT'S $X_1(k)$ and $X_2(k)$. This decimation-in-time process can be continued until 2-point DFT's are left to be evaluated. The DFT of the two points f(0) and f(1) can then be evaluated as

$$F(0) = f(0) + e^{0}f(1) = f(0) + f(1) \tag{2.29}$$

and

$$F(1) = f(0) + e^{j(2\pi/2)}f(1) = f(0) - f(1) \tag{2.30}$$

The number of complex multiplications and additions involved using direct evaluation of the Fourier transform is of the order of N^2, while for the decimation-in-time algorithm described above, it is of the order of $N\log_2(N)$. Hence by using the fast algorithm, there are tremenduous savings in the number of computations to be carried out, especially when N is large, and the speed of computation of the transform is also increased.

2.3.3 Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) of a finite duration sequence $\{x(n), 0 \leqslant n \leqslant N-1\}$ is defined as

$$G(k) = \frac{2c(k)}{N}\sum_{n=0}^{N-1} x(n)\cos[\pi(2n+1)k/2N] \quad k = 0,1,\ldots,N-1$$

where $c(k) = 1/\sqrt{2}$ when $k = 0$

$= 1$ when $k = 1,2,\ldots,N-1$ (2.31)

The inverse cosine transform is defined as

$$x(n) = \sum_{k=0}^{N-1} c(k)G(k)\cos[\pi(2n+1)k/2N] \quad n = 0,1,\ldots,N-1 \tag{2.32}$$

Several efficient methods of computing the DCT are available [74-78]. It can be evaluated using a 2N-point DFT, or through an N-point inverse DFT. A method that is approximately six times as fast as the approach using a double-size DFT has been developed by Chen et al [76]. The method proposed by Narasimha and Peterson [77] using an N-point inverse DFT will now be described as it can be easily implemented.

By defining a new N-point sequence {y(n)}, assuming N is even, such that

$$y(n) = x(2n)$$

$$y(N-1-n) = x(2n+1) \qquad n = 0,1,\ldots,(N/2)-1 \tag{2.33}$$

then
$$G(k) = \frac{2c(k)}{N}\{\sum_{n=0}^{N/2-1} y(n)\cos[\pi(4n+1)k/2N] + \sum_{n=0}^{N/2-1} y(N-1-n)\cos[\pi(4n+3)k/2N]\} \tag{2.34}$$

Denoting $n' = N-1-n$, the following equation can be obtained by reordering the above summation.

$$G(k) = \frac{2c(k)}{N}\sum_{n=0}^{N-1} y(n)\cos[\pi(4n+1)k/2N] \tag{2.35}$$

Equation (2.35) can be shown to be equivalent to

$$G(k) = \frac{2c(k)}{N}\mathrm{Re}[U(k)] \qquad k = 0,1,\ldots,N-1 \tag{2.36}$$

where
$$U(k) = e^{j(\pi k/2N)}\sum_{n=0}^{N-1} y(n)e^{j(2\pi nk/N)} = e^{j(\pi k/2N)}Y(k) \tag{2.37}$$

in which Y(k) is the inverse DFT of the sequence $\{y(n), 0 \leq n \leq N-1\}$.

Alternatively, as $U(N-k) = j[U^*(k)]$, where $U^*(k)$ denotes the conjugate of $U(k)$, $G(k)$ can be obtained by computing $U(k)$ for $k=0,1,\ldots,N/2$ and separating the real and imaginary parts as follows:

$$G(k) = \frac{2c(k)}{N}\mathrm{Re}[U(k)] \qquad k = 0,1,\ldots,(N/2)-1$$

and

$$G(N-k) = \frac{2c(k)}{N}\mathrm{Im}[U(k)] \qquad k = 1,2,\ldots,N/2 \tag{2.38}$$

For the inverse transform, the inverse DFT of the complex sequence $\{c(k)G(k)e^{j\pi k/2N},\ 0 \leqslant k \leqslant N-1\}$ is first evaluated. Let these be $\{V(n),\ 0 \leqslant n \leqslant N-1\}$. The even data points are then determined as

$$x(2n) = \mathrm{Re}\{V(n)\} \qquad n = 0,1,\ldots,N-1 \tag{2.39}$$

The odd data points can be found from the even data points using the relation

$$x(2n+1) = x[2(N-1-n)] \qquad n = 0,1,\ldots,(N/2)-1 \tag{2.40}$$

2.4 TRANSFORM CODING

In the block transform coding of speech signals, a windowed segment of the signal consisting of N samples $\{x(n),n=0,\ldots,N-1\}$ is linearly transformed into a new set of sequence $\{y(n)\}$ using a unitary matrix $[T]$. Mathematically,

$$[y] = [T][x] \tag{2.41}$$

The elements of $\{y(n)\}$ are then independently quantized, giving a vector $[\hat{y}]$. This quantized sequence is subsequently transmitted to the receiver. An estimate $[\hat{x}]$ of $[x]$ is obtained by taking the inverse transform of $[\hat{y}]$ giving

$$[\hat{x}] = [T]^{-1}[\hat{y}] \tag{2.42}$$

By joining the successive segments, an estimate of the original signal is obtained.

For unitary matrices, the average mean-square overall distortion, $\bar{D}$, is equal to

$$\bar{D} = \frac{1}{N} E[(x-\hat{x})^t(x-\hat{x})] = \frac{1}{N} E[(y-\hat{y})^t(y-\hat{y})] \qquad (2.43)$$

where the superscript t denotes the transpose of the vector.

If the average mean-square distortion is used as a criterion of goodness of the coding scheme, then (i) a transform [T] and (ii) a method of quantization of [y] that will give the minimum distortion can be found.

For block coding of signals where the signal amplitudes have a Gaussian distribution, the optimum quantization solution requires that the coefficients be quantized independently with varying number of bits. It has been found [79] that the number of bits, b_i, to be allocated to code a coefficient of variance σ_i^2 is related to the average mean-squared distortion D_i by

$$b_i = \theta + \frac{1}{2} \log_2(\sigma_i^2/D_i) \qquad i = 0,1,\ldots,N-1 \qquad (2.44)$$

where θ is a constant whose value depends on the quantizer used.

Suppose that the average number of bits available per block is equal to $\bar{b}$ bit/sample. It can be shown that to minimize the average distortion $\bar{D} = (1/N) \sum_{i=0}^{N-1} D_i$, all transform values should have an identical distortion equal to $\bar{D}$. Under this condition,

$$b_i = \bar{b} + \frac{1}{2} \log_2\{\sigma_i^2/[\prod_{j=0}^{N-1} \sigma_j^2]^{(1/N)}\} \qquad (2.45)$$

and $$\bar{D} = [\prod_{j=0}^{N-1} \sigma_j^2]^{(1/N)} 2^{(2\theta - 2\bar{b})} \quad (2.46)$$

With this bit assignment, the noise spectrum will be flat.

If transform coefficients have the same probability distribution as the signal samples, then θ does not change irrespective of whether the signal samples, or the transform coefficients are quantized. If the signal samples are quantized using a PCM scheme, then the average distortion, $\bar{D}_{PCM}$, due to quantization is given by

$$\bar{D}_{PCM} = 2^{2\theta - 2\bar{b}} \sigma^2 \quad (2.47)$$

where σ^2 is the signal variance.

The gain of a transform coding scheme as compared with a PCM coding scheme can be defined as the ratio of $\bar{D}_{PCM}$ to $\bar{D}$, that is,

$$G_{TC} = (\bar{D}_{PCM}/\bar{D}) = \{\sigma^2/[\prod_{j=0}^{N-1} \sigma_j^2]^{(1/N)}\} \quad (2.48)$$

For a unitary transform, σ^2 is related to the variances of the transform coefficients by

$$\sigma^2 = \frac{1}{N}\sum_{j=0}^{N-1} \sigma_j^2 \quad (2.49)$$

Thus, G_{TC} is the ratio of the arithmetic mean to the geometric mean of the variances of the transform coefficients. If all the σ_j^2's are equal, then G_{TC}=1 and there is no gain. Thus transform coding has an advantage over PCM coding only if there is variation within the variances of the transform coefficients.

It has been shown [79] that the best transform to be used is the Karhunen-Loeve transform (KLT). However, the KLT suffers from the setback that the basis vectors are signal dependent and these make computation more complicated. Furthermore, no fast algorithm is available for its computation.

A comparison of the DFT, the WHT and the KLT for full-band speech coding has been carried out by Campanella and Robinson [80] based on the improvement of signal-to-quantizing-noise over PCM. It was reported that bit savings of 13.5 kbit/s, 10 kbit/s and 7.5 kbit/s can be achieved over 56 kbit/s logarithmically companded PCM using the KLT, the DFT and the WHT respectively. The gains in signal-to-quantizing-noise over PCM for the three transforms have been found to be independent of bit rates and they are :KLT 9 dB, DFT 5 dB and WHT 3 dB. Using stationary tenth-order Markov processes with the first ten autocorrelation coefficients made identical to those of segments of speech signal, Zelinski and Noll have shown [81] that of the transforms that have fixed basis vectors, the discrete cosine transform (DCT) is closest to the KLT in performance for full-band speech coding. It was also observed that both the DFT and the DCT approach the KLT performance for large transform block sizes. The WHT and the discrete slant transform (DST), however, do not show an asymptotically optimum performance.

Although the WHT does not show optimal performance for full-band speech coding when the signal-to-noise ratio is used as a criterion, its computation is very much simpler than that of the DCT. Furthermore, at high bit rates, speech signals coded using the WHT are of very good quality [82-83]. Also, for the same degree of accuracy,

the word-length required for block-floating-point hardware implementation of the WHT is very much smaller than that required for the DFT and the DCT.. Details of this comparison are further highlighted in Chapter 7.

The advantage that can be gained over PCM in using a transform technique lies in the fact that the energy content of segments of speech signals are concentrated in a few transform coefficients. Thus, more bits can be allocated to code the most significant coefficients. For this method of coding, both the transmitter and the receiver must have information about the number of bits allocated and the step-size used to encode the different transform coefficients. This information is often referred to [81] as side-information.

Several methods of bit assignments have been reported. With schemes in which a fixed bit allocation procedure is adopted, the number of bits allocated and the step-size used are predetermined on the basis of the statistics of long sections of speech signals. With fixed allocation scheme, transmission of the 'side-information' is not necessary. However, as speech is a highly nonstationary process, the fixed allocation methods are unable to exploit to the full extent the redundancy in speech signals.

For adaptive transform coding using the discrete cosine transform, a simple method of coding the side-information on a frame-to-frame basis has been proposed by Zelinski and Noll [81]. The technique makes use of the smoothness of the spectra of the logarithm of cosine transform coefficients of speech signals. An estimate of the envelope of the spectra is then encoded and transmitted. The number of

bits assigned and the step-size used to encode the transform coefficients at both the transmitter and receiver are based on the estimated spectral envelope. Specifically, the variances σ_k^2 of the coefficients are estimated using the information about the envelope of the actual variances. The number of bits to be assigned for coding a coefficient is then determined according to Equation (2.45).

The method proposed by Zelinski and Noll is illustrated in Fig.2.8. The spectral envelope can be described by a set of equally spaced samples of spectral estimates called 'group variances'. As shown in the figure, these group variances are determined by a local averaging of the DCT coefficients around them. The logarithms of the group variances are quantized and transmitted. Estimates of the variances of the transform coefficients can then be obtained by geometric interpolation using the logarithms of the group variances. According to the method proposed by Zelinski and Noll [81], approximately 2 kbit/s is required for side-information when 16 group variances are used.

The method is adequate for speech coding at 16 kbit/s and above. The reason is that after deducting the number of bits used for side-information, there still remains a sufficient number of bits for adequate coding of most of the transform coefficients with the result that the pitch harmonic structure is adequately represented. However, at bit rates below around 14 kbit/s, the low level portions at high frequencies are not coded, therefore the speech sounds muffled. At 8 kbit/s, the pitch information is no longer sufficiently preserved and the reconstructed signal is significantly degraded with either a perceptible burbling or 'click-type' distortion appearing.

A more accurate method of estimating the spectral structure has been proposed by Tribolet and Crochiere [6]. It is called the vocoder-driven adaptation method as the spectral structure is described by an all-pole model of the formant of speech together with a pitch model to represent the pitch striations in the speech spectrum. It should be noted that the estimated spectral structure is used only as 'side-information' to determine the bit assignment and step-size for the encoding and decoding of the transform coefficients and is not used directly in coding the actual speech signal. Another vocoder-driven adaptation method using a homomorphic vocoder model has been investigated by Cox and Crochiere [84].

Although these methods give more accurate estimates of the variances of the transform coefficients, and an improvement in the quality of the reproduced speech, time consuming computations are involved and hence the complexity of the system is increased.

The bit assignment according to Equation (2.45) aims to maximize the overall SNR value and results in a flat noise spectrum. However, it is known from experiments on auditory masking that noise in the regions of the spectrum where speech energy is concentrated, such as the formant regions, is masked by the speech sound. Perceptually, noise is more noticeable when it occurs in frequency regions in which the speech signal is low. To make use of this perceptual masking characteristics, the noise spectrum can be shaped such that there is lower noise in region where signal level is low. This can be achieved by modifying the bit assignment rule in the following manner:

$$b_i = \bar{b} + \frac{1}{2}\log_2(W_i\sigma_i^2/\bar{D}) \qquad (2.50)$$

where W_i represents a positive weighting function. One class of weighting functions suggested [6] is

$$W_i = \sigma_i^{-2\gamma} \tag{2.51}$$

where γ is a constant between 0 and 1, and σ_i^2 denotes the variances of the transform coefficients. As shown in Fig.2.9, for γ=0, this reduces to the case of a flat noise spectrum. The value γ=1 implies that the same number of bits is assigned to code the coefficients, hence the noise spectrum will follow the input signal spectrum. Noise spectra between the two extremes can be obtained by choosing a value of γ between 0 and 1.

It has been found [6] that improved speech quality can be obtained in the range from 8 kbit/s to 16 kbit/s as compared with the non-speech adaptive transform coder. This improvement in quality can be attributed to the fact that with the vocoder-driven adaptive transform coder, a more parsimonious allocation of bits to encode the side-information can be made, hence there are more bits available to encode the transform coefficients proper. Furthermore, the addition of the pitch model in the side-information means that more weights are given to the components at the pitch harmonics in the allocation of bits; and in consequence, the pitch structure is better preserved. However, at lower bit rates, there is insufficient number of bits to adequately code all the transform values and to convey the side-information. In fact, at 4.8 kbit/s, significant degradations and click noise can also result with the vocoder-driven adaptive transform coder.

2.5 VARIABLE-BIT-RATE CODING OF SPEECH SIGNALS

As mentioned previously, speech is a nonstatonary and intermittent process. In the broad sense, speech signals can be segmented into voiced speech, unvoiced speech and silence with each having different features. These properties are utilized in communication systems such as TASI (Time Assignment Speech Interpolation) [86,87] to obtain a more efficient use of the transmission channel. Efficient coding schemes should also seek to exploit these different characteristics by coding the different segments of speech signals differently with more bits being assigned to code the perceptually more significant segments. Alternatively, different numbers of bits can be used to code different classes of speech segments. If these are carried out, the source rate is then variable. Various schemes using variable source rate and exploiting the nonstationarity of speech signals have been proposed [88-93]. Some of the schemes, such as extremal coding, time encoding and adaptive aperture coding inherently result in variable-bit-rate generation, while other techniques such as the variable-rate coding technique proposed by Crochiere and Dubnoski [93] exploits the advantage of coding the different segments of speech signal at different bit rates by adaptive differential PCM. Since the source activity is variable, information is generated and extracted at irregular intervals; and in order to avoid distortion and to maintain a low constant transmission rate, time delays must be allowed and storages must be provided for the encoding and decoding of data. Details of a theoretical study of the problems associated with constant-rate transmission of variable-rate source information can be found in [94,95].

2.5.1 Extremal Coding

In extremal coding as proposed by Mathews [88], only the amplitudes and the time intervals between the extremes of the speech waveform are determined and coded (see Fig.2.10(a)). As illustrated in Fig.2.10(b), at the receiver, a smooth function is used to interpolate between the extremal amplitudes. To further reduce the bit rate, extremes resulting from noise during silent intervals and speech wave 'bumps' which are small compared with the surrounding signal waveform are not coded.

As there are more extremal points for unvoiced sounds than for voiced sounds, the source activity is not constant. Based on a small amount of speech material, Mathews has shown that on the average, there are 1400 extremal points per second in speech signals. The actual number of extremes depends not only on the speech material and the speaker but also on the type of microphone used. It has been found [88] that at an average bit rate of 15 kbit/s, the quality of speech produced by the use of extremal coding is much the same as that produced by the use of 30 kbit/s PCM. However, even when the bit rate is increased, the inherent degradation in the extremal interpolation function does not allow for significant improvement in speech quality. Moreover, it is also estimated [88] that a buffer of 15000 bits and associated transmission time delay of about 1 second is required for a single channel coder.

2.5.2 Time Encoding

The general waveforms between zero crossings of a speech signal appear to be very simple and consist of only a small number of peaks. Thus a speech signal can be encoded in terms of the patterns of the waveform between zero crossings. The speech produced using this method of coding is referred to [89,90] as 'time encoded speech' (TES) by King and Gosling, and these same authors have found that good quality speech can be produced at low bit rates.

In the system proposed by King and Gosling, the speech signal is encoded using a set of alphabets (waveforms, patterns) which define the waveform of the speech signal in terms of two parameters, namely, the time interval between adjacent zero-crossings T_i and the number of absolute maxima of the waveform between these zero crossings N_i. The set of alphabets thus consists of the duplets (T_i, N_i), $i = 1,\ldots,N$; with the alphabet that best matches the particular segment of speech signals being used to encode that segment of speech.

In addition, to provide information about the average amplitude of the signal, an N-bit word is also transmitted intermittently. The average transmission bit rate depends on the number of alphabets used, the number of bits per alphabet, and also the number of zero crossings in the particular signal. The number of alphabets to be used can be decided based on the probability of occurrence of the duplets (T_i, N_i). It has been found [89] that good intelligible speech can be achieved when 23 alphabets are used. As for the number of zero crossings, it has been found [94] that for speech signals, it varies widely depending on the speaker, the speech material and the type of

microphone used. If it is assumed that the average zero crossing rate is 900 per second and 5 bits are assigned to code each alphabet and 5 bits are allocated for coding the r.m.s. amplitude of every 4 zero-crossing segments, then the average transmission rate is [900x5 + (900x5/4)] = 5625 bit/s.

Studies [94] have shown that the source activity is highly non-uniform and that if the alphabets are coded using constant-length codes each of n bits, and if the transmission rate is set to be n times the average rate of generation of the alphabets, then for distortionless transmission, delays of the order of 1 second are required to convert the variable-rate-source data into constant-transmission-rate data.

Various methods of reducing the delays have been investigated: Mason and Balston have shown [96] that if the transmission rate is increased by a factor of 1.9, distortionless transmission can be achieved at 200 ms delay; and that at this delay, little or no distortion occurs even if the transmission rate is increased only by a factor of 1.5. By using suboptimum bounded entropic codes in place of constant-length codes, King and Holbeche have shown [97] that the delays may be reduced by an order of magnitude.

Holmes [98] proposed that by omitting a large proportion of the unvoiced waveform description and replacing the missing portions with repeats of waveform segments already sent, the buffer requirements can be greatly reduced; and that by suitable control of the omission, a great economy in the transmission rate can also be achieved.

2.5.3 Adaptive Aperture Coding

In aperture coding as proposed by Jayant and Christensen [91,92], the samples of a speech signal are coded only when their amplitudes exceed the limit of the aperture centered around the last encoded value (see Fig.2.11). For a slowly varying waveform and using wide aperture, the number of samples to be coded in this way are very much reduced below the number of Shannon samples, thus bit-rate reduction can be achieved. By using a converging aperture of finite length L, the information of the time instant of the samples that are coded are implicitly transmitted. In the situation illustrated in Fig.2.11, a transmitted sample at time 0 is followed by two redundant samples. The first sample to cross the aperture is X_3, which is encoded as amplitude P_3. This method of coding carries two types of information: the first is that the last two samples are not coded, that is, the run-length R=3; and second, the amplitude of the sample X_3 is approximately P_3. The value P3 is then chosen as the center point for the next aperture. The aperture length can be made finite by deciding that the level is 0 if no samples cross the aperture after L samples. Jayant and Christensen found that the converging apertures can be formalized into the class

$$A(t) = A_o 2^{-Jt} \tag{2.52}$$

where A(t) is the aperture width at time t relative to the starting point, A_o is the aperture width at the starting point and J is a constant. For speech signals, J is chosen such that $0.5 \leqslant J \leqslant 1$. The starting aperture width A_o can be made adaptive depending on the history of run-length R. If a large number of the preceding

run-lengths are long, indicating that the aperture may be too wide, A_o can be reduced. Similarly, if a large number of preceding run-lengths are short, A_o can be increased.

At the receiver, the non-transmitted samples are interpolated based on the value of the last transmitted sample and the aperture shape. The shape of the aperture with first order prediction a_1=1 and that with more general prediction are shown in Fig.2.12. In the case where the predictor coefficient a_1 is equal to 1, the non-transmitted or redundant samples are predicted with values equal to the last transmitted value as indicated by the dashed horizontal line running through the aperture area. In general, a speech specific predictor with a_1 = 0.85 or higher order predictor is used to reconstruct the redundant samples and to predict nonredundant samples prior to updating.

Using adaptive apertures with first order prediction and L = 2,3, it has been found [91] that average encoding rate of 1.2 bit/sample is possible without much deterioration in the output speech quality. This means that for sampling rates of 8 ksample/s and 12 ksample/s, the average bit rates are 9.6 kbit/s and 14.4 kbit/s respectively. The reconstructed speech has been judged to be better than adaptive delta-modulated speech at similar bit rates. It has also been reported [91] that encoding delays in the range of 35 ms to 80 ms can be achieved depending on the degree of sophistication.

2.5.4 Variable-Rate ADPCM Coding

The coding of speech where different numbers of bits are used to code different segments of speech signal, with more bits being assigned to segments with larger variances, has been examined by Dubnowski and Crochiere [93]. In this scheme, the segments are coded using adaptive differential pulse-coded-modulation (ADPCM).

Specifically, the output of the ADPCM coder is framed into packets of typically 60 bits with 2-bit header preceding each packet. The variance of the speech signal is determined once per packet based on the step-size used in the adaptive one-word memory quantizer. Depending on the value of the signal variance, 2,3,4 or 5 bits are allocated to code each of the samples. With a fixed length packets of 60 bits, these correspond respectively to 30, 20, 15 and 12 samples of data per packet.

As the variances of segments of speech signal vary widely for different speech signals, buffer overflow or underflow will result if the number of bits used to encode the samples is determined strictly according to the variance. In the coder proposed by Dubnowski and Crochiere [93], a dynamic buffer approach is adopted to deal with the problem of buffer overflow/underflow. Using this approach, the number of bits per sample used to code the signal is dynamically adjusted. For example, if the source activity is overly high, resulting in an excess demand on bit rate, the number of bits per sample used may be reduced. Conversely, if the source activity is very low, the number of bits per sample used may be increased.

2.6 MOTIVATION FOR HYBRID FREQUENCY-DOMAIN CODING

It has been reported [15] that good communication quality speech can be produced at 7.2 kbit/s using a vocoder-driven adaptive transform coding technique [6,84] and that similar quality speech can be obtained at 9.6 kbit/s using sub-band coding. Based on a relative count of the number of logic gates used, the complexity of the adaptive transform coder is estimated [15] to be approximately ten times that of the sub-band coder implemented using charge-couple-device filters.

Transform coding has the advantage over waveform coding such as PCM in that a reduced number of bits is required. With waveform coding, it is necessary to sample the signal at a rate which is at least twice the highest frequency contained in the signal and then to encode each sample with at least one bit. With transform coding, on the other hand, the signal is handled in blocks of samples and it is possible in the transform domain, to use a reduced average number of bits per coefficient as compared with the average number of bits per sample used in waveform coding. Besides, it is also possible in certain cases to neglect certain of the transform coefficients. Furthermore, the noise spectrum can be shaped according to a perceptual criterion by coding the transform coefficients with different number of bits.

For coding a full-band speech signal using fixed basis transforms, the cosine transform is the best choice for achieving good quality speech at low bit rates [81]. However, the computational load

associated with the cosine transform is considerable and this adds to the complexity of the method.

The sub-band coding technique has the advantage of confining the noise to the sub-bands, but the coding of the sub-band signals using adaptive PCM is, however, somewhat inflexible. As explained above, a number of bits equal to at least twice the bandwidth of the sub-bands has to be assigned for coding the sub-bands if frequency aliasing is to be avoided; and at low transmission bit rates, this limits the capabilities of dynamic bit assignment and noise spectral shaping.

In addition, if the sub-bands are not contiguous, that is, if gaps are left in the speech spectrum in order to conserve bandwidth and hence to reduce the number of samples to be coded, the resulting speech will have a reverberant quality. Furthermore, the method of interband aliasing cancellation using quadrature mirror filters cannot be applied if spectral gaps are to be imposed and very long-tap sharp cutoff filters have to be used.

The hybrid frequency-domain coding technique examined in this thesis combines the salient features of sub-band coding and transform coding. Through the use of sub-band coding, the noise can be restricted to the sub-bands. For signals in the upper sub-bands, where the available number of bits is low, if a transform coding technique is used instead of adaptive PCM, then each transform component can be separately weighted. Through transform coding of the upper sub-band signals, a dynamic bit allocation strategy can be employed. This not only allows for effective use of the available number of bits, but it also permits noise shaping and the reduction of the residual aliasing

components due to downsampling and quantization of the filtered signals in an implementation using quadrature mirror filters. In addition, as the correlations among samples of a sub-band signal are very much lower than those of a full-band speech signal, there should be minimal effect of the type of transform on the quality of the speech produced; the Walsh-Hadamard transform which is computationally much simpler compared with other transforms can be effectively applied for coding the sub-band signals.

The property of the nonstationarity of speech signals can be exploited to improve speech quality while maintaining the same average transmission bit rate. The hybrid frequency-domain coding scheme can be easily adapted for variable-bit-rate coding, and a tradeoff can be made between time delays and speech quality by coding different time segments of speech signals using different numbers of bits.

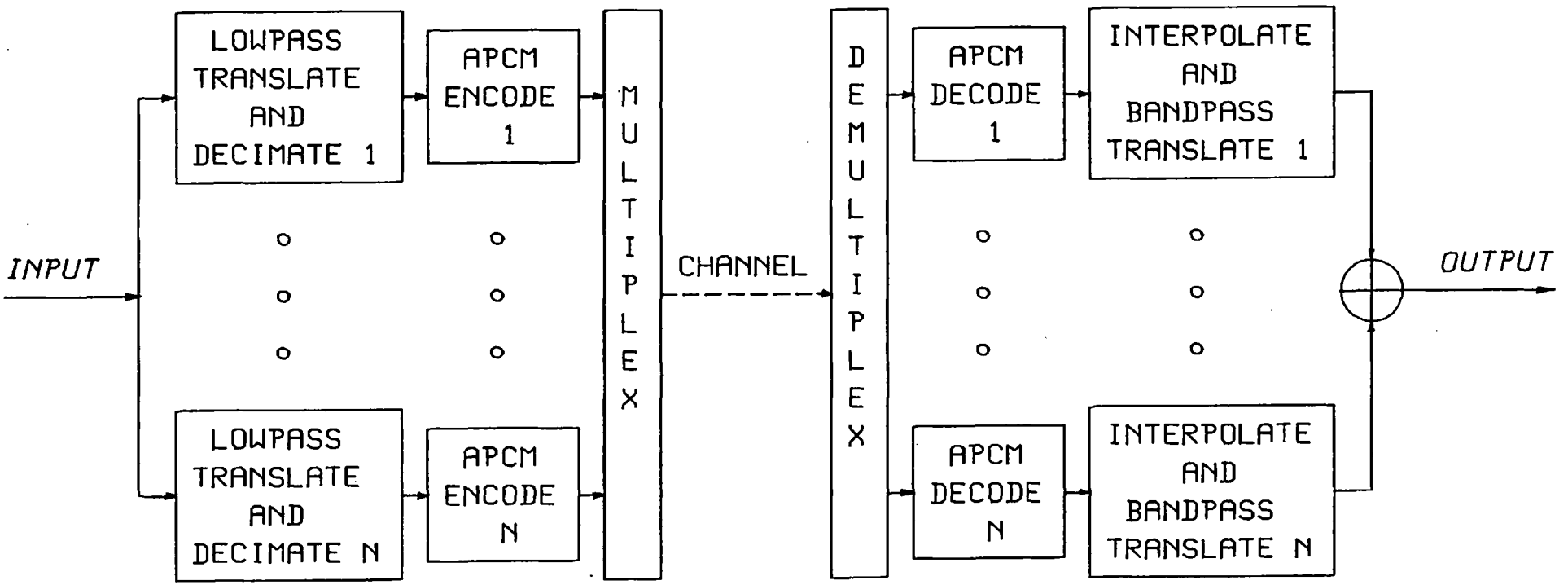


Fig.2.1 Block diagram of a general sub-band coding system.

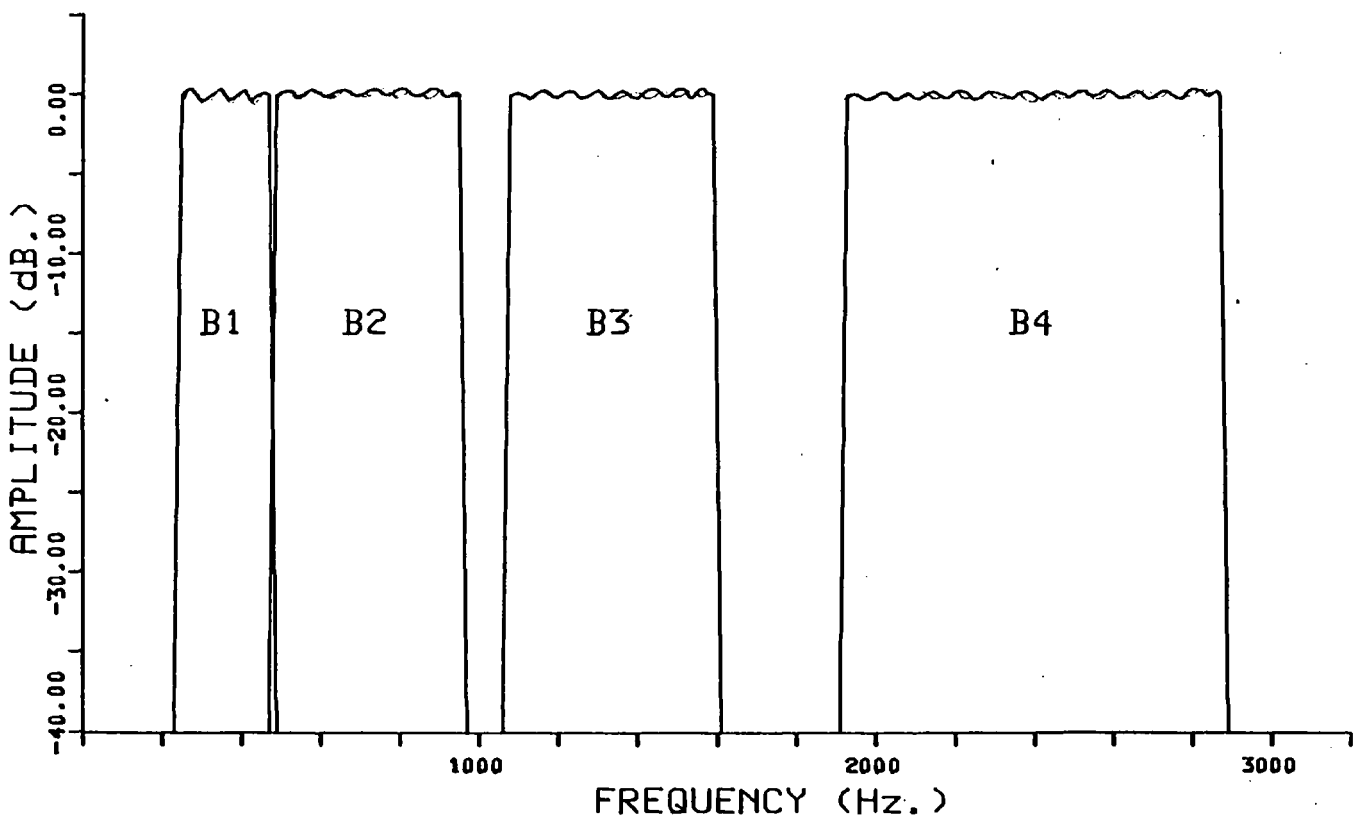
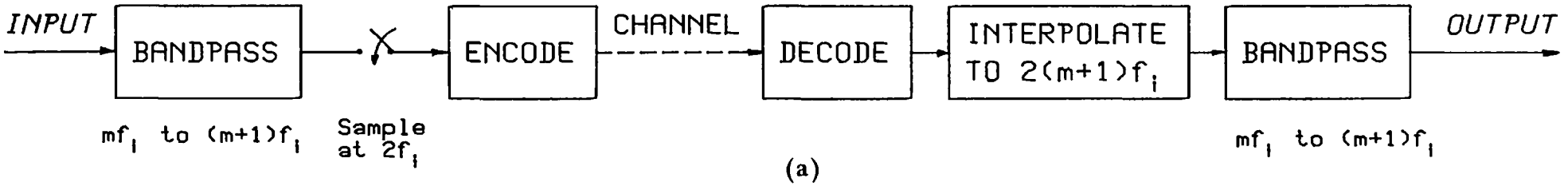
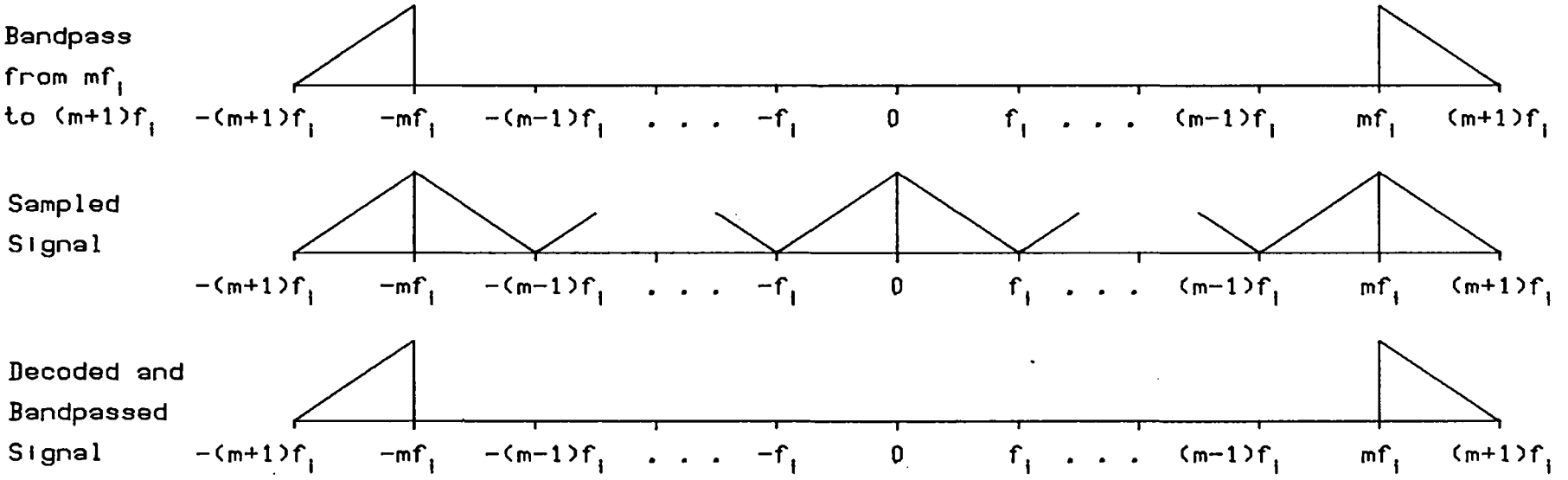


Fig.2.2. Partition of speech spectrum into bands.



(a)



(b)

Fig.2.3 Illustration of integer band sampling technique for one sub-band; (a) Block diagram; (b) Amplitude spectra.

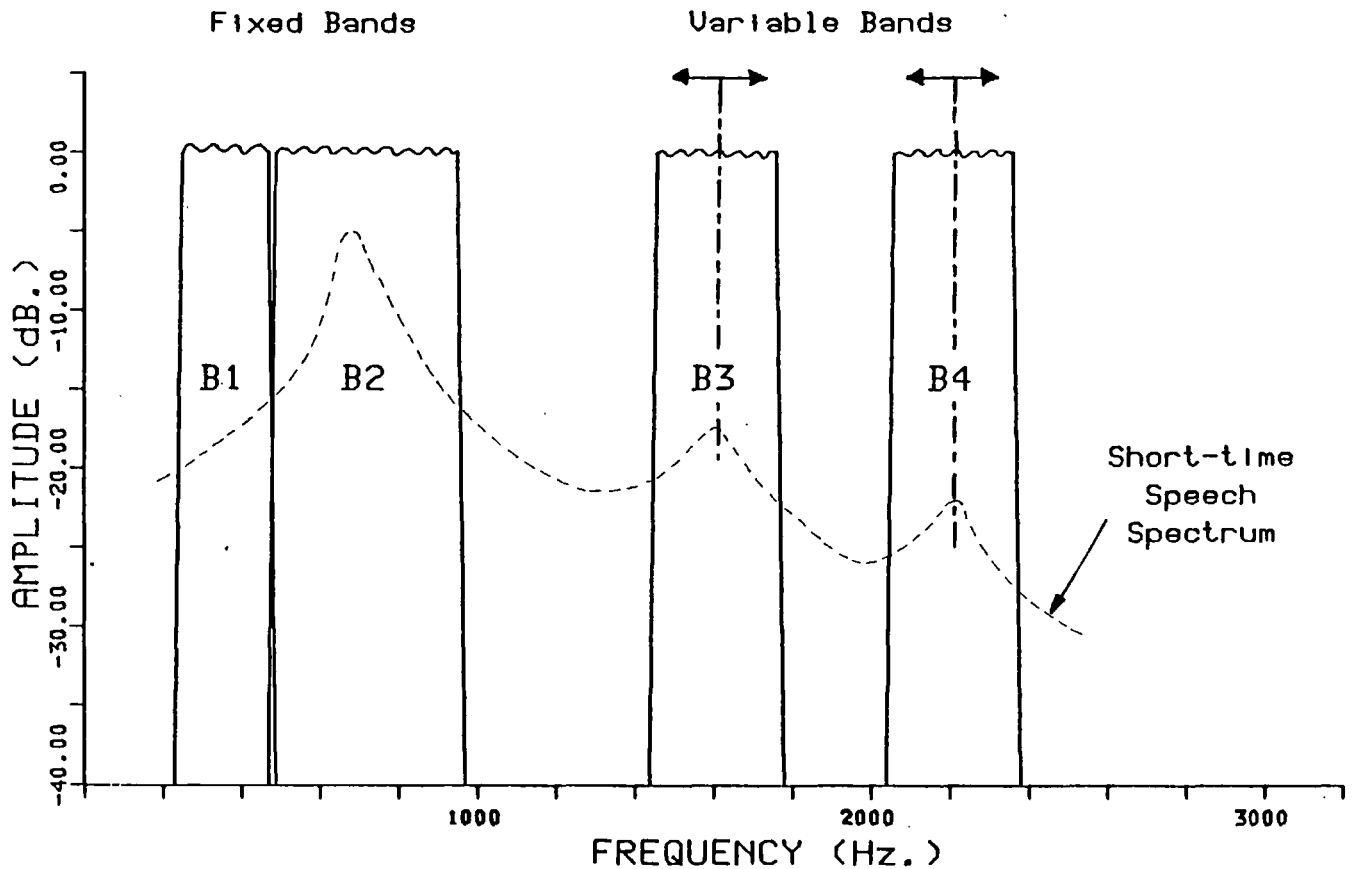


Fig.2.4 Frequency domain illustration of the movable upper sub-bands.

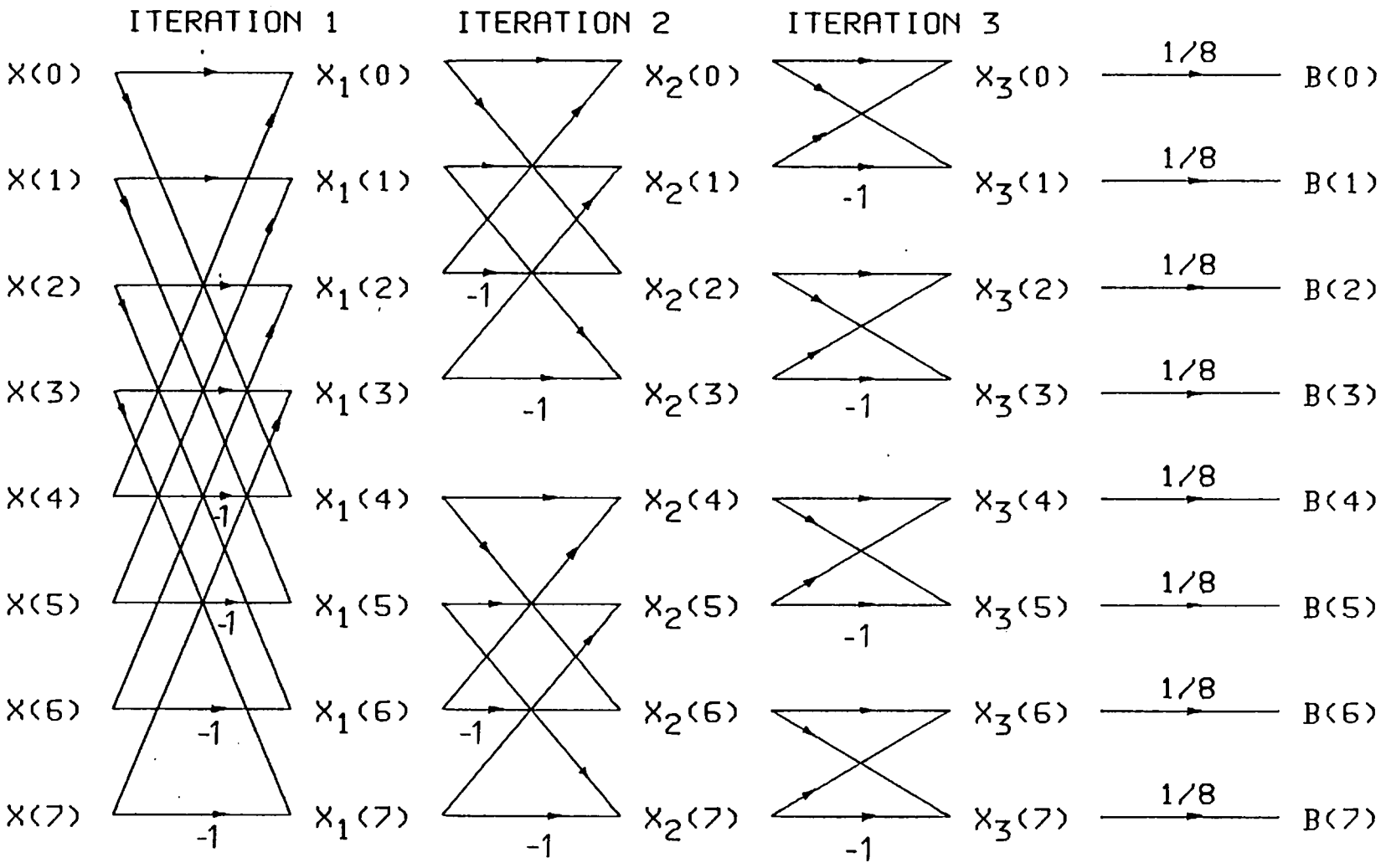


Fig.2.5 Signal flow graph of fast Hadamard-ordered Walsh-Hadamard transform.

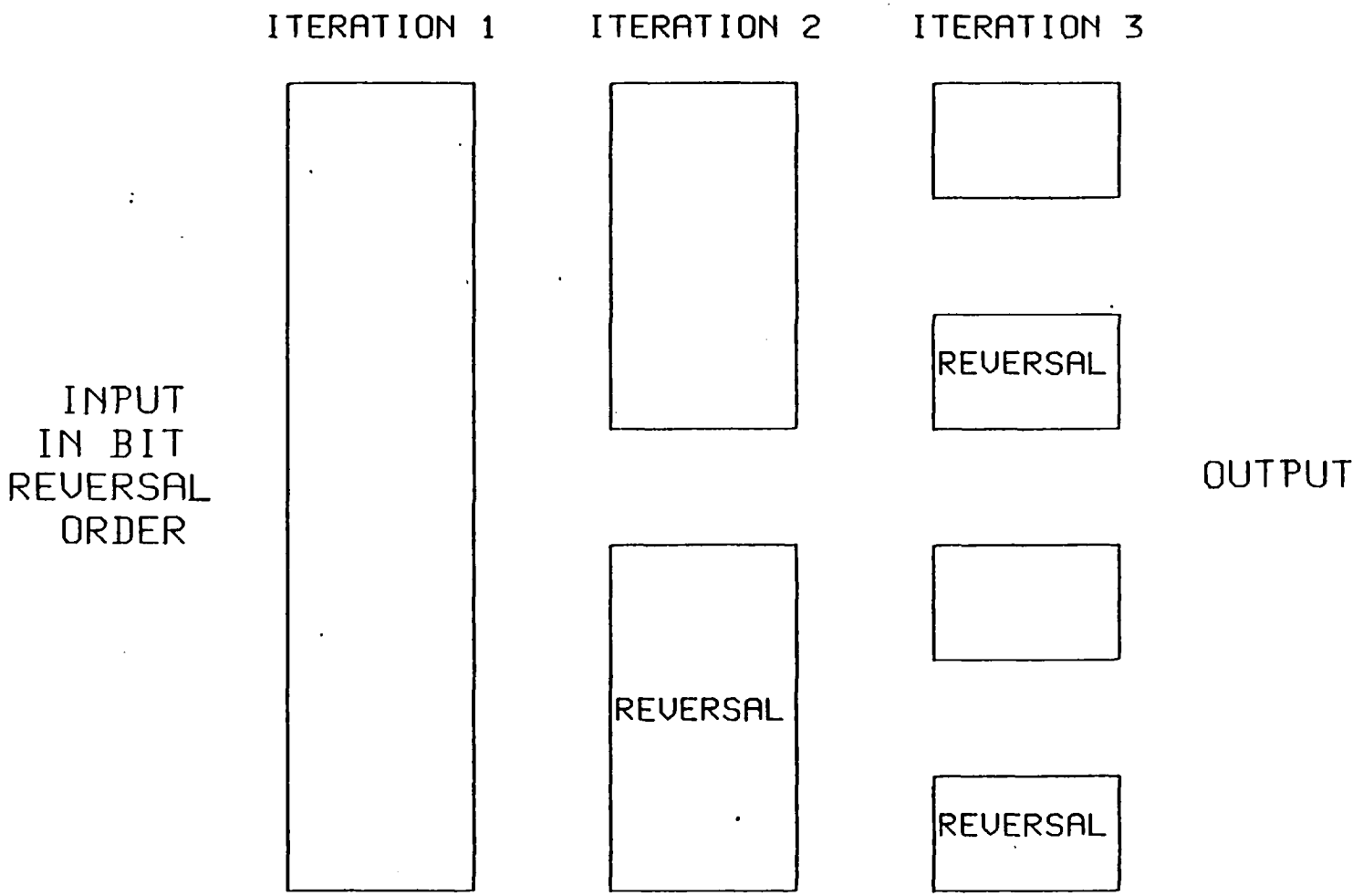


Fig.2.6 Illustration of block reversal.

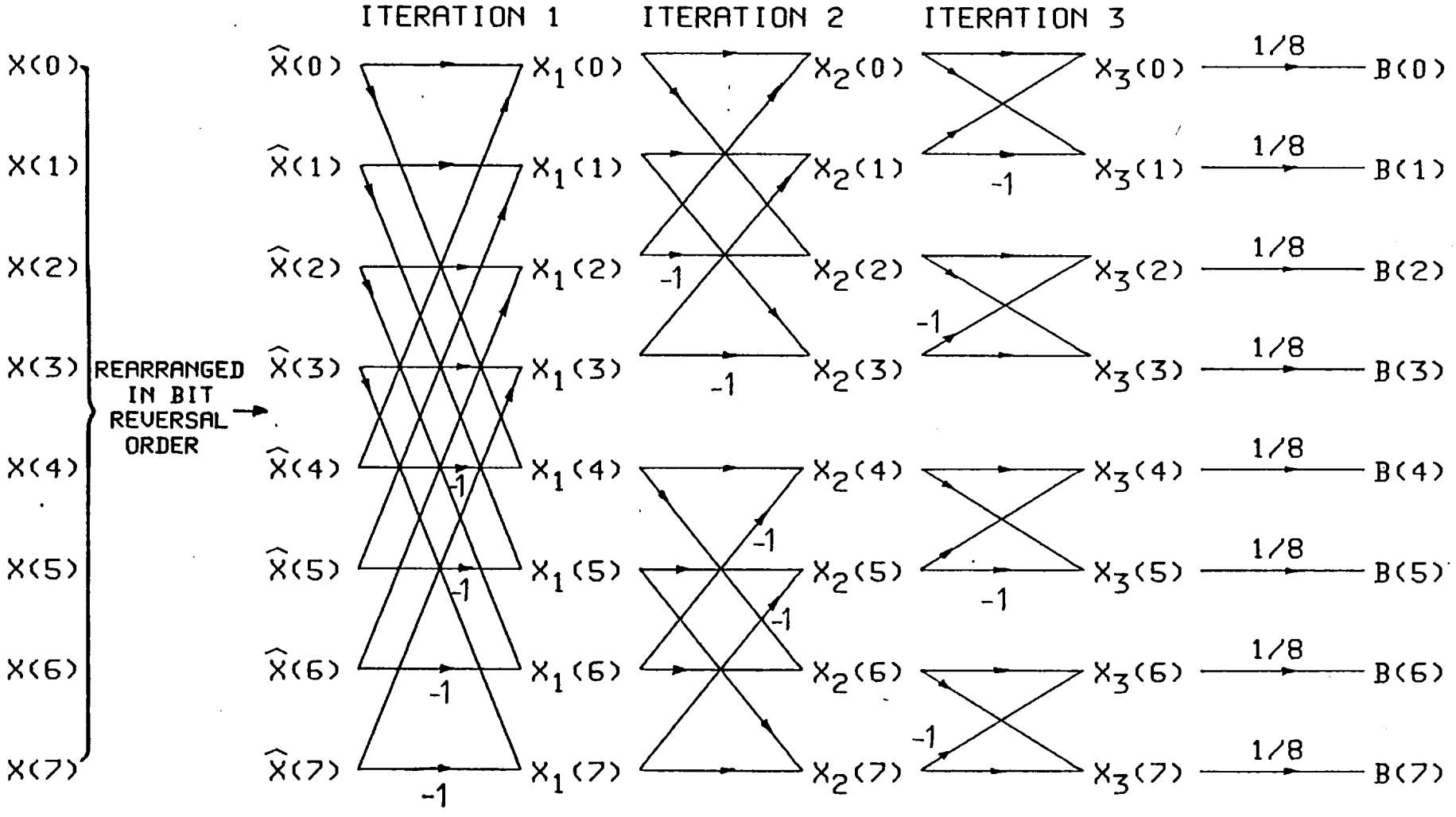


Fig.2.7 Signal flow graph of fast Walsh-ordered Walsh-Hadamard transform.

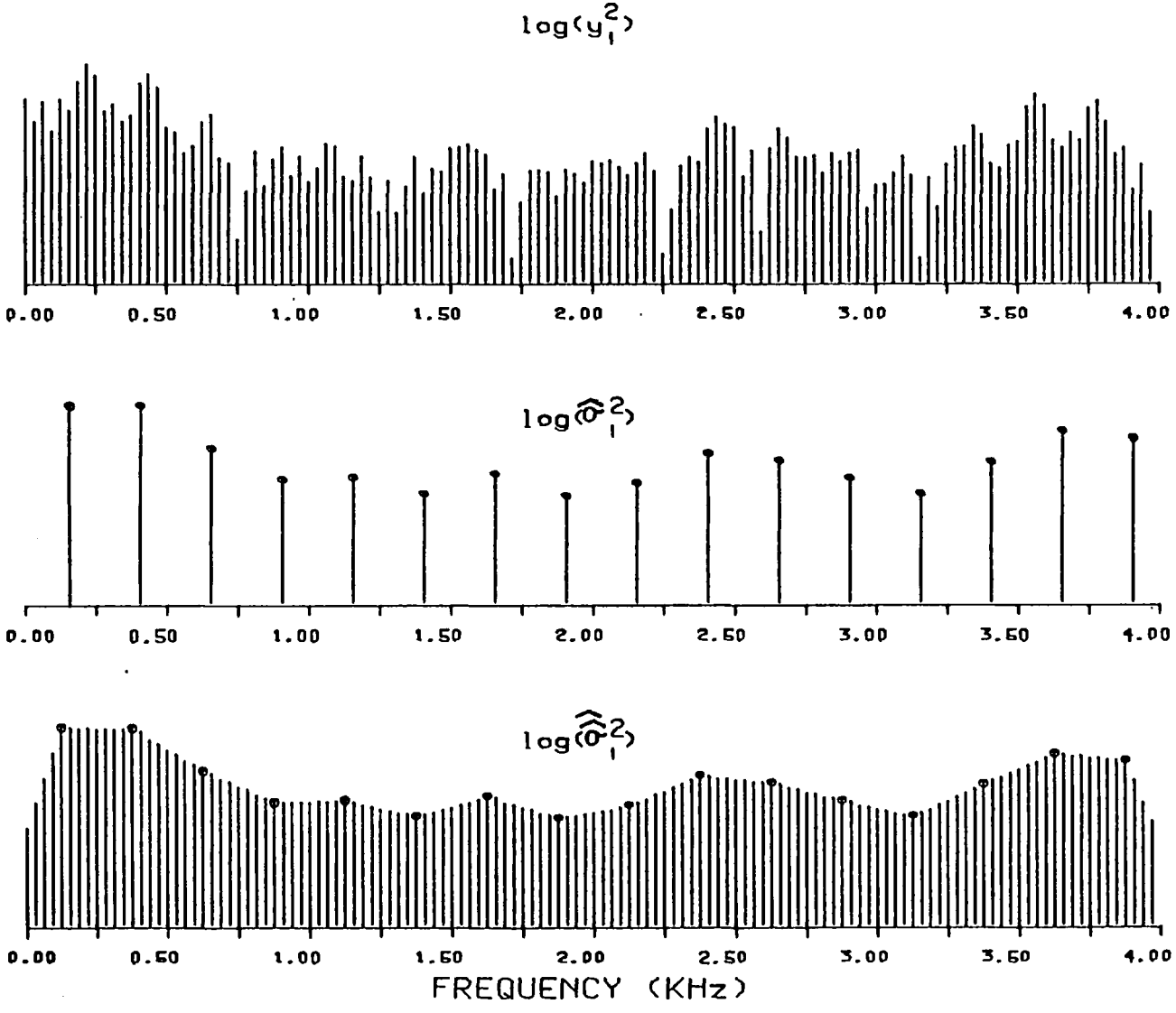


Fig.2.8 Illustration of coding and estimation of spectral envelope.

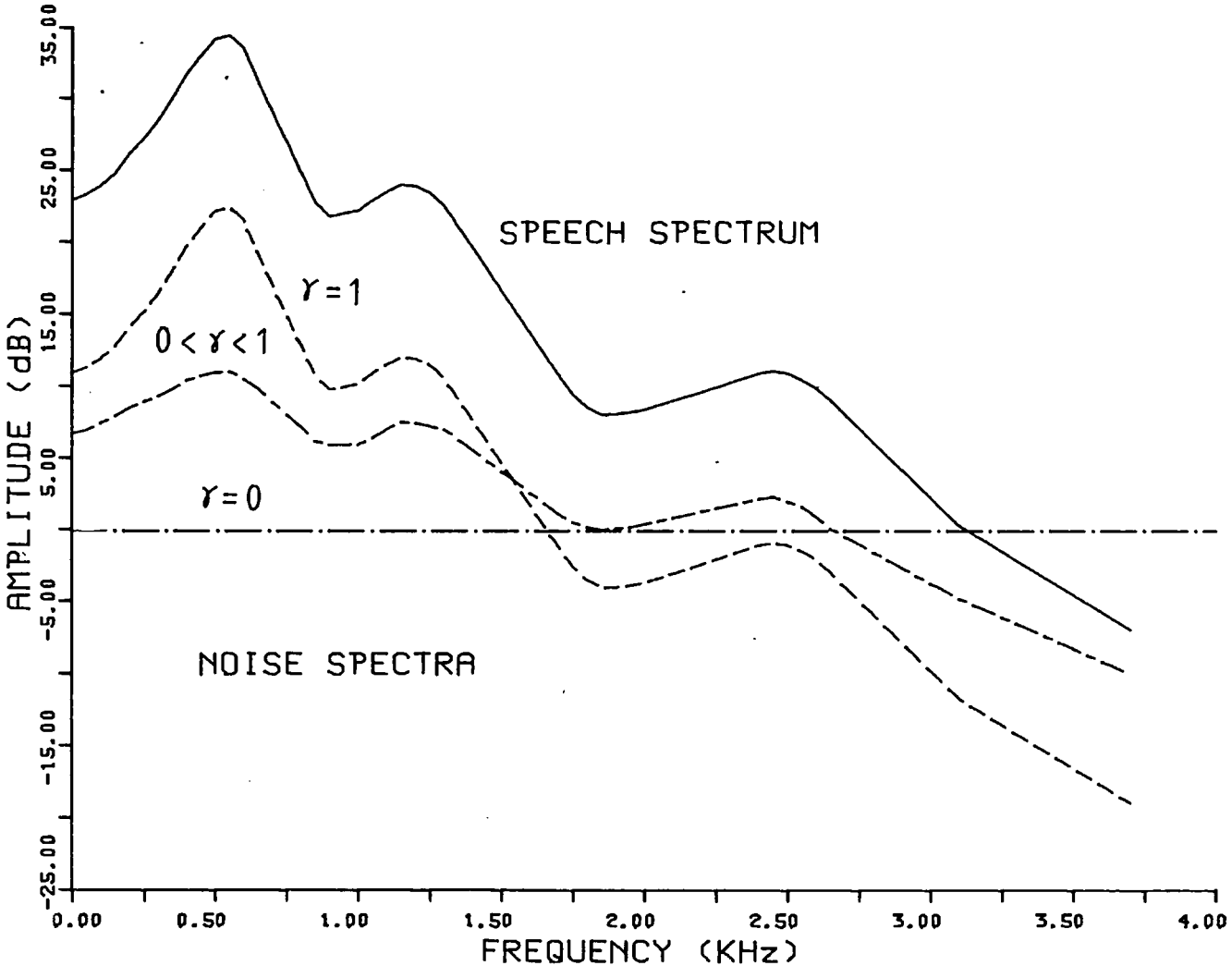
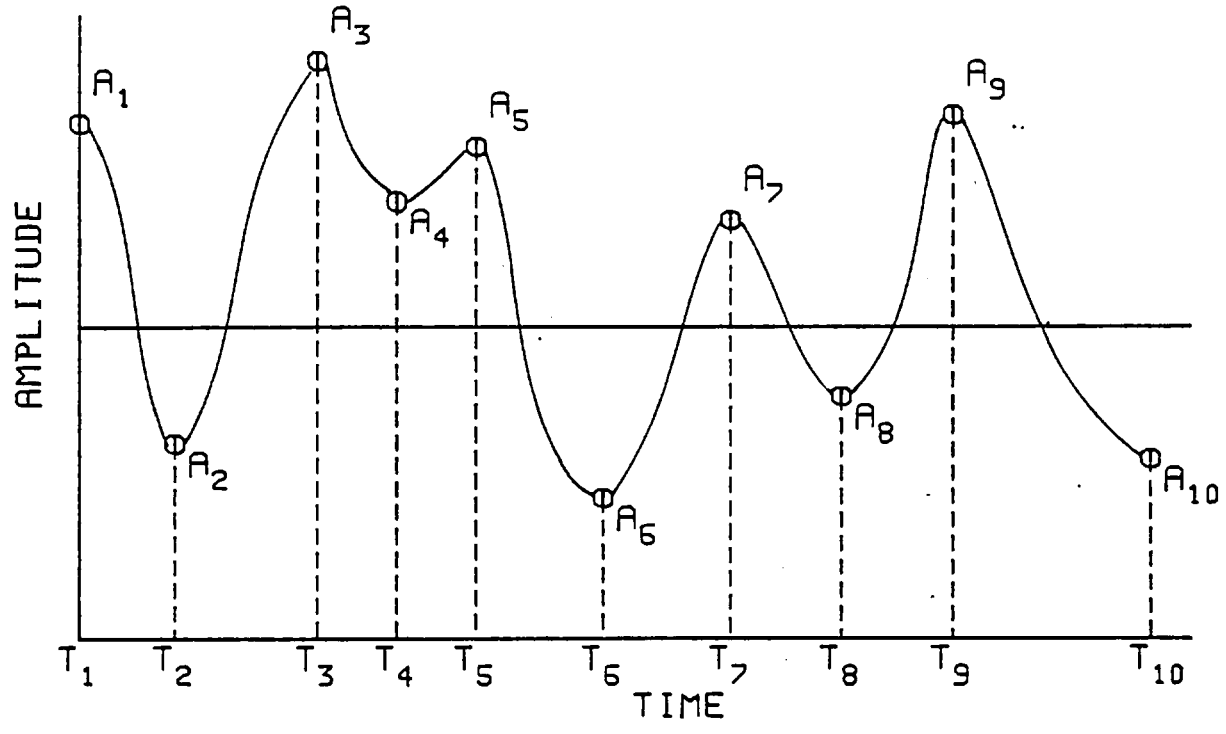
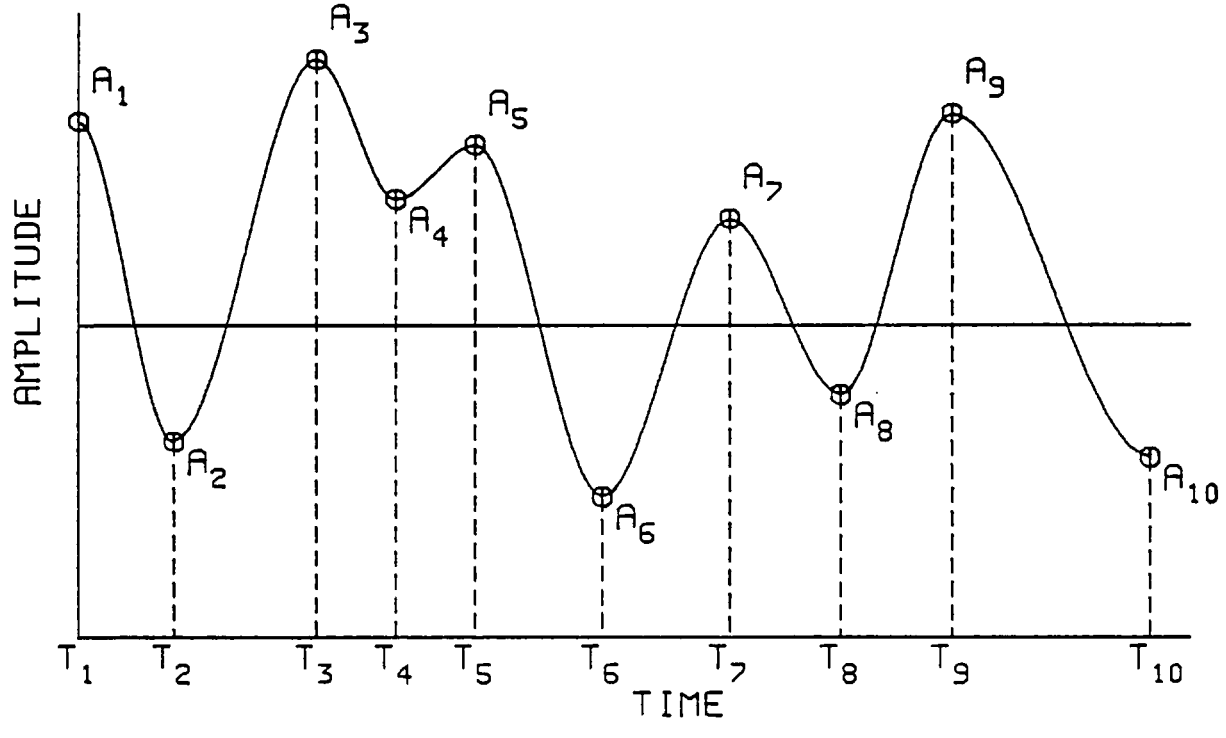


Fig.2.9 Variation of noise spectrum with parameter γ.



(a)



(b)

Fig.2.10 Illustration of extremal coding; (a) Original waveform; (b) Interpolated waveform.

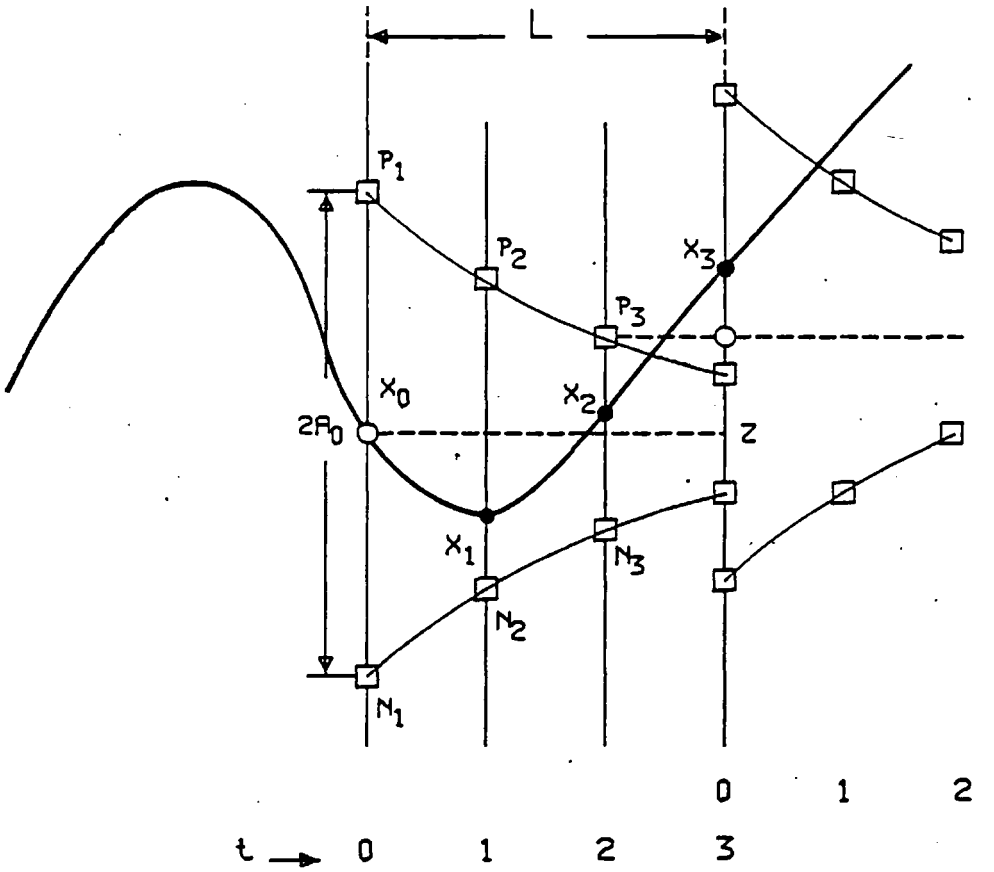


Fig.2.11 Illustration of aperture coding.

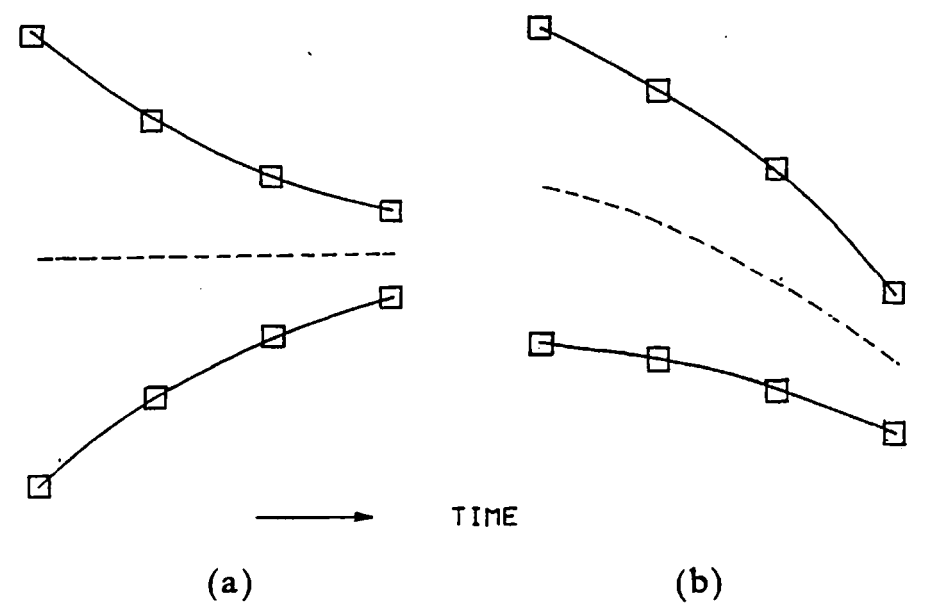


Fig.2.12 Aperture with (a) first-order prediction : $a_1 = 1$;
(b) more general prediction.

CHAPTER 3

FIXED-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS

CHAPTER 3

FIXED-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS

3.1 INTRODUCTION

In this chapter, a fixed-rate hybrid frequency-domain coding technique is discussed. In this scheme, blocks of speech signals are coded using a fixed number of bits.

A schematic diagram of the proposed hybrid frequency-domain coder is shown in Fig.3.1. The speech spectrum is partitioned into four contiguous bands labelled B1,B2,B3 and B4 as shown in Fig.3.2, with sub-bands B1 and B2 having narrower bandwidths than sub-bands B3 and B4. The sub-band signals are lowpass translated and downsampled at the Shannon sampling rate [17]. The sampled signals in the bands B1 and B2 are coded using an adaptive quantizer and the signals in the bands B3 and B4 are separately transformed using a fast transform algorithm after which they are coded in the transform domain.

At the receiver, the sub-band signals are decoded: B1 and B2 in the time domain and B3 and B4 in the transform domain. The decoded signals are then bandpass translated back to the original bands and the speech signal is regenerated by the summation of these signals.

For additional adaptation to the speech signal, and better utilization of the available number of bits, a voiced/unvoiced

detector can be used with the bits being assigned according to whether the block is classified as voiced or unvoiced.

Although in this chapter, attention is focused on hybrid frequency-domain coding at bit rates of 7.2kbit/s and 9.6kbit/s, the method can also be applied effectively for higher bit rates, though tests would be needed in order to determine how the resulting speech quality compared with other techniques.

Several important issues have to be considered in the design of the hybrid frequency-domain coder. These include: (i) the partitioning of the speech spectrum and the design of the filter banks, (ii) the allocation of bits to the various sub-bands, (iii) the choice of transform, (iv) the transform block size, and (v) the method of quantizing the signals. These aspects of design are considered in Sections 3.2 to 3.6. The subject of system delay is discussed in Section 3.7 and in Section 3.8, the results of subjective tests are presented in which the hybrid frequency-domain coding system is compared with continuously variable slope delta modulation (CVSD).

3.2 THE PARTITIONING OF THE SPECTRUM AND THE DESIGN OF THE FILTER BANKS

Preferably, the partitioning of the speech spectrum should be carried out according to a perceptual criterion. Crochiere et al [57] proposed that the partitioning be carried out such that an equal contribution to the articulation index [99] is made by the signals in each sub-band.

Several methods are available for carrying out the band-splitting, the associated downsampling, and the upsampling. Among them, integer-band sampling and quadrature mirror filtering are the most attractive from an implementation point of view.

In the design using quadrature mirror filters, spectrum partitioning can be carried out using a series of band-splitting operations. In each band-splitting operation, the signal is filtered into lower and upper bands by two filters which are mirror images of each other. The signals in the two bands are then downsampled to half the input sampling rate. Quadrature mirror filters are structured so that frequency aliasing resulting from down-sampling is removed from the reconstructed signal. On account of this aliasing cancellation, lower-order filters with relatively wide transition widths can be used to cover the entire speech band of interest. Although in theory aliasing can be removed through the use of quadrature mirror filters, it is shown in Chapter 6 that when signal quantization is employed, significant aliasing can still remain. However, the use of the hybrid frequency-domain coding technique allows the aliasing components to be reduced.

For ease of implementation using quadrature mirror filters, the frequency ranges of the sub-bands used for the work described in this chapter are as shown in Table 3.1. In the table, the approximate percent contributions to the articulation index by signals of the sub-bands are also shown.

Table 3.1 Partition of Speech Band

Band	Frequency (Hz)	Sampling Rate (sample/s)	Approx. Percent Contribution to Articulation Index
B1	200-400	400	10
B2	400-800	800	15
B3	800-1600	1600	25
B4	1600-3200	3200	30
Total contribution to articulation index : approx.			80

With the partitioning as given in Table 3.1, the system is easy to implement when the input speech signal is sampled at 6400 sample/s. A block diagram of a scheme for band splitting using a cascade of quadrature mirror filters is shown in Fig.3.3.

3.3 THE ALLOCATION OF BITS TO SUB-BANDS

It has been found [5] that, for high-quality speech, on the average the signal energy in bands B1 and B2 are very much higher than those in bands B3 and B4, and that it is only during periods of unvoiced sounds that the energy in the upper bands are higher than those in the lower bands. Thus, since voiced sounds tend to predominate, and since coarse quantization of voiced sounds is more perceptible than coarse quantization of unvoiced sounds, it is sensible to assign a greater number of bits per sample to the lower bands.

A theoretical guideline for the allocation of bits to the sub-bands based on the average energy of the sub-band signals is given in the following.

In a given time segment, let the numbers of signal samples in the four frequency bands be N_1, N_2, N_3 and N_4 respectively. Let $\{y_{mj}(i), i=0,\ldots,N_j-1; j=1,\ldots,4\}$ be the set representing the N_j samples of the m^{th} segment of signal of the j^{th} band and $\{\hat{y}_{mj}(i), i=0,\ldots,N_j-1; j=1,\ldots,4\}$ be the set representing the N_j samples of the corresponding segment of signal quantized with an average of b_j bit per sample. The average mean-square signal energy, S_j^2, and the average mean-square distortion, E_j^2, within segments of the j^{th} band are defined as

$$S_j^2 = \frac{1}{N_j} \sum_{i=0}^{N_j-1} y_{mi}^2 \tag{3.1}$$

$$E_j^2 = \frac{1}{N_j} \sum_{i=0}^{N_j-1} (y_{mi} - \hat{y}_{mi})^2 \tag{3.2}$$

Assume that the average signal-to-noise ratio is proportional to $2^{(2b_j)}$, then

$$\frac{S_j^2}{E_j^2} = K2^{(2b_j)} \tag{3.3}$$

where K is a constant.

Let the total number of bits available for encoding the total number of samples from all the four bands be b_t, then

$$b_t = \sum_{j=1}^{4} N_j b_j \tag{3.4}$$

From Equation (3.3), the average distortion of the j^{th} band can be written as

$$E_j^2 = S_j^2 K^{-1} 2^{(-2b_j)} \tag{3.5}$$

and the total distortion, E^2, of the signal is then given by

$$E^2 = \sum_{j=1}^{4} N_j E_j^2 = \sum_{j=1}^{4} N_j S_j^2 K^{-1} 2^{(-2b_j)} \tag{3.6}$$

By denoting $X_j = K^{-1}2^{(-2b_j)}$, the problem can be formulated as the minimization of the overall distortion E, where

$$E^2 = \sum_{j=1}^{4} N_j S_j^2 X_j \tag{3.7}$$

subject to the constraint

$$\prod_{j=1}^{4} X_j^{N_j} = \prod_{j=1}^{4} K^{-1}2^{(-2N_j b_j)} = K^{-4}2^{(-2b_t)} = \text{constant} = C \tag{3.8}$$

This minimization problem can be solved by using Lagrange multipliers. The solution of the set of equations

$$\frac{d}{dX_j}(E^2 + \lambda C) = 0 \qquad j=1,..,4 \tag{3.9}$$

gives

$$S_j^2 X_j = -\lambda C \tag{3.10}$$

from which it can easily be shown that

$$b_j = \frac{b_t}{N} + \frac{1}{2}\log_2 S_j^2 - \frac{1}{2N}\sum_{i=1}^{4} N_i \log_2 S_i^2 \tag{3.11}$$

where $N = N_1+N_2+N_3+N_4$.

As the long-term average energy is more concentrated in bands B1 and B2, it is clear from Equation (3.11) that a greater number of bits per sample should be allocated to the lower two bands than the upper two bands; and since the energy distribution is biased towards the upper bands for the unvoiced segment, a voiced/unvoiced detector can be incorporated to alter the bit assignment. Subjective tests have shown (Appendix C) that the quality of the regenerated speech is slightly improved when the voiced and unvoiced segments are coded differently.

3.4 THE SELECTION OF TRANSFORM

Of the orthogonal transforms that have fast computation algorithms, the discrete cosine transform and the Walsh-Hadamard transform have been investigated as a means for encoding the signals in bands B3 and B4.

As a means of obtaining an objective measure of the comparative performance of the discrete cosine transform and the Walsh-Hadamard transform in the coding of signals of the upper sub-bands, the average segmental signal-to-noise ratio as described by Noll [100] has been used. Average segmental signal-to-noise ratio was chosen in preference to the straight-forward signal-to-noise ratio since it has been found [100] to relate more closely to subjective performance. A comparison of the performance of the two transforms based on the average segmental signal-to-noise ratio at different bit rates is given in Figs.3.4 and 3.5. The speech material used consists of 5 seconds of male speech and 5 seconds of female speech. The average segmental signal-to-noise ratio was computed in the following manner.

Considering a particular frequency band (B3 or B4), let the input samples of a segment of the sub-band signals be represented by $\{x_i, i=0,\ldots,N_j-1\}$, and the corresponding estimated output values be $\{\hat{x}_i, i=0,\ldots,N_j-1\}$. If the distortion e_i is defined by $e_i = (x_i - \hat{x}_i)$, then the segmental signal-to-noise ratio, $SSNR_m$, associated with the m^{th} segment is

$$SSNR_m = 10\log\left(\sum_{i=0}^{N_j-1} x_i^2 / \sum_{i=0}^{N_j-1} e_i^2 \right) \qquad (3.12)$$

and the average over M segments is

$$SSNR = \frac{1}{M} \sum_{m=1}^{M} SSNR_m \qquad (3.13)$$

For signals representing silence, the signal energy is essentially zero, with the result that a large negative value of $SSNR_m$ may be obtained. This problem is overcome by excluding those segments that have low energy. Also, to prevent the average value being dominated by large values of $SSNR_m$, a lower and an upper limit are imposed on the $SSNR_m$ value. For the measurements carried out, the $SSNR_m$ values were constrained to lie in the range 0 dB to 50 dB. Since the bandwidth of band B4 is twice that of band B3, the number of samples corresponding to a given speech segment is twice as large in the case of band B4 as in the case of band B3. Block sizes of 64 samples and 128 samples are used in Figs.3.4 and 3.5 for bands B3 and B4 respectively. These correspond to a speech segment of 40 ms. In these figures, the number of bits allocated to the block is marked along the horizontal axis. The allocated number of bits does not include those bits required to send information relating to the envelope of the transform coefficients, that is, the side-information. Details of the extraction and encoding of the side-information are given in Section 3.6.

The average segmental signal-to-noise ratios (SSNR's) obtained by using the Walsh-Hadamard transform (WHT) and the discrete cosine transform (DCT) are shown in Figs.3.4 and 3.5. Fig.3.4 shows the SSNR's of signals in band B3 and Fig.3.5 the SSNR's of signals in band B4. In Figs.3.4(a) and 3.5(a), three different SSNR curves are shown for each transform technique. The first curve, which is labelled 'exact', deals with the case in which the true values of the transform coefficients were used to control the bit allocation and quantization step-size. The other two curves which are labelled '9 grp' and '5 grp'

deal with the cases in which the unquantized side-information obtained by grouping the transform coefficients into 9 groups and 5 groups respectively were used for bit allocation and control of quantization step-size. In Figs.3.4(b) and 3.5(b), the SSNR's were obtained when quantized side-information were used for bit allocation and step-size control.

An investigation into the time dependence of the r.m.s. energy and $SSNR_m$'s of the sub-band signals for one sentence was also carried out and the results are presented in Figs.3.6 and 3.7 for band B3 and band B4 respectively. The sentence used was 'Airmail requires a special stamp' spoken by a male speaker.

Figs.3.6(a) and 3.7(a) give the r.m.s. energy of sub-band signals as a function of time. Figs.3.6(b-d) and 3.7(b-d) give the variation of $SSNR_m$'s as a function of time for sub-band signals coded with the DCT and the WHT using different methods of bit allocation and step-size adaptation. From these figures, it can be seen that for both bands B3 and B4, the $SSNR_m$'s associated with the WHT do exceed those associated with the DCT for segments with different r.m.s. energy.

Using a tenth-order Markov process that simulates full-band speech signals, and with the magnitude of signal-to-noise gain over PCM as a criterion, Zelinski and Noll have shown [81] that for a block size of 128 samples, the gain achieved with the DCT is slightly above 9 dB whereas the gain obtained with the WHT is only 3 dB. In this case, the gain over PCM is calculated as the ratio of the arithmetic mean to the geometric mean of the variances of the transform coefficients.

However, if real speech are used and the signals within sub-bands are considered, then, when using the adaptive bit allocation and quantization methods as described in Section 3.6, it can be seen from Figs.3.4 to 3.7 that the SSNR's associated with the Walsh-Hadamard transform and the discrete cosine transform are similar for the range of bit rates under consideration. The figures also show that when using 9-group or 5-group side-information, in some cases, the Walsh-Hadamard transform appears to give a slightly larger SSNR than the discrete cosine transform.

The different results obtained for sub-band signals and full-band signals can be explained as follows.

(a) The SSNR's obtained with full-band speech signals coded using the DCT and the WHT are presented in Fig.3.8. The bit allocation and step-size adaptation were carried out using the estimated spectral envelope described by Zelinski and Noll [81]. From Fig.3.8, it can be seen that even though the SSNR's of signals coded with the DCT are consistently higher than those of signals coded with the WHT, the improvement is not as much as in the case when the SNR gain over PCM is used as a criterion.

(b) For sub-band signals, the correlations among the signal samples are lower than those for full-band signals, and using the estimated spectral envelope for bit allocation and step-size control as described in Section 3.6, the SSNR's of signals coded with the WHT can be higher than those of signals coded with the DCT.

As the Walsh-Hadamard transform is computationally very much simpler than the discrete cosine transform, it was used in the design

of the hybrid frequency-domain coder. The Walsh-ordered Walsh-Hadamard transform was used rather than the Hadamard-ordered transform since sequency order is involved in the Walsh-ordered transform and this results in a smoother envelope similar to the normal frequency spectrum of the signal. As explained in [81], a smooth envelope is desirable for the adaptive coding of the transform coefficients.

3.5 THE SELECTION OF BLOCK SIZE

The effect of block size on the SSNR for transform coding of signals in bands B3 and B4 will now be considered.

The SSNR values obtained through transform coding of signals in bands B3 and B4 shown in Fig.3.9, are expressed as a function of block size. It can be seen that the increase in SSNR for increasing block size is quite small. It is likely that this is due to the low correlation of the signals in the sub-bands. In the investigations reported in this chapter, speech segments of 40 ms were used. The use of long segments results in long system delay and requires a large size buffer. The use of short segments, on the other hand, does not permit the efficient use of transform techniques.

3.6 THE METHODS OF SIGNAL QUANTIZATION

3.6.1 Quantization of Signals in Bands B1 and B2

For signals in bands B1 and B2, the numbers of samples in a segment of 40 ms are small (16 and 32 samples), and as the sample-to-sample correlation of the signals in the sub-bands is low,

encoding in bands B1 and B2 is done using an adaptive quantizer of the type proposed by Jayant [64]. With this method, the quantization step-size is adaptively adjusted for each new sample according to the level of the previous sample value. Mathematically, if x_k is the k^{th} sample to be quantized and if Δ_k represents the step-size to be used for this sample, the quantization level, L_k, taken by the k^{th} sample with this step-size, is given by

$$L_k = \text{Integer part of } (\frac{x_k}{\Delta_k}) + 1 \qquad (3.14)$$

The step-size for the next sample is calculated as

$$\Delta_{k+1} = M(L_k)\Delta_k \qquad (3.15)$$

The step-size is constrained to lie within a specified minimum Δ_{min} and a specified maximum Δ_{max}, that is,

$$\Delta_{min} \leqslant \Delta_i \leqslant \Delta_{max} \qquad \text{for all } i \qquad (3.16)$$

The decoded value of x_k is then given by

$$x_k = (L_k - 0.5)\Delta_k \qquad (3.17)$$

In Equation (3.15), $M(L_k)$ is a multiplication factor which is a function of L_k. $M(L_k)$ is also dependent on the type of signal and the number of bits used for the quantization. The values of $M(L_k)$ obtained experimentally for the signals in bands B1 and B2 are given in Table 3.2. This set of values maximizes the long-term signal-to-noise ratio of the signals.

Table 3.2 Multiplication Factor for Step-Size Adaptation

Bits per Sample	M_1	M_2	M_3	M_4
3	0.85	1.0	1.0	1.5
2	0.82	1.6	-	-
1.5	0.92	1.4	-	-

In order to improve the idle channel performance, a mid-rise/mid-tread switch [101] is also used in the decoder. The switching from a mid-rise quantizer characteristics to a mid-tread quantizer characteristics occurs when the step-size Δ falls below a threshold set equal to $2\Delta_{min}$. That is, when $\Delta \leqslant 2\Delta_{min}$, the two lowest levels of the mid-rise quantizer in the decoder are switched to zero.

3.6.2 Quantization of Signals in Bands B3 and B4

For the coding of the transform coefficients of the signals in bands B3 and B4, a method similar to that introduced by Zelinski and Noll [81] has been adopted. The data blocks $\{x_{1i}, i=0,\ldots,N_3-1\}$ and $\{x_{2j}, j=0,\ldots,N_4-1\}$ of signals in bands B3 and B4 are separately transformed using the Walsh-Hadamard transform to obtain the sets $\{y_{1i}\}$ and $\{y_{2j}\}$ respectively as shown in Fig.3.10(a). The squares of these values give the variances $\{y_{1i}^2\}$ and $\{y_{2j}^2\}$. The spectrum is partitioned into n groups. For the case in which n=5, the grouping is as shown in Fig.3.10(b). For each of the bands B3 and B4, an estimate of the envelope of the spectrum can be made using the respective average values, g_{1l}^2 and g_{2l}^2, $l=1,\ldots,5$, of the variances in each group. The encoded values of $\{g_{1l}^2\}$ and $\{g_{2l}^2\}$ form the side-information which is transmitted, together with the quantized transform coefficients. At the receiver, the envelopes of the spectra of the

signals in the bands B3 and B4 are estimated using the side-information. At both transmitter and receiver the side-information is used both in the determination of the optimum number of bits to be used for encoding the individual coefficient and in controlling the quantizer step-size.

In order to obtain smoother approximations to the transform spectrum, the spectrum shown in Fig.3.10(c) is obtained by first taking the logarithm of the g_{kl}^2, for k=1,2; l =1,...,5, and then estimating all variances by using a simple straight line interpolation between the logarithms of the appropriate g_{kl}^2.

Instead of allocating a fixed number of bits per block to each of the bands B3 and B4 and carrying out the bit assignment in the two bands separately, the bands can be considered as a whole and the bits distributed accordingly. When the two bands are considered together, allocation of bits to the i^{th} transform coefficients, b_i, can be made according to the equation

$$b_i = \bar{b} + \frac{1}{2}\log_2(\sigma_i^2) - \frac{1}{2N_u}\sum_{j=1}^{N_u}\log_2(\sigma_j^2) \qquad (3.18)$$

where, for a given time segment, N_u is the total number of samples in the two bands, $\bar{b}$ is the average number of bits per sample allocated to the two bands, and σ_i^2 denotes the variance of the i^{th} coefficient. Equation (3.18) can be derived by applying the Rate Distortion Theory [102]. The bit allocation scheme based on Equation (3.18) results in a quantization noise that is flat across the spectrum. It has been found that, from the point of view of speech perception, this method of allocation is not the most effective and the subjective quality can be improved if a weighting factor is introduced. For full-band speech

coding, Tribolet and Crochiere have found [6] that if all the estimated variances σ_i^2 of the coefficients are modified by their respective multiplicative factors $\sigma_i^{-2\gamma}$, with $\gamma=0.125$, then improved speech quality can be obtained.

In assigning bits, Tribolet and Crochiere's modification was used. In addition, since the resulting values, b_i, calculated may not be an integer, it is necessary to adjust the values. This was done in accordance with the reoptimization procedure of Zelinski and Noll [103].

3.6.3 Coding of Side-Information

Let the average variances of the blocks of samples from bands B3 and B4 be denoted by $\bar{y}_1^2$ and $\bar{y}_2^2$ respectively, where

$$\bar{y}_1^2 = (1/N_3) \sum_{i=0}^{N_3-1} y_{1i}^2 \tag{3.19}$$

and

$$\bar{y}_2^2 = (1/N_4) \sum_{j=0}^{N_4-1} y_{2j}^2 \tag{3.20}$$

in which N_3 and N_4 are the numbers of samples per block in the respective bands. A study of the long-term statistics of speech signals with the amplitudes of the signals being restricted to the range -2048 to 2047, and with 'silent' segments being excluded, has shown that the logarithms of $\bar{y}_1$ and $\bar{y}_2$ lie between -1.6 and 1.6.

Let the ratio of the standard deviation of the transform coefficients within a group to the standard deviation of the transform coefficients of the complete block, that is, $g_{kl}/\bar{y}_k$, for k=1,2; l=1,...,5, be denoted by R_{kl}. The distributions of R_{kl}'s, determined using the same speech material as mentioned in Section 3.4, are

plotted in Fig.3.11(a) for band B3 and in Fig.3.11(b) for band B4. From these figures, it can be seen that the distributions can be approximated by the positive part of the normalized Gaussian distribution and thus a Gaussian quantizer adapted from that proposed by Max [104] was used.

In the simulation of the hybrid frequency-domain coding system, the encoding of the side-information was performed by first quantizing the logarithm of each of $\bar{y}_1$ and $\bar{y}_2$ using a 5-bit linear quantizer and then, for each of the five groups, using a 2-bit Gaussian quantizer to code the ratio R_{kl}. The block average standard deviation used in determining the above ratio was obtained by taking the anti-logarithm of the corresponding quantized value of $\log(\bar{y}_1)$ or $\log(\bar{y}_2)$.

The effect resulting from the use of exact side-information and from the quantized side-information can be seen from Figs.3.4 and 3.5. From Figs.3.4(a) and 3.5(a), it is clear that if side-information rather than 'exact' coefficient information is used, a loss in SSNR of approximately 3 dB results for both the Walsh-Hadamard and the discrete cosine transforms. Comparisons of Figs.3.4(a) and 3.4(b) and Figs.3.5(a) and 3.5(b) respectively shows that if encoded rather than exact side-information is used then a further loss in SSNR of approximately 1 dB is incurred for both transforms. The figures also show that a slight loss is involved if 5-group side-information rather than 9-group side-information is used. This loss is so little that it is sensible to use 5-group side-information rather than 9-group side-information. By using 5-group rather than 9-group side-information, bits can be saved, and the bits so saved can be used to provide more accurate encoding of the transform coefficients.

3.6.4 Quantizer for Transform Coefficients

The probability distributions of the transform coefficients in bands B3 and B4, normalized by using the respective standard deviation in each block, are shown in Figs.3.12(a) and 3.12(b) and the appropriately normalized Gaussian and Laplacian distributions are also plotted in the figures. Using the same speech material as for Figs.3.4 and 3.5, the SSNR's obtained using Gaussian and Laplacian quantizers for different numbers of bits allocated per block are shown in Figs.3.13(a) and 3.13(b). From these figures, it is clear that the Gaussian quantizer gives better results.

3.7 SYSTEM DELAY CONSIDERATION

The filtering operations used in the hybrid frequency-domain coding system are shown in Figs.3.3(a) and 3.3(b). Fig.3.3(a) relates to the transmitter part of the system and Fig.3.3(b) to the receiver. These filtering operations are the main sources of signal delay encountered with the system and this section is devoted primarily to a consideration of these filtering delays. From Fig.3.3(a), it can be seen that in order to obtain signals in the band B1, the input signal has to pass through a cascade of four quadrature mirror filters. To obtain signals in the band B4, the input signal has to pass through only a single filter. The result of this is that signals in the various bands are delayed by different amounts.

If the delay associated with the k^{th} stage of quadrature mirror filtering, d_k, is expressed in terms of the number of samples by which the signal is delayed, then for a quadrature mirror filter of length

L_k, $d_k=(L_k-1)/2$. Since L_k is constrained to be an even integer, d_k is not an integer.

If at the transmitter, encoded signals drawn from the same time segment are to be multiplexed together, then in order to maintain time synchronisation, it is necessary that the signals within the separate sub-bands are delayed by different amounts.

Let D_i be the number of additional samples by which a signal in the i^{th} band is delayed prior to transmission, and let D_i be chosen so that the total time delay ∂_i associated with the i^{th} band is such that $\partial_1 = \partial_2 = \partial_3 = \partial_4$. Thus,

$$\partial_i = [\sum_{k=1}^{5-i} d_k T_{(k-1)}] + D_i T_{(5-i)} \qquad i=1,\ldots,4 \qquad (3.21)$$

where T_k, $k=1,\ldots,4$ is the sampling interval associated with the downsampled output from the k^{th} stage of quadrature mirror filtering and T_0 is the sampling interval relating to the original input signal. The interval T_k, $k=1,\ldots,4$ is related to T_0 by $T_4=2T_3=4T_2=8T_1=16T_0$.

Now, suppose that $\partial_4 = \partial_3$

then $$d_1T_0 + D_4T_1 = d_1T_0 + d_2T_1 + D_3T_2$$

which simplifies to $$2(D_4-2D_3) = L_2-1 \qquad (3.22)$$

Since L_k, $k=1,\ldots,4$ is an even integer, it is clear that Equation (3.22) can never be satisfied and hence it is not possible to make all ∂_i, $i=1,\ldots,4$ equal. In view of this, the best that can be done is to minimize the difference between ∂_i and ∂_j for all i,j; $i \neq j$.

If it is required that the multiplexed data contain one block of encoded data from each band, then since the filtering delay suffered by signals in band B1 is the longest, the overall system delay can be minimized by choosing the additional delay D_1 such that $D_1=(T_b/T_4)=D_b$, where T_b is the duration of the data block. For $L_k=32$, $k=1,\ldots,4$, the smallest difference between ∂_i and ∂_j for all i,j; $i\neq j$, can then be achieved by setting $D_1=D_b$, $D_2=15+2D_b$, $D_3=46+4D_b$ and $D_4=108+8D_b$. With this selection, $\partial_1=T_b+232.5T_0$, $\partial_2=T_b+228.5T_0$, $\partial_3=T_b+230.5T_0$ and $\partial_4=T_b+231.5T_0$.

It should be noted that the performance of the system is unaffected by the small differences between ∂_i and ∂_j for all i,j with $i \neq j$.

At the receiver, which is shown in Fig.3.3(b), delay adjustment has also to be carried out before the signals from the different bands are added together to arrive at the resultant output signal. Let the delay associated with signals received from the i^{th} band be ρ_i s with respect to the original signal. If the received signals in the bands B1,B2,B3 and B4 are delayed by D_1',D_2',D_3' and D_4' samples respectively before interpolation and subsequent filtering, then provided that:

$$D_1' = 0 \tag{3.23}$$

$$D_2' = [(\rho_1 - \rho_2)/8T_0] + d_4 \tag{3.24}$$

$$D_3' = [(\rho_1 - \rho_3)/4T_0] + 2d_4 + d_3 \tag{3.25}$$

and

$$D_4' = [(\rho_1 - \rho_4)/2T_0] + 4d_4 + 2d_3 + d_2 \tag{3.26}$$

the overall end-to-end delay will be the same for signals in all bands.

At the receiver, block processing of signals in bands B3 and B4 can only commence after a complete block of encoded data is received. The magnitude of this delay depends on the way in which the transmitted data are multiplexed. If it is assumed that the transmission rate is the same as the rate of data generation at the source, then the maximum additional delay is equal to the duration of one segment of data, that is, T_b.

Thus, if the encoding and decoding time is neglected, the maximum overall system delay is twice the sum of the filtering delay and the segment duration, that is, $2[(d_1 + 2d_2 + 4d_3 + 8d_4)T_0 + T_b]$. In the system reported on in this paper, $d_1=d_2=d_3=d_4=(32-1)/2$ and T_b=40 ms, so the maximum overall system delay is 153 ms.

In a practical implementation of the scheme, the delay could be considerably reduced by prefiltering the input and postfiltering the output full-band signals using a recursive highpass filter with cutoff frequency at 200 Hz. This would dispense with the need for quadrature mirror filters in stage 4 of both the transmitter and the receiver parts of the system. Reduction of delay could also be achieved both

through the use of alternative methods of multiplexing the encoded data and through the use of shorter-length filters in the higher stages.

3.8 RESULTS OF SUBJECTIVE TESTS

It is well known that when coding speech at low-bit-rates, objective criteria in general are unreliable measures of speech quality. With this in mind, a series of subjective listening tests were carried out in an attempt to provide a more meaningful statement of the quality of the speech produced by the hybrid frequency-domain coding system operating at bit rates of 7.2 kbit/s and 9.6 kbit/s.

Since little is available by way of standard speech material against which the hybrid frequency-domain coding system could be compared and through which the system could be compared with other speech coding systems, a decision was taken to compare the system with continuously variable slope delta modulation (CVSD) [105] operating at bit rates ranging from 14 kbit/s to 26 kbit/s. The reasons for this choice were : (i) the CVSD system is simple to implement; its bit rate can be varied in small increments; (ii) Crochiere et al [57] have previously compared sub-band coding systems with adaptive delta modulation.

Details of the CVSD system used are given in Appendix A. The bit assignments used in the four sub-bands of the hybrid frequency-domain coding system when it was operating at data rates of 7.2 kbit/s and 9.6 kbit/s are set out in Table 3.3 below.

Table 3.3 Bit Allocation for HFDC (bit/s)

Rate	Type	Band 1	Band 2	Band 3+4	Header+Side Infor.
7200	Voiced	800	1600	4000	800
	Unvoiced	600	1200	4600	800
9600	Voiced	1200	2400	5200	800
	Unvoiced	800	1600	6400	800

The subjective tests used to compare the hybrid frequency-domain coding system and CVSD coding system were designed as factorial experiments using a randomized balance design so that any effects not specifically included would not distort the results. Specifically, the tests were designed as 4X4 experiments with the factors being (i) the sentences used and (ii) the coder pairs $\{(C_0,C_i),\ i=1,\ldots,4\}$ where C_0 denotes the hybrid frequency-domain coder and C_i denotes the CVSD coder operating at one of the four bit rates.

Two sentences by a male speaker and two different sentences by a female speaker were used. Different sentences were used to reduce the number of repetitions and make listenings more interesting.

Two separate experiments were carried out each involving twelve listeners, with the same listeners being used in the two experiments. In each of the experiments, the listeners were asked to compare pairs of sentences and to indicate which of the two sentences they preferred.

In Experiment 1, the material used consists of each of the sentences encoded using the hybrid frequency-domain coder operating at 7.2-kbit/s, together with each of the sentences encoded using the CVSD system operating at bit rates of 14, 17, 20 and 23 kbit/s.

In Experiment 2, the material used consists of each of the sentences encoded using the hybrid frequency-domain coding system operating at 9.6-kbit/s, together with each of the sentences encoded using the CVSD system operating at 17, 20, 23 and 26 kbit/s.

During the course of each experiment, the group of listeners was presented with one sentence coded using the hybrid frequency-domain coding system and with the same sentence encoded using the CVSD system operating at one of the four bit rates. The order in which the pair of sentences were presented was randomized as was the order in which the basic sentence material was presented.

The results of the tests were analysed using the analysis of variance method and the effect of the different sentences on the comparison of coders was found not to be significant. The results of the analysis which are summarized in Tables 3.4 and 3.5 below, are presented in graphical form in Figs.3.14(a) and 3.14(b).

Table 3.4 Comparison of HFDC at 7.2 kbit/s and CVSD

Bit Rate of CVSD (kbit/s)	Percentage Preference For	
	CVSD	HFDC
14	18.75	81.25
17	31.25	68.75
20	58.33	41.67
23	85.42	14.58

Table 3.5 Comparison of HFDC at 9.6 kbit/s and CVSD

Bit Rate of CVSD (kbit/s)	Percentage Preference For	
	CVSD	HFDC
17	14.58	85.42
20	27.08	72.92
23	52.08	47.92
26	81.25	18.75

If the crossover point of the curves of listener preference is taken as the point at which the two coders have approximately equal subjective quality, then from Fig.3.14(a), it can be seen that the subjective quality of 7.2 kbit/s HFDC is comparable with 19 kbit/s CVSD and, from Fig.3.14(b), that 9.6 kbit/s HFDC is comparable with 23 kbit/s CVSD.

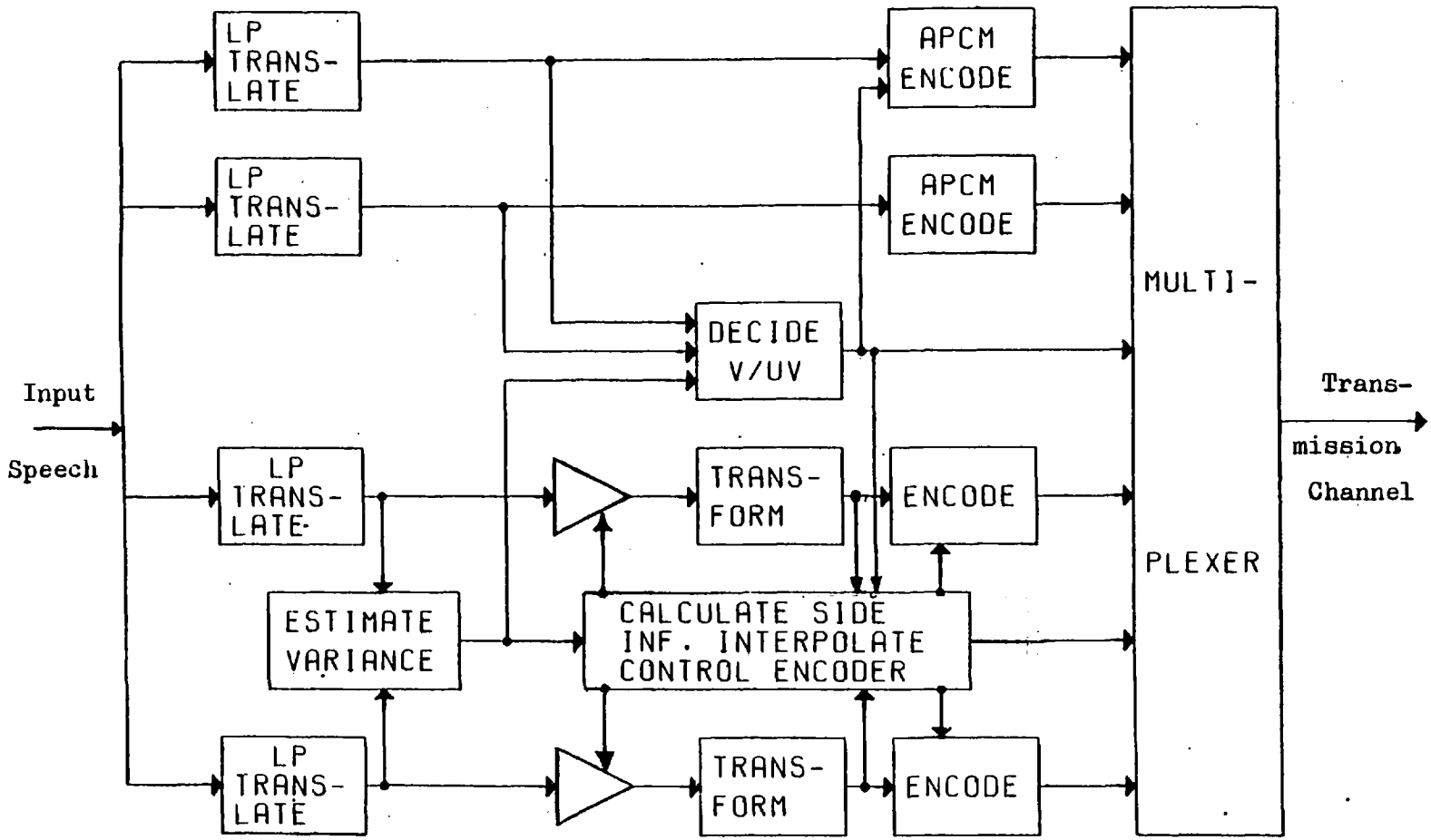


Fig.3.1 (a) Block diagram of the encoder for hybrid frequency-domain coding scheme.

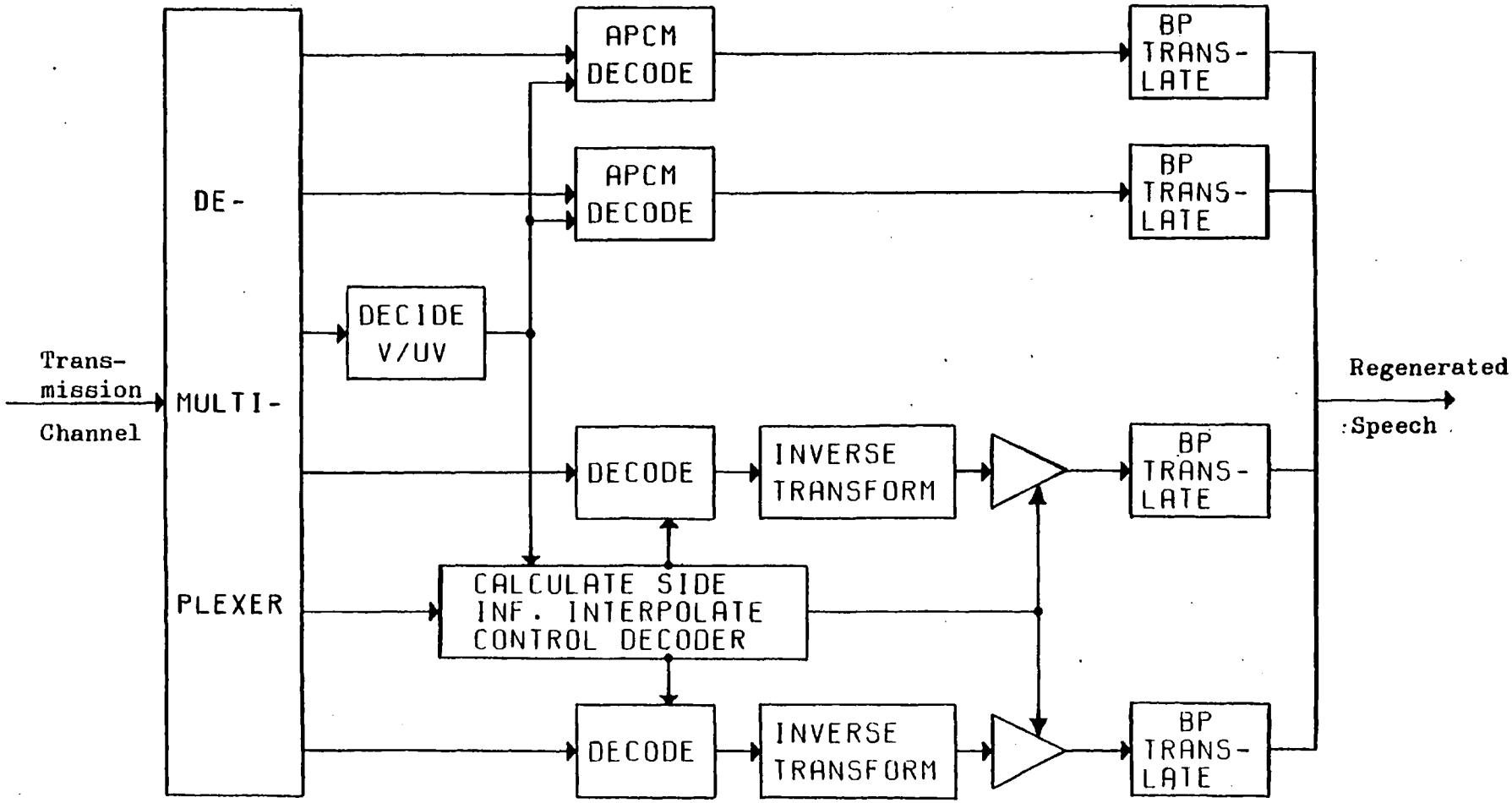


Fig.3.1 (b) Block diagram of the decoder for hybrid frequency-domain coding scheme.

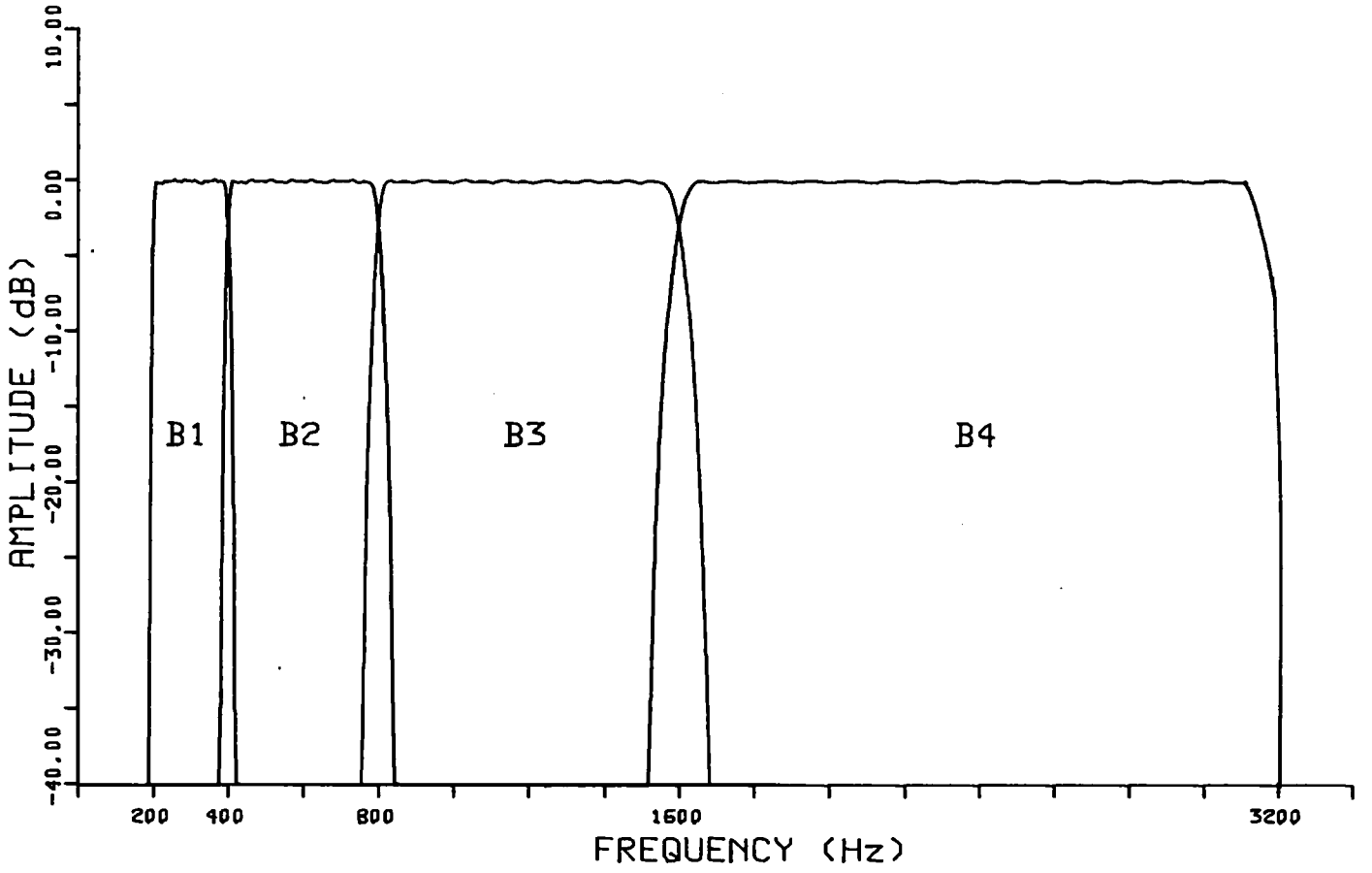


Fig.3.2 Frequency domain illustration of partition of speech spectrum.

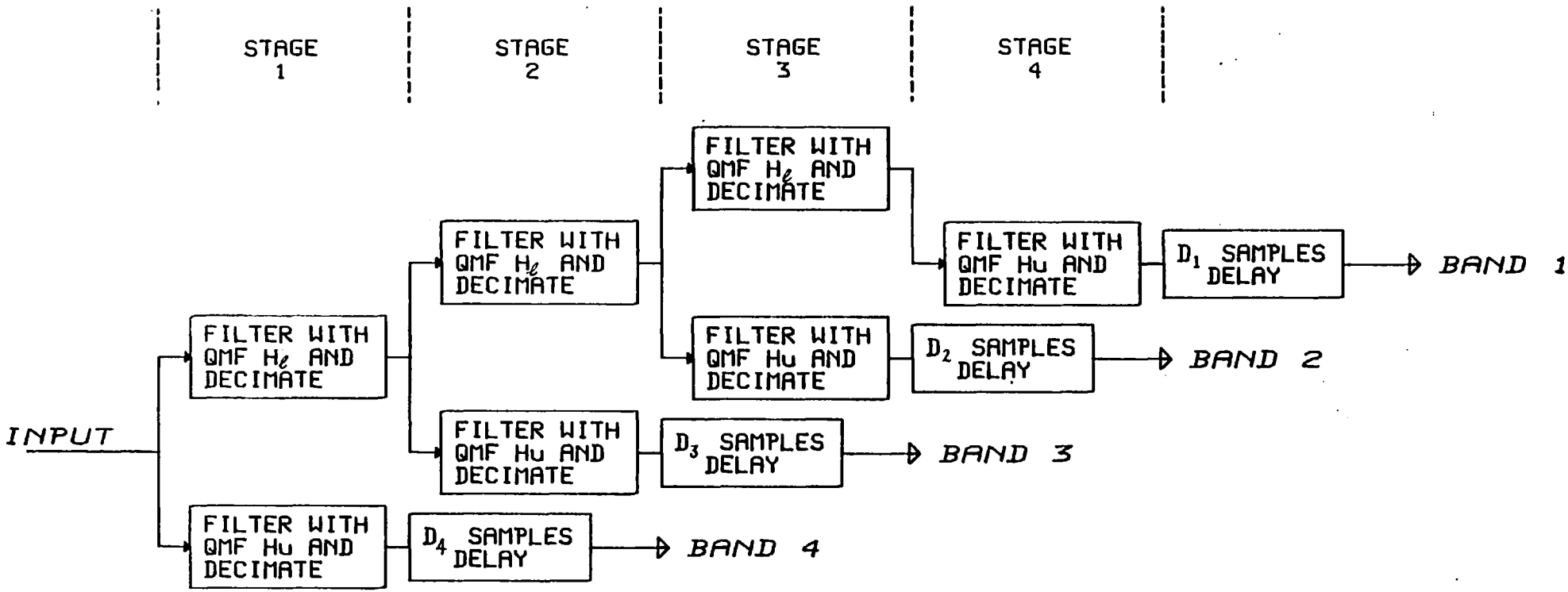


Fig.3.3 (a) Band-splitting at the transmitter.

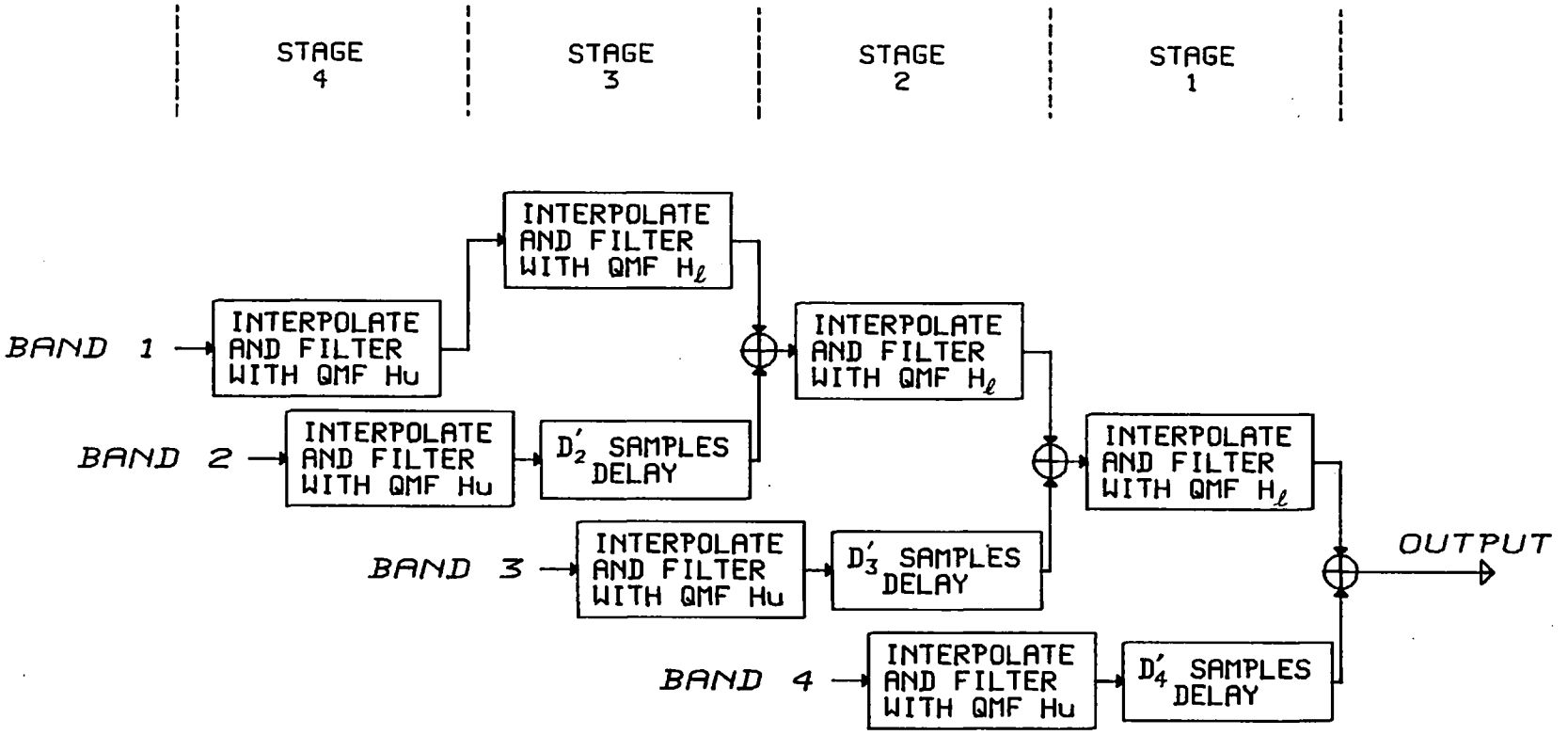


Fig.3.3 (b) Band-combining at the receiver.

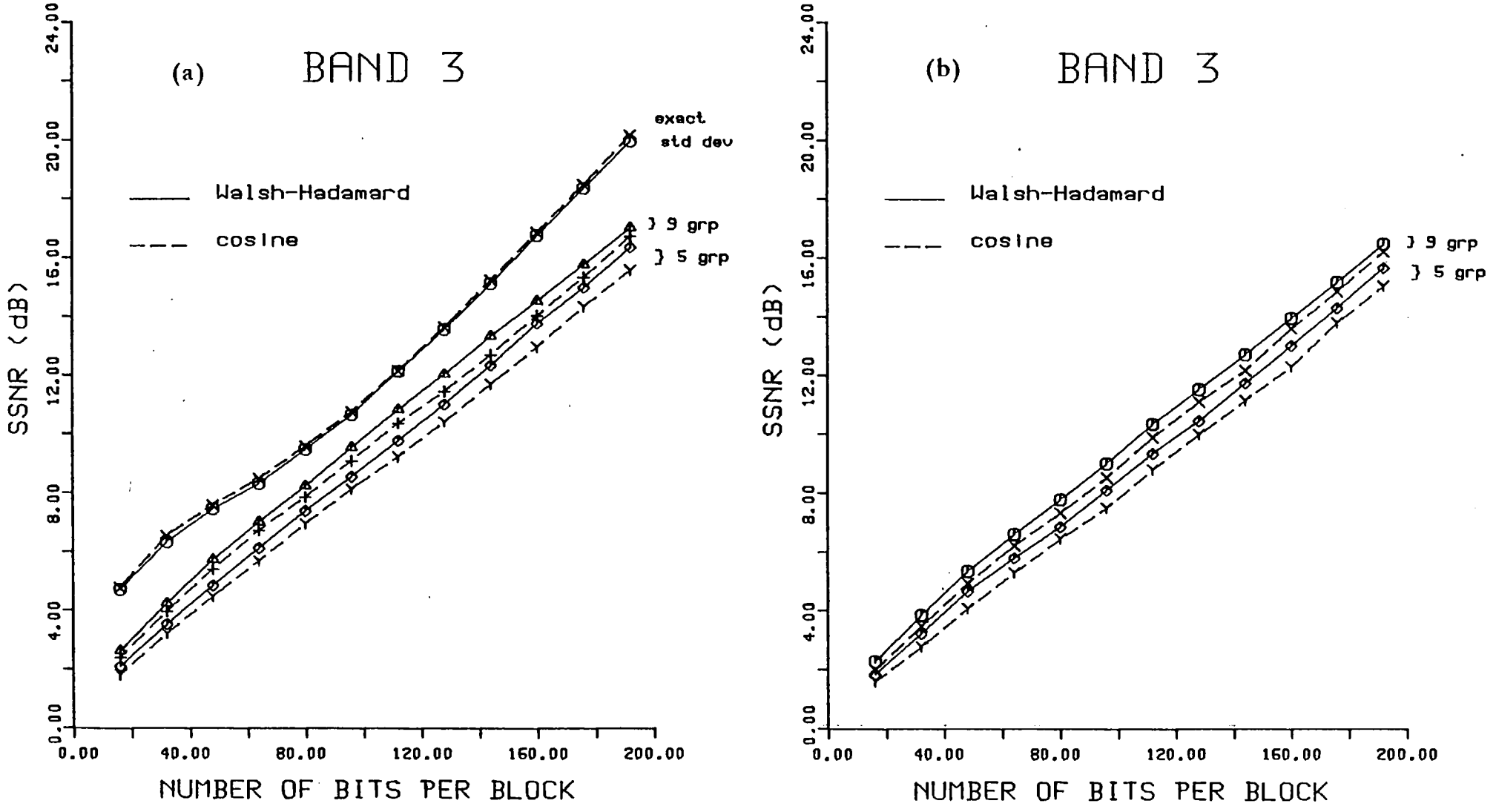


Fig.3.4 Effects of different transforms on SSNR of Band 3 signals. (a) Side-information unquantized; (b) Side-information quantized.

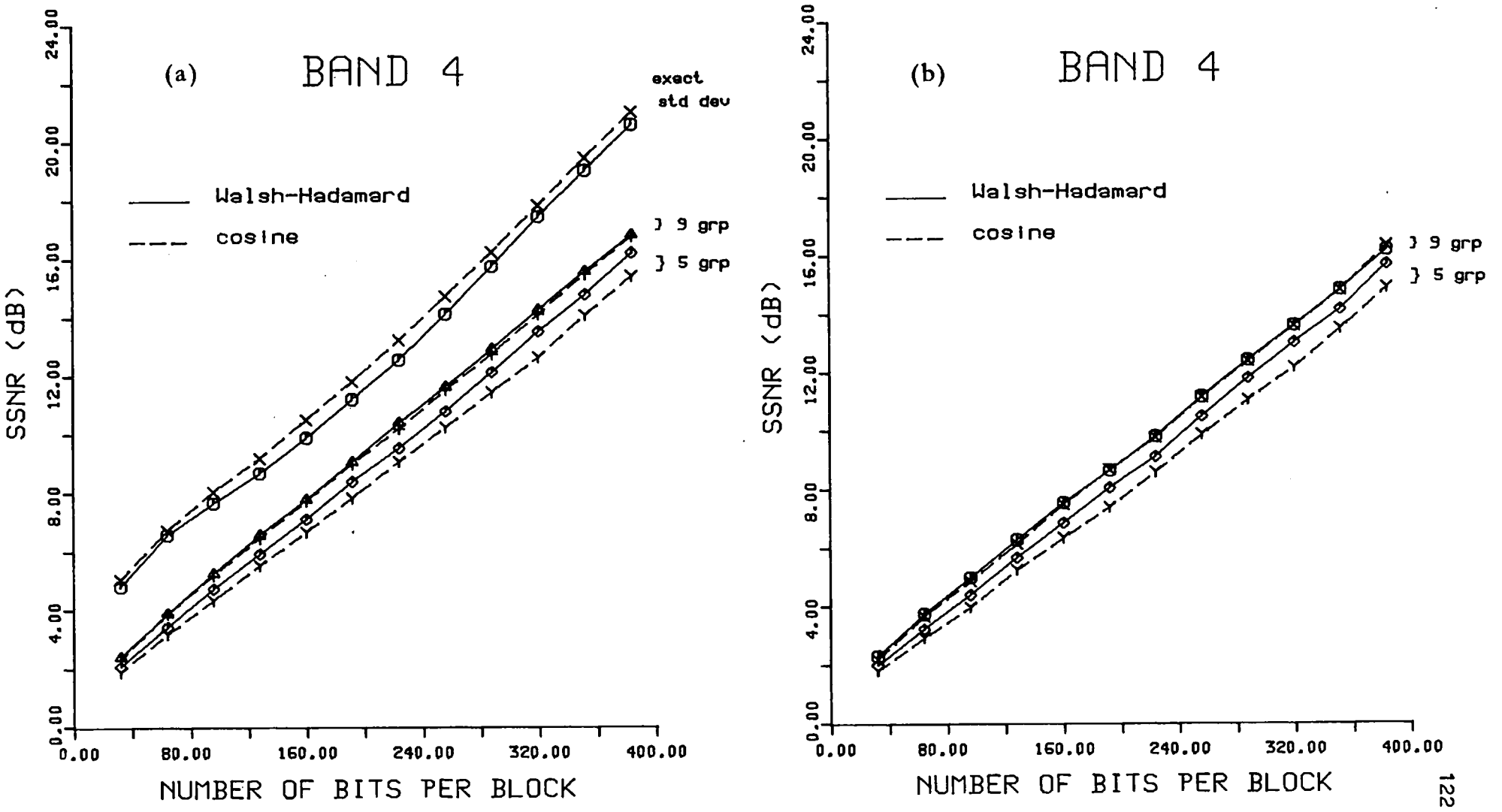


Fig.3.5 Effects of different transforms on SSNR of Band 4 signals. (a) Side-information unquantized; (b) Side-information quantized.

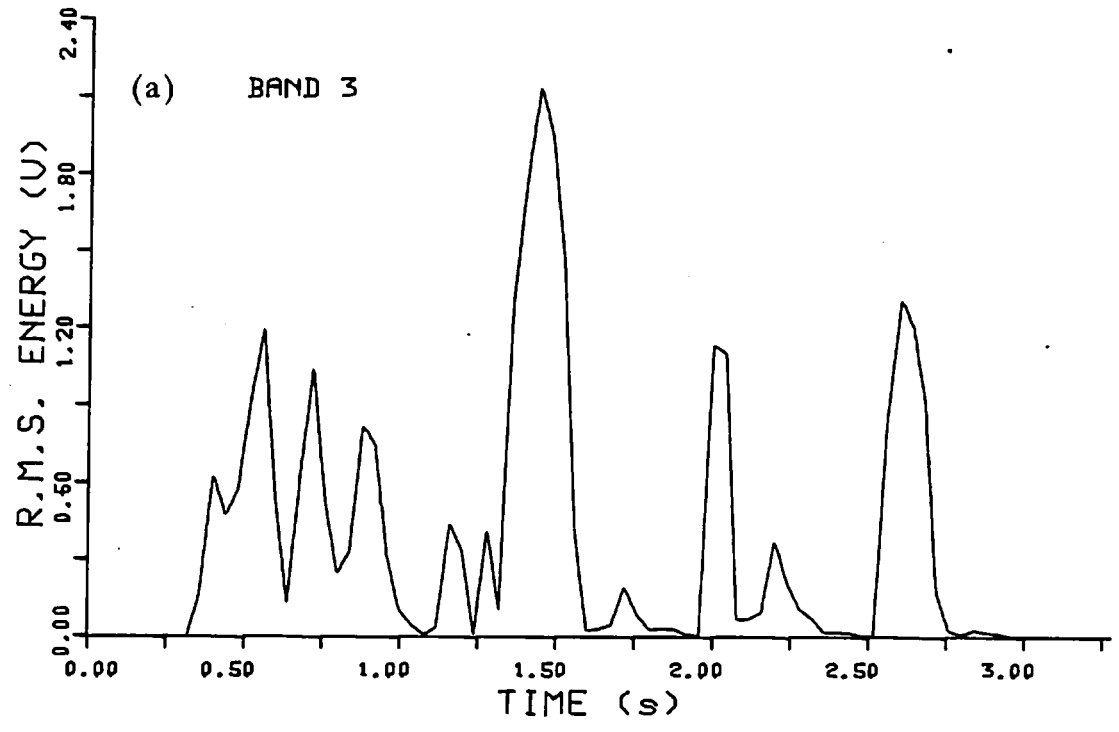


Fig.3.6 (a) Variation of r.m.s. energy within segments of a sub-band (B3) signal of speech.

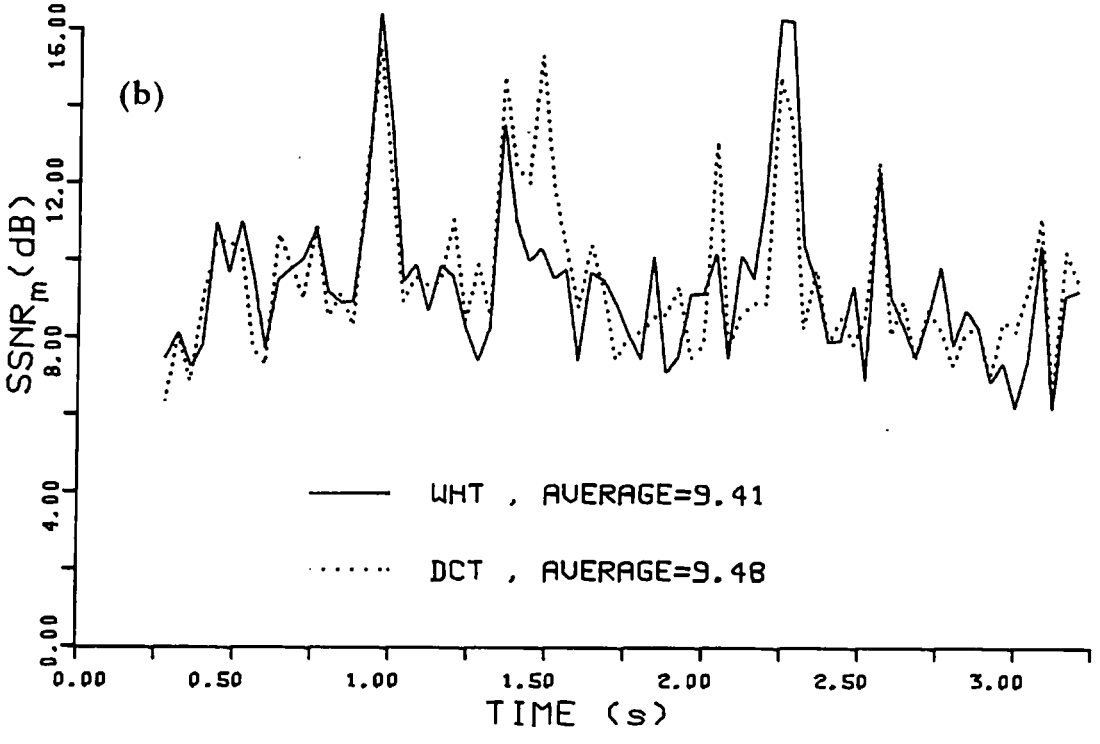


Fig.3.6 (b)

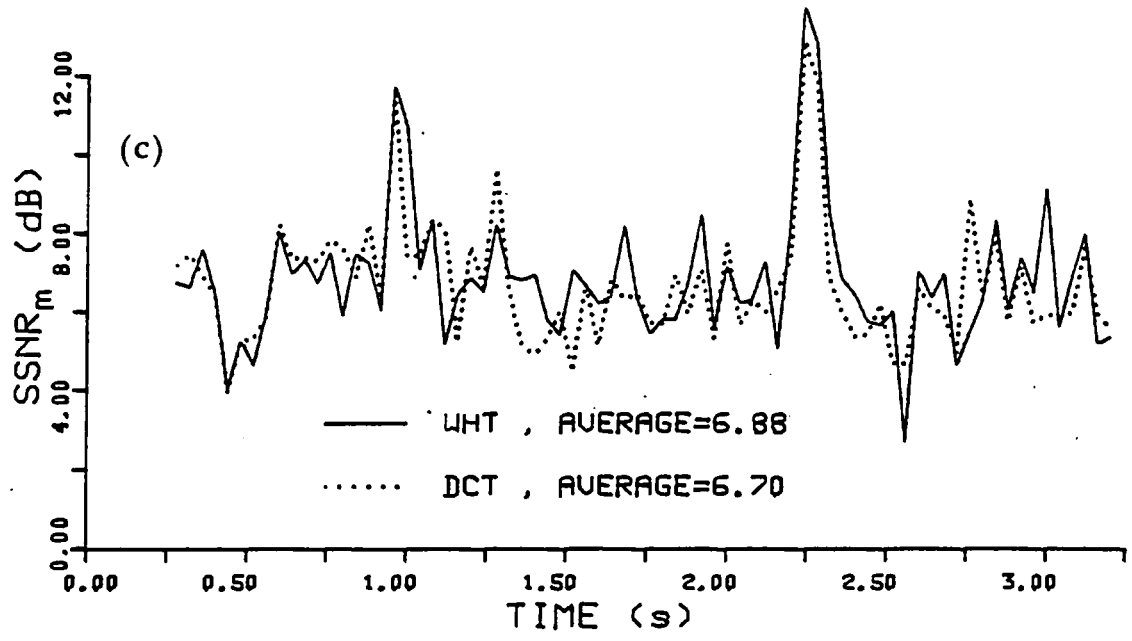


Fig.3.6(c)

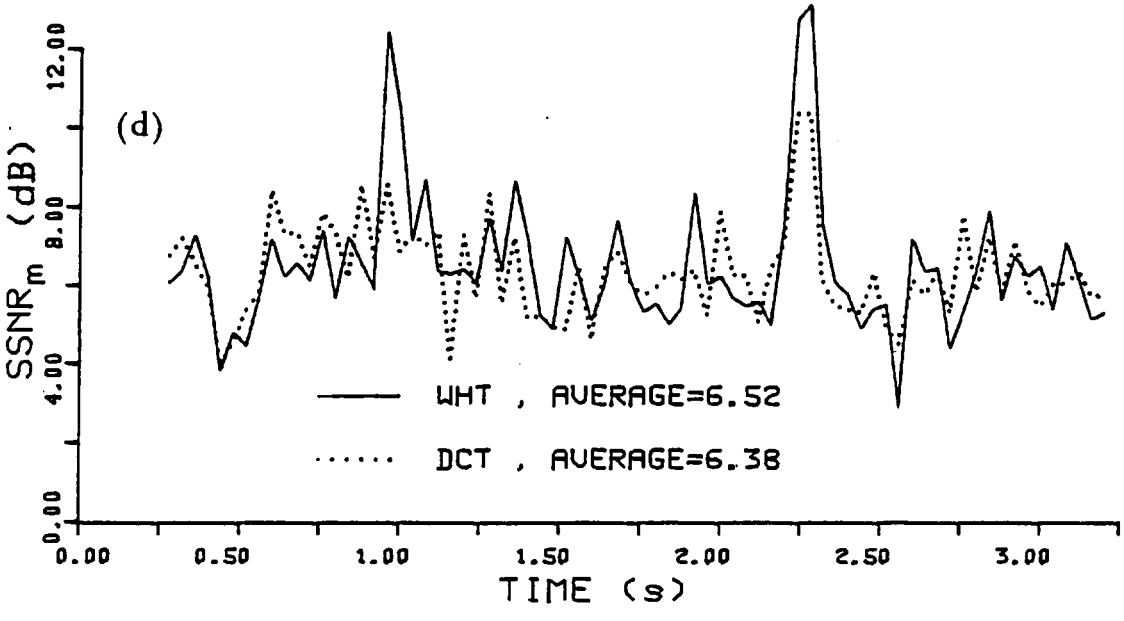


Fig.3.6(d)

Fig.3.6(b-d) Comparisons of $SSNR_m$'s of the sub-band signal coded using the DCT and those of the signal coded using the WHT. Block length = 40 ms. Number of samples per block = 64. Total number of bits per block allocated for quantization of the transform coefficients = 80. (b) using exact standard deviation; (c) using unquantized side-information with 5 group values; (d) using quantized side-information with 5 group values.

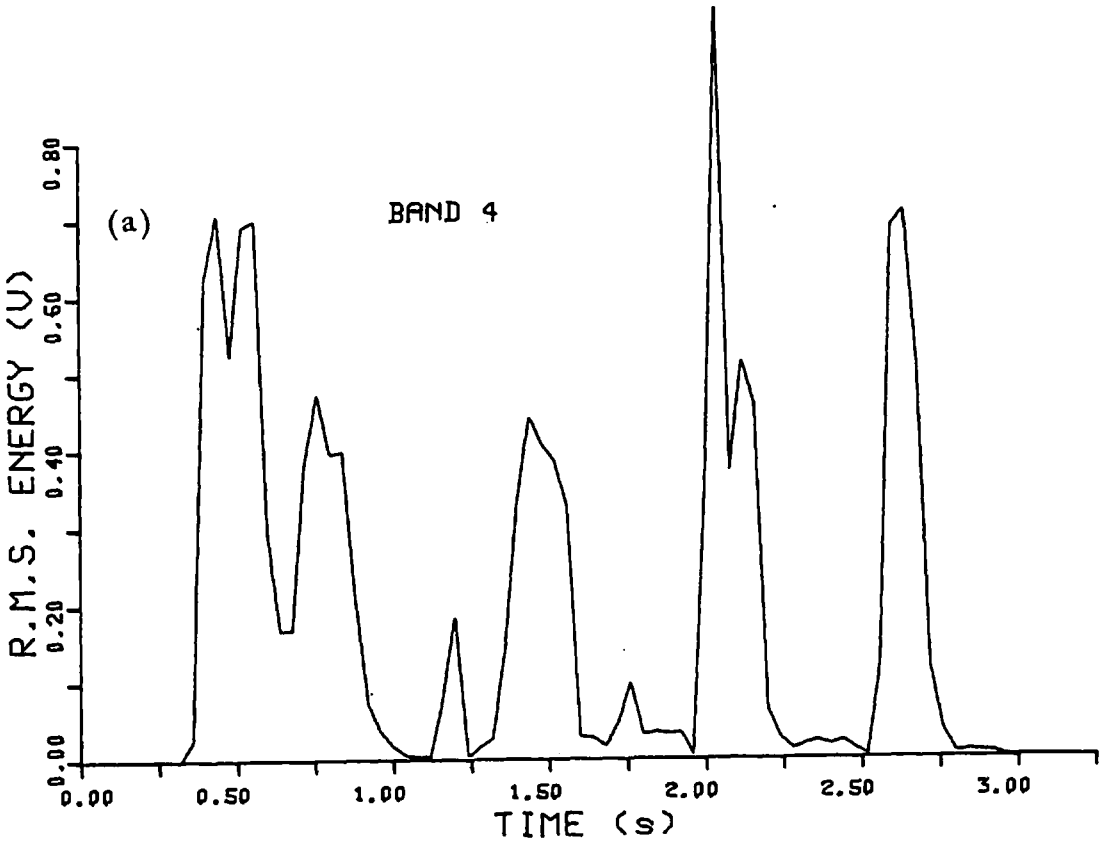


Fig.3.7 (a) Variation of r.m.s. energy within segments of a sub-band (B4) signal of speech.

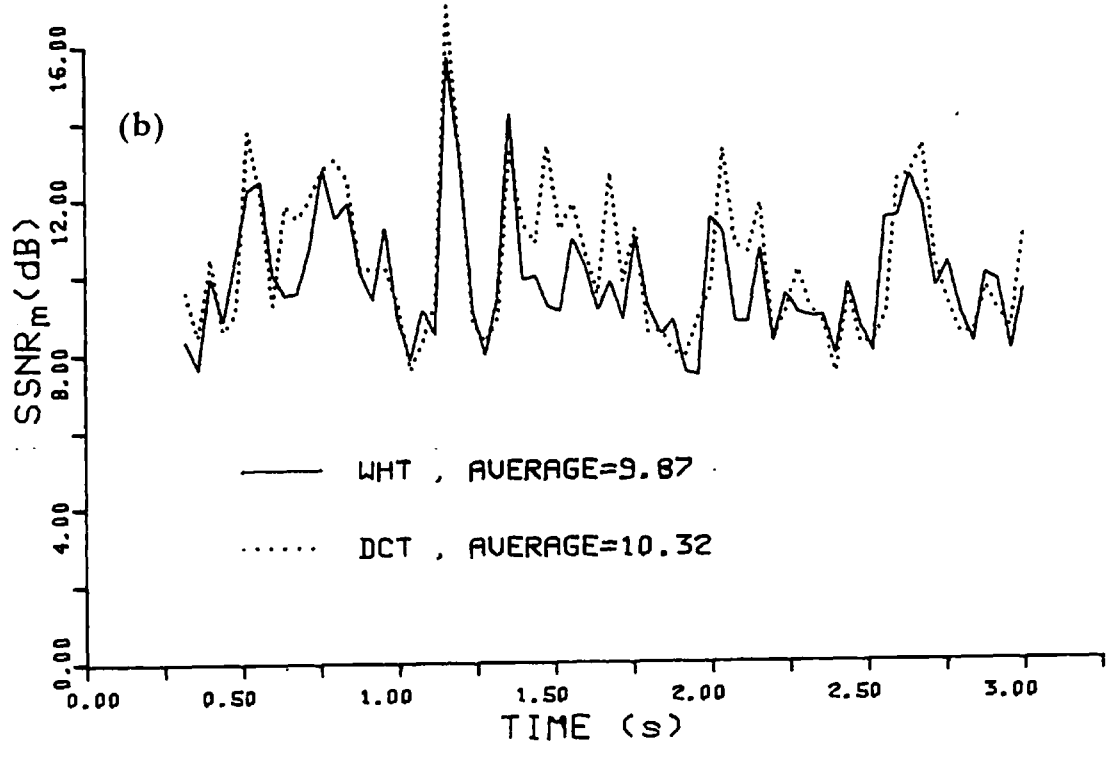


Fig.3.7 (b)

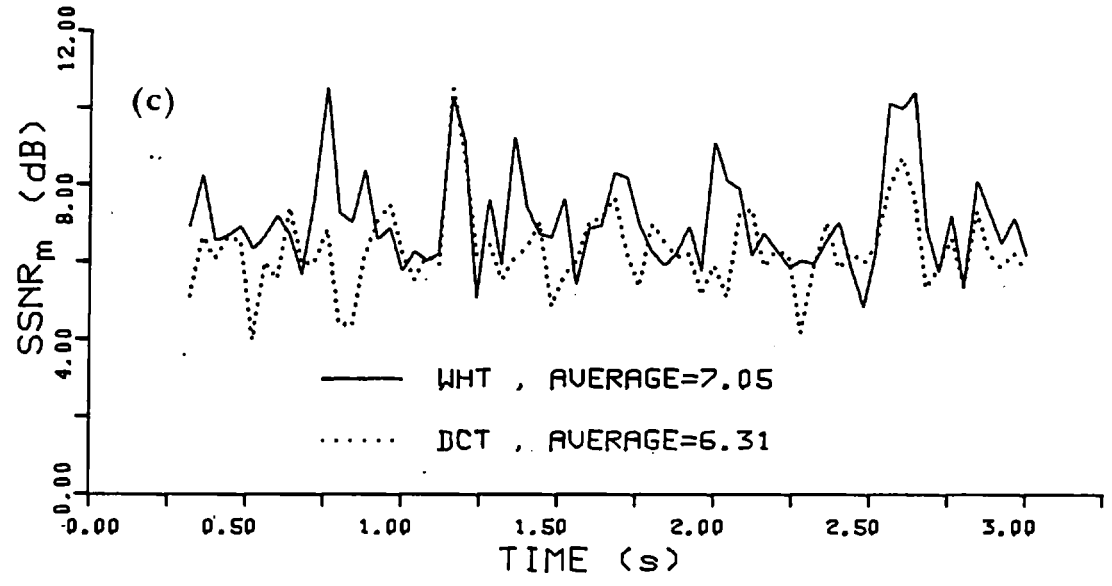


Fig.3.7(c)

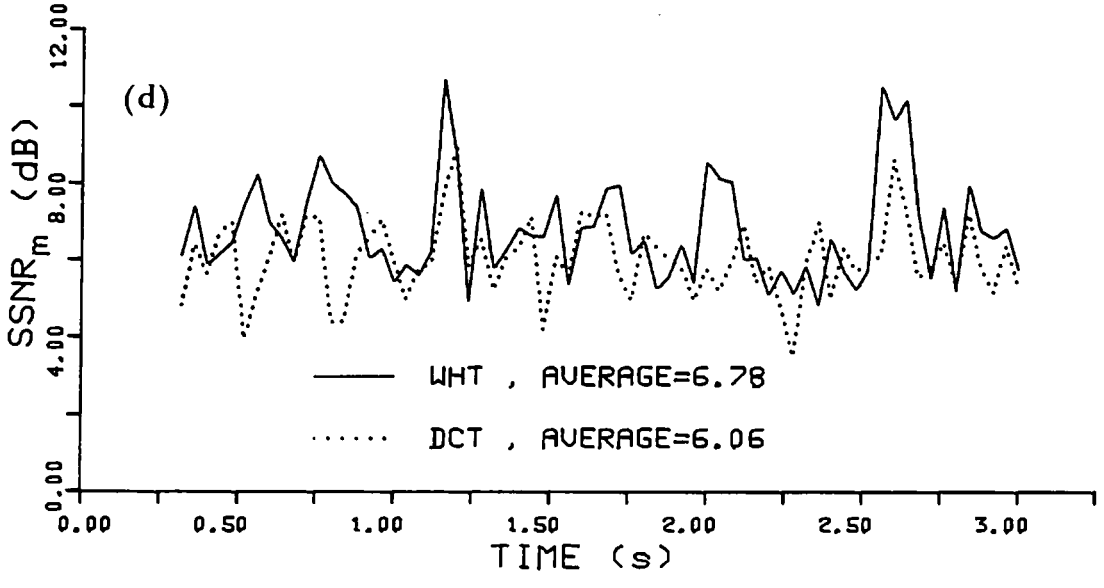


Fig.3.7 (d)

Fig.3.7 (b-d) Comparisons of $SSNR_m$'s of the sub-band signal coded using the DCT and those of the signal coded using the WHT. Block length = 40 ms. Number of samples per block = 128. Total number of bits per block allocated for quantization of the transform coefficients = 160. (b) using exact standard deviation; (c) using unquantized side-information with 5 group values; (d) using quantized side-information with 5 group values.

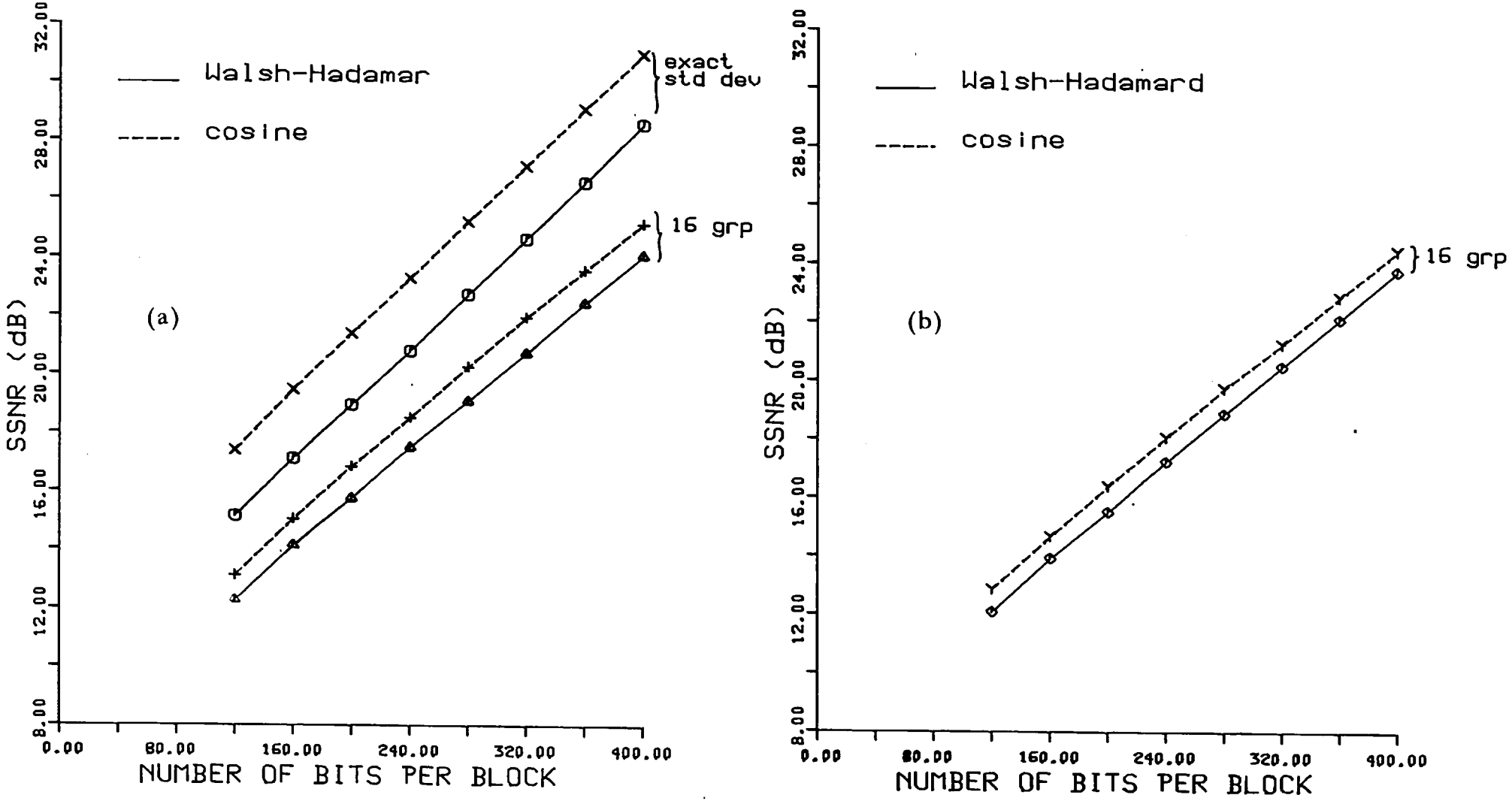


Fig.3.8 Effects of different coding methods on SSNR of full-band signals. (a) Side-information unquantized; (b) Side-information quantized.

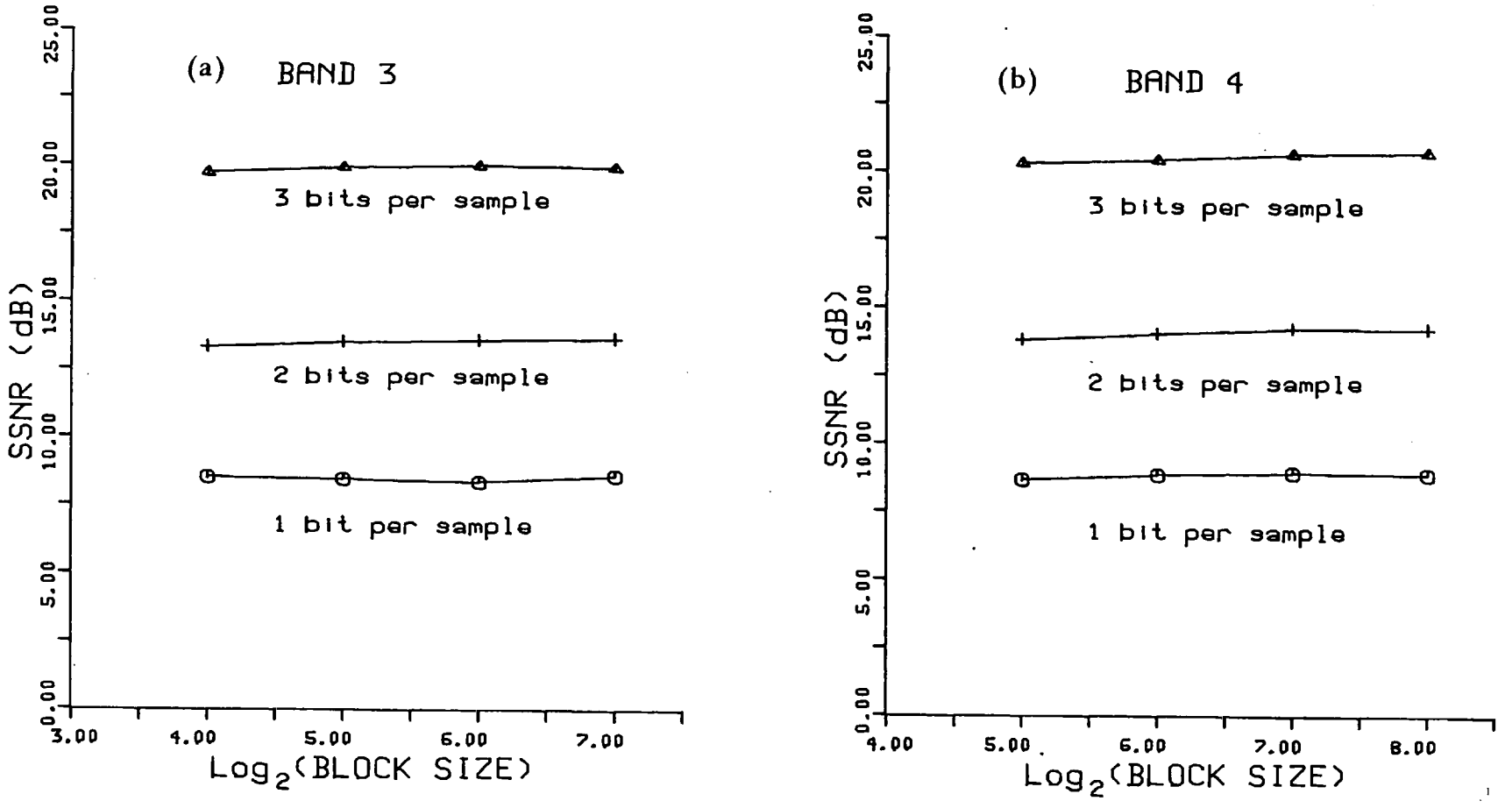
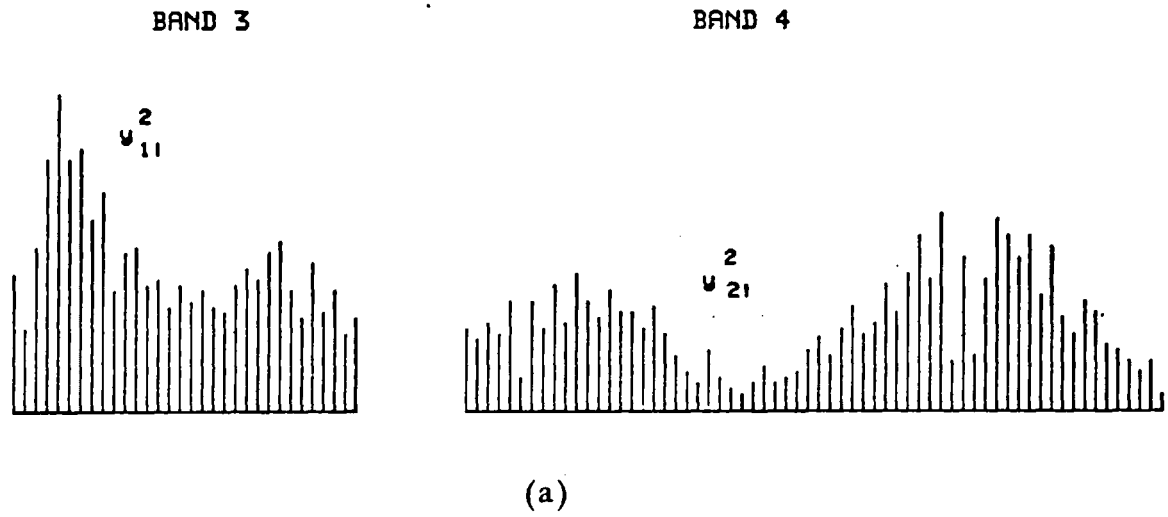
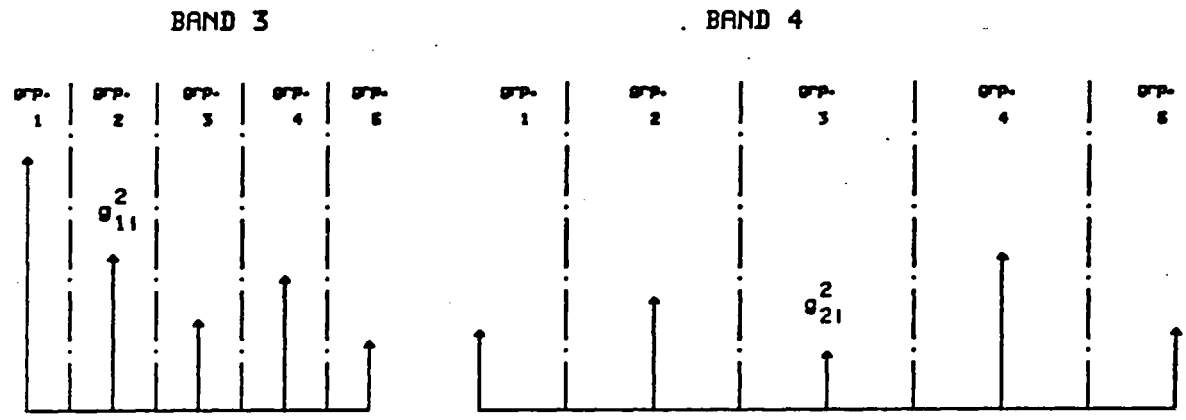


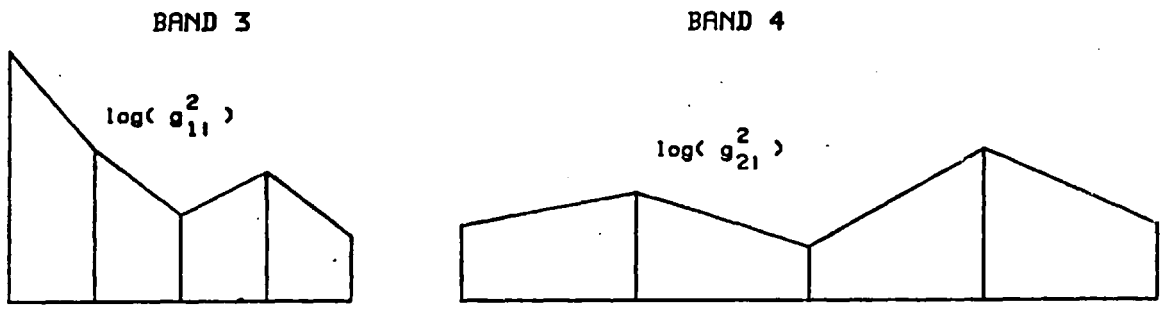
Fig.3.9 Illustration of the effects of transform block size on SSNR; (a) Band 3; (b) Band 4.



(a)



(b)



(c)

Fig.3.10 Illustration of coding and estimation of spectral envelope; (a) Original spectra; (b) Averages of groups; (c) Estimation by interpolation.

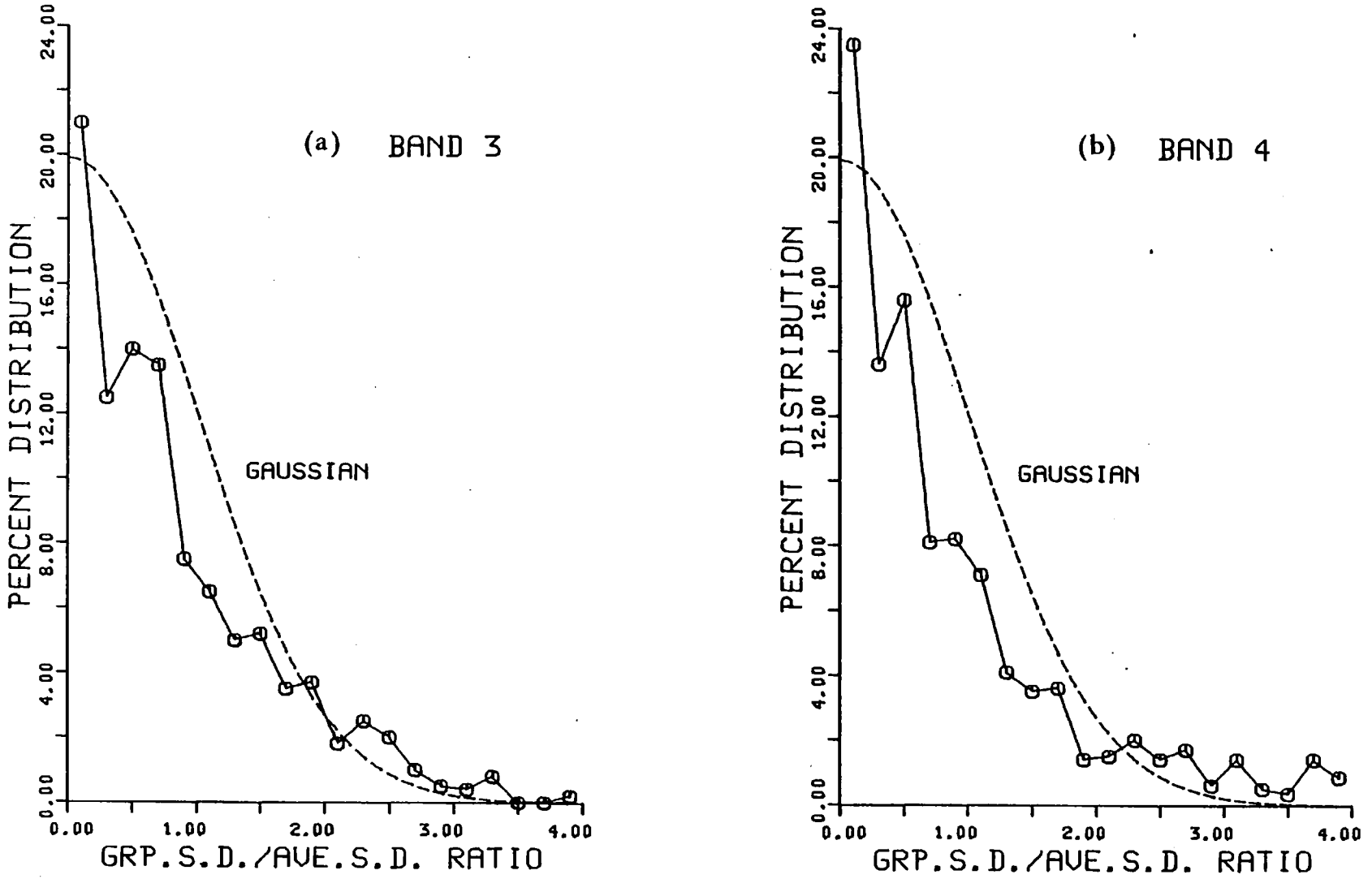


Fig.3.11 Distribution of ratios of group standard deviation to average block standard deviation; (a) Band 3; (b) Band 4.

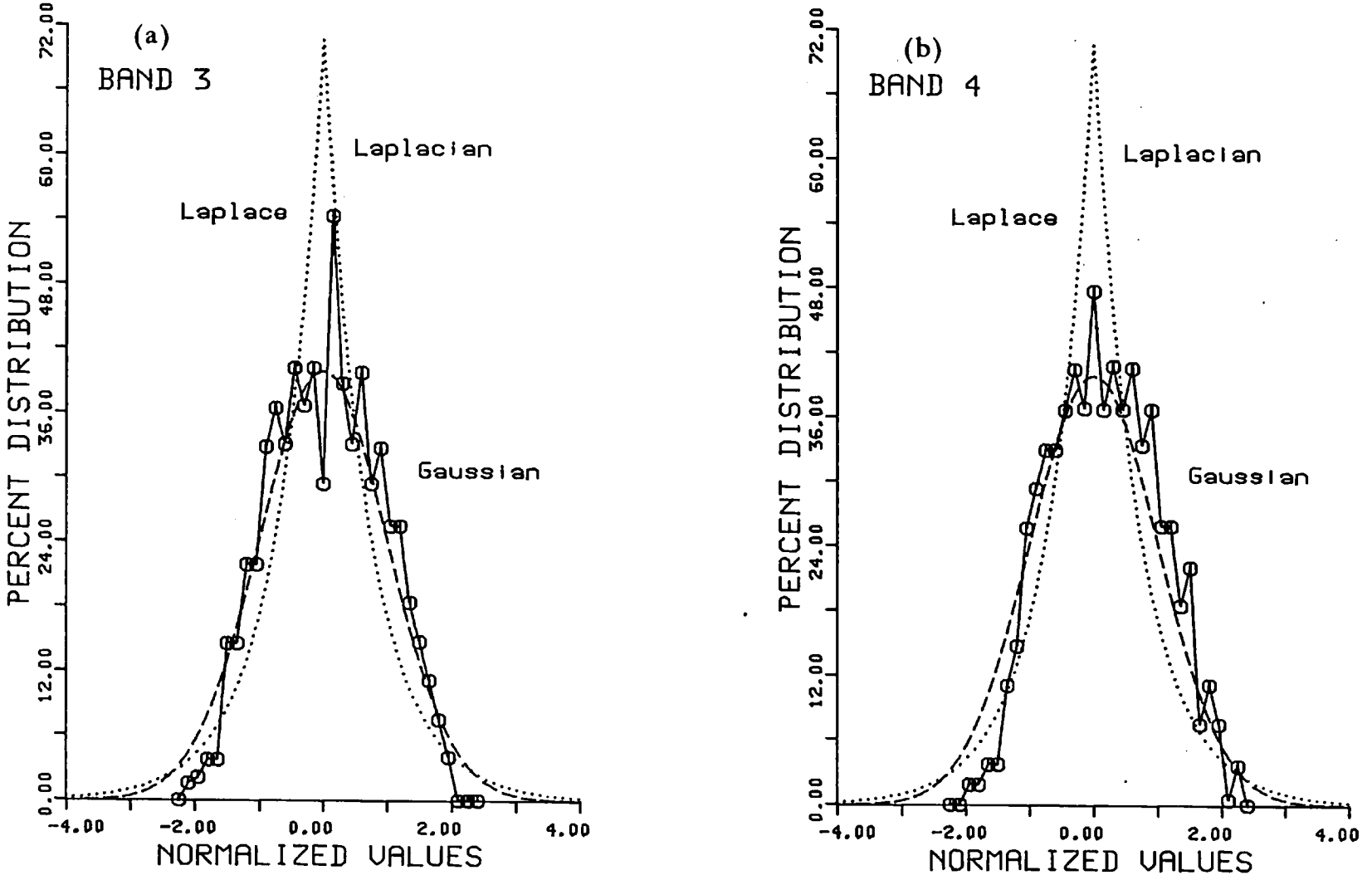


Fig.3.12 Probability density functions for normalized transform values of signals; (a) Band 3; (b) Band 4.

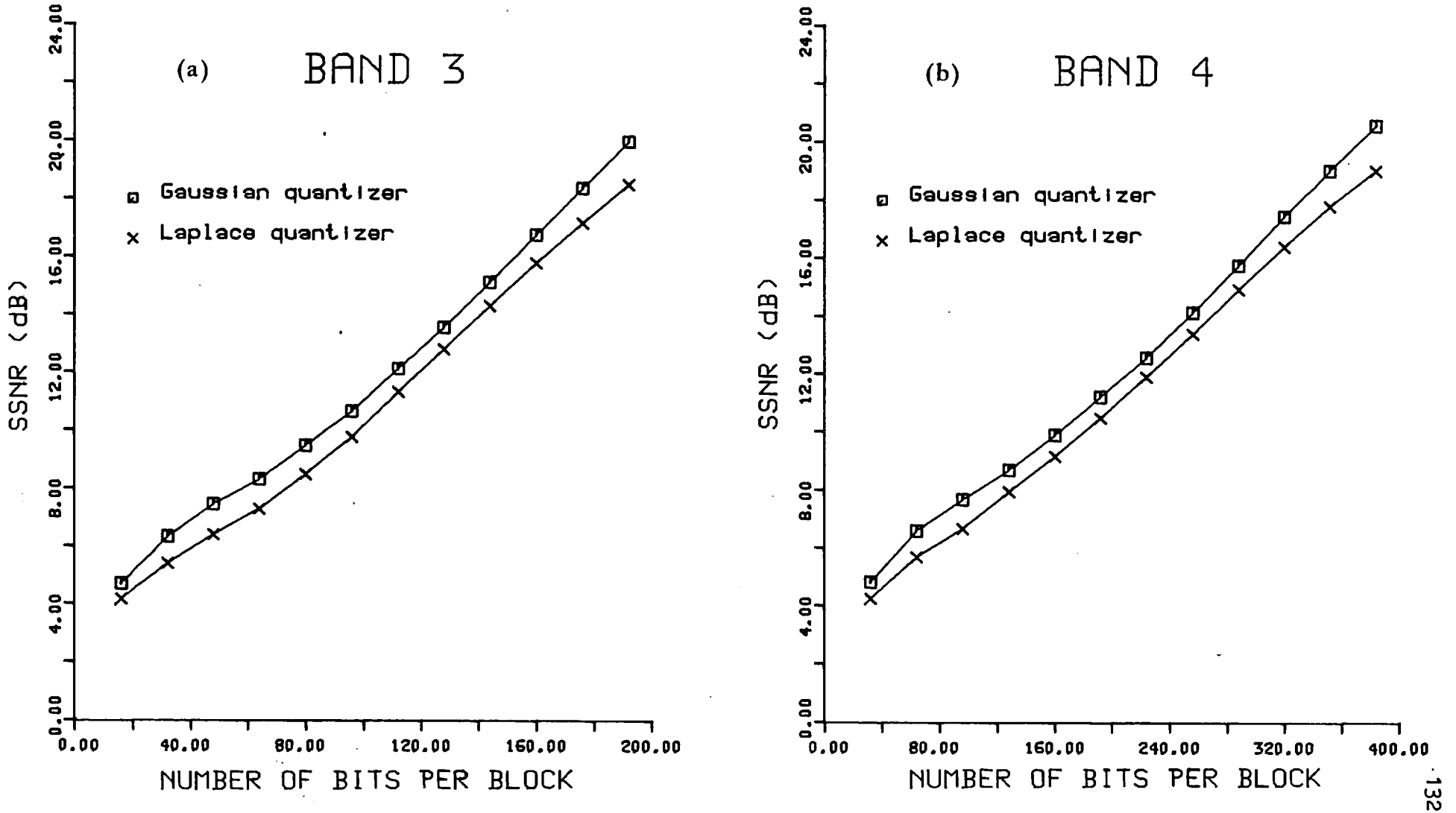


Fig.3.13 Effects of types of quantizer on SSNR; (a) Band 3; (b) Band 4.

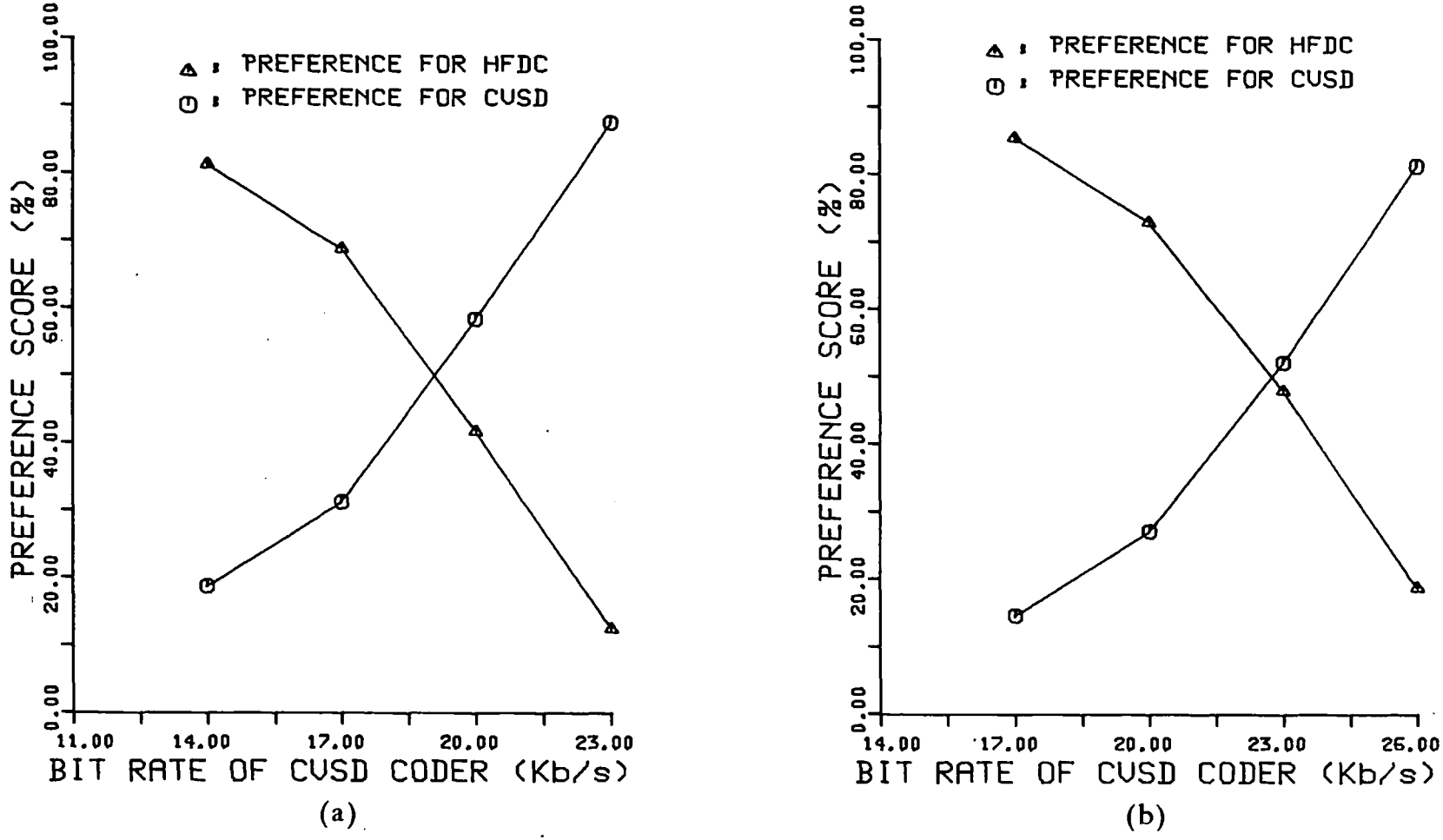


Fig.3.14 Relative subjective preference of quality of speech produced by HFDC and that produced by CVSD at different bit rates. (a) Experiment 1 (HFDC at 7.2 kbit/s); (b) Experiment 2 (HFDC at 9.6 kbit/s).

CHAPTER 4

VARIABLE-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS

CHAPTER 4

VARIABLE-RATE HYBRID FREQUENCY-DOMAIN CODING OF SPEECH SIGNALS

4.1 INTRODUCTION

It is possible, by appropriate segmentation, to divide speech signals into voiced, unvoiced and silent intervals. Voiced speech is produced by air from the lungs being modulated by the vibration of the vocal cords, the resulting signal is usually quasi-periodic. Unvoiced speech, on the other hand, is produced when the air flow from the lungs becomes turbulent as the air passes through constriction in the vocal tract, with the consequence that the signal is more noise-like. The intervals during which no utterance is made are classified as silence. The very short intervals of silence within sentences are usually referred to as micro-silence to distinguish them from the long pauses between sentences. The distribution of the energy within segments of voiced speech, unvoiced speech and silence is shown in Fig.4.1. From this figure, it can be seen that, on the average, there is a larger energy content within a segment of voiced sound than within a segment of unvoiced sound. Speech can be considered to be a concatenation of voiced sounds, unvoiced sounds and silence. The order and the sustained duration of the sounds depends very much on the actual utterance. For examples, the variation of the r.m.s. value of

samples within segments of 40 ms of a conversation of duration 36 s is illustrated in Fig.4.2 and that of a news announcement of about 24 s in Fig.4.3. The distributions of the r.m.s. values of segments of the two signals are shown in Figs.4.4 and 4.5. From these figures, it can be seen that the percentage distribution of the energy depends very much on the actual speech material. Thus, speech is a highly nonstationary process with the power of speech signal changing considerably over a period of time and, in consequence, variable-bit-rate coding is generally found to be more efficient than fixed-bit-rate coding.

Dubnowski and Crochiere [93] have described a variable-bit-rate speech coding technique in which the speech samples in different time segments are coded with different numbers of bits per sample. The technique is based on the use of adaptive differential pulse-code-modulation (ADPCM). Other interesting variable-bit-rate speech coding methods include: Extremal Encoding [88] in which the time intervals between the extremes of the signals and the amplitudes at the extremes are encoded and transmitted; Time Encoding [89,90] in which a fixed alphabet is used to indicate the shape of the waveform and the time interval between successive pairs of zero crossings; and Adaptive Aperture Coding [91,92] in which samples are coded only if the amplitudes cross an appropriately wide aperture centered around the last encoded value. These techniques are described in detail in Chapter 2.

In this chapter, a new technique of variable-bit-rate coding is put forward. The technique is based on the breaking down of the speech signal into fixed-length segments which are then classified according

to the energy content. The technique, which is of a quite general nature, has been applied to hybrid frequency-domain coding and the resulting variable-rate hybrid frequency-domain coding technique has been compared with fixed-rate hybrid frequency-domain coding. The new variable-bit-rate coding technique is described in Section 4.4. Details of its application to hybrid frequency-domain coding are given in Section 4.5 and the results of subjective tests are presented in Section 4.6

4.2 THEORETICAL BASIS OF BIT ALLOCATION TO SEGMENTS OF SPEECH SIGNALS

There are several ways in which the nonstationarity of speech can be exploited. One way is to divide the signals into segments and code the segments using different numbers of bits. A theoretical guideline for the allocation of bits based on the energy content within segments of speech signals will now be derived.

Let $\{y_{mi}, i=0,\ldots,N-1\}$ be the set representing the N samples of the m^{th} segment of input signal and $\{\hat{y}_{mi}, i=0,\ldots,N-1\}$ be the set representing the N samples of the corresponding segment of signal quantized with an average of b_m bits per sample. The average mean-square signal power of the m^{th} segment, denoted by S_m^2, is then given by

$$S_m^2 = \frac{1}{N} \sum_{i=0}^{N-1} y_{mi}^2 \tag{4.1}$$

and the average mean-square overall distortion of the m^{th} segment, denoted by E_m^2, is given by

$$E_m^2 = \frac{1}{N} \sum_{i=0}^{N-1} (y_{mi} - \hat{y}_{mi})^2 \tag{4.2}$$

For all practical purposes, it can be assumed that the average segmental signal-to-noise ratio is proportional to $2^{(2b_m)}$, where b_m is the average number of bits assigned to a sample of the m^{th} segment. Thus,

$$\frac{S_m^2}{E_m^2} = K2^{(2b_m)} \tag{4.3}$$

where K is a constant.

Consider a sufficiently long duration of signal consisting of M segments with each segment containing N samples of the signal. The problem is how to allocate the total number of bits, B_t, available to code these M segments such that the total distortion is minimized.

The total distortion of the M segments of signal, denoted by E^2, can be written as

$$E^2 = \sum_{m=1}^{M} E_m^2 = \sum_{m=1}^{M} S_m^2 K^{-1} 2^{(-2b_m)} \tag{4.4}$$

Denoting $X_m = K^{-1}2^{(-2b_m)}$, the problem can be formulated as the minimization of the overall distortion

$$E^2 = \sum_{m=1}^{M} S_m^2 X_m \tag{4.5}$$

subject to the constraint

$$\prod_{m=1}^{M} X_m = \prod_{m=1}^{M} K^{-1}2^{(-2b_m)} = K^{-M}2^{(-2B_t/N)} = \text{constant} = C \tag{4.6}$$

This minimization problem can be solved by using the Lagrange multipliers. The solution of the set of equations

$$\frac{d}{dX_m}(E^2 + \lambda C) = 0 \qquad m = 1,\ldots,M \tag{4.7}$$

gives

$$S_m^2 X_m = -\lambda C \qquad m = 1,\ldots,M \tag{4.8}$$

from which it can be derived that

$$b_m = \frac{B_t}{NM} + \frac{1}{2}\log_2 S_m^2 - \frac{1}{2M}\sum_{i=1}^{M} \log_2 S_i^2 \tag{4.9}$$

4.3 METHODS OF VARIABLE-BIT-RATE CODING BASED ON SIGNAL SEGMENTATION

4.3.1 A Super Block Processing Approach

From the analysis given in the previous section, it is clear that to minimize the total distortion, the total available number of bits should be allocated in such a way that segments with relatively higher energy are coded with more bits. One approach is to consider a sufficiently long duration of the speech consisting of, say L segments together and distribute the total available number of bits according to the relative average energy content within the segments.

However, if sufficient gain is to be obtained to make this method worthwhile for a single user system, then very long-duration segment of speech signal must be used. That is, the L segments must cover the typical duration of phonemes and micro-silence of speech. As processing can commence only after all the samples of the L segments are available, very long transmission delay results.

4.3.2 A 'Limited-Type' Approach

In order to gain an insight into the problem of tradeoff between time delay and quality improvement for a single user system, the classification of the segments into two types, namely silence and non-silence which includes voiced and unvoiced segments, is considered next. Let the proportion of non-silent intervals over the total signal duration be α and that of the silent intervals be $(1-\alpha)$. If the average transmission bit rate is I_n and the silent intervals are not coded, then all the bits saved from the silent intervals can be allocated for coding the non-silent intervals. It can be shown that if this is carried out, the non-silent intervals can then be coded at the rate of $I_v=(I_n/\alpha)$, which corresponds to an increase by a factor of $1/\alpha$.

For the conversation speech mentioned in Section 4.1, it was found that the non-silent periods comprise 60 percent of the total signal duration, thus $\alpha=0.6$, and $I_v=(5/3)I_n$, which is a significant increase. However, most of the silent intervals occur between sentences, and to encode the signal using this method, delays of up to 2-3 seconds are involved.

The delay can be considerably reduced, however, if advantage is taken mainly of the micro-silent rather than the silent intervals between sentences. Statistics based on long duration of speech signals [106] show that micro-silent intervals comprise about 15 percent of connected speech. Thus $\alpha=0.85$, and $I_v=(20/17)I_n$. The increase is small as the bits saved from a small proportion of segments are distributed over a large proportion of segments. The delay, however, is reduced to less than 1 second.

By classifying speech signals into two types, the actual delay cannot be suitably controlled without drastic deterioration in the quality. For instance, when the number of non-silent intervals exceeds the expected limit, to avoid overflow or to limit the time delay or both, subsequent non-silent intervals will have to be treated as silent intervals resulting in severe degradation of speech quality. If, however, segments are classified into one of three different types, a control strategy can be designed to limit the delays and buffer requirements to below predetermined limits without introducing significant distortion into the speech. This new technique is described in detail in the following sections.

4.4 A NEW TECHNIQUE OF VARIABLE-BIT-RATE CODING

4.4.1 General Comments

In the new method of variable-bit-rate coding set out in this chapter, speech signals are divided into fixed-length segments, which are then classified according to their energy content. If the r.m.s. value of samples within a segment is less than a threshold V_u, the segment is classified as being Type-1. If, on the other hand, the r.m.s. value of the segment is greater than a threshold V_v ($V_v > V_u$), then the segment is classified as being Type-3. Those segments whose r.m.s. values are between the two thresholds are classified as Type-2. Signals of Type-2 and Type-3 tend to correspond to voiced sounds and signals that are classified as Type-1 tend to correspond to unvoiced sounds or silence.

An essential feature in the transmission of a variable-bit-rate source data over a fixed-bit-rate communication channel is the provision of buffering at the transmitter and/or the receiver. It has been shown [94] that if the effects of long period of low source activity are to be overcome, then the transmitter buffer must introduce a delay δ, where δ satisfies the condition

$$(t-\delta)R \leqslant \min\{C(t)\} \qquad t > \delta \tag{4.10}$$

in which R is the constant transmission rate in bit/s and C(t) is the cumulative total number of information bits generated up to time t seconds. It is also shown in [94] that if the effects of high source activity are to be overcome, the receiver buffer must introduce a delay Δ where

$$(\Delta+t)R \geqslant \max\{U(t)\} \tag{4.11}$$

in which U(t) is the cumulative total number of information bits used up to time t for signal reconstruction.

In the analysis given in [94], no attention was given to the possibility of controlling the rate at which the source generates its information with the result that the delays and buffer requirements were considered to be fixed by the long-term statistical properties of the source. By adopting a control strategy as described below, the communication delays and buffer requirements can be kept below fixed pre-determined limits.

4.4.2 A Control Strategy for Variable-Bit-Rate Coding

Consider a speech waveform to be divided into segments each of duration T seconds, and let the fixed transmission bit rate be R bit/s with I_2 bits being transmitted in the T second interval, that is $I_2 = RT$. Let the number of bits used to encode Type-1, Type-2 and Type-3 segments be I_1, I_2 and I_3 respectively. Thus if the condition of constant-rate transmission is to be satisfied then it is necessary for

$$I_2 = P_1 I_1 + P_2 I_2 + P_3 I_3 \tag{4.12}$$

where P_1, P_2 and P_3 are the respective probabilities associated with Type-1, Type-2 and Type-3 segments.

If $L_1(t)$, $L_2(t)$ and $L_3(t)$ are the cumulative total numbers of segments of Type-1, Type-2 and Type-3 respectively to time t, then the total number of bits generated up to time t is given by $L_1(t)I_1+L_2(t)I_2+L_3(t)I_3$ and the number of bits transmitted up to the same time t is $[L_1(t)+L_2(t)+L_3(t)]I_2$. Also, let

$$J_u = I_2 - I_1 \tag{4.13}$$

$$J_v = I_3 - I_2 \tag{4.14}$$

From Equation (4.10), it follows that if buffer underflow is to be avoided, then the minimum delay, δ, introduced by the transmission buffer must satisfy the condition

$$\delta \geq \frac{1}{R} \max\{[L_1(t)+L_2(t)+L_3(t)]I_2-[L_1(t)I_1+L_2(t)I_2+L_3(t)I_3]\} \tag{4.15}$$

i.e. $$\delta \geq \frac{T}{I_2} \max\{L_1(t)J_u - L_3(t)J_v\} \quad (4.16)$$

The maximum value of $\{L_1(t)J_u - L_3(t)J_v\}$, that is, $\max\{L_1(t)J_u - L_3(t)J_v\}$ can be confined to be less than a predetermined value B_u by using I_2 bits rather than I_1 bits to code those segments with r.m.s. values $< V_u$, which if they had been coded as Type-1, that is, with I_1 bits, would have made the $\max\{L_1(t)J_u - L_3(t)J_v\}$ exceed B_u. In this way, the delay δ can be guaranteed to be less than $(T/I_2)B_u$.

Using a similar argument, it can be shown that if buffer underflow at the receiver is to be prevented, then the minimum delay Δ has to satisfy the condition

$$\Delta \geq \frac{T}{I_2} \max[L_3(t)J_v - L_1(t)J_u] \quad (4.17)$$

and by using I_2 bits rather than I_3 bits to code those segments with r.m.s. values $> V_v$, which if they had been coded as Type-3, that is, with I_3 bits, would have resulted in $\max[L_3(t)J_v - L_1(t)J_u]$ exceeding the value B_v, the delay Δ is guaranteed to be less than $(T/I_2)B_v$.

The general variable-bit-rate strategy described above can be simplified if J_u is set equal to J_v. If this is done, then Equations (4.16) and (4.17) become,

$$\delta \geq \frac{T}{I_2} [L_1(t) - L_3(t)]J_u \quad (4.18)$$

and $$\Delta \geq \frac{T}{I_2} [L_3(t) - L_1(t)]J_u \quad (4.19)$$

respectively.

If $L_u(t)$ is defined to be $L_1(t) - L_3(t)$; $L_v(t)$ is defined to be

$L_3(t)-L_1(t)$ and K_u and K_v are the specified thresholds (number of segments), then the simplified strategy involves coding Type-1 segments with I_2 rather than I_1 bits when $L_u(t) > K_u$ and coding Type-3 segments with I_2 rather than I_3 bits when $L_v(t) > K_v$. The delays δ and Δ are then limited according to the following relations:

$$\delta \leqslant (T/I_2)K_u J_u \tag{4.20}$$

$$\Delta \leqslant (T/I_2)K_v J_u \tag{4.21}$$

4.4.3 Buffer Requirements

The buffer requirements necessary for implementing the above strategy can be determined as follows:

At the transmitter, the sets of digits representing each segment have to be stored in the buffer for an interval of δ seconds. In addition, capacity has to be provided in the transmitter buffer to cope with periods of high source activity when the rate of information generation exceeds the transmission rate. Since the maximum number by which the number of bits generated exceeds the number of bits transmitted is constrained to be less than ΔR, it follows that the required maximum capacity, Q_t, of the transmitter buffer is given by

$$Q_t = \delta R + \Delta R = B_u + B_v \tag{4.22}$$

Using similar arguments, the maximum capacity required for the receiver buffer, Q_r, can be seen to be

$$Q_r = \Delta R + \delta R = B_v + B_u \tag{4.23}$$

It is clear from Equations (4.22) and (4.23) that the same size buffer is required at the transmitter and the receiver, and that by suitable selection of B_u and B_v, the necessary buffer sizes can be restricted.

In the case of the simplified strategy in which J_u is set equal to J_v, the maximum buffer required at both the transmitter and the receiver is $(K_u+K_v)J_u$.

4.4.4 Selection of the Parameters

The parameter V_u, which is a threshold determining the percentage of segments classified as Type-1, that is, in the absence of any control strategy, the percentage of segments coded using the lowest number of bits I_1, has to be determined by observing the effects on the perceptual quality of speech of varying I_1 and V_u. The lowest value of I_1 giving an acceptable quality speech depends on the actual scheme used for coding the speech segments and also on the number of segments that are actually classified as Type-1. In the case of systems operating at low bit rates, a useful guide in selecting V_u is to choose V_u such that a high percentage of the total number of segments classified as Type-1 are in fact silent segments. In the case of the simplified strategy in which $J_u=J_v$, that is, $I_2-I_1=I_3-I_2$, once I_1 and V_u are determined, I_3 can be calulated since I_2 is fixed directly by the average transmission bit rate. P_3 can then be determined from Equation (4.12). The value of V_v which results in P_3 of the segments being classified as Type-3, can then be determined from the cumulative distribution function of the r.m.s. values of the speech signal measured over a long period.

The values of the parameters K_u and K_v are related to the transmitter and receiver delays in the manner indicated in Equations (4.20) and (4.21); and hence having determined J_u and I_2, K_u and K_v can be selected to give the desired values of delays. It should be remembered that the parameters K_u and K_v, in addition to controlling the delays, also affect the quality of the coded speech by virtue of the fact that they control the number of Type-3 segments coded using I_2 bits, as well as the number of Type-1 segments coded using I_2 bits.

A selection of the parameters V_u, V_v, K_u and K_v based on the long-term statistics of speech may not lead to optimal performance and it may be necessary to adopt a dynamic control process. One such approach may be to select and fix V_u commensurate with an acceptable speech quality, and then adjust the threshold V_v in the following manner: When K_u is exceeded, the fraction of segments coded with I_1 bits is compared with P_1 and if this fraction is found to be consistently less than P_1, thus indicating that the selected value of V_v is too high, V_v is reduced. In the same way, when K_v is exceeded, the fraction of the segments coded with I_3 bits is checked and compared with P_3. If this fraction is found to be consistently greater than P_3, then V_v is increased.

As an example, the percentage of segments coded as Type-3 for different values of K_u and K_v are given in Table 4.1. The speech material used consists of 24 s of news announcement by a male speaker.

Table 4.1 Percentage of Segments Coded as Type 3

Speech Material: Radio Announcement

LIMITS		PERCENTAGE OF SEGMENTS CODED AS TYPE 3
K_u	K_v	
(no limit)	(no limit)	19.1
15	15	17.1
10	10	16.3
5	5	15.4

From this table, it can be seen that when no limit is imposed on K_u and K_v, 19.1 % of the segments are coded as Type-3 and the percentage of segments coded in this way is reduced to 16.3 % when $K_u = K_v = 10$.

4.5 APPLICATION OF THE VARIABLE-BIT-RATE TECHNIQUE TO HYBRID FREQUENCY-DOMAIN CODING

Fig.4.6 is a block diagram of a hybrid frequency-domain coder which makes use of the proposed variable-bit-rate technique.

In the transmitter shown in Fig.4.6(a), the speech is segmented and the r.m.s. value associated with each segment is determined in a power measuring unit. The measured r.m.s. value is then compared in a threshold unit with the thresholds, V_u and V_v, in order to classify the segment as being of Type-1, 2 or 3. The information from the segment classifier is then fed to a counter unit which keeps count of the difference between the number of segments that have been classified as Type-1 and the number which have been classified as Type-3, that is, the unit keeps count of $L_u(t)$.

The information from the counter is fed to a decision unit in which $L_u(t)$ is compared with the predetermined threshold K_u and $-L_u(t)$ $[=L_v(t)]$ is compared with K_v. If $L_u(t)$ is found to be greater than K_u, then an instruction is sent to the hybrid frequency-domain coder to code the segment with I_2 rather than I_1 bits. If, on the other hand, the mutually exclusive event in which $-L_u(t)$ is found to be greater than K_v, then an instruction is sent to the hybrid frequency-domain coder to code the segment using I_2 rather than I_3 bits.

The buffer, which is shown connected between the hybrid frequency-domain coder and the transmission channel, and which is of size $(K_u+K_v)J_u$ bits, introduces a fixed delay of $(T/I_2)K_uJ_u$ seconds.

The receiver part of the system is shown in Fig.4.6(b). The incoming data are stored in the receiver buffer. After a delay of Δ s, the header information in the data is examined in order to determine the Type number that was assigned to the segment by the transmitter. This information is supplied to the hybrid frequency-domain coder which then operates in the manner described in the previous chapter to deliver an output.

4.6 SIMULATION OF A VARIABLE-BIT-RATE HYBRID FREQUENCY DOMAIN CODER AND RESULTS OF SUBJECTIVE TESTS

In order to assess the performance of the variable-bit-rate technique, a computer simulation was carried out in which variable-rate hybrid frequency domain coders operating at transmission bit rates of 7.2 kbit/s and 9.6 kbit/s were compared with fixed-rate hybrid frequency-domain coders operating at the same bit rates. The

effects of changes in the overall buffer delay on the performance of the variable-bit-rate scheme were also investigated.

4.6.1 Bit Allocation

In the case of the variable-rate hybrid frequency domain coder, the way in which the total number of bits was assigned to the different frequency bands into which a segment of speech was partitioned is indicated in Tables 4.2 and 4.3.

Table 4.2 Allocation of Bits for Average Rate of 7.2 kbit/s

Segment Classified as	Total No. of bit/s	Distribution of bit/s			
		Band 1	Band 2	Bands 3+4	Side Infor.+Header
Type-1	4000	0	0	3200	800
Type-2	7200	800	1600	4000	800
Type-3	10400	1200	2400	6000	800

Table 4.3 Allocation of Bits for Average Rate of 9.6 kbit/s

Segment Classified as	Total No. of bit/s	Distribution of bit/s			
		Band 1	Band 2	Bands 3+4	Side Infor+Header
Type-1	4000	0	0	3200	800
Type-2	9600	1200	2400	5200	800
Type-3	15200	1600	3200	9600	800

4.6.2 Method of Quantization

For the hybrid frequency-domain coder, the method of quantization of the sub-band signals remains the same as described in the previous chapter except that the step-size used for the adaptive PCM is adjusted when the number of quantization levels available for coding the signal samples is changed. The reason for and method of adjustment are as follows. From Tables 4.2 and 4.3, it can be seen that more bits per sample are allocated to encode the signals in bands B1 and B2 when a segment of the signals is classified to a higher type compared with the preceding segment. As a result, the number of quantization levels available to code the samples also increases. Thus, the quantization step-size should be adjusted when segments of a different type are to be coded. It is found that instead of using the method as proposed by Dubnowski and Crochiere [93], an increase in SNR can be achieved if the ratio of the step-sizes, R_{on}, is such that

$$R_{on} = \frac{S_n}{S_o} 2^{(b_o - b_n)} \qquad b_o, b_n > 0 \tag{4.24}$$

where S_n and S_o denote the average r.m.s. values of segments of the two types; and b_n and b_o the numbers of bits per sample available for encoding the respective segments.

For example, in the simulation carried out, the average r.m.s. value of Type-3 segments is approximately 1.26 times that of Type-2 segments. Thus, in changing from the quantization of Type-2 to Type-3 segments, when the number of bits allocated per sample is increased by 1, the step-size should be reduced by a factor of 1.26/2=0.63. Conversely, the step-size should be increased by a factor of 1/0.63=1.59 in changing from the coding of Type-3 to Type-2 segments.

4.6.3 Results of Subjective Tests

In the subjective tests, two separate experiments were carried out. The first experiment was devoted to operation at 7.2 kbit/s and the second to operation at 9.6 kbit/s. In each experiment, twelve listeners were asked to compare pairs of sentences and to indicate which they preferred. One sentence was coded using the variable-bit-rate technique and the other using the fixed-bit-rate technique. With the variable-bit-rate technique, K_u was set equal to K_v and three different values: 5, 10 and ∞; were used. The speech material consists of three news announcements: two of which were by male speakers and one by a female speaker, with the duration of each announcement being approximately 20 s.

Pairs of sentences were selected and presented at random in a carefully balanced experiment. Each pair of sentences was presented twice and the subjects were instructed to record their preference after having heard the pair for the second time. During the course of each of the two experiments, each subject was presented with 9 different pairs for comparisons. Further details of the subjective tests are given in Appendix C.

In Tables 4.4 and 4.5, the mean percentage preference scores, $\emptyset$, for the variable-rate HFDC over fixed-rate HFDC are given together with the 90% confidence interval $[\emptyset_1, \emptyset_2]$. The values $[\emptyset_1, \emptyset_2]$ were obtained assuming that the distribution of the observations was normal with expectation $M\emptyset$ and variance $M\emptyset(1 - \emptyset)$ where M is the total number of listeners.

Table 4.4. Percentage Preference Score for Experiment 1
Average Bit Rate : 7.2 kbit/s

K_u	K_v	$\emptyset$	$\emptyset_1$	$\emptyset_2$
5	5	58.3	44.7	70.8
10	10	69.4	54.8	80.3
∞	∞	77.8	64.7	87.0

Table 4.5. Percentage Preference Score for Experiment 2
Average Bit Rate : 9.6 kbit/s

K_u	K_v	$\emptyset$	$\emptyset_1$	$\emptyset_2$
5	5	58.3	44.7	70.8
10	10	66.7	53.0	78.0
∞	∞	72.2	58.7	82.6

4.7 CONCLUSIONS

The results contained in Tables 4.4 and 4.5 show that using subjective preference as a measure, improvement can be realised through the use of the variable-rate hybrid frequency-domain coding scheme, especially when K_u and K_v are relatively large.

For the system operating at 7.2 kbit/s, a value of K_u=K_v=5 corresponds to an overall buffer delay of 178 ms, which is to be added to the encoding and decoding delays associated with the basic hybrid frequency-domain coding. Although the delay of 178 ms still remains quite long, it constitutes a significant reduction compared with the delay that would have occurred in the absence of the control strategy. As was indicated in Table 4.1, when K_u and K_v are large, which corresponds to an uncontrolled situation, 19.1 % of the segments of a

sample of speech were coded as Type-3 and examination shows that in this case, the delay is 900 ms. The control strategy thus leads in this particular case, to a reduction in delay by a factor of approximately 5.

In general, in the absence of the use of the control strategy, or some other techniques for reducing delays, the buffer delay can be in excess of 1 second, depending on the statistics of the actual speech sample. This long delay would make two-way communications difficult if not impossible. However, with the proposed scheme, it is possible to achieve quality improvement and at the same time, limit the buffer delay to a value commensurate with the system requirements. In particular, the technique described is of value in one-way speech transmission systems.

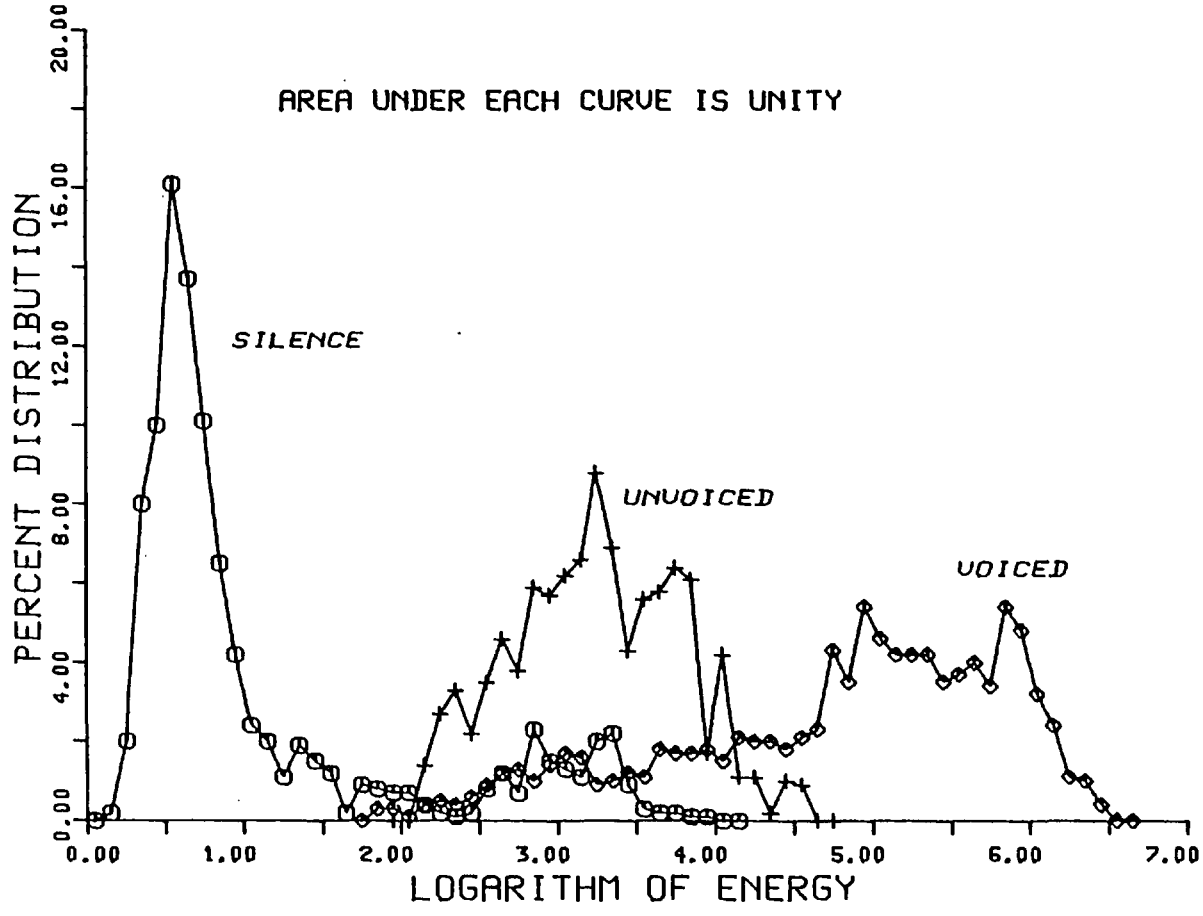


Fig.4.1 Distributions of energy within segments of voiced speech, unvoiced speech and silence. Segment length = 10 ms.

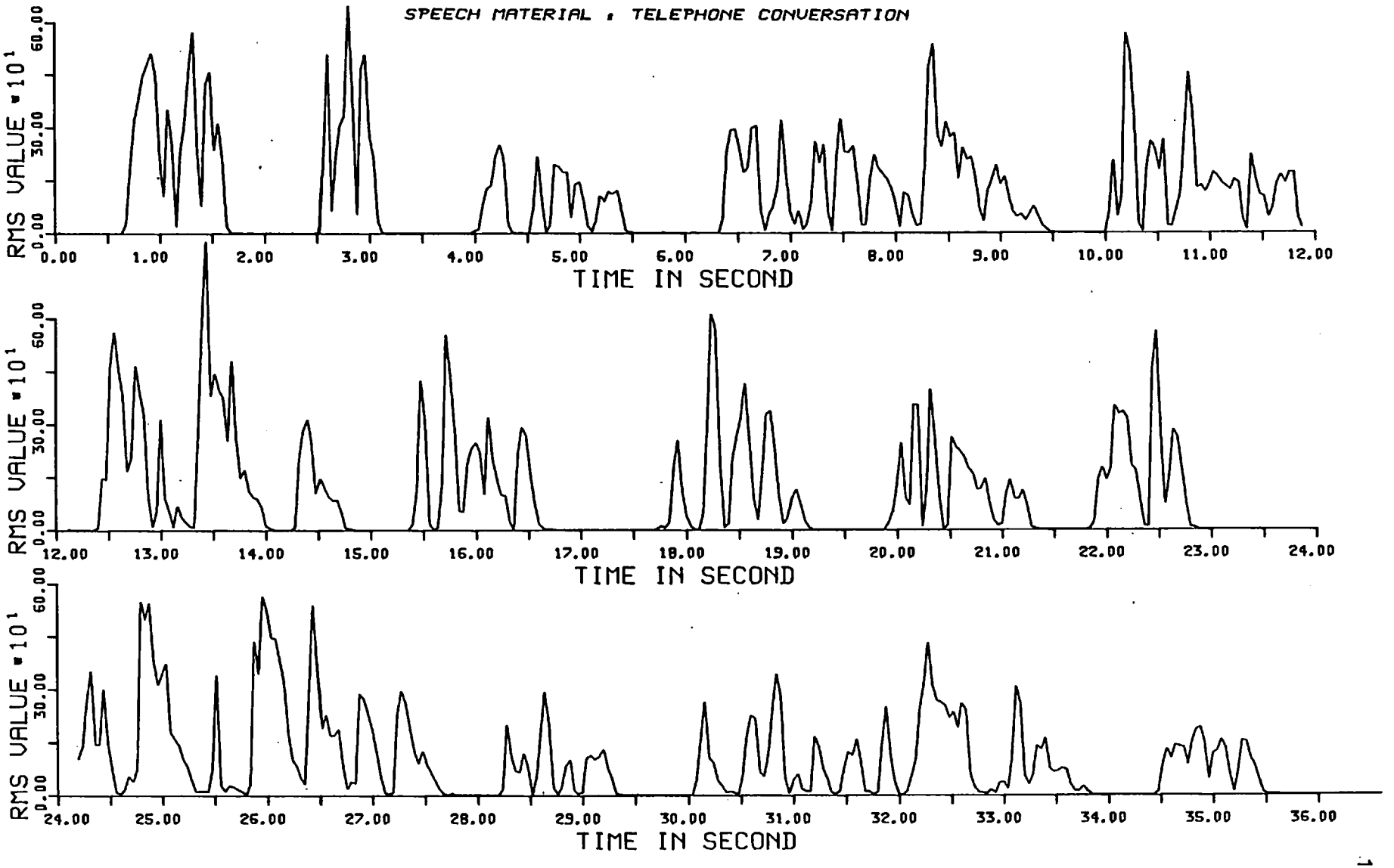


Fig.4.2 Variation of r.m.s. values of segments of the telephone conversation.

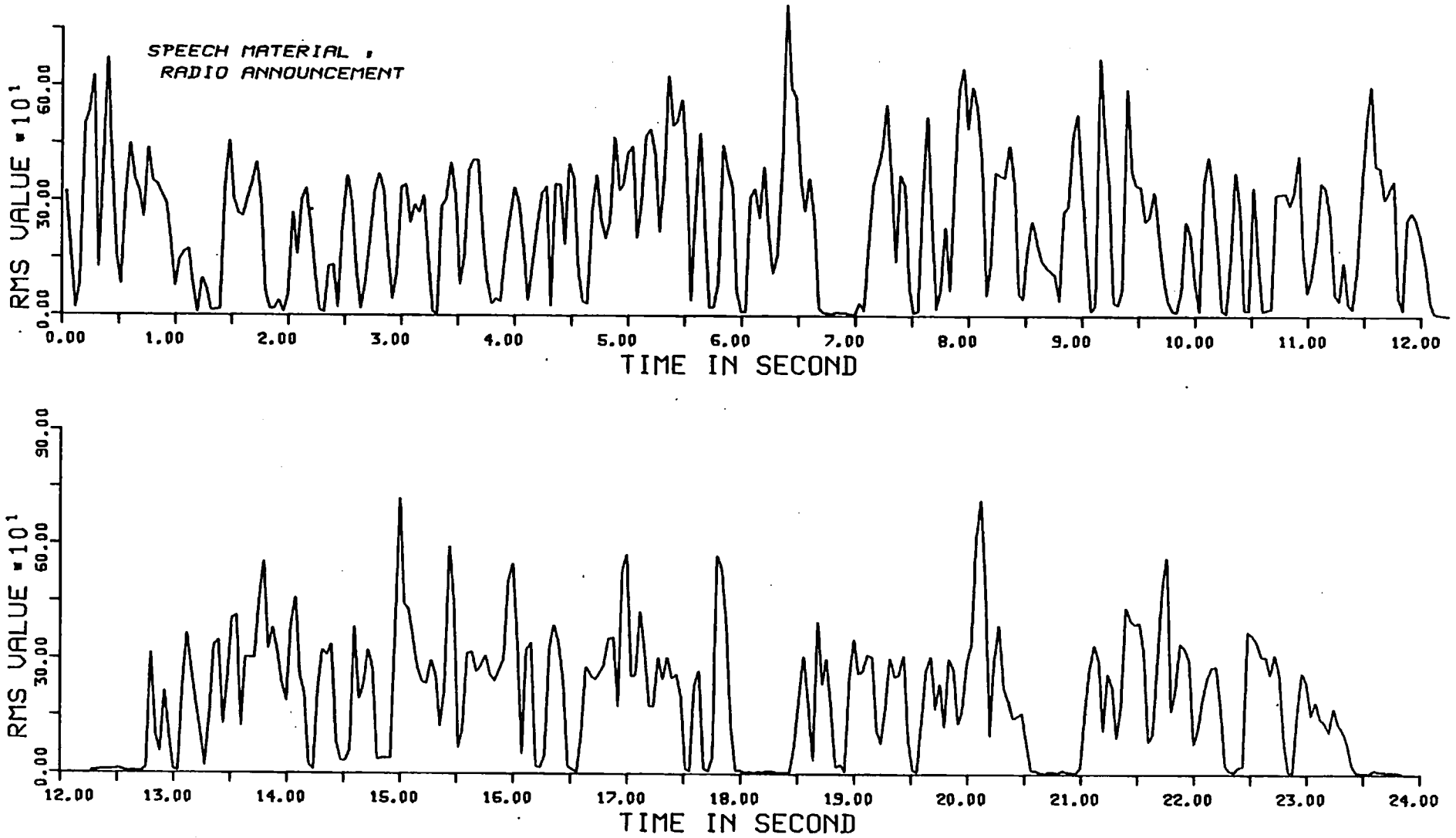


Fig.4.3 Variation of r.m.s. values of segments of the radio announcement.

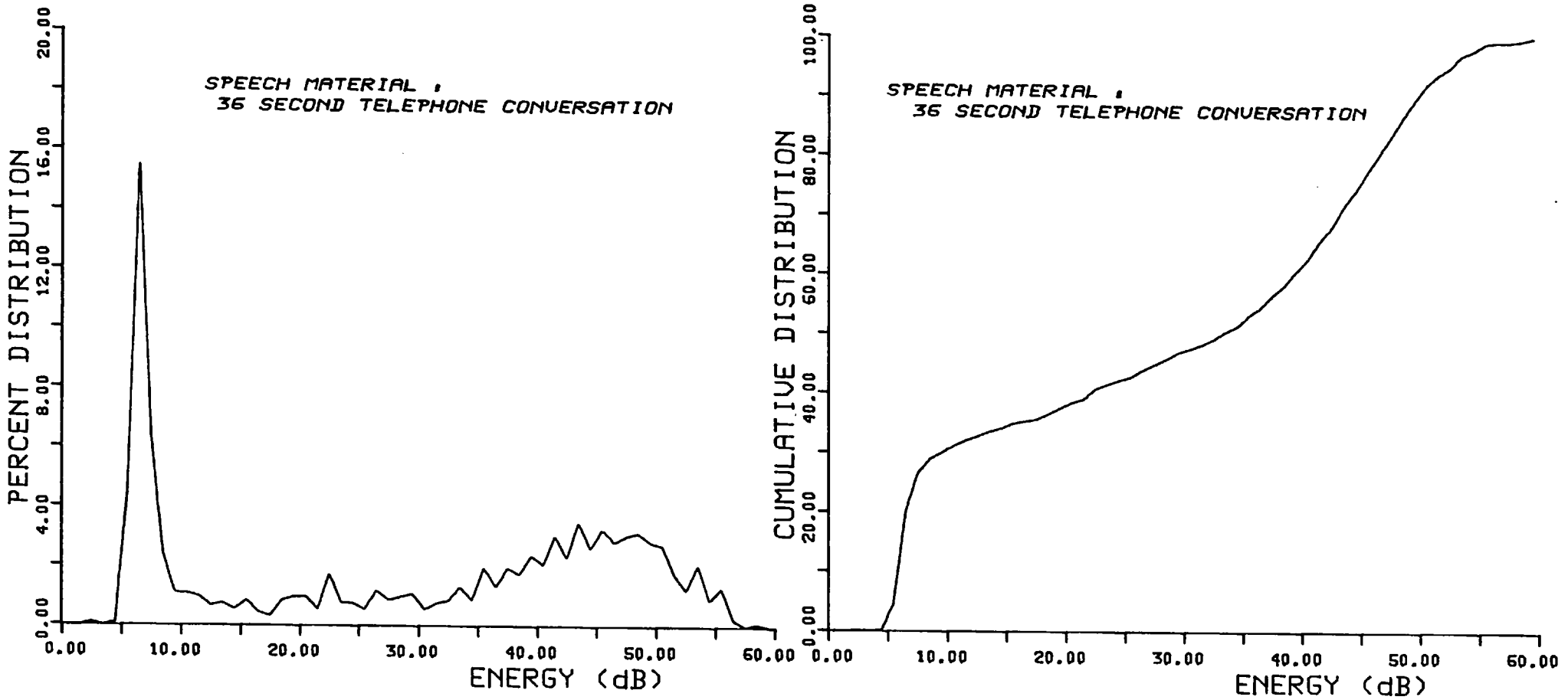


Fig.4.4 Distribution of energy within segments of the telephone conversation. Segment length = 10 ms.

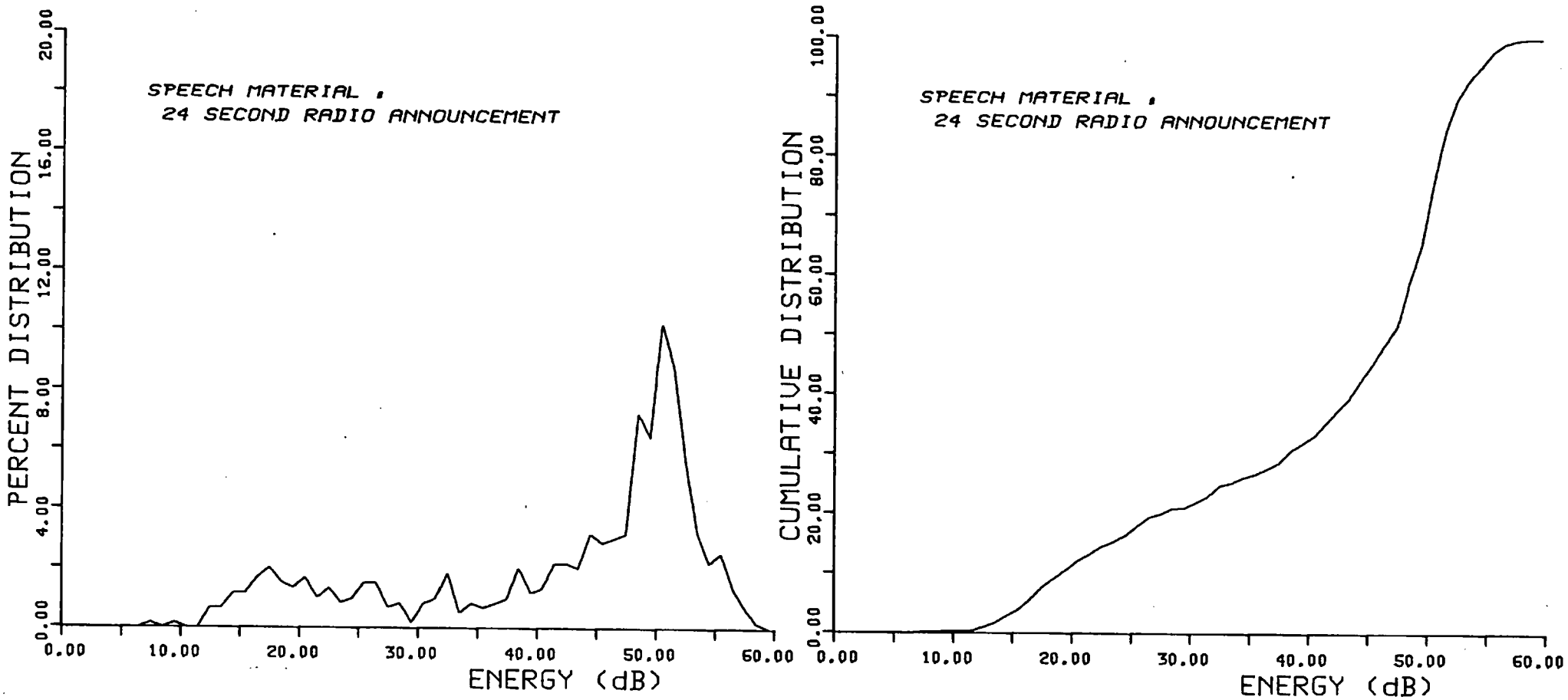
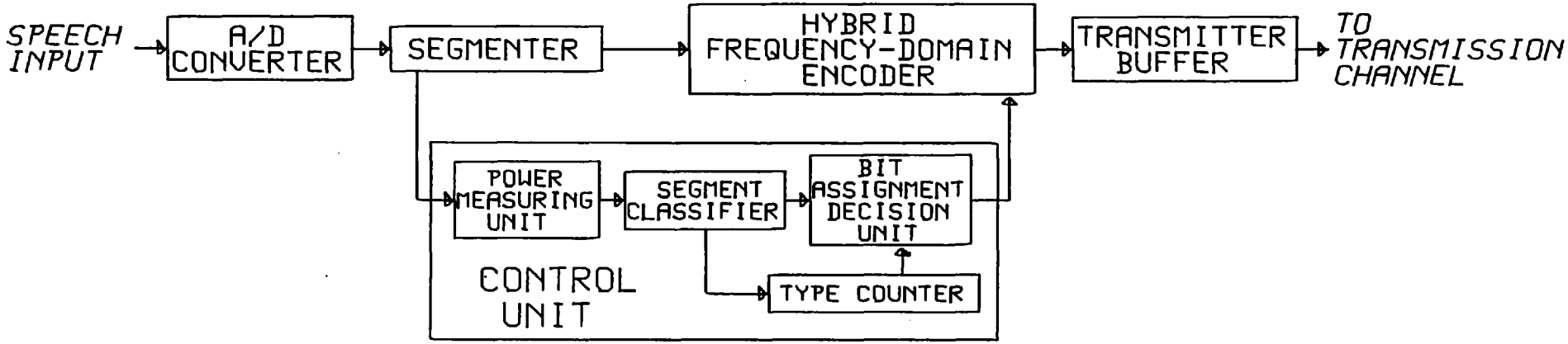
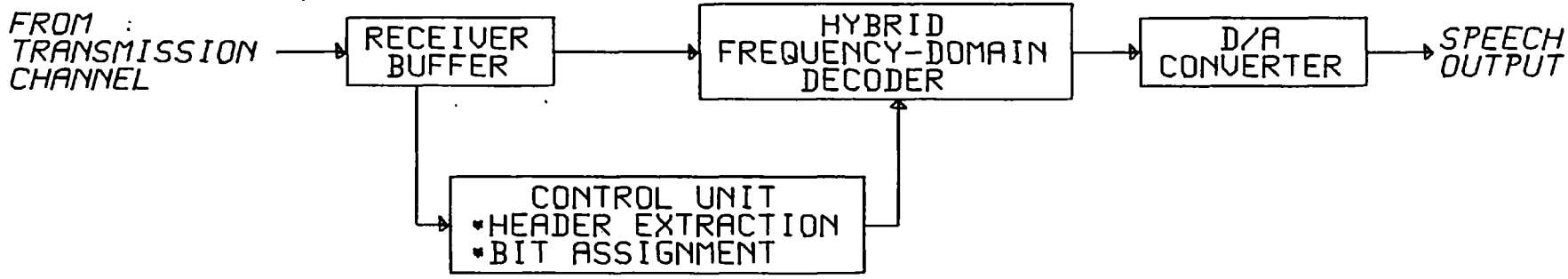


Fig.4.5 Distribution of energy within segments of the radio announcement. Segment length = 10 ms.



(a) Encoder



(b) Decoder

Fig.4.6 Schematic diagram of the variable-rate hybrid frequency-domain coder; (a) Encoder; (b) Decoder.

CHAPTER 5

DESIGN OF NON-RECURSIVE QUADRATURE MIRROR FILTERS

CHAPTER 5

DESIGN OF NON-RECURSIVE QUADRATURE MIRROR FILTERS

5.1 INTRODUCTION

In the sub-band coding of speech, the speech signal is filtered into several contiguous sub-bands and the signals in the different sub-bands are processed and coded differently before transmission. In this way a reduction of transmission bit rate is effected while the quality of the original speech signal is preserved.

As part of the signal processing involved within each sub-band, the signals are sampled at a rate less than the Shannon rate with the result that aliasing can occur when practical filters are employed. Aliasing cancellation can be achieved by using quadrature mirror filters (QMF) to perform the band splitting [62]. Although quadrature mirror filters can be implemented in a recursive form [107], and the length of filter for a given amount of out-of-band attenuation can thereby be reduced, there is the disadvantage that the phase of the signal is distorted. Non-recursive structures have linear phase characteristics and hence they better preserve the phase of the signal.

In this chapter a new method of designing non-recursive quadrature mirror filters is described. By the inclusion of

additional linear constraints, the method takes into consideration the deviation of the overall system frequency response as well as the passband ripple and stopband attenuation.

For completeness, the theory of quadrature mirror filters is described in Section 5.2. In Section 5.3, the general problem of quadrature mirror filter design is formulated and the various parameters involved are discussed. In Section 5.4, the specific design of quadrature mirror filters with two and four taps is examined and a discussion of the specific design of filters of larger order having parameters of particular interest is given in Section 5.5. Some conclusions are presented in Section 5.6.

5.2 THEORY OF QUADRATURE MIRROR FILTERS

The block diagram of a split-band system employing quadrature mirror filters as proposed by Estaban and Galand [62] is shown in Fig.5.1. The input signal x(t) is lowpass filtered to a bandwidth of B $(=f_s/2)$ and sampled at the Shannon rate f_s. The lowpass filter H_l and the highpass filter H_u are then used to split the spectrum of x(n) into lower and upper parts of equal bandwidth. The filtered signals $x_l(n)$ and $x_u(n)$, each of which occupies a bandwidth of approximately half of B, are decimated (down-sampled) by retaining only alternate samples. Compression encoding is usually carried out on the decimated signals $y_l(n)$ and $y_u(n)$ before transmission to the receiver.

At the receiver of the system, signals $y_l(n)$ and $y_u(n)$ are interpolated to obtain the signals $u_l(n)$ and $u_u(n)$. The interpolation takes the form of the addition of zeros between successive samples of

signals $y_l(n)$ and $y_u(n)$ respectively. The interpolated signals $u_l(n)$ and $u_u(n)$ are then filtered by P_l and P_u respectively and combined to obtain s(n) which is intended to be a replica of the original signal x(n).

The signal spectrum can be partitioned into several sub-bands by cascading the split-band system in a tree structure. An example is given in Fig.3.3.

For applications where it is desirable to implement the bandpass filter for each sub-band based on a single impulse response rather than a cascade of quadrature mirror filters, the equivalent filters can be derived based on successive convolutions of the impulse responses of the quadrature mirror filters with due considerations of the downsampling and upsampling. The equivalent filters for partitioning the signal spectrum into eight sub-bands of equal bandwidths were examined by Cheung and Winslow [108]. The underlying design in this case is still that of the basic quadrature mirror filters for the two-band system.

The use of quadrature mirror filters provides a practical method whereby interband aliasing distortion can be completely eliminated and an overall flat frequency response obtained.

Filters H_l and H_u whose frequency responses are mirror images of each other are referred to [62] as quadrature mirror filters. Thus if the z-transform $H_l(z)$ of H_l expressed in terms of its impulse response $h_l(n)$ is

$$H_l(z) = \sum_{n=0}^{N-1} h_l(n) z^{-n} \qquad (5.1)$$

then the z-transform $H_u(z)$ of H_u is given by

$$H_u(z) = \sum_{n=0}^{N-1} h_l(n)z^{-n}(-1)^n = H_l(-z) \tag{5.2}$$

If the z-transforms of $x(n)$ and $x_l(n)$ are $X(z)$ and $X_l(z)$ respectively, then

$$X_l(z) = H_l(z)X(z) \tag{5.3}$$

The z-transform $Y_l(z)$ of the decimated signal $y_l(n)$ and the z-transform $U_l(z)$ of the interpolated signal $u_l(n)$ are related [109] to the z-transform $X_l(z)$ respectively by

$$Y_l(z) = \frac{1}{2}\{X_l(z^{1/2}) + X_l(-z^{1/2})\} \tag{5.4}$$

$$U_l(z) = Y_l(z^2) \tag{5.5}$$

The z-transform $T_l(z)$ of the filtered signal $t_l(n)$ is related to $U_l(z)$ and the z-transform $P_l(z)$ of the filter P_l by

$$T_l(z) = P_l(z)U_l(z) \tag{5.6}$$

By combining (5.3) to (5.6), we obtain,

$$T_l(z) = \frac{1}{2}\{H_l(z)X(z) + H_l(-z)X(-z)\}P_l(z) \tag{5.7}$$

Similarly, the z-transform $T_u(z)$ of $t_u(n)$ is given by

$$T_u(z) = \frac{1}{2}\{H_u(z)X(z) + H_u(-z)X(-z)\}P_u(z) \tag{5.8}$$

By addition of $T_l(z)$ and $T_u(z)$, the z-transform $S(z)$ of the signal $s(n)$ is found to be

$$S(z) = R(z)X(z) + A(z)X(-z) \tag{5.9}$$

where

$$R(z) = \frac{1}{2}\{H_l(z)P_l(z) + H_u(z)P_u(z)\} \tag{5.10}$$

and

$$A(z) = \frac{1}{2}\{H_l(-z)P_l(z) + H_u(-z)P_u(z)\} \tag{5.11}$$

The term $A(z)X(z)$ in (5.9) arises from frequency aliasing.

By choosing $$P_l(z) = H_l(z) \tag{5.12}$$

and $$P_u(z) = -H_u(z) = -H_l(-z) \tag{5.13}$$

$$\begin{aligned} A(z) &= \tfrac{1}{2}\{H_l(-z)H_l(z) + H_u(-z)H_u(z)\} \\ &= \tfrac{1}{2}\{H_l(-z)H_l(z) - H_l(z)H_l(-z)\} \\ &= 0 \end{aligned} \tag{5.14}$$

the aliasing is eliminated.

When (5.12) and (5.13) are satisfied, the z-transform S(z) of the overall system is

$$S(z) = R(z)X(z) \tag{5.15}$$

where $$R(z) = \tfrac{1}{2}\{H_l^2(z) - H_l^2(-z)\} \tag{5.16}$$

The Fourier transform of R(z) is obtained by making the substitution $z = e^{j\omega}$ in (5.16).

Thus $$R(e^{j\omega}) = \tfrac{1}{2}\{H_l^2(e^{j\omega}) - H_l^2(e^{j(\pi-\omega)}\} \tag{5.17}$$

If H_l is a symmetrical filter, then its Fourier transform is given by,

$$H_l(e^{j\omega}) = H_l^*(e^{j\omega})\exp(-j(N-1)\omega/2) \tag{5.18}$$

and hence

$$\begin{aligned} R(e^{j\omega}) &= \tfrac{1}{2}\{H_l^{*2}(e^{j\omega})e^{[-j2(N-1)\omega/2]} - H_l^{*2}(e^{j(\pi-\omega)})e^{[-j2(N-1)(\omega+\pi)/2]}\} \\ &= \tfrac{1}{2}\{H_l^{*2}(e^{j\omega}) - H_l^{*2}(e^{j(\pi-\omega)})e^{[-j(N-1)\pi]}\}\exp[-j(N-1)\omega] \end{aligned} \tag{5.19}$$

The term $\exp(-j(N-1)\pi) = \cos(N-1)\pi - j\sin(N-1)\pi$ is equal to -1 for N even and +1 for N odd.

Hence for N even,

$$R(e^{j\omega}) = \frac{1}{2}\{H_l^{*2}(e^{j\omega})+H_l^{*2}(e^{j(\pi-\omega)})\}\exp[-j(N-1)\omega] \tag{5.20}$$

If $$H_l^{*2}(e^{j\omega}) + H_l^{*2}(e^{j(\pi-\omega)}) = 2 \tag{5.21}$$

then $$R(e^{j\omega}) = \exp(-j(N-1)\omega) \tag{5.22}$$

and from (5.15) $$S(e^{j\omega}) = \exp(-j(N-1)\omega)X(e^{j\omega}) \tag{5.23}$$

So $$S(z) = X(z)z^{-(N-1)} \tag{5.24}$$

or $$S(n) = x(n-N+1) \tag{5.25}$$

Thus, apart from a delay of N-1 samples, a perfect reconstruction of the original signal is obtained. The design constraints for filters yielding an undistorted reconstruction of the original signal can thus be summarized as follows:

(a) H_l : a symmetrical non-recursive lowpass filter of even order

(b) $H_u(z) = H_l(-z)$, where $H_u(z)$ is the z-transform of H_u

(c) $P_l(z) = H_l(z)$, where $P_l(z)$ is the z-transform of P_l

(d) $P_u(z) = -H_u(z)$, where $P_u(z)$ is the z-transform of P_u

(e) $H_l^{*2}(e^{j\omega}) + H_l^{*2}(e^{j(\pi-\omega)}) = 2$

where $H_l^{*}(e^{j\omega})$ is related to the Fourier transform $H_l(e^{j\omega})$, of $H_l(z)$ by $H_l^{*}(e^{j\omega}) = H_l(e^{j\omega})\exp(+j(N-1)\omega/2)$

When these constraints are satisfied, the output signal s(n) is an exact replica of x(n), apart from a delay of (N-1) samples. That is, s(n) = x(n-N+1).

Constraints (a) to (d) are necessary for the cancellation of

frequency aliasing, with the additional constraint (e) being necessary in order to ensure a flat overall response. Constraint (e) is difficult to satisfy in practice and it is this constraint that is considered in detail in this chapter.

5.3 PROBLEM FORMULATION

The aim is to design non-recursive quadrature mirror filters which, for a given number of taps and a given transition width, minimize the passband ripple, maximize the stopband attenuation and minimize the overall system deviation. Only the design of the lowpass filters will be considered since the design of other filters can be readily derived from the lowpass design.

Let B_p be the set of passband frequencies f for which $\{0 \leqslant f \leqslant F_p\}$, and let B_s be the set of stopband frequencies f for which $\{F_s \leqslant f \leqslant 0.5\}$, where, F_p and F_s are the cutoff frequencies of the passband and stopband respectively of the lowpass filter.

Further, let B represent the set of frequencies f such that $\{0 \leqslant f \leqslant 0.5\}$.

If, in the passband, the desired magnitude is $\sqrt{2}$, and δ_p is the maximum deviation allowed, where $\delta_p \geqslant 0$, then it is desired that

$$\sqrt{2} - \delta_p \leqslant H_l^*(e^{j2\pi f}) \leqslant \sqrt{2} + \delta_p \qquad f \in B_p \qquad (5.26)$$

If in the stopband, the desired magnitude is 0, and δ_s denotes the maximum allowable ripple, with $\delta_s \geqslant 0$, then it is desired that

$$-\delta_s \leqslant H_l^*(e^{j2\pi f}) \leqslant \delta_s \qquad f \in B_s \qquad (5.27)$$

Further, it is desired that

$$2 - d_o \leqslant [H_l^*(e^{j2\pi f})]^2 + [H_l^*(e^{j\pi(1-2f)})]^2 \leqslant 2 + d_o \quad f \in B \qquad (5.28)$$

where d_o is a nonnegative number which denotes the maximum deviation in the overall response.

The objective is to find the impulse response h(n) that minimizes the quantity $\max\{\delta_p, \delta_s, d_o\}$ for a given set of N, F_p and F_s.

On account of the non-linear constraints set out in (5.28) above, linear programming methods [110,111] cannot be used to solve the global optimization problem. A method will now be described which provides a better solution than one in which the constraints of (5.28) are ignored.

At the mid-point of the filter frequency scale, that is, when $f_m = f_s/4 = 1/4$ for $f_s = 1$, the middle term of (5.28) becomes

$$[H_l^*(e^{j2\pi/4})]^2 + [H_l^*(e^{j\pi(1-2(1/4))})]^2$$

$$= [H_l^*(e^{j\pi/2})]^2 + [H_l^*(e^{j\pi/2})]^2$$

$$= 2[H_l^*(e^{j\pi/2})]^2$$

and constraints (5.28) become

$$2 - d_o \leqslant 2[H_l^*(e^{j\pi/2})]^2 \leqslant 2 + d_o$$

or

$$1 - d_o/2 \leqslant [H_l^*(e^{j\pi/2})]^2 \leqslant 1 + d_o/2 \qquad (5.29)$$

On adding $(d_o/4)^2$ to the three terms in (5.29), the following result is obtained

$$1 - (d_o/2) + (d_o/4)^2 \leqslant [H_l^*(e^{j\pi/2})]^2 + (d_o/4)^2$$

$$\leqslant 1 + (d_o/4) + (d_o/4)^2$$

i.e. $$(1 - (d_o/4))^2 \leqslant [H_l^*(e^{j\pi/2})]^2 + (d_o/4)^2$$

$$\leqslant (1 + (d_o/4))^2 \qquad (5.30)$$

For cases of practical interests, $(d_o/4)^2 << [H_l^*(e^{j\pi/2})]^2$,

so $$(1 - (d_o/4)) \leqslant |H_l(e^{j\pi/2})| \leqslant (1 + (d_o/4)) \qquad (5.31)$$

It is important to note that the constraints of (5.31) are linear constraints.

By imposing the linear constraints of (5.31) instead of the non-linear constraints of (5.28), a solution can be obtained using linear programming technique. The solution minimizes d_o as well as δ_s and δ_p, although it may not be a global minimum. As will be seen later, the solution is very satisfactory for most practical purposes and should further optimization be required, the solution can be used as a starting point in an iterative process.

The deviation d_o of the overall system for frequencies less than F_p, is related to δ_s and δ_p, as is the overall deviation d_o for frequencies greater than F_s. That is, for $f \in B_p \cup B_s$,

$$(\sqrt{2} + \delta_p)^2 + \delta_s^2 \leqslant 2 + d_o \qquad (5.32)$$

$$2 - d_o \leqslant (\sqrt{2} - \delta_p)^2 + \delta_s^2 \qquad (5.33)$$

On expanding (5.32), it can be seen that,

$$d_o \geq 2\sqrt{2}\,\delta_p + \delta_p^2 + \delta_s^2 \tag{5.34}$$

In most practical cases, $\delta_p^2 + \delta_s^2 \ll 2\sqrt{2}\delta_p$, thus the expected minimum d_o is approximately $2\sqrt{2}\delta_p$.

In order that the problem can be formulated in a form suitable for solution by linear programming, δ_p is set to be $K\delta_s$, where K is a constant. A set of frequencies $\{f_{pc}\}$ and $\{f_{sc}\}$ are chosen from B_p and B_s respectively. The greater the number of frequencies chosen, that is, the higher the grid density, the more accurate are the results of the computation. However, the complexity of the computation also increases with increasing grid density.

Using the notations in Appendix D, the problem can be formulated as : Find the set of b(i), i=1, 2,...(N/2), subject to

(i) $$\delta_s \geq 0 \tag{5.35}$$

(ii) $$-K\delta_s + \sum_{n=1}^{N/2} b(n)\cos\{2\pi f(n-(1/2))\} \leq \sqrt{2} \qquad f \in \{f_{pc}\} \tag{5.36}$$

(iii) $$-K\delta_s - \sum_{n=1}^{N/2} b(n)\cos\{2\pi f(n-(1/2))\} \leq -\sqrt{2} \qquad f \in \{f_{pc}\} \tag{5.37}$$

(iv) $$-\delta_s + \sum_{n=1}^{N/2} b(n)\cos\{2\pi f(n-(1/2))\} \leq 0 \qquad f \in \{f_{sc}\} \tag{5.38}$$

(v) $$-\delta_s - \sum_{n=1}^{N/2} b(n)\cos\{2\pi f(n-(1/2))\} \leq 0 \qquad f \in \{f_{sc}\} \tag{5.39}$$

(vi) $$-(1/\sqrt{2})\,K\delta_s + \sum_{n=1}^{N/2} b(n)\cos\{\pi f_s(n-(1/2))/2\} \leq 1 \tag{5.40}$$

(vii) $$-(1/\sqrt{2})\,K\delta_s - \sum_{n=1}^{N/2} b(n)\cos\{\pi f_s(n-(1/2))/2\} \leq -1 \tag{5.41}$$

such that $-\delta_s$ is maximized, i.e. δ_s is minimized.

In (5.35) to (5.41) above, N is an even integer and $f_s = 1$.

Assume that δ_s is the minimum value obtained for a given set of F_p, F_s and N, then from (5.34) the lower bound for d_o is,

$$\min\{d_o\} \geq 2\sqrt{2}K\delta_s + \delta_s^2 + K^2\delta_s^2 \qquad (5.42)$$

In the discussions which follow, the solution with constraints (5.35) to (5.39) only will be referred to as solution A, and the solution with the complete set of constraints (5.35) to (5.41) will be referred to as solution B. It should be noted that constraints (5.40) and (5.41) are based on constraint (e) of Section 5.2.

5.4 QUADRATURE MIRROR FILTERS OF TWO AND FOUR TAPS

The design of quadrature mirror filters with N=2 and N=4 will now be considered. These designs provide an insight into the general design of quadrature mirror filters, in addition to being valuable in their own right.

When N=2, there is only one design variable, viz. h(0), since h(1)=h(0).

In this case, the magnitude of the frequency response is given by

$$H_l^*(e^{j\omega}) = 2h(0)\cos(\omega/2) \qquad (5.43)$$

and the magnitude of the frequency response of the mirror filter is

$$H_u^*(e^{j\omega}) = 2h(0)\cos[(\pi-\omega)/2]$$

$$= 2h(0)\cos[(\pi/2)-(\omega/2)]$$

$$= 2h(0)\sin(\omega/2) \qquad (5.44)$$

The magnitude of the overall response is

$$[H_l^*(e^{j\omega})]^2 + [H_u^*(e^{j\omega})]^2$$

$$= [2h(0)\cos(\omega/2)]^2 + [2h(0)\sin(\omega/2)]^2$$

$$= 4h^2(0)[\cos^2(\omega/2) + \sin^2(\omega/2)]$$

$$= 4h^2(0) \qquad (5.45)$$

If $h(0) = (1/\sqrt{2})$ (5.46)

then $R^*(e^{j\omega}) = 2$ for all ω

and a perfect solution is obtained.

When N=4, there are two variables, h(0) and h(1). The discrete frequency spectrum is as shown in Fig.5.2. As explained in Appendix D, point 3 in the figure is fixed at zero as a characteristics of symmetrical FIR filters. Thus the values at two points only are to be decided upon. As point 2 represents the magnitude at the mid-point of the filter frequency scale, this point is chosen to be equal to unity and point 1 being in the passband, is chosen to be $\sqrt{2}$. With this selection,

$$H_l^*(e^{j0}) = 2h(0) + 2h(1) = \sqrt{2} \qquad (5.47)$$

$$H_1^*(e^{j\pi/2}) = 2h(0)\cos(3\pi/4) + 2h(1)\cos(\pi/4)$$

$$= [\ -2h(0) + 2h(1)\ ]/\sqrt{2} = 1$$

i.e. $$-2h(0) + 2h(1) = \sqrt{2} \qquad (5.48)$$

On solving (5.47) and (5.48), it is found that $h(0) = 0$ and $h(1) = 1/\sqrt{2}$ which reduces to the case for N=2.

Thus, if the size of the transition width, the stopband attenuation and the passband ripple are unimportant, then a perfect quadrature mirror filter that gives the required overall response can always be obtained.

5.5 QUADRATURE MIRROR FILTERS OF LARGER NUMBER OF TAPS

As explained above, perfect quadrature mirror filters are obtainable with the desired overall response if the passband ripple, stopband attenuation and the size of the transition width of the filters are ignored. In practice, these parameters cannot be ignored and it is desirable to have filters that not only keep the overall response close to the desired value, but also have a low passband ripple, a large stopband attenuation and a small transition width.

The design technique as proposed and formulated in Section 5.3 has been used in the design of quadrature mirror filters for N=8, 16 and 32 for various values of transition width and various values of K's of interest. Filters of any even order can be designed using this method, although for large N, the complexity of the problem increases and the time taken for computer solution increases rapidly. The design for the case of N=16 will now be considered in some detail.

The grid density used in the solution was chosen to be 10, that is, the frequency scale normalized with respect to the sampling frequency f_s, was divided into 16x10=160 divisions for the case in which N=16.

The normalized transition width (TW), defined as $(F_s-F_p)/f_s$, was varied in steps from 1/8 to 1/4 and the cutoff frequency, F_p, of the filters was varied over the range of interest.

Of particular interest is the peak-to-peak deviation of the overall response from the desired value as well as the stopband attenuation. The stopband attenuation, denoted by E_s, was calculated as

$$E_s = 20(\log_{10}\delta_s - \log_{10}\sqrt{2}) \qquad (5.49)$$

A measure of the peak-to-peak deviation of the overall response was obtained as follows:

The frequency response of the filter was obtained by first adding 256-(16/2)=248 zeros to both ends of the impulse response as shown in Fig.5.3. The Fourier transform of this augmented impulse response gives 512 samples of the frequency response of the filter over the frequency range from 0 to 1. By adding the square of the mirror image of this frequency response to the square of the response itself, the response of the overall system is obtained. Mathematically,

$$R_m = |H(e^{j2\pi n/512})|^2 + |H(e^{j2\pi(256-n)/512})|^2$$

$$n=0,1,\ldots,256 \qquad (5.50)$$

Let R_{max} and R_{min} represent the maximum and the minimum values respectively of the resulting overall response over the entire filter frequency range. The maximum peak-to-peak deviation of the overall response, denoted by D_o, is defined as

$$D_o = 20(\log_{10}R_{max} - \log_{10}R_{min}) \tag{5.51}$$

The results of D_o and E_s could be presented as a function of the cutoff frequency. However, a better way of presenting the data is obtained by transforming the cutoff frequency in terms of percentage of transition width deviation from the mid-point, i.e. f_m=0.25, of the filter frequency scale. This value, denoted by f_c, is calculated from the normalized cutoff frequency F_p as

$$f_c = 100(0.25 - F_p)/TW \tag{5.52}$$

The results of D_o and E_s as a function of f_c, for various values of transition width (TW), with K being chosen to be 1, are shown in Figs.5.4 and 5.5 respectively.

To give some idea of the results obtained when the mid-point constraints are neglected (solution A), the corrresponding D_o and E_s are shown in Figs.5.6 and 5.7.

For various values of K, with TW fixed, the variation of D_o and E_s as a function of f_c is shown in Figs.5.8 and 5.9 under the conditions for solution B and in Figs.5.10 and 5.11 under the conditions for solution A.

Fig.5.12 provides a comparison of the maximum one-sided deviation D_{os}, with K=1 and TW=5/32 for solutions A and B, and

Fig.5.13 provides a comparison of E_s with K=1 and TW=5/32 for solutions A and B. The maximum one-sided deviation D_{os} is calculated using

$$D_{os} = \text{maximum of } \{20\log_{10}(R_{max}/2); 20\log_{10}(2/R_{min})\} \tag{5.53}$$

and the bounds are calculated as $20[\log_{10}((2 + d_a)/2)]$,

where $$d_a = 2\sqrt{2K}\delta_s + \delta_s^2 + K^2\delta_s^2 \tag{5.54}$$

From these figures, the following important observations can be made:

(1) The minimum values of D_o for a given K and TW occur at about the same cutoff frequency for both solution A and solution B, but the minimum values of D_o are not necessarily the same at the points of minimum.

(2) As TW is increased, the cutoff frequency, f_c, at which the minimum D_o occurs moves closer to 50%. It should be noted that a value of f_c of 50% means that the transition band is symmetrical about the mid-point of the filter frequency scale.

(3) In general, for a given value of K, increasing TW results in a decreasing D_o for solution B, but this is not so for solution A.

(4) For solution B with a given value of TW, decreasing K results in a smaller attenuation (i.e. larger E_s) and also a decreasing value of D_o. For solution A, decreasing K results in a smaller stopband attenuation but the value of D_o does not necessarily decrease.

(5) Solution B achieves a lower value of D_o at the expense of a large increase of E_s when compared with solution A. The graphs of E_s against f_c are concave upwards for solution B, while those for solution A are concave downwards (convex).

(6) From Fig.5.12 it can be seen that if the mid-point constraints are included, then the minimum one-sided deviation, D_{os}, is very close to the global minimum since the global optimal value of D_{os}, for the range of f_c of interests, lies between the curve representing the results for solution B and the curve for the bound to solution B. The difference between these curves is small. If, however, the mid-point constraints are not imposed (solution A), then the difference between the results and the bound is large.

Figures 5.14 to 5.27 show the results of simulation for the cases of N=8 and N=32. It can be seen that the above observations also apply in these cases.

As an example, the lowpass filter response and the overall system response for a particular case in which N=32 are shown in Figs.5.28 and 5.29 respectively.

5.6 CONCLUSIONS

The introduction of additional linear constraints at the mid-point of the frequency scale in the linear programming formulation of the design of the quadrature mirror filters provides a much better overall response than that obtained when the mid-point constraints are neglected. The deviation of the overall response is very much reduced but the reduction of the ripple in the overall response has the effect

of reducing the stopband attenuation of the filter. Different combinations of stopband attenuation and deviation of the overall response can easily be obtained by using a different multiplication factor K.

A global optimal solution to the design involves non-linear constraints, and in an approach to the problem, Johnston [112] made use of the Hookes and Jeaves direct search algorithm [113,114]. The algorithm essentially employs a 'brute force' approach by sequential examination of the trial solutions involving comparison of each trial solution with the best obtained up to that time. The choice of the starting point and step sizes are critical and manual intervention is required to ensure that the solution is the right one. The method introduced in this chapter involves only linear constraints, is relatively simple, and gives a solution which is satisfactory in most cases of practical interest, and is very close to the global optimum. Furthermore, the results can be used as a starting point for non-linear algorithm should a global optimal solution be required.

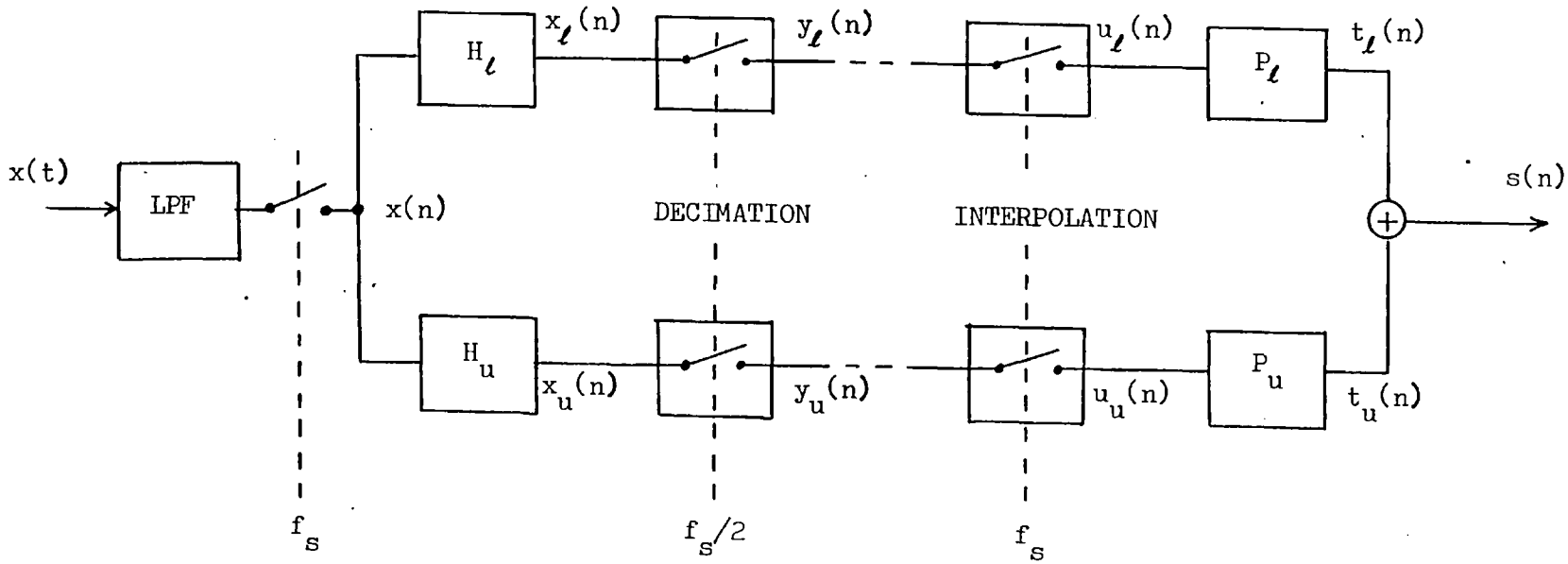


Fig.5.1 Illustration of two-band splitting using quadrature mirror filters.

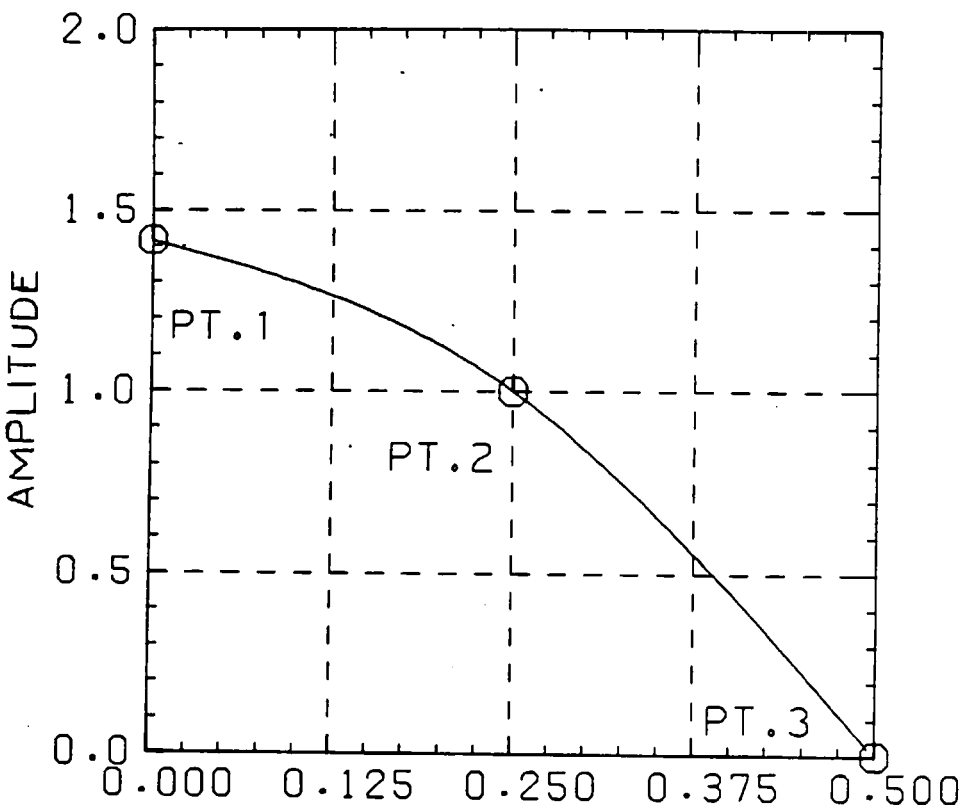


Fig.5.2 Frequency response of 4-point FIR lowpass filter.

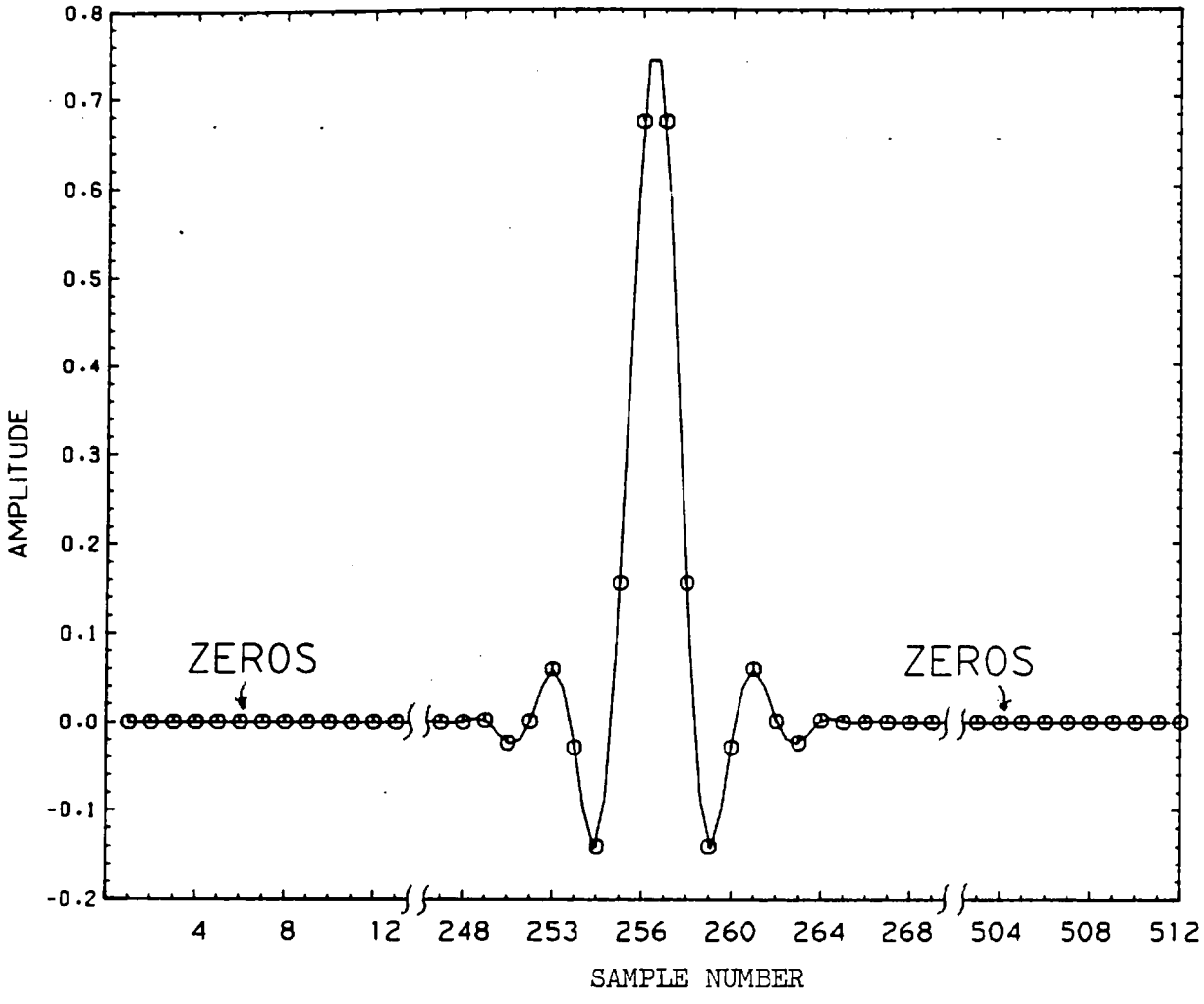


Fig.5.3 Augmented impulse response.

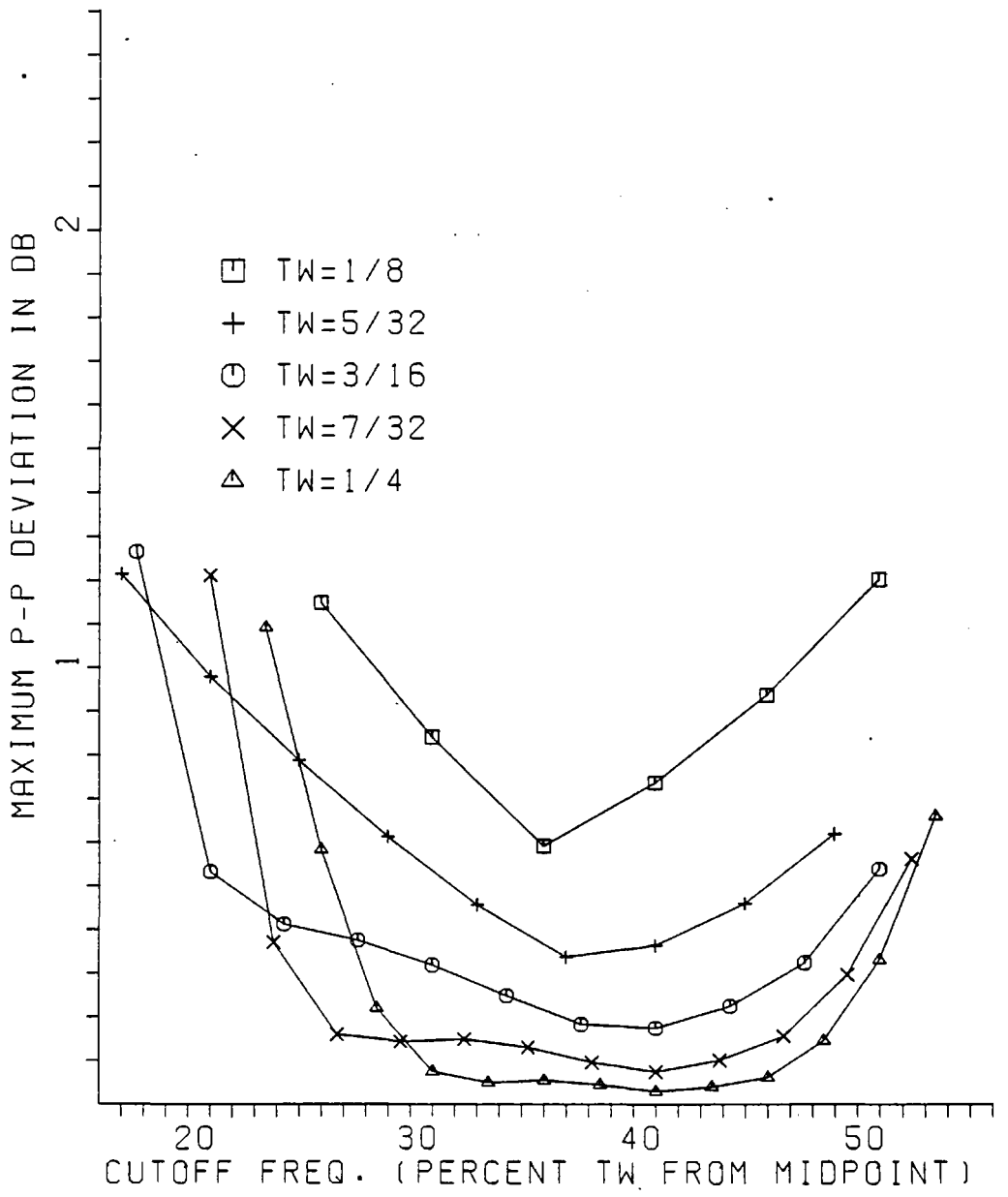


Fig.5.4 Maximum peak-to-peak deviation of the overall response as a function of cutoff frequency expressed in percentage of transition width from $f_s/4$. (solution B, N=16, K=1).

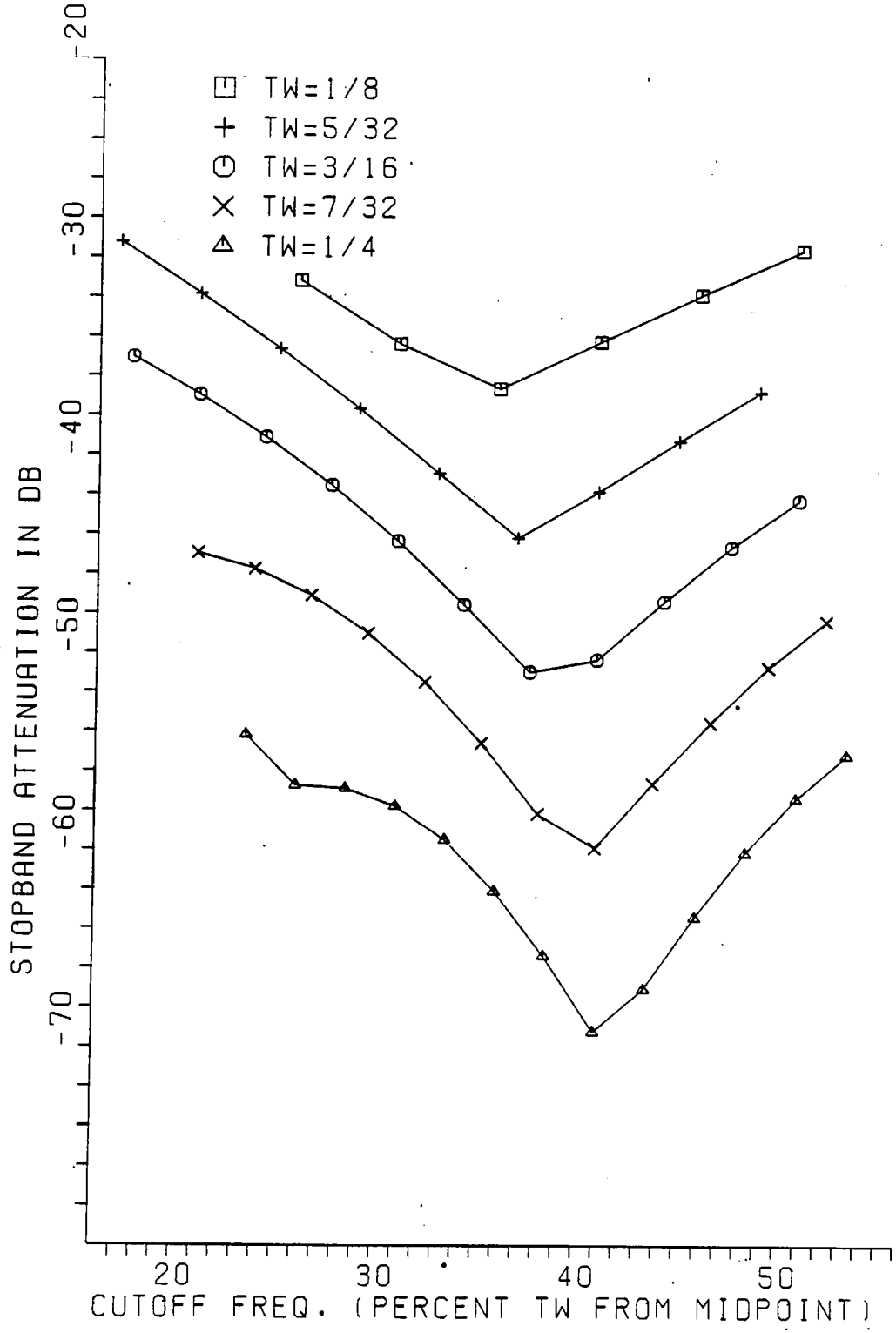


Fig.5.5 Stopband attenuation as a function of cutoff frequency expressed in percentage of transition width from $f_s/4$. (solution B, N=16, K=1).

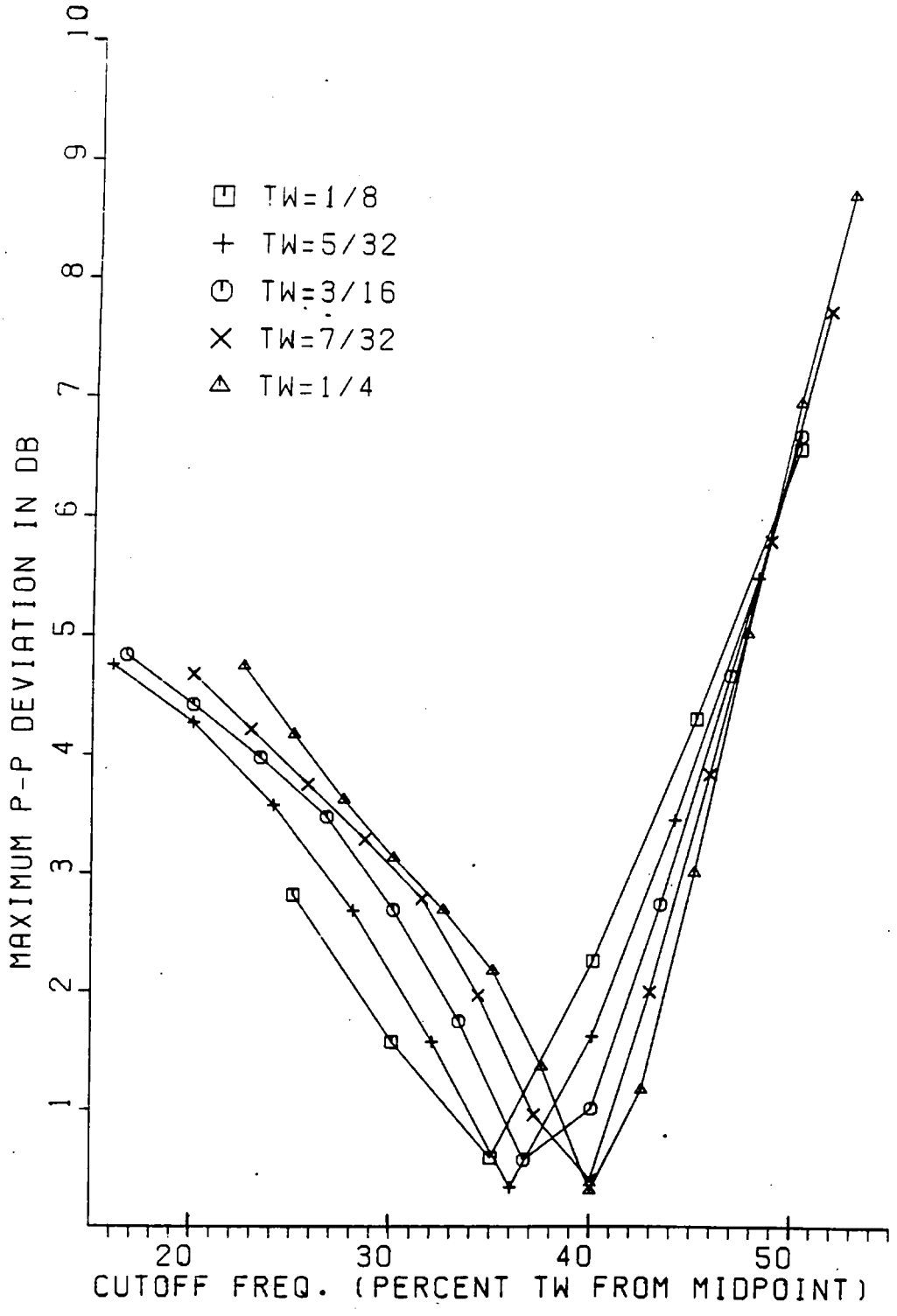


Fig.5.6 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution A, N=16, K=1).

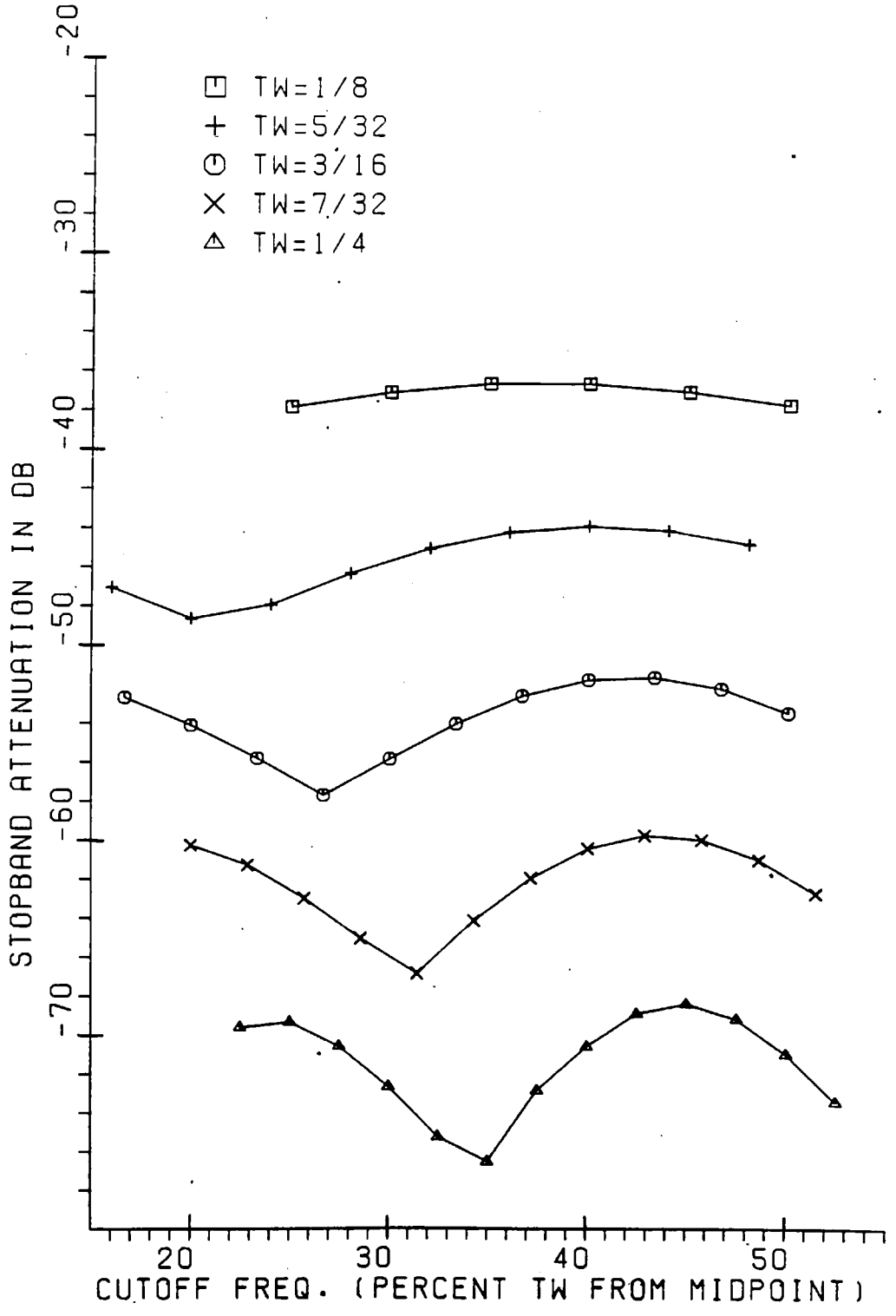


Fig.5.7 Stopband attenuation for various transition widths. (solution A, N=16, K=1).

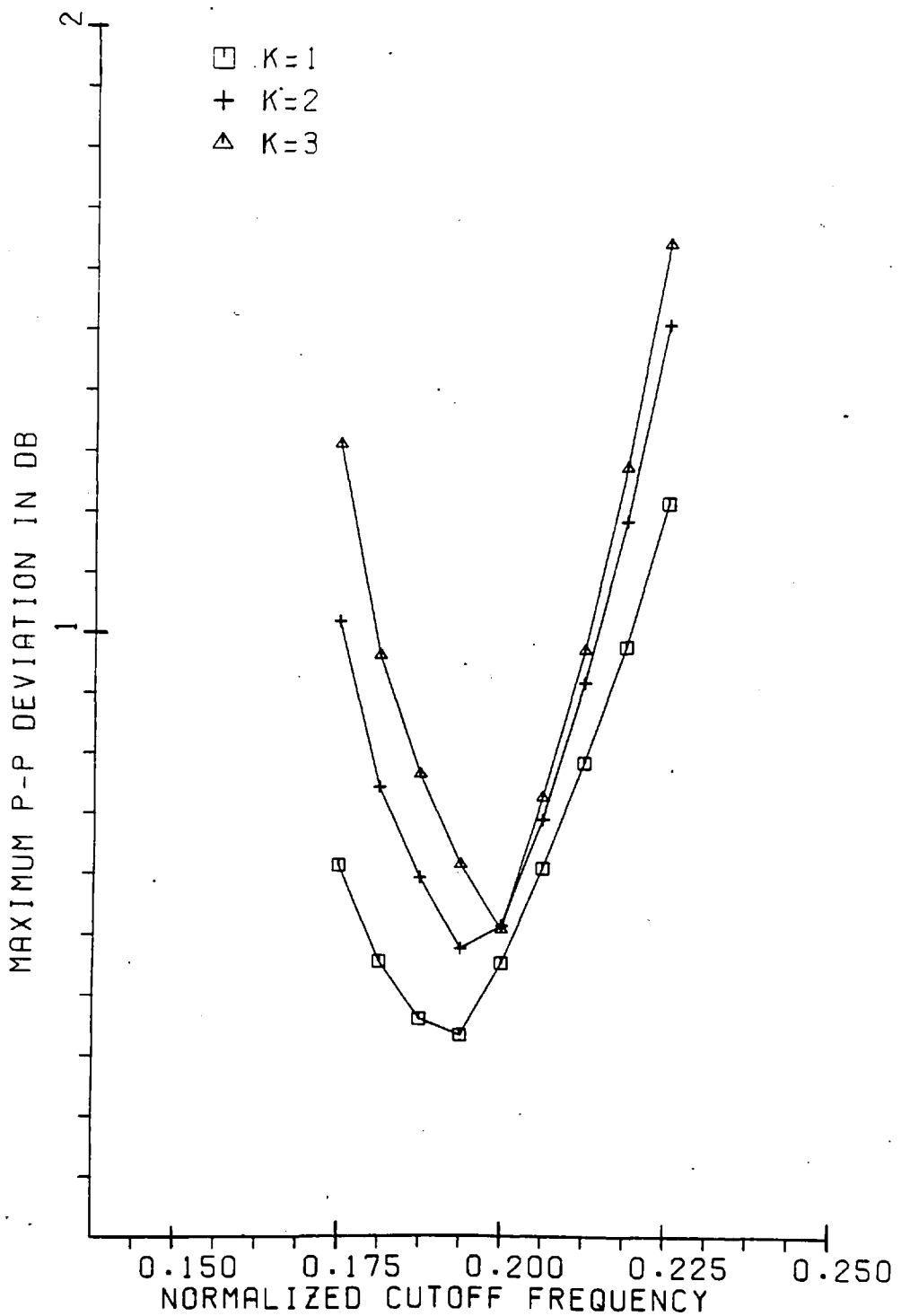


Fig.5.8 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N = 16, TW = 5/32).

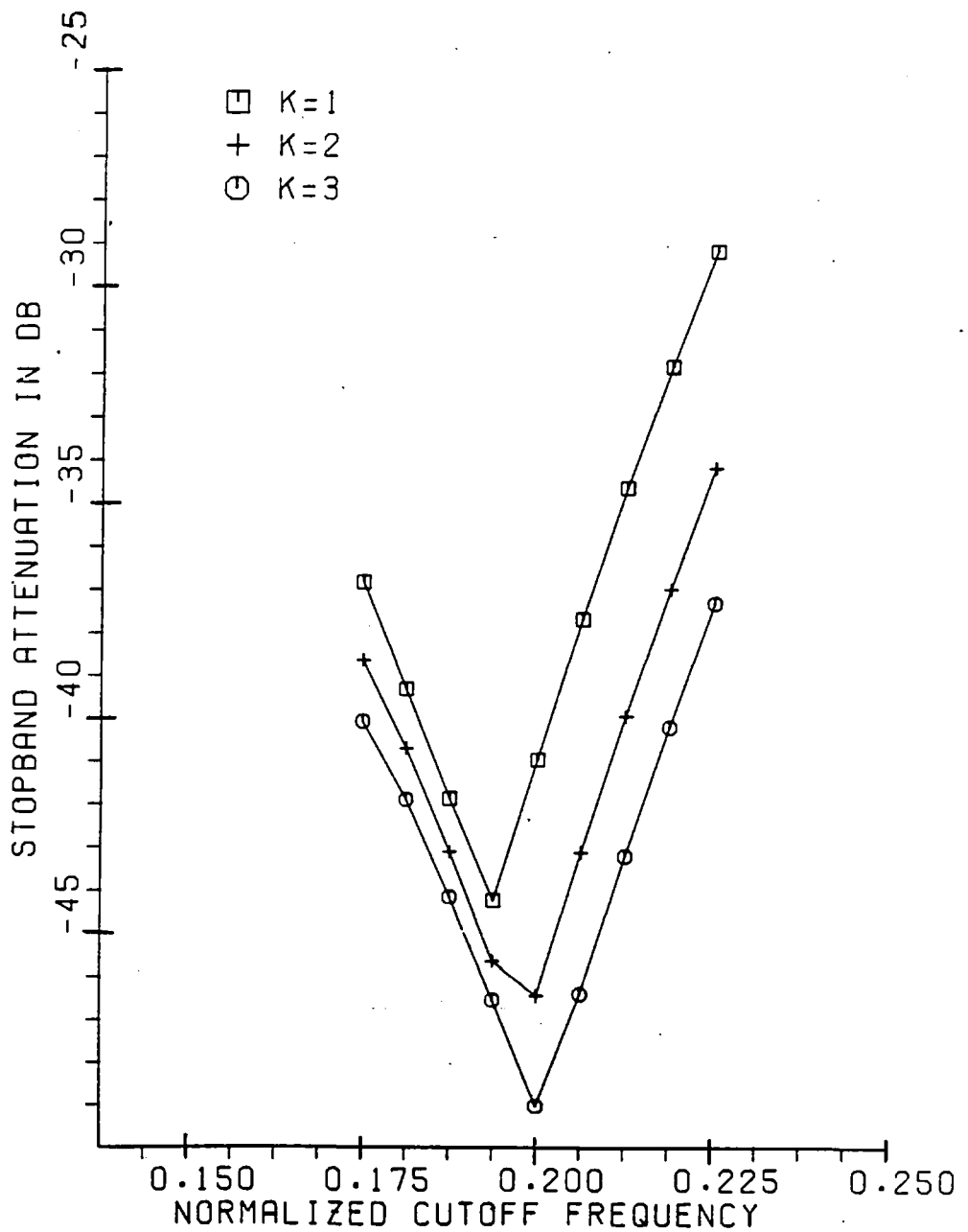


Fig.5.9 Stopband attenuation for various values of K. (solution B, N=16, TW=5/32).

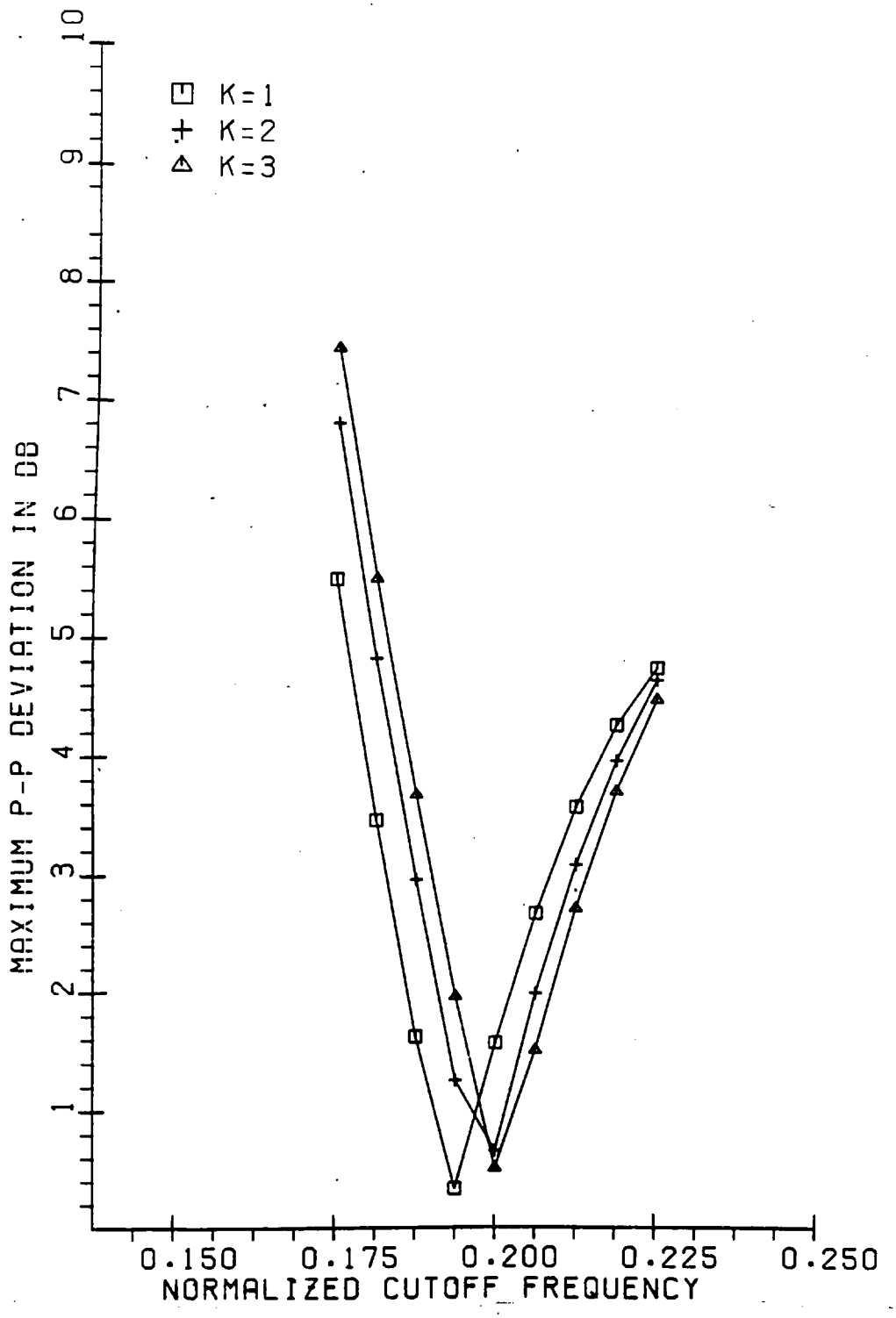


Fig.5.10 Maximum peak-to-peak deviation of the overall response for various values of K. (solution A, N=16, TW=5/32).

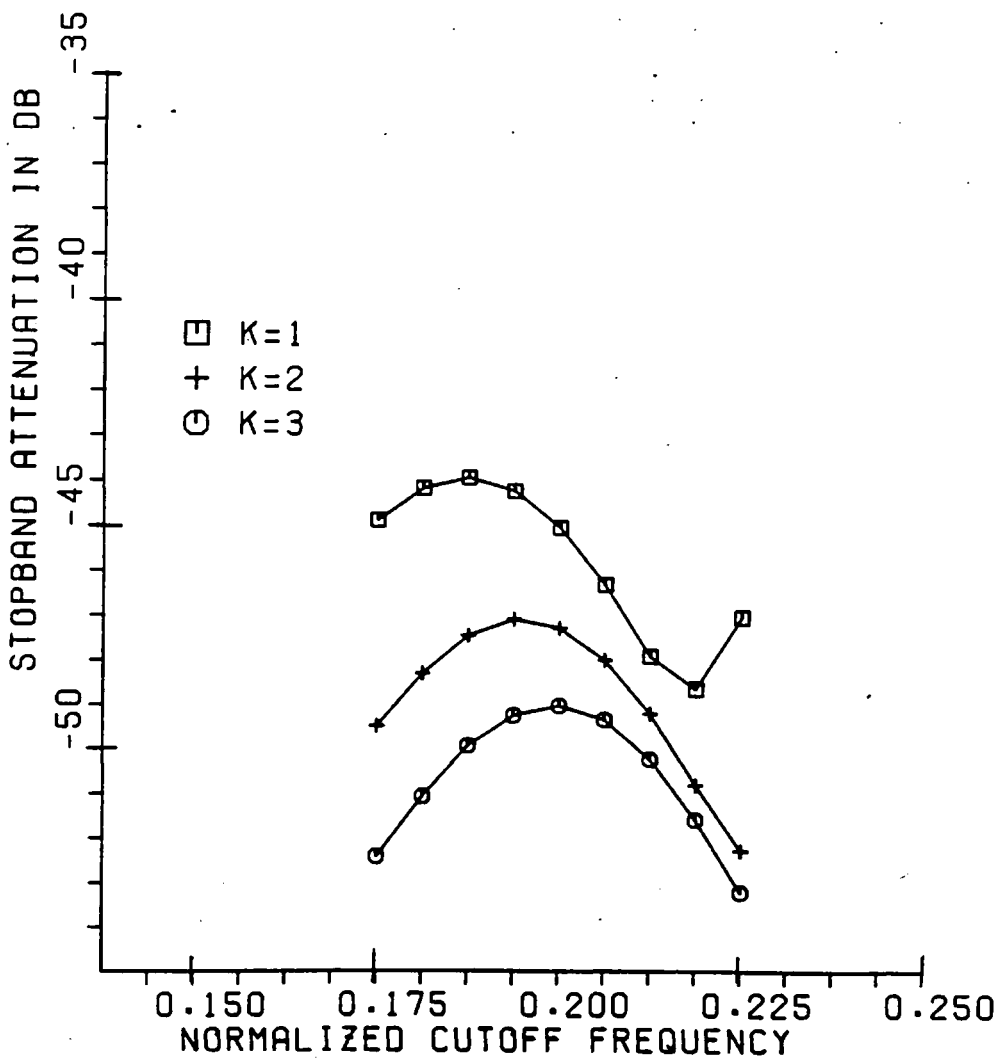


Fig.5.11 Stopband attenuation for various values of K. (solution A, N=16, TW=5/32).

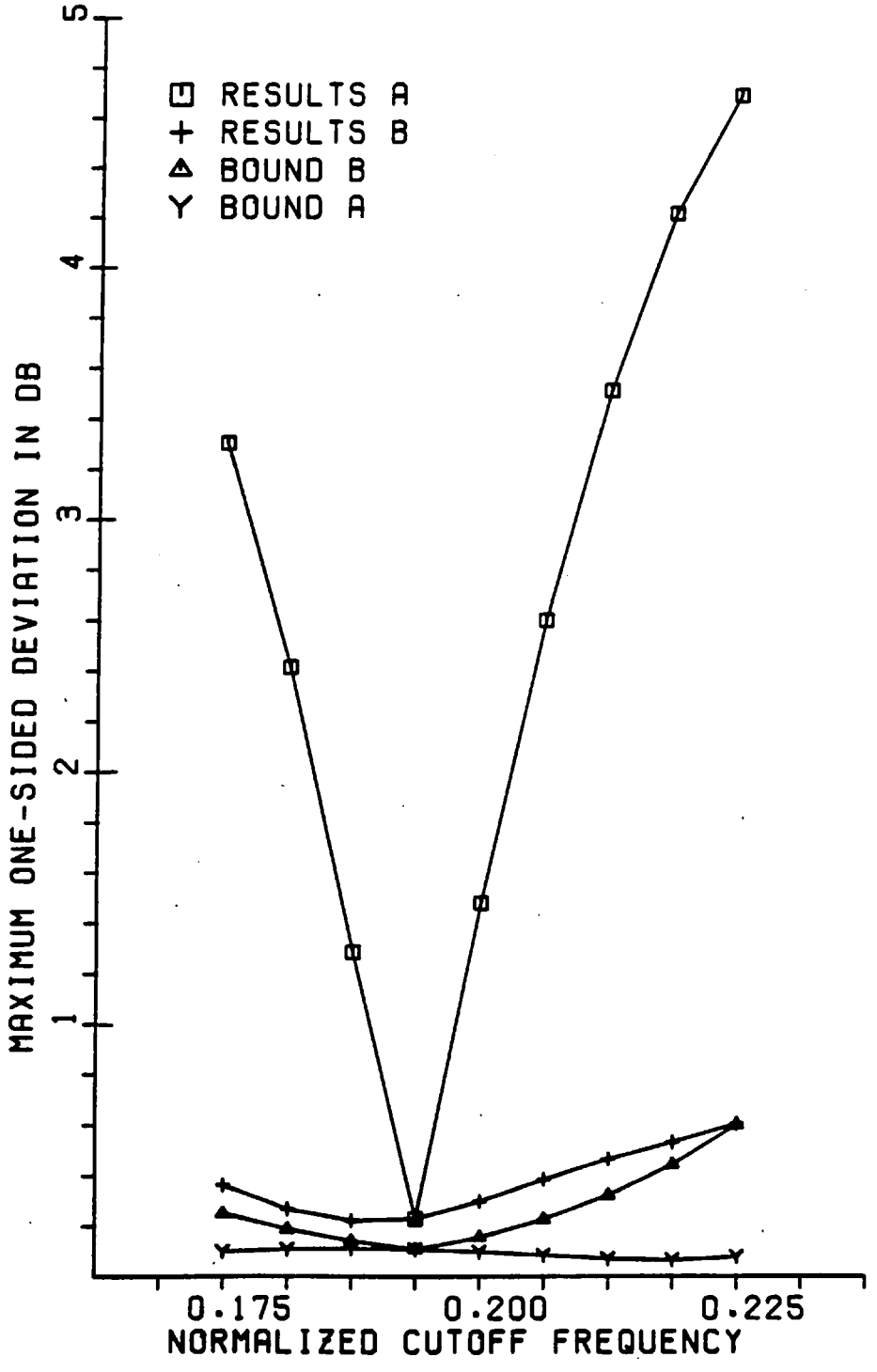


Fig.5.12 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=16, K=1, TW=5/32).

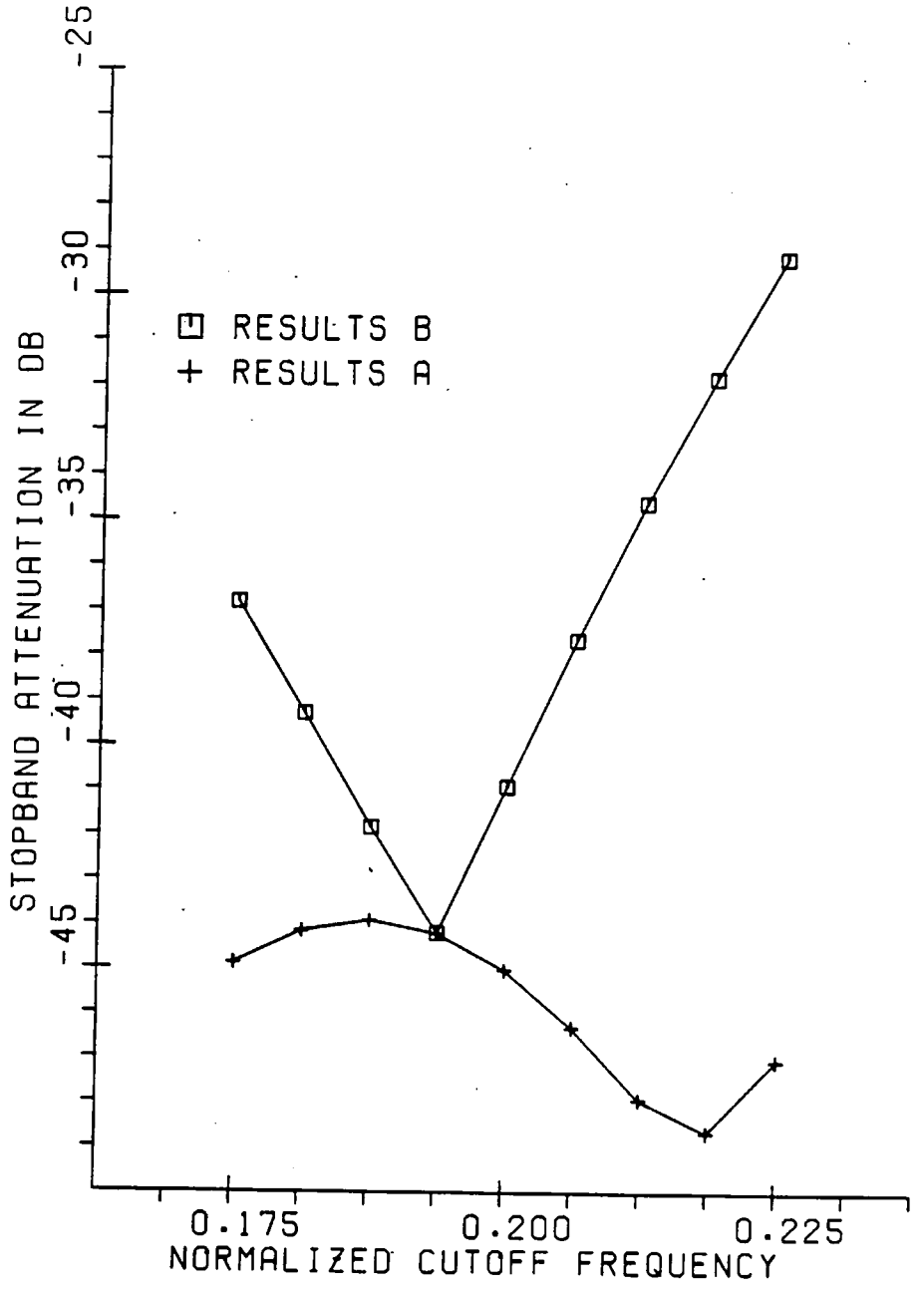


Fig.5.13 Comparison of stopband attenuation for solution A and solution B. (N=16, K=1, TW=5/32).

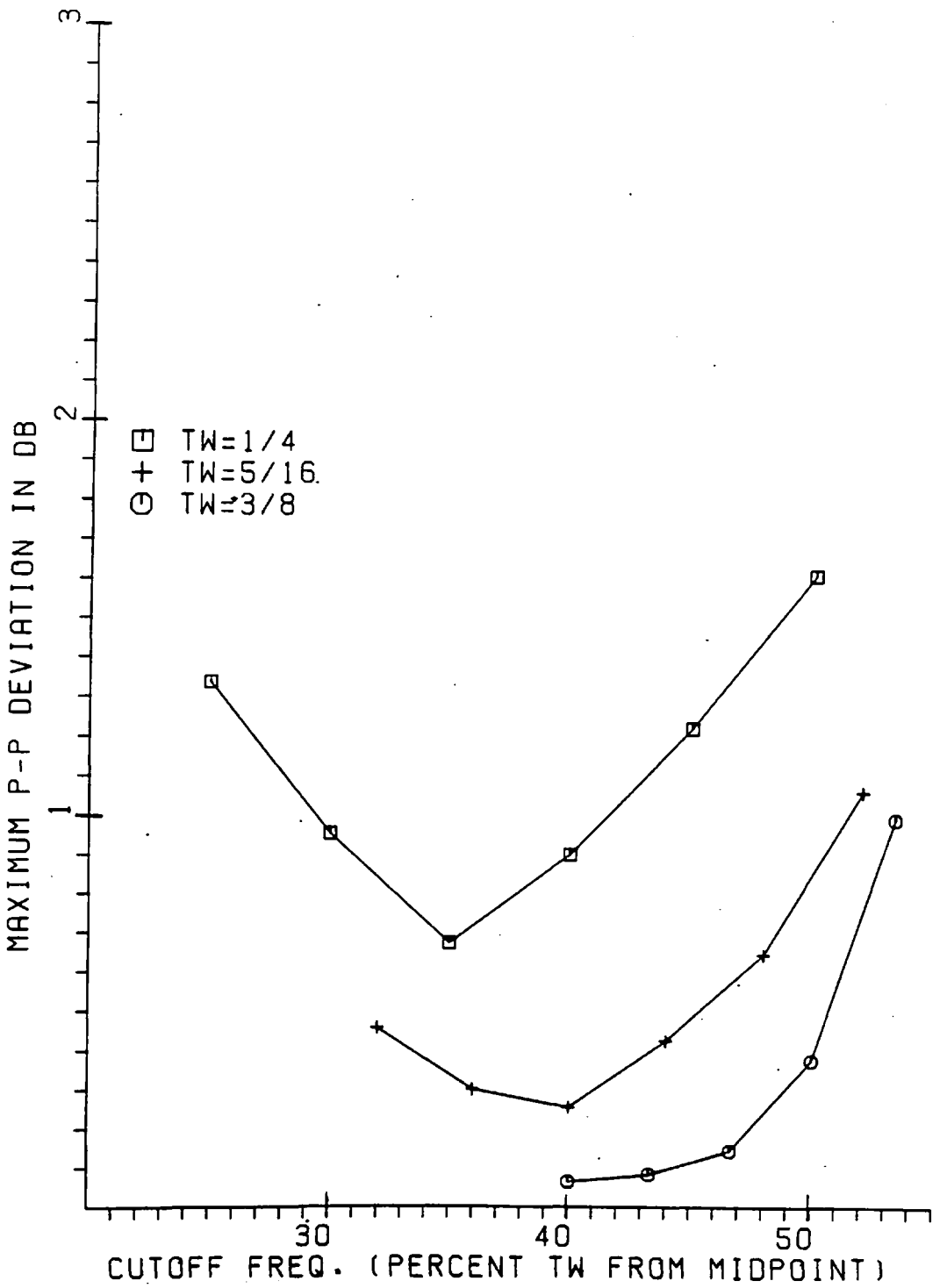


Fig.5.14 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution B, N=8, K=1).

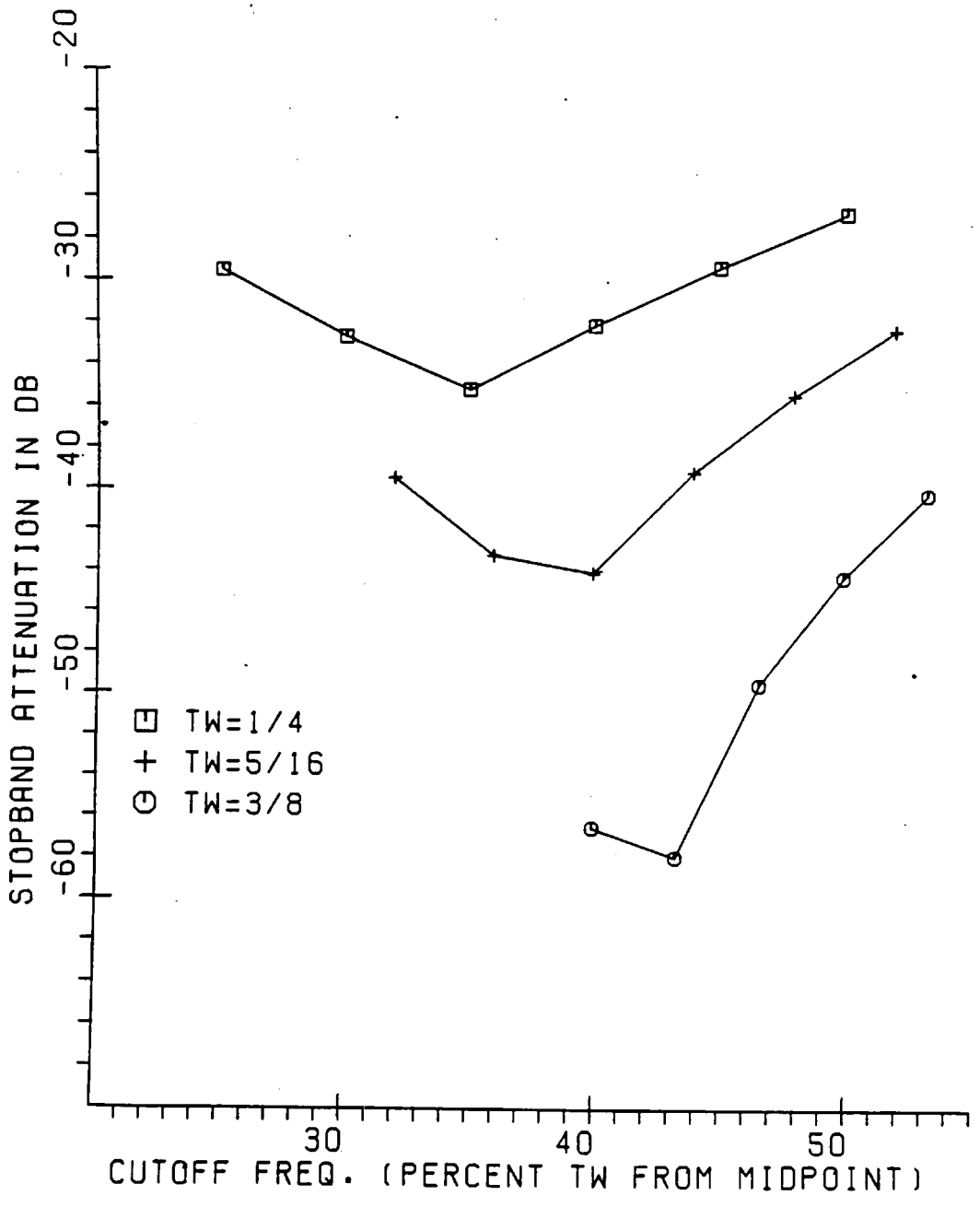


Fig.5.15 Stopband attenuation for various transition widths. (solution B, N=8, K=1).

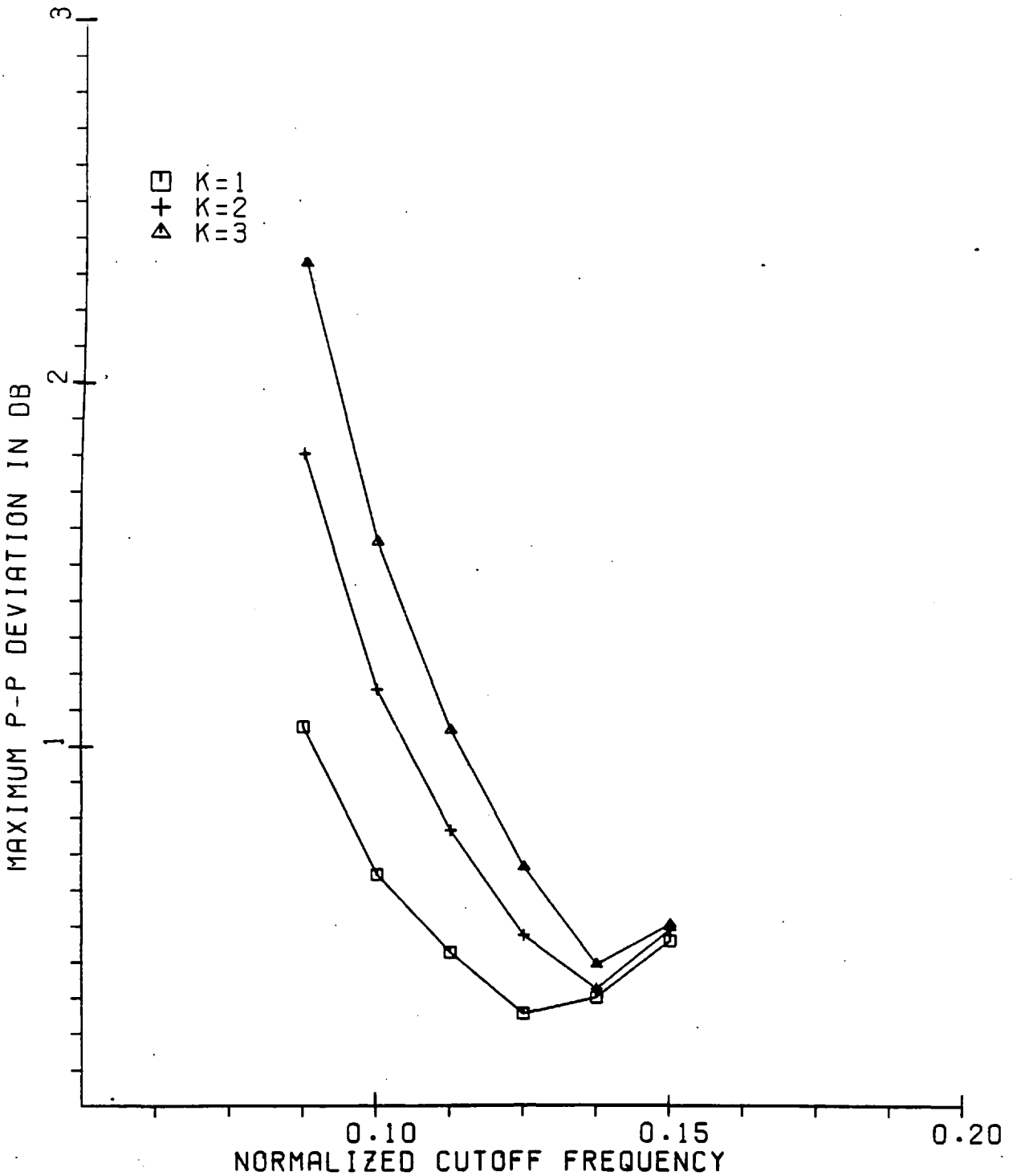


Fig.5.16 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N=8, TW=5/16).

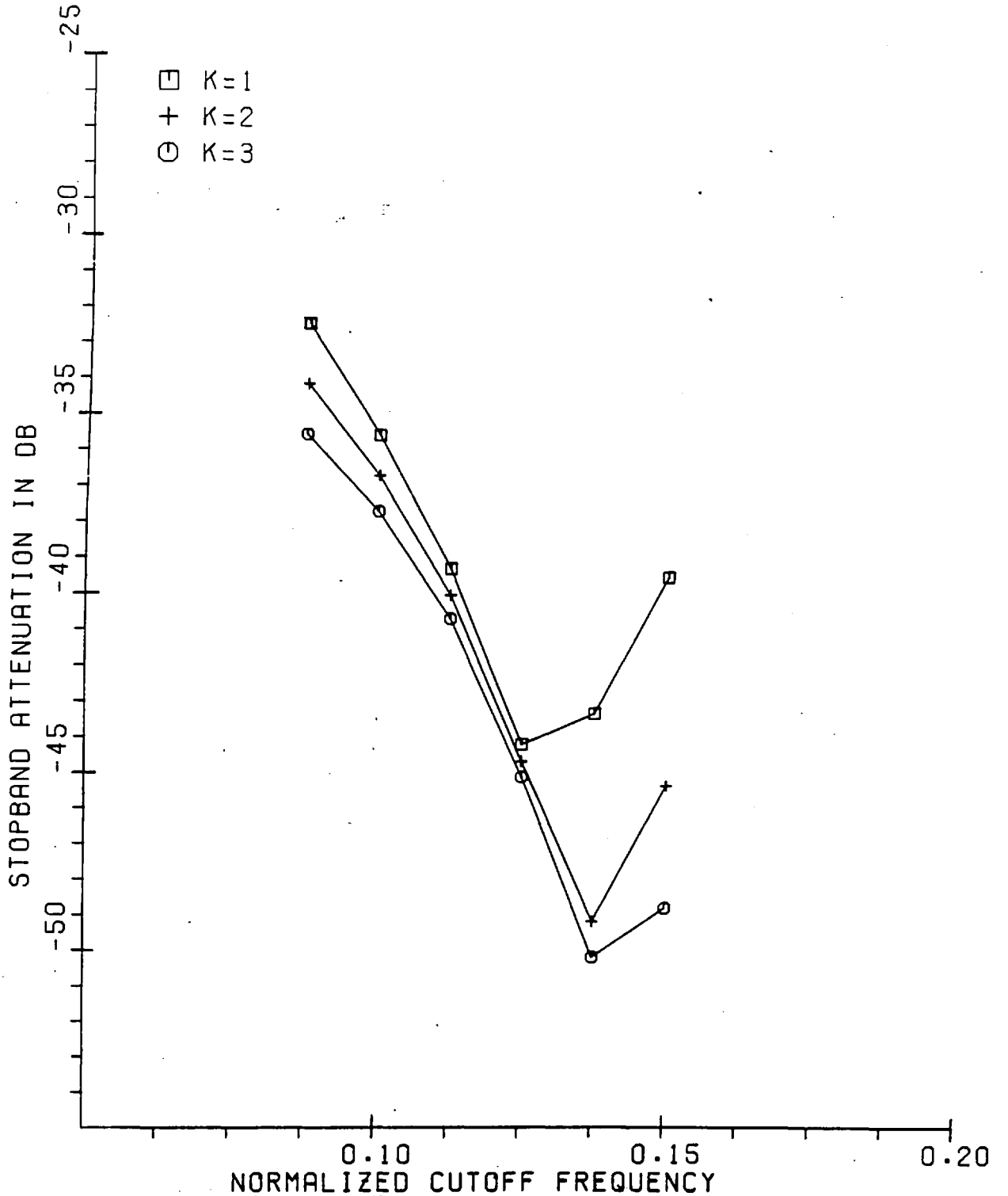


Fig.5.17 Stopband attenuation for various values of K. (solution B, N=8, TW=5/16).

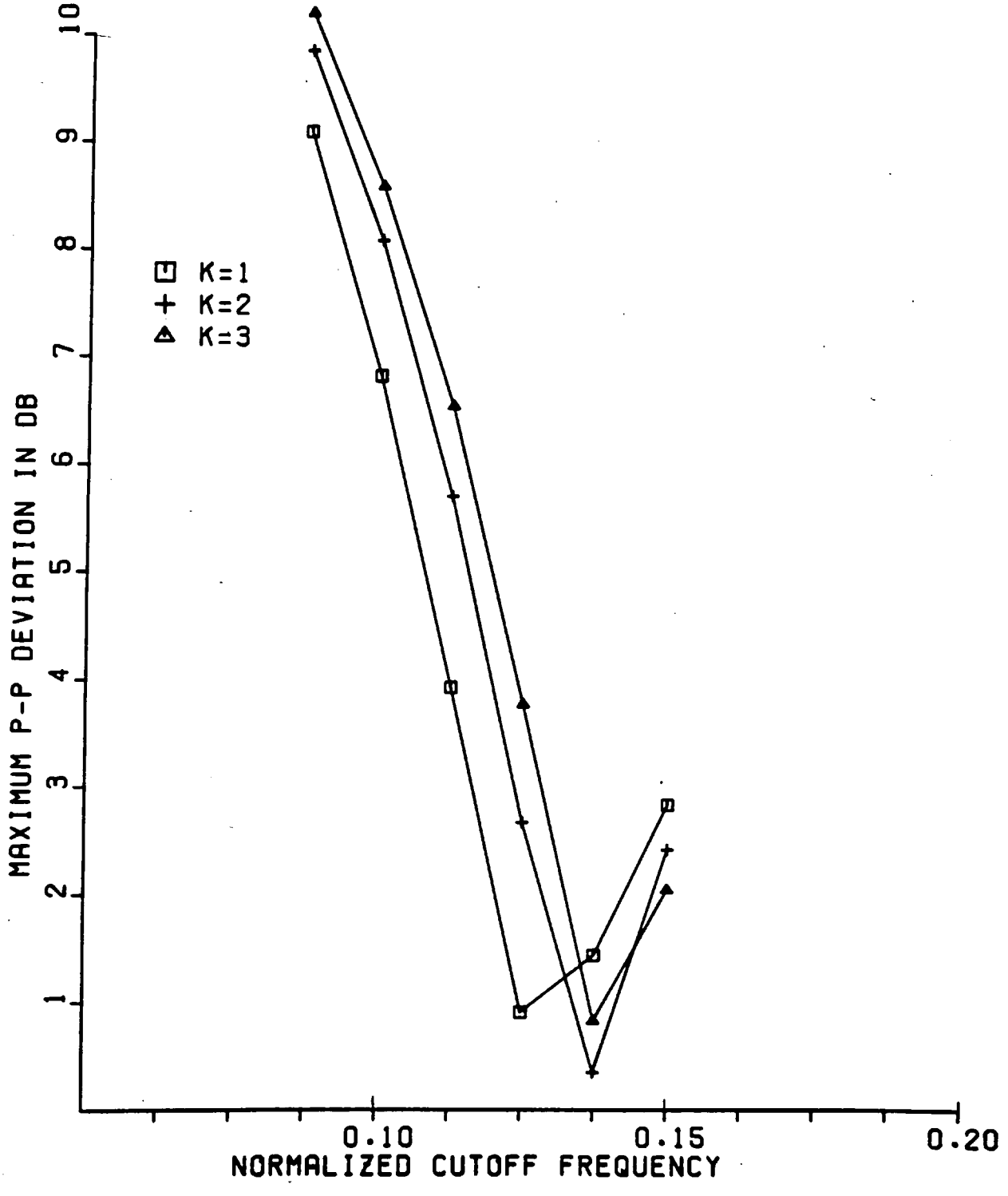


Fig.5.18 Maximum peak-to-peak deviation of the overall response for various values of K. (solution A, N=8, TW=5/16).

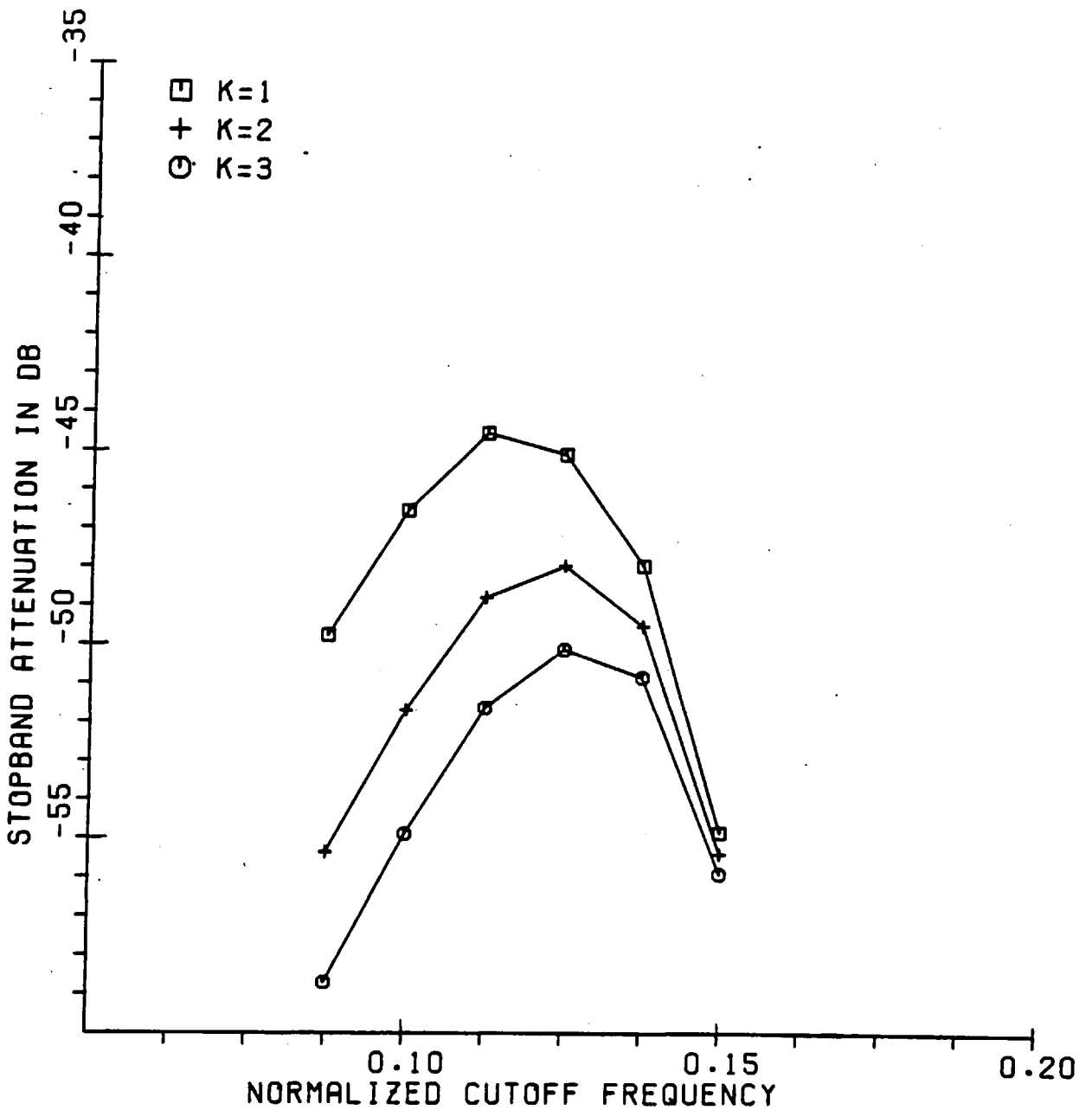


Fig.5.19 Stopband attenuation for various values of K. (solution A, N=8, TW=5/16).

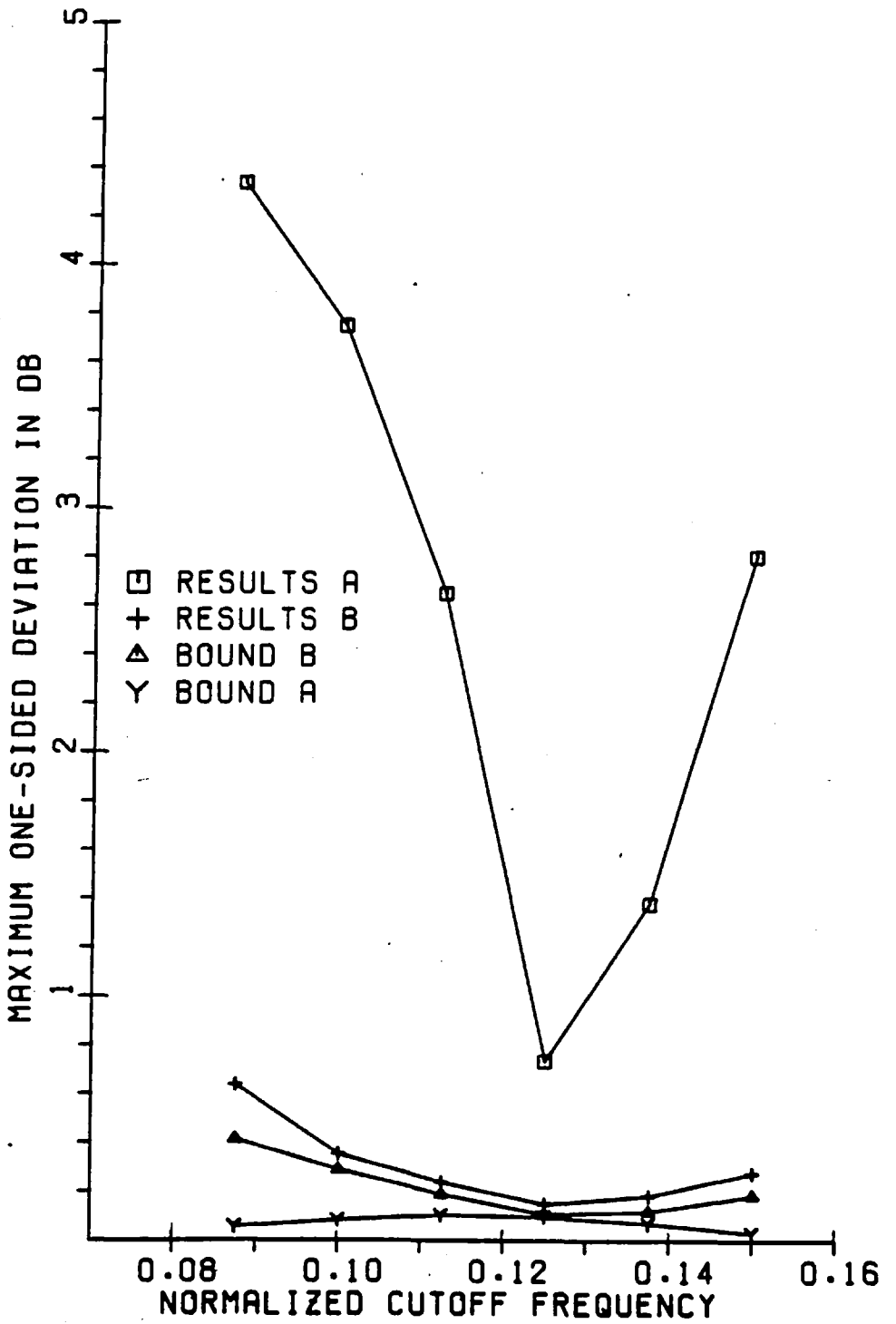


Fig.5.20 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=8, K=1, TW=5/16).

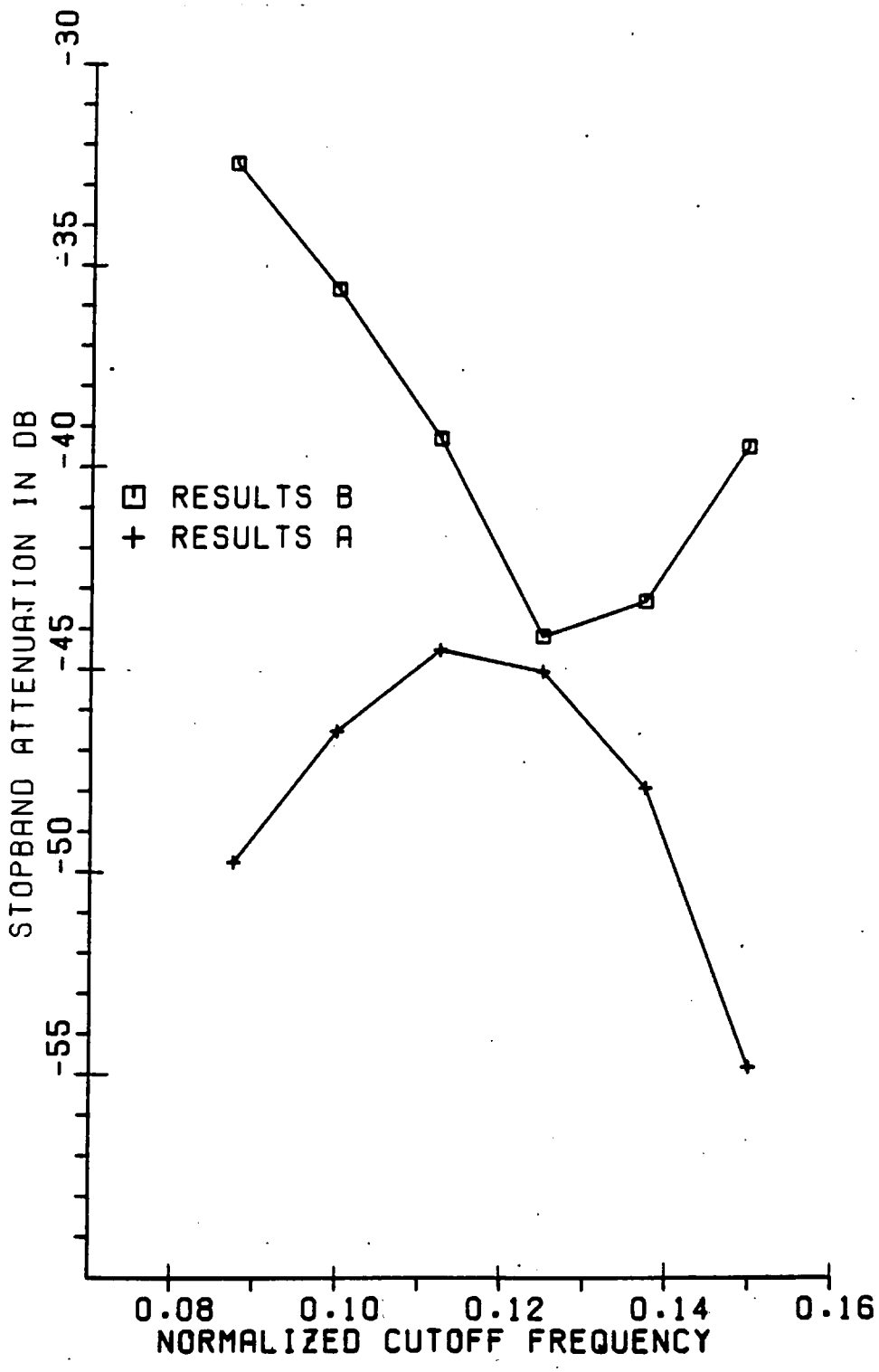


Fig.5.21 Comparison of stopband attenuation for solution A and solution B. (N=8, K=1, TW=5/16).

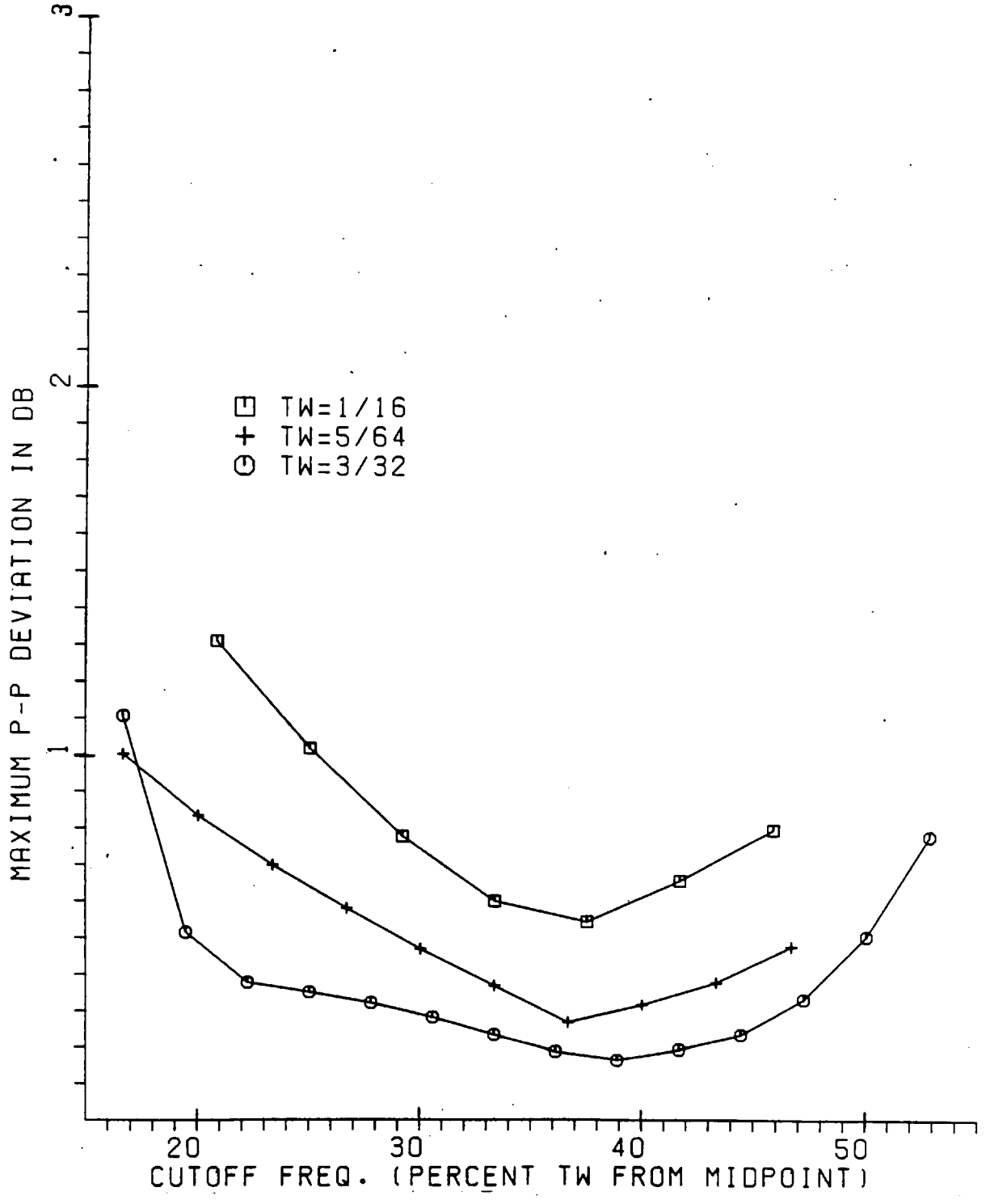


Fig.5.22 Maximum peak-to-peak deviation of the overall response for various transition widths. (solution B, N=32, K=1).

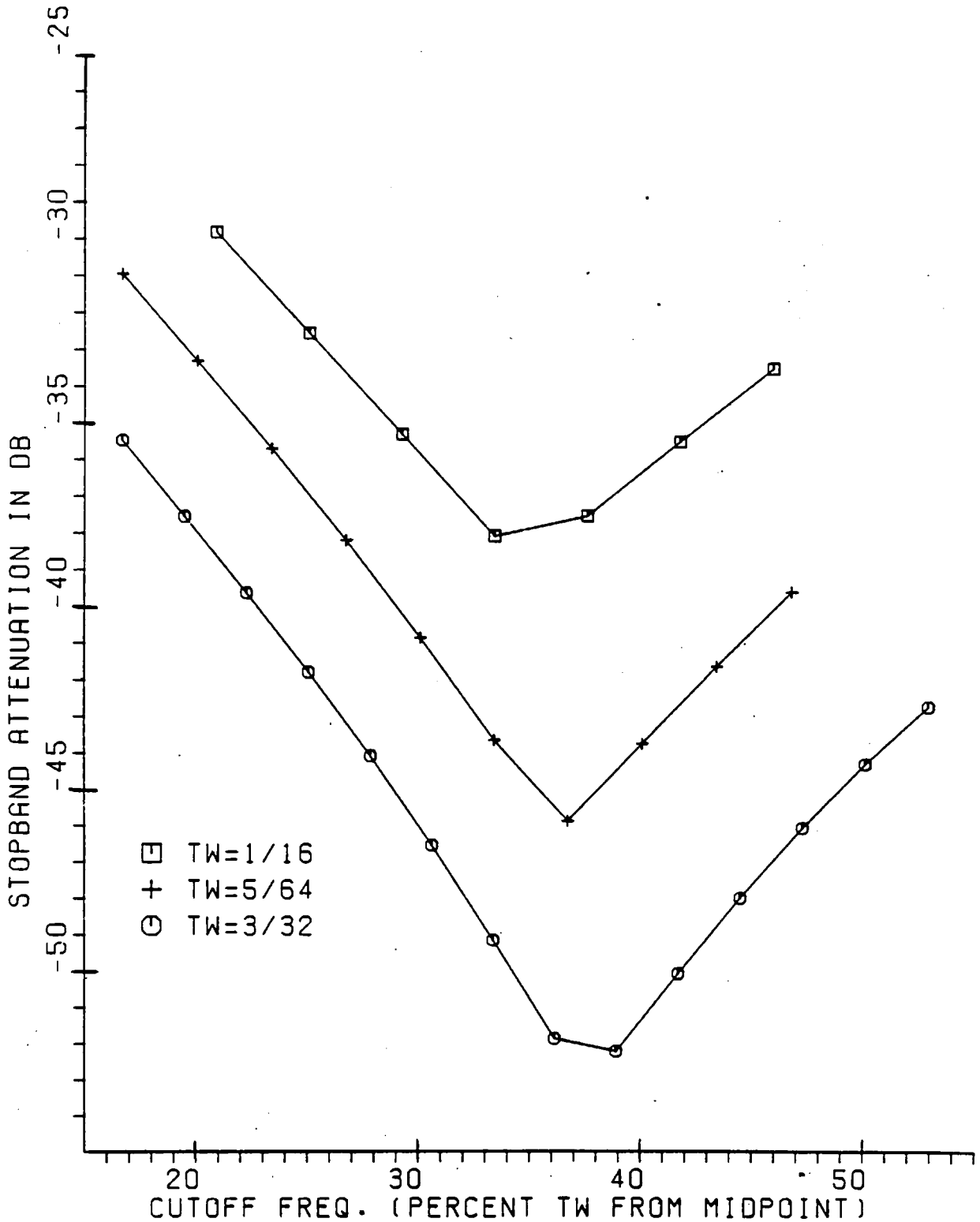


Fig.5.23 Stopband attenuation for various transition widths. (solution B, N=32, K=1).

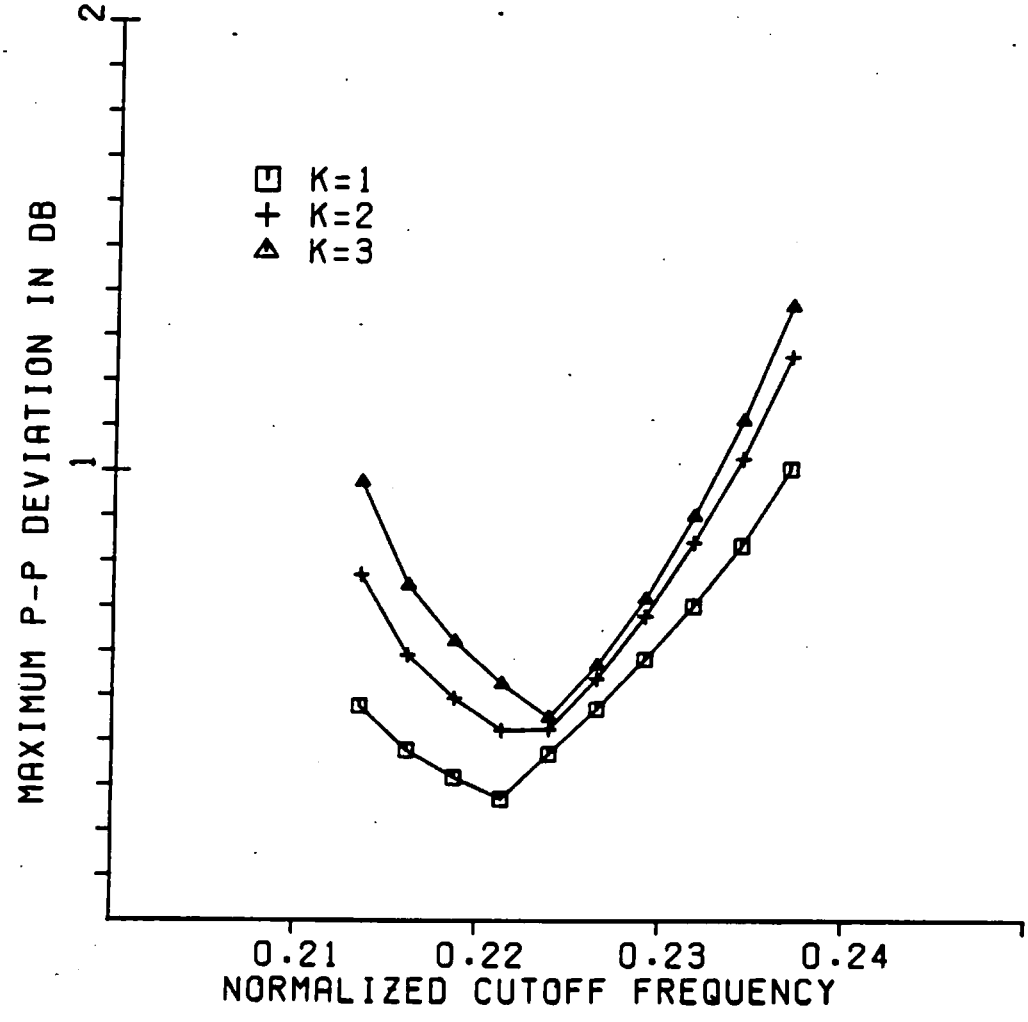


Fig.5.24 Maximum peak-to-peak deviation of the overall response for various values of K. (solution B, N=32, TW=5/64).

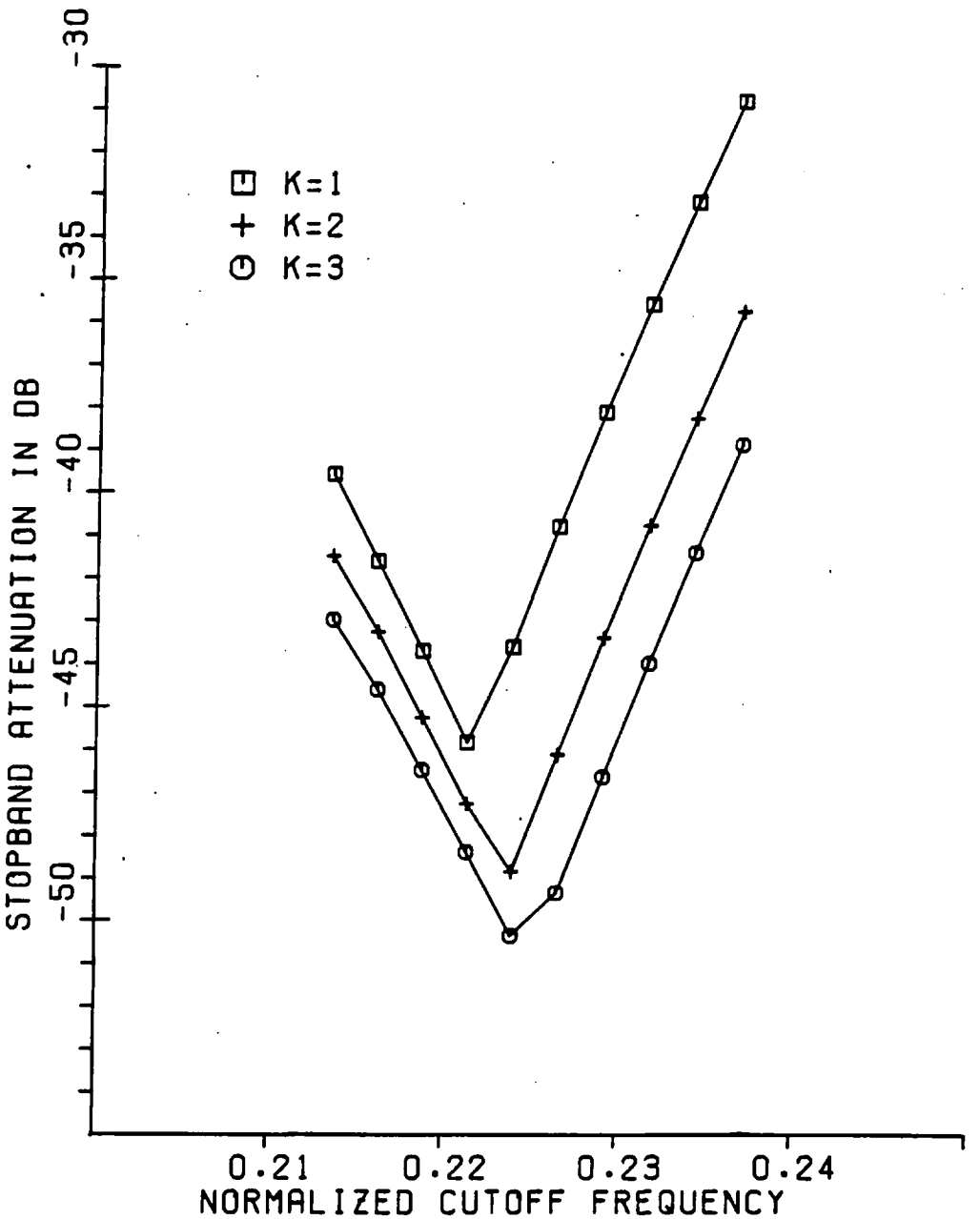


Fig.5.25 Stopband attenuation for various values of K. (solution B, N=32, TW=5/64).

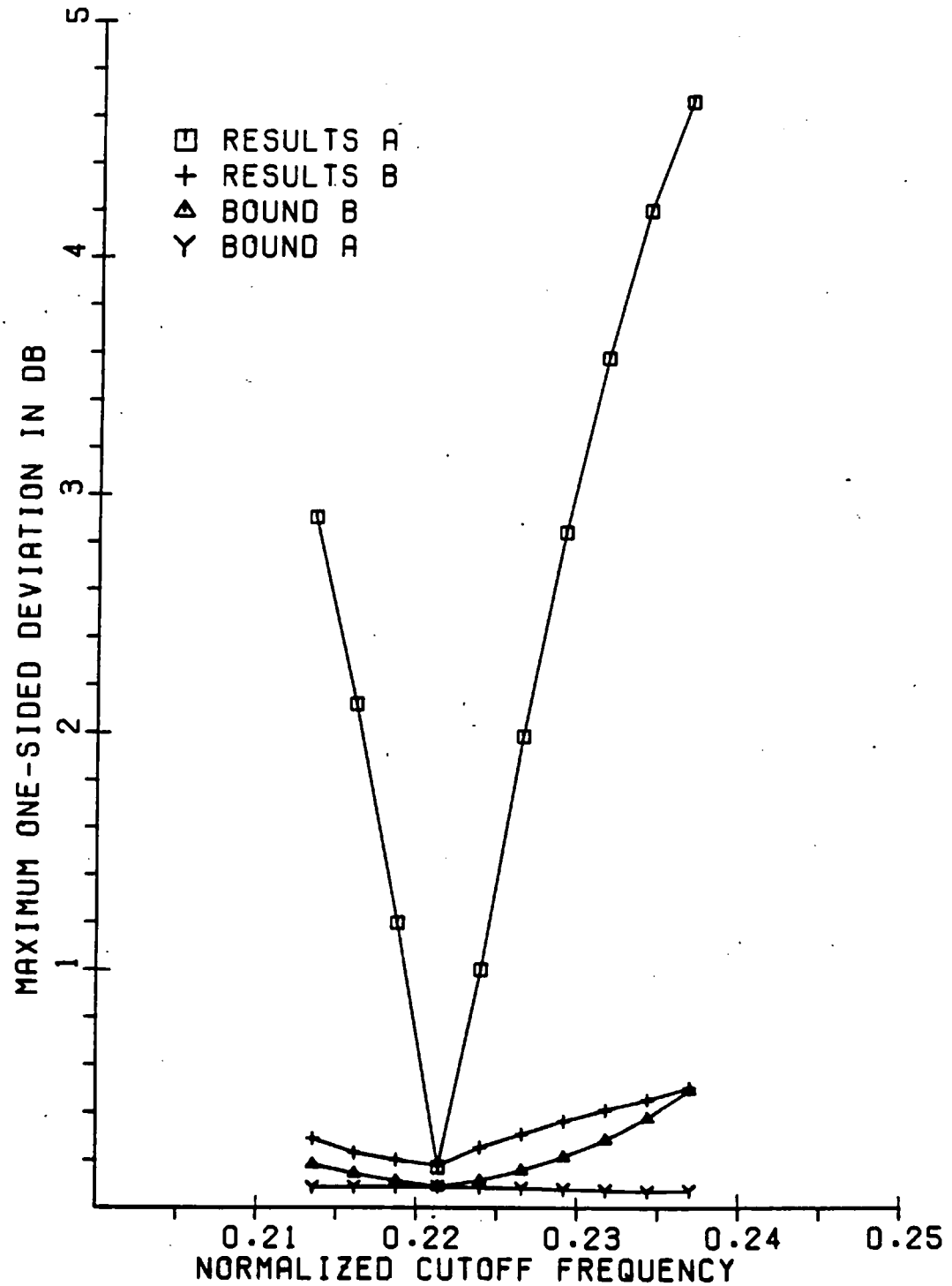


Fig.5.26 Comparison of maximum one-sided deviation of the overall response for solution A and solution B. (N=32, K=1, TW=5/64).

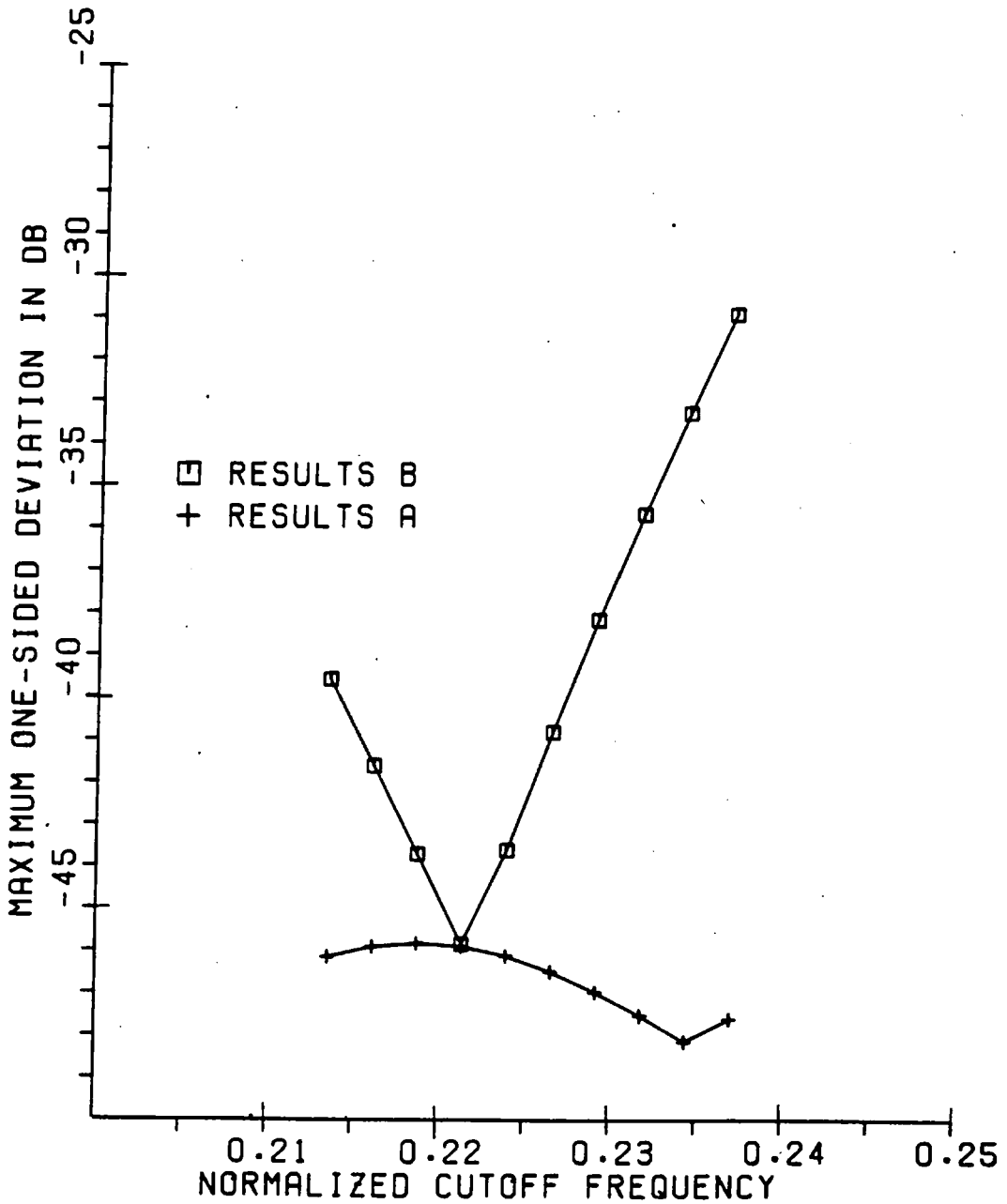


Fig.5.27 Comparison of stopband attenuation for solution A and solution B. (N=32, K=1, TW=5/64).

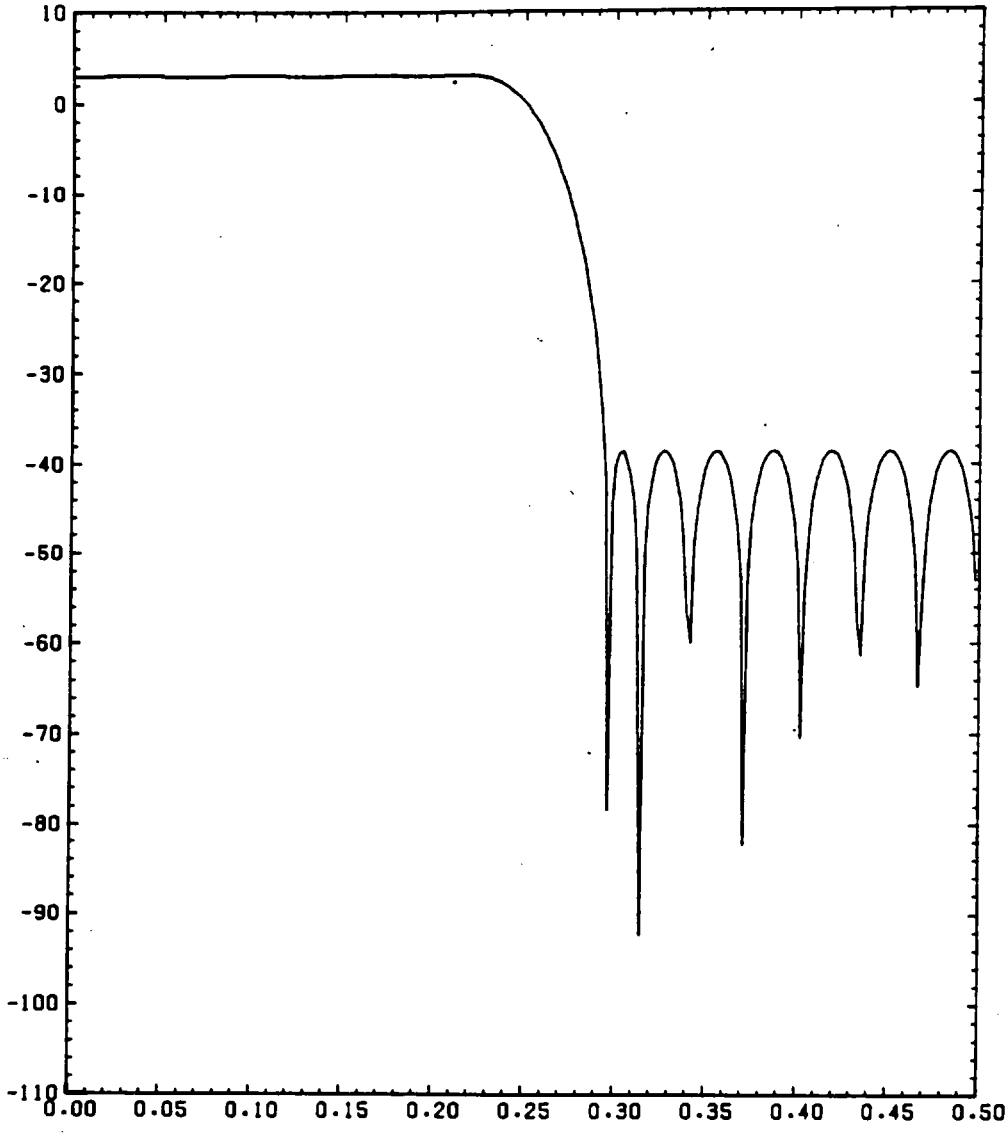


Fig.5.28 Frequency response of a 32-point FIR lowpass filter.

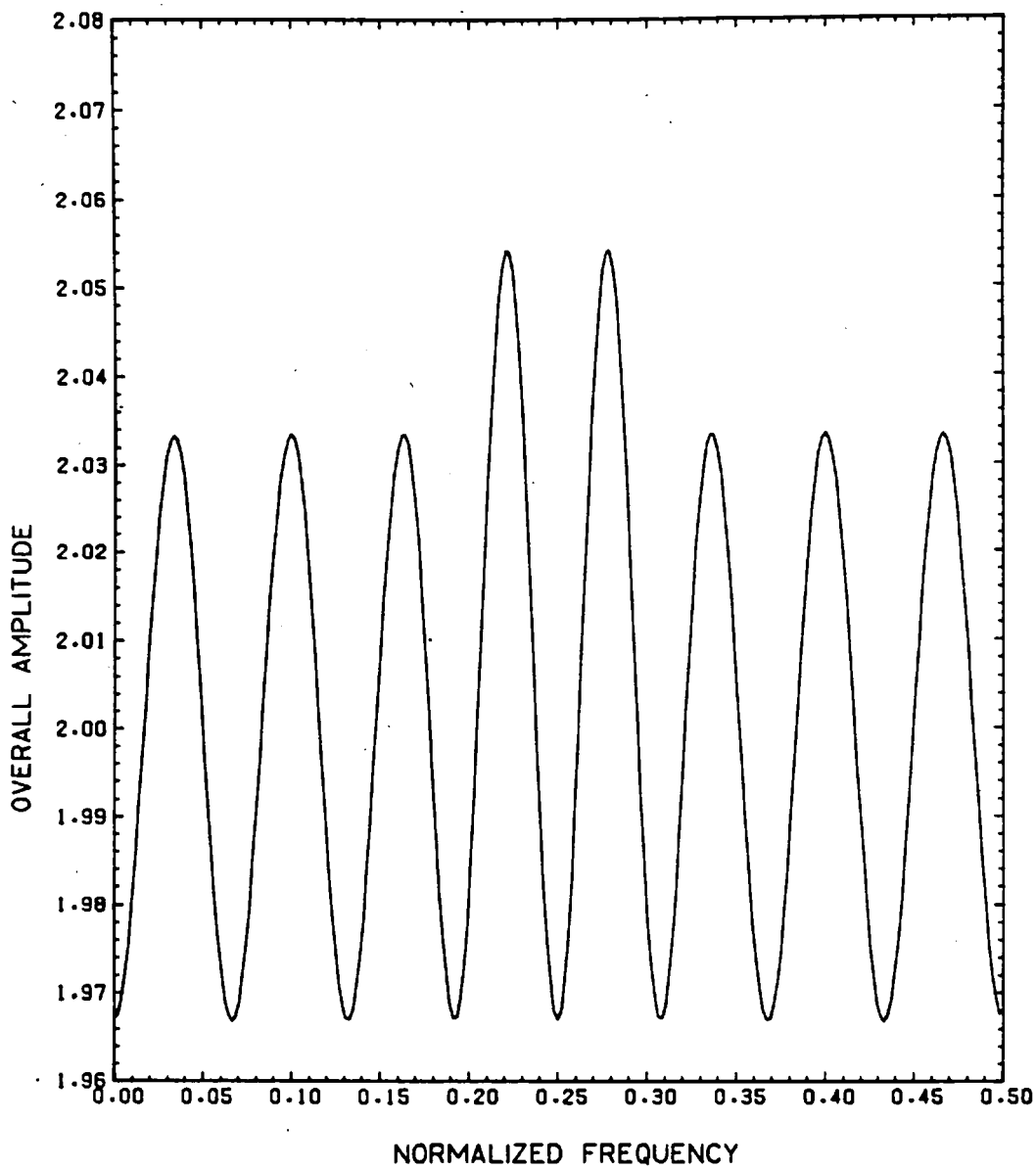


Fig.5.29 Frequency response of the overall system based on filters with lowpass frequency response shown in Fig.5.28.

CHAPTER 6

EFFECTS OF FINITE-WORD-LENGTH ON THE PERFORMANCE OF SPLIT-BAND SYSTEMS USING QUADRATURE MIRROR FILTERS

CHAPTER 6

EFFECTS OF FINITE-WORD-LENGTH ON THE PERFORMANCE OF SPLIT-BAND SYSTEMS USING QUADRATURE MIRROR FILTERS

6.1 INTRODUCTION

In the previous chapter, a method was put forward for the design of quadrature mirror filters so that the peak-to-peak ripple of the frequency response of a split-band system using quadrature mirror filters could be minimized. In this chapter, the errors resulting from the use of finite-word-length registers is considered and the results of an experimental verification of the theory derived are presented. The analysis carried out in this chapter provides some useful guidelines for the implementation of sub-band coding systems using quadrature mirror filters; and in particular, provides information as to the word-length that should be used in storage and arithmetic operations.

A discussion of the number representation and the noise model used in the subsequent analysis is given in Section 6.2. In Section 6.3, the structure and the coefficient scaling of quadrature mirror filters are discussed. The effects of quantization at various points of the system are presented in Section 6.4 and some experimental results are given in Section 6.5.

6.2 NUMBER REPRESENTATION AND NOISE MODEL USED

A number x with $|x| < 1$ can be represented by a sequence of binary digits $b_0 b_1 b_2 \ldots$ with x being related to the binary digits according to the expression:

$$x = -b_0 + \sum_{k=1}^{\infty} b_k 2^{-k}$$

in which $b_k = 1$ or 0 for $k \geq 0$

If it is assumed that only L bits are available for the representation of x, then

$$\hat{x} = -\hat{b}_0 + \sum_{k=1}^{L-1} \hat{b}_k 2^{-k}$$

where $\hat{b}_k = 1$ or 0 for $0 \leq k \leq (L-1)$.

In this case, the resulting truncation error is $(x-\hat{x}) = \sum_{k=L}^{\infty} b_k 2^{-k}$ provided that $\hat{b}_i = b_i$ for $i=0,1,\ldots,(L-1)$. This error will in general be larger than that resulting from rounding in which a 1 or a 0 is added to the L^{th} bit depending on whether the $(L+1)^{th}$ bit is 1 or 0 respectively. When x is rounded to L bits, $|x-\hat{x}| < 2^{-L}$.

In the analysis which follows, it is assumed that all quantized values are represented by fixed-point binary numbers of L bits, that negative numbers are represented in 2's complement form, and that all numbers are left justified with a magnitude less than unity.

The representation of negative numbers in 2's complement form has the advantages that (i) subtraction can be obtained through the operation of addition; (ii) since the representation is circular, the results of addition and/or subtraction of numbers is exact, irrespective of the number and the order of operations, so long as the

final result is within the range represented by the given number of digits, even if overflows occur in the sequence of operations [115].

The representation of numbers as fractions has the benefit that since the product of the two fractions is less than unity, overflow does not occur as a result of multiplication.

The effects on the accuracy of systems using a finite-word-length number representation can be studied if it is assumed that the error introduced by rounding to L bits is a random noise having a uniform probability distribution in the interval (-q/2,q/2), in which q ($=2^{-(L-1)}$) is the quantization step size. In order to simplify the following analysis, it will be assumed that :

(i) any two different samples from the same noise source are uncorrelated ;

(ii) any two samples from different noise sources are uncorrelated;

and (iii) each noise source is uncorrelated with the input.

This means that the noise source is modelled as a discrete stationary white noise process with uniform power density spectrum of magnitude $\eta^2 = q^2/12$.

In the case of split-band systems using quadrature mirror filters as shown in Fig.5.1, the following four categories of noise source can be identified :

(i) the quantization of samples of the original analogue input, {x(t)}, into a set of discrete levels;

(ii) the representation of the coefficients of the filters by a finite number of binary digits;

(iii) the rounding of values resulting from multiplications of numbers;

and (iv) the compression coding of the split-band signals $\{y_l(n)\}$ and $\{y_u(n)\}$.

6.3 STRUCTURE AND COEFFICIENT SCALING OF THE FILTERS

In order that all the numbers and data are less than unity, it is necessary that scaling of the input data and all filter coefficients be carried out. If it is assumed that the values of the sequence $\{x(n)\}$ have magnitude less than unity, then in order to ensure that the magnitude of the filtered sequence $\{x_l(n)\}$ is less than unity, 'sum scaling' [116] has to be performed on the coefficients $\{h'_l(k)\}$. That is, a new set of coefficients $\{h_l(k)\}$ is obtained in which

$$h_l(k) = h'_l(k) / \sum_{k=0}^{N-1} |h'_l(k)| \quad \text{for} \quad k = 0,1,\ldots,N-1 \tag{6.1}$$

Thus $$\sum_{k=0}^{N-1} |h_l(k)| = \sum_{k=0}^{N-1} |h'_l(k)| / \sum_{k=0}^{N-1} |h'_l(k)| = 1 \tag{6.2}$$

Since $h_l(k) = h_l(N-1-k)$ for $k = 0,1,\ldots,(N-1)$

it follows that $$\sum_{k=0}^{N/2-1} 2|h_l(k)| = 1 \tag{6.3}$$

and that $$\sum_{k=0}^{N/2-1} |h_l(k)| = 1/2 \qquad \text{for} \qquad N \geqslant 2 \tag{6.4}$$

This implies that $|h_l(k)| \leqslant 1/2$ for $k = 0,1,\ldots,N-1$ (6.5)

so that $$h_l^2(k) \leqslant (1/2)|h_l(k)| \quad \text{for} \quad k = 0,1,\ldots,N-1 \tag{6.6}$$

$$\sum_{k=0}^{N-1} h_l^2(k) \leqslant \sum_{k=0}^{N-1} (1/2)|h_l(k)| = 1/2 \quad \text{for} \quad N \geqslant 2 \tag{6.7}$$

and $$\sum_{k=0}^{N/2-1} h_l^2(k) \leqslant (1/2)(1/2) = 1/4 \tag{6.8}$$

If, in implementing the filters, a direct-form shared-multiplier structure as shown in Fig.6.1 is used, then the filtered output is given by

$$x_l(n) = \sum_{k=0}^{N/2-1} h_l(k)[x(n-k)+x(n-N+1+k)] \tag{6.9}$$

and hence
$$|x_l(n)| \leq \sum_{k=0}^{N/2-1} |h_l(k)|[|x(n-k)|+|x(n-N+1+k)|]$$
$$< \sum_{k=0}^{N/2-1} |h_l(k)|(2) = 1 \quad \text{as} \quad |x(n)| < 1 \tag{6.10}$$

This means that $|x_l(n)| < 1$ for all n and the filtered sequence $\{x_l(n)\}$ thus has values less than unity.

It should be noted that although the partial sum $[x(n-k)+x(n-N+1+k)]$ may exceed unity, the product $h_l(k)[x(n-k)+x(n-N+1+k)]$ is always less than unity and the shared-multiplier structure can thus be adopted when all the numbers are represented in 2's complement form. It is important to note that the number of multipliers required with a structure of this kind is only half the number of stages in the filter.

The same arguments also apply to the filter H_u.

At the receiver, the interpolated sequence $\{u_l(n)\}$, has alternate zero values and if a shared-multiplier structure is used for the filter, then the filtered output is given by

$$t_l(n) = \sum_{k=0}^{N/2-1} p_l(k)[u_l(n-k)+u_l(n-N+1+k)] \tag{6.11}$$

Since one of the terms of $[u_l(n-k) + u_l(n-N+1+k)]$ is zero,

$$|u_l(n-k)|+|u_l(n-N+1+k)| < 1 \tag{6.12}$$

and thus $$|t_l(n)| < \sum_{k=0}^{N/2-1} |p_l(k)| \tag{6.13}$$

This means that $|t_l(n)| < 1$ if $\sum_{k=0}^{N/2-1} |p_l(k)| = 1$ which can be achieved by dividing the original set of coefficients $\{p_l'(k)\}$ by $\sum_{k=0}^{N/2-1} |p_l'(k)|$, or equivalently, by (1/2) $\sum_{k=0}^{N-1} |p_l'(k)|$.

When the coefficients are so scaled,

$$|p_l(k)| \leqslant 1 \qquad \text{for} \qquad k = 0,1,\ldots,N-1 \tag{6.14}$$

$$\sum_{k=0}^{N-1} |p_l(k)| = 2 \tag{6.15}$$

and $$\sum_{k=0}^{N-1} p_l^2(k) \leqslant 2 \qquad \text{for} \qquad N \geqslant 2 \tag{6.16}$$

Similar arguments can also be applied to the filter P_u. When scaling is carried out as proposed, the output sequence has to be multiplied by $[\sum_{k=0}^{N-1} |h_l'(k)|][(1/2) \sum_{k=0}^{N-1} |p_l'(k)|]$ so that its values are of the same order of magnitude as those of the input sequence.

6.4 EFFECTS OF QUANTIZATION AND FINITE-WORD-LENGTH IN SYSTEMS USING QUADRATURE MIRROR FILTERS

6.4.1 Effects of Quantization of the Input Signal

Let $X(z)$ represent the z-transform of the sequence $\{x(n)\}$ where $X(z) = \sum_{k=0}^{\infty} x(k)z^{-k}$ and let $\{\hat{x}(n)\}$ be the quantized input sequence and $\hat{X}(z)$ its z-transform. The quantization error is then given by

$e(k) = \hat{x}(k) - x(k)$ for all $k \geq 0$. Let $E(z)$ be the z-transform of the sequence $\{e(k)\}$ and let the random variable consisting of the sequence $\{e(k)\}$ be represented by E. Also, denote the mean and variance of E by $\bar{E}$ and var(E) respectively. The analog-to-digital conversion process introduces quantization noise. Let this quantization noise be modelled as a white noise source $\{e_x(n)\}$ whose samples are uniformly distributed in the range $-q/2$ to $q/2$, where q is the quantization step size. This quantization noise is added to the input $\{x(n)\}$ to give the quantized input sequence $\{\hat{x}(n)\}$. Thus

$$\hat{X}(z) = X(z) + E_x(z) \tag{6.17}$$

From the theory of quadrature mirror filters, the z-transform of the noisy output sequence $\{s_x(n)\}$ is seen to be

$$\begin{aligned} S_x(z) &= R(z)\hat{X}(z) \\ &= R(z)X(z) + R(z)E_x(z) \\ &= S(z) + \mathcal{E}_s(z) \end{aligned} \tag{6.18}$$

where $$\mathcal{E}_s(z) = R(z)E_x(z) \tag{6.19}$$

is the z-transform of the noise sequence imposed on the output.

If $R(z) = z^{-(N-1)}$, that is, if the frequency response of the system is flat, then

$$S_x(z) = \hat{X}(z)z^{-(N-1)} \tag{6.20}$$

and the output is the exact replica of the quantized input except for the delay and the signal-to-noise ratio at the output is exactly the same as that at the input.

6.4.2 Effects of Quantization of Filter Coefficients

Deviation of the filter coefficients from their 'infinitely-accurate' values produces a distortion of the overall system response and this poses the question as to the level of distortion resulting from the use of finite-word-length coefficients. The analysis below provides the designer with a method of estimating the expected deviation for a given accuracy of coefficient representation.

The effects of coefficient quantization on (i) the resultant frequency response $R(e^{j\omega})$ and (ii) the output sequence $\{s(n)\}$ will now be considered.

Let the 'infinitely accurate' coefficients associated with the filters H_l and P_l be $h_l(k)$ and $p_l(k)$ respectively, and let $\hat{h}_l(k)$ and $\hat{p}_l(k)$, $k=0,1,\ldots,N-1$ denote these coefficients when they have been rounded to L bits. The quantization errors are then seen to be,

$$e_h(k) = \hat{h}_l(k) - h_l(k) \tag{6.21}$$

and

$$e_p(k) = \hat{p}_l(k) - p_l(k) \quad \text{for} \quad k = 0,1,\ldots,N-1 \tag{6.22}$$

where

$$|e_h(k)| < q/2 \tag{6.23}$$

and

$$|e_p(k)| < q/2 \tag{6.24}$$

Let E_h and E_p represent the random variables consisting of $\{e_h\}$ and $\{e_p\}$ respectively. From the assumed noise model, it follows that the variance of both E_h and E_p is equal to η^2, where $\eta^2 = q^2/12$.

If it is assumed that

$$\hat{H}_u(z) = \hat{H}_l(-z) = H_l(-z) + E_h(-z) \tag{6.25}$$

$$\hat{P}_l(z) = P_l(z)+E_p(z) = 2\hat{H}_l(z) \tag{6.26}$$

and $$\hat{P}_u(z) = -2\hat{H}_u(z) = -2\hat{H}_l(-z) = -\hat{P}_l(-z) \tag{6.27}$$

where $E_h(z)$ and $E_p(z)$ are the z-transforms of the the respective sequences $\{e_h\}$ and $\{e_p\}$, then it can be shown that the z-transform of the noisy output sequence is given by

$$\begin{aligned} S_c(z) &= (1/2)\{\hat{H}_l(z)\hat{P}_l(z)-\hat{H}_l(-z)\hat{P}_l(-z)\}X(z) \\ &= (1/2)\{[H_l(z)+E_h(z)][P_l(z)+E_p(z)] \\ &\quad -[H_l(-z)+E_h(-z)][P_l(-z)+E_p(-z)]\}X(z) \\ &= R(z)X(z) + \mathcal{E}_{cr}(z)X(z) \end{aligned} \tag{6.28}$$

where $$\mathcal{E}_{cr}(z) = (1/2)\{H_l(z)E_p(z)+P_l(z)E_h(z)+E_h(z)E_p(z) - H_l(-z)E_p(-z)-P_l(-z)E_h(-z)-E_h(-z)E_p(-z)\} \tag{6.29}$$

The z-transform of the time-domain output error sequence, denoted by $\mathcal{E}_{cs}(z)$, is then given by the following relation:

$$\mathcal{E}_{cs}(z) = S_c(z)-S(z) = \mathcal{E}_{cr}(z)X(z) \tag{6.30}$$

It should be noted that there is no term representing the z-transform of a frequency folded sequence and hence cancellation of frequency aliasing is retained even when finite-word-length coefficients are used.

6.4.3 Error Bounds and Statistics Associated with Finite-Word-Length Coefficients

The analysis of the error $\mathcal{E}_{cr}$ can best be carried out by considering its various component terms separately and this has been done in Appendix E. If it is assumed that

$$\mathcal{E}_{cr}(z) = \sum_{k=0}^{2N-2} \mathcal{E}_{cr}(k) z^{-k} \tag{6.31}$$

then using the results derived in Appendix E, it can be seen that the absolute bound on the sequence $\{\epsilon_{cr}(k)\}$ is given by

$$\begin{aligned}
|\epsilon_{cr}(k)| &\leqslant (1/2)\{\sum_{i=1}^{6} |\epsilon_{ci}(k)|\} \text{ for } 0 \leqslant k \leqslant 2N-2 \\
&< (1/2)(2)\{(q/4)+(q/2)+(q^2/4)(k+1)\}=(q/4)[3+q(k+1)] \\
&\qquad \text{for } 0 \leqslant k \leqslant (N/2)-1 \\
&< (1/2)(2)\{(q/2)+(q)+(q^2/4)(k+1)\}=(q/4)[6+q(k+1)] \\
&\qquad \text{for } (N/2) \leqslant k \leqslant N-1 \\
&< (1/2)(2)\{(q/2)+(q)+(q^2/4)(2N-1-k)\} \\
&\quad =(q/4)[6+q(2N-1-k)] \\
&\qquad \text{for } N-1 \leqslant k \leqslant (3N/2)-2 \\
&< (1/2)(2)\{(q/4)+(q/2)+(q^2/4)(2N-1-k)\} \\
&\quad =(q/4)\{3+q(2N-1-k)] \\
&\qquad \text{for } (3N/2)-1 \leqslant k \leqslant 2N-2
\end{aligned} \tag{6.32}$$

The largest value of the bound is $(3q/2)[1+(Nq/6)]$, and thus, in general,

$$|\epsilon_{cr}(k)| < (3q/2)[1+(Nq/6)] \quad \text{for} \quad 0 \leqslant k \leqslant 2N-2 \tag{6.33}$$

Since $\mathcal{E}_{cr}$ is a linear function of the random variables $\mathcal{E}_{ci}$, i = 1, 6 (defined in Appendix E) which were assumed to be uncorrelated, superposition applies and hence the mean of $\mathcal{E}_{cr}$, is given by

$$\overline{\mathcal{E}_{cr}} = (1/2) \sum_{i=1}^{6} (\overline{\mathcal{E}_{ci}}) = 0 \tag{6.34}$$

and its variance is

$$\begin{aligned} \text{var}(\mathcal{E}_{cr}) &= (1/2)^2 \sum_{i=1}^{6} [\text{var}(\mathcal{E}_{ci})] \\ &\leq (1/4)(2)\{[N/(2(2N-1))]+[2N/(2N-1)]+[N^2\eta^2/(2N-1)]\}\eta^2 \\ &= [5N/(4(2N-1))]\{1+(2N\eta^2/5)\}(\eta^2) \end{aligned} \tag{6.35}$$

It should be noted that the contributions to the above variance from $\mathcal{E}_{c5}$ and $\mathcal{E}_{c6}$ are small for all practical cases where η^2 are small.

If the error of the overall frequency response is defined to be

$$\mathcal{E}^*_{cr}(e^{j\omega}) = \sum_{n=0}^{2N-2} \epsilon_{cr}(n)e^{-j\omega n} \tag{6.36}$$

then since $\epsilon_{cr}(k) = \epsilon_{cr}(2N-2-k)$, k = 0,1,...,2N-2, Equation (6.36) can be expressed as

$$\mathcal{E}^*_{cr}(e^{j\omega}) = \{ \sum_{n=0}^{N-2} 2\epsilon_{cr}(n)\cos[(N-1-n)\omega] + \epsilon_{cr}(N-1)\}e^{-j\omega(N-1)} \tag{6.37}$$

and since the factor $e^{-j\omega(N-1)}$ does not affect the magnitude of the response, it can be neglected.

Let $$\mathcal{E}_{cr}(e^{j\omega}) = \mathcal{E}^{*}_{cr}(e^{j\omega})e^{j\omega(N-1)} \tag{6.38}$$

then $$|\mathcal{E}_{cr}(e^{j\omega})| \leq \sum_{n=0}^{N-2} 2|\epsilon_{cr}(n)||\cos[(N-1-n)\omega]| + |\epsilon_{cr}(N-1)|$$

$$\leq \sum_{n=0}^{2N-2} |\epsilon_{cr}(n)| \leq \sum_{n=0}^{2N-2} [(1/2) \sum_{i=1}^{6} |\epsilon_{ci}(n)|]$$

$$\leq \sum_{j=0}^{N-1} \{ \sum_{k=0}^{N-1} |h(k)||e_p(j)| + \sum_{k=0}^{N-1} |p(k)||e_h(j)| + \sum_{k=0}^{N-1} |e_h(k)||e_p(j)| \}$$

$$< \sum_{j=0}^{N-1} \{(q/2) \sum_{k=0}^{N-1} |h(k)| + (q/2) \sum_{k=0}^{N-1} |p(k)| + (q/2)(q/2) \sum_{k=0}^{N-1} (1)\}$$

$$= (N)\{(q/2)(1) + (q/2)(2) + (q/2)^2(N)\}$$

$$= (3Nq/2)[1 + (Nq/6)] \tag{6.39}$$

In order to find the statistics of the error response, an approach similar to that of Chan and Rabiner [117] can be adopted. It can be shown that the error in the overall frequency response has a zero mean and a variance given by

$$\sigma_{cr}^{2}(\omega) = \sum_{n=0}^{N-2} 4\overline{\epsilon_{cr}^{2}(n)}\cos^2[(N-1-n)\omega] + \overline{\epsilon_{cr}^{2}(N-1)}$$

$$= \text{var}(\mathcal{E}_{cr})[1 + \sum_{n=1}^{N-1} 4\cos^2(n\omega)]$$

If a weighting function $W(\omega)$ is defined to be

$$W(\omega) = \{[1/(4N-3)][1 + 4 \sum_{n=1}^{N-1} \cos^2\omega n]\}^{(1/2)} \tag{6.40}$$

then $$\sigma_{cr}^{2}(\omega) = \text{var}(\mathcal{E}_{cr})(4N-3)W^2(\omega) \tag{6.41}$$

$W^2(\omega)$ can be expressed in closed form by summing the series in (6.40). Details of the derivation are given in Appendix F. Thus

$$W^2(\omega) = \frac{1}{(4N-3)}\{2(N-1) + [\sin(2N-1)\omega/\sin(\omega)]\} \tag{6.42}$$

Equation (6.42) shows that $W^2(0) = W^2(\pi) = 1$, and $0 < W^2(\omega) \leqslant 1$ for all N, therefore, it follows that

$$\sigma^2_{cr}(\omega) \leqslant \text{var}(\mathcal{E}_{cr})(4N-3)$$

$$\leqslant [5N(4N-3)/4(2N-1)][1+(2N\eta^2/5)]\eta^2 \qquad (6.43)$$

In the limit for large N and $0 < \omega < \pi$,

$$\lim_{N\to\infty} W^2(\omega) = \frac{1}{2}$$

The convergence of $W^2(\omega)$ to 1/2 is not uniform for all ω. However, for ω in the range $[\delta,\pi-\delta]$, for any $\delta > 0$, the convergence is uniform as can be seen from Figs.6.2(a) and 6.2(b). Thus, in the range $0 < \omega < \pi$, the bound given by (6.43) for σ^2_{cr} is overestimated by a factor of approximately 2 for large N.

Next, consider the time domain error in the output sequence.

From (6.30), $\mathcal{E}_{cs}(z) = \sum_{k=0}^{\infty} \epsilon_{cs}(k)z^{-k} = \mathcal{E}_{cr}(z)X(z)$

and so, using (6.33),

$$|\epsilon_{cs}(n)| \leqslant \sum_{k=0}^{2N-2} |\epsilon_{cr}(k)||x(n-k)|$$

$$< (3q/2)[1+(Nq/6)] \sum_{k=0}^{2N-2} |x(n-k)| \qquad (6.44)$$

From (6.44) it can thus be seen that the error $\epsilon_{cs}(n)$ depends on the specific input sequence.

In order to obtain an expression for the error which is independent of $\{x(n)\}$, it should be noted that as $|x(n)| < 1$,

$$|\epsilon_{cs}(n)| < (3q/2)[1+(Nq/6)] \sum_{k=0}^{2N-2} (1) = (3q/2)[1+(Nq/6)](2N-1) \quad (6.45)$$

The random variable $\mathcal{E}_{cs}$ consisting of $\{\epsilon_{cs}(k)\}$ has a mean equal to zero and a variance given by

$$\text{var}(\mathcal{E}_{cs}) = \sum_{k=0}^{2N-2} \epsilon_{cr}^2(k)x^2(n-k)$$

$$= \text{var}(\mathcal{E}_{cr}) \sum_{k=0}^{2N-2} x^2(n-k) \quad (6.46)$$

Alternatively, as $x^2(*) < 1$,

$$\text{var}(\mathcal{E}_{cs}) < \text{var}(\mathcal{E}_{cr})(2N-1) \lesssim (5N/4)[1+(2N\eta^2/5)]\eta^2 \quad (6.47)$$

6.4.4 Effects of Arithmetic Roundoff

Errors result when the product of two L-bit numbers exceeds L bits and rounding has to be performed.For the systems implemented in the manner described in Section 6.3, multiplications are only performed within the filters, and noise is thus added to the output sequences $\{x_l(n)\}$, $\{x_u(n)\}$, $\{t_l(n)\}$ and $\{t_u(n)\}$ from the four filters. Let $X_l(z)$, $X_u(z)$, $T_l(z)$ and $T_u(z)$ represent respectively the z-transform of these sequences and let $E_{xl}(z)$, $E_{xu}(z)$, $E_{tl}(z)$ and $E_{tu}(z)$ represent the z-transform of the respective noise sequences. The z-transform of the noisy filter outputs are thus given by

$$\hat{X}_l(z) = X_l(z)+E_{xl}(z) \quad (6.48)$$

$$\hat{X}_u(z) = X_u(z)+E_{xu}(z) \quad (6.49)$$

$$\hat{T}_l(z) = T_l(z)+E_{tl}(z) \quad (6.50)$$

and

$$\hat{T}_u(z) = T_u(z)+E_{tu}(z) \quad (6.51)$$

Using $\hat{X}_l(z)$, $\hat{X}_u(z)$, $\hat{T}_l(z)$ and $\hat{T}_u(z)$ in place of $X_l(z)$, $X_u(z)$, $T_l(z)$ and $T_u(z)$, the z-transform of the noisy output sequence can be shown to be

$$S_a(z) = S(z)+(1/2)[P_l(z)E_{xl}(z)+P_l(z)E_{xl}(-z)]+E_{tl}(z)$$
$$+(1/2)[P_u(z)E_{xu}(z)+P_u(z)E_{xu}(-z)]+E_{tu}(z) \qquad (6.52)$$

Thus the z-transform of the output noise sequence, $\mathcal{E}_{as}(z)$, is given by

$$\mathcal{E}_{as}(z) = S_a(z)-S(z) = (1/2)[P_l(z)E_{xl}(z)+P_l(z)E_{xl}(-z)]$$
$$+E_{tl}(z)+(1/2)[P_u(z)E_{xu}(z)+P_u(z)E_{xu}(-z)]+E_{tu}(z) \qquad (6.53)$$

6.4.5 Analysis of Errors Arising from Arithmetic Roundoff

The magnitudes of the errors E_{xl}, E_{xu}, E_{tl} and E_{tu} depend on the structure of the filters and on the positions within the filters where rounding is carried out. Two possibilities exist. The first possibility is one in which all multiplication products are represented exactly and rounding is performed only after they are summed. In this case, noise is only superimposed on the filter output. The output roundoff noise $\{e_q(n)\}$ is uniformly distributed between (-q/2, q/2) with zero mean and variance η^2. This error is independent of the form and structure of the filters. With the second possibility, all multiplication products are rounded before they are summed with the result that different structures will give rise to different error magnitudes.

The second possibility for rounding will now be considered in some detail. A shared-multiplier structure as described in Section 6.3 is assumed.

Consider the z-transform $E_{x_l}(z) = \sum_{k=0}^{\infty} e_{x_l}(k)z^{-k}$, in which each $e_{x_l}(k)$ is equal to the sum of errors contributed by the rounding of the products, that is,

$$e_{x_l}(k) = \sum_{n=0}^{N/2-1} e_q(n) \qquad \text{for all } k \geq 0 \tag{6.54}$$

and
$$|e_{x_l}(k)| \leq \sum_{n=0}^{N/2-1} |e_q(n)| < (q/2)(N/2)=Nq/4 \tag{6.55}$$

E_{x_l} has zero mean and a variance given by

$$\text{var}(E_{x_l}) = \sum_{n=0}^{N/2-1} \overline{e_q^2(n)} = (N/2)(\eta^2) \tag{6.56}$$

Similar results can be obtained for $E_{xu}(z)$, $E_{t_l}(z)$ and $E_{tu}(z)$.

The analysis of the resultant error can be carried out by considering the components separately.

Let
$$\mathcal{E}_{a1}(z) = \sum_{k=0}^{\infty} \epsilon_{a1}(k)z^{-k} = P_l(z)E_{x_l}(z) \tag{6.57}$$

thus
$$\epsilon_{a1}(k) = \sum_{n=0}^{N-1} p_l(n)e_{x_l}(k-n) \tag{6.58}$$

$$|\epsilon_{a1}(k)| \leq \sum_{n=0}^{N-1} |p_l(n)||e_{x_l}(k-n)| < (Nq/4) \sum_{n=0}^{N-1} |p_l(n)|$$

and by using (6.15), it can be seen that,

$$|\epsilon_{a1}(k)| \leq (Nq/4)(2) = Nq/2 \tag{6.59}$$

The mean of $\mathcal{E}_{a1}$ is given by

$$\mathcal{E}_{a1} = \sum_{n=0}^{N-1} p_l(n)\overline{e_{x_l}(k-n)} = 0 \tag{6.60}$$

and its variance is given by

$$\text{var}(\mathcal{E}_{a1}) = \sum_{n=0}^{N-1} p_l^2(n)\overline{e_{x_l}^2(k-n)}$$

$$= (N/2)(\eta^2) \sum_{n=0}^{N-1} p_l^2(n)$$

$$\leq (N/2)(\eta^2)(2) = N\eta^2 \qquad (6.61)$$

Similar results can be obtained for $\mathcal{E}_{a2}(z) = P_l(z)E_{x_l}(-z)$, $\mathcal{E}_{a3}(z) = P_u(z)E_{xu}(z)$ and $\mathcal{E}_{a4}(z) = P_u(z)E_{xu}(-z)$.

Thus if $\mathcal{E}_{as}(z) = \sum_{k=0}^{\infty} \epsilon_{as}(k)z^{-k}$

then $$|\epsilon_{as}(k)| < (1/2)\{|\epsilon_{a1}(k)|+|\epsilon_{a2}(k)|+|\epsilon_{a3}(k)|+|\epsilon_{a4}(k)|\}$$

$$+|e_{tl}(k)|+|e_{tu}(k)|$$

$$= (3Nq/2) \qquad (6.62)$$

If it is assumed that $\mathcal{E}_{a1}$, $\mathcal{E}_{a2}$, $\mathcal{E}_{a3}$, $\mathcal{E}_{a4}$, E_{t_l} and E_{tu} are uncorrelated, then the mean of the random variable $\mathcal{E}_{as}$ is zero and its variance is given by

$$\text{var}(\mathcal{E}_{as}) \leq (1/4)\{N\eta^2+N\eta^2+N\eta^2+N\eta^2\}+(N/2)\eta^2+(N/2)\eta^2$$

$$= 2N\eta^2 = Nq^2/6 \qquad (6.63)$$

which is independent of the magnitude of the input signal.

6.4.6 Effects of Combined Coefficient Quantization and Arithmetic Roundoff

In this section, the errors resulting from both coefficient quantization and the arithmatic roundoff are analysed. A case of

practical importance in which the same word-length is used throughout for data storage and arithmetic operations will now be considered.

The z-transform of the noisy output sequence is

$$\begin{aligned} S_b(z) = {} & (1/2)[\hat{H}_l(z)\hat{P}_l(z)-\hat{H}_l(-z)\hat{P}_l(-z)]X(z) \\ & +(1/2)\hat{P}_l(z)[E_{x_l}(z)+E_{x_l}(-z)]+E_{t_l}(z) \\ & -(1/2)\hat{P}_l(-z)[E_{xu}(z)+E_{xu}(-z)]+E_{tu}(z) \\ = {} & S(z)+ \bar{\mathcal{E}}_{cs}(z)+ \mathcal{E}_{as}(z)+ \mathcal{E}_{is}(z) \end{aligned} \tag{6.64}$$

where $$\mathcal{E}_{is}(z) = (1/2)E_p(z)[E_{x_l}(z)+E_{x_l}(-z)]$$

$$-(1/2)E_p(-z)[E_{xu}(z)+E_{xu}(-z)] \tag{6.65}$$

The z-transform of the output noise sequence, due to combined coefficient quantization and arithmetic roundoff is thus

$$\mathcal{E}_{bs}(z) = \mathcal{E}_{cs}(z) + \mathcal{E}_{as}(z) + \mathcal{E}_{is}(z)$$

Note that this consists of separate contributions from coefficient quantization and arithmetic rounding and from the interaction term $\mathcal{E}_{is}(z)$.

6.4.7 Analysis of Errors Associated with Coefficient Quantization and Arithmetic Roundoff

Consider the interaction term $\mathcal{E}_{is}(z)$.

Let $$\mathcal{E}_{i1}(z) = \sum_{k=0}^{\infty} \epsilon_{i1}(k)z^{-k} = E_p(z)E_{x_l}(z) \tag{6.66}$$

where $$\epsilon_{i1}(k) = \sum_{n=0}^{N-1} e_p(n)e_{x_l}(k-n) \tag{6.67}$$

Using (6.55), and assuming that the random variables E_{p_l} and E_{x_l} are independent, it can be seen that,

$$|\epsilon_{i1}(k)| \leq \sum_{n=0}^{N-1} |e_p(n)||e_{x_l}(k-n)|$$

$$< (q/2)(Nq/4)(N) = (Nq)^2/8 \tag{6.68}$$

Similar results can be obtained for the other three terms of $\mathcal{E}_{is}(z)$.

Therefore, $$|\epsilon_{is}(k)| < 4(1/2)(Nq)^2/8 = (Nq)^2/4 \tag{6.69}$$

and $$|\epsilon_{bs}(k)| < |\epsilon_{cs}(k)|+|\epsilon_{as}(k)|+|\epsilon_{is}(k)|$$

$$< (q/4)(18N+3N^2q-Nq-6) \tag{6.70}$$

The random variable $\mathcal{E}_{i1}$, has zero mean, and its variance is given by

$$\text{var}(\mathcal{E}_{i1}) = (\eta^2)(N\eta^2/2) \sum_{n=0}^{N-1} (1) = (N/2)(\eta^2)^2(N)$$

$$= (N^2/2)(\eta^4) \tag{6.71}$$

Similar results can also be obtained for the other three product terms of $\mathcal{E}_{is}(z)$. The random variable $\mathcal{E}_{is}$ consists of a linear combination of four random variables which are assumed to be uncorrelated. The mean of $\mathcal{E}_{is}$ is thus zero and its variance is

$$\text{var}(\mathcal{E}_{is}) = 4(1/4)(N^2/2)(\eta^4) = (N^2/2)(\eta^4) \tag{6.72}$$

Also if it is assumed that $\mathcal{E}_{cs}$, $\mathcal{E}_{as}$ and $\mathcal{E}_{is}$ are uncorrelated, then the mean of $\mathcal{E}_{bs}$ is

$$\overline{\mathcal{E}_{bs}} = \overline{\mathcal{E}_{cs}} + \overline{\mathcal{E}_{as}} + \overline{\mathcal{E}_{is}} = 0 \tag{6.73}$$

and its variance is

$$\text{var}(\mathcal{E}_{bs}) = \text{var}(\mathcal{E}_{cs})+\text{var}(\mathcal{E}_{as})+\text{var}(\mathcal{E}_{is})$$

$$< (5N/4)[1+(2N\eta^2/5)](\eta^2)+(2N)(\eta^2)+(N^2/2)(\eta^4)$$

$$= (13N/4)(\eta^2)+(N^2\eta^4) \qquad (6.74)$$

As η^2 is usually small, the contribution of the error variance due to interaction can be ignored. Var($\mathcal{E}_{bs}$) can be effectively obtained by simple addition of the error variances of the component sources.

6.4.8 Effects of Quantization for Transmission Bit Rate Reduction

In order to transmit the split-band signal at low bit rates, the decimated sequences $\{y_l(n)\}$ and $\{y_u(n)\}$ are compression coded using only very few bits. In the following analysis, linear quantization is assumed for each individual element of the sequences. With this assumption, the quantization of the decimated sequences is the same as that obtained through the quantization and subsequent decimation of the sequences $\{x_l(n)\}$, $\{x_u(n)\}$.

Let $\hat{X}_l(z)$ and $\hat{X}_u(z)$ denote the z-transforms of the quantized sequences $\{\hat{x}_l(n)\}$ and $\{\hat{x}_u(n)\}$ respectively and define $E_{pl}(z)$ and $E_{pu}(z)$ such that

$$E_{pl}(z) = \hat{X}_l(z)-X_l(z) \qquad (6.75)$$

$$E_{pu}(z) = \hat{X}_u(z)- X_u(z) \qquad (6.76)$$

The z-transform of the noisy output sequence is then

$$S_p(z) = S(z)+(1/2)P_l(z)[E_{pl}(z)+E_{pl}(-z)]$$

$$-(1/2)P_l(-z)[E_{pu}(z)+E_{pu}(-z)] \qquad (6.77)$$

and if the z-transform of the noise sequence imposed on the output is defined to be $\mathcal{E}_{ps}(z)$,

then $$\mathcal{E}_{ps}(z) = (1/2)P_l(z)[E_{pl}(z)+E_{pl}(-z)]$$

$$-(1/2)P_l(-z)[E_{pu}(z)+E_{pu}(-z)] \qquad (6.78)$$

From Equation (6.78) it should be noted that part of the noise is in fact aliasing noise. Thus frequency aliasing cannot be completely eliminated if the signals $\{y_l(n)\}$ and $\{y_u(n)\}$ are quantized. The frequency spectrum of the output noise including the aliasing noise depends on the spectra of $\{y_l(n)\}$, $\{y_u(n)\}$ as well as the frequency responses of the filters P_l and P_u.

6.4.9 Analysis of Errors Associated with Quantization for Transmission-Bit-Rate Reduction

Consider

$$\mathcal{E}_{p1}(z) = P_l(z)E_{pl}(z) \qquad (6.79)$$

As $$\epsilon_{p1}(n) = \sum_{n=0}^{N-1} p_l(k)e_{pl}(n-k) \qquad (6.80)$$

so $$|\epsilon_{p1}(n)| \leq \sum_{n=0}^{N-1} |p_l(k)||e_{pl}(n-k)|$$

$$< (q/2)\sum_{n=0}^{N-1} |p_l(k)| \leq (q/2)(2) = q \qquad (6.81)$$

The random variable $\mathcal{E}_{p1}$ has zero mean and a variance given by

$$\text{var}(\mathcal{E}_{p1}) = \sum_{n=0}^{N-1} p_l^2(k)\overline{e_{p_l}^2(n-k)}$$

$$= (\eta^2) \sum_{n=0}^{N-1} p_l^2(k) \leqslant (\eta^2)(2) \qquad (6.82)$$

Similar results can be obtained for the other terms.

Thus, if $\mathcal{E}_{ps}(z) = \sum_{k=0}^{\infty} \epsilon_{ps}(k)z^{-k}$ (6.83)

then $|\epsilon_{ps}(k)| < (1/2)(q + q + q + q) = 2q$ (6.84)

and if it is assumed that the four terms of $\mathcal{E}_{ps}(z)$ are uncorrelated, then the mean of $\mathcal{E}_{ps}$ is zero and its variance is

$$\text{var}(\mathcal{E}_{ps}) \leqslant 4(1/2)^2(2\eta^2) = 2\eta^2 = q^2/6 \qquad (6.85)$$

The error variance var($\mathcal{E}_{ps}$) is thus independent of N, the length of the filters.

6.5 SIMULATION RESULTS

A computer simulation was carried out using a CDC 6500 computer which operates to 28 significant digits when the machine is used in the double-precision mode. The results obtained from computations using double-precision arithmetic were taken as being 'infinitely-accurate'. The system structure used was as described in Section 6.3. The impulse response and characteristics of the filters used are given in Appendix G.

The error variance $\sigma_{cr}^2(\omega)$ of the frequency response of the overall system using finite-word-length coefficients with filters of

lengths N = 8, 16 and 32 are shown in Figs.6.3(a), (b) and (c) respectively. The input used in this case was an impulse.

As indicated in Section 6.4.3, the bound (6.43) is overestimated by a factor of approximately 2 for large N. Thus, most of the points lie at a distance below the bound, especially when N is large.

Using the theoretical estimate of variance, lines of deviations of $\sigma_{cr}(\omega)$ and $2\sigma_{cr}(\omega)$ from the 'infinitely accurate' maximum peak-to-peak ripple have been derived and in Figs.6.4(a), (b) and (c) these are superimposed on the graphs showing the maximum peak-to-peak ripple of the overall frequency response of the systems when the filter coefficients are quantized with different numbers of bits. It can be seen that the actual peak-to-peak ripple of the system frequency response for coefficients quantized with the number of bits shown are within $2\sigma_{cr}(\omega)$ of the nominal values for the three cases investigated with filters of lengths N=8, 16 and 32. The peak-to-peak ripple in some cases are smaller than the nominal values as the method adopted for the filter design does not necessarily provide a global optimal solution.

The error variance of ξ_{cs}, defined in Section 6.4.3, was computed using a) speech, b) Gaussian distributed random data and c) uniformly distributed random data. The computed variances, together with the theoretical bounds (6.47) are shown in Figs.6.5(a), (b) and (c) for systems with filters of lengths 8, 16 and 32 respectively. The speech input comprised 512 samples obtained by the analogue-to-digital conversion of a segment of voiced speech. The speech was sampled at 8 ksample/s. The mean-square value of these

samples (after gain adjustment) was .036039 . The Gaussian random samples were obtained from a normal distribution having zero mean and a standard deviation of 0.25, with samples having a magnitude greater than 1 being discarded. The mean-square value of the 512 samples was 0.0640976. The uniformly distributed data ranged from -1 to 1 and the mean-square value of the 512 samples used was 0.344866.

From Figs.6.5(a), (b) and (c), it can be seen that the error variances computed from simulation results are below the theoretical bounds. For a given word-length, the magnitude of the error variance for the various inputs is seen to be approximately in proportion to the mean-square value of the respective input. This is in agreement with the prediction given in (6.46).

Simulation results of errors arising from arithmetic roundoff together with the theoretical values calculated using (6.63) for systems using filters of lengths 8, 16 and 32 are shown in Figs.6.6(a), (b) and (c) respectively. It can be seen that the experimental results using the three different input sequences are in excellent agreement with the theoretical bounds.

Figs.6.7(a), (b) and (c) show the results of a computer simulation when both the coefficients and the multiplication products are rounded. The results are shown for the three types of input data and filter lengths of 8, 16 and 32. Also shown in the figures are the theoretical bounds of var($\mathcal{E}_{bs}$), the error variance due to coefficient quantization and arithmetic roundoff.

It should be noted that as the values of var($\mathcal{E}_{cs}$) are dependent on the values of the input as indicated by Equation (6.46), its bound

(6.47), derived on the assumption that all input data take on the maximum values, is overestimated for most practical cases. As var($\mathcal{E}_{cs}$) is one of the constituent terms of var($\mathcal{E}_{bs}$), the bound on var($\mathcal{E}_{bs}$), given by Equation (6.74), is also overestimated for most practical cases. For errors arising from quantization associated with transmission bit rate reduction, the theoretical values and results calculated from experimental data for various values of L are compared in Fig.6.8. It can be observed that there is excellent agreement between the two sets of values for the range of L considered, especially for L larger than 4.

6.6 CONCLUSIONS

Problems relating to the implementation of split-band coding systems employing quadrature mirror filters have been investigated and theoretical bounds and statistics have been derived for errors arising from the various sources of quantization involved. The results derived provide guidelines for the design and implementation of split-band systems. A simulation has been performed and excellent agreement has been found between the simulation results and those predicted theoretically. The closeness of the agreement indicates that the assumptions relating to the various noise models hold good in practice.

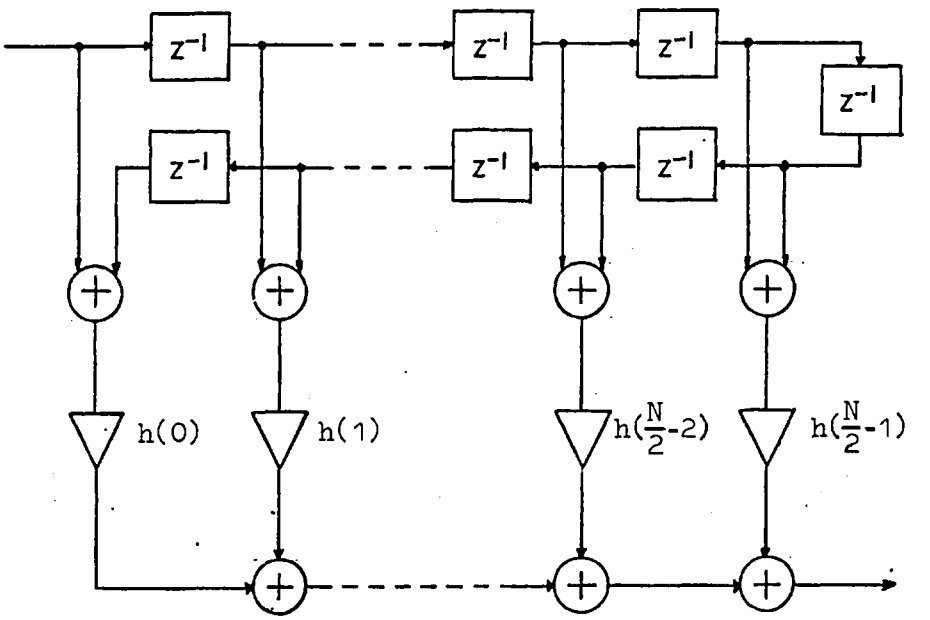
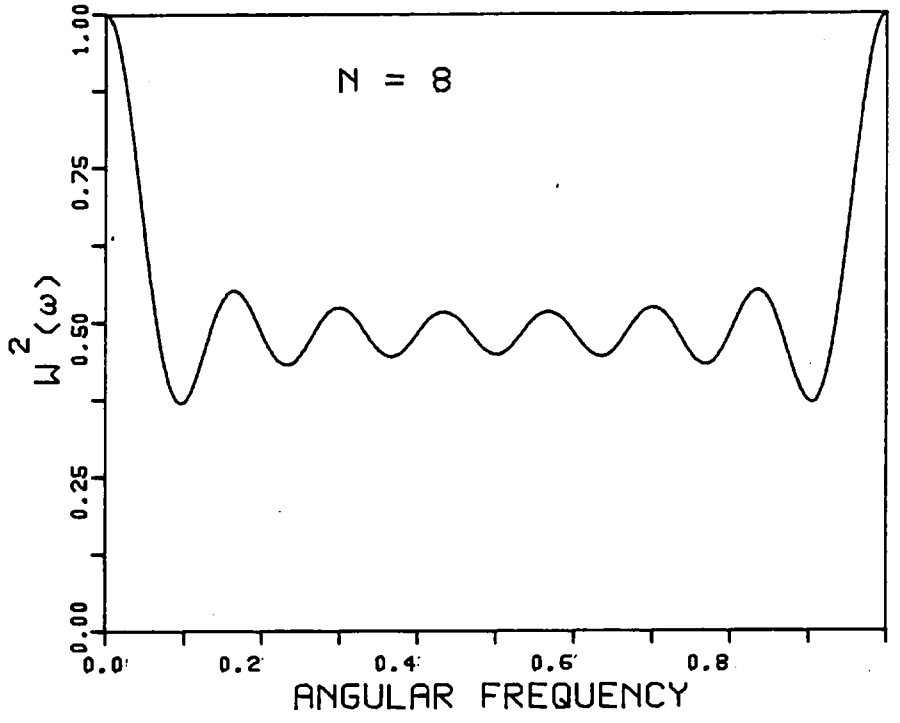
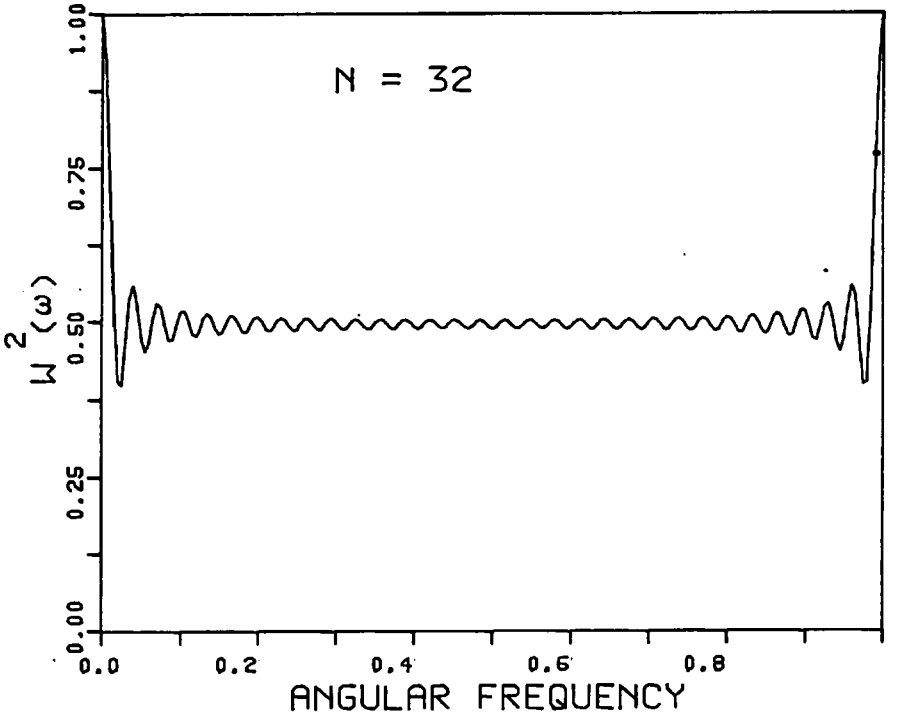


Fig.6.1 Block diagram of a shared multiplier structure.

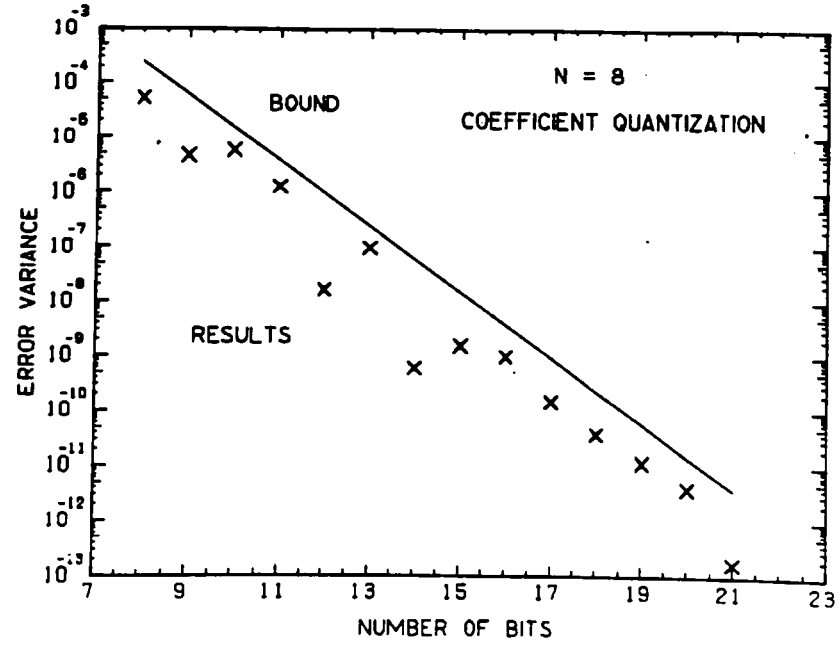


(a)

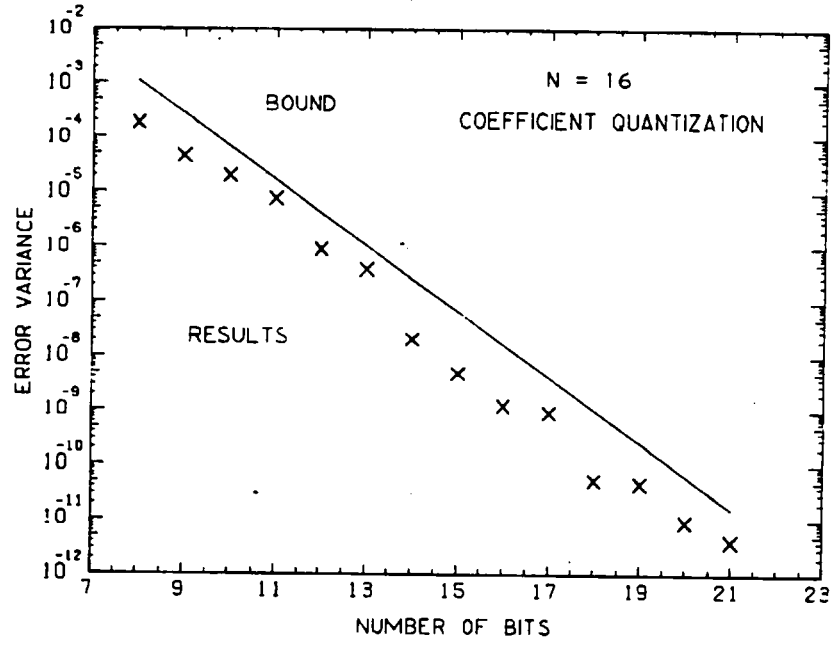


(b)

Fig.6.2 Variation of the weighting function with angular frequency; (a) N=4; (b) N=32.



(a)



(b)

Fig.6.3 Error variance $\sigma^2_{cr}(\omega)$ of the frequency response of systems due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32.

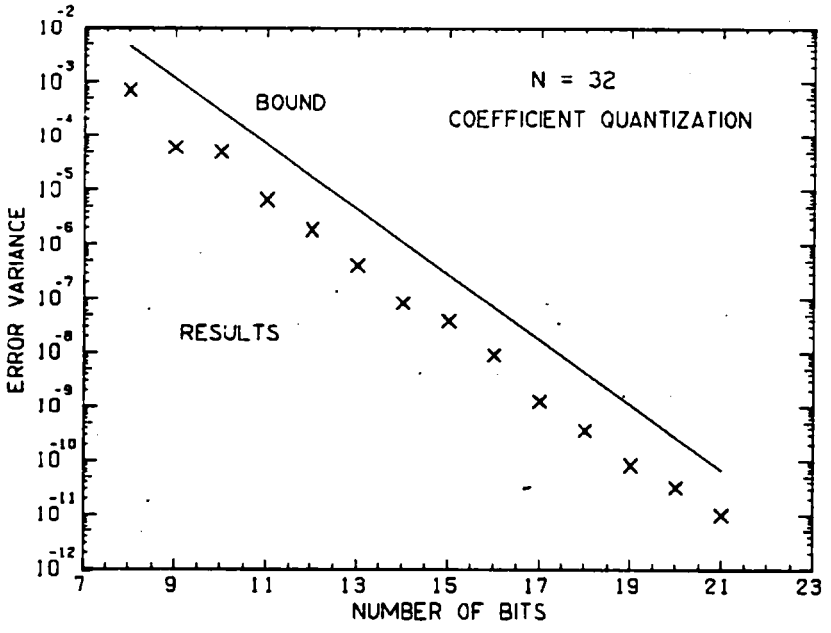


Fig.6.3(c)

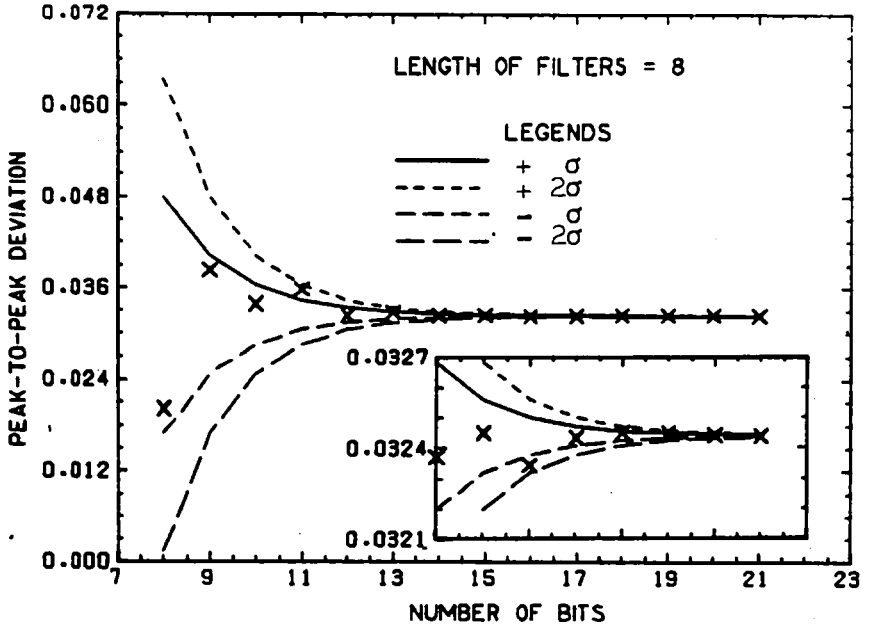
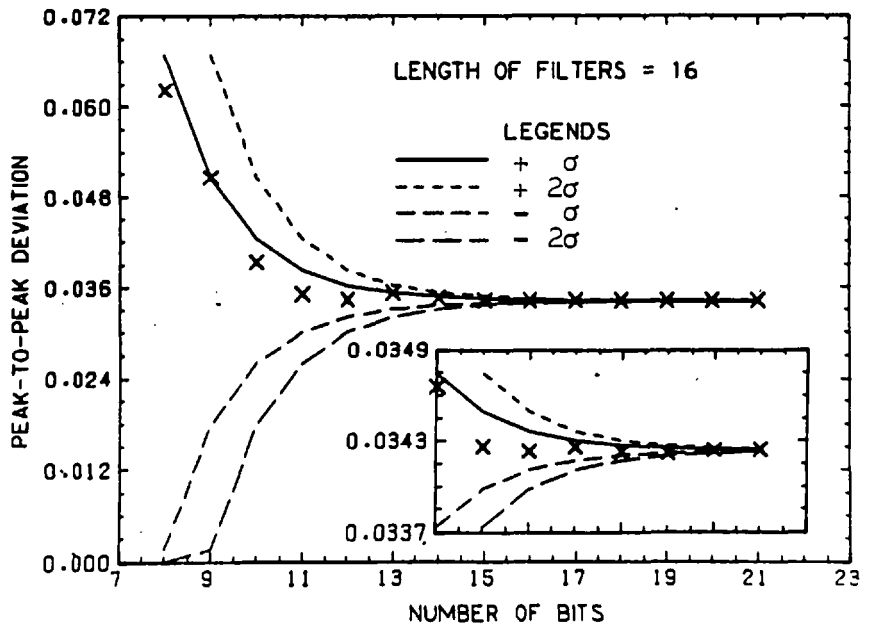
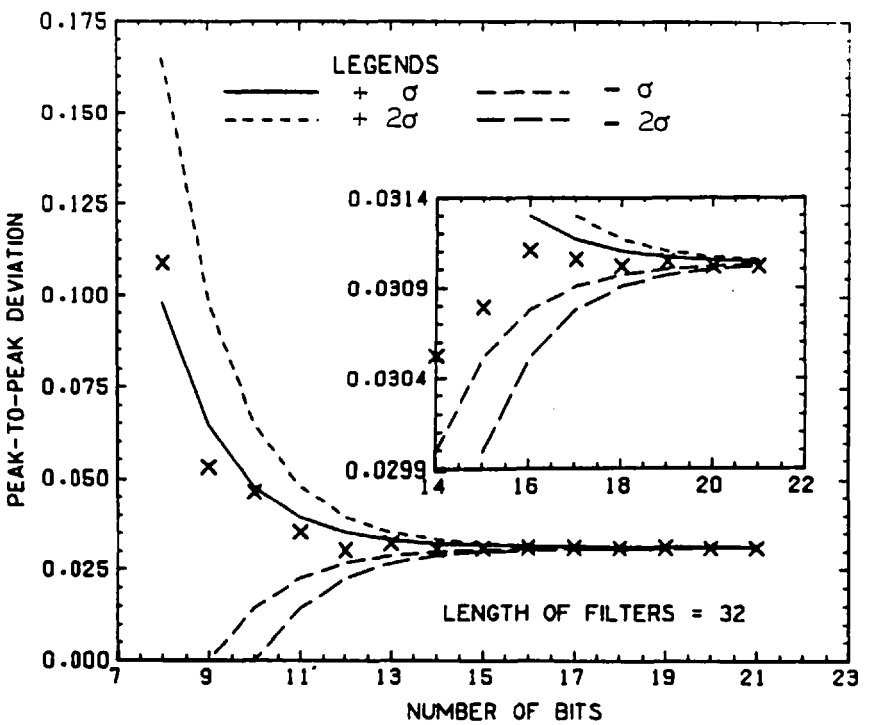


Fig.6.4(a)



(b)



(c)

Fig.6.4 Illustration of deviation of peak-to-peak ripple from nominal value due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32.

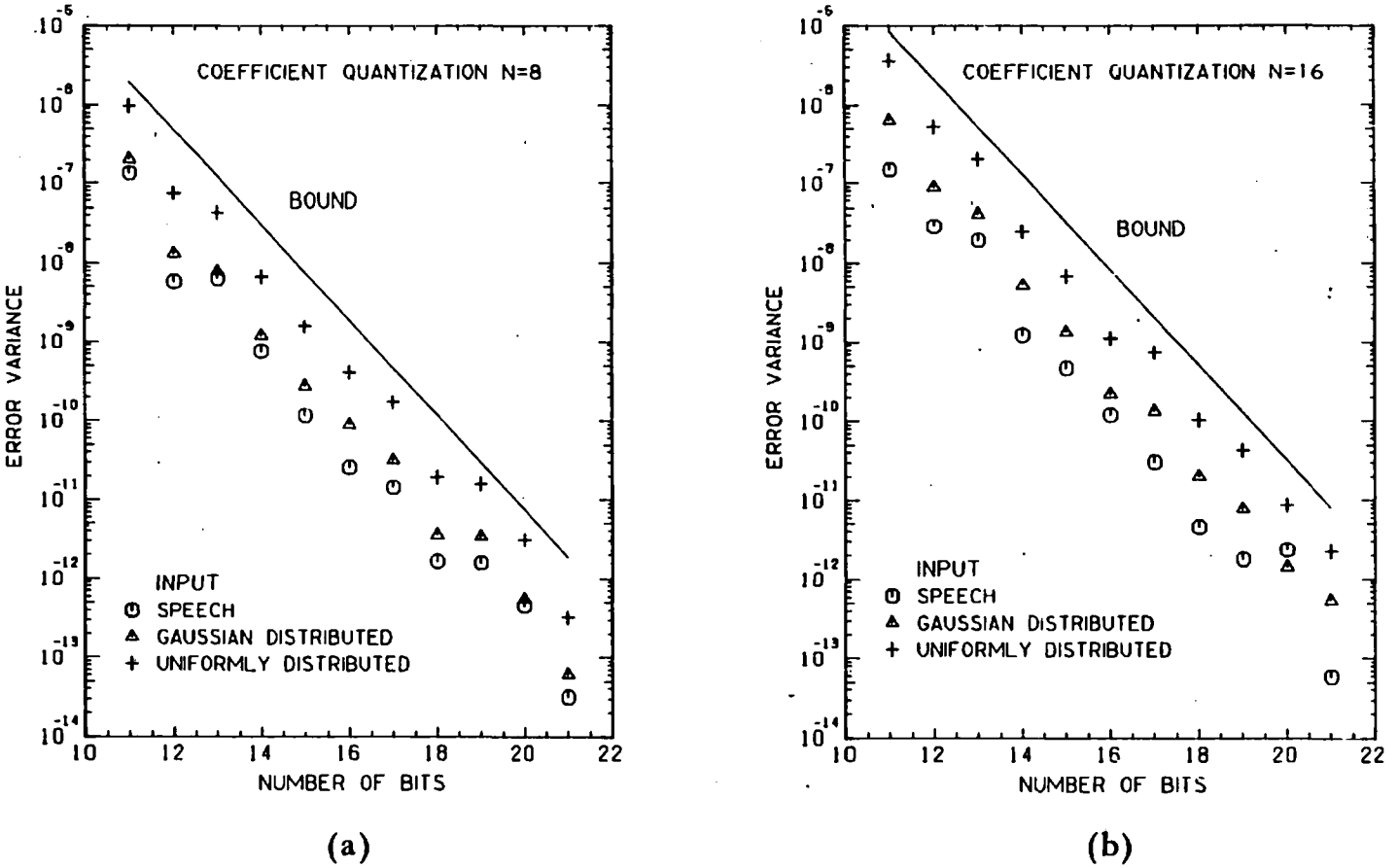


Fig.6.5 Error variances of output due to coefficient quantization; (a) N=8; (b) N=16; (c) N=32.

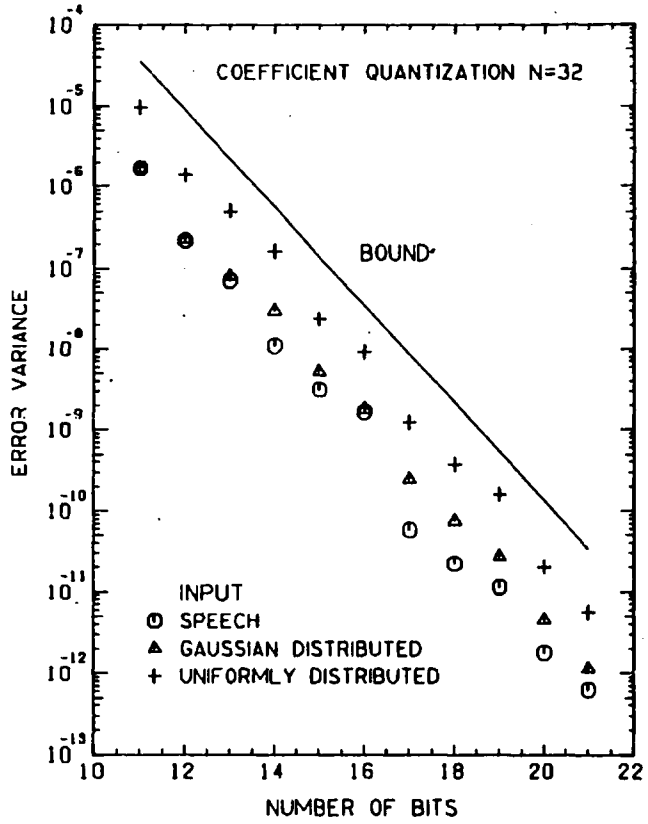


Fig.6.5(c)

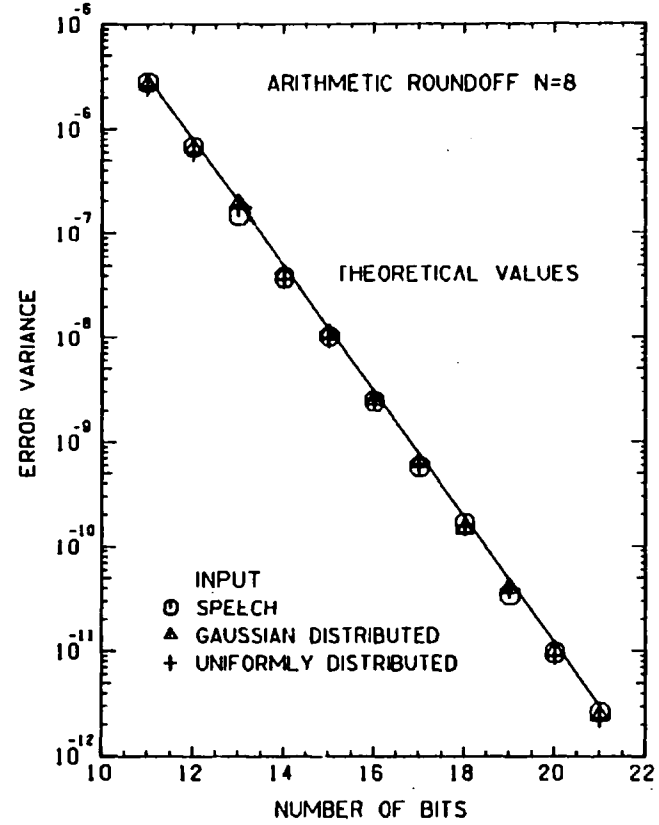
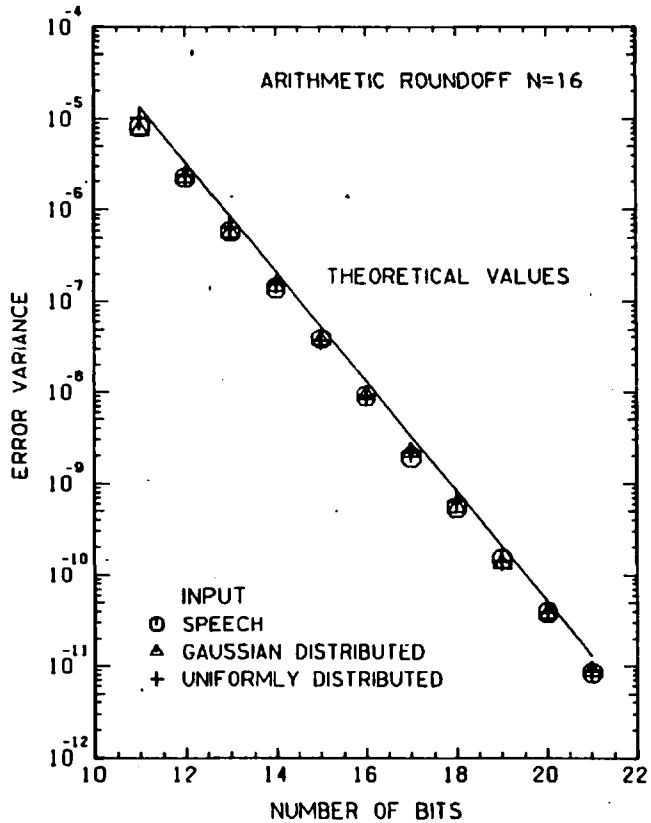


Fig.6.6(a)



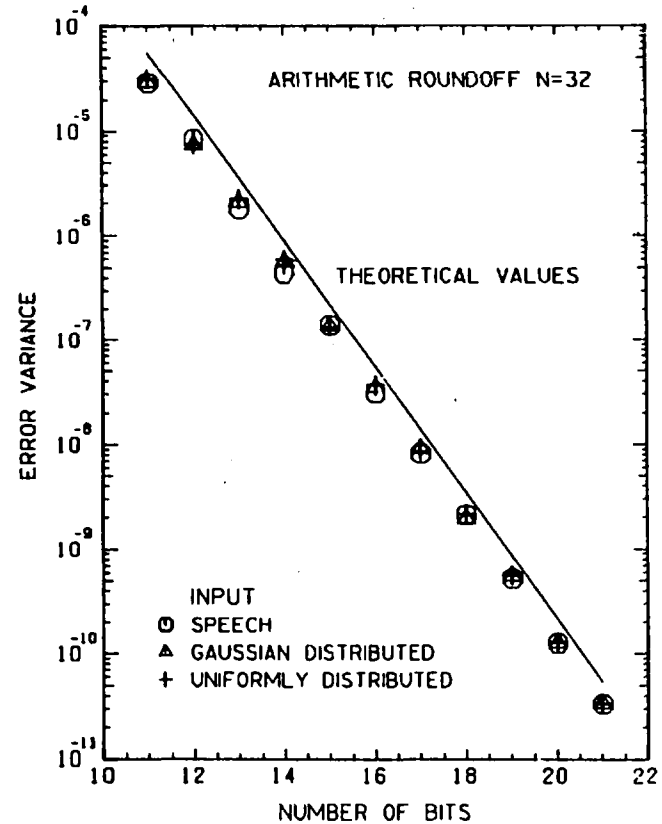


Fig.6.6 Error variances of output due to arithmetic roundoff; (a) N=8; (b) N=16; (c) N=32.

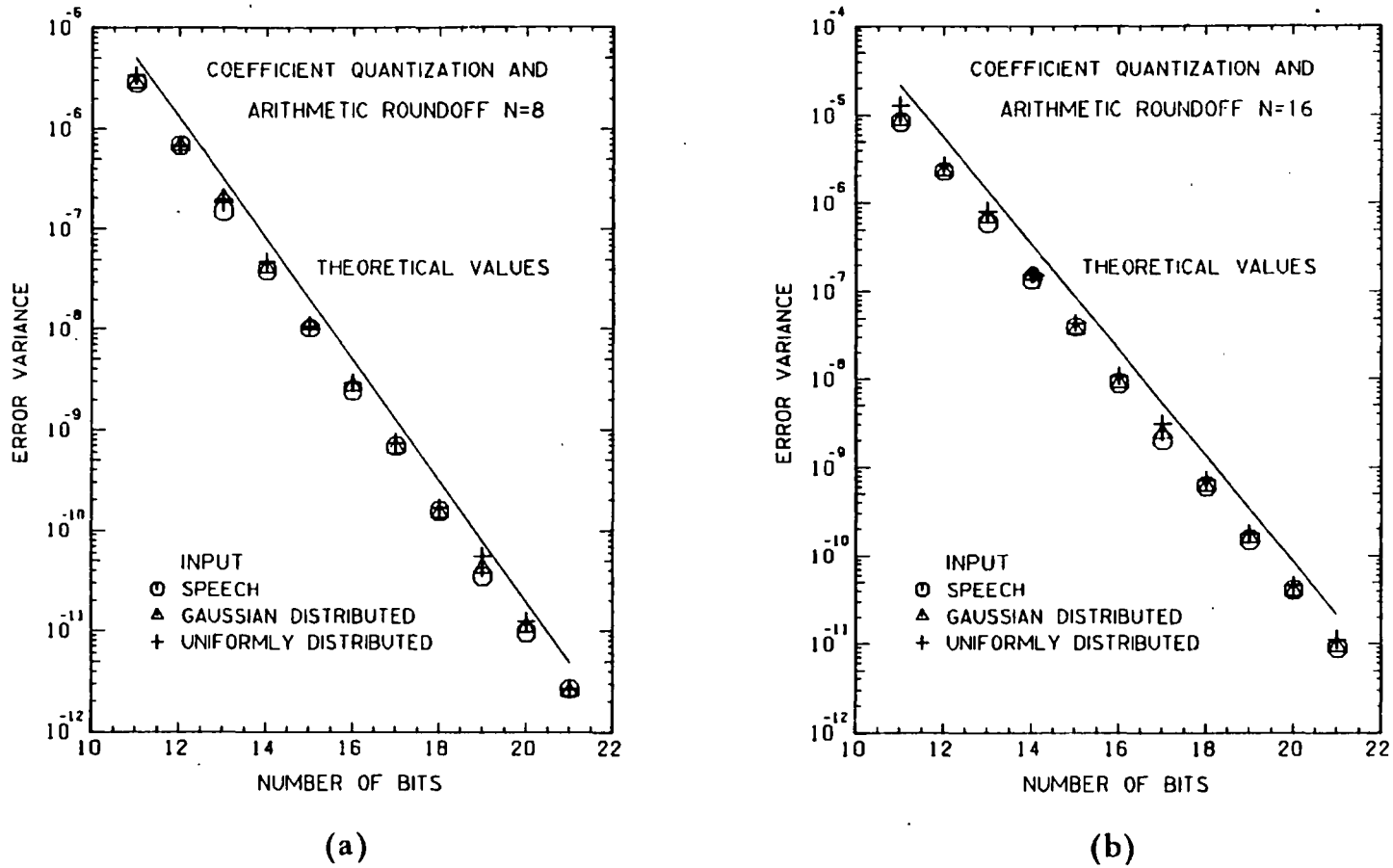


Fig.6.7 Error variances of output due to coefficient quantization and arithmetic roundoff; (a) N=8; (b) N=16; (c) N=32.

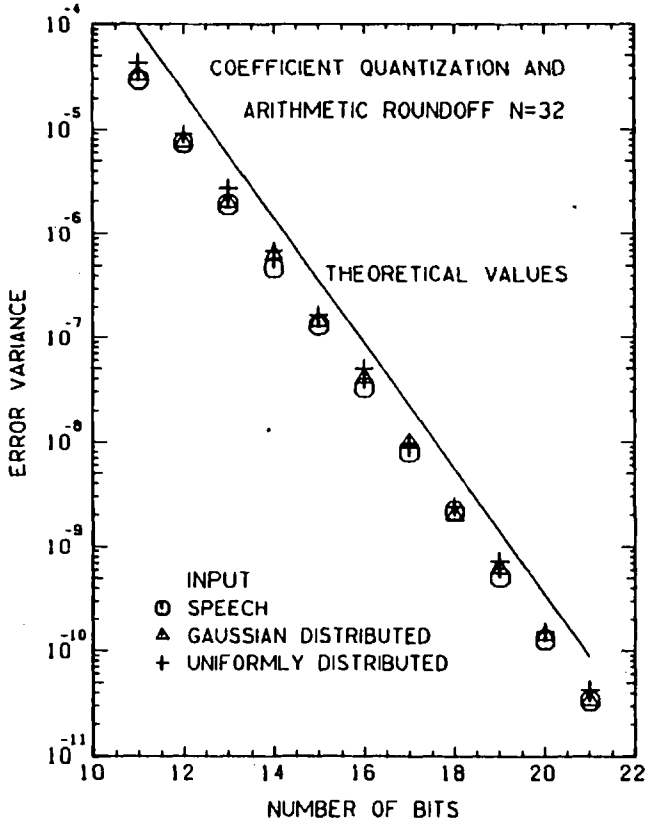


Fig.6.7(c)

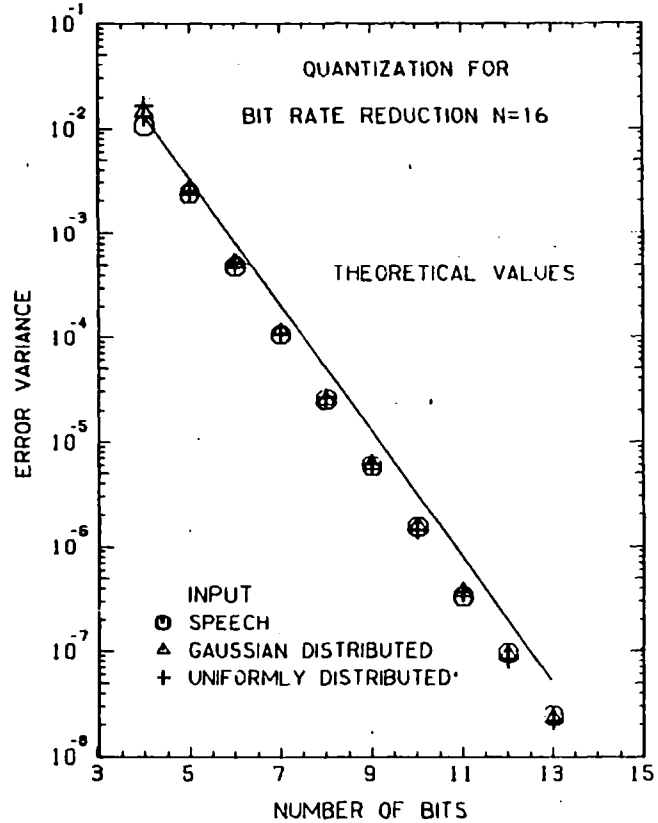


Fig.6.8 Error variances of output due to quantization for bit rate reduction.

CHAPTER 7

ERROR ANALYSIS OF BLOCK-FLOATING-POINT FAST WALSH-HADAMARD TRANSFORMS

CHAPTER 7

ERROR ANALYSIS OF BLOCK-FLOATING-POINT FAST WALSH-HADAMARD-TRANSFORMS

7.1 INTRODUCTION

The one-dimensional discrete Walsh-ordered Walsh-Hadamard transform of size N of the real signal x(t) can be defined as

$$W(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)Wal(k,n) \qquad k=0,1,\ldots,N-1 \qquad (7.1)$$

where W(k) is the k^{th} coefficient, x(n) the n^{th} sample of x(t) and Wal(k,n) is the Walsh function.

Similarly, the k^{th} coefficient B(k) of the discrete Hadamard-ordered Walsh-Hadamard transform is related to the Hadamard function Had(k,n) by

$$B(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)Had(k,n) \qquad k=0,1,\ldots,N-1 \qquad (7.2)$$

The two transforms differ only in the ordering of the transform coefficients. Also, for both the Hadamard-ordered and the Walsh-ordered transforms, apart from the multiplier (1/N), the inverse transform matrix is identical to the transform matrix.

In performing the fast Walsh-Hadamard transform (FWHT) [74], the values NW(k) or NB(k) are computed through a sequence of intermediate stages and the final coefficients W(k) or B(k) are obtained by dividing NW(k) and NB(k) by N. By applying Parseval's theorem, it can be shown that

$$\sum_{k=0}^{N-1} [NB(k)]^2 = \sum_{k=0}^{N-1} [NW(k)]^2 = N \sum_{n=0}^{N-1} [x(n)]^2 \qquad (7.3)$$

and hence, from Equation (7.3), it is clear that the mean-square values of NW(k) and NB(k) are greater than the mean-square value of the input x(n). Therefore, some intermediate values derived for NW(k) and NB(k) must be larger than the input values x(n), and overflow will thus result if finite-word-length storage is used for all the intermediate data and no adjustment is made. At each stage of the algorithm in the execution of the FWHT algorithm, the following two types of operation are performed:

$$X_{m+1}(i) = X_m(i) + X_m(j) \qquad (7.4)$$

$$X_{m+1}(j) = X_m(i) - X_m(j) \qquad (7.5)$$

with the result that if the size of the transform is $N = 2^M$, where M is the number of stages of such operations, then (N/2) additions and (N/2) subtractions are performed at each stage. As a result of this, the maximum absolute value of the data increases from stage to stage. The maximum absolute value, however, will not increase by more than a factor of 2. Thus, using Equations (7.4) and (7.5), it can easily be verified that,

$$\max\{|X_m(i)|\} \leq \max\{|X_{m+1}(i)|\} \leq 2\max\{|X_m(i)|\} \quad i=0,1,\ldots,N-1 \qquad (7.6)$$

7.2 ERROR ANALYSIS

In this section, two methods that are commonly used for reducing the effects of overflow are examined and an analysis is presented which provides a guideline in the selection of the word-length when the methods are used in conjunction with block-floating-point arithmetic to obtain the fast Walsh-Hadamard transform. The two methods are as follows:

Method 1. At each stage of the iteration, all intermediate binary data are shifted to the right by 1 bit, that is, divided by 2. This guarantees that there is no overflow and also dispenses with the need for the register to keep track of the number of shift operations. The weakness of this method is that a loss of accuracy can result since overflow will not necessarily occur at every stage of the iteration.

Method 2. Shifting is carried out only when there is an overflow at that stage. From Equation (7.6), it is clear that at most one shift is required at each stage. The method is one of the most efficient ways of performing the FWHT when the data array is represented in block-floating-point form.

In number representations using the block-floating-point form, in addition to a fixed-point mantissa assigned for each individual number, one fixed-length exponent is used for a whole block, or array, of numbers. The value of the fixed exponent is obtained by examining all the numbers in the block and representing the largest number as an ordinary floating-point number with a normalized mantissa. Memory storage is thus considerably reduced as only a single exponent is required for a large block, or array, of numbers.

In the analysis which follows, it is assumed that the data are represented in block-floating-point form and that the mantissa associated with each element of the input array {x(n)} is represented by L bits, inclusive of a sign bit. L thus denotes the word-length. The analysis is carried out for the values NW(k) and NB(k), as division by N, where N is a power of 2, can be achieved by simply changing the value of the exponent.

The sources of errors that propagate through the computation are:

(i) error due to inaccuracy of the input data, for example, analog-to-digital conversion error. Let the variance of this error be denoted by q^2.

(ii) error due to the dropping of the last bit when shifting is carried out as a result of overflow. If the bit that is dropped is 0, there is no error. If, however,it is 1, then an error equal to $2^{-(L-1)}$ or $-2^{-(L-1)}$ results depending on the sign of the data. If it is assumed that the probabilities of dropping a 1 or a 0 are equal, then the variance of the error, denoted by e^2, is equal to $2^{-2(L-1)}/2 = 2^{-2L+1}$.

An absolute upper bound to the error associated with any of the coefficients can be derived assuming that the error occurring at each stage in the computation is a maximum.

If the maximum absolute value of errors due to inaccuracy of input data is q_0, then at the first stage, addition or subtraction of two numbers each with error q_0 is $2q_0$ and shifting the sum/difference to the right by one bit results in a maximum error of $e_0 = 2^{-(L-1)}$. At

the second stage, the error due to the addition or subtraction of two numbers each with error $2q_0+e_0$ is at most $2(2q_0+e_0)$, and the maximum error due to the dropping of the last bit of the resulting number is $2[2^{-(L-1)}] = 2e_0$, so the maximum total error at the second stage is $4q_0+4e_0$. By induction, it follows that the maximum error associated with each coefficient for a transform of size 2^M is $2^Mq_0+M[2^{(M-1)}]e_0$.

Using Equation (7.3), the mean-square value of the transform $Q = (1/N)\sum_{n=0}^{N-1}[NW(n)]^2$ is related to the mean-square value of the input, $P = (1/N)\sum_{n=0}^{N-1}[x(n)]^2$ by $Q = 2^MP$. The maximum error, E_{max}, normalised by the root-mean-square value of the transform becomes

$$E_{max} = [2^Mq_0+M2^{(M-1)}2^{-(L-1)}]/(2^MP)^{1/2}$$

$$= (q_0+M2^{-L})_2(M/2)_P(-1/2) \qquad (7.7)$$

This absolute bound is useful in determining the word-length to be used when the error of the individual transform coefficient is not to exceed a specified value. However, the absolute bound is in general too large for cases where only the average error is of interest and, in these cases, an estimate of the error variance is more useful. An estimate of the error variance can be readily derived if it is assumed that there is an overflow at each stage.

At each stage of the FWHT algorithm, the error results from two sources: (i) the addition or subtraction of two numbers with their associated errors; and (ii) the dropping of one bit from the number representing the sum or difference. Since both addition and subtraction are performed on the same two numbers at each stage, the error variance of the sum/difference of two numbers with their

associated errors is equal to the sum of the error variances of the individual numbers. This can be verified as follows. Using the notations given in Equations (7.4) and (7.5), the error variance due to addition at the $(m+1)^{th}$ stage, σ_a^2, is given by

$$\sigma_a^2 = \text{var}(X_m(i)) + \text{var}(X_m(j)) + 2\rho_{ij}\sqrt{\text{var}(X_m(i))\text{var}(X_m(j))} \quad (7.8)$$

where $\text{var}(X_m(i))$ denotes the error variance of $X_m(i)$ and ρ_{ij} is the correlation coefficient of the errors associated with $X_m(i)$ and $X_m(j)$. Similarly, the error variance due to subtraction, σ_s^2, is given by

$$\sigma_s^2 = \text{var}(X_m(i)) + \text{var}(X_m(j)) - 2\rho_{ij}\sqrt{\text{var}(X_m(i))\text{var}(X_m(j))} \quad (7.9)$$

As both addition and subtraction are present at every stage, so the part of error variance, σ_o^2, arising from the addition/subtraction operation is the average of σ_a^2 and σ_s^2, that is,

$$\sigma_o^2 = \text{var}(X_m(i)) + \text{var}(X_m(j)) \quad (7.10)$$

Since $$\text{var}(X_m(i)) = \text{var}(X_m(j)) = \text{var}(X_m) \quad (7.11)$$

so $$\sigma_o^2 = 2\text{var}(X_m) \quad (7.12)$$

The error variance at the $(m+1)^{th}$ stage, $\text{var}(X_{m+1})$, can then be written as

$$\text{var}(X_{m+1}) = 2\text{var}(X_m) + 2^{(m-1)}e^2 + K\sqrt{\text{var}(X_m)2^{(m-1)}e^2} \quad (7.13)$$

The first term on the right-hand-side of Equation (7.13) is due to addition/subtraction, the second term is due to the shift operation and the third term is due to the correlation between the two sources of errors [118].

In order that the analysis can be simplified, it is assumed that the errors arising from the two sources are independent, that is, K=0. Thus, at the first stage, the variance of the error is given by

$$\text{var}(X_1) = 2q^2 + e^2 \tag{7.14}$$

in which the first term is due to addition/subtraction of input numbers each with error variance q^2 and the second term is due to the shift operation. Similarly, at the second stage,

$$\begin{aligned}\text{var}(X_2) &= 2\text{var}(X_1) + 2^2e^2 \\ &= 2^2q^2 + 2e^2(1+2)\end{aligned} \tag{7.15}$$

and, by induction, the error variance at the m^{th} stage is given by

$$\begin{aligned}\text{var}(X_m) &= 2^mq^2 + 2^{(m-1)}e^2(1+2+\ldots+2^{(m-1)}) \\ &= 2^mq^2 + 2^{(m-1)}e^2(2^m-1)\end{aligned} \tag{7.16}$$

For a transform of size 2^M, the ratio, σ_M, of the root-mean-square error to the root-mean-square value of the transform coefficients, is given by

$$\begin{aligned}\sigma_M &= \sqrt{\text{var}(X_M)/(2^MP)} \\ &= \sqrt{\{[2q^2 + (2^M - 1)e^2]/2P\}}\end{aligned} \tag{7.17}$$

From Equation (7.17), it can be seen that for large M, the normalized standard deviation σ_M increases approximately in proportion to $\sqrt{2^M}$.

7.3 SIMULATION RESULTS

A simulation of block-floating-point WHT was carried out on a CDC computer which gives an accuracy of 28 digits when operated in the double precision mode. The errors d(n),n=0,...,N-1 were determined as the difference between the block-floating-point numbers and the numbers obtained using double-precision computation.

In the simulation, 2^{10} data samples were used, with the data being obtained from a) a uniform distribution with random numbers between -1 and 1; b) a uniform distribution with random numbers between 0 and 1; c) a Gaussian distribution N(0,1/4) with random numbers limited to between -1 and 1, and d) a segment of voiced speech signal. The r.m.s. values of the data are a) 0.574, b) 0.584, c) 0.25 and d) 0.283 respectively. For each type of data and each transform, and for a given transform size, the value

$$R_e = \sqrt{[\frac{1}{N} \sum_{n=0}^{N-1} d^2(n)] / \sum_{n=0}^{N-1} x^2(n)} \qquad (7.18)$$

was evaluated when using Methods 1 and 2. In the case of a transform of size 2^M, a total of $2^{(10-M)}$ transforms were evaluated and the average R_e was determined. In the simulation, the errors associated with the input data were set to be zero.

The results obtained with both the Walsh-ordered and Hadamard-ordered transforms are shown in Figs.7.1 to 7.4. Fig.7.1 relates to data drawn from the uniform distribution with numbers between -1 and 1; Fig.7.2 to the data from the uniform distribution with values between 0 and 1; Fig.7.3 to the data from the Gaussian distribution and Fig.7.4 to the speech signal. In these figures, R_e is

plotted as a function of the transform size. The maximum error bounds and the estimated normalized standard deviations, σ_M, and $2\sigma_M$ are also plotted on the figures.

From these figures, the following observations can be made:

(i) As expected, for a given transform size and a given input, the R_e is very much lower than the maximum error bound.

(ii) The R_e's associated with Method 1 lie above the estimated σ_M. This discrepancy can be attributed to the simplifying assumption that the random variable associated with error resulting from addition/subtraction and the random variable associated with the error resulting from the shift operation are independent. However, as can be seen from the figures, the R_e's lie below $2\sigma_M$.

(iii) For a given transform size and a given input, the R_e associated with Method 2 is lower than that associated with Method 1, except in the case shown in Fig.7.2 where for N=1024, they are the same. This shows that in general, overflow does not occur at every stage and hence Method 2 is a much better technique than Method 1.

(iv) Although theoretically, the Walsh-ordered WHT and the Hadamard-ordered WHT contain the same set of transform values, the ordering of the elements does affect the accuracy of the final results. However, under identical conditions, the order of magnitude of errors are the same for the two transforms.

A further comparison has been made between the errors resulting when the block-floating-point arithmetic is used in obtaining the fast Fourier transform (FFT) and the FWHT. In the case of the FFT, the 'decimation-in-time' algorithm was used and the method of scaling by testing for an overflow as described by Welch [119] was adopted. The

results of the comparisons are shown in Figs.7.5 and 7.6 for L=10 and L=16. The input data used for Figs.7.5 and 7.6 were the same as those used for Figs.7.1 and 7.3 respectively. From these figures, it is clear that for a given word-length and a given transform size, the error associated with the FFT is greater than that associated with the FWHT. The figures also show, as expected, that the errors in both cases decrease with increasing word-length.

7.4 CONCLUSIONS

Theoretical analysis and simulation results show that the error introduced through the use of the block-floating-point fast Walsh-Hadamard transform increases with the size of the transform and decreases with increasing word-length. Further, for a given word-length and a large transform size, the error resulting from the fast Fourier transform is at least an order of magnitude larger than the error introduced by the fast Walsh-Hadamard transform using the same input data.

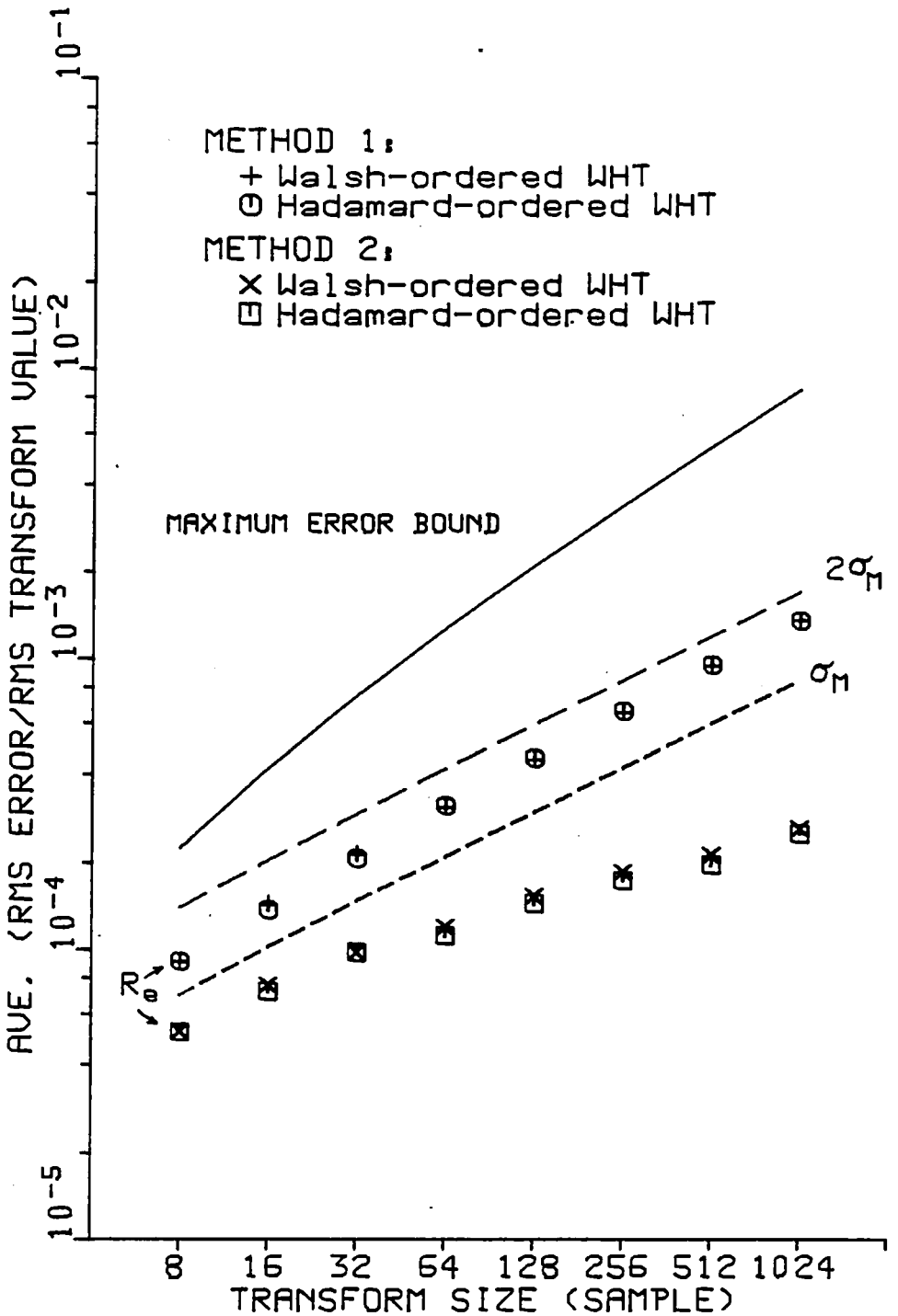


Fig.7.1 Experimental error results : random numbers from a uniform distribution; -1<random numbers<1; word-length = 16 bits.

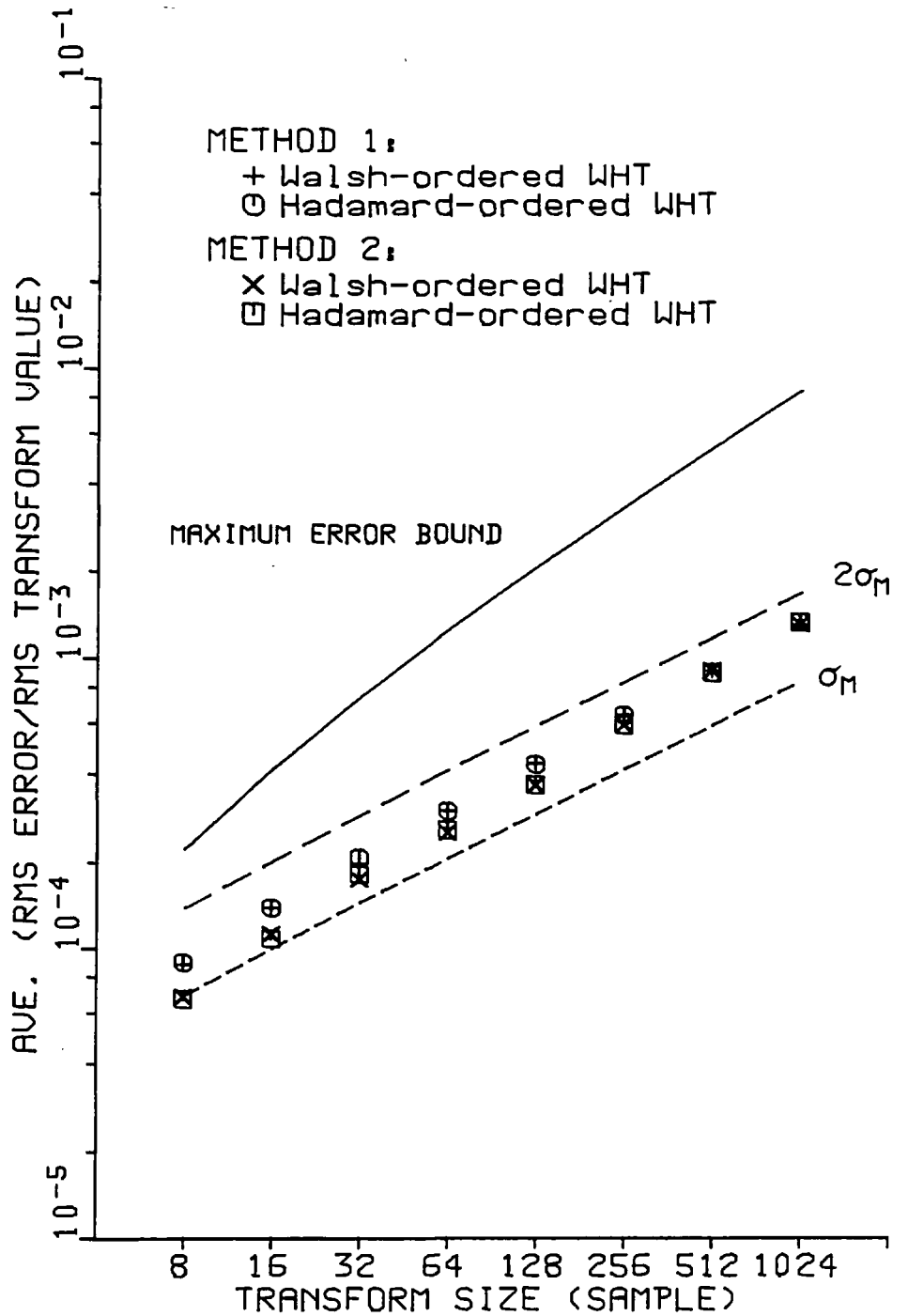


Fig.7.2 Experimental error results: random numbers from a uniform distribution; $0 \leqslant$ random numbers < 1; word-length $=$ 16 bits.

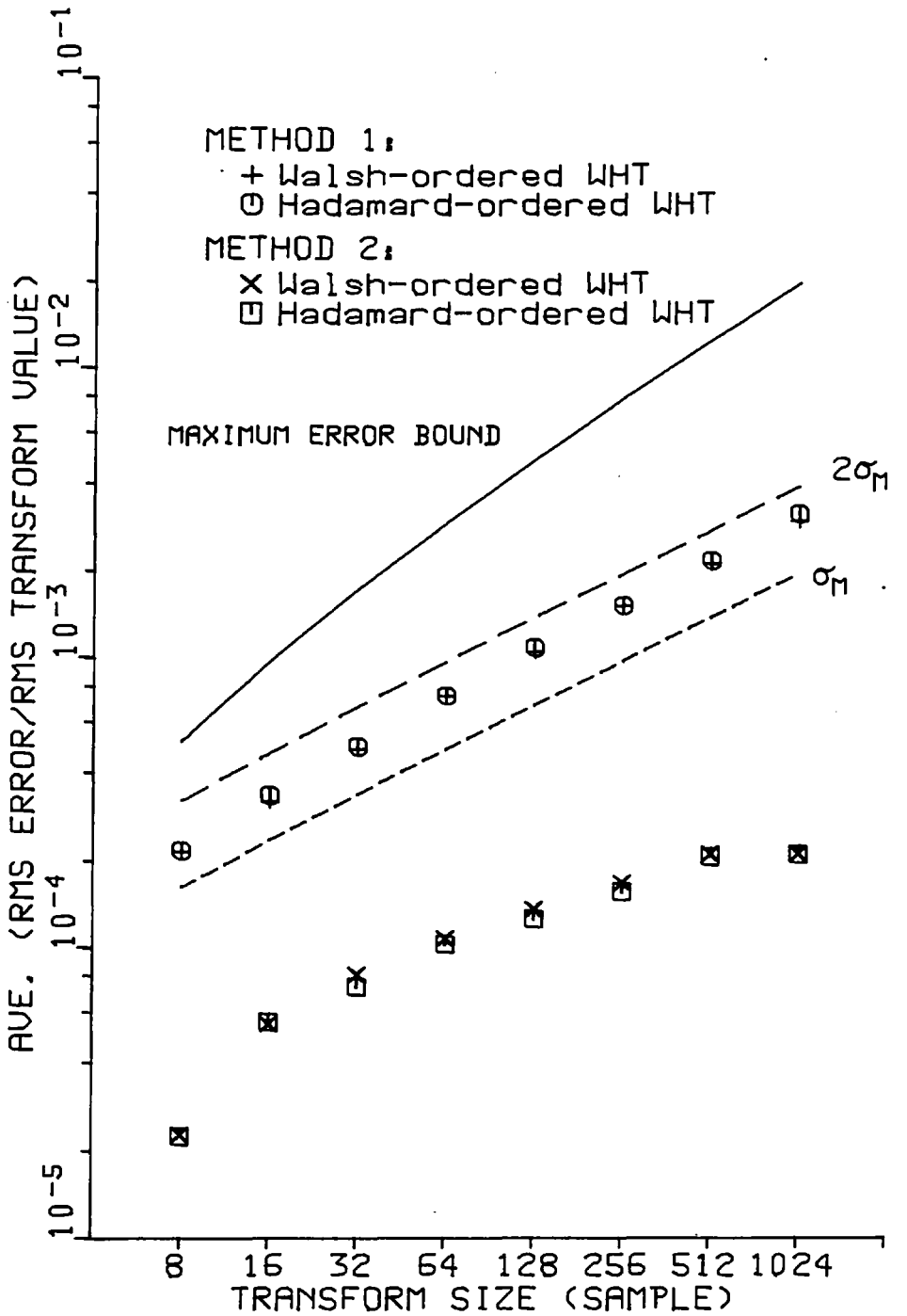


Fig.7.3 Experimental error results: random numbers from a Gaussian distribution N(0.,0.25); -1<random numbers<1; word-length = 16 bits.

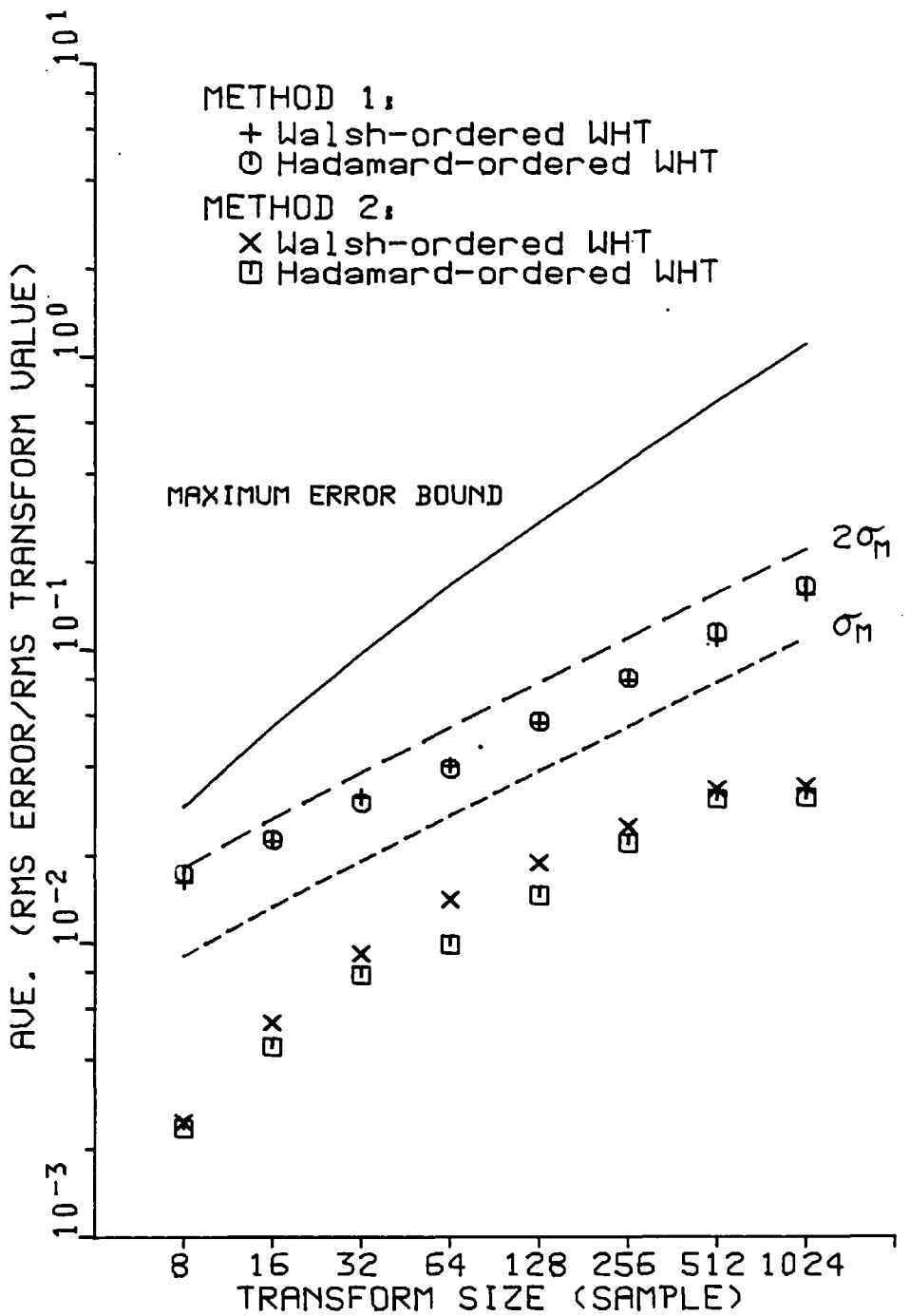


Fig.7.4 Normalized errors of block-floating-point WHT as a function of block size using data from a segment of speech signals. Word-length = 10 bits.

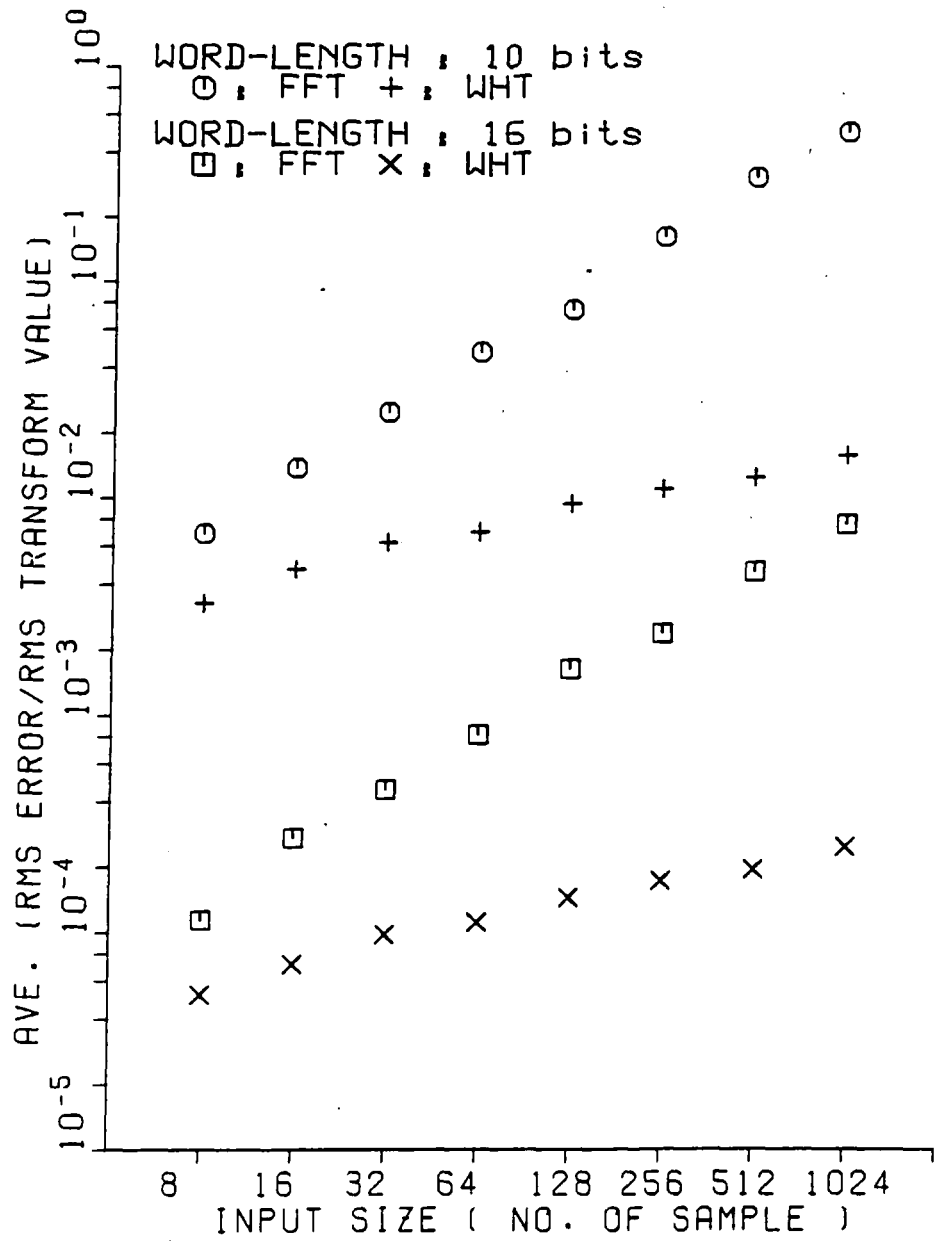


Fig.7.5 Comparisons of errors resulting from block-floating-point WHT and those from block-floating-point FFT using random numbers from uniform distribution as input.

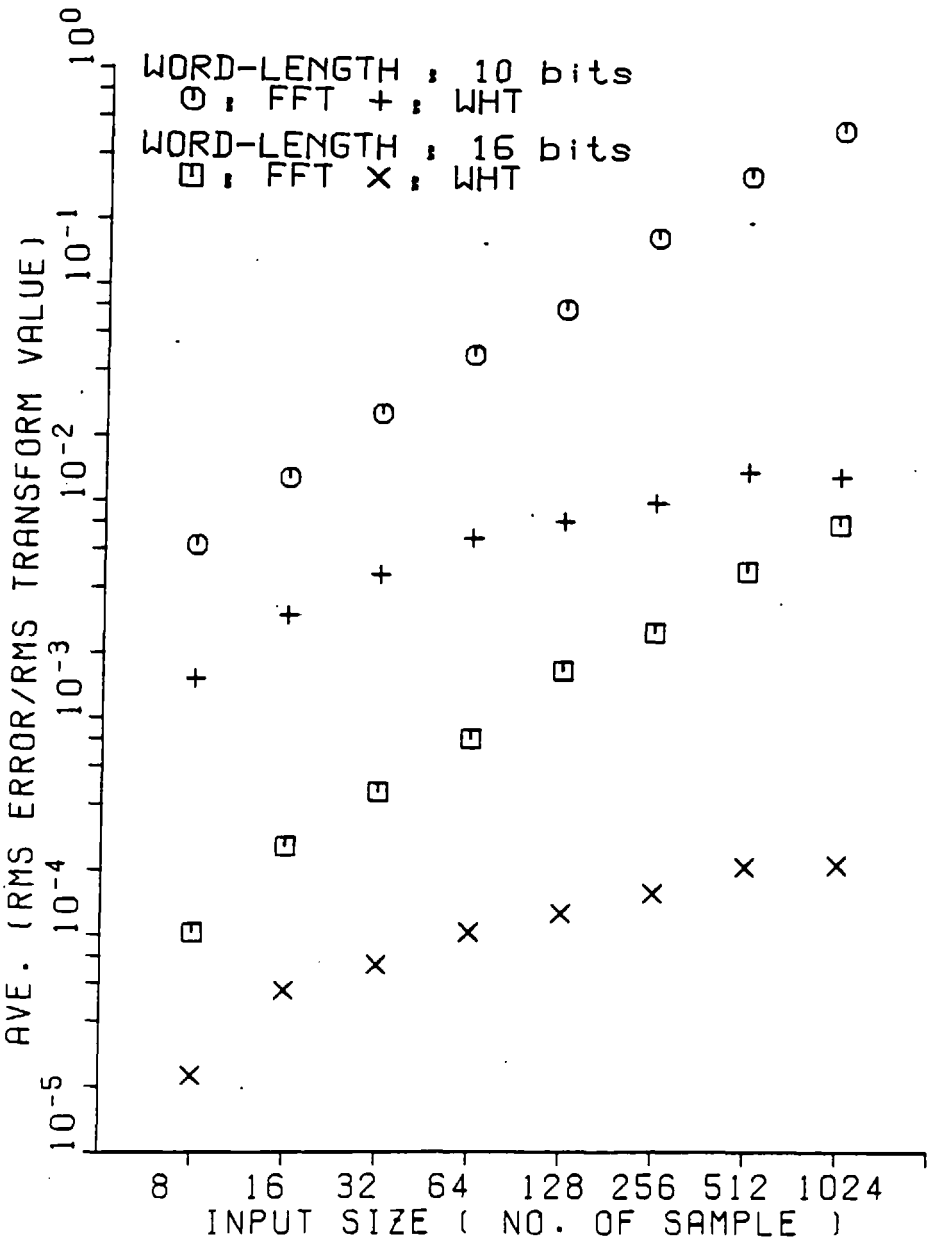


Fig.7.6 Comparisons of errors resulting from block-floating-point WHT and those from block-floating-point FFT using random numbers from Gaussian distribution as input.

CHAPTER 8

APPLICATIONS OF SUB-BAND TECHNIQUE TO THE CLASSIFICATION OF VOICED, UNVOICED AND SILENCE SEGMENTS OF SPEECH

CHAPTER 8

APPLICATIONS OF SUB-BAND TECHNIQUE TO THE CLASSIFICATION OF VOICED, UNVOICED AND SILENCE SEGMENTS OF SPEECH

8.1 INTRODUCTION

Voiced speech is produced when air is forced through the glottis thereby forcing the vocal cords into vibration. Unvoiced speech, on the other hand, is generated by the turbulent flow of air resulting from the presence of a constriction in the vocal tract. Human speech consists of a concatenation of voiced speech, unvoiced speech and intervals of silence. The transition from one class to another, however, is not abrupt; and one of the difficulties in the voiced-unvoiced-silence decision is due to this gentle transition.

The reliable classification of segments of a speech signal into voiced speech, unvoiced speech and silence is an important and difficult problem in speech analysis. A good classifier for voiced sounds, unvoiced sounds and silence should be able to perform the task accurately for a wide class of speech material. However, the task is made more difficult by the large dynamic range of speech signals within the speech of a single talker, variation in the recording conditions and the dissimilarities among different talkers. The problem is further compounded in the presence of noise.

Earlier methods for discriminating speech sounds from silent intervals were based on the energy level of signals, with voiced sounds and unvoiced sounds being differentiated according to the number of zero crossings [120], or the presence or absence of a distinct pitch [121]. Pitch detection, however, is itself a difficult problem.

A pattern classification approach was investigated by Atal and Rabiner [122] in which five signal parameters were used. These included the energy, zero crossing rate, autocorrelation coefficient at unit sample delay, first predictor coefficient and energy of the prediction error. In order to make the decisions, these parameters were combined into a non-Euclidean distance metric. A similar approach to classification was proposed by Rabiner and Sambur [123] using a nonlinear combination of Linear Predictive Coding distance measure and an energy distance. One disadvantage of the above methods is the need for a set of data to be used in training the algorithms so that they can be adjusted to suit the existing conditions.

A method aimed specifically at making the three-way classification from speech signals that have been contaminated by wide-band noise was proposed by Cox and Timothy [124] using non-parametric rank-order statistics. The method has the merit that a training set of data is not required, the thresholds being calculated from theory. However, a large number of comparisons has to be made in order to obtain the required statistics.

In applications where discrimination of voiced speech and nonvoiced speech (either unvoiced sounds or silence) only is required,

the problem is much simplified. Many methods [125-127] have been investigated, some of which are modifications of the three-way classification techniques. The error rate of misclassification for the two-way decision is in general very much lower than that for the three-way classification.

In this chapter, a method for three-way classification is put forward which makes use of the ratio of energy of the sub-band signals of speech.

8.2 DESCRIPTION OF THE ALGORITHM

The proposed method makes use of the different spectral characteristics of the three classes of speech signals. Any noise involved is considered to be white Gaussian noise.

It is known that, in general, the spectral density of voiced speech is higher at low frequencies than at high frequencies; whereas the spectral density of unvoiced speech is higher at high frequencies than at low frequencies; and white noise, or silence, has a relatively flat spectrum.

The proposed method is designed to take advantage of these facts. The signal is filtered into four contiguous bands B1, B2, B3 and B4. During a fixed time segment, the mean-square values of each of the four sub-band signals is determined. Let these values be denoted by S_1, S_2, S_3 and S_4. The largest, $S_{max} = \max\{S_1, S_2, S_3, S_4\}$ and the smallest $S_{min} = \min\{S_1, S_2, S_3, S_4\}$ are then sorted and the ratio $R = S_{max}/S_{min}$ is computed. If S_{max} belongs to the lower two bands, that is, if $S_{max} \in \{S_1, S_2\}$, and in addition $R > T_v$, where T_v is a

threshold, then the segment is classified as voiced; otherwise it is classified as silence. If S_{max} belongs to the upper two frequency bands, that is, $S_{max} \in \{S_3, S_4\}$, and in addition, $R > T_u$, where T_u is a threshold which is different from T_v, then the segment is classified as unvoiced; otherwise it is classified as silence. The thresholds T_v and T_u can be determined from the probability distributions of R for the three classes.

It should be noted that this method makes use of the ratio of the sub-band energy and disregards the absolute energy level of the original signal. Mathematically, if the energy level of the original signal is S_o and $S_i = k_i S_o$, $i=1,\ldots,4$, where $k_1,\ldots,k_4$ are constants, then $R = k_m S_o / k_n S_o = k_m / k_n$ with $m,n \in \{1,2,3,4\}$. Thus, R is independent of S_o and so long as the signal amplitude is not distorted, the values of the ratio, and hence the accuracy of the method, are not affected by the dynamic range of the signal.

8.3 DETERMINATION OF THRESHOLDS

In order to carry out the classification using the proposed algorithm, the thresholds T_u and T_v have to be determined first. For this purpose, two utterances each from a male and a female speaker, with the utterances being recorded through a microphone, were used. By inspection of the waveform, segments of 10 ms of speech signals were manually classified as voiced speech, unvoiced speech and silence. Using only those segments which were unambiguous, the distributions of the sub-band energy ratio for the three classes were then determined. From these distributions, the values of T_u and T_v which result in a minimum misclassification error were then chosen as the thresholds.

8.4 EXPERIMENTS AND RESULTS

A computer simulation of the technique described in Section 8.2 has been carried out using speech signals recorded with different signal-to-noise ratios. The speech signals were lowpass filtered to 4.5 kHz, sampled at 10 ksample/s and quantized using a 12 bit analog-to-digital converter. The first test signal consists of about 9 s of male speech and 5 s of female speech recorded with a signal-to-noise ratio of approximately 40 dB. The second test signal consists of about 5 s of female speech and 3 s of male speech, with a signal-to-noise ratio of approximately 25 dB. Different speakers were used to obtain the two test signals, and different speech material were also used in the two cases. The test signals used in the simulation were different from that used in the determination of the thresholds T_u and T_v.

The sampled signals were filtered into four sub-bands covering the bandwidth of 200-3200 Hz. Band 1 covered the range 200 Hz to 700 Hz, band 2 the range 700 Hz to 1310 Hz, band 3 the range 1310 Hz to 2020 Hz and band 4 the range 2020 Hz to 3200 Hz. The widths of the sub-bands were chosen so that the signals in each sub-band made equal contribution to the articulation index. The four filters used were designed as 181-tap FIR digital filters with a minimum stopband attenuation of 40 dB and a passband ripple of less than 0.5 dB.

The signals within each sub-band were grouped into sets of samples, and experiments were performed using sets of 50, 100 and 150 samples, which corresponded to segments of 5 ms, 10 ms and 15 ms respectively. For each segment, or group of samples, the average

energy in each sub-band was computed as the mean value of the sum of squares of the sample values within the segment.

In order to avoid the problem associated with the division by small numbers when deriving values of the ratio R, a limit of 1 was used for average energy values when the estimate of the average value was found to be less than 1. The thresholds, T_u and T_v, determined as described in Section 8.3, were 5 and 15 respectively, and these values were used in the experiments.

A 'reference' classification against which automatic classification was judged to be correct, or otherwise, was obtained by manual inspection of the signal waveform.

An example is given in Figs.8.1 and 8.2 to illustrate the variation in the energy content within segments of the sub-band signals. The waveform of a typical speech signal of 1.2 s duration is shown in Fig.8.1. The variation of the energy content within segments of 5 ms of the corresponding sub-band signals is shown in Fig.8.2.

The results obtained from the simulation of the proposed classification scheme are set out in the following tables. Table 8.1 shows the results obtained when using test signal 1 and Table 8.2 the results obtained using test signal 2. The tables show the way in which the performance of the proposed scheme varied as length of the segments used in the energy estimate was increased.

Table 8.1 Confusion Matrices for Test Signal 1 (SNR=40dB)

Segment Length	5 ms.			10 ms.			15 ms.		
Classified as / Actual	V	UV	S	V	UV	S	V	UV	S
Voiced	1692	4	72	851	3	26	573	2	16
(in percent)	95.7	0.2	4.1	96.7	0.3	3.0	97.0	0.3	2.7
Unvoiced	1	82	15	1	37	1	0	27	1
(in percent)	1.0	83.7	15.3	2.6	94.8	2.6	0.0	96.4	3.6
Silence	9	11	975	4	5	502	3	3	328
(in percent)	0.9	1.1	98.0	0.8	1.0	98.2	0.9	0.9	98.2

Table 8.2 Confusion Matrices for Test Signal 2 (SNR=25dB)

Segment Length	5 ms.			10 ms.			15 ms.		
Classified as / Actual	V	UV	S	V	UV	S	V	UV	S
Voiced	1360	4	24	682	3	6	457	2	4
(in percent)	98.0	0.3	1.7	98.7	0.4	0.9	98.8	0.4	0.8
Unvoiced	1	78	5	0	40	1	0	24	1
(in percent)	1.2	92.9	5.9	0.0	97.5	2.5	0.0	96.0	4.0
Silence	15	10	113	2	3	68	1	1	47
(in percent)	10.9	7.2	81.9	2.7	4.1	93.2	2.0	2.0	96.0

It can be seen from the tables that higher percentages of misclassification occured when 5 ms segments were used, and that improved classifications resulted from the use of either 10 ms or 15 ms segments. It is thus clear that with the proposed scheme, as with other existing schemes, the need to make rapid classification decisions has to be weighed against the desired accuracy of decision making.

The distributions of the sub-band energy ratios for the two test signals are shown in Figs.8.3(a) and 8.3(b). In these figures, the segments used are of length 40 ms and the 'silence' class as shown is classified as uv-type when S_{max} belongs to bands B3 or B4 and as v-type when S_{max} belongs to bands B1 or B2.

To highlight the value of the method, the distributions of the energy in 10 ms segments of the test signals were determined for the four different classes of speech signals and these are given in Figs.8.4(a) and 8.4(b). For test signal 1, the three classes of speech signals, namely, voiced speech, unvoiced speech and silence are easily separable based on energy levels, but for test signal 2 the unvoiced and 'silence' segments are almost indistinguishable if classification is based on energy level alone.

8.5 COMPUTATIONAL LOAD

Let $N=N_1+N_2+N_3+N_4$ where N_1, N_2, N_3 and N_4 denote the numbers of samples within a given segment of the signal for the respective sub-bands B1, B2, B3 and B4. For the technique described by Cox and Timothy [124], it is estimated that on the order of N^2 operations are required to perform the ranking operation. For signals sampled at 10 ksample/s and using segments of 10 ms duration, N=400, and approximately 160,000 operations are to be performed.

The method proposed in this chapter requires N squaring operations, N additions and a few comparisons. Thus the number of operations required is considerably less than that using the technique of Cox and Timothy.

8.6 CONCLUSIONS

Experimental results indicate that with the proposed method of speech classification, a very high percentage (in excess of 95%) of correct classification into voiced speech, unvoiced speech and silence can be achieved and that the level of performance can be maintained for signals with high dynamic range which may be contaminated with white noise.

The error rate compares favorably with the average misclassification rate of about 6% using the method proposed by Cox and Timothy [124]. For the method described by Atal et al. [122], an overall error recognition rate of 6.4% was reported for signals that had been transmitted over telephone lines; and for the technique investigated by Rabiner and Sambur [123], an overall error rate of 5% was reported.

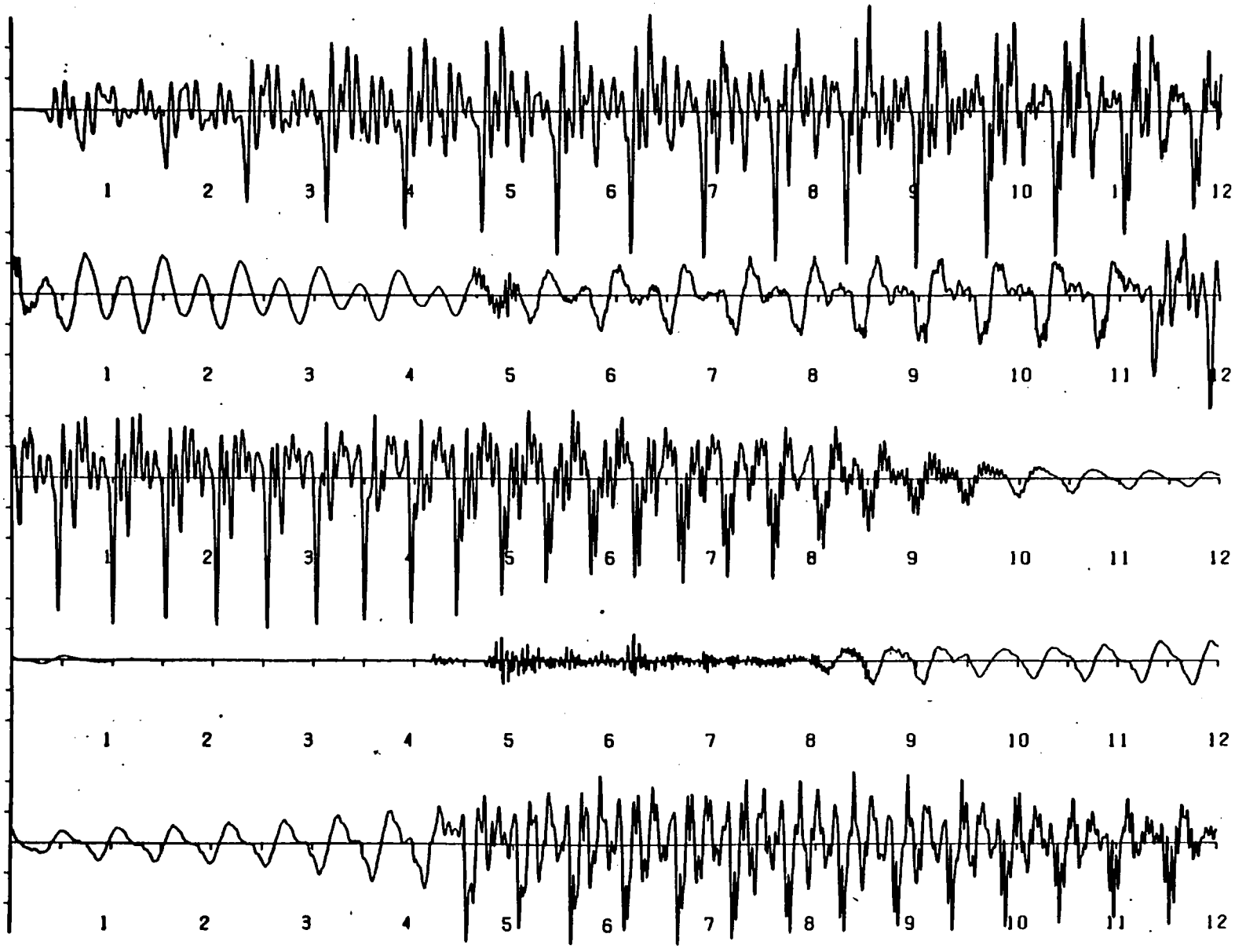


Fig.8.1 Waveform of a segment of speech signal.

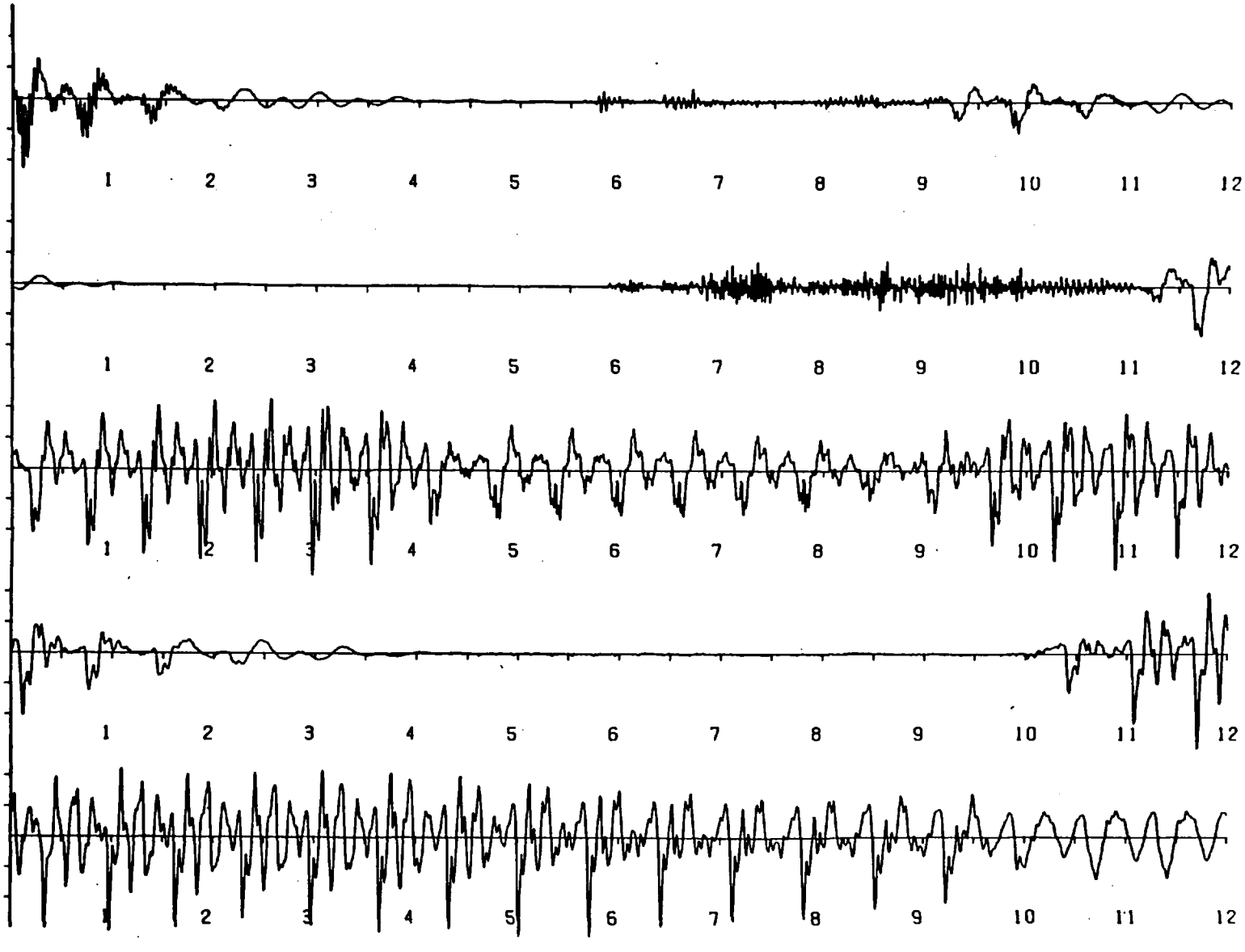


Fig.8.1 (continued).

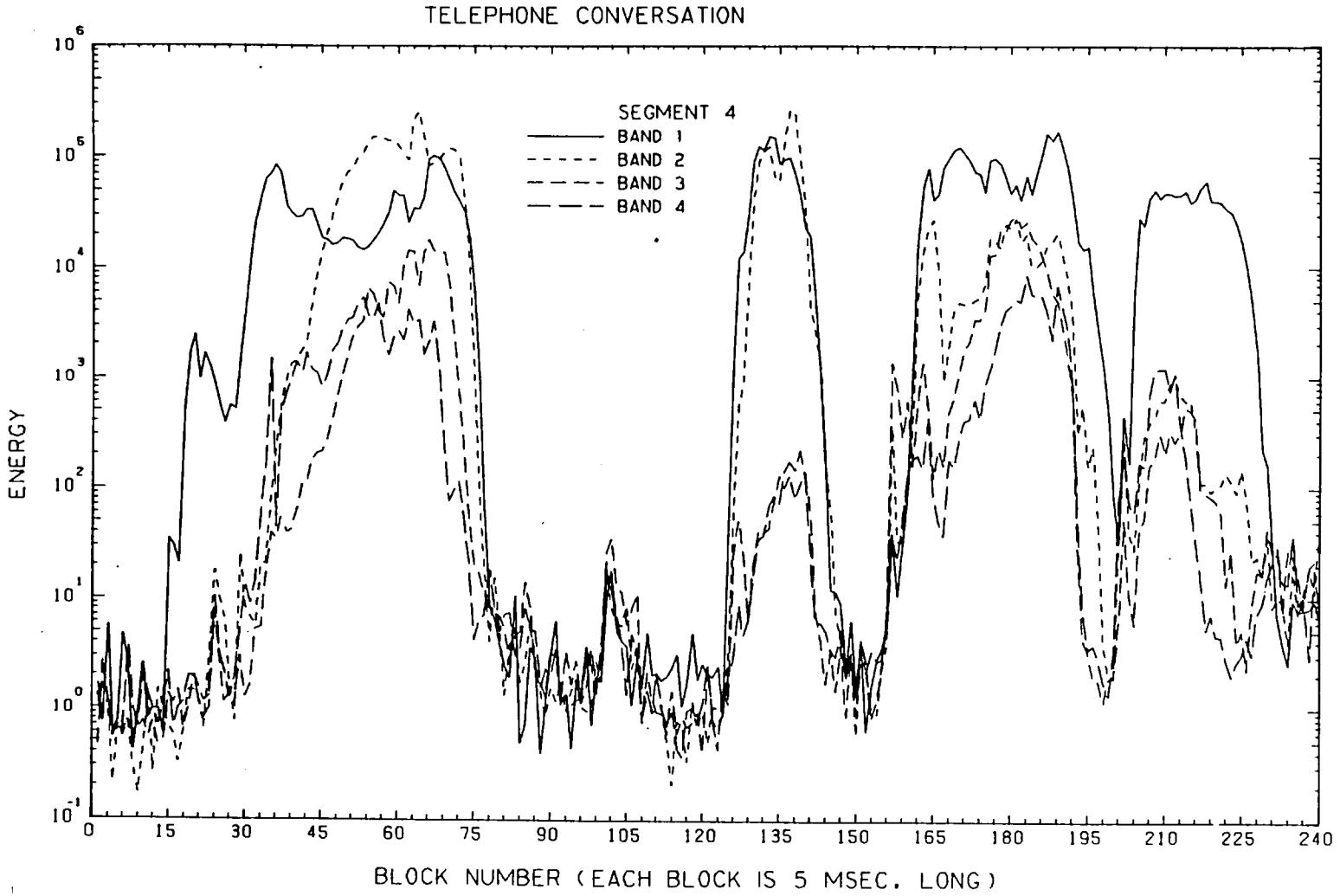
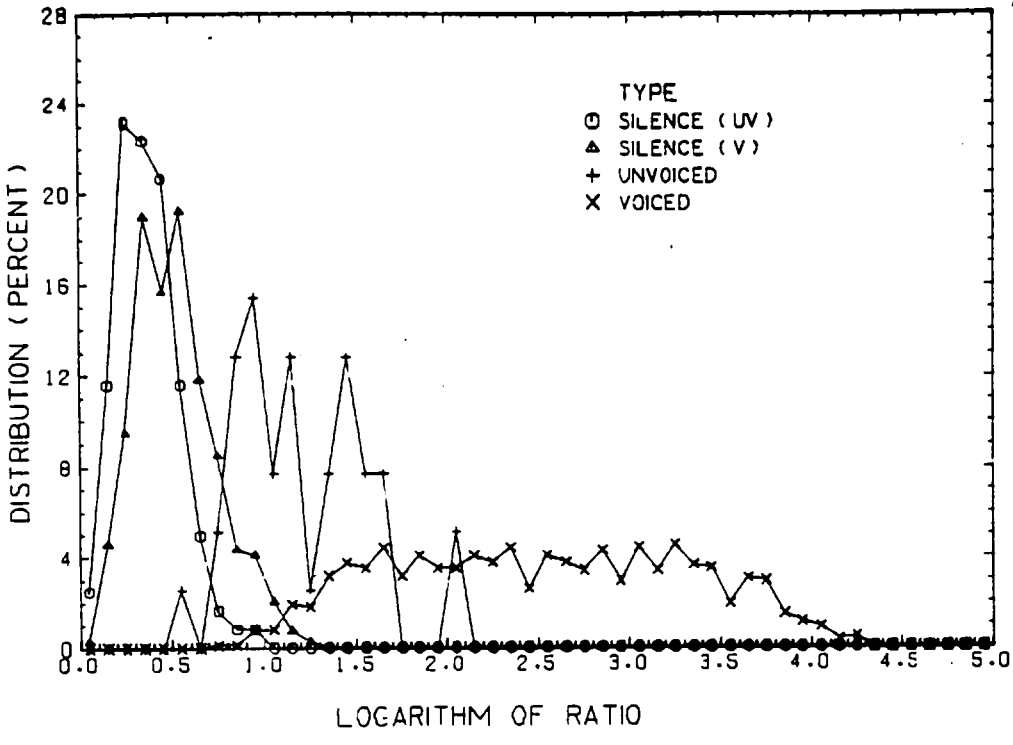
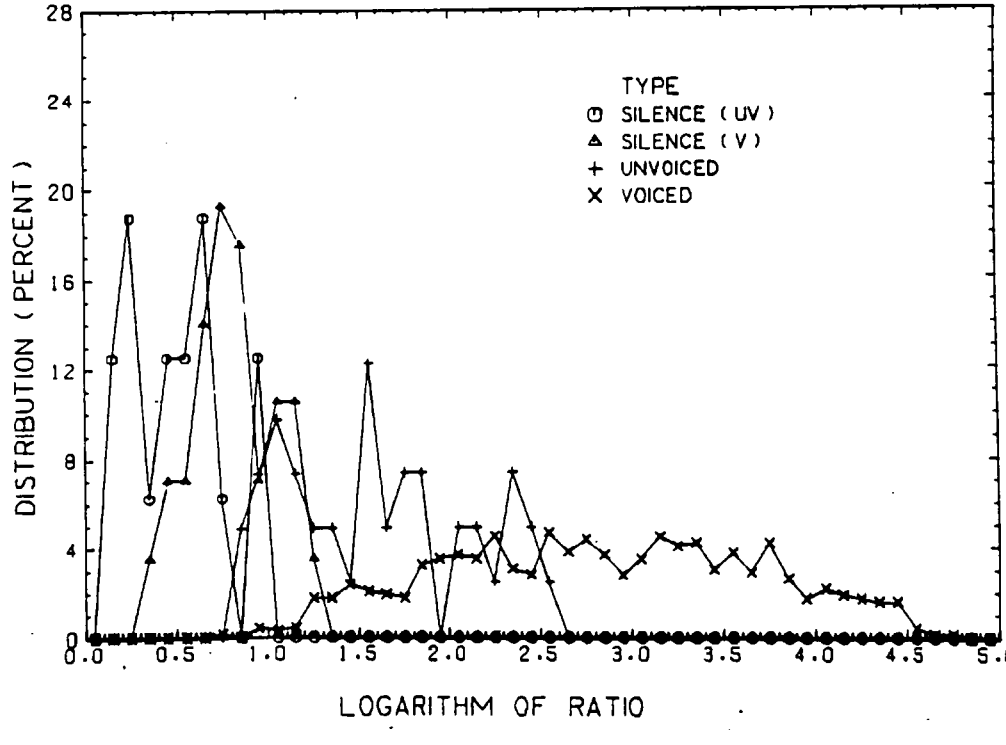


Fig.8.2 Variation of the energy content within segments of sub-band speech signals. Segment length = 5 ms.

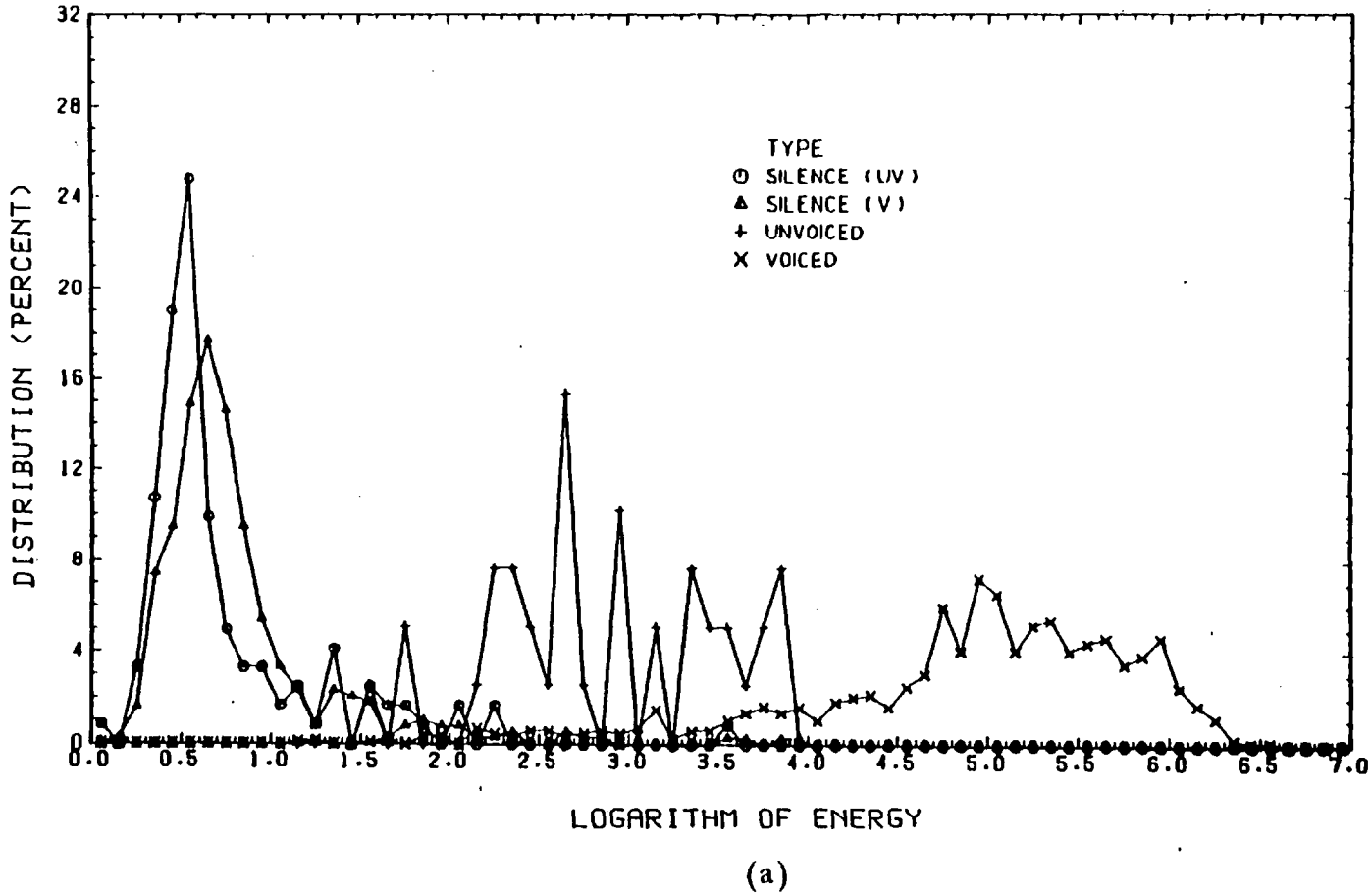


(a)



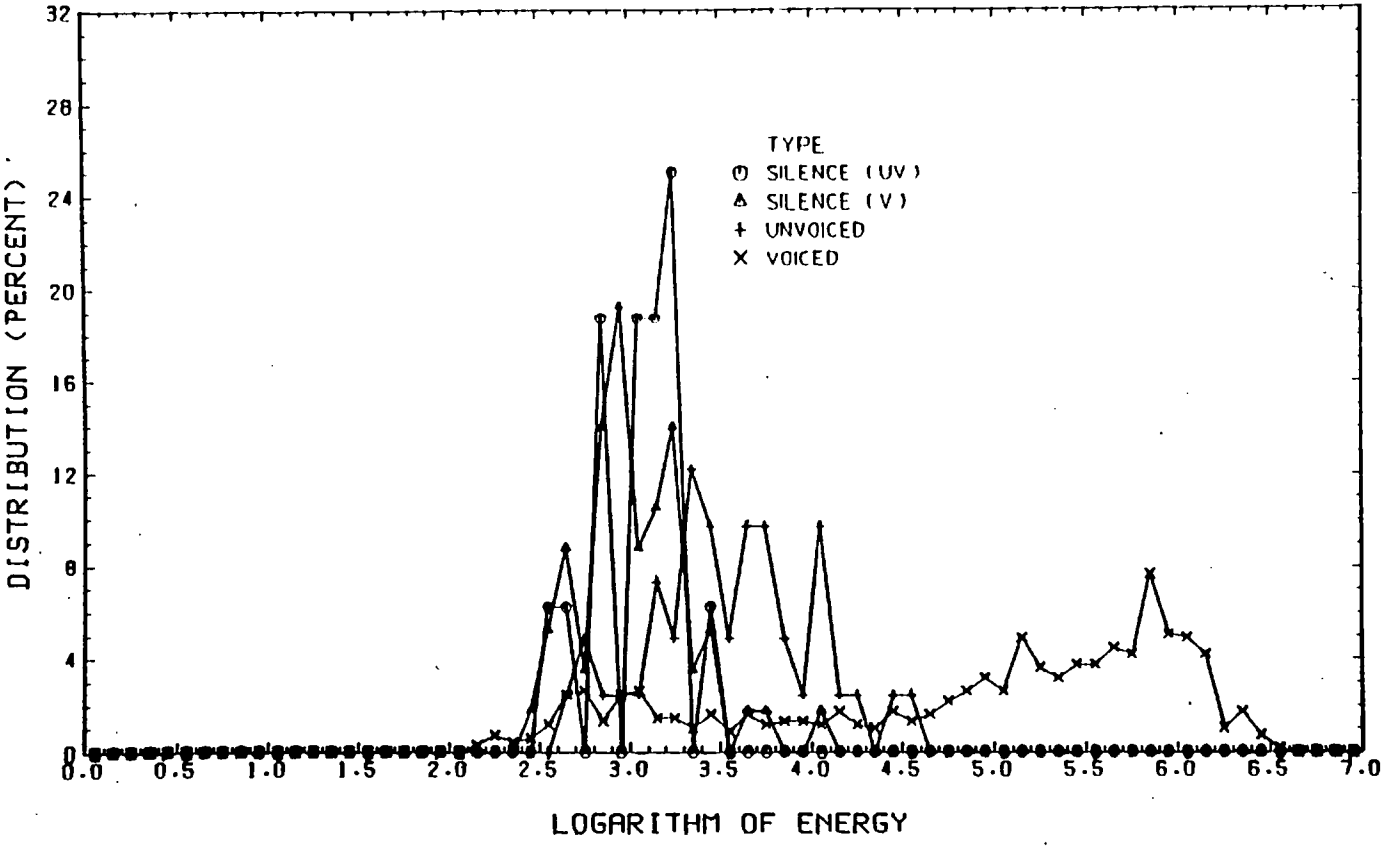
(b)

Fig.8.3 Distribution of logarithm of energy ratio. Segment length = 10 ms. (a) Test signal 1; (b) Test signal 2.



(a)

Fig.8.4 Distribution of logarithm of energy within segments of speech signals. Segment length = 10 ms. (a) Test signal 1; (b) Test signal 2.



(b)

Fig.8.4 (continued).

CHAPTER 9

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

CHAPTER 9

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

9.1 CONCLUSIONS

Hybrid frequency-domain coding (HFDC), a technique that combines the salient features of sub-band coding and transform coding, has been investigated for speech coding at transmission bit rates of 7.2 kbit/s and 9.6 kbit/s.

In hybrid frequency-domain coding, the full-band input signal is filtered into four sub-bands through a series of band splitting operations using quadrature mirror filters. The two lower sub-band signals are coded using adaptive PCM and the upper two bands are coded using an adaptive Walsh-Hadamard transform.

Two versions of hybrid frequency-domain coding scheme have been examined. In one version, referred to as fixed-rate HFDC, blocks of speech signal are coded using the same number of bits. In another version, referred to as variable-rate HFDC, the number of bits allocated to encode the block is made to depend on the characteristics of the block.

Subjective tests have been carried out which show that at 7.2 kbit/s and 9.6 kbit/s, the quality of speech reproduced by the fixed-rate hybrid coding scheme is comparable respectively with the

quality reproduced by 19 kbit/s and 23 kbit/s continuously variable slope delta modulation (CVSD). Based on subjective comparison of the quality of speech produced, Crochiere et al have shown [57] that sub-band coders operating at 7.2 kbit/s and 9.6 kbit/s are approximately equivalent to adaptive delta modulation systems operating at bit rates of 18 kbit/s and 19 kbit/s respectively.

The computation of the Walsh-Hadamard transform involves only subtractions and additions and hence the speed of computation is very much faster than that of the discrete cosine transform. The speed of computation required for real-time application using HFDC is well within the reach of present day technology. Furthermore, by using quadrature mirror filters for band splitting, short filters can be used. The complexity of the hybrid frequency-domain coder is thus estimated to be between that of the sub-band coder and the adaptive transform coder.

The improvement in the quality of speech produced by HFDC over that produced by sub-band coding at the same bit rate can be attributed to the following reasons: (i) Through coding in sub-bands, quantization noise is confined to the sub-bands; (ii) The sub-band signals can be coded with different numbers of bits based on a perceptual criterion; (iii) The residual correlation of the sub-band signals can be exploited using transform coding. Transform coding also enables dynamic bit allocation and spectral noise shaping to be applied; (iv) The use of quadrature mirror filters in band-splitting together with dynamic bit allocation for coding the spectral components in the sub-bands reduces the aliasing noise due to downsampling and quantization of the sub-band signals.

In Chapter 4, a new method of variable-bit-rate coding was investigated. In the new method, different time segments of speech signals are coded with different numbers of bits. Segments of speech signals which are more significant are coded with more than the average number of bits while less than the averge number of bits are assigned to code those segments such as silent intervals and unvoiced segments which carry less information.

By continuously checking the average rate of bits generated and reverting to average rate of coding when it is anticipated that either the upper or lower limit is exceeded, the maximum time delays and buffer storages required can be predetermined.

The new method was applied to the hybrid frequency-domain coding and results of subjective tests carried out on speech produced using variable-rate hybrid frequency-domain coding show that some improvement in speech quality over the fixed-rate hybrid frequency-domain coding can be achieved with an additional communication delay as low as 200 ms.

The performance of any system depends very much on the component parts. One of the major components of sub-band coding, the quadrature mirror filters, were examined in detail in Chapters 5 and 6.

A new design method of non-recursive quadrature mirror filters was described in Chapter 5. It was shown that by reducing the non-linear constraint imposed on the design of quadrature mirror filters to a linear form, a solution very close to the optimal one can then be obtained using linear programming techniques. The new method provides for the control of the amount of ripple in the overall system

response as well as the control of other design parameters such as stopband attenuation and width of transition band. The solution can be obtained without human intervention unlike other techniques such as the direct search approach described by Johnston [112] where Hookes and Jeaves algorithm is applied.

In the simulation and implementation of systems using digital techniques, invariably, the word-length to be used for data storage and computation requires careful consideration. Problems of implementation of the split-band system with quadrature mirror filters and the effects of finite-word-length on the performance of such system were analysed in Chapter 6.

The analysis shows that complete cancellation of frequency aliasing is not affected by quantization of the filter coefficients or the rounding of arithmetic operations performed within the filters. However, complete cancellation of frequency aliasing cannot be achieved when the filtered and downsampled sub-band signals are quantized for compression encoding.

Useful formulae are derived for the bounds and variances of errors due to coefficient quantization, arithmetic roundoff and quantization for compression encoding. These formulae serve as useful guidelines for the selection of word-length in the simulation and implementation of systems using quadrature mirror filters.

The effects of finite-word-length on the computation of the Walsh-Hadamard transform were analysed in Chapter 7. The analysis shows that using the same word-length for data storage and arithmetic operations, the accuracy attainable with the fast Walsh-Hadamard

transform is much higher than that attainable with the fast Fourier transform.

In systems where classification of voiced speech, unvoiced speech and silence is involved, erroneous classification can result in degradation of system performance. The automatic classification of speech signals into voiced speech, unvoiced speech and silence is an important and a difficult problem. A new approach based on ratios of energy of sub-band signals was described in Chapter 8. The new method is particularly useful in dealing with signals contaminated with wideband noise. The results of simulation show that very high probability of correct classification can be achieved using this method.

9.2 SUGGESTIONS FOR FURTHER RESEARCH

In order to apply the principles of adaptive bit allocation, information about the spectral envelope must be coded and transmitted. For the method described in Chapter 3, the logarithm of the root-mean-square value of the block of transform coefficients is coded with 5 bits. The quantized average value is then obtained by taking the anti-log. The block of transform coefficients is divided into five groups and the ratios of the group standard deviations to the quantized average value of the block are then coded with 2 bits each. Using this method, the number of bits required for coding the side-information is approximately 750 bit/s. For coding at low bit rates, this represents a significant proportion of the transmission bit rate. As the samples of the sub-band signals are in general not highly correlated, the spectral patterns are relatively simple.

Methods of describing the spectral envelopes such as pattern classification which would require the use of relatively few bits should be investigated in order that more bits would be available for coding the transform coefficients proper.

For those segments of voiced speech that show quasi-periodicity, the samples of the signal at pitch intervals are highly correlated in the time domain. In the frequency domain, this is reflected in the spectral fine structure manifested as narrow lobes spaced at the harmonics of the pitch frequency. This spectral fine structure is exploited in the vocoder-driven adaptive transform coding [6,84] by superimposing the pitch spectrum on the envelope of the short time frequency spectrum to obtain a more accurate estimate of the actual spectrum. The pitch spectrum is obtained through the Fourier transform of a windowed impulse train with exponentially decaying amplitude and spaced at pitch period apart. Mathematically, the pitch spectrum is obtained by Fourier transform of the sequence

$$p(n) = h(n) \sum_{m=0}^{\infty} G^m \delta(n-m\ell) \qquad n=0,1,\ldots,N-1 \qquad (9.1)$$

where h(n) is the window function, G is the pitch gain defined as the ratio of the autocorrelation function of the signal at the pitch period to its value at the origin, $\delta(*)$ is the Kronecker delta, ℓ is the pitch period and N is the transform size.

The possibility of incorporating a pitch model in the Walsh-Hadamard transform domain will now be discussed.

Fig.9.1 shows the waveform of a typical voiced segment of speech signal. The variation of the absolute values of the Walsh-Hadamard

transform coefficients of the segment is shown in Fig.9.2 in ascending Walsh order number. Figs.9.3 and 9.4 respectively show the values of the cal and sal components of the same segment in sequency order. The power spectrum of the Walsh-ordered WHT is given in Fig.9.5. It can be seen that although there is an appearance of spectral fine structure similar to that of the corresponding power spectrum of the discrete Fourier transform shown in Fig.9.6, the lobes are not as prominent and are not necessarily spaced at the pitch harmonics as in the case of the frequency spectrum. Furthermore, it has been shown [74] that while the amplitudes of the discrete Fourier transform coefficients are invariant to cyclic shifts of the input sequence, this is not the case for Walsh-ordered Walsh-Hadamard transform coefficients even though the absolute values of the Walsh-Hadamard transform coefficients and the spectral power of the Walsh-ordered WHT of the data sequence are invariant to dyadic shifts. The implication of this finding is that in the Walsh domain, different spectra will be obtained depending on the exact positions of the pitch pulses in the pitch model. This makes it more difficult to obtain the correct pitch spectrum in the Walsh-domain if one does exist.

The above arguments show that it is not an easy task to exploit the pitch correlation in the Walsh transform domain. Furthermore, even if a method can be found, the question of whether the added computational load is worthwhile for the additional improvement obtained remains to be answered.

The robustness of the hybrid coder is now considered. In the case of voiced speech, the most important information lies in band 1 and band 2 which are coded using adaptive quantizer. The simple

adaptive quantizer makes use of past quantizing information to predict the step-size; any error in the step-size information will therefore propagate. A robust adaptive quantizer as described by Goodman and Wilkinson [66] can be adopted for transmission over noisy channel. With this quantizer, the step-size at sampling instant (k+1), denoted by Δ_{k+1}, is adapted from the preceeding step size Δ_k using

$$\Delta_{k+1} = M'(k)(\Delta_k)^{\beta} \tag{9.2}$$

For $0 < \beta < 1$, the error in Δ_k will be dissipated with time. A fast rate of decay can be obtained with smaller β, but the performance of the quantizer will also degrade when β is chosen to be much less than 1. In order to optimize the performance of the quantizer, different sets of $M'(k)$ will have to be used for different β.

For a signal coded with the adaptive transform technique, any error in the side-information can seriously distort the whole frame of decoded data. Thus error protection must be provided for the side-information. As long as the side-information is correctly coded, error in decoding one or two transform coefficients is unlikely to give rise to any significant degradation of quality.

Simulation can be carried out to determine the perceptual degradation of speech quality resulting from the presence of transmission channel errors when hybrid coders are used without error protections. Various forms of error protection might also be examined.

In Chapter 6, the effects of finite-word-length on the performance of systems using quadrature mirror filters are studied based on rounding off the 'infinitely accurate' values of the filter coefficients. This approach provides significant insight into the accuracy attainable for a given word-length. In actual design when finite-word-length is the main concern, the more complicated integer programming technique [128] can be adopted. Comparative studies of the effects on the performance of systems using quadrature mirror filters when the finite word-length values of the filter coefficients are obtained by rounding the 'infinitely accurate' values and, when they are obtained by integer programming technique, is worthy of investigation.

The compression of speech signals for transmission at bit rates below 16 kbit/s is a complex problem of considerable importance and the interest shown in this area is reflected in the vast amount of literature available on the subject. In this thesis, new techniques were explored and studies of topics related to speech coding were also made. Although these contributions have added to the knowledge of some aspects of speech coding, there is still room for further improvement in speech quality; and this is an area worthy of further investigation. One factor that has an important bearing on the techniques of speech coding is the state of the technology. With very-large-scale-integration (VLSI), system complexity becomes less of a constraint, and it may well be that emphasis should be placed on the quality of the speech produced, almost regardless of system complexity.

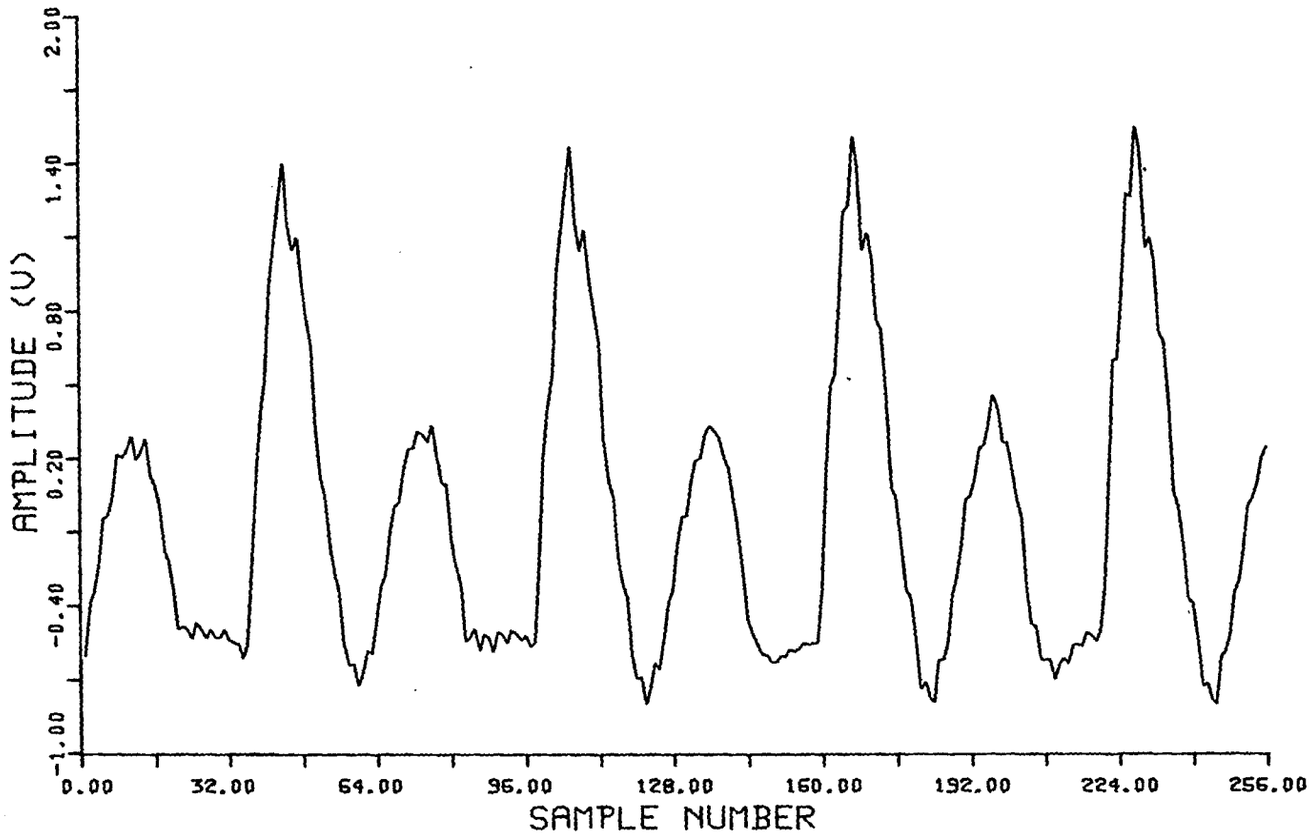


Fig.9.1 Waveform of a voiced segment of speech signal.

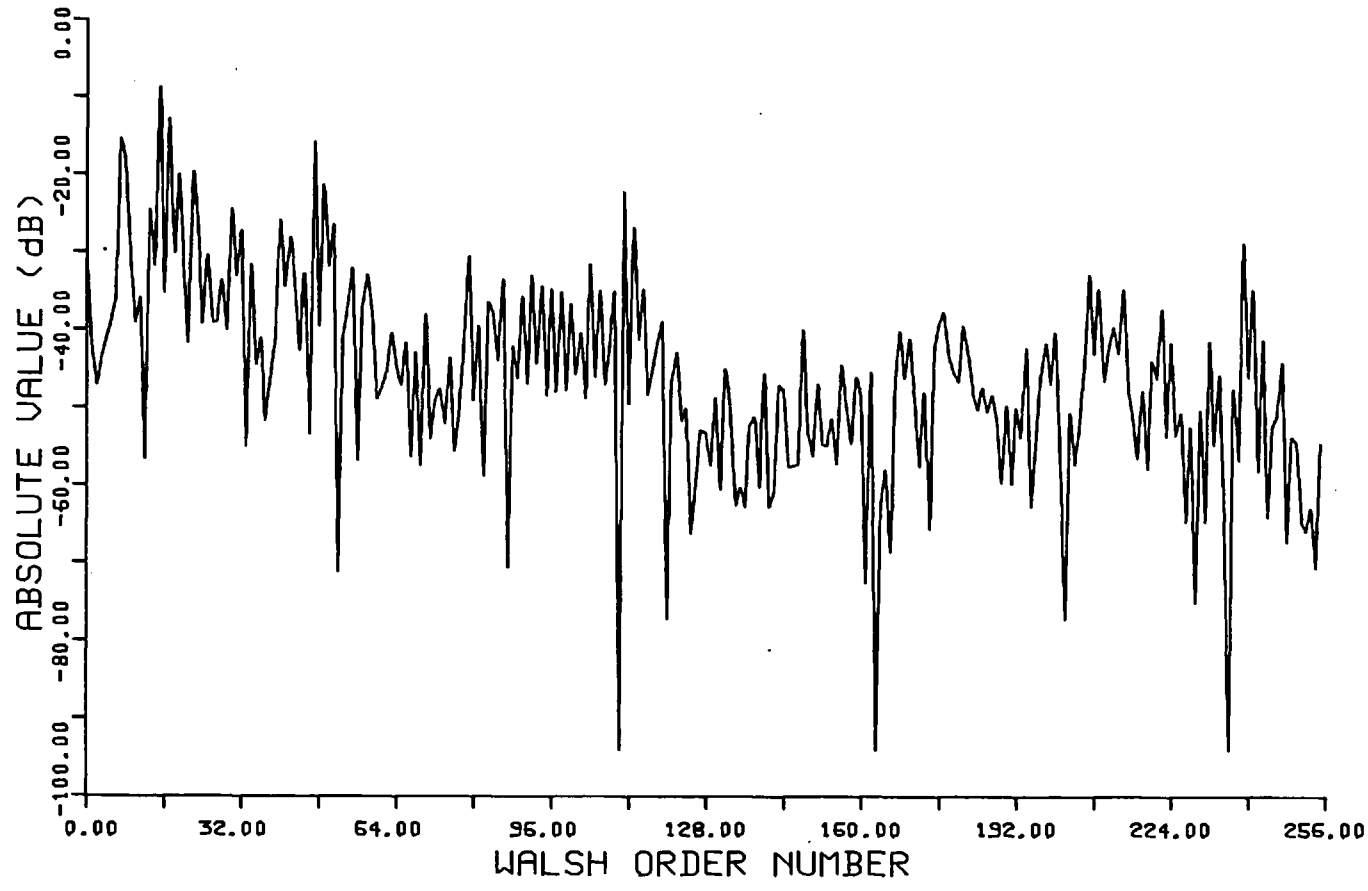


Fig.9.2 Absolute values of the Walsh-Hadamard transform coefficients.

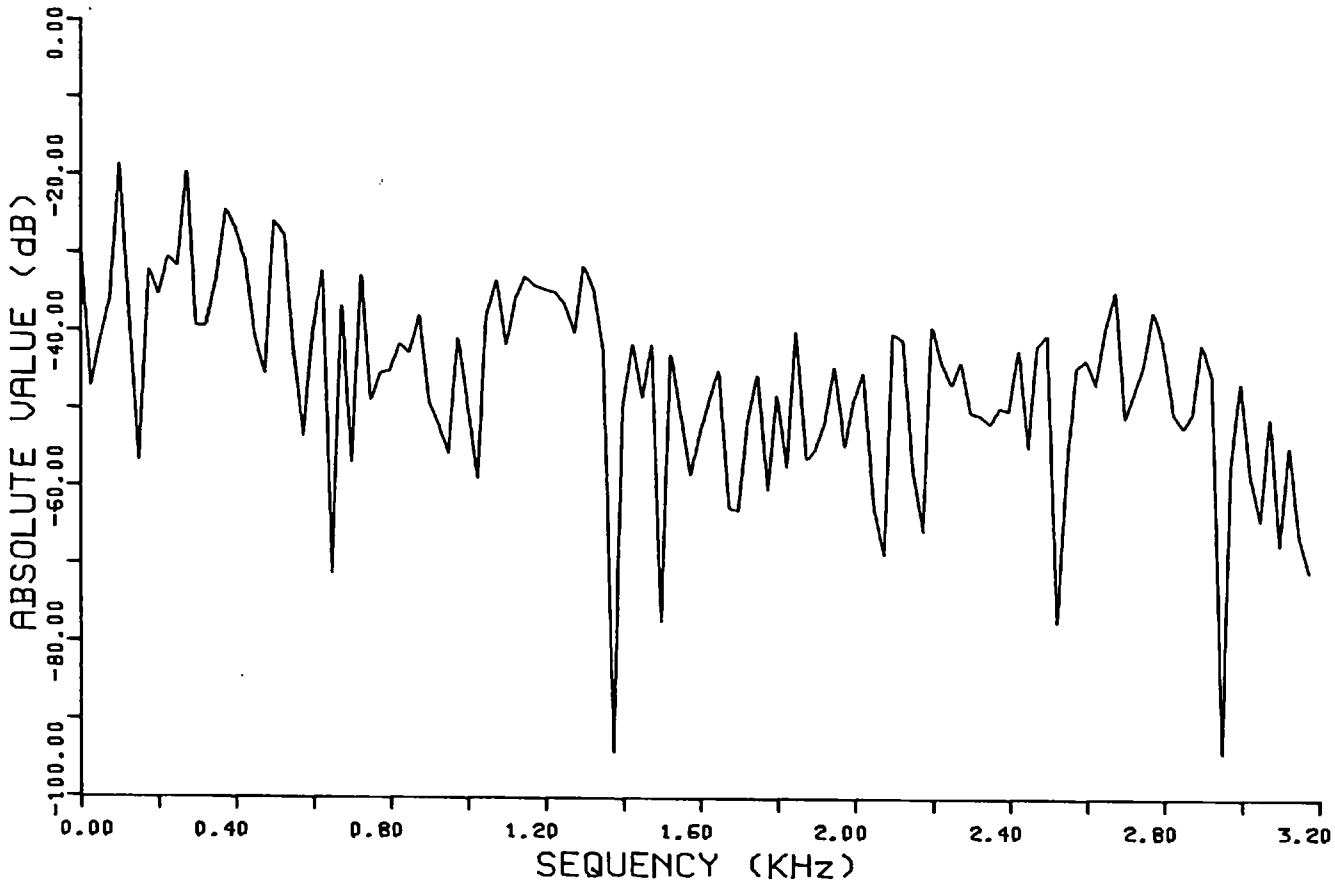


Fig.9.3 Absolute values of the cal components of the Walsh-Hadamard transform.

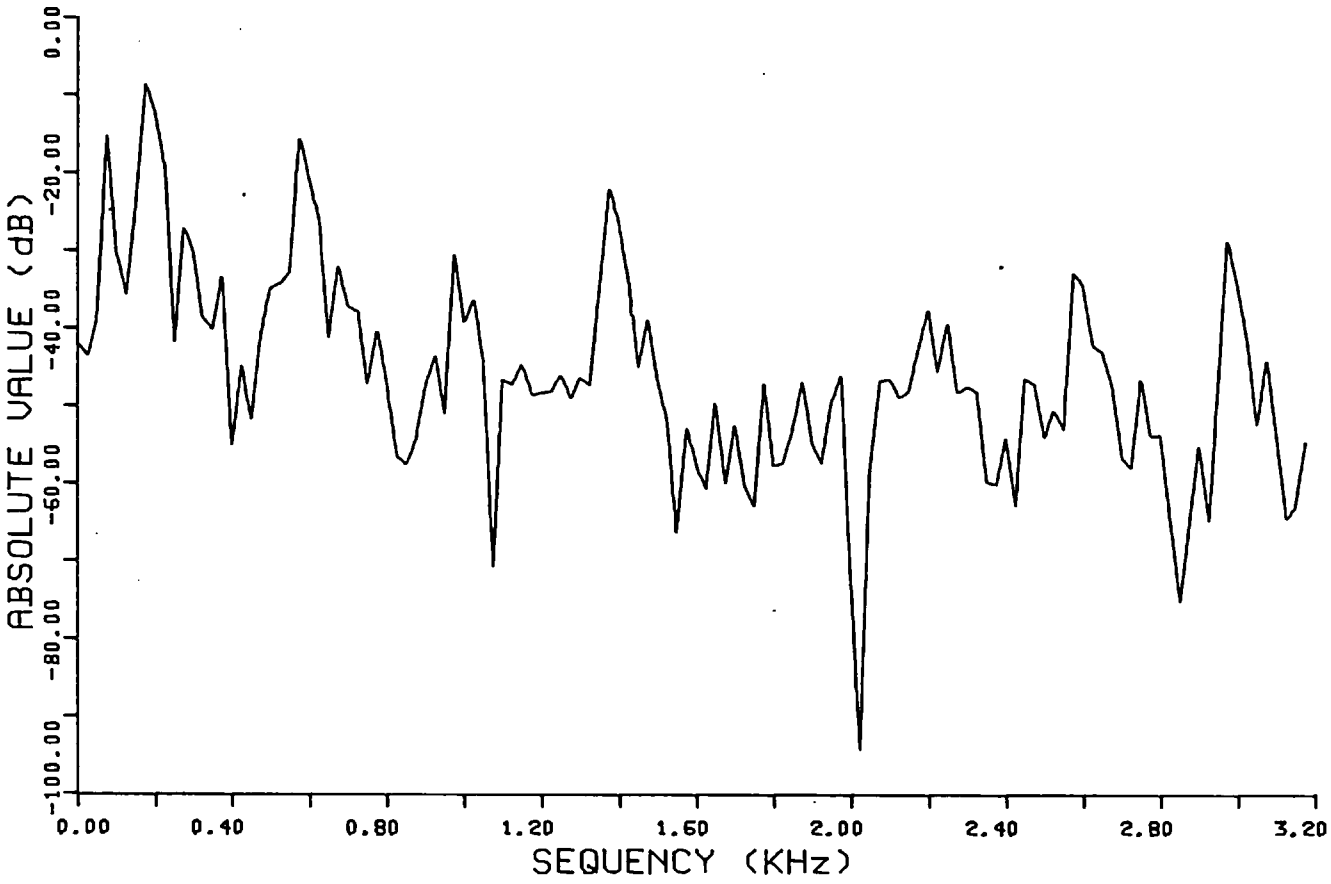


Fig.9.4 Absolute values of the sal components of the Walsh-Hadamard transform.

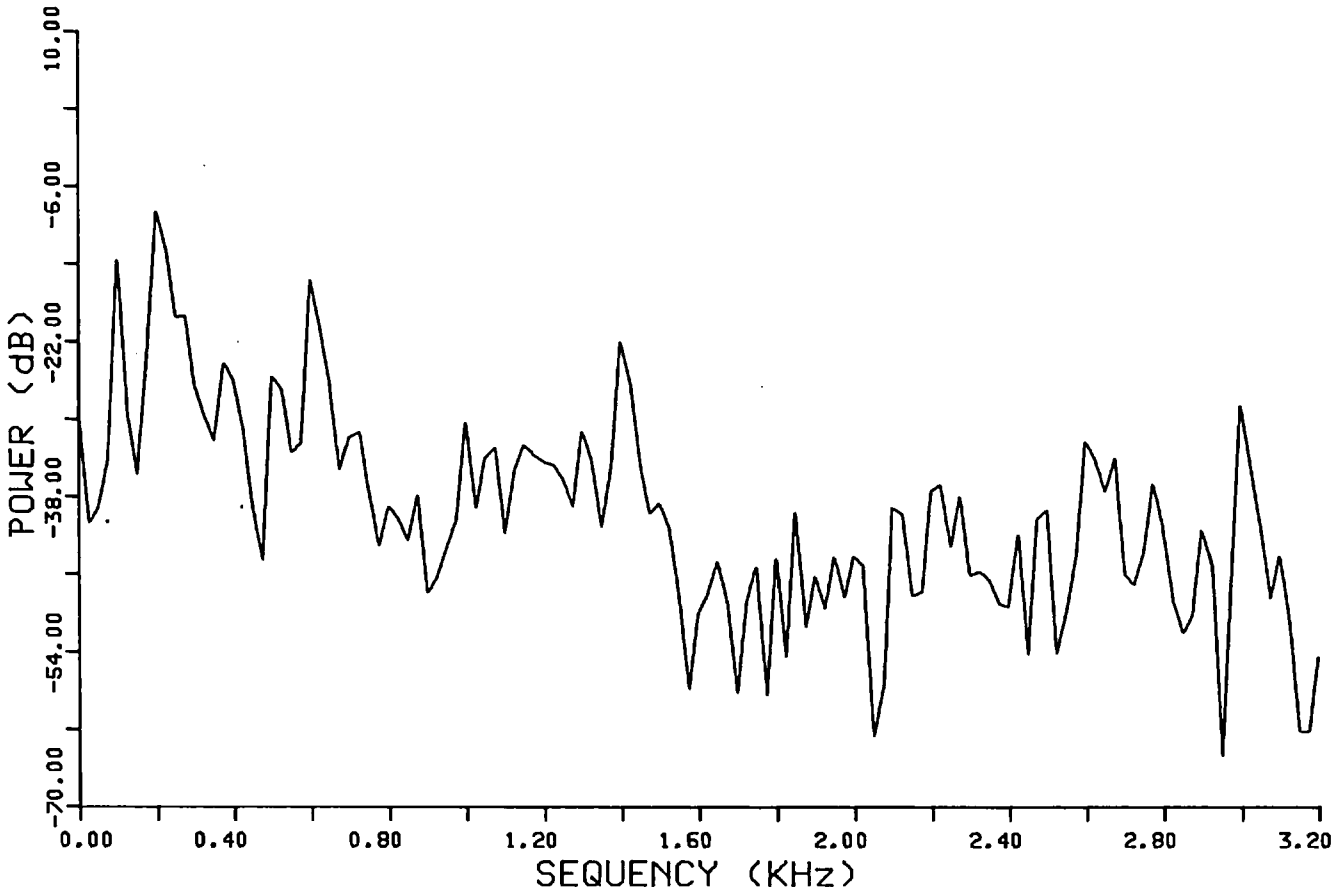


Fig.9.5 Power spectrum of the Walsh-ordered WHT of the voiced segment shown in Fig.9.1.

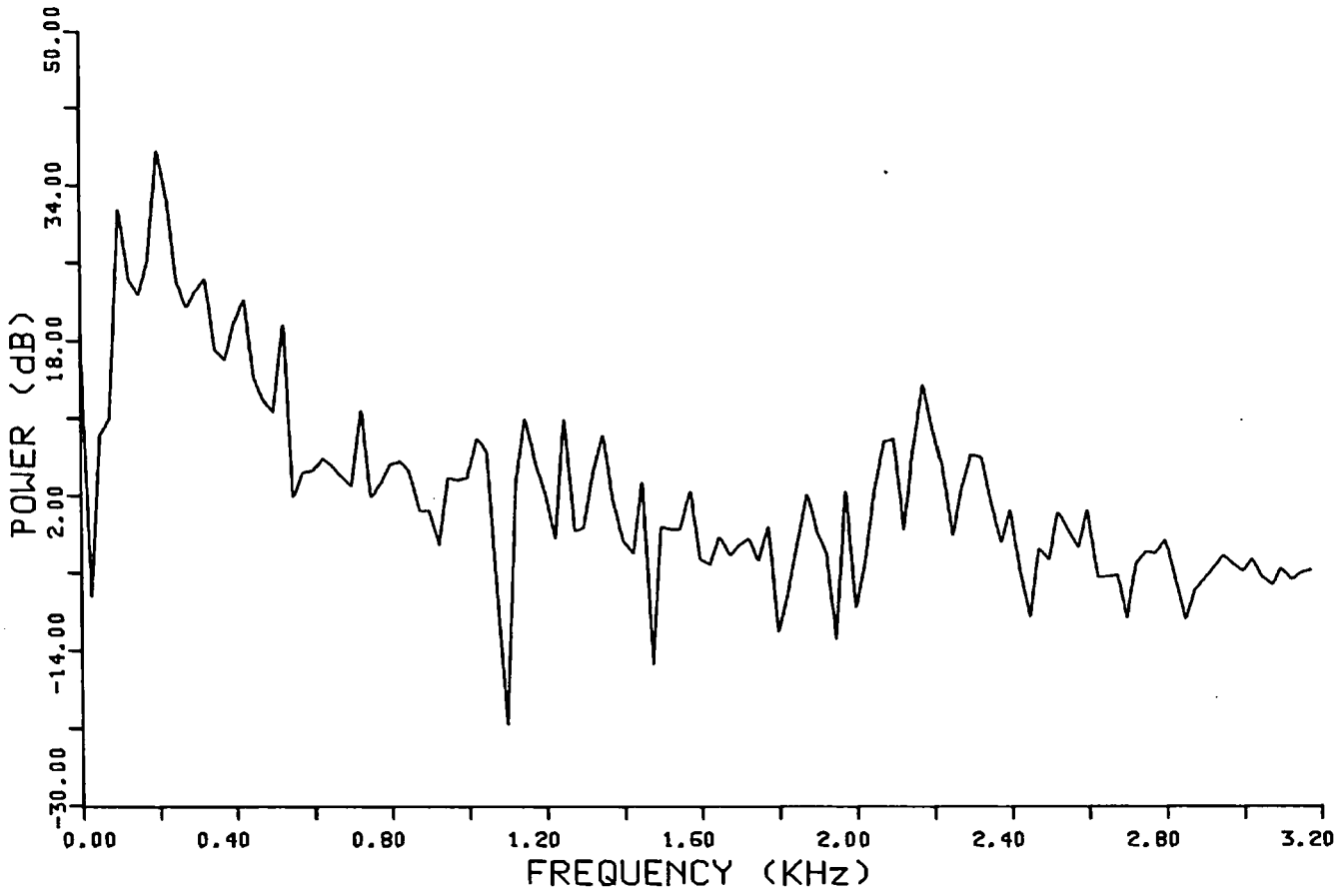


Fig.9.6 Power spectrum of the discrete Fourier transform of the voiced segment shown in Fig.9.1.

REFERENCES

REFERENCES

[1] Scharf,B.
'Critical Bands'
in Tobias,J.V.(ed.) 'Foundations of modern auditory theory'
Academic Press, 1970, pp.157-202.

[2] Paez,M.D. and Glisson,T.H.
'Minimum Mean-Squared-Error Quantization in Speech PCM and DPCM Systems'
IEEE Trans. Commun., pp.225-230, April 1972.

[3] Noll,P.
'Non-adaptive and Adaptive DPCM of Speech Signals'
Polytech., Tijdschr. Ed. Elektrotech/Elektron
(The Netherlands),No.19,1972.

[4] Rabiner,L.R. and Schafer,R.W.
'Digital Processing of Speech Signals'
Prentice Hall, 1978.

[5] Dunn,H.K. and White,S.D.
'Statistical Measurements on Conversational Speech'
J.Acoust.Soc.Amer.,Vol.11,pp.278-288,Jan 1940.

[6] Tribolet,J.M. and Crochiere,R.E.
'Frequency Domain Coding of Speech'
IEEE Trans.Acoust.,Speech and Signal Processing,
Vol.27,No.5,pp.512-530,Oct.,1979.

[7] Gabor,D.
'New Possibilities in Speech Transmission'
J.IEE,94 Pt.III,pp.369-390,1947.

[8] Malah,D.
'Time Domain Algorithm for Harmonic Bandwidth Reduction and Time Scaling of Speech Signals'
IEEE Trans.Acoust.,Speech and Signal Processing,
Vol.27,pp.121-133,Apr.1979.

[9] Schroeder,M.R.,Logan,B.F. and Prestigiacomo,A.J.
'New Methods for Speech Analysis-Synthesis and Bandwidth Compressions'
Proc.Speech Comm.Seminar, Stockholm,1962,paper G8.

[10] Malah,D. and Flanagan,J.L.
'Frequency Scaling of Speech Signals by Transform Techniques'
Bell Syst. Tech J.,Vol.60,No.9,pp.2107-2156, Nov.1981.

[11] Dewey,G.
'Relative frequency of English Speech Sounds'
Cambridge, Massachusetts, Harvard University Press, 1923.

[12] Flanagan,J.L.
'Speech Analysis, Synthesis and Perception'
Springer-Verlag, Berlin-Heidelberg-New York, 1972.

[13] Shannon,C.E.
'A Mathematical Theory of Communication'
Bell Syst.Tech. J., Vol.27,pp.623-656,October 1968.

[14] Haskell,B.G. and Steele,R.
'Audio and Video Bit-Rate Reduction'
Proc.IEEE,Vol.69,No.2,pp.252-262,Feb.,1981.

[15] Flanagan,J.L.,Schroeder,M.R.,Atal,B.S.,Crochiere,R.E.,
Jayant,N.S. and Tribolet,J.M.
'Speech Coding'
IEEE Trans.Comm.,Vol.27,No.4,pp.710-737,April 1979.

[16] Holmes,J.N.
'A Survey of Methods for Digitally Encoding Speech Signals'
IERE Conf.on Digital Processing of Signals in Comm.,
Loughborough, Apr.7-10,1981,Proc.No.49,pp.291-306,1981.

[17] Shannon,C.E.
'Communication in the Presence of Noise'
Proc.IRE,Vol.37,No.1,pp.10-22,Jan.,1949.

[18] Cattermole,K.W.
'Principles of Pulse Code Modulation'
Iliffe Books Ltd.,London 1969.

[19] Cummiskey,P.,Jayant,N.S. and Flanagan,J.L.
'Adaptive Quantization in Differential PCM Coding of Speech'
Bell Syst.Tech.J.,Vol.52,No.7,pp.1105-1118,Sep.,1973.

[20] Stroh,R.W.
'Optimum and Adaptive Differential PCM'
Ph.D. Dissertation,Polytesh.Inst.Brooklyn,Farmingdale,N.Y.,1970.

[21] Atal,B.S. and Schroeder,M.R.
'Predictive Coding of Speech Signals'
Bell Syst.Tech.J.,Vol.49,pp.1973-1986,October,1970.

[22] Atal,B.S. and Schroeder,M.R.
'Adaptive Predictive Coding of Speech Signals' Bell
Syst.Tech.J.,Vol.49,No.8,pp.1973-1986,October,1970.

[23] Noll,P.
'A Comparative Study of Various Schemes for Speech Encoding'
Bell Syst.Tech.J.,Vol.54,No.9,pp.1597-1614,November 1975.

[24] DeJager,F.E.
'Delta Modulation, a Method of PCM Transmission Using a 1-Unit Code' Phillips Res. Rep.,pp.442-466,December 1952.

[25] Greefkes,J.A. and Jager,F.de
'Continuous Delta Modulation'
Phillips Res. Rep.,Vol.23,No.2,pp.233-246,April,1968.

[26] Greefkes,J.A.
'A Digitally Companded Delta Modulation Modem for Speech Transmission'
Proc.IEEE Int.Conf.Comm.,pp.7-33 to 7-48,June 1970.

[27] Greefkes,J.A. and Riemens,K.
'Code Modulation with Digitally Controlled Companding for Speech Transmission'
Phillips Tech.Rev.,Vol.31,No.11/12,pp.335-353,1970.

[28] Jayant,N.S.
'Adaptive Delta Modulation With a One-Bit Memory'
Bell Syst.Tech.J.,pp.321-342,March 1970.

[29] Steele,R.
'Delta Modulation Systems'
Halsted Press, London, 1975.

[30] Atal,B.S. and Hanauer,S.L.
'Speech Analysis and Synthesis by Linear Prediction of the Speech Wave'
J.Acoust.Soc.Amer.,Vol.50,No.2,pp.637-655,August 1971.

[31] Tierney,J.,Gold,B.,Sferrino,V.,Dumanian,J.A. and Aho,E.
'Channel Vocoder with Digital Pitch Extractor'
J.Acoust.Soc.Amer.,Vol.36,No.10,pp.1901-1905,October,1964.

[32] Gold,B. and Rader,C.M.
'The Channel Vocoder'
IEEE Trans. Audio and Electroacoust.,
Vol.15,pp.148-160,Dec.1967.

[33] Holmes,J.N.
'The JSRU Channel Vocoder'
IEE Proc.Vol.127,Part F,No.1,pp.53-60,February 1980.

[34] Munson,W.A. and Montgomery,H.C.
'A Speech Analyzer and Synthesizer'
J.Acoust.Soc.Amer.,Vol.22,678(A),1950.

[35] Flanagan,J.L. and House,A.S.
'Development and Testing of a Formant-Coding Speech Compression System'
J.Acoust.Soc.Amer.,Vol.28,pp.1099-1106,1956.

[36] Stead,L.G. and Jones,E.T.
'The S.R.D.E. Speech Bandwidth Compression Project'
Report No.1133,Signals Research and Development Establishment,
Christchurch,England, March 1961.

[37] Coker,C.H.
'Real-Time Formant Vocoder, Using a Filter Bank, a General Purpose Digital Computer, and an Analog Synthesizer'
J.Acoust.Soc.Amer.,Vol.38,p940,1965.

[38] Rabiner,L.R.
'Digital-Formant Synthesizer for Speech-Synthesis Studies'
J.Acoust.Soc.Amer.,Vol.43,No.4,pp.822-828,1968.

[39] Rabiner,L.R.,Jackson,L.B.,Schafer,R.W. and Coker,C.H.
'A Hardware Realization of a Digital Formant Speech Synthesizer'
IEEE Trans. Comm.Tech.,Vol.19,No.6,pp.1016-1020,Dec.1971.

[40] Holmes,J.N.
'Parallel Formant Vocoders'
Proc.IEEE EASCON Conf.,Washington,D.C.,pp.713-718,1978.

[41] Flanagan,J.L.
'Resonance-Vocoder and Baseband Complement'
IRE Trans.Audio Vol.8,pp.95-102,1960.

[42] Flanagan,J.L. and Golden,R.M.
'Phase Vocoder'
Bell Syst.Tech.J.,Vol.45,pp.1493-1509,1966.

[43] Oppenheim,A.V. and Schafer,R.W.
'Homomorphic Analysis of Speech'
IEEE Trans. Audio and Electroacoust.,
Vol.16,No.2,pp.221-226,June 1968.

[44] Markel,J.D. and Gray,A.H.,Jr.
'Linear Prediction of Speech'
Springer-Verlag, 1980.

[45] Schroeder,M.R. and Crystal,T.H.
'Auto-Correlation Vocoder'
J.Acoust.Soc.Amer.,Vol.32,No.7,p.913,July,1960.

[46] Crowther,W.R. and Rader,C.M.
'Efficient Coding of Vocoder Channel Signals Using Linear Transformation'
Proc.IEEE,Vol.54,No.11,pp.1594-1595,November 1966.

[47] Sambur,M.R.
'An Efficient Linear-Prediction Coder'
Bell Syst.Tech.J.,Vol.54,No.10,pp.1693-1723,Dec.1975.

[48] McLarnon,E.
'A Method for Reducing the Transmission Rate of a Channel Vocoder by Using Frame Interpolation'
Proc.IEEE Int.Conf.on Acoustics,Speech and Signal Processing, Tulsa,Oklahoma,pp.458-461,1978.

[49] Holmes,J.N.
'Variable-Frame-Rate Coding Scheme for Speech-Analysis/Synthesis Systems'
Electron.Lett.,Vol.10,No.7, pp.101-102, 1974.

[50] Smith,C.P.
'Voice-Communications Method Using Pattern Matching for Data Compression'
J.Acoust.Soc.Amer.,Vol.35,No.5,805(A) May,1963.

[51] Smith,C.P.
'Perception of Vocoder Speech Processed by Pattern Matching'
J.Acoust.Soc.Amer.,Vol.46,No.6,pp.1562-1571,1969.

[52] Buzo,A.,Gray,A.H.,Gray,R.M. and Markel,J.D.
'Speech Coding Based Upon Vector Quantization'
IEEE Trans. Acoust.,Speech and Signal Processing, Vol.28,No.5,pp.562-574,Oct.1980.

[53] Buzo,L.A.
'Optimum Vector Quantization for Linear Predictive Coded Speech'
Ph.D. dissertation, Stanford University, Stanford,C.A. Aug.1978.

[54] David,E.E.Jr.,Schroeder,M.R.,Logan,B.F. and Prestigiacomo,A.J.
'New Applications of Voiced-Excitation to Vocoders'
Proc.Speech Comm.Seminar, Stockholm, 1962, paper G3.

[55] Gold,B. and Tierney,J.
'Digitalized Voiced-Excited Vocoder for Telephone Quality Inputs Using Bandpass Sampling of the Baseband Signal'
J.Acoust.Soc.Amer.,Vol.73,pp.753-754,1965.

[56] Malah,D.,Crochiere,R.E. and Cox,R.V.
'Performance of Transform and Subband Coding Systems Combined With Harmonic Scaling of Speech'
IEEE Trans.Acoust.,Speech and Signal Processing, Vol.29, pp.273-283, April 1981.

[57] Crochiere,R.E.,Webber,S.A. and Flanagan,J.L.
'Digital Coding of Speech in Sub-bands'
Bell Syst.Tech.J.,Vol.55,No.8,pp.1069-1085,Oct.1976.

[58] Crochiere,R.E.
'On the Design of Sub-band Coders for Low-Bit-Rate Speech Communication'
Bell Syst.Tech.J.,Vol.56,No.5,pp.747-770,May-June,1977.

[59] Crochiere,R.E. and Sambur,M.R.
'A Variable-Band Coding Scheme for Speech Encoding at 4.8 Kb/s.'
Bell Syst.Tech.J.,Vol.56,No.5,pp.771-779,May-June,1977.

[60] Crochiere,R.E.
'A Novel Approach for Implementing Pitch Prediction in Subband Coding' Proc.IEEE Int.Conf.on Acoust.,Speech and Signal Processing,
Washington,D.C. pp.526-529,April2-4,1979.

[61] Barabell,A.J. and Crochiere,R.E.
'Subband Coder Design Incorporating Quadrature Filters and Pitch Prediction'
Proc.IEEE Int.Conf.on Acoust.,Speech and Signal Processing,
Washington,D.C. pp.530-533,Apr.2-4,1979.

[62] Estaban,D. and Galand,C.
'Application of Quadrature Mirror Filters to Split Band Voice Coding Schemes'
Proc.IEEE Int.Conf.on Acoust.,Speech and Signal Processing,
Hartford,Conn. pp.191-195,May 9-11,1977.

[63] Grauel,C.
Subband Coding With Adaptive Bit Allocation'
Signal Processing, Vol.2,pp.23-30,1980.

[64] Jayant,N.S.
'Adaptive Quantization With a One-Word Memory'
Bell Syst.Tech.J.,Vol.52,No.7,pp.1119-1143,Sep.1973.

[65] Goodman,D.J. and Gersho,A.
'Theory of an Adaptive Quantizer'
IEEE Trans. Comm. Vol.22,No.8,pp.1037-1045,Aug.1974.

[66] Goodman,D.J. and Wilkinson,R.M.
'A Robust Adaptive Quantizer'
IEEE Trans. Comm.,Vol.23,pp.1362-1365,Nov.1975.

[67] Ramstad,T.A.
'Sub-band Coder With a Simple Adaptive Bit Allocation Algorithm: A Possible Candidate for Digital Mobile Telephony ?'
Proc.IEEE Int.Conf.Acoust.,Speech and Signal Processing,
Paris,pp.203-207,Apr.1982.

[68] Ahmed,N. and Rao,K.R.
'Orthogonal Transform for Digital Signal Processing'
Springer-Verlag, 1975.

[69] Harmuth,H.F.
'Transmission of Information by Orthogonal Functions'
Springer-Verlag,Berlin-Heidelberg-New York, 1972.

[70] Manz,J.W.
'A Sequency-Ordered Fast Walsh Transform'
IEEE Trans.Audio Electroacoust.,Vol.20,No.3,pp.204-205,Aug.1972.

[71] Cooley,J.W. and Tukey,J.W.
'An Algorithm for the Machine Computation of Complex Fourier Series'
Math.Computation,Vol.19,pp.297-381,April 1965.

[72] Singleton,R.C.
'An Algorithm for Computing the Mixed Radix Fast Fourier Transform'
IEEE Trans. Audio and Electoacoust.,
Vol.17,No.2,pp.93-103,June 1969.

[73] Rabiner,L.R. and Gold,B.
'Theory and Application of Digital Signal Processing'
Prentice Hall, Englewood Cliffs, N.J., 1975.

[74] Ahmed,N.,Natarajan,T. and Rao,K.R.
'Discrete Cosine Transform'
IEEE Trans. Computers,Vol.23,No.1,pp.90-93,Jan.1974.

[75] Haralick,R.M.
'A Storage Efficient Way to Implement the Discrete Cosine Transform'
IEEE Trans. Computers,Vol.25,No.7,pp.764-765,July 1976.

[76] Chen,W.H.,Smith,C.H. and Fralick,S.C.
'A Fast Computational Algorithm for the Discrete Cosine Transform'
IEEE Trans. Comm.,Vol.25,No.9,pp.1004-1009,Sep.1977.

[77] Narasimha,M.J. and Peterson,A.M.
'On the Computation of the Discrete Cosine Transform'
IEEE Trans. Comm.,Vol.26,No.6,pp.934-936,1978.

[78] Makhoul,J.
'A Fast Cosine Transform in One and Two Dimensions'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.28,No.1,pp.27-34,Feb.1980.

[79] Huang,J.J.Y. and Schultheiss,P.M.
'Block Quantization of Correlated Gaussian Random Variables'
IEEE Trans. Comm.Systems,Vol.11,pp.289-296,Sep.1963.

[80] Campanella,S.J. and Robinson,G.S.
'A Comparison of Orthogonal Transformations for Digital Signal Processing'
IEEE Trans. Communication Technology,
Vol.19,No.6,pp.1045-1049,Dec.1971.

[81] Zelinski,R. and Noll,P.
'Adaptive Transform Coding of Speech Signals'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.25,No.4,pp.299-309, August 1977.

[82] Shum,Y.Y.,Elliott,A.R. and Brown,O.W.
'Speech Processing With Walsh-Hadamard Transforms'
IEEE Trans. Audio,Vol.21,pp.174-179,1973.

[83] Frangoulis,E. and Turner,L.F.
'Hadamard-Transformation Technique of Speech Coding: Some Further Results'
Proc.IEE,Vol.124,No.10 pp.845-852,Oct.1977.

[84] Cox,R.V. and Crochiere,R.E.
'Real Time Simulation of Adaptive Transform Coding'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.29,No.2,pp.147-154,Apr.1981.

[85] Zelinski,R. and Noll,P.
'Approach to Adaptive Transform Speech Coding at Low Bit Rates'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.27,No.1,pp.89-95,Feb.1979.

[86] Bullington,K. and Fraser,J.M.
'Engineering Aspects of TASI'
Bell Syst.Tech.J.,pp.353-364,Mar 1959.

[87] Fraser,J.M.,Bullock,D.B. and Long,N.G.
'Overall Characteristics of a TASI system'
Bell Syst.Tech.J.,pp.1439-1454,July 1962.

[88] Mathews,M.V.
'Extremal Coding for Speech Transmission'
IRE Trans. Information Theory,Vol.5,No.3,pp.129-136,Sep.1959.

[89] King,R.A. and Gosling,W.
'Time Encoded Speech'
Electronics Letters,Vol.14,No.15,pp.456-457,July 1978.

[90] King,R.A. and Gosling,W.
'Time Encoded Speech'
1979 Peebles Symposium,IEE Conf. Publication No.180,
pp.140-143,1979.

[91] Jayant,N.S. and Christensen,S.W.
'Adaptive Aperture Coding for Speech Waveforms - I'
Bell Syst.Tech.J.,Vol.58,No.7,pp.1631-1645,Sept.1979.

[92] Jayant,N.S. and Christensen,S.W.
'Adaptive Aperture Coding for Speech Waveforms - II'
Bell Syst.Tech.J.,Vol.59,No.3,pp.471-477,March 1980.

[93] Dubnowski,J.J. and Crochiere,R.E.
'Variable Rate Coding of Speech'
Bell Syst.Tech.J.,Vol.58,No.3,pp.577-600,March 1979.

[94] Turner,L.F.,Frangoulis,E. and Alcaim,A.
'Some Considerations Relating to the Performance of Variable-Information-Rate-Source to Constant-Transmission-Rate Scheme of Data Compression'
Proc.IEE Computers and Digital Techniques,Vol.2,pp.134-141,1979.

[95] Turner,L.F.,Frangoulis,E. and Alcaim,A.
'Further Results on Relationship between System Delay and Transmission Rate in Time Encoded Speech Type Systems'
Electron.Lett., Vol.16, No.25, pp.947-948, 1980.

[96] Mason,D.C. and Balston,D.M.
'Relationship between System Delay and Transmission Rate in Time-Encoded Speech'
Electron.Lett., Vol.16, No.4, pp.128-130, 1980.

[97] King,R.A. and Holbeche,J.
'The Impact of Entropic Coding on the Time Delays Associated with Time-Encoded Speech (TES) Systems'
Electron.Lett., Vol.17, pp.394-396, 1981.

[98] Holmes,J.N.
'Buffering Strategy for Variable-Symbol-Rate Speech Waveform Coders'
Electron.Lett., Vol.18, No.16, pp.691-693, 1982.

[99] Beranek,L.L.
'The Design of Speech Communication Systems'
Proc.of the IRE, pp.880-890, Sept. 1947.

[100] Noll,P.
'Adaptive Quantizing in Speech Coding Systems'
Int. Zurich Seminar, B3(1-6), March 1974.

[101] Crochiere,R.E.
'A Mid-Rise/Mid-Tread Quantizer Switch for Improved Idle-Channel Performance in Adaptive Coders.'
Bell Syst.Tech.J.,Vol.57,No.8,pp.2953-2955,Oct.1978.

[102] Berger,T.
'Rate Distortion Theory - A Mathematical Basis for Data Compression'
Prentice Hall, 1971.

[103] Zelinski,R. and Noll,P.
'Adaptive Block Quantization of Speech Signals (in German)'
Heinrich-Hertz-Institut,Berlin,Germany,Tech.Rep.181,1975.

[104] Max,J.
'Quantizing for Minimum Distortion'
IRE Trans. Information Theory,Vol.6,pp.7-12,March 1960.

[105] Crochiere,R.E.,Goodman,D.J.,Rabiner,L.R. and Sambur,M.R.
'Tandem Connections of Wideband and Narrowband Speech Communications Systems: Part 1 - Narrow-to-Wideband Link'
Bell Syst.Tech.J.,Vol.56,No.9,pp.1701-1721,1977.

[106] Webber,S.A.,Harris,C.J. and Flanagan,J.L.
'Use of Variable-Quality Coding and Time-Interval Modification in Packet Transmission of Speech'
Bell Syst.Tech.J.,Vol.56,pp.1569-1573,1977.

[107] Ramstad,T.A. and Foss,O.
'Subband Coder Design Using Recursive Quadrature Mirror Filters'
Proc.Conf.of European Association for Signal Processing, Lausanne, Switzerland, pp.747-752, 1980.

[108] Cheung,R.S. and Winslow,R.L.
'High Quality 16 Kb/s Voice Transmission: The Subband Coder Approach'
Proc.IEEE Int.Conf.Acoust.,Speech and Signal Processing, Denver,CO,pp.319-322,Apr.1980.

[109] Schaffer,R. and Rabiner,L.R.
'A Digital Signal Processing Approach to Interpolation'
Proc.IEEE,Vol.61,pp.692-702,June 1973.

[110] Gass,S.I.
'Linear Programming : Methods and Applications'
McGraw Hill Book Company, 1969.

[111] Simonnard,M.
'Linear Programming'
Prentice Hall,Inc., Englewood Cliffs,N.J., 1966.

[112] Johnston,J.D.
'A Filter Family Designed for Use in Quadrature Mirror Filter Bank'
Proc.IEEE Int.Conf.Acoust.,Speech and Signal Processing, Denver,CO,pp.291-294,Apr.1980.

[113] Hooke,R. and Jeaves,T.
'Direct Search Solution of Numerical and Statistical Problems'
J.Ass.Comp.Mach.,Vol.8,pp.212-229,Apr.1961.

[114] Kuester,J.L. and Mize,J.H.
'Optimization Techniques with Fortran'
McGraw Hill,N.Y.,1973.

[115] Hamacher,V.C.,Vranesic,Z.G. and Zaky,S.G.
'Computer Organisation'
McGraw Hill,1978.

[116] Chan,D.S.K. and Rabiner,L.R.
'Theory of Roundoff Noise in Cascade Realizations of Finite Impulse Response Digital Filters'
Bell Syst.Teeh.J.,Vol.52,pp.329-245,Mar.1973.

[117] Chan,D.S.K. and Rabiner,L.R.
'Analysis of Quantization Errors in the Direct Form for Finite Impulse Response Digital Filters'
IEEE Trans.Audio Electroacoust.,Vol.21,No.4,pp.354-366,1973.

[118] Hogg,R.V. and Graig,A.T.
'Introduction to Mathematical Statistics'
Macmillan 1971.

[119] Welch,P.D.
'A Fixed-Point Fast Fourier Transform Error Analysis'
IEEE Trans.Audio Electroacoust.,Vol.17,No.2,pp.151-157,Jun.1969.

[120] Rabiner,L.R. and Sambur,M.R.
'An Algorithm for Determining the Endpoints of Isolated Utterances'
Bell Syst.Tech.J.,Vol.54,No.2,pp.297-315,Feb.1975.

[121] Gold,B.
'Note on Buzz-Hiss Detection'
J.Acoust.Soc.Amer.,Vol.36,pp.1659-1661,1964.

[122] Atal,B.S. and Rabiner,L.R.
'A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.24,No.3,pp.201-212,June 1976.

[123] Rabiner,L.R. and Sambur,M.R.
'Application of an LPC Distance Measure to the Voiced-Unvoiced-Silence Detection Problem'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.25,No.4,pp.338-343,Aug.1977.

[124] Cox,B.V. and Timothy,L.M.K.
'Non-parametric Rank-Order Statistics Applied to Robust Voiced-Unvoiced-Silence Classification'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.28,No.5,pp.550-561,Oct.1980.

[125] Knorr,S.G.
'Reliable Voiced/Unvoiced Decision'
IEEE Trans.Acoust.,Speech and Signal Processing,
Vol.27,No.3,pp.263-267, June 1979.

[126] Siegel,L.J.
'A Procedure for Using Pattern Classification Techniques to Obtain a Voiced/Unvoiced Classifier'
IEEE Trans. Acoust.,Speech and Signal Processing,
Vol.27,No.1,pp.83-88,Feb.1979.

[127] Ramamoorthy,V.
'A Simple Time Domain Algorithm for Voiced-Unvoiced Detection'
Proc.Conf.of European Association for Signal Processing,
Lausanne, Switzerland, pp.737-742, 1980.

[128] Taha,H.
'Integer Programming : Theory, Applications, and Computations'
Academic Press, New York, 1975.

[129] Cochran,W.G. and Cox,G.M.
'Experimental Designs'
John Wiley and Sons,Inc. 1957.

[130] Cox,D.R.
'Planning of Experiments'
John Wiley and Sons,Inc. 1958.

[131] Kempthorn,O.
'The Design and Analysis of Experiments'
Robert E Krieger Publishing Company, 1973.

[132] Scheffe,H.
'The Analysis of Variance'
John Wiley, N.Y., 1959.

[133] Wetherill,G.B.
'Elementary Statistical Methods'
Chapman and Hall, 1972.

[134] Cox,D.R.
'The Analysis of Binary Data'
Methuen and Co. Ltd., 1970.

[135] Tukey,J.W.
'One Degree of Freedom for Non-additivity'
Biometrics, Vol.5,pp.232-242,1949.

[136] Anscombe,F.J. and Tukey,J.W.
'The Examination and Analysis of Residuals'
Technometrics, Vol.5, pp.141-160, 1963.

[137] Chernoff,H. and Lieberman,G.J.
'Use of Normal Probability Paper'
J.Amer.Statist.Ass.,Vol.49,pp.778-785,1954.

APPENDICES

APPENDIX A

PARAMETERS USED FOR THE CVSD SCHEME

The block diagram of a CVSD coder is shown in Fig.A.1. The relevant parameters and equations used are described briefly below.

At the k^{th} sampling instant, for example, the input sample, x_k, of the speech signal is compared with an estimated signal sample y_k. The transmitted signal u_k is then determined as follows:

$$\begin{aligned} u_k &= 1 \quad \text{if } x_k > y_k \\ &= -1 \quad \text{if } x_k \leq y_k \end{aligned} \tag{A.1}$$

The next estimated signal sample y_{k+1} is then computed from y_k and u_k as

$$y_{k+1} = 0.94y_k + .18u_k\Delta_k \tag{A.2}$$

where Δ_k is the k^{th} step-size.

The step-size is obtained from the step-size integrator which processes the output of an overload detector. When the three previous output samples are identical (either all 1 or all -1), the output of the overload detector is V; otherwise it is 0. A small quantity V_1 is added to the output of the overload detector to ensure that the minimum step-size is non-zero.

As the speech samples are represented to a 12-bit accuracy in the computer simulations used to produce the test material, V is chosen to be 2047. To provide for a wide dynamic range of step-sizes, V/V_1 is set equal to 200 and thus $V_1 = 2047/200 = 10.24$. The step-size is then updated according to the following relation,

$$\begin{aligned} \Delta_{k+1} &= 0.99\Delta_k + 20.57 \quad \text{when } u_{k-2} = u_{k-1} = u_k \\ &= 0.99\Delta_k + 0.1024 \quad \text{otherwise} \end{aligned} \tag{A.3}$$

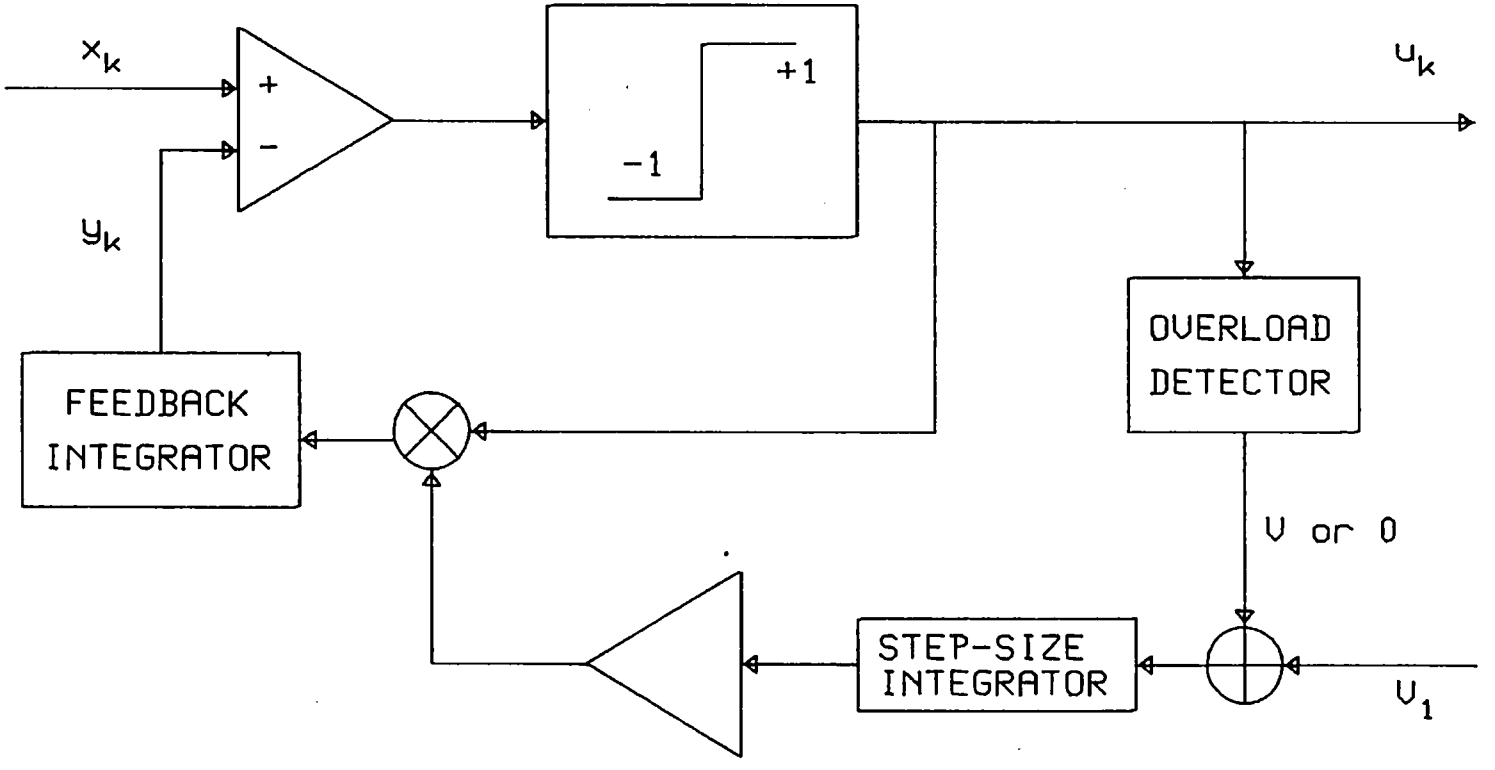


Fig.A.1 Block diagram of CVSD coder.

APPENDIX B

SUBJECTIVE TESTS OF FIXED-RATE HFDC

B.1 DESIGN OF EXPERIMENTS

The objective of the experiments is to compare the subjective quality of speech reproduced by hybrid frequency-domain coding technique (HFDC) with that reproduced by the continuously variable slope delta-modulation (CVSD).

Two separate experiments were carried out. For experiment 1, comparisons were made between HFDC at 7.2 kbit/s and CVSD at bit rates of 14 kbit/s, 17 kbit/s, 20 kbit/s and 23 kbit/s. For experiment 2, comparisons were made between HFDC at bit rate of 9.6 kbit/s and CVSD at 17 kbit/s, 20 kbit/s, 23 kbit/s and 26 kbit/s.

As the nature of the experiments are similar, the same kind of design and analysis was used. Let the main coding scheme (HFDC) be denoted by C_0 and the CVSD coding schemes operating at four different bit rates by C_1, C_2, C_3 and C_4. A forced decision paired-comparison method was adopted. The subjective tests were designed as factorial experiments [129-131] using a randomized balance design so that any effects not specifically included would not distort the results. Specifically, the tests were designed as 4X4 experiments with the factors being (i) sentences used and (ii) coder pairs $\{(C_0, C_i), i=1,2,3,4\}$.

Two sentences by a male speaker (M_1 and M_2) and two different sentences by a female speaker (F_1 and F_2) were used. Different sentences were used to reduce the number of repetitions and make listenings more interesting.

As C_0 is to be compared with C_1, C_2, C_3 and C_4 using four different sentences (M_1, M_2, F_1 and F_2), a total of 16 comparisons was to be made. Each combination of (sentence, coder pair) was assigned a number from 1 to 16. The order of presentation of the (sentence, coder pair) to the listeners was based on a sequence of randomly generated number from 1 to 16. For each experiment, two orders of presentation were used.

For further randomization, the presentation within the pair was ordered as (C_0, C_i) for sentences M_1 and F_1; and in the order (C_i, C_0) for sentences M_2 and F_2, with i=1,2,3,4.

The actual arrangements used are given in Table B.1. The first element in each group indicates the sentence used, the second and third elements indicate the coders used in the order given.

Each pair of sentences together with the serial number was recorded twice on analog magnetic tapes with a short interval of less than 1 second between them and an additional interval of about 15 seconds was allowed between the different comparisons.

Table B.1 Arrangements of Sentences on Tapes

Number	Experiment 1 (7.2 kbit/s)		Experiment 2 (9.6 kbit/s)	
	Tape A	Tape B	Tape C	Tape D
1	F_1,C_0,C_4	F_2,C_0,C_3	M_1,C_0,C_1	M_2,C_0,C_3
2	M_2,C_0,C_1	F_1,C_1,C_0	F_2,C_0,C_2	M_1,C_1,C_0
3	F_2,C_2,C_0	M_1,C_0,C_3	F_1,C_2,C_0	F_2,C_3,C_0
4	M_1,C_2,C_0	M_2,C_4,C_0	M_2,C_3,C_0	M_2,C_0,C_4
5	F_2,C_3,C_0	F_2,C_0,C_2	F_1,C_4,C_0	F_1,C_0,C_3
6	F_1,C_0,C_3	M_1,C_0,C_1	F_2,C_0,C_4	F_2,C_1,C_0
7	M_1,C_1,C_0	F_1,C_2,C_0	M_1,C_0,C_3	F_1,C_0,C_4
8	M_2,C_0,C_3	M_1,C_0,C_4	F_1,C_3,C_0	M_1,C_4,C_0
9	F_1,C_0,C_1	F_2,C_0,C_4	M_1,C_0,C_4	F_2,C_4,C_0
10	F_2,C_4,C_0	M_2,C_3,C_0	M_2,C_4,C_0	M_2,C_0,C_2
11	M_2,C_0,C_4	F_1,C_3,C_0	F_2,C_0,C_1	F_1,C_0,C_1
12	M_1,C_3,C_0	F_2,C_0,C_1	M_2,C_2,C_0	M_1,C_3,C_0
13	M_2,C_0,C_2	M_2,C_1,C_0	M_1,C_0,C_2	F_2,C_2,C_0
14	F_1,C_0,C_2	M_1,C_0,C_2	F_1,C_1,C_0	F_1,C_0,C_2
15	F_2,C_1,C_0	M_2,C_2,C_0	F_2,C_0,C_3	M_2,C_0,C_1
16	M_1,C_4,C_0	F_1,C_4,C_0	M_2,C_1,C_0	M_1,C_2,C_0

The sentences used were:

M_1: Airmail requires a special stamp.

M_2: The guilty one should take the blame.

F_1: The hockey player scored a goal.

F_2: He wiped the sink with a sponge.

B.2 CONDUCT OF EXPERIMENTS

The subjects for both experiments consisted of twelve undergraduate and postgraduate students, ten male and two female, with English as their mother tongue and with no previous experience in this kind of test. As an incentive to their efforts, each was paid one sterling pound. Six subjects listened to Tape A and Tape C, and six others to Tape B and Tape D. The experiments were conducted in a quiet room and listenings were made through headphones. Each experiment lasted for about 8 minutes.

For each comparison (sentence A vs. sentence B), the listeners were asked to decide whether (i) sentence A is better than sentence B; or (ii) sentence B is better than sentence A.

The responses of all the 12 listeners for each experiment were grouped together and these are tabulated and shown in Tables B.2 and B.3. In these tables, the data are shown in the format (m/M), where M denotes the total votes available for a given combination and m is the number of votes given in favour of the coding scheme specified.

Table B.2 Results of Experiment 1

Preference for	M_1	M_2	F_1	F_2
C_1	2/12	1/12	4/12	2/12
C_2	3/12	5/12	3/12	4/12
C_3	8/12	7/12	6/12	7/12
C_4	10/12	11/12	10/12	11/12

Table B.3 Results of Experiment 2

Preference for	M_1	M_2	F_1	F_2
C_1	3/12	2/12	1/12	1/12
C_2	3/12	3/12	2/12	5/12
C_3	7/12	5/12	6/12	7/12
C_4	8/12	11/12	10/12	9/12

B.3 ANALYSIS OF EXPERIMENTAL DATA

The results of the experiments were analysed using the analysis of variance method [132].

As it is expected that the interaction effect of the coding schemes and the sentences used is negligible, the analysis of variance made use of the two factor model and assumed that the observation in the $(i,j)^{th}$ cell is

$$y_{ij} = \mu + (c_i - \mu) + (s_j - \mu) + e_{ij} \qquad i,j=1,2,3,4 \qquad (B.1)$$

That is, y_{ij} consists of i) a portion μ common to all observations; ii) the difference between the expected value c_i due to the i^{th} coding scheme and μ, iii) the difference between the expected value s_j due to the j^{th} sentence and μ, and iv) a deviation e_{ij} particular to the j^{th} individual of the i^{th} group. e_{ij} is sometimes referred to as the residual.

One of the assumptions of the analysis of variance method is that the variances of all the observations are identical. Each of the sixteen observations in Tables B.2 and B.3 is the number of preference votes out of a total of 12. It is reasonable to assume that the

preference vote in the $(i,j)^{th}$ cell follows a binomial distribution with index 12 and probability ϕ_{ij}, where i denotes the row and j the column. The expectation and variance of observations y_{ij} for each cell are therefore equal to ϕ_{ij} and $\phi_{ij}(1 - \phi_{ij})$ respectively. As the expectation of the cell values are different, so the variances cannot be assumed to be identical in general. However, these can be overcome by first transforming the data before analysis.

Several methods of transformation such as angular transformation [133] or logistic transformation [134], can be applied. The angular transformation was chosen for its simplicity. Specifically, the value of the proportion (m/M), was transformed into t where

$$t = \sin^{-1}\sqrt{(m/M)} \qquad (B.2)$$

The variances of the transformed data are then approximately constant, being equal to 1/(4M).

The transformed data for the two experiments are given in Tables B.4 and B.5.

Table B.4 Transformed Data of Experiment 1

Preference for	M_1	M_2	F_1	F_2
C_1	0.42053	0.29284	0.61548	0.42053
C_2	0.52360	0.70167	0.52360	0.61548
C_3	0.95532	0.86912	0.78540	0.86912
C_4	1.15026	1.27795	1.15026	1.27795

Table B.5 Transformed Data of Experiment 2

Preference for	M_1	M_2	F_1	F_2
C_1	0.52360	0.42053	0.29284	0.29284
C_2	0.52360	0.52360	0.42053	0.70167
C_3	0.86912	0.70167	0.78540	0.86912
C_4	0.95532	1.27795	1.15026	1.04720

For p coding schemes and q sentences, the analysis of variance table can be constructed as shown in the following table.

Table B.6 Analysis of Variance Table

Due to	Degree of Freedom (D.F.)	Sum of Squares (S.S.)
Mean	1	$SSM = (\sum_{i=1}^{p}\sum_{j=1}^{q} t_{ij})^2/(pq)$
Coding schemes	(p-1)	$\sum_{i=1}^{p} (\sum_{j=1}^{q} t_{ij})^2/q - SSM$
Sentences	(q-1)	$\sum_{j=1}^{q} (\sum_{i=1}^{p} t_{ij})^2/p - SSM$
Remainder	(p-1)(q-1)	By difference
Total	pq	$\sum_{i=1}^{p}\sum_{j=1}^{q} t_{ij}^2$

The mean square (M.S.) for each factor is then calculated as the ratio of the S.S. to the D.F. due to that factor. The various variance ratios (VR) are obtained by dividing the respective mean squares by the mean square due to the remainder. F-tests are then used to determine the significance of the various effects.

The ANOVA tables for the data shown in Tables B.4 and B.5 are given in Tables B.7 and B.8 respectively. In these tables, the non-additivity effects with one degree of freedom [135] were also extracted from the sums of squares due to remainder 1.

Table B.7 ANOVA Table for Experiment 1

Due to	D.F.	S.S.	M.S.	VR1	VR2
Mean	1	9.68631	9.68631		
Coding scheme	3	1.39835	0.46612	40.71	41.10
Sentence	3	0.00280	0.00093	0.081	0.082
Remainder 1	9	0.10305	0.01145		
Non-Additivity	1	0.01232	0.01232		1.086
Remainder 2	8	0.09073	0.01134		
Total about zero	16	11.19051			

Grand mean : 0.77807

Means for coding schemes

C_1:0.43735, C_2:0.59109 C_3:0.86974, C_4:1.21411

Means for sentences

M_1:0.76243, M_2:0.78540 F_1:0.76869, F_2:0.79577

Table B.8 ANOVA Table for Experiment 2

Due to	D.F.	S.S.	M.S.	VR1	VR2
Mean	1	8.05889	8.05889		
Coding scheme	3	1.21129	0.40376	25.42	24.37
Sentence	3	0.01238	0.00413	0.260	0.250
Remainder 1	9	0.14294	0.01588		
Non-Additivity	1	0.01038	0.01038		0.626
Remainder 2	8	0.13256	0.01657		
Total about zero	16	9.4255			

Grand mean : 0.70970

Means for coding schemes

C_1:0.38245, C_2:0.54235 C_3:0.80633, C_4:1.10768

Means for sentences

M_1:0.71791, M_2:0.73094 F_1:0.66226, F_2:0.72771

B.4 VERIFICATION OF ASSUMPTIONS

The following assumptions were made in the analysis of variance method.

(i) Homogeneity of variances;

(ii) Additivity of effects;

(iii) Independence of residuals and

(iv) Normality of distribution of residuals.

These assumptions can be checked by an appropriate analysis of the residuals [136]. Of these, the most important are the assumptions on Additivity and Independence. The inferences made from the ANOVA tables are not invalidated for slight departure from the Normality and Homogeneity of variance assumptions [132]. As described in Section B.3, the Homogeneity of variance is satisfied by using the transformed values instead of the original data in the form of proportions.

The additivity of effects can be checked using the method of isolating one degree of freedom for non-additivity as described by Tukey [135]. These procedures were carried out and the results are shown in Tables B.7 and B.8 for experiment 1 and 2 respectively. In Table B.7, the variance ratio for non-additivity is 0.909 and in Table B.8, it is 0.304. Both are not significant even at 50% level.

The Independence and Normality assumptions are next considered by examining and analysing the residuals of the individual cell value. For the two factor model, the residual for the $(i,j)^{th}$ cell, e_{ij}, is given by

$$e_{ij} = t_{ij} - \bar{t}_{ij} \tag{B.3}$$

where $\bar{t}_{ij}$ is the fitted value for the $(i,j)^{th}$ cell given by

$$\bar{t}_{ij} = c_i + s_j - \mu \tag{B.4}$$

These values are given in Tables B.9 and B.10 respectively for the data in Tables B.4 and B.5.

Table B.9 Residuals for Experiment 1

Preference for	M_1	M_2	F_1	F_2
C_1	-0.00117	-0.15183	0.18752	-0.03452
C_2	-0.05185	0.10326	-0.05810	0.00669
C_3	0.10122	-0.00795	-0.07496	-0.01832
C_4	-0.04820	0.05652	-0.05446	0.04614

Table B.10 Residuals for Experiment 2

Preference for	M_1	M_2	F_1	F_2
C_1	0.13294	0.01684	-0.04217	-0.10762
C_2	-0.02696	-0.03999	-0.07437	0.14132
C_3	0.05459	-0.12589	0.02651	0.04479
C_4	-0.16057	0.14904	0.09002	-0.07849

Figs.B.1(a) and B.1(b) are the graphs of the residuals plotted against the fitted values for Experiments 1 and 2 respectively. From these scatter plots, it can be seen that there is no conspicuous trend or dependence of the residuals on the fitted values.

To study the normality of the residuals, the method described by Chernoff and Lieberman [137] was adopted. For each experiment, the residuals are ordered in ascending magnitudes. These were then plotted against the values of unit-normal-cumulative corresponding to the probabilities $(i-0.5)/16$, for $i=1,\ldots,16$, as shown in Fig.B.2(a) for Experiment 1 and Fig.B.2(b) for Experiment 2.

If the underlying distribution of the residuals is normal, a straight line is expected. From Figs.B.2(a) and B.2(b), it can be seen that the points fit quite closely to the least mean square lines. Thus it can be said that the residuals are approximately normally distributed.

B.5 INTERPRETATION OF THE RESULTS

From standard F-tables for (3,9) degrees of freedom, the value at 50% level is 0.852, at 25% level is 1.63 and at 0.1% level is 13.90.

Comparing these values with the variance ratios (VR1) obtained in Table B.7, it can be seen that the effect of CVSD coding schemes at different bit rates on the subjective preference score is significant well below 0.1% level, while the effect of different sentences used is significant only at more than 25% level. For experiment 2, the results in Table B.8 again show that the effect of CVSD coding schemes at different bit rates on the subjective preference score is significant at less than 0.1% level. However, the effect of different sentences used is not significant even at 50% level.

Next, the confidence intervals for the percentage preference scores of the various coding schemes were determined as follows.

The distribution of m, the number of preference votes, can be assumed to be binomial with index M and parameter ϕ, where ϕ is the probability of success and M is the total number of available votes. For large M, the binomial distribution approaches Normal distribution with mean $M\phi$ and variance $M\phi(1-\phi)$, or equivalently, the

distribution of $\{(m-M\phi)/\sqrt{[M\phi(1-\phi)]}\}$ is approximately Normal with mean 0 and variance 1.

The confidence intervals of ϕ, at a specified level of significance, denoted by $[\phi_1, \phi_2]$, can then be obtained given m and M. For example, if the values for 90% confidence interval are desired, the value c for which

$$\text{Probability of}\{ -c \leq \{(m-M\phi)/\sqrt{[M\phi(1-\phi)]}\} \leq c \} = 0.90 \quad \text{(B.5)}$$

is first found from tables of Normal distribution, after which ϕ_1 and ϕ_2 can be obtained from

$$\phi_1 = \frac{M(2m+c^2)-\sqrt{M^2(2m+c^2)^2-4M(M+c^2)m^2}}{2M(M+c^2)} \quad \text{(B.6)}$$

and $$\phi_2 = \frac{M(2m+c^2)+\sqrt{M^2(2m+c^2)^2-4M(M+c^2)m^2}}{2M(M+c^2)} \quad \text{(B.7)}$$

The mean percentage preference scores for the CVSD coding schemes at the four rates over all the four sentences used and the 90% confidence intervals for these means were determined and given in Tables B.11 and B.12. They are also represented in graphical form in Figs.B.3(a) and B.3(b).

A mean preference score of 50% implies that the votes of the listener group are equally divided betweeen the two coding schemes under comparison and the two can then be considered as being comparable in the subjective quality of the speech reproduced.

From Fig.B.3(a), it can thus be inferred that the subjective quality of 7.2 kbit/s HFDC is comparable to 19 kbit/s CVSD. Statistically speaking, at 90% confidence level, the subjective quality of 7.2 kbit/s HFDC is better than CVSD at bit rates more than 17.8 kbit/s but worse than CVSD at bit rates more than 20.4 kbit/s.

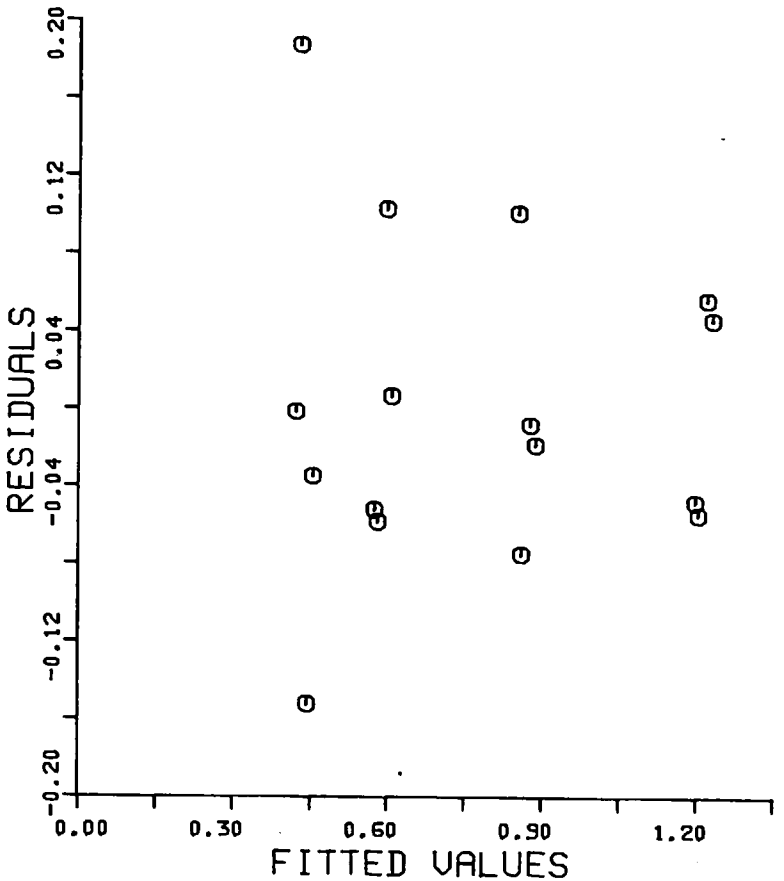
Similarly, from Fig.B.3(b), it can be said that the subjective quality of 9.6 kbit/s HFDC is comparable with approximately 23 kbit/s CVSD. Statistically speaking, at 90% confidence level, it is better than CVSD at less than 21.5 kbit/s but worse than CVSD at more than 24.1 kbit/s.

Table B.11 Percentage Preference Score for CVSD (Experiment 1: HFDC at 7.2 kbit/s)

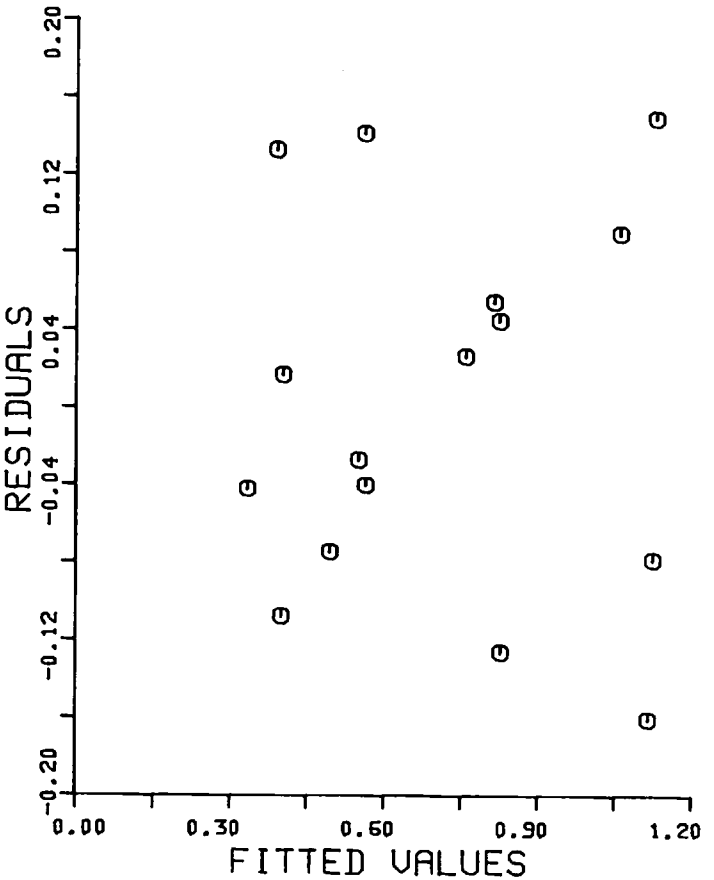
Bit Rate of CVSD (kbit/s)	mean ϕ	ϕ_1	ϕ_2
14	18.75	11.23	29.62
17	31.25	21.47	43.04
20	58.33	46.46	69.32
23	87.50	77.57	93.41

Table B.12 Percentage Preference Score for CVSD (Experiment 2: HFDC at 9.6 kbit/s)

Bit Rate of CVSD (kbit/s)	mean ϕ	ϕ_1	ϕ_2
17	14.58	8.09	24.88
20	27.08	17.95	38.68
23	52.08	40.40	63.55
26	79.16	68.06	87.14



(a)



(b)

Fig.B.1 Scatter plots for residuals; (a) Experiment 1; (b) Experiment 2.

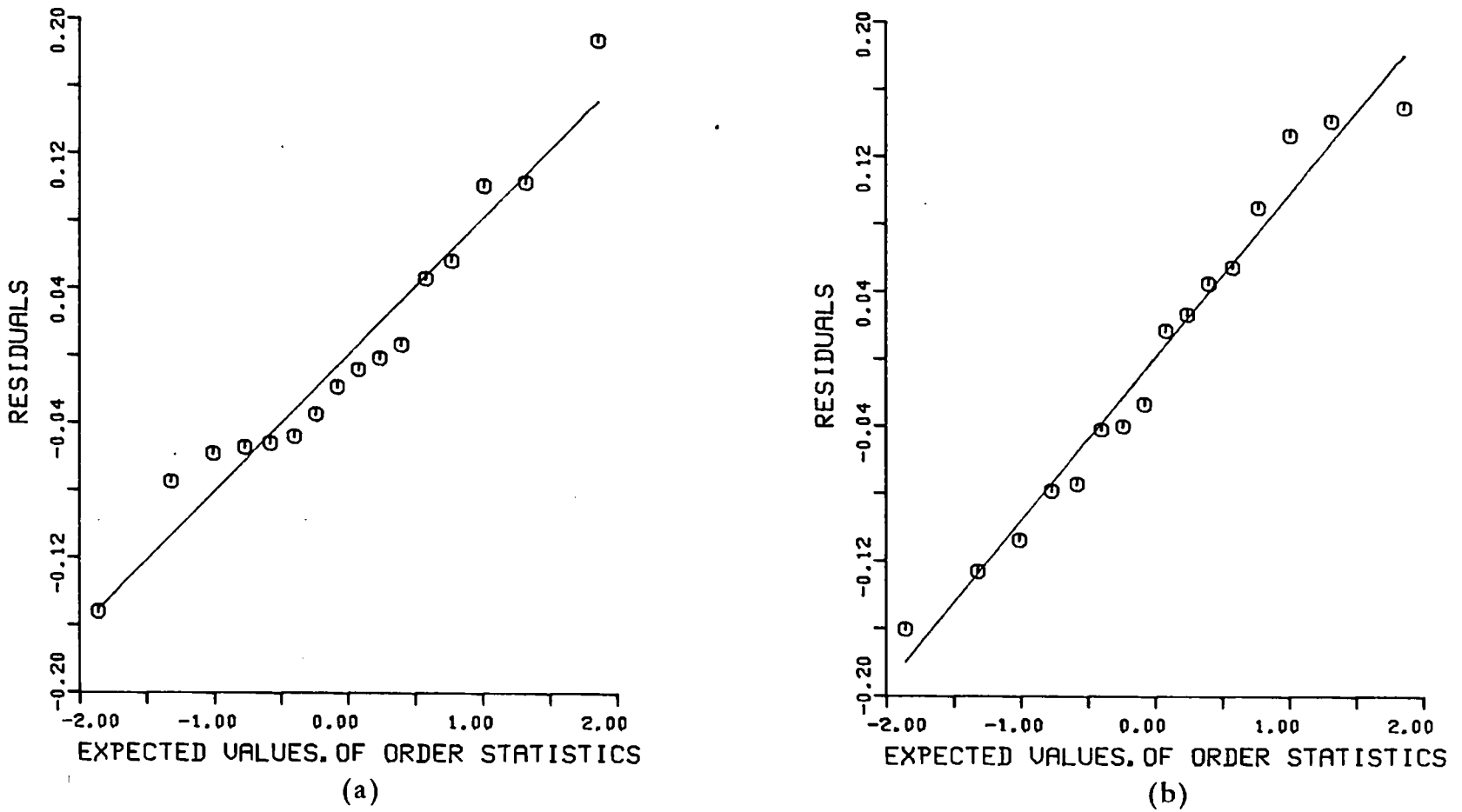
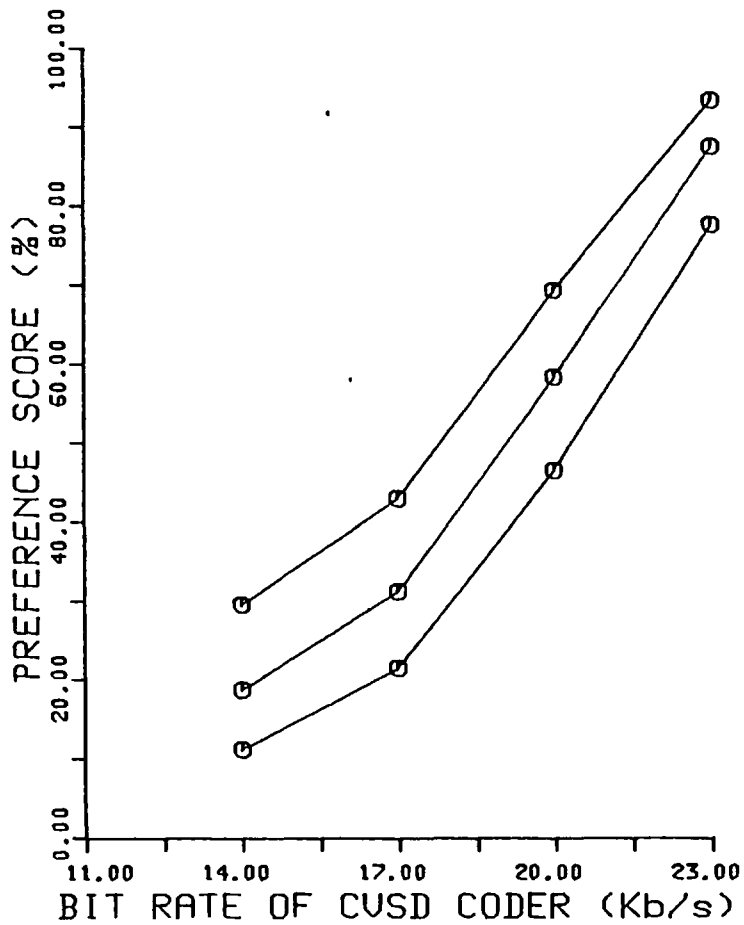
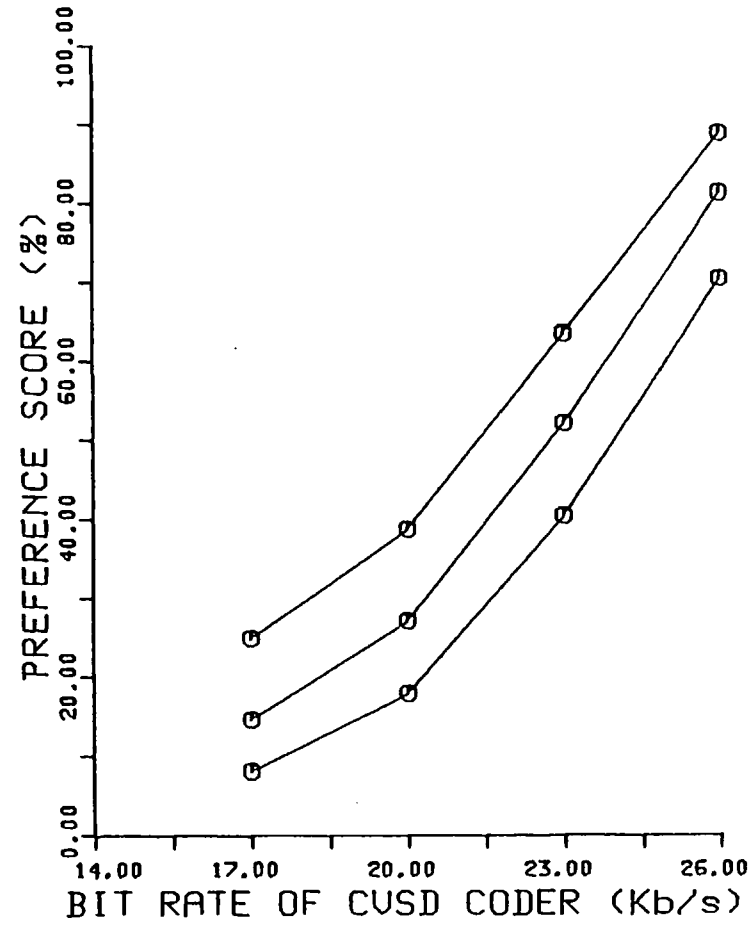


Fig.B.2 Normal plots for residuals; (a) Experiment 1; (b) Experiment 2.



(a)



(b)

Fig.B.3 Subjective preference of quality of speech produced by CVSD against that produced by HFDC with the 90% confidence limits imposed; (a) Experiment 1; (b) Experiment 2.

APPENDIX C

SUBJECTIVE TESTS OF VARIABLE-RATE HFDC

C.1 DESIGN OF EXPERIMENTS

One of the objectives of the experiments is to obtain an indication of the improvement in the subjective quality of speech that can be obtained over the fixed-rate hybrid frequency domain coding technique by using the variable-rate hybrid frequency domain coding technique.

The subjective quality of speech produced by fixed-rate HFDC when the voiced and unvoiced segments are coded differently and that produced by fixed-rate HFDC with no distinction in coding the voiced and unvoiced segments are also compared.

Two experiments were carried out. A forced decision paired-comparison method was adopted. For both experiments, comparisons were made between fixed-rate HFDC with distinction between voiced and unvoiced segments (C_0) and the following schemes

C_1: Variable-rate HFDC with K_u and K_v set to be 5;

C_2: Variable-rate HFDC with K_u and K_v set to be 10;

C_3: Variable-rate HFDC with K_u and K_v set to ∞,

C_4: Fixed-rate HFDC with no distinction in coding the voiced and unvoiced segments.

The two experiments differ only in the average transmission bit rate chosen. The average bit rate of 7.2 kbit/s was used for Experiment 1 and for Experiment 2, 9.6 kbit/s was used.

Two sentences by male speakers (M_1 and M_2) and one sentence by a female speaker (F_1) were used. The sentences are comprised of radio announcements which lasted from 17 seconds to 24 seconds.

As C_0 is to be compared with C_1, C_2, C_3 and C_4 using 3 different sentences (M_1, M_2 and F_1), a total of 12 comparisons was to be made. Each combination of (sentence, coder pair) was assigned a number from 1 to 12. The order of presentation of the (sentence, coder pair) to the listeners was based on a sequence of randomly generated numbers from 1 to 12. As two orders of presentation were used for each experiment, so four different sequences were generated. For further randomization, the order of the coder pair was reversed for one presentation in each experiment. The actual arrangements used are set out in Table C.1.

Each comparison was recorded twice on the analog magnetic tapes with less than 1 second interval between sentences of the same comparison and 15 second interval between comparisons.

Table C.1 Arrangements of Sentences on Tapes

Number	Experiment 1 (7.2 kbit/s)		Experiment 2 (9.6 kbit/s)	
	Tape A	Tape B	Tape C	Tape D
1	M_2,C_2,C_0	F_1,C_0,C_1	M_2,C_0,C_1	M_1,C_0,C_2
2	M_1,C_0,C_3	M_2,C_0,C_2	F_1,C_0,C_4	M_2,C_1,C_0
3	F_1,C_2,C_0	F_1,C_0,C_3	M_1,C_4,C_0	F_1,C_2,C_0
4	M_2,C_3,C_0	M_1,C_1,C_0	F_1,C_0,C_1	M_1,C_0,C_3
5	F_1,C_3,C_0	M_2,C_0,C_1	M_2,C_0,C_3	M_2,C_4,C_0
6	M_1,C_0,C_1	F_1,C_0,C_4	M_1,C_1,C_0	F_1,C_1,C_0
7	F_1,C_4,C_0	M_1,C_4,C_0	M_2,C_0,C_2	M_2,C_3,C_0
8	M_2,C_1,C_0	M_2,C_0,C_3	F_1,C_0,C_3	F_1,C_3,C_0
9	F_1,C_1,C_0	M_1,C_2,C_0	M_1,C_3,C_0	M_1,C_0,C_4
10	M_1,C_0,C_2	M_2,C_0,C_4	M_2,C_0,C_4	M_2,C_2,C_0
11	M_2,C_4,C_0	F_1,C_0,C_2	F_1,C_0,C_2	F_1,C_4,C_0
12	M_1,C_0,C_4	M_1,C_3,C_0	M_1,C_2,C_0	M_1,C_0,C_1

The sentences used were:

M_1 : Soviet radio has described Sadat's assassination as the results of opposition among ordinary Egyptians to his rule and his peace agreement with Israel; but in Peking, his death was described as a tremenduous loss. Mrs. Thatcher, who was in Melbourne, said she's sure Israel and Egypt will strive even harder for peace, and she described vice president Husni Mubarak as a very able and experienced man.

M_2 : Last week the pound rose quite steeply, seven or eight cents to its highest level for three months supported by Britain's high interest rate and the resurgence of oil as a potential important factor. Today sterling was down again even though our interest rate remains high and things in the Middle East remain tense.

F_1 : There's intense anxiety too about the so-called Laser beam approach adopted by the Japanese to penetrate and lay waste certain targeted area in the British market. Through the sixties and seventies, the motorcycle market was a classic case. The Japanese now provides more than ninety two percent of all machines sold in this country.

C.2 CONDUCT OF EXPERIMENT

Twelve subjects for the experiments were chosen from among the undergraduate and postgraduate students of the Department of Electrical Engineering. All of them have English as their mother tongue. Six subjects listened to Tape A and Tape C, and six others listened to Tape B and Tape D. The listeners were seated in a quiet room and were asked to listen to the tapes through the headphones. They were instructed to indicate, for each comparison (sentence A, sentence B), the sentence they preferred.

The results were compiled and presented in Tables C.2 and C.3. The data are shown in the format (m/M) where M denotes the total votes

available per cell and m is the number of votes given in favour of the coding scheme specified.

Table C.2 Results of Experiment 1

Preference for	M_1	M_2	F_1
C_1	8/12	7/12	6/12
C_2	8/12	8/12	9/12
C_3	10/12	9/12	9/12
C_4	4/12	3/12	5/12

Table C.3 Results for Experiment 2

Preference for	M_1	M_2	F_1
C_1	7/12	8/12	6/12
C_2	7/12	9/12	8/12
C_3	9/12	9/12	8/12
C_4	3/12	6/12	4/12

C.3 ANALYSIS OF EXPERIMENTAL DATA

The results of the experiments for the comparisons of variable-rate HFDC (C_1-C_3) and fixed-rate HFDC (C_0) were analysed separately from the comparison of C_4 and C_0.

To find out the effect of the different variable-rate HFDC schemes on the results, the analysis of variance method was used.

As detailed in Appendix B, to satisfy the homogeneity of variance assumption, the data were first transformed using the angular transformation

$$t = \sin^{-1}\sqrt{(m/M)} \qquad (C.1)$$

The transformed values for the two experiments are tabulated in Tables C.4 and C.5.

Table C.4 Transformed Data for Experiment 1

Preference for	M_1	M_2	F_1
C_1	0.95532	0.86912	0.78540
C_2	0.95532	0.95532	1.04720
C_3	1.15026	1.04720	1.04720

Table C.5 Transformed Data for Experiment 2

Preference for	M_1	M_2	F_1
C_1	0.86912	0.95532	0.78540
C_2	0.86912	1.04720	0.95532
C_3	1.04720	1.04720	0.95532

For the two factor model, the transformed value in each cell can be represented as a summation of the fitted value $\bar{t}_{ij}$ and the residual e_{ij}

where
$$t_{ij} = \bar{t}_{ij} + e_{ij} \qquad (C.2)$$

$$\bar{t}_{ij} = c_i + s_j - \mu \qquad i,j=1,2,3 \qquad (C.3)$$

in which μ denotes the grand mean, c_i denotes the effect due to the i^{th} coding schemes and s_j denotes the effect due to the j^{th} sentence.

The ANOVA tables for the data shown in Tables C.4 and C.5 are given in Tables C.6 and C.7 respectively. In these tables, the sum of squares due to the non-additivity effect [135] extracted from that due to remainder 1 is also given. The variance ratios VR1 are found by dividing the M.S. due to the specified effect by the M.S. due to

remainder 1. The variance ratios VR2 are similarly determined using M.S. due to remainder 2 in place of M.S. due to remainder 1.

Table C.6 ANOVA Table for Experiment 1

Due to	D.F.	S.S.	M.S.	VR1	VR2
Mean	1	8.62857	8.62857		
Coding scheme	2	0.06737	0.03369	6.904	5.272
Sentence	2	0.00763	0.00382	0.783	0.598
Remainder 1	4	0.01952	0.00488		
Non-Additivity	1	0.00035	0.00035		0.0548
Remainder 2	3	0.01917	0.00639		
Total about zero	9	8.72309			

Grand mean = 0.97915

Means for coding schemes C_1:0.86995, C_2:0.98595, C_3:1.08155

Means for sentences M_1:1.02030, M_2:0.95721, F_1:0.95993

Table C.7 ANOVA Table for Experiment 2

Due to	D.F.	S.S.	M.S.	VR1	VR2
Mean	1	8.08679	8.08679		
Coding scheme	2	0.03264	0.01632	4.872	4.280
Sentence	2	0.02255	0.01128	3.367	2.958
Remainder 1	4	0.01338	0.00335		
Non-Additivity	1	0.00194	0.00194		0.509
Remainder 2	3	0.01144	0.003813		
Total about zero	9	8.15536			

Grand mean : 0.94791

Means for coding schemes C_1:0.86995, C_2:0.95721, C_3:1.01657

Means for sentences M_1:0.92848, M_2:1.01657, F_1:0.89868

C.4 VERIFICATION OF ASSUMPTIONS

The method of analysis of variance is derived based on the assumptions of homogeneity of variance, the additivity of the main effects, and the independence and normality of the residuals. The statistical analysis of the experimental results will not be complete without verifying whether the assumptions are indeed satisfied.

The homogeneity of variance assumption is satisfied through the use of the angular transformation of the original data.

Using the method of isolating one degree of freedom for non-additivity as described by Tukey [135], the variance ratios given in Tables C.6 and C.7 show that the effects of non-additivity are not significant for both experiments.

The residuals, computed as the differences between the actual values and the fitted values, using the data in Tables C.4 and C.5 are given in Tables C.8 and C.9 respectively.

Table C.8 Residuals for Experiment 1

Preference for	M_1	M_2	F_1
C_1	0.04422	0.02111	-0.06533
C_2	-0.07178	-0.00869	0.08047
C_3	0.02756	-0.01242	-0.01514

Table C.9 Residuals for Experiment 2

Preference for	M_1	M_2	F_1
C_1	0.01861	0.01671	-0.03532
C_2	-0.06866	0.02132	0.04734
C_3	0.05006	-0.03803	-0.01202

The scatter plots of the residuals against the fitted values are shown in Figs.C.1(a) and C.1(b) for Experiments 1 and 2 respectively. Inspection of these plots show that there is no discernable trend or dependence of the residuals on the fitted values.

The normality of residuals can be checked using the method described by Chernoff and Lieberman [137] where the residuals, ordered in ascending magnitudes, are plotted against the expected values of order statistics $Gau^{-1}[(i-0.5)/9]$ where Gau(*) denotes the unit-normal-cumulative. These are shown in Fig.C.2(a) for Experiment 1 and Fig.C.2(b) for Experiment 2. From these figures, it can be seen that the points fall more or less around the least square lines for both experiments. Considering the fact that only 9 points are available, it can be said that the residuals are approximately normally distributed.

C.5 INTERPRETATION OF THE RESULTS

The F-value for 2 and 4 degrees of freedom as found from standard statistical table is 0.828 at 50% level, 2.00 at 25% level, 4.32 at 10% level and 6.94 at 5% level. From the ANOVA table for Experiment 1, the variance ratios VR1 show that the effect of using different values of K_u and K_v is significant at slightly above 5% level, while the effect of different sentences used is significant

only at more than 50% level. For Experiment 2, the results show that the effect of using different values of K_u and K_v is not significant at 5% level but at 10% confidence level, it becomes significant. On the other hand, the effect of sentences, is significant only at 25% level.

Assuming that the distribution of the true percentage score is normal, and using the method detailed in Appendix B, the confidence interval at a given significant level can be computed. The 90% confidence intervals so determined together with the mean percentage preference scores for the four coding schemes (C_1-C_4) over all the three sentences are given in Tables C.10 and C.11.

Table C.10 Percentage Preference Score for Experiment 1

Preference for	mean ϕ	ϕ_1	ϕ_2
Variable-rate HFDC with $K_u=K_v=5$	58.33	44.70	70.80
Variable-rate HFDC with $K_u=K_v=10$	69.44	55.83	80.34
Variable-rate HFDC with $K_u=K_v=\infty$	77.78	64.67	87.00
Fixed-rate HFDC with no UV/V distinction	33.33	21.98	47.02

Table C.11 Percentage Preference Score for Experiment 2

Preference for	mean ϕ	ϕ_1	ϕ_2
Variable-rate HFDC with $K_u=K_v=5$	58.33	44.70	70.80
Variable-rate HFDC with $K_u=K_v=10$	66.67	52.98	78.02
Variable-rate HFDC with $K_u=K_v=\infty$	72.22	58.73	82.61
Fixed-rate HFDC with no UV/V distinction	36.11	24.35	49.82

A mean preference score of 50% implies that the votes of the listener group are equally divided between the two schemes under comparison. If the preference score is more than 50%, it indicates that the particular scheme is much preferred. Judging from the results, it can be concluded that there is a significant improvement in the subjective quality of speech obtained using the variable-rate HFDC over the fixed-rate scheme, especially when the values of K_u and K_v are set to be equal to or larger than 10. Furthermore, the fixed-rate HFDC with distinction between voiced and unvoiced segments (C_0) is much preferred to the fixed-rate HFDC with no distinction in coding the voiced and unvoiced segments (C_4).

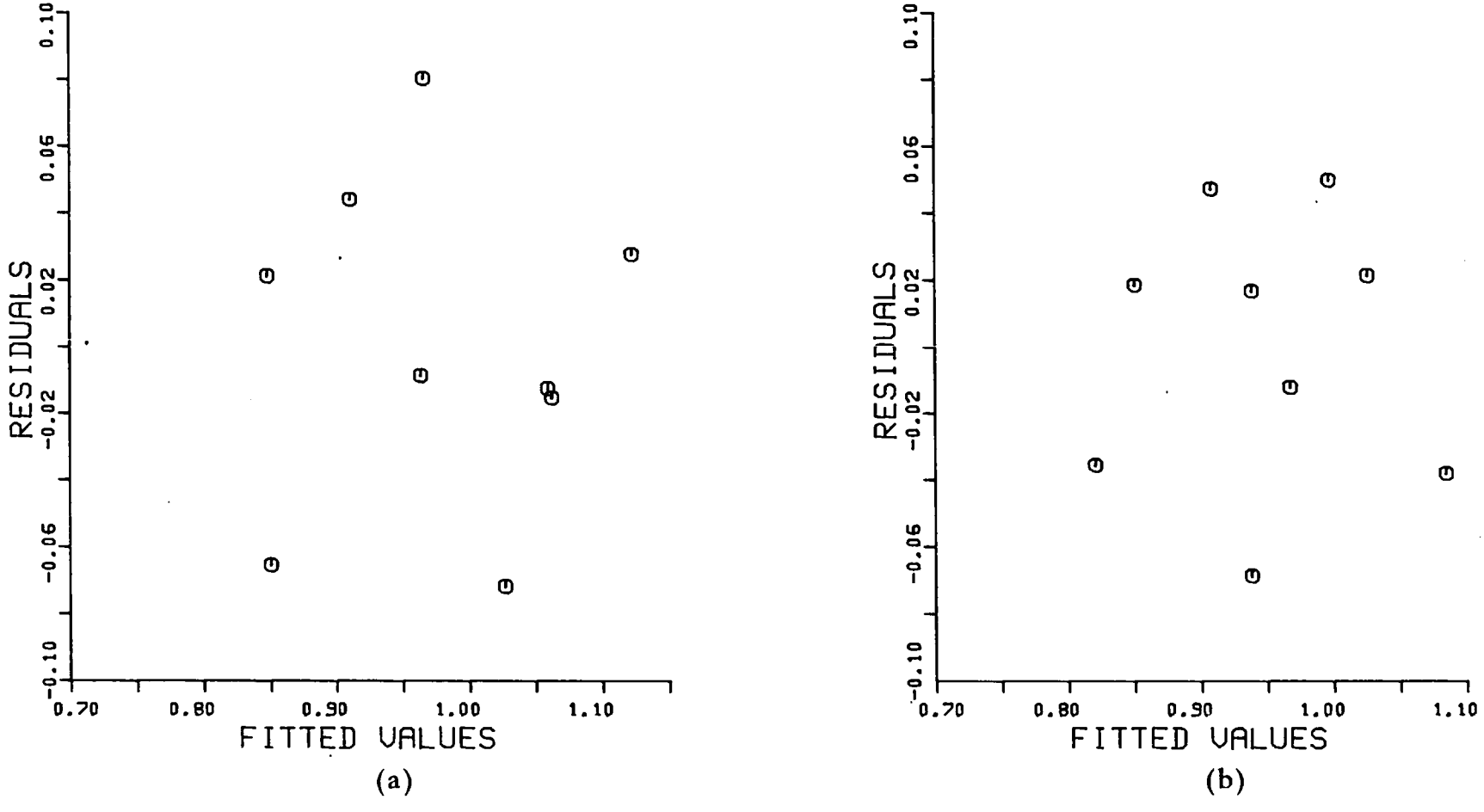
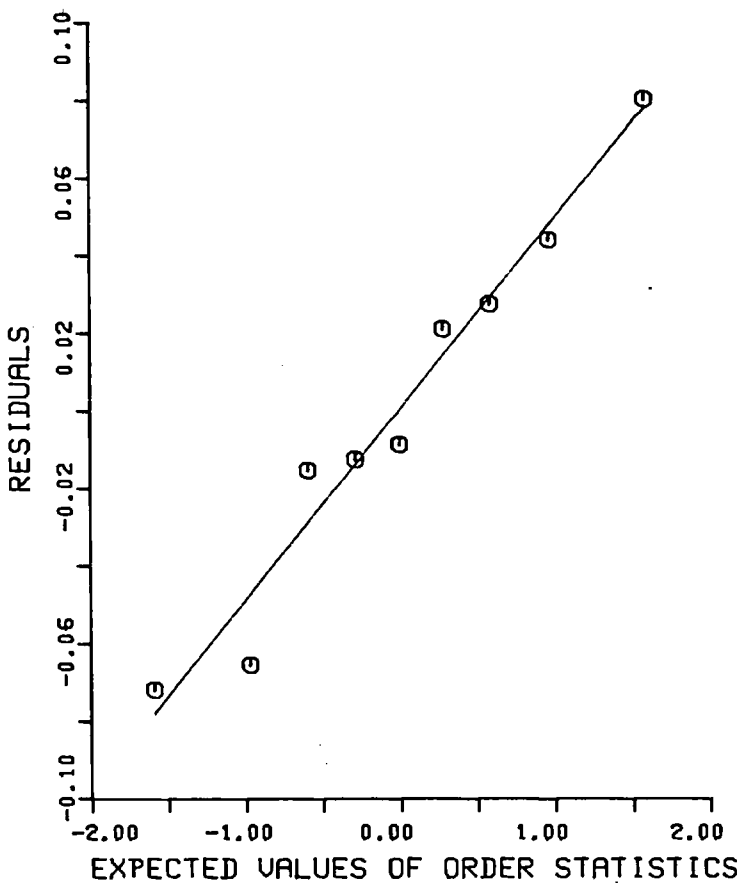
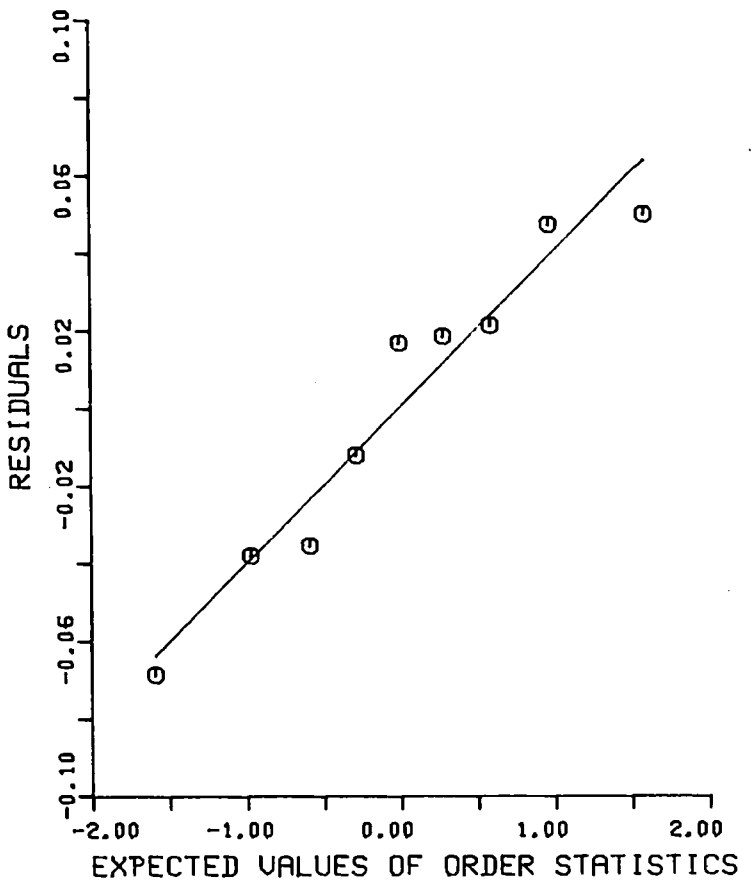


Fig.C.1 Scatter plots for residuals; (a) Experiment 1; (b) Experiment 2.



(a)



(b)

Fig.C.2 Normal plots for residuals; (a) Experiment 1; (b) Experiment 2.

APPENDIX D

NON-RECURSIVE FILTERS OF EVEN ORDER

For symmetrical FIR filters of even order, the impulse response coefficients are related by,

$$h(n) = h(N-1-n) \qquad 0 \leq n \leq (N/2)-1 \qquad (D.1)$$

Under this condition, the frequency response of the filter is given by

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n}$$

$$= \sum_{n=0}^{N/2-1} h(n)e^{j\omega n} + \sum_{m=N/2}^{N-1} h(m)e^{-j\omega m}$$

By putting m = (N-1-n), we have,

$$H(e^{j\omega}) = \sum_{n=0}^{N/2-1} h(n)e^{-j\omega n} + \sum_{n=N/2-1}^{0} h(N-1-n)e^{-j\omega(N-1-n)}$$

$$= \sum_{n=0}^{N/2-1} [h(n)e^{-j\omega n} + h(N-1-n)e^{-j\omega(N-1-n)}]$$

$$= \sum_{n=0}^{N/2-1} h(n)[e^{-j\omega n} + e^{-j\omega(N-1-n)}]$$

$$= e^{-j\omega(N-1)/2} \sum_{n=0}^{N/2-1} 2h(n)\cos[\omega((N-1)/2)-n)] \qquad (D.2)$$

To reduce the argument of the cosine term into a form independent of N, let

$$b(n) = 2h((N/2)-n) \qquad n=1,2,\ldots,N/2 \tag{D.3}$$

then $$H(e^{j\omega}) = e^{-j\omega(N-1)/2}H^{*}(e^{j\omega}) \tag{D.4}$$

where $$H^{*}(e^{j\omega}) = \sum_{n=1}^{N/2} b(n)\cos[\omega(n-(1/2)] \tag{D.5}$$

It should be noted that $H^{*}(e^{j\omega})$ is a real function taking on both positive and negative values, and the magnitude of $H(e^{j\omega})$ is the same as that of $H^{*}(e^{j\omega})$. Furthermore, at $\omega=\pi$, $H^{*}(e^{j\omega}) = 0$ and thus $H(e^{j\omega})$ is independent of b(n) or h(n). This implies that the filters with a frequency response that is nonzero at $\omega=\pi$, (e.g. a highpass filter) cannot be satisfactorily designed with this technique.

APPENDIX E

THE ANALYSIS OF THE ERROR $\mathcal{E}_{cr}$

Let $$\mathcal{E}_{c1}(z) = \sum_{k=0}^{2N-2} \epsilon_{c1}(k)z^{-k} = H_l(z)E_p(z) \tag{E.1}$$

Then since $$H_l(z)E_p(z) = \Big[\sum_{i=0}^{N-1} h_l(i)z^{-i}\Big]\Big[\sum_{j=0}^{N-1} e_p(j)z^{-j}\Big]$$

$$= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} h_l(i)e_p(j)z^{-i-j} \tag{E.2}$$

it follows that $$\epsilon_{c1}(k) = \sum_{i,j} h_l(i)e_p(j) \tag{E.3}$$

where the summation is over all i, j such that $0 \leqslant i \leqslant N-1$, $0 \leqslant j \leqslant N-1$ and $i+j = k$.

In general, for $0 \leqslant k \leqslant N-1$,

$$\mathcal{E}_{c1}(k) = \sum_{m=0}^{k} h_l(m)e_p(k-m) \tag{E.4a}$$

and for $N-1 \leqslant k \leqslant 2N-2$,

$$\epsilon_{c1}(k) = \sum_{m=0}^{2N-2-k} h_l(N-1-m)e_p(k-N+1+m) \tag{E.4b}$$

The following points should be noted

(i) $\mathcal{E}_{c1}(z)$ contains an odd number (2N-1) terms.

(ii) Since $h_l(n) = h_l(N-1-n)$, $n = 0,1,\ldots,N-1$ is symmetrical, it follows that $\epsilon_{c1}(k) = \epsilon_{c1}(2N-2-k)$, $k = 0,1,\ldots,2N-2$,

and (iii) $\epsilon_{c1}(N-1)$ has the maximum number of terms of the form $h_l(*)e_p(*)$ of all the $\epsilon_{c1}(k)$, $k = 0,1,\ldots,2N-2$.

From (E.4a), it can be seen that

$$|\epsilon_{c1}(k)| \leq \sum_{m=0}^{k} |h_l(m)||e_p(k-m)| < (q/2) \sum_{m=0}^{k} |h_l(m)|$$

$$\leq \begin{cases} (q/2)(1/2)=q/4 & \text{for} \quad 0 \leq k \leq (N/2)-1 \quad \text{(E.5a)} \\ (q/2)(1)=q/2 & \text{for} \quad (N/2) \leq k \leq N-1 \quad \text{(E.5b)} \end{cases}$$

and similarly, from (E.4b)

$$|\epsilon_{c1}(k)| \leq \sum_{m=0}^{2N-2-k} |h_l(N-1-m)||e_p(k-N+1+m)| < (q/2) \sum_{m=0}^{2N-2-k} |h_l(N-1-m)|$$

$$\leq \begin{cases} (q/2)(1)=q/2 & \text{for} \quad N-1 \leq k \leq (3N/2)-2 \quad \text{(E.5c)} \\ (q/2)(1/2)=q/4 & \text{for} \quad (3N/2)-1 \leq k \leq 2N-2 \quad \text{(E.5d)} \end{cases}$$

The mean of $\mathcal{E}_{c1}$ is given by

$$\overline{\mathcal{E}}_{c1} = \left(\sum_{k=0}^{2N-2} \overline{\epsilon_{c1}(k)}\right)/(2N-1)$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} h_l(i)\overline{e_p(j)} = 0 \qquad \text{(E.6)}$$

and its variance is given by

$$\text{var}(\mathcal{E}_{c1}) = [1/(2N-1)] \sum_{k=0}^{2N-2} \overline{\epsilon_{c1}^2(k)}$$

$$= [1/(2N-1)] \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} h_l^2(i)\overline{e_p^2(j)}$$

$$= [1/(2N-1)](N\eta^2) \sum_{i=0}^{N-1} h_l^2(i)$$

$$< [N/2(2N-1)]\eta^2 \qquad \text{(E.7)}$$

Similar results can be derived for the term

$$\mathcal{E}_{c2}(z) = H_l(-z)E_p(-z)$$

In the case of $\mathcal{E}_{c3}(z)$, on using

$$\mathcal{E}_{c3}(z) = \sum_{k=0}^{2N-2} \epsilon_{c3}(k) z^{-k} = P_l(z) E_h(z) \tag{E.8}$$

and applying similar arguments to those used in connection with $\mathcal{E}_{c1}(z)$, it can be seen that

$$|\epsilon_{c3}(k)| \leqslant \sum_{m=0}^{k} |p_l(m)||e_p(k-m)| < (q/2) \sum_{m=0}^{k} |p_l(m)|$$

$$\leqslant \begin{cases} (q/2)(1)=q/2 & \text{for} \quad 0 \leqslant k \leqslant (N/2)-1 \qquad \text{(E.9a)} \\ (q/2)(2)=q & \text{for} \quad (N/2) \leqslant k \leqslant N-1 \qquad \text{(E.9b)} \end{cases}$$

and $$|\epsilon_{c3}(k)| \leqslant \sum_{m=0}^{2N-2-k} |p_l(N-1-m)||e_p(k-N+1+m)| < (q/2) \sum_{m=0}^{2N-2-k} |p_l(N-1-m)|$$

$$\leqslant \begin{cases} (q/2)(2)=q & \text{for} \quad N-1 \leqslant k \leqslant (3N/2)-2 \qquad \text{(E.9c)} \\ (q/2)(1)=q/2 & \text{for} \quad (3N/2)-1 \leqslant k \leqslant 2N-2 \qquad \text{(E.9d)} \end{cases}$$

The mean of the random variable $\mathcal{E}_{c3}$ is zero, its variance is given by

$$\text{var}(\mathcal{E}_{c3}) = [1/(2N-1)] \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p_l^2(i) \overline{e_p^2(j)}$$

$$= [1/(2N-1)](\eta^2)(N) \sum_{i=0}^{N-1} p_l^2(i)$$

$$\leqslant [1/(2N-1)](\eta^2)(N)(2)$$

$$= [2N/(2N-1)](\eta^2) \tag{E.10}$$

Similar results can be obtained for the term

$$\mathcal{E}_{c4}(z) = P_l(-z) E_h(-z) \tag{E.11}$$

Next, if $\mathcal{E}_{c5}(z) = \sum_{k=0}^{2N-2} \epsilon_{c5}(k)z^{-k} = E_h(z)E_p(z)$ (E.12)

then since $E_h(z)E_p(z) = [\sum_{i=0}^{N-1} e_h(i)z^{-i}][\sum_{j=0}^{N-1} e_p(j)z^{-j}]$

$$= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} e_h(i)e_p(j)z^{-i-j}$$

it follows that

$$\epsilon_{c5}(k) = \sum_{m=0}^{k} e_h(m)e_p(k-m) \qquad \text{for} \quad 0 \leqslant k \leqslant N-1 \qquad \text{(E.13a)}$$

$$= \sum_{m=0}^{2N-2-k} e_h(N-1-m)e_p(k-N+1+m) \quad \text{for} \quad N-1 \leqslant k \leqslant 2N-2 \qquad \text{(E.13b)}$$

and hence $|\epsilon_{c5}(k)| \leqslant \sum_{m=0}^{k} |e_h(m)||e_p(k-m)|$

$$< (q^2/4)\sum_{m=0}^{k} (1)$$

$$= (q^2/4)(k+1) \qquad \text{for} \quad 0 \leqslant k \leqslant N-1 \qquad \text{(E.14a)}$$

Similarly, $|\epsilon_{c5}(k)| < (q^2/4)(2N-1-k)$ for $N-1 \leqslant k \leqslant 2N-2$ (E.14b)

If E_h and E_p are assumed to be independent, rather than just uncorrelated, then the mean of the random variable E_hE_p is equal to the product of the means of E_h and E_p and is thus zero. With independent E_h and E_p, the variance is then equal to the product of the variances of E_h and E_p and is thus equal to $(\eta^2)^2 = \eta^4$.

Hence $\mathcal{E}_{c5}$ has zero mean, and a variance

$$\text{var}(\mathcal{E}_{c5}) = [1/(2N-1)] \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \overline{e_h^2(i)} \overline{e_p^2(j)}$$

$$= [1/(2N-1)](\eta^4)(N.N)$$

$$= [N^2/(2N-1)](\eta^4) \qquad \text{(E.15)}$$

Similar results can be obtained for

$$\mathcal{E}_{c6}(z) = E_h(-z)E_p(-z) \qquad \text{(E.16)}$$

APPENDIX F

EXPRESSION OF $W^2(\omega)$ IN CLOSED FORM

The following two identities will be used in the derivation:

$$(I1)\quad \cos^2 A = \frac{1}{2}[1 + \cos 2A]$$

and

$$(I2)\quad 2\cos A \sin B = \sin(A+B) - \sin(A-b)$$

Multiplying $(4N-3)W^2(\omega)$ by $\sin\omega$, then

$$(4N-3)W^2(\omega)\sin\omega = (1 + \sum_{n=1}^{N-1} 4\cos^2 n\omega)\sin\omega$$

$$= [1 + \sum_{n=1}^{N-1} 2(1+\cos 2n\omega)]\sin\omega$$

$$= \sin\omega + \sum_{n=1}^{N-1} (2\sin\omega + 2\cos 2n\omega \sin\omega)$$

$$= \sin\omega + \sum_{n=1}^{N-1} 2\sin\omega + \sum_{n=1}^{N-1} [\sin(2n+1)\omega - \sin(2n-1)\omega]$$

$$= \sin\omega + (N-1)2\sin\omega + \sum_{n=1}^{N-2} \sin(2n+1)\omega + \sin[2(N-1)+1]\omega$$

$$-\sin(2-1)\omega - \sum_{n=2}^{N-1} \sin(2n-1)\omega$$

$$= 2(N-1)\sin\omega + \sin(2N-1)\omega + \sum_{n=1}^{N-2} \sin(2n+1)\omega - \sum_{n=2}^{N-1} \sin(2n-1)\omega$$

Let $m = n - 1$, then the last term in the above expression becomes $\sum_{m=1}^{N-2} \sin(2m+1)\omega$ and hence

$$W^2(\omega) = [2(N-1)\sin\omega + \sin(2N-1)\omega]/[(4N-3)\sin\omega]$$

$$= \{2(N-1) + [\sin(2N-1)\omega/\sin\omega]\}/(4N-3)$$

By applying l'Hospital's rule, in the limit when $\omega = 0$, we have

$$\lim_{\omega \to 0} [\sin(2N-1)\omega/\sin\omega] = \lim [(2N-1)\cos(2N-1)\omega]/\cos\omega = (2N-1)$$

Thus, $$W^2(0) = W^2(\pi) = [2(N-1) + (2N-1)]/(4N-3) = 1$$

Also, for large N and for ω in the range $0 < \omega < \pi$,

$$\lim_{N \to \infty} W^2(\omega) = \lim_{N \to \infty} \{\{2[1-(1/N)] + [\sin(2N-1)\omega/N\sin\omega]\}/[4-(3/N)]\}$$
$$= \frac{1}{2}$$

APPENDIX G

CHARACTERISTICS OF QUADRATURE MIRROR FILTERS USED IN THE SIMULATION

The following filters are designed using the method proposed in Chapter 5. The grid density used is 12. Nominal passband response is 0 dB. Only the low pass responses, $\{h_1(k), k=1,\ldots,N\}$, are given as the response of the other filters in the split-band system can readily be derived from knowledge of the low pass responses.

(1) Length of filter (N) : 8

Normalized passband cutoff frequency : 14/96 = 0.1458

Normalized stopband cutoff frequency : 44/96 = 0.4583

h(0) = +.5492098065813584706 2E-02 h(1) = -.6444113308312221180 9E-01

h(2) = +.6126546715846820949 0E-01 h(3) = +.4945079019341864152 9E+00

Stopband attenuation : .6351331849308004636 8E-02 = -43.94 dB

Maximum normalized peak-to-peak ripple of the system response : 0.0324388426

(2) Length of filter (N) : 16

Normalized passband cutoff frequency : 38/192 = 0.1979

Normalized stopband cutoff frequency : 68/192 = 0.3542

h(0) = +.1203698623297996026 2E-02 h(1) = -.1394718391004968998 2E-01

h(2) = +.8687845275095119732 5E-03 h(3) = +.3954135313549292352 9E-01

h(4) = -.2012380039365773995 0E-01 h(5) = -.9716653133911092190 7E-01

h(6) = +.1102449307216510999 2E+00 h(7) = +.4762772400176805068 0E+00

Stopband attenuation : .6203017234372627187 0E-02 = -44.15dB

Maximum normalized peak-to-peak ripple of the system response : 0.0342363782

(3) Length of filter (N) : 32

Normalized passband cutoff frequency : 86/384 = 0.2240

Normalized stopband cutoff frequency : 116/384 = 0.3021

h(0) = +.537700360614379350 40E-03 h(1) = -.543743124017354153 46E-02

h(2) = -.790287018440096224 02E-03 h(3) = +.819814639027282437 69E-02

h(4) = -.250204666216893146 79E-03 h(5) = -.141235288851787760 32E-01

h(6) = +.283297730054966983 88E-02 h(7) = +.225912388682262836 81E-01

h(8) = -.836158265001390885 30E-02 h(9) = -.349129552655987885 86E-01

h(10) = +.194955166034240588 72E-01 h(11) = +.549373206970827238 37E-01

h(12) = -.442654080155410340 44E-01 h(13) = -.996266281469382751 04E-01

h(14) = +.132562336652355748 77E+00 h(15) = +.464066967249628842 17E+00

Stopband attenuation : .509164353189356523 83E-02 = -45.86dB

Maximum normalized peak-to-peak ripple of the system response : 0.0310370518