



הטכניון – מכון טכנולוגי לישראל
Technion – Israel Institute of Technology

ספריות הטכניון
The Technion Libraries

בית הספר ללימודי מוסמכים ע"ש ארווין וג'ואן ג'ייקובס
Irwin and Joan Jacobs Graduate School

©

All rights reserved

*This work, in whole or in part, may not be copied (in any media), printed, translated, stored in a retrieval system, transmitted via the internet or other electronic means, except for "fair use" of brief quotations for academic instruction, criticism, or research purposes only.
Commercial use of this material is completely prohibited.*

©

כל הזכויות שמורות

אין להעתיק (במדיה כלשהי), להדפיס, לתרגם, לאחסן במאגר מידע, להפיץ באינטרנט, חיבור זה או כל חלק ממנו, למעט "שימוש הוגן" בקטעים קצרים מן החיבור למטרות לימוד, הוראה, ביקורת או מחקר. שימוש מסחרי בחומר הכלול בחיבור זה אסור בהחלט.

מקודד ADPCM ממותג לאותות דיבור ונתונים

בקצב שידור של 16 kbps

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר

מגיסטר במדעים

בהנדסת חשמל

10/11/87

מאת

אמיר זינתי

44.925 2050763



000000872331

24.12.88

הוגש לסנט הטכניון - מכון טכנולוגי לישראל

יולי 1987

חיפה

תמוז תשמ"ז



000001012187

לאשתי היקרה עפרה אשר עזרה לי באירגון העבודה,

ולבני היקר אשר אכל בחנאה חלקים ממנה.

המחקר נערך בהנחייתו של פרופסור דוד מלאך
בפקולטה להנדסת השמל.

תודתי העמוקה נתונה לפרופ' דוד מלאך על הצעת הנושא, על
הדרכתו והערותיו המועילות ועל סבלנותו הרבה במשך העבודה.

תודתי נתונה לחברת תדיראן - הטיבת התיקשורת שתמכה
במחקר.

ברצוני להודות גם לזיוה אבני אשר על המחשב, ליורם אור-חן
מהנדס המעבדה, ולחברי על עזרתם ותמיכתם.

תוכן ענינים

עמוד

1	ת ק צ י ר
2	רשימת סמלים וקיצורים
5	פרק 1 - מ ו א
5	1.1 מטרת העבודה
6	1.2 קריטריוני טיב
6	1.2.1 הערכת טיב אותות דיבור
7	1.2.2 הערכת טיב אותות נתונים
7	1.3 מבנה החיבור
9	פרק 2 - קוונטיזציה
10	2.1 קוונטיזר אופטימלי
11	2.2 קוונטיזר לוגריתמי
13	2.3 קוונטיזר אדפטיבי
13	2.3.1 אדפטציה קדמית
13	2.3.2 אדפטציה אחורית
19	פרק 3 - מקודדי אות הפרש (ADPCM)
20	3.1 מסנן קטבים
21	3.1.1 שערך קדמי
24	3.1.2 שערך אחורי
26	3.2 מסנן אפסים
27	פרק 4 - מקודד ADPCM סטנדרטי לקצב 32 kbps
27	4.1 דרישות התכנון
27	4.2 תאור האלגוריתם
28	4.2.1 קוונטיזר אדפטיבי
30	4.2.2 חזאי אדפטיבי
32	4.2.3 יחידת הסינכרון
32	4.3 תוצאות סימולציה
32	4.3.1 קידוד אותות דיבור
33	4.3.2 קידוד אותות נתונים

תוכן ענינים (המשך)

<u>עמוד</u>		
35 עיצוב ספקטרום שגיאת הקוונטיזציה	פרק 5 -
36 Noise Feedback Coding	5.1
37 Postfiltering	5.2
39 השוואת ביצועי המערכות השונות	5.3
41 אלגוריתם הפרדה בין אותות דיבור ונתונים	פרק 6 -
42 חציות אפס ונקודות אקסטרמום	6.1
44 6.1.1 קצב חציות האפס והאקסטרמום של אותות דיבור נתונים	
47 עוצמת האות	6.2
47 6.2.1 רמת סף קבועה - T_c	
48 6.2.2 רמת סף אדפטיבית - T_A	
48 סיכום אלגוריתם הסיווג	6.3
52 מבנה מערכת הקידוד המוצעת וביצועיה	פרק 7 -
53 חזזאי הממותג	7.1
53 שיפור סוביקטיבי של איכות האות המשוחזר	7.2
55 הקוונטיזר	7.3
56 7.3.1 התחום הדינמי של הקוונטיזר	
56 7.3.2 פרמטרי הקוונטיזר	
59 ביצועי המערכת המוצעת	7.4
59 7.4.1 קידוד אותות דיבור	
62 7.4.2 קידוד אותות נתונים	
64 סכום ומסקנות	פרק 8 -
66 כיוונים למחקר המשך	9.1
67 מקורות	
I תקציר (באנגלית)	

ת ק צ י ר

בעבודה זו אנו מציגים מערכת קידוד ADPCM משופרת לקצב 16 kbps המיועדת לאותות דיבור ונתונים (אותות נתונים הם אותות ספרתיים המאופננים לקבלת אות אנלוגי בתחום רוחב הסרט של אות הדיבור, VBD – Voiceband Data. אנו מטפלים באותות VBD המתקבלים על-ידי אפנון – FSK Frequency Shift Keying) – PSK – Phase Shift Keying בקצבי נתונים של עד 2400 bps.

להבדיל ממערכות קידוד קיימות לקצב הנ"ל, המותאמות לאותות דיבור בלבד, המערכת המוצגת ממותגת בין שני אופני עבודה המתאימים לשני סוגי האותות הנדונים – דיבור/VBD:

ברירת המחדל של המערכת היא אופן פעולה המתאים לאותות דיבור. במצב זה המערכת מבוססת על אלגוריתם ה-ADPCM הסטנדרטי לקצב 32 kbps אשר נקבע על-ידי CCITT. בקצב שידור של 16 kbps מתקבלת, על-ידי אלגוריתם ה-ADPCM, שגיאת קוונטיזציה המקטינה במידה ניכרת את איכות האות המשוחזר. בספרות מוצעות מספר דרכים לשיפור סוביקטיבי באיכות האות המשוחזר, כגון אלגוריתמי עיצוב ספקטרום שגיאת הקוונטיזציה בעזרת מסנן משוב (NFC – Noise Feedback Coding) וכן סינון האות המשוחזר על-ידי מסנן עיצוב הנמצא במוצא המפענח (Postfiltering). במערכת המוצגת אנו משתמשים ב-Postfilter אדפטיבי, שבעזרתו התקבל שיפור סוביקטיבי משמעותי באיכות האות המשוחזר.

עבור אותות VBD השילוב של ADPCM בתוספת Postfilter התגלה כלא מתאים. לעומתם הביא מקודד APCM לביצועים טובים יותר ואף טובים מביצועי מקודד ADPCM בערוץ ספרתי רועש.

המיתוג בין שני אופני העבודה מבוצע בעזרת אלגוריתם המסווג את אות הכניסה לאחת משתי הקטגוריות דיבור/VBD. אלגוריתם הסיווג מבוסס על אלגוריתמים המוצגים בספרות, תוך הכנסת השינויים הדרושים לשם התאמתם למערכת הקידוד המוצגת, בה לא מועברת כל אינפורמציה צד מהמקדד למפענח. פרמטרי האות המשמשים את האלגוריתם הם קצב חציות האפס, עוצמתו, מידת חסטיזציות וכן היחס בין מכסימום העוצמה לערכה הממוצע כמדד לצורת האות.

ביצועי מערכת הקידוד המוצגת, עבור אותות דיבור בקצב 16 kbps, עולים במידה ניכרת על הביצועים המתקבלים על-ידי מערכת קידוד ADPCM מקובלת באותו קצב שידור. זאת למרות הנחת התדרים הגבוהים המתקבלת על-ידי ה-Postfilter. בנוכחות אותות VBD בקצב 2400 bps מתקבלים ביצועים דומים לאלה המתקבלים על-ידי אלגוריתם ה-ADPCM הסטנדרטי לקצב שידור 32 kbps.

רשימת סמלים וקיצורים

ממוצע הערך המוחלט של האות.	-	AVE
מייצג התמרת Z של a ו- b בחתאמה.	-	$B(Z), A(Z)$
וקטור מקדמי מסנן חיזוי מסוג all-pole ו-all-zero בחתאמה.	-	$\underline{b}, \underline{a}$
אוסף מקדמי מסנן.	-	$\{a_i\}, \{b_i\}$
פונקציה המבצעת Compression.	-	$C(\cdot)$
אות ההפרש במקודד ADPCM.	-	$d(t)$
וקטור אות ההפרש.	-	$\underline{d}(t)$
אות ההפרש האחורי במוצא הדרגה ה-n-ית.	-	d_n^b
אות ההפרש הקדמי במוצא הדרגה ה-n-ית.	-	d_n^f
אות ההפרש במוצא הקוונטייזר.	-	$d_q(\cdot)$
תוחלת.	-	$E[\cdot]$
שגיאת הקוונטיזציה.	-	e_q
מסנן משוב לשגיאת הקוונטיזציה.	-	$F(Z)$
תדר דגימה.	-	f_s
פונקציות צפיפות החסתברות.	-	$f_x(\cdot)$
פונקציות הפילוג המשותפת לתהליך ולנגזרתו.	-	$f_{xx}(x\dot{x})$
תגובת התדר של מסנן גוזר דיסקרטי.	-	$H(\theta)$
יציאת המסווג.	-	ID
מילת הקוד ביציאת הקוונטייזר.	-	$i(t)$
מקדמי ההחזרה (מקדמי הקורלציה החלקיים).	-	$\{K_n\}$
מספר רמות היצוג של הקוונטייזר.	-	L
כופלי הקוונטייזר.	-	$\{M_i\}$
גורם הנירמול של מסנן ה-PF.	-	N
גורם נירמול קבוע של מסנן ה-PF שנקבע לאותות דיבור.	-	N_s
מסנן חיזוי.	-	$P(Z)$
המקדם ה-i של מסנן החיזוי.	-	P_i
פונקציות התמסורת של מסנן ה-PF במישור Z.	-	$PF(Z)$
מייצג פעולת קוונטיזציה.	-	$Q(\cdot)$
שגיאת הקוונטיזציה.	-	q
מטריצת האוטוקורלציה.	-	R
פונקציות האוטוקורלציה של תהליך רציף.	-	$R(\tau)$

הנגזרת ה-n-ית של $R(\tau)$	-	$R(\tau)^{(n)}$
וקטור האוטוקורלציה.	-	r
וריאנס אות ההפרש d_n^f .	-	r_n^f
ספקטרום החספק של תהליך רציף.	-	$S(f)$
ספקטרום החספק של תהליך בדיד.	-	$S(\theta)$
הפעולה המתמטית של חוצאת סימן.	-	$\text{sign}(\cdot)$
רמת סף אדפטיבית.	-	T_A
רמת סף קבועה.	-	T_c
משעריך וריאנס האות.	-	V_ℓ
משעריך הוריאנס לזמן קצר.	-	V_s
אות כניסה.	-	$X(t)$
האות המשוערך במסנן החיזוי.	-	X_e
תפוקת מסנן הקטבים.	-	X_{ep}
תפוקת מסנן האפסים.	-	X_{ez}
נקודת הגלישה המגדירה את תחום העבודה של הקוונטייזר.	-	X_{ol}
נגזרת האות $X(t)$.	-	$\dot{X}(t)$
רמות החלטה של הקוונטייזר.	-	$\{X_i\}$
אות מוצא.	-	$Y(t)$
רמות היצוג (היציאה) של הקוונטייזר.	-	$\{Y_i\}$
מייצג התמרת Z .	-	$Z(\cdot)$
מונה מספר חציות האפס.	-	Z_o
מונה מספר נקודות האקסטרמום.	-	Z_1
רמת יחוס של הספק, מיוצגת על-ידי הערך 4000.	-	0 dBm0
משתנה המתאר את מידת הסטציונריות של האות.	-	$\alpha(t)$
גדלים הקובעים את מידת הסינון של מס ה-PF.	-	α, β
גדלים הקובעים את מידת הסינון של מסנן ה-NFC.	-	α_F, β_F
וריאנס אות ההפרש.	-	σ_d^2
וריאנס שגיאת הקוונטיזציה.	-	σ_q^2
$\log(\cdot)$ של צעד הקוונטיזציה.	-	∇
$\log(\cdot)$ של צעד הקוונטיזציה בעל קבוע הזמן האיטי ביותר.	-	∇_ℓ
$\log(\cdot)$ של צעד הקוונטיזציה בעל קבוע הזמן המהיר ביותר.	-	∇_u
גרדיאנט וריאנס אות ההפרש לפי a .	-	$\nabla \sigma_d^2(a)$
צעד הקוונטיזציה (גורם הנירמול).	-	Δ
צעד הקוונטיזציה המכסימלי.	-	Δ_{\max}
צעד הקוונטיזציה המינימלי.	-	Δ_{\min}

מקדם הקורלציה המנורמל בין שני דגמים סמוכים.	-	ρ_1
רמות החלטה מנורמלות וקבועות של הקוונטייזר.	-	$\{\xi_i\}$
רמות יצוג מנורמלות וקבועות של הקוונטייזר.	-	$\{\eta_i\}$
.Adaptive Delta Modulation	-	ADM
.Adaptive Differential PCM	-	ADPCM
.ADPCM with Postfilter	-	ADPCM-PF
.Adaptive Predictive Coding	-	APC
.Adaptive PCM	-	APCM
.Adaptive Transform Coding	-	ATC
.Bit Error Rate	-	BER
.Block Error Rate	-	BLER
.bit per second	-	bps
Comité Consultatif International Téléphonique et .Télégraphique	-	CCITT
.Dynamic Locking of the Quantization step size	-	DLQ
.Frequency Shift Keying	-	FSK
kilobit per second	-	kbps
.Least Mean Square	-	LMS
.Low Pass Filter	-	LPF
.Minimum MSE	-	MMSE
.Mean Opinion Score	-	MOS
.Mean Square Error	-	MSE
.Noise Feedback Coding	-	NFC
.Pulse Code Modulation	-	PCM
.Postfilter	-	PF
.Phase Shift Keying	-	PSK
.Quadrature PSK	-	QPSK
.Subband Coding	-	SBC
.Segmental SNR (Seg. size = 256)	-	SEG-SNR
.Signal to Noise Ratio	-	SNR
.Voice Band Data	-	VBD

פרק 1 – מבוא

עבודה זו עוסקת בקידוד אותות לשם העברתם בערוץ תקשורת ספרתי ברוחב סרט כשל ערוץ טלפון. פעולת הקידוד מהווה שם כולל לתהליך ההמרה מהתחום הרציף (Analog) אל התחום הבדיד (Digital), לתהליך הדחיסה (הקטנת קצב השידור) ולהמרה ההפוכה. בעבודה זו אנו מניחים כי ההמרה לתחום הבדיד, הכוללת דגימה וקוונטיזציה ראשוניים PCM - Pulse Code Modulation, כבר בוצעה ונתרכז בפעולת הדחיסה בלבד.

נפתח דיון זה בסקירה קצרה על סוגי המקודדים השונים ופירוט מטרות העבודה, תוך ציון מקומו של המקודד המוצג במשפחת המקודדים. אחר נציין מהן אמות המידה על-פיהן אנו קובעים את טיב האות המקודד, ונסיים בתאור מבנה החיבור.

1.1 מטרת העבודה

ניתן לחלק את משפחת מקודדי אותות הדיבור על-פי אופן פעולתם וקצב השידור שלהם. לחשגת איכות גבוהה (toll quality) ואיכות תקשורת (communication quality) משתמשים במקודדי צורת גל (Waveform coders) בתחום קצבי השידור 16:64 kbps. מקודדים אלה מנסים לשחזר את צורת האות. ניתן לחלק את מקודדי צורת גל למקודדים הפועלים בתחום הזמן (כגון: Adaptive PCM, Adaptive ADPCM - Differential PCM, Adaptive ADM - Adaptive Delta Modulation, APC - Predictive Coders) ולמקודדים הפועלים בתחום התדר (SBC - Sub Band Coders, ATC - Adaptive Transform Coders), ראה פירוט ב-[Jayant84]. איכות זיבור סינטטית מתקבלת על-ידי מקודדי מקור (Source Coders או Vocoders) בתחום הקצבים 1.2:2.4 kbps, ראה פירוט ב-[Rabiner78]. מקודדים אלה מבוססים על בניית מודל למערכת הפקת אות דיבור.

- מטרת עבודה זו היא פיתוח אלגוריתם קידוד לקצב שידור - 16 kbps העונה על הדרישות הבאות:
1. אלגוריתם הקידוד יותאם לא רק לאותות דיבור אלא גם לאותות נתונים (VBD).
 2. האלגוריתם יהיה בעל סיבוכיות נמוכה על מנת לאפשר מימוש ב- real time. רצוי גם כי כמות הזיכרון הנדרשת תהיה נמוכה.
 3. חשיית האות המשוחזר על-ידי המקודד תהיה מינימלית. כך תמנענה הבעיות הנובעות מהדים חוזרים.
 4. דרושה אדפטציה אחורית של פרמטרי המקודד. בכך ימנע הצורך בסינכרון המקלט למשדר, ובשידור פרמטרי המקודד כאינפורמציה צד.
 5. לצורך שימוש באלגוריתם זה בערוץ תקשורת, עליו להיות מסוגל לפעול גם בנוכחות שגיאות ערוץ ורעש נלווה לאות בכניסת המקודד.

מסיבות אלו בחרנו לעבוד עם מקודד צורת גל אדפטיבי הפועל בתחום הזמן, שהוא בעל סיבוכיות נמוכה ורגישות קטנה יחסית לפרמטרי האות המקודד ולרעש ערוץ. אנו מנצלים בעבודה זו את הנסיון הרב, אשר נרכש בעת פיתוח אלגוריתם ה-ADPCM לקצב שידור - 32 kbps, ומבססים עליו את האלגוריתם המוצג תוך הכנסת השינויים הדרושים לצורך התאמתו לקצב שידור - 16 kbps.

1.2 קריטריוני טיב

נסקור כעת מספר מדדים המשמשים אותנו לקביעת טיב האותות המשוחזרים על-ידי מקודדים. שיטות המדידה המקובלות ביותר מבוססות על מדידת השגיאה הריבועית הממוצעת (Mean Square Error - MSE). במקרים רבים אנו מעוניינים בדיעת ה-MSE המנורמל על-פי עוצמת האות, או לחלופין יחס האות לרעש (SNR - Signal to Noise Ratio). גדלים אלה נוחים מאוד למדידה, והצבתם כפונקציה מטרה בעת האופטימיזציה מוביל במקרים רבים למערכת לינארית. למרות יתרונות אלה אין הן משקפות נאמנה את האיכות הסובייקטיבית של אותות דיבור ואת מספר שגיאות הפיענוח של אותות VBD. לצורך הערכת טיב האותות הנידונים יש צורך בהשלמת מדידת ה-SNR במדידות נוספות כדלקמן:

1.2.1 הערכת טיב אותות דיבור

מחלקים את סוגי המדידות השונות להערכת טיב אותות הדיבור לשתי קבוצות:

מדידות אובייקטיביות: מדידת SNR המתאימה יותר לאותות דיבור (שהם אי-סטציונריים) היא ה-Segmental SNR (SEGSNR) הנתונה על-ידי הביטוי הבא:

$$\text{SEGSNR} = \frac{1}{M} \sum_{i=1}^M \text{SNR}_i$$

כאשר SNR_i הוא ה-SNR (ביחידות של dB) המתקבל בבלוק ה-i מתוך M בלוקים. אורך בלוק מקובל הוא בתחום 8-32 msec. לעתים מגבילים את ערכי ה-SNR לתחום מותר. באופן זה נמנע מצב בו לבלוק בודד השפעה גדולה מידי על ה-SEGSNR הכללי. נעיר כי SNR נמוך מאוד מתקבל בקטעי שקט. אנו מגבילים את ה-SNR מלמטה על-ידי קביעת רמת סף, אשר מתחת לה אין מחשבים את ה-SNR בבלוק הרלוונטי (רמת סף זו היא בגובה של כ-42 dBm0, המיוצגת על-ידי מילת מחשב שערכה - 30).

קיימות מדידות מורכבות יותר המביאות בחשבון את ההשפעות הסובייקטיביות השונות של סוגי השגיאה הבאים:

1. שגיאות גלישה (overload) ושגיאות גרגריות (granular).
2. שגיאות קורלטיביות ושאינן קורלטיביות לאות המקודד.
3. שגיאות בתחומי תדר שונים.

מדידות סוביקטיביות: המגוון הרב של שיטות מדידת שגיאת הקידוד (אשר חלקן חובא לעיל) מעיד למעשה על כך שלא קיימת שיטה אחת אופטימלית בעזרתה ניתן להעריך את איכות האות או את המובנות שלו. לכן יש צורך בהערכות סוביקטיביות הנקבעות על-פי שמיעה.

בדיקת מובנות היא מדידה סטטיסטית השימושית בעיקר עבור מקודדים לקצבי שידור נמוכים ו/או לאותות המלווים בהפרעות. בתחום קצבי השידור בו אנו עוסקים המדד הרלוונטי הוא איכות האות, המבטאת את מידת "הנעימות" הסוביקטיבית שלו.

שתי השיטות המקובלות ביותר לקביעת איכות האות הן:

1. דירוג איכות האות בסולם סוביקטיבי בן 5 דרגות (ממציין עד בלתי נסבל).
 2. השוואה בין זוגות אותות לקביעת הטוב מביניהם.
- בעבודה זו אנו משתמשים בשיטה השנייה, הן משום חוסר הנסיון של המאזינים שהשתתפו בהערכת איכות האות (בשיטה הראשונה דרוש נסיון רב יותר), והן משום שהמטרה בעת ביצוע מדידות אלה היתה לקבל תוצאות השוואתיות למערכות קידוד קיימות.

1.2.2 הערכת טיב אותות נתונים

המדד לקביעת איכות האות המשוחזר הוא כמות שגיאות הפיענוח הנגרמות כתוצאה מפעולת הקידוד. גם כאן לא ניתן להעריך את איכות האות על-פי ה-SNR. עבור אותות PSK מודדים לעיתים את הוריאנס של הזזת הפזה הנגרמת על-ידי המקודד אך גם מדידה זו לא מספקת תשובה מלאה. עליכן יש צורך למדוד באופן ישיר את שכיחות שגיאות הפיענוח. בדרך-כלל מודדים זאת במונחים של Block Error Rate – BLER. במטרה להקטין את כמות הסימולציות הדרושה נחליף מדידה זו במדידת שכיחות הסיביות השגויות Bit Error Rate – BER.

1.3 מבנה החיבור

פרק 2 מוקדש לנושא הקוונטיזציה. תחילה מוגדרת פעולת הקוונטיזציה, ואחר מוצג הקוונטייזר האופטימלי על-פי קריטריון ה-MSE. כמו כן מוצג הקוונטייזר הלוגריתמי וקוונטייזרים אדפטיבים. דגש הושם על קוונטיזציה אדפטיבית אחורית.

בפרק 3 מבוצעת סקירה על מסנני חיזוי אדפטיבים קדמיים ואחוריים. מוצגים מספר קירובים בעזרתם מגיעים לאלגוריתם אדפטיבי אחורי בעל סיבוכיות נמוכה לחישוב מסנן מסוג all-pole ומסנן מסוג all-zero.

בפרק 4 מתואר בפירוט אלגוריתם ה-ADPCM הסטנדרטי לקצב 32 kbps, מוצגות המטרות שעמדו בפני המתכננים וכן ביצועי המקודד. תאור זה מוגש שכן אלגוריתם זה שימש אותנו כבסיס למערכת המוצעת בעבודה זו.

בפרקים 5 ו-6 אנו מתארים את התוספות והשינויים העיקריים שחונקסו באלגוריתם הסטנדרטי לקצב 32 kbps, אשר איפשרו חודת הקצב ל-16 kbps תוך פגיעה קטנה ככל האפשר בביצועי המקודד. בפרק 5 אנו מתארים שתי שיטות לשיפור סוביקטיבי של אותות דיבור. האחת על-ידי עיצוב ספקטרום שגיאת הקוונטיזציה בעזרת מסנן משוב NFC – Noise Feedback Coding, והשנייה על-ידי סינון האות המשוחזר בעזרת מסנן אדפטיבי הנמצא במוצא המפענח Postfiltering. שיטת ה-NFC מבוססת על ההנחה כי באוזן האדם מבוצעת חלוקת האות לפסי תדר בהם האות עובר טיפול נפרד. השפעת עיוותים על איכות האות המשוחזר תלויה בצורת ספקטרום העיוותים וביחסה לצורת ספקטרום האות. בשיטת ה-Postfiltering מנצלים דווקא את ההשפעה ההדדית שבין פסי התדר השונים. אות בעוצמה גבוהה בפס תדר אחד מקטין את רגישות האוזן בפסי התדר האחרים. על-ידי הגברת האות המשוחזר בתחום תדרים בהם ה-SNR גבוה על השבון הנחתת האות המשוחזר בתחום תדרים בהם ה-SNR נמוך, ניתן למסך את העיוותים בתחום התדרים האחרון. עקב השוואה סוביקטיבית של שתי שיטות אלה בחרנו להשתמש בשיטת ה-Postfiltering.

מקודד ה-ADPCM בתוספת ה-Postfilter (ADPCM-PF) חתברר כלא מתאים לקידוד אותות VBD. ביצועים טובים יותר באופן משמעותי התקבלו על-ידי מקודד APCM (PCM עם קוונטיזציה אדפטיבית). מכאן נבע הצורך במקודד הממותג בין שני אופני עבודה: ADPCM-PF עבור אותות דיבור ו-APCM עבור אותות VBD. קביעת אופן העבודה מבוצעת באופן אוטומטי על-ידי אלגוריתם המסווג את אות הכניסה לאחת משתי הקטגוריות – דיבור/VBD. אלגוריתם זה מבוסס על פרמטרי האות הבאים: קצב חציות האפס של האות כמדד לתדירות הדומיננטית של האות; עוצמת האות; קצב ירידת עוצמת האות ויחס מכסימום העוצמה לערכה הממוצע כמדד לצורת האות. שלבי אלגוריתם הסיווג מופיעים בפרק 6.

בפרק 7 מוצגת מערכת ה-ADPCM המוצעת, הממותגת בין שני אופני העבודה המתאימים לאותות דיבור ו-VBD. כמו כן מוגש סיכום תוצאות הסימולציה שבוצעו לצורך בחינת ביצועי המערכת תוך השוואתם עם מקודדים קיימים.

סיכום החיבור מופיע בפרק 8, וכן מוצעים שם מספר כיוונים למחקר המשך לשיפור מערכת הקידוד.

פרק 2 – קוונטיזציה

קוונטיזציה היא המרה של תחום עוצמות רציף לתחום בדיד על-פיה קטעים שונים בתחום הרצף מיוצגים על-ידי רמות בדידות המתאימות להם כמוגדר באופן הבא:

$$(2.1a) \quad \begin{cases} Q : X \rightarrow Y, \quad X = \mathbb{R}, \quad Y \subset X \\ Q(I_i) = y_i, \quad I_i = \{x \mid x_{i-1} \leq x < x_i\}, \quad y_i \in I_i \\ \bigcup_i I_i = X, \quad \bigcap_i I_i = \emptyset \end{cases}$$

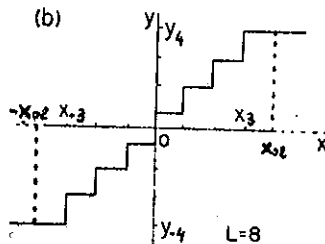
כאשר X ו- Y הם כניסת ויציאת הקוונטייזר Q בהתאמה, $\{x_i\}$ הן רמות החלטה המגדירות את הקטעים $\{I_i\}$ ו- $\{y_i\}$ הן רמות היצוג (היציאה) של הקוונטייזר. נגדיר את צעד הקוונטיזציה Δ כהפרש בין זוג רמות החלטה עוקבות

$$(2.1b) \quad \Delta_i = x_i - x_{i-1}$$

אנו מצמצמים את הדיון לקוונטייזרים סימטריים וזוגיים בעלי $L = 2^R$ רמות ייצוג R הוא מספר הסיביות לדגם ביצוג בינארי) כמוגדר להלן:

$$(2.2) \quad \begin{cases} x_{-i} = -x_i, \quad y_{-i} = -y_i \quad i=1, \dots, L/2 \\ x_0 = 0, \quad x_{L/2} = \infty \\ |\pm x_{oi}| = x_{L/2-1} + \Delta_{L/2-1} \end{cases}$$

x_{oi} היא נקודת הגלישה המגדירה את תחום העבודה של הקוונטייזר (ראח איור 2.1), ו- i הוא מספר רמת הקוונטייזר.



איור 2.1 – תאור של קוונטייזר סימטרי זוגי בעל $L=8$ רמות ייצוג.

Figure 2.1 – Description of an even and simetric quantizer with $L=8$ representation levels.

אנו פותחים פרק זה בהגדרת קריטריון טיב לקוונטיזר במובן שגיאה ריבועית ממוצעת (mse), והצגת הקוונטיזר האופטימלי לפילוג אחיד. לאחר מכן נסקור שתי דרכים להרחבת התחום הדינמי של הקוונטיזר.

2.1 קוונטיזר אופטימלי

נסמן ב- q את שגיאת הקוונטיזציה, היא ההפרש בין אות הכניסה לאות היציאה של הקוונטיזר. $q = X - Y$. שגיאה בתחום העבודה של הקוונטיזר $|X| \leq x_{0\ell}$ - נקראת שגיאה גרגירית (granular error), ושגיאה בתחום הערכים $|X| > x_{0\ell}$ נקראת שגיאת גלישה (overload error). עבור שגיאת קוונטיזציה עם ממוצע אפס, וריאנס שגיאה הכוללת (גרגירית וגלישה) - σ_q^2 נתון על-ידי:

$$(2.3) \quad \sigma_q^2 = E[(X - Y)^2] = 2 \cdot \sum_{i=1}^{L/2} \int_{x_{i-1}}^{x_i} (x - y_i)^2 f_x(x) dx$$

כאשר $f_x(\cdot)$ היא פונקציית צפיפות ההסתברות (pdf) של אות הכניסה X .

וריאנס שגיאת הקוונטיזציה או ה-mse הוא קריטריון טיב מקובל. הקוונטיזר האופטימלי הוא זה אשר עבורו מתקבל mse מינימלי (mmse). תנאי הכרחי לקבלת mmse הוא איפוס כל הנגזרות החלקיות של σ_q^2 לפי $\{x_i\}$ ו- $\{y_i\}$. בנוסף, עבור $\log[f_x(x)]$ קמורה תנאים אלו מספיקים ומובילים לקוונטיזר Max [Jayant84] על-ידי:

$$(2.4) \quad \begin{cases} x_i = \frac{1}{2}(y_i + y_{i+1}) \\ y_i = \frac{\int_{x_{i-1}}^{x_i} x f_x(x) dx}{\int_{x_{i-1}}^{x_i} f_x(x) dx} = E[X | X \in I_i] \end{cases}$$

רמת החלטה x_i נמצאת במרכז שתי רמות היצוג הסמוכות לה משני צידיה, ורמת היצוג y_i היא התוחלת המותנית של X בהנתן שייכות לקטע I_i . עבור קוונטיזר עם $L > 2$ רמות, הפתרון האופטימלי ניתן להשגה באופן איטרטיבי.

מקרה פרטי חשוב הוא הקוונטיזר האופטימלי לפילוג אחיד. קוונטיזר זה מוגדר חד ערכית על-ידי צעד הקוונטיזציה ומספר הרמות.

$$(2.5) \quad \begin{cases} \Delta_i = \Delta & , i=1, \dots, L/2-1 \\ x_i = i \cdot \Delta & , i=1, \dots, L/2-1 \\ x_o = 0, x_{L/2} = \infty, x_{ol} = L/2 \cdot \Delta \\ y_i = (i - \frac{1}{2}) \cdot \Delta & , i=1, \dots, L/2 \end{cases}$$

בשל פשטותו, השימוש בו מקובל במקרים רבים גם לקידוד אותות שאינם בעלי פילוג אחיד. קוונטייזר אחיד אופטימלי מתקבל בתנאי איפוס נגזרת σ_q^2 לפי Δ .

אנו מודדים את ביצועי הקוונטייזר על-פי היחס שבין וריאנס אות הכניסה לוריאנס השגיאה אשר תסומן ב- SNR_Q .

$$(2.6) \quad SNR_Q [dB] = 10 \log (\sigma_x^2 / \sigma_q^2)$$

SNR_Q מכסימלי מתקבל על-ידי הקוונטייזר האופטימלי (עבור פילוג אחיד מתקבל $SNR_Q \text{ optimum} = 6.02 \cdot R \text{ dB}$, מספר הסיביות לדגם). אי התאמה של הקוונטייזר לצורת פילוג האות או לעוצמתו תקטין את ה- SNR_Q המתקבל. קוונטייזר תת אופטימלי לפילוג האות, אך חרגיש פחות לשינויים בעוצמתו, הוא הקוונטייזר הלוגריתמי המתואר בסעיף הבא.

2.2 קוונטייזר לוגריתמי

קוונטייזר זה מתוכנן להשגת רגישות נמוכה לוריאנס האות (תחום דינמי רחב), בו צעד הקוונטיזציה יחסי לעוצמת האות. במלים אחרות, צעדי קוונטיזציה קטנים עבור הרמות הנמוכות וצעדי קוונטיזציה גדולים עבור הרמות הגבוהות.

$$(2.7) \quad \Delta_i = \Delta(x_i) \propto x_i$$

דיאגרמת מלבנים למימוש קוונטייזר זה מתוארת באיור 2.2. בדיאגרמה זו האות עובר דחיסה (compression) על-ידי פונקציה אי לינארית, קוונטיזציה בעזרת קוונטייזר אחיד והרחבה (expansion) על-ידי פונקציה הופכית לפונקציה הדחיסה.

$$(2.8) \quad Y = C^{-1} \{ Q_{uniform} [C(X)] \}$$

כאשר $C(\cdot)$ - compression function
 $C^{-1}(\cdot)$ - expanding function } companding law

לשם מילוי (2.7) ותנאי השפה: $C(0) = 0, C(x_{ol}) = x_{ol}$ דרושה פונקציה דחיסה לוגריתמית עם קירוב לינארי בראשית.

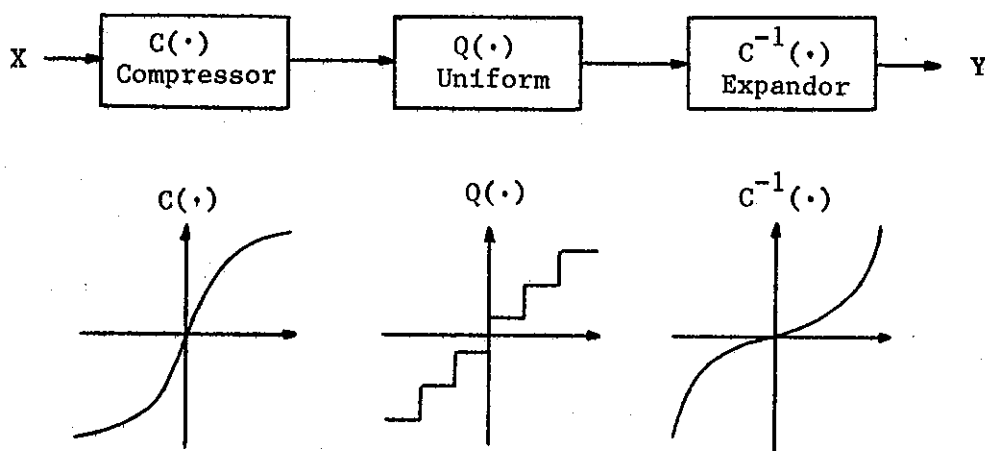
שני קירובים המקובלים בתחום התקשורת הספרתית הם:

- 1) A Law Companding ($A = 87.56$)

$$C(x) = \begin{cases} \frac{Ax}{1 + \ln A} & , \quad \frac{|x|}{x_{ol}} \leq \frac{1}{A} \\ \frac{x_{ol}}{1 + \ln A} \cdot [1 + \ln(A \cdot \frac{|x|}{x_{ol}})] \cdot \text{sign}(x) & , \quad \text{else} \end{cases}$$

- 2) μ Law Companding ($\mu = 255$)

$$C(x) = x_{ol} \cdot \frac{\ln(1 + \mu \frac{|x|}{x_{ol}})}{\ln(1 + \mu)} \cdot \text{sign}(x)$$



איור 2.2 - דיאגרמת מלבנים של קוונטייזר לוגריתמי בעזרת פונקצית דחיסה/חריכה [Jayant84].

Figure 2.2 - Block diagram of a logarithmic quantizer using a companding function [Jayant84].

2.3 קוונטיזר אדפטיבי

דרך נוספת להגדלת התחום הדינמי היא בהתאמת צעדי הקוונטיזציה לוריאנס האות. את הקוונטיזר נייצג על-ידי רמות החלטה $\{\xi_i\}$ ורמות יצוג $\{\eta_i\}$ מנורמלות וקבועות, המוכפלות בגורם נירמול (צעד קוונטיזציה) Δ .⁽¹⁾

$$(2.9) \quad x_i = \xi_i \cdot \Delta, \quad y_i = \eta_i \cdot \Delta$$

לכל צמד פילוג-קוונטיזר מתאים קבוע ϕ המבטא את יחס גורם הנירמול האופטימלי לשרש וריאנס האות

$$(2.10) \quad \Delta_{opt} = \phi \cdot \sigma_x$$

הרחבת התחום הדינמי מבוצעת על-ידי התאמת גורם הנירמול Δ לוריאנס האות על-פי (2.10). מכאן, כל שנותר הוא שיערוך הוריאנס.

2.3.1 אדפטיציה קדמית

וריאנס האות משוערך מתוך דגמיו על-ידי חלוקת סדרת הדגמים למסגרות זמן (בלוקים) באורך M

דגמים באופן הבא:

$$(2.11) \quad \hat{\alpha}_x^2(n) = \frac{1}{M} \sum_{t=0}^{M-1} X^2(n \cdot M + t)$$

כאשר n הוא מספר הבלוק ו- t משתנה הזמן חבדיד. גודל חבלוק - M נקבע תוך שיקלול הגורמים הבאים: מהירות תגובת הקוונטיזר, זמן ההשהייה של האות, גודל הזכרון הדרוש וכן הצורך בהעברת צעד הקוונטיזציה מהמקדד למפענח מידי כל בלוק.

לאותות דיבור הדגומים בתדר $f_s = 8 \text{ KHz}$ אופייני גודל בלוק - M בתחום (16-32 msec) 128:256 [Jayant84]

2.3.2 אדפטיציה אחורית

שיערוך וריאנס האות מבוצע על-סמך דגמי המוצא של הקוונטיזר - $Y(t)$. ניתן לשערך את הוריאנס תוך הפעלת חלון מלבני על דגמי האות כב-(2.11), אך ביטוי נוח יותר מתקבל בעזרת חלון אקספוננציאלי.

$$(2.12a) \quad \hat{\alpha}_{y(t)}^2 = \alpha_q \sum_{k=0}^{\infty} (1-\alpha_q)^k Y^2(t-k), \quad 0 < \alpha_q \ll 1$$

(1) עבור הבחירה $\xi_1 = 1$ מתקבל גורם נירמול השווה בגודלו לצעד הקוונטיזציה הראשון כמוגדר

ב-(2.2) $\Delta = x_1 = \Delta_1$

ובצורה רקורסיבית

$$(2.12b) \quad \hat{\alpha}_y^2(t) = (1 - \alpha_q) \hat{\alpha}_y^2(t-1) + \alpha_q Y^2(t)$$

α_q קובע את מהירות תגובת המשערך. בחצבה $\alpha_q = 2/M+1$, מתקבל משערך עם מומנטים מסדר ראשון ושני הזחים לאלו המתקבלים על-ידי משערך בעל חלון מלבני באורך M דגמים - עבור תחליך אקראי סטציונרי וחסר קורלציה [Jayant84].

עם הצבת משתני הקוונטייזר המופיעים במשוואות (2.9) ו-(2.10) ב- $Y(t)$ מתקבל אלגוריתם האדפטציה הבא [Jayant73]:

$$(2.13a) \quad \Delta(t+1) = M [i(t)] \cdot \Delta(t)$$

$$(2.13b) \quad M [i(t)] = [1 - \alpha_q + \alpha_q \cdot \phi^2 \cdot \eta_i^2]^{1/2}$$

בהנחה $\hat{\alpha}_y \approx \sigma_x$ מתקבל כי עבור קוונטייזר בעל אוסף רמות $\{i_j\}$, $\{n_j\}$ נתון, $\{M[i(t)]\}$ הוא אוסף סופי של $L/2$ כופלים הקבועים בזמן. מעתה ואילך נשתמש בסימון $M_1 = M[i(t)]$.

ההנחה $\hat{\alpha}_y \approx \sigma_x$ נכונה עבור שינויים איטיים של וריאנס האות ביחס למהירות העקיבה של הקוונטייזר וכן עבור שגיאת קוונטיזציה קטנה. לצורך התאמת הקוונטייזר למצבים בהם קיים הבדל ניכר בין צעד הקוונטיזציה לצעד הקוונטיזציה האופטימלי, דרוש תיקון אוסף הכופלים. ב-[Jayant73] מוצע להגדיל את הכופל המתאים לרמה הגבוהה ביותר - $M_{L/2}$, מעבר לגודל הנקבע ב-(2.13b). לשם התאמת הקוונטייזר למצב בו $\Delta < \Delta_{opt}$, באופן דומה ניתן להקטין את הכופל המתאים לרמה הנמוכה ביותר - M_1 כפיצוי למקרה בו $\Delta > \Delta_{opt}$.

גישה שונה לחישוב אוסף הכופלים נובעת משיקולי יציבות. גישה זו נוחה לפתרון בתחום הלוגריתמי. נשתמש בסימונים הבאים:

$$(2.14) \quad \nabla = \log \Delta, \quad W_i = \log M_i$$

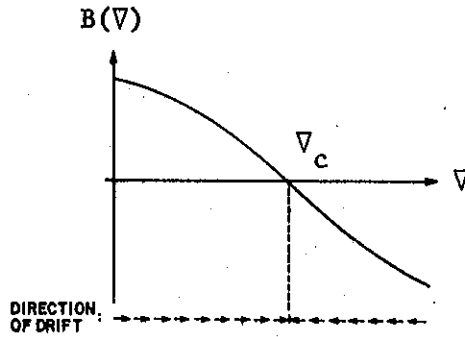
נגדיר את פונקצית ההטיה $B(\cdot)$ של הלוגריתם של צעד הקוונטיזציה - ∇ באופן הבא:

$$(2.15) \quad B[\nabla(t)] = E[\nabla(t+1)] - \nabla(t)$$

דרישה טבעית מתחליך האדפטציה היא הגדלת צעד הקוונטיזציה כאשר הוא נמוך מדי (מצב בו דומיננטיות שגיאות גלישה), והקטנתו כאשר הוא גבוה מדי (מצב בו דומיננטיות שגיאות גרגריות). כלומר, על אוסף הכופלים לקיים את התנאי -

$$(2.16) \quad 1 > M_1 \leq M_2 \leq \dots \leq M_{L/2} > 1$$

מדרישה זו נובע כי $B(\cdot)$ היא פונקציה מונוטונית יורדת. זאת ועוד, עבור ערכי ∇ נמוכים $B(\nabla) > 0$, ועבור ערכי ∇ גדולים $B(\nabla) < 0$. מכאן של- $B(\nabla)$ שורש יחיד, שיקרא לוגריתם צעד הקוונטיזציה חמרכזי - ∇_c (ראו איור 2.3).



איור 2.3 - פונקציית החטייה $B(\cdot)$ וכיוון השינוי ב- ∇ . השורש של $B(\cdot)$ הוא לוגריתם צעד הקוונטיזציה המרכזי - ∇_c [Debasis79]

Figure 2.3 - Bias function $B(\cdot)$ and the direction of ∇ 's drift. The root of $B(\cdot)$ is the central log step size - ∇_c [Debasis79].

עם הצבת $\nabla = \nabla_c$ ב-(2.15), ובהצבה $\nabla_{(t+1)} = \nabla_{(t)} + W_i$ נקבל את תנאי היציבות שעל אוסף הכופלים לקיים [Jayant73]

$$(2.17) \quad \sum_i W_i P_i = 0$$

כאשר P_i היא ההסתברות לאיכלוס רמת הקוונטיזציה ה- i -ית. אוסף ההסתברויות $\{P_i\}$ מתקבל מתוך ה-pdf של אות הכניסה ורמות ההחלטה של הקוונטיזר עבור $\nabla = \nabla_c$.

נפרט מספר תכונות של המשערך האחורי המהוות יתרון על-פני השיערוך הקידמי:

- * דרוש זכרון באורך מילה אחת בלבד לאיחסון מילת הקוד האחרונה.
- * קידוד האות מבוצע ללא השחייה נוספת על השחיית הדגימה.
- * אין צורך בהעברת צעד הקוונטיזציה כאינפורמציה צד, שכן האדפטציה מבוצעת על-סמך מילות הקוד של הקוונטיזר הקיימות גם במפענח.

ביתרונו של המשערך האחורי הפותר את הצורך בהעברת אינפורמציה צד למפענח, טמון גם חסרונו כאשר הקידוד נועד למטרות שידור בערוצי תקשורת. למרות דרישת המימוש לזכרון באורך מילה אחת בלבד לאלגוריתם זכרון אינסופי. צעד הקוונטיזציה ברגע $t+1$ הוא פונקציה של הצעד החתלתי וכל מילות הקוד עד לרגע t , ומכאן רגישותו הרבה לשגיאות ערוץ. שגיאת ערוץ ברגע t_0 פירושה אי שיויון בין מילות הקוד של המקודד והמפענח. עקב כך מתקבל הפרש קבוע בין הלוגריתם של צעד הקוונטיזציה של המקודד והמפענח החל מרגע t_0+1 ואילך.

$$\nabla(t) - \nabla'(t) = W(t_0) - W'(t_0), \quad t = t_0+1, t_0+2, \dots$$

כאשר ∇, W הם פרמטרי המקודד ו- ∇', W' הם פרמטרי המפענח.

ב-[Goodman75] מוצעת דרך להגבלת זכרון הקוונטייזר לצורך השגת עמידות לשגיאות ערוץ. במקום אלגוריתם האדפטציה המתואר ב-(2.13a) מוצע האלגוריתם הבא:

$$(2.18a) \quad \Delta(t+1) = M_i \cdot \Delta^\beta(t), \quad 0 \ll \beta < 1$$

ובתחום הלוגריתמי נקבל

$$(2.18b) \quad \nabla(t+1) = \beta \cdot \nabla(t) + W_i$$

כאשר קיימת שגיאת ערוץ בודדת המתרחשת ברגע t_0 נקבל

$$\nabla(t) - \nabla'(t) = \beta^{t-(t_0+1)} [W(t_0) - W'(t_0)], \quad t = t_0+1, t_0+2, \dots$$

β הוא גורם דעיכה אקספוננציאלי הגורם לזיכרון סופי (מעשית) של הקוונטייזר. עבור $\beta=1$ מקבלים את (2.13a) ועבור β קטן נקבל דעיכה מחירת.

תנאי היציבות של הקוונטייזר (לצורך חישוב אוסף הכופלים), מתקבל מהצבת (2.18b) ב-(2.15) עבור

$$\nabla_t = \nabla_c$$

$$(2.19a) \quad \sum_i W_i P_i = (1-\beta) \nabla_c$$

בהנתן אוסף כופלי הקוונטייזר בעל הזכרון האינסופי, המתקבל מ-(2.17), אשר יסומן ב- $\{M_i^\infty\}$, ניתן לקבל מתוכו בדרך נוספת אוסף כופלים המקיים את (2.19a) עבור $\beta < 1$ על-ידי התיקון הבא [Goodman75]

$$(2.19b) \quad M_i = M_i^\infty \cdot \Delta_c^{(1-\beta)}$$

תיקון זה מתקבל מתוך השוואת צעדי הקוונטיזציה ברגע $t+1$ בשני האלגוריתמים עבור תנאי ההתחלה

$$\Delta_t = \Delta_c$$

אלגוריתם זה רגיש פחות מקודמו לשגיאות ערוץ, אך זאת על חשבון הקטנת התחום הדינמי. קל לראות מתוך הצגה לוגריתמית של (2.18), כי ∇ מוגבל לתחום -

$$\frac{W_1}{1-\beta} \leq \nabla \leq \frac{W_{L/2}}{1-\beta}$$

זאת ועוד, שלא כבאלגוריתם הקודם, ה-SNR תלוי בווריאנס האות (ראה איור 2.4).

ב-[Einarsson81] מוצעת דרך להרחבת התחום הדינמי על-ידי הוספת משערך לורינס האות.

$$(2.20) \quad \begin{cases} \Delta(t+1) = M_i \cdot \Delta^\beta(t) \cdot \delta^\gamma(t), & 0 < \beta < 1 \quad 0 < \gamma < 1 - \beta \\ \delta(t) = \alpha \delta(t-1) + (1-\alpha)C |Y(t)|, & 0 < \alpha < 1 \end{cases}$$

כאשר: β, γ הם גורמי הדעיכה של הקוונטייזר.

α הוא קבוע הזמן של δ .

C הוא קבוע המתאם את $E|y| \rightarrow \sqrt{E y^2}$.

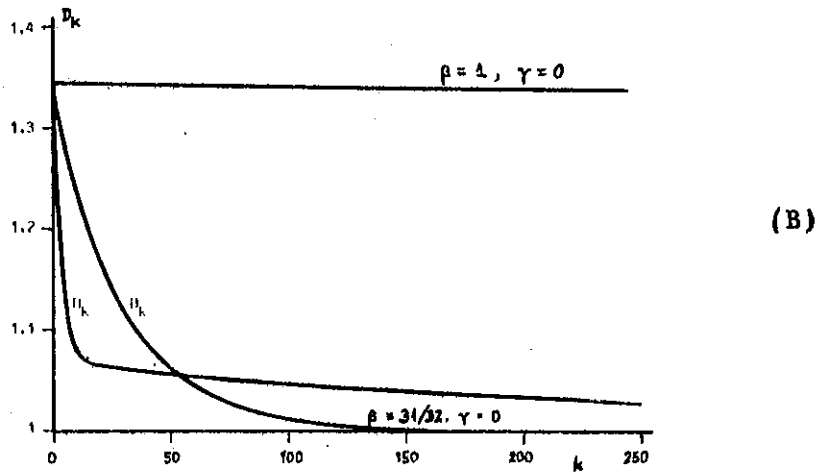
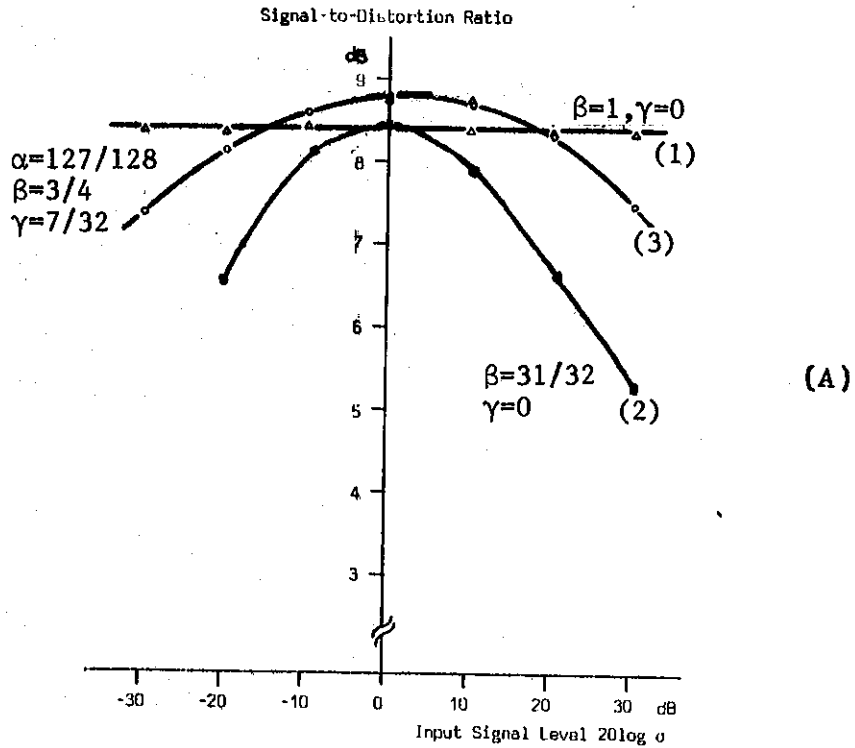
תנאי היציבות לצורך חישוב אוסף הכופלים מתקבל באופן דומה ל-(2.19a)

$$(2.21) \quad \sum_i W_i P_i = (1-\beta) \nabla_c - \gamma E [\log \delta | \nabla(t) = \nabla_c]$$

התחום הדינמי של הקוונטייזר מתקבל עם הצבת (2.9) ב-(2.20)

$$\frac{W_1 + \gamma [\log(c \eta_1)]}{1-\beta-\gamma} \leq \nabla \leq \frac{W_{L/2} + \gamma [\log(c \eta_{L/2})]}{1-\beta-\gamma}$$

באיור 2.4 מוצעת השוואת התחומים הדינמיים והתגובה לשגיאת ערוץ בודדת של אלגוריתמי האדפטציה האחורית שהוצגו.



איור 2.4 - השוואה של התחומים הדינמיים (A) ותגובה לשגיאת ערוץ בודדת (B) של אלגוריתמי האדפטציה הבאים [Einarsson81]:

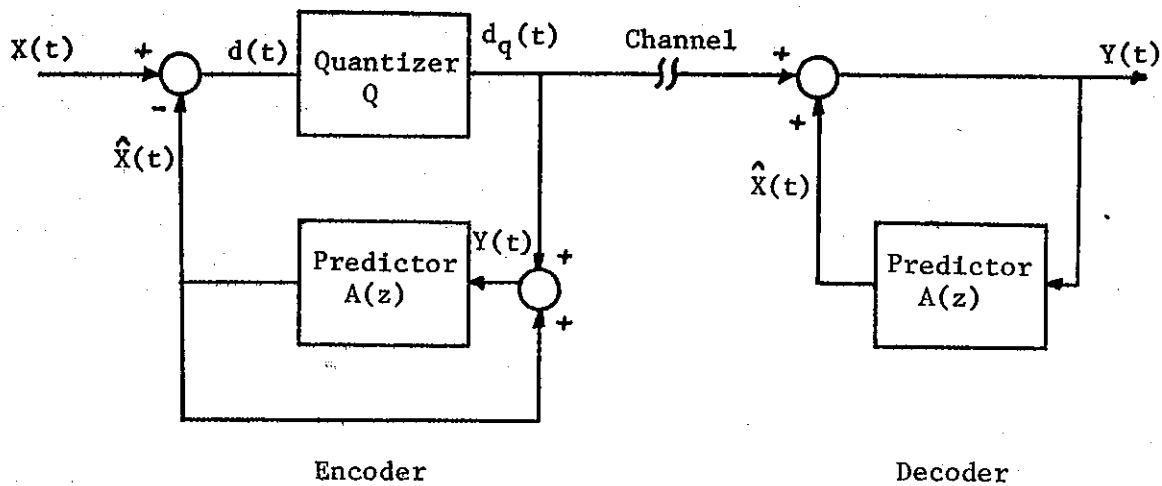
- 1 - קוונטייזר בעל זיכרון אינסופי - משוואה (2.13a).
- 2 - קוונטייזר עמיד לשגיאות ערוץ - משוואה (2.18).
- 3 - קוונטייזר עמיד לשגיאות ערוץ עם תחום דינמי מורחב - משוואה (2.20).

Figure 2.4 - Comparison of dynamic range (A) and single error effect (B) of the following adaptation algorithms [Einarsson81]:

- 1 - Quantizer with infinite memory - equation (2.13a)
- 2 - Robast quantizer to channel errors - equation (2.18).
- 3 - Robast quantizer to channel errors with an extended dynamic range - equation (2.20).

פרק 3 – מקודדי אות הפרש (ADPCM)

במקודדי אות הפרש מנוצלת יתירות האות בתחום הזמן, לצורך חורדת קצב השידור. האות המקודד במקודדים אלו הוא ההפרש שבין האות לחיזויו. אנו בוחנים בעבודתנו מקודדים בהם החיזוי מבוצע בעזרת מסננים לינארים מתואמים (אדפטיבים). תרשים מלבנים עקרוני של מקודד אות הפרש בעל מסנן חיזוי המדמה קטבים באות הכניסה מתואר באיור 3.1. מסנן חיזוי זה מכונה מסנן קטבים (all pole predictor).



איור 3.1 – תרשים מלבנים של מקודד ADPCM עם מסנן חיזוי מסוג all-pole.
 Figure 3.1 – Block diagram of an ADPCM coder with an all-pole predictor.

בהעדר שגיאות ערוץ, האות המשוחזר ביציאת המפענח $Y(t)$ נבדל מאות הכניסה $X(t)$ בשגיאת הקוונטיזציה שבמקודד $q(t)$.

$$(3.1) \quad X(t) - Y(t) = d(t) - d_q(t) = q(t)$$

כאשר $d(t)$ הוא אות ההפרש ו- $d_q(t)$ הוא אות ההפרש במוצא הקוונטייזר. בקוונטייזר בעל צעד קוונטיזציה מתואם (אדפטיבי), וריאנס שגיאת הקוונטיזציה יחסי ישר לוריאנס האות בכניסה אליו. מכאן, ככל שוריאנס אות ההפרש σ_d^2 יהיה קטן יותר, כך יהיה קטן גם וריאנס שגיאת הקוונטיזציה σ_q^2 (הוא וריאנס שגיאת השיחזור). וריאנס האות בכניסה לקוונטייזר, מוקטן במקודדי אות הפרש בעזרת מסנני חיזוי. אנו נציג מספר אלגוריתמי אדפטיבי של מסנני חיזוי, המסתמכים על שערך קדמי ואחורי. אלגוריתמי האדפטיבי של מסנני החיזוי במקודד ה-ADPCM הסטנדרטי לקצב 32 kbps בו אנו משתמשים מוצג בפרק 4.

3.1 מסנן קטבים

נניח כי אות הכניסה למקודד X הוא תהליך אקראי סטציונרי מרקובי מסדר P הנתון על-ידי:

$$(3.2) \quad X(t) = \sum_{i=1}^P \alpha_i X(t-i) + U(t)$$

כאשר $\{\alpha_i\}$ הוא אוסף מקדמי התהליך ו-U הוא תהליך חסר קורלציה עם ממוצע אפס ווריאנס σ_u^2 . בשימוש במסנן חיזוי מסדר N (לאו דוקא שווה ל-P) בעל אוסף המקדמים $\{a_i\}$ המתואר באיור 3.1, ובהזנחת שגיאת הקוונטיזציה¹, כלומר $Y(t) \approx X(t)$, אות ההפרש d(t) מתקבל על-פי הביטוי הבא:

$$(3.3) \quad d(t) = X(t) - \sum_{i=1}^N a_i X(t-i)$$

ווריאנס אות ההפרש הוא:

$$(3.4) \quad \sigma_d^2 = \sigma_x^2 - 2\mathbf{a}^T \mathbf{r} + \mathbf{a}^T \mathbf{R} \mathbf{a}$$

כאשר \mathbf{a} הוא וקטור מקדמי המסנן, \mathbf{r} הוא וקטור האוטוקורלציה ו-R היא מטריצת האוטוקורלציה המוגדרים באופן הבא:

$$\begin{aligned} \mathbf{a} &= (a_1, \dots, a_N)^T \\ \mathbf{r} &= (r_1, \dots, r_N)^T, \quad r_i = E[X(t)X(t-i)] \\ \mathbf{R} &= \{R_{ij}\}, \quad R_{ij} = r_{|i-j|} \end{aligned}$$

מינימום של σ_d^2 יתקבל על-ידי אוסף מקדמים $\{a_i\}$, אשר יאפס את הגרדיאנט $\nabla \sigma_d^2(\mathbf{a})$. אוסף מקדמים זה הוא האופטימלי במובן מינימום שגיאה ריבועית ויסומן ב- \mathbf{a}_{opt} . ממשוואה (3.4) מקבלים

$$(3.5) \quad \mathbf{a}_{opt} = \mathbf{R}^{-1} \mathbf{r}$$

עבור $N=P$ מקבלים $\forall i \ a_{iopt} = \alpha_i$. אות ההפרש d(t) המתקבל במקרה זה, בהזנחת שגיאת הקוונטיזציה, הוא חידוש התהליך U(t), אשר לו ספקטרום שטוח (לבן). מכאן גם נקרא המסנן $1-A(z)$ "מסנן מלבין". כאשר A(z) הוא התמרת Z של הוקטור \mathbf{a} :

$$(3.6) \quad A(z) = \sum_{i=1}^N a_i z^{-i}$$

(1) ההצדקה להזנחה זו, היא ההפרש הנמוך ב-SNR, המתקבל עקב הכנסת הקוונטייזר לחוג המשוב. עבור תהליך אקראי גאוסי מרקובי מסדר ראשון עם $p_1 = 0.85$ (ערך אופייני לאותות דיבור), מתקבל הפרש הקטן מ-0.4 dB עבור קוונטייזר בעל 4 רמות [Jayant84].

3.1.1 שיערוך קידמי

בדרך כלל \underline{x} ו- R אינם נתונים מראש, והם משוערכים על-סמך דגמי האות. שיטת שיערוך בה מובטחת יציבות מסנני החיזוי (המקודד והמפענח) היא שיטת האוטוקורלציה, לפיה:

$$(3.7) \quad \hat{r}_i = \frac{1}{M} \sum_{t=t_0+i}^{t_0+M-1} X(t)X(t-i), \quad \hat{R}_{ij} = \hat{r}_{|i-j|}, \quad i, j=1, \dots, N$$

כאשר M הוא גודל הבלוק על פיו מבוצע השיערוך ו- t_0 הוא רגע התחלתו. השיקולים בבחירת גודל הבלוק דומים לאלו שהוזכרו בסעיף 2.3.1 הן בשיערוך וריאנס קדמי.

המטריצה \hat{R} המתקבלת היא טופליץ סימטרית בעלת מספר תכונות מועילות:

- (i) קיימות שיטות חישוביות מהירות להיפוכה, כמתבקש במשוואה (3.5).
- (ii) כל ערכי המטריצה מצויים בשורה/עמודה הראשונה. מכאן שלצורך אגירת מטריצה בגודל $N \times N$ זרושים N תאי זכרון בלבד.

כאמור לעיל, בשיטת האוטוקורלציה מובטחת יציבות המסננים. אולם באותות דיבור עבורם מתקבלים קטבים קרובים למעגל היחידה במישור Z , עלולים להתקבל מסננים לא יציבים עקב בעיות אורך מילה סופי. עבור מסננים מסדר ראשון ושני בעיה זו ניתנת לפתרון פשוט על-ידי הגבלת וקטור המקדמים \underline{a} לתחום ערכים יציב [Jayant84]. עבור מסננים מסדר גבוה יותר, מסנן סריג - Lattice Filter מחזוה פתרון אלגנטי ויעיל מבחינה חישובית שאינו רגיש לבעיות אורך מילה סופי. מסנן זה מבוסס על הפתרון הרקורסיבי של Levinson & Durbin כמתואר להלן [Porat86]:

הפתרון האופטימלי מסדר $n+1$ נבנה על-סמך הפתרון האופטימלי מסדר n . נסמן את וקטור מקדמי מסנן החיזוי האופטימלי מסדר n ב- \underline{a}_n . מתוך משוואות (3.4) ו-(3.5) מקבלים:

$$(3.8) \quad (1, -\underline{a}_n^T) \hat{R}_{n+1} = (r_n^f, \underline{0}_n^T)$$

כאשר \hat{R}_n היא מטריצת האוטוקורלציה המשוערכת מסדר n ו- $\underline{0}_n$ הוא וקטור אפסים באורך n ו- r_n^f הוא וריאנס אות ההפרש - d_n^f המתקבל ביציאת המסנן האופטימלי מסדר n , ו- $(\underline{a}, \underline{b}) = (a_1, a_2, \dots, b_1, b_2, \dots)$.

$$(3.9) \quad (1, -\underline{a}_{n+1}^T) \hat{R}_{n+2} = (r_{n+1}^f, \underline{0}_{n+1})$$

כעת נחפש פתרון מסדר $n+1$ אשר יקיים את השוויון הבא:

כאמור, \hat{R} היא מטריצה טופליץ סימטרית. לכן הכפלתה בוקטור מסובב בסדר הפוך נותנת תוצאה המסובבת בסדר הפוך. יהי \underline{a}_n וקטור מקדמי המסנן האופטימלי מסדר n בסדר הפוך, כלומר

$$\underline{a}_n = (a_{n,n}, a_{n,n-1}, \dots, a_{n,1})^T$$

אזי:

$$(-\underline{a}_n^T, 1) \hat{R}_{n+1} = (\underline{0}_n^T, r_n^f)$$

ניחוש פיתרון מסדר $n+1$ מהצורה $(\underline{a}_n, 0)$ והצבתו ב-(3.9) מוביל לשוויון הבא:

$$(3.10) \quad \begin{bmatrix} 1, -\underline{a}_n^T, 0 \\ 0, -\underline{a}_n^T, 1 \end{bmatrix} \hat{R}_{n+2} = \begin{bmatrix} r_n^f, \underline{0}_n, \gamma \\ \gamma, \underline{0}_n, r_n^f \end{bmatrix}$$

$$\gamma = r_{n+1} - \sum_{i=1}^n a_{n,i} r_{n+1-i}$$

כאשר

עם הכפלה משמאל של שני האגפים בוקטור $(1, -k_{n+1})$ כאשר $k_{n+1} = \gamma/r_n^f$ מתקבל הפתרון הרצוי כמתואר במשוואה (3.9). מכאן -

$$(3.11a) \quad (1, -\underline{a}_{n+1}^T) = (1, -k_{n+1}) \begin{bmatrix} 1, -\underline{a}_n^T, 0 \\ 0, -\underline{a}_n^T, 1 \end{bmatrix}$$

וריאנס אות השארית במוצא הדרגה ה- $n+1$ הוא

$$(3.11b) \quad r_{n+1}^f = r_n^f (1 - k_{n+1}^2)$$

מסנן אופטימלי מסדר N מתקבל על-ידי הפעלת (3.11) פעמים עם תנאי החתלה של מסנן מסדר אפס, עבורו וריאנס אות השארית הוא וריאנס האות - $r_o^f = r_o$. המקדמים $\{k_n\}$ נקראים מקדמי החזרה או מקדמי הקורלציה החלקיים (parcor). ממשוואות (3.10) ו-(3.11) נובע כי $|k_n| < 1$ תנאי אשר קל לשמרו גם לאחר קוונטיזציה המקדמים. ממשוואה (3.11b) מקבלים כמוצר לוואי גם את וריאנסי אותות השארית במוצא כל הדרגות. ניתן לנצל מידע זה לצורך התאמת צעד הקוונטיזציה, וכן יכול לשמש כפרמטר לקביעת סדר המסנן. הצעדים המבוצעים במקודד ובמפענח, כאשר וקטור מקדמי החזרה k משודר כאינפורמציה צד למפענח, ניתנים לסיכום באופן הבא [Jayant84]:

$$\begin{array}{l} \text{מקודד} : \quad [X(t)] \rightarrow \hat{R}_{N+1} \rightarrow \underline{a}_N, k_N \rightarrow Q(k_N) \rightarrow Q(\underline{a}_N) \\ \text{אינפורמציה צד} : \quad Q(k_N) \\ \text{מפענח} : \quad Q(k_N) \rightarrow Q(\underline{a}_N) \end{array}$$

$Q(\cdot)$ מייצג פעולת קוונטיזציה.

השלב האחרון - המבוצע במקודד ובמפענח, כרוך בסדרת הפעולות הבאות:

$$(3.12) \quad \begin{array}{l} a_{1,1} = k_1 \\ a_{n,j} = a_{n-1,j} - k_n a_{n-1,n-j} \quad j=1, \dots, n-1 \quad n=2, \dots, N \end{array}$$

ניתן לוותר על שלב אחרון זה על-ידי שימוש במסנן סריגי המתואר לחלק:

נרחיב את משוואה (3.11a) באופן הבא:

$$(3.13) \begin{bmatrix} 1 & -a_{n+1} \\ -a_{n+1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -k_{n+1} \\ -k_{n+1} & 1 \end{bmatrix} \begin{bmatrix} 1 & a_n & 0 \\ 0 & a_n & 1 \end{bmatrix}$$

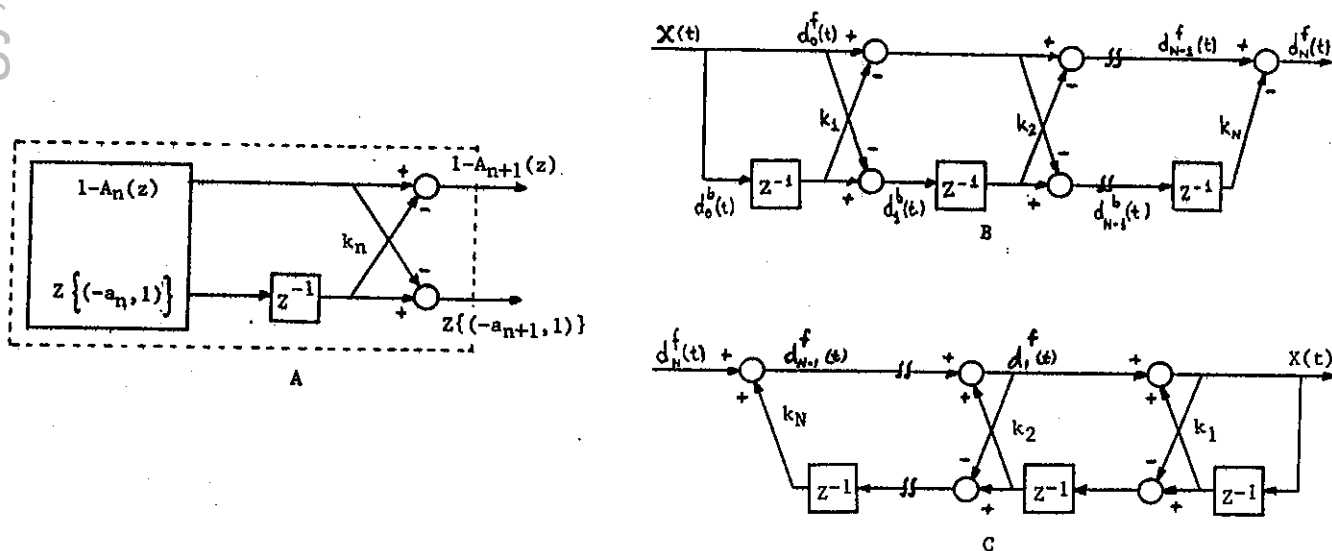
כעת נכפול את שני אגפי המשוואה מימין בוקטור Z_{n+1} הבא -

$$Z_{n+1} = [1, z^{-1}, \dots, z^{-n-1}]^T$$

לקבלת [Porat86]

$$(3.14) \begin{bmatrix} 1 - A_{n+1}(z) \\ Z\{(-a_{n+1}, 1)\} \end{bmatrix} = \begin{bmatrix} 1 & -k_{n+1} \\ -k_{n+1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} 1 - A_n(z) \\ Z\{(-a_n, 1)\} \end{bmatrix}$$

כאשר $Z(\cdot)$ מייצג התמרת Z ו- $A_n(z)$ היא התמרת Z של a_n . ממוש משוואה (3.14) מתואר באיור 3.2.



איור 3.2 - תרשים מלבנים של מסנן סריג. (A) תרשים של משוואה 3.14. (B) מסנן שיחזור הנמצא במפענת. d_n^b ו- d_n^f הם אותות החפרש הקידמי והאחורי במוצא הדרגה ה- n לית בהתאמה.

Figure 3.2 - Block diagram of a lattice filter. (A) block diagram of equation 3.14. (B) whitening filter in the encoder. (C) reconstruction filter in the decoder. d_n^f and d_n^b are the forward and backward difference signals at the n^{th} stage respectively.

במטרה להקטין את עומס החישוב הנדרש, מוצע ב- [Nakamura84] אלגוריתם תת אופטימלי לחישוב מקדמי ההחזרה. אלגוריתם זה משתמש בגישה חלקוחה מאלגוריתם ה-LMS אשר יתואר בסעיף הבא. מקדמי ההחזרה מעודכנים בכל דגם בכיוון הפוך לניגזרת של רבוע אות החפרש הקדמי במטרה למזער אותו. נסמן ב- d_n^f וב- d_n^b את אותות החפרש הקדמי והאחורי במוצא הדרגה ה- n ית-בחתאמה (ראה איור 3.2), כלומר -

$$(3.15) \quad \begin{aligned} d_n^f(t) &= d_{n-1}^f(t) - k_n d_{n-1}^b(t-1) \\ d_n^b(t) &= d_{n-1}^b(t-1) - k_n d_{n-1}^f(t) \end{aligned}$$

אזי, אלגוריתם האדפטציה נתון על-ידי:

$$(3.16) \quad k_n(t+1) = k_n(t) - \mu \frac{\partial d_n^f(t)}{\partial k_n} = k_n(t) + 2\mu d_n^f(t) d_{n-1}^b(t-1)$$

כאשר μ הוא גודל הקובע את מהירות ההתכנסות ומידת התנודות של אוסף המקדמים סביב נקודת המנוחה. במאמר מוצגים קירובים נוספים המקטינים במידה נוספת את העומס החישובי. קירובים דומים קיימים גם עבור אלגוריתם ה-LMS ועל כן לא יפורטו פה.

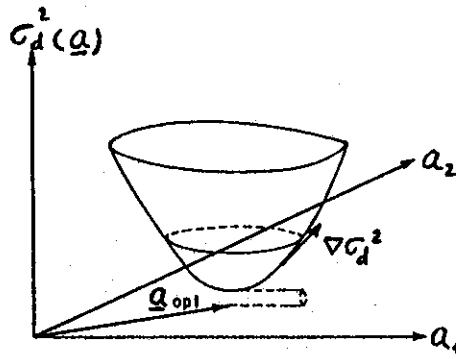
3.1.2 שערך אחורי

בשערך אחורי מחושבים מקדמי מסנני החיזוי על-סמך דגמי האות המקודד. על-ידי כך נחסך הצורך בשידור המקדמים כאינפורמציה צד, שכן הם ניתנים לחישוב גם במפענח. זאת ועוד, המקדמים מעודכנים מידי כל דגם ובכך נמנעת השחיית האות. במקום חישוב מלא של אוסף המקדמים, מבוצע תיקון ערכיהם בלבד על-פי אחד מתוך קריטריוני הטיב הבאים:

(i) ווריאנס אות החפרש - מתוך משוואה (3.4) רואים כי σ_d^2 היא תבנית ריבועית של \underline{a} , אותה ניתן לדמיין כמשטח פרבולואידי במרחב ה- N מימדי (כאשר N הוא סדר המסנן). נקודת המינימום של σ_d^2 נמצאת בכיוון מנוגד לגרדיאנט - $\nabla \sigma_d^2(\underline{a})$ (ראה איור 3.3), לכן וקטור המקדמים \underline{a} מתוקן בכיוון הפוך לו.

$$(3.17) \quad \underline{a}(t+1) = \underline{a}(t) - \mu \nabla \sigma_d^2(t) = \underline{a}(t) - 2\mu(R\underline{a} - \underline{r})$$

אלגוריתם עידכון זה אינו שימושי, משום שהוא כרוך בעומס חישובי רב. עומס זה נובע מהצורך לשערך מידי דגם את R ו- \underline{r} .



איור 3.3 - σ_d^2 כפונקציה של \underline{a} עבור מסנן מסדר שני. $\nabla \sigma_d^2$ הוא בכיוון הפוך לנקודת המינימום של σ_d^2 [Jayant84]

Figure 3.3 - σ_d^2 as a function of \underline{a} for a second order predictor. $\nabla \sigma_d^2$ is in opposite direction to the minimum point of σ_d^2 [Jayant84].

(ii) אות החפרש - בדרישה למיזעור $d^2(t)$ במקום σ_d^2 , מתקבלת הפחתה משמעותית של העומס החישובי. באלגוריתם זה, המכונה אלגוריתם ה-LMS (Least Mean Square), וקטור המקדמים מתוקן בכיוון הפוך לגרדיאנט של $d^2(t)$.

$$(3.18) \quad \underline{a}(t+1) = \underline{a}(t) - \mu \nabla d^2(t) = \underline{a}(t) + 2\mu d(t) \underline{X}(t)$$

$$\underline{X}(t) = [X(t-1), \dots, X(t-N)]^T$$

כאשר

כאמור, תהליך האדפטציה מבוצע על סמך הדגמים המקודדים \underline{X}, d . ב-(3.18) מוחלפים באות החפרש במוצא הקוונטייזר - d_q ובאות המשוחזר \underline{Y} בהתאמה.

$$(3.19) \quad \underline{a}(t+1) = \underline{a}(t) + 2\mu d_q(t) \underline{Y}(t)$$

ניתן לחקטין במידה נוספת את העומס החישובי בשיטות התת אופטימליות הבאות:
א. במקום מיזעור של $d_q^2(t)$ מנסים למזער את $|d_q(t)|$ לקבלת:

$$(3.20) \quad \underline{a}(t+1) = \underline{a}(t) - \mu \nabla |d_q(t)| = \underline{a}(t) + \mu \text{sign}[d_q(t)] \underline{Y}(t)$$

ב. במשוואה (3.19) מוחלף הביטוי לקורלציה שבין d_q ל- \underline{Y} בקורלציה שבין סימני האותות - $\text{sign}[d_q(t)] \cdot \text{sign}[\underline{Y}(t)]$. אנו נכנה אלגוריתם זה בשם אלגוריתם ה-"sign", או קירוב ה- $\text{sign}(\cdot)$ לאלגוריתם הגרדיאנט.

מאחר שאוסף המקדמים מחושב במפענח על-סמך האות המשודר, האלגוריתם רגיש לשגיאות ערוץ. בעיה זו מטופלת על-ידי חוספת גורם דעיכה ל- a , כפי שמתואר להלן עבור אלגוריתם ה- $\text{sign}(\)$

$$(3.21) \quad a_i(t+1) = (1-\alpha)a_i(t) + \mu \text{sign}[d_q(t)]\text{sign}[Y(t-i)]$$

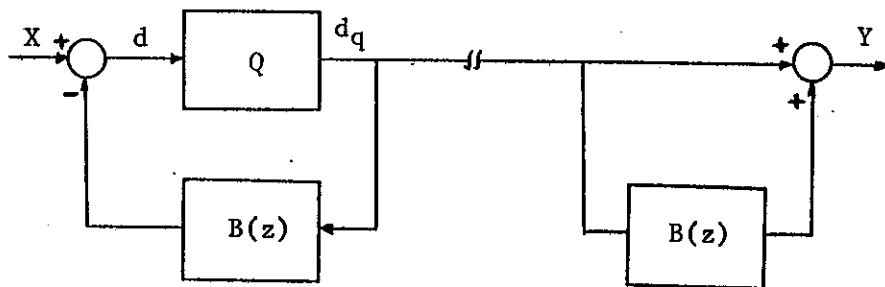
$$0 \leq \alpha, \mu < 1, \quad i=1, \dots, N$$

α הוא גודל הקובע את זמן הדעיכה.

3.2 מסנן אפסים

עד כה עסקנו במסנני חיזוי המדמים קטבים באות הכניסה. בשל הקושי הכרוך בשמירה על יציבות המסננים, מגבילים לעתים את הסדר שלהם ל-2. לקבלת מסנן חיזוי מסדר גבוה יותר, ניתן להשלים את מסנן הקטבים במסנן חיזוי המדמה אפסים (all zero predictor) אשר בו לא קיימת בעיית היציבות. מסנן מסוג זה מתואר במשוואה (3.22) ובאיור 3.4.

$$(3.22) \quad \begin{cases} d(t) = X(t) - \sum_{i=1}^N b_i d_q(t-i) \\ Y(t) = d_q(t) + \sum_{i=1}^N b_i d_q(t-i) \end{cases}$$



איור 3.4 - מקודד אות הפרש עם חיזוי מסוג all-zero. $B(z)$ הוא התמרת Z של \underline{b} .

Figure 3.4 - A coder with an all zero predictor. $B(z)$ is the Z transform of \underline{b} .

ב- [Nishitani82] מוצג אלגוריתם אדפטציה מקדמי המסנן הנגזר מתוך השאיפה למזער את ריבוע אות ההפרש $d(t)$. ביצוע קירובים דומים לאלו אשר בוצעו בביטוי (3.21) מובילים לאלגוריתם האדפטציה הבא:

$$(3.23) \quad b_j(t+1) = (1-\alpha)b_j(t) + \mu \text{sign}[d_q(t)]\text{sign}[d_q(t-j)], \quad j=1, \dots, N$$

פרק 4 – מקודד ADPCM סטנדרטי לקצב 32 kbps

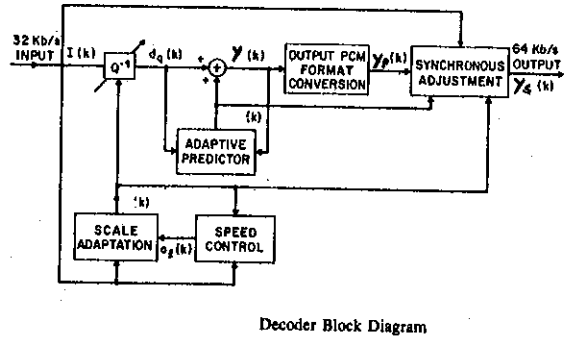
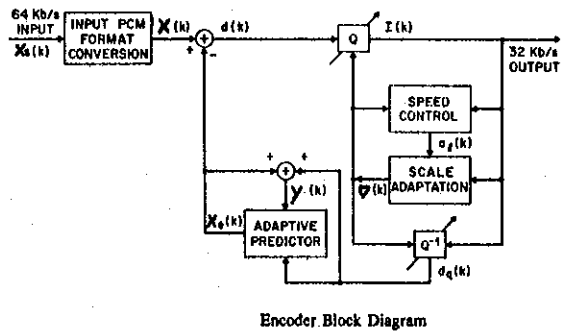
בפרק זה אנו מגישים סקירה על מקודד ADPCM לקצב 32 kbps, שאושר כתקן חדש לשימוש בערוץ הטלפון על-ידי CCITT באוקטובר 1984. תחילה מוצגות דרישות התכנון, לאחר מכן מתואר מבנה המקודד ותוצאות סימולציות שבוצעו. הסקירה המוגשת כאן בנושא זה רחבה, שכן מקודד זה שימש כבסיס לתיכנון מערכת הקידוד המוצגת בעבודה זו.

4.1 דרישות התיכנון

- ביוני 1982 מונתה ועדת מומחים של CCITT לקביעת דרישות התיכנון למקודד ADPCM, לשם דחיסה של 2:1 בקצב השידור, ביחס לתקן הקיים לקצב 64 kbps. דרישות התיכנון היו:
- (1) המקודד יותאם לקידוד אות הדגום בתדר 8 KHz, ומיוצג על-ידי 8 bit/sample, תוך שימוש בקוונטיזר לוגריתמי עם פונקציית דחיסה A-Law או μ -Law ($A=87.6$ ו- $\mu=255$). יציאת המקודד היא אות, הדגום באותו הקצב המיוצג על-ידי 4 bit/sample.
 - (2) למניעת הצורך בשידור אינפורמציה צד וכן להשגת זמן השחיה מינימלי של האות המקודד, תהיה אדפטציה מקדמי המסנן וצעד הקוונטיזציה אחורית באופייה.
 - (3) על המקודד להתאים לסוגי האותות השונים המועברים בערוץ הטלפון, והחשובים שביניהם הם אותות דיבור ו-VBD.
 - (4) על האלגוריתם להיות עמיד בפני שגיאות ערוץ. כלומר, דרושה עקיבה של פרמטרי המפענח אחר פרמטרי המקודד גם בנוכחות שגיאות ערוץ.
 - (5) האלגוריתם יכיל מנגנון אשר יבטל סכימת שגיאות הנובעות מחיבור טורי של מספר דרגות ברשת ספרתית (tandeming).

4.2 תאור האלגוריתם

תרשים מלבנים של מקודד ה-ADPCM הסטנדרטי לקצב 32 kbps מתואר באיור 4.1. המקודד מורכב מקוונטיזר וחזאי אדפטיביים: הקוונטיזר מתואם לפילוג גאוסי בעל 16 רמות, תחום דינמי $\Delta_{\max}/\Delta_{\min} = 491$ ומהירות אדפטציה הנקבעת על-פי מידת סטציונריות האות. החזאי מסוג pole-zero, מורכב ממסנן אפסים מסדר שיש, ומסנן קטבים מסדר שני. למקדמי המסנן וצעד הקוונטיזציה נוסף גורם דעיכה באלגוריתם האדפטציה, המאפשר לפרמטרים אלו במפענח להתכנס לערכי פרמטרי המקודד גם בנוכחות שגיאות ערוץ. המקודד והמפענח מכילים יחידות המרה (Format conversion) linear to log, log to linear בהתאמה. בנוסף, המפענח מכיל יחידת סינכרון למניעת סכימת שגיאות הנובעות מחיבור טורי של מספר דרגות.



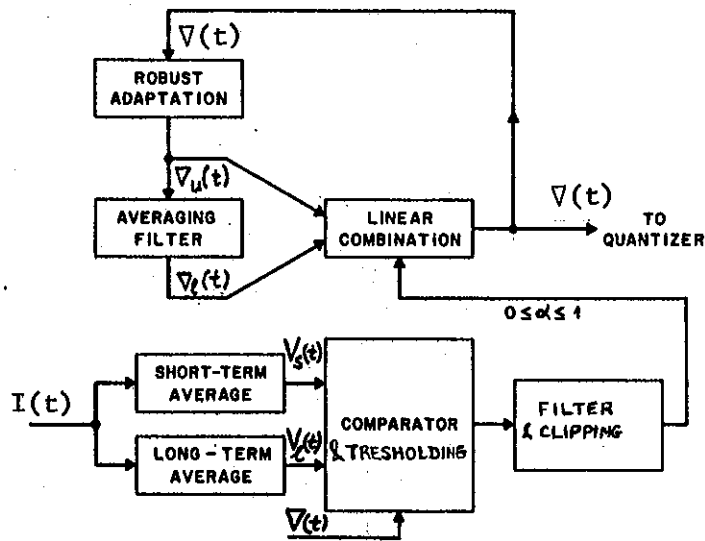
איור 4.1 - תרשים מלבנים של מקודד ה-ADPCM הסטנדרטי לקצב 32 kbps [Daumer84].
 Figure 4.1 - Block diagram of the standard 32 kbps ADPCM coder [Daumer84].

4.2.1 קוונטיזציה אדפטיבית

הקוונטיזציה מורכב מקוונטיזציה קבוע בזמן, כאשר אות הכניסה אליו מנורמל על-ידי גורם נירמול (המכונה בעבודה זו גם כצעד הקוונטיזציה) לקבלת אות בעל וריאנס קבוע. במטרה להקטין את עומס החישובים, הקוונטיזציה והאדפטיציה מבוצעים בתחום הלוגריתמי לפי בסיס 2.

רמות הקוונטיזציה: נעשה שימוש בקוונטיזציה Max לפילוג גאוסי המתאים לאותות VBD (אותות Modem בקצב 4800 bps). איכות אות הדיבור המתקבל לא הושפעה באופן משמעותי מבחירה זו [Daumer84].

האדפטיציה: תהליך האדפטיציה מבוסס על משעריך וריאנס אחורי, עמיד לשגיאות ערוץ הדומה לזה אשר מתואר במשוואה (2.18). המתכננים עמדו בפני בעיה בעת קביעת מהירות האדפטיציה. עבור אותות דיבור, שהם אי סטציונרים ובעלי תחום דינמי רחב, דרושה אדפטיציה מהירה. לעומת זאת, עבור אותות VBD, שהם בעלי עוצמה קבועה יחסית, מתאים קוונטיזציה בעל צעד קוונטיזציה מכוון הקבוע בזמן [Petr82]. פשרה בקביעת מהירות האדפטיציה גוררת הרעה בביצועים, הן עבור אותות דיבור והן עבור אותות VBD. [Petr82] מציע פתרון חדשני בו מותאמת מהירות האדפטיציה למידת סטציונריות האות. אופן עבודה במהירות אדפטיציה מינימלית, המתאימה לאותות VBD, מכונה - locked; אופן עבודה במהירות אדפטיציה מכסימלית, המתאימה לאותות דיבור, מכונה - unlocked ומערכת הבקרה על מהירות האדפטיציה מכונה על-ידי המחבר (DLQ Dynamic Locking of the Quantization step size). מערכת זו מתוארת באיור 4.2.



איור 4.2 - סכימת מלבנים של מערכת ה-DLQ.
 Figure 4.2 - Block diagram of the DLQ system.

נסמן ב- ∇ ; ∇_u ; ∇_l את ה- $\log_2(\cdot)$ של: צעד הקוונטיזציה, הצעד בעל קבוע זמן המחיר ביותר והצעד בעל קבוע זמן האיטי ביותר בהתאמה.
 ∇_u מתקבל בדומה לאלגוריתם האדפטציה המתואר ב-(2.18), תוך שימוש באוסף כופלים ממנו נגזר קבוע זמן מחיר:

$$(4.1) \quad \nabla_u(t) = \beta \cdot \nabla(t-1) + W(i_t), \quad \beta = 31/32$$

∇_l מתקבל מתוך מיצוע של ∇_u (העברת ∇_u דרך מסנן מעביר נמוכים).

$$(4.2) \quad \nabla_l(t) = 31/32 \nabla_l(t-1) + 1/32 \nabla_u(t)$$

שילוב שתי מהירויות האדפטציה מתקבל על-ידי שיקלולם הליניארי בהתאם למידת סטציונריות האות.

$$(4.3) \quad \nabla(t) = \alpha(t) \cdot \nabla_u(t) + [1-\alpha(t)] \cdot \nabla_l(t), \quad 0 \leq \alpha(t) \leq 1$$

$\alpha(t)$ מתאר את מידת סטציונריות האות. עבור אות סטציונרי $\alpha(t) = 0$. במצב זה $\nabla(t) = \nabla_l(t)$ והקוונטיזציה נמצא במצב locked עבור אות אי-סטציונרי $\alpha(t) = 1$. במצב זה $\nabla(t) = \nabla_u(t)$ והקוונטיזציה נמצא במצב unlocked. הקוונטיזציה נמצא במצב unlocked גם כאשר הערוץ נמצא במצב idle, לשם חכנת הקוונטיזציה להופעת אות חדש.

חישוב $\alpha(t)$: מידת סטציונריות האות, נקבעת ע"פ השינויים בוריאנס אות החפרש (הוא האות בכניסה לקוונטיזציה) לזמן קצר. נסמן ב- V_s את משעריך הוריאנס לזמן קצר. אות הכניסה הוא תהליך אקראי, מכאן שגם V_s הוא תהליך אקראי. ככל שהאות סטציונרי יותר, כך הוריאנס המנורמל של V_s קטן יותר. כלומר, עוצמת התנודות של V_s סביב נקודת המנוחה שלו תהינה קטנות יותר. $\alpha(t)$ הוא מדד

רגעי לעוצמת תנודות אלו.

באלגוריתם זה נעשה שימוש במשעריך וריאנס פשוט העמיד לשגיאות ערוץ:

$$(4.4) \quad V_s(t) = 31/32 V_s(t-1) + 1/32 F(i_t)$$

$\{F(i)\}$ הוא אוסף ערכים המחליף את הגודל d_q^2 .

משעריך וריאנס האות V_ℓ מתקבל ע"י ביטוי דומה ל-(4.4), אך תוך שימוש בקבוע זמן ארוך יותר

$$(4.5) \quad V_l(t) = 127/128 V_l(t-1) + 1/128 F(i_t)$$

ניתן לשערך את עוצמת התנודות המנורמלות של V_s סביב נקודת המנוחה שלו על-ידי סכום הריבועים -

$\{|V_s(t) - V_\ell(t)| / V_\ell(t)\}^2$. במקום זאת מבוצעת סכימה פשוטה היחסית להפרשים - $|V_s(t) - V_\ell(t)|$

$$(4.6) \quad \alpha_p(t) = \begin{cases} 15/16 \alpha_p(t-1) + 2^{-3}, & \nabla(t) < 3 \\ 15/16 \alpha_p(t-1) + 2^{-3}, & |V_s(t) - V_l(t)| > 2^{-3} V_l(t) \\ 15/16 \alpha_p(t-1), & \text{else} \end{cases}$$

$\alpha_p(t)$ הוא מדד ראשוני לסטציונריות האות, הנע בתחום ערכים $0 \leq \alpha_p(t) < 1$, ממנו נגזר $\alpha(t)$.

$$(4.7) \quad \alpha(t) = \begin{cases} \alpha_p(t), & \alpha_p(t) \leq 1 \\ 1, & \alpha_p(t) > 1 \end{cases}$$

המקרה של $\nabla(t) < 3$ מתאים למצב בו הערוץ ב-idle.

4.2.2 חזאי אדפטיבי

מסנן החיזוי מורכב משני מסננים: מסנן all-zero מסדר שני, ומסנן all-pole מסדר שני. הבחירה הסופית של מבנה זה נבעה מחשיקולים הבאים: ביצועים עבור אותות צרי סרט ורחבי סרט כאחד, יציבות, עקיבה נכונה של המפענח אחר המקודד גם בנוכחות שגיאות ערוץ ויעילות חישובית. מסנן חיזוי בעל קטבים ואפסים, התקבל כיעיל יותר מבחינה חישובית ממסנן בעל אפסים בלבד, כאשר שאר הביצועים נלקחים בחשבון [Daumer84]. האות המשוערך - X_e מתקבל כסכום של תפוקת מסנן הקטבים - X_{ep} ותפוסת מסנן האפסים - X_{ez} .

$$(4.8) \quad \begin{cases} X_e(t) = X_{ep}(t) + X_{ez}(t) \\ X_{ep}(t) = \sum_{i=1}^2 a_i(t) Y(t-i) \\ X_{ez}(t) = \sum_{i=1}^6 b_i(t) d_q(t-i) \end{cases}$$

כאשר $Y(t)$ הוא האות המשוחרר הנתון על-ידי

$$(4.9) \quad Y(t) = X_e(t) + d_q(t)$$

מקדמי מסנן האפסים: אדפטציה מקדמי מסנן האפסים מבוצעת על-ידי קירוב ה- $\text{sign}(\cdot)$ לאלגוריתם הגרדיאנט כמתואר במשוואה (3.23).

$$(4.10) \quad b_i(t+1) = (1-\delta_1) \cdot b_i(t) + \delta_2 \cdot \text{sign}[d_q(t)] \cdot \text{sign}[d_q(t-i)]$$

$$i = 1, \dots, 6 \quad \delta_1 = 2^{-8} \quad \delta_2 = 2^{-7}$$

מקדמי מסנן הקטבים: יציבות ועקיבה⁽¹⁾ הן בעיות מרכזיות במסנן קטבים אדפטיבי. יציבות המסנן מובטחת על-ידי שימוש במסנן מסדר שני בלבד והגבלת ערכי a_1 ו- a_2 לתחום מסוים. שורש בעית העקיבה נעוץ בכך שמקדמי מסנן הקטבים הם פונקציה של האות המשוחרר $Y(t)$. אות האוגר בחובו את כל העבר כולל שגיאות הערוץ שחתרשו. ב-[Miller84] מוצע פתרון לנושא זה המתקבל על-ידי אדפטציה מקדמי מסנן הקטבים מתוך האות בכניסה למסנן (כדוגמת מסנן האפסים) במקום האות המשוחרר כפי שיוסבר להלן:

תגובת מסנן הקטבים במישור Z , ניתנת לביטוי על-ידי מסנן בעל אינסוף אפסים

$$(4.11) \quad [1 - \sum_{i=1}^2 a_i z^{-i}]^{-1} = 1 + \sum_{i=1}^{\infty} c_i z^{-i}$$

$\{C_i\}$ הוא אוסף מקדמי מסנן האפסים הדואלי. מקדמים אלו מעודכנים על-פי אלגוריתם אדפטציה כמתואר במשוואה (4.10), אשר יתואר בקצור באופן הבא:

$$(4.12a) \quad c_i(t+1) = (1 - \gamma_i) c_i(t) + \Delta c_i(t), \quad \gamma_i \ll 1$$

הצבה דומה נבצע עבור מקדמי מסנן הקטבים

$$(4.12b) \quad a_i(t+1) = (1 - \alpha_i) a_i(t) + \Delta a_i(t), \quad \alpha_i \ll 1$$

כעת נציב את Δa_i כפונקציה של ΔC_j . מתוך השוואת מקדמים של Z^{-i} במשוואה (4.11), בהנחות $\alpha_i, \gamma_i \approx 0$, $\Delta C_1 \ll a_1$, וכן תוך קיצוץ הטור $\sum C_i Z^{-i}$ לשני האיברים הראשונים מקבלים

$$(4.12c) \quad \Delta a_1 = \Delta c_1, \quad \Delta a_2 = \Delta c_2 - 2a_1 \Delta a_1$$

(1) בעיית העקיבה קיימת כאשר למסנן הקטבים יש מספר נקודות מנוחה יציבות. כידוע, אלגוריתם הגרדיאנט מחפש אחר מינימום מקומי של וריאנס אות החפרש. לאותות כניסה בעלי מספר תדירויות תהודה מרוחקות זו מזו, קיימות מספר נקודות מנוחה של המסנן. באופן זה, עבור תנאי התחלה שונים או שגיאות ערוץ, מסנן החיזוי במפענח עלול להגיע לנקודת מנוחה השונה מזו של המקודד.

עם הוספת גורמי הדעיכה $\{\alpha_i\}$ בחישוב $\{a_i\}$ נקבל את אלגוריתם האדפטציה הבא:

$$(4.13) \quad a_1(t+1) = (1-\alpha_{11}) \cdot a_1(t) + \alpha_{12} \cdot \text{sign}[Y_Z(t)] \cdot \text{sign}[Y_Z(t-1)]$$

$$a_2(t+1) = (1-\alpha_{21}) \cdot a_2(t) + \alpha_{22} \cdot \{\text{sign}[Y_Z(t)] \cdot \text{sign}[Y_Z(t-2)] - f[a_1(t)] \cdot \text{sign}[Y_Z(t)] \cdot \text{sign}[Y_Z(t-1)]\}$$

$$\alpha_{11} = 2^{-8}, \quad \alpha_{21} = 2^{-7}, \quad \alpha_{12} = 3 \cdot 2^{-8}, \quad \alpha_{22} = 2^{-7}$$

על-פי משוואה (4.12c), $f(a_1) = 2a_1$, אך מתוך מדידות לקבלת SNR מכסימלי לאות כניסה סינוסי בתחום תדרים רחב נקבע $f(a_1)$ הבא:

$$f(a_1) = \begin{cases} 4a_1 & , |a_1| \leq 1/2 \\ 2 \cdot \text{sign}(a_1) & , |a_1| > 1/2 \end{cases}$$

להבטחת יציבות המסנן, מוגבלים (על-ידי קיצוץ) מקדמי המסנן לתחום הערכים:

$$|a_2| \leq 0.75, \quad |a_1| \leq 15/16 - a_2$$

4.2.3 יחידת הסינכרון

תפקיד יחידה זו הוא למנוע סכימת שגיאות הנובעות מחיבור טורי של מספר דרגות מקודד/מפענח ברשת ספרתית. אופן פעולת יחידה זו הוא כדלקמן: ביציאת המפענח מבוצעת הדמייה של דרגת הקידוד הבאה לקבלת מילת הקוד $i(t)$ שלה. מילת קוד זו משווית עם מילת הקוד של המפענח. במקרה של אי שיוויון בין השתיים, מבוצע תיקון מילת האות המשוחזר ביציאת המפענח.

4.3 תוצאות סימולציה

בסעיף זה אנו מגישים תוצאות סימולציה אשר בוצעו בטרם אושר האלגוריתם כתקן חדש. האלגוריתם נבחן בקידוד אותות דיבור, אותות סינטטיים ואותות נתונים, בערוץ שקט ורועש, תוך השוואת הביצועים עם מקודד PCM לקצב 64 kbps.

4.3.1 קידוד אותות דיבור

המקודד שהוצע נבחן בקידוד אותות דיבור, המכילים מספר רב של משפטים, אשר נאמרו על-ידי דוברים ודוברות בשש שפות שונות. תוצאות אלו מוצגות במונחים של MOS (Mean Opinion Score). בין קבוצות המאזינים השונות התקבל פיזור ערכי MOS רב. יחד עם זאת, היתה עקביות בשינוי ה-MOS

כפונקציה של מספר הדרגות הטוריות, שגיאות הערוץ, וכן ביחס שבין תוצאות ה-MOS שהתקבלו על-ידי מקודד ה-ADPCM לבין מקודד ה-PCM. באיור 4.3 מוצגות תוצאות ה-MOS כפונקציה של מספר הדרגות האסינכרוניות בטור (בחיבור אסינכרוני קיים ערוץ אנלוגי בין דרגה אחת לשניה), ובאיור 4.4 מוצגות תוצאות MOS כפונקציה של הסתברות השגיאות בערוץ (שבין המקודד למפענח). נסכם תוצאות אלה:

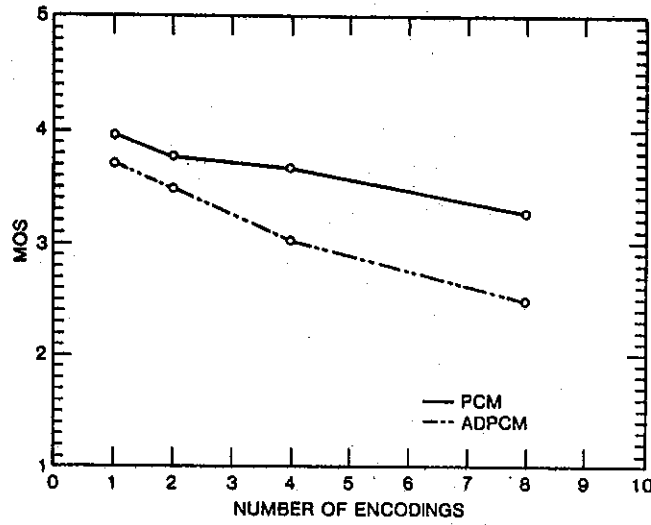
- (1) איכות אות החיבור המתקבלת במוצא מקודד ה-ADPCM, שקולה לזו המתקבלת במוצא שלוש עד ארבע דרגות PCM בחיבור טורי אסינכרוני. תוצאות אשר לדברי המחברים הן למעלה ממשביעות רצון עבור רוב השימושים במערכות התקשורת.
- (2) המקודד עמיד לחיבור טורי סינכרוני. תכונה זו חשובה ביותר כאשר נשקלת התקנת מערכות אלו ברשתות תקשורת ספרתיות.
- (3) מקודד ה-ADPCM רגיש פחות לשגיאות ערוץ מאשר מקודד ה-PCM.

4.3.2 קידוד אותות נתונים

תוצאות הסימולציה המדווחות ב-[Belfield84] מתייחסות ברובן לאותות VBD בקצב נתונים 1200:9600 bps. המקודד נבדק גם בקידוד אותות סינטטיים למדידת ה-SNR, הגבר והזזת מופע כפונקציה של תדר ועוצמת האות. המחברים מעירים כי למרות המאמץ שנעשה לכיול המדידות השונות, אין להתייחס אליהן באופן בלתי תלוי, אלא תוך השוואת ביצועי ה-ADPCM ביחס למקודד ה-PCM. במהלך מדידות אלה נבדקו הביצועים המתקבלים ביציאת דרגה בודדת עד לחיבור של 4 דרגות אסינכרוניות, השפעת רעש עיוותים והסחת מופע. פירוט התוצאות עבור האפננים (Modems) השונים בקצבים הנידונים מופיע ב-[Belfield84]. להלן מוגש סיכום התוצאות הנמדדות ב-BLER (Block Error Rate).

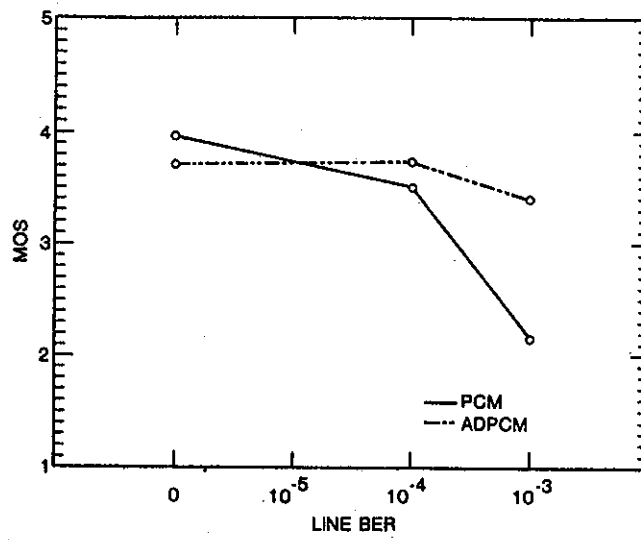
- (1) בקצבי נתונים קטנים מ-2400 bps התקבלו הבדלים מזעריים בלבד בין המקודדים השונים.
- (2) בקצב נתונים של 2400 bps התקבל על-ידי מקודד ה-ADPCM BLER הגבוה בממוצע פי 5 מזה המתקבל על-ידי מקודד ה-PCM.
- (3) בקצב 4800 bps ה-BLER הממוצע המתקבל על-ידי ה-ADPCM גבוה מזה המתקבל על-ידי ה-PCM פי 15.
- (4) בקצב 9600 bps מתקבל על-ידי ה-ADPCM BLER גבוה אשר אינו מאפשר שימוש בו ברוב המקרים. שימוש במקודד מסוג זה יצריך מחקר נוסף.

-
- (2) תוצאות מקורבות אלה נלקחו מתוך הגרפים ב-[Belfield84], תוך התחשבות בתוצאות עבורם מתקבל BLER בתחום 10^{-4} ל- 10^{-1} על-ידי מקודד ה-PCM ומיצוע התוצאות המתקבלות עבור האפננים השונים שנבדקו באותו הקצב.



איור 4.3 - מיצוע תוצאות MOS כפונקציה של מספר הדרגות האסינכרוניות. [Suyderhoud84]

Figure 4.3 - Average MOS vs. number of Encodings.



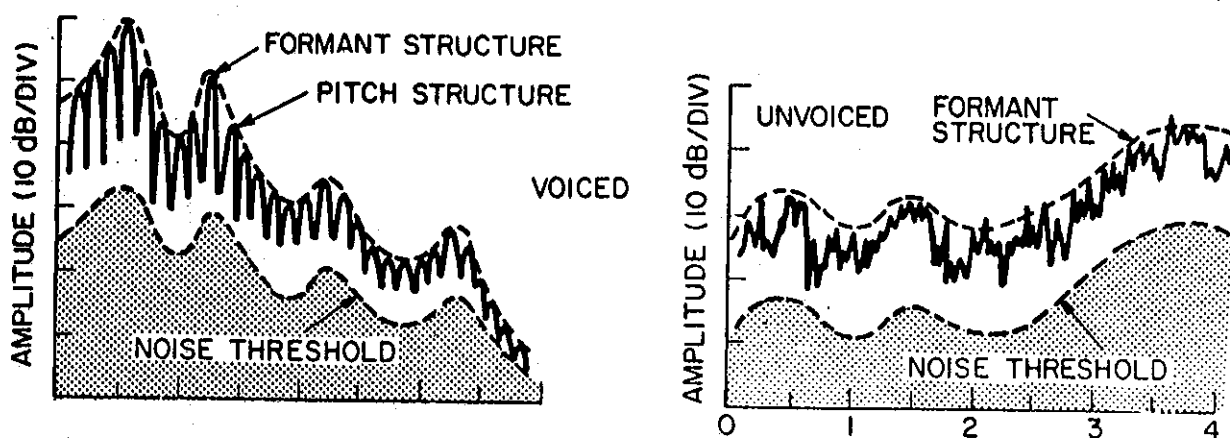
איור 4.4 - מיצוע תוצאות MOS כפונקציה של שכיחות שגיאות הערוץ. [Suyderhoud84]

Figure 4.4 - Average MOS vs. channel BER.

פרק 5 – עיצוב ספקטרום שגיאת הקוונטיזציה

במקודד ADPCM אופייני נמצא תקוונטיזציה בחוג החיזוי כמתואר באיור 3.1. הראינו כי שגיאת השחזור המתקבלת במקודד זה היא שגיאת הקוונטיזציה. בקוונטיזציה מתואם וריאנס שגיאת הקוונטיזציה יחסי ישר לוריאנס האות בכניסה אליו. מכאן, שחזאי אופטימלי, המביא למינימום את וריאנס אות החפרש, הוא זה אשר מביא גם למינימום את וריאנס שגיאת השחזור. אולם, האיכות הסוביקטיבית של אות הדיבור המשוחזר אינה תלויה אך ורק בוריאנס שגיאת השחזור. להשגת איכות גבוהה יש צורך להתחשב גם בספקטרום השגיאה ביחס לספקטרום אות הדיבור.

ב-[Crochiere82] מוצג מודל לאוזן האנושית המורכב ממערך מסננים. בכל פס תדר (ערוץ) במערך זה קיימות בקרת הגבר ובקרת סף עוצמה בהתאם לעוצמת האות באותו הערוץ. אות (רעש) אשר עוצמתו נמוכה מרמת הסף ממוסך ואינו מורגש. בנוסף לכך – עקב מבנה האוזן, אות בעוצמה גבוהה בתחום התדרים הנמוכים, מוריד את רגישות האוזן בתדרים הגבוהים. מסיבה זו ניתן לשאת רעש ב-SNR נמוך יותר בתחום התדרים הגבוה. באיור 5.1 מתואר ספקטרום הספק אופייני של אותות דיבור קוליים ואי-קוליים, וכן תחום עוצמות הרעש הנלווה הממוסך על-ידי האות.



איור 5.1 – ספקטרום הספק אופייני של אותות דיבור קוליים (a) ואי-קוליים (b). האיזור המוצל הוא תחום עוצמת הרעש הממוסך על-ידי האות [Crochiere82].

Figure 5.1 – Typical power spectrum of voiced (a) and unvoiced (b) speech signals. The shaded area is the power ranged, masked by the signal [Crochiere82].

בפרק זה אנו מציגים ומשווים שתי שיטות מקובלות לשיפור איכות האות המשוחזר המתקבל על-ידי מערכת קידוד ADPCM. בשיטה אחת הקרויה Noise Feedback Coding (NFC) מעוצב ספקטרום שגיאת הקוונטיזציה בעזרת מסנן משוב, ובשיטה שניה מועבר האות המשוחזר במוצא המפענת דרך מסנן שחזור נוסף – Postfiller (PF).

5.1 Noise Feedback Coding

בפרק 3 הראינו כי ללא שגיאות ערוץ האות המשוחזר Y הוא $Y = X - e_q$, כאשר e_q היא שגיאת הקוונטיזציה המוגדרת על-ידי $e_q = d - d_q$. מכאן קל לראות שהתמרת Z של אות החפרש $d(t)$ הוא אות הכניסה של הקוונטייזר (ראה איור 3.1):

$$(5.1) \quad d(z) = X(z)[1 - P(z)] + e_q(z)P(z)$$

לפי ביטוי זה אות החפרש $d(t)$ הוא סיכום של אות הכניסה המועבר דרך מסנן מלבין $1 - P(z)$, ושגיאת הקוונטיזציה המועברת דרך המסנן $P(z)$. באיור 5.2a מוצג תרשים מלבנים המתאר באופן ישיר משוואה זו. בתרשים זה הוצא הקוונטייזר אל מחוץ לחוג החיזוי. חסינון אותו עוברת שגיאת הקוונטיזציה (כצד המקדד) מתואר בעזרת מסנן משוב נפרד.

אם נחליף את מסנן משוב שגיאת הקוונטיזציה במסנן $G(z)$, $G(z) \neq P(z)$ (כמתואר באיור 5.2b) אזי יתקבל האות המשוחזר הבא:

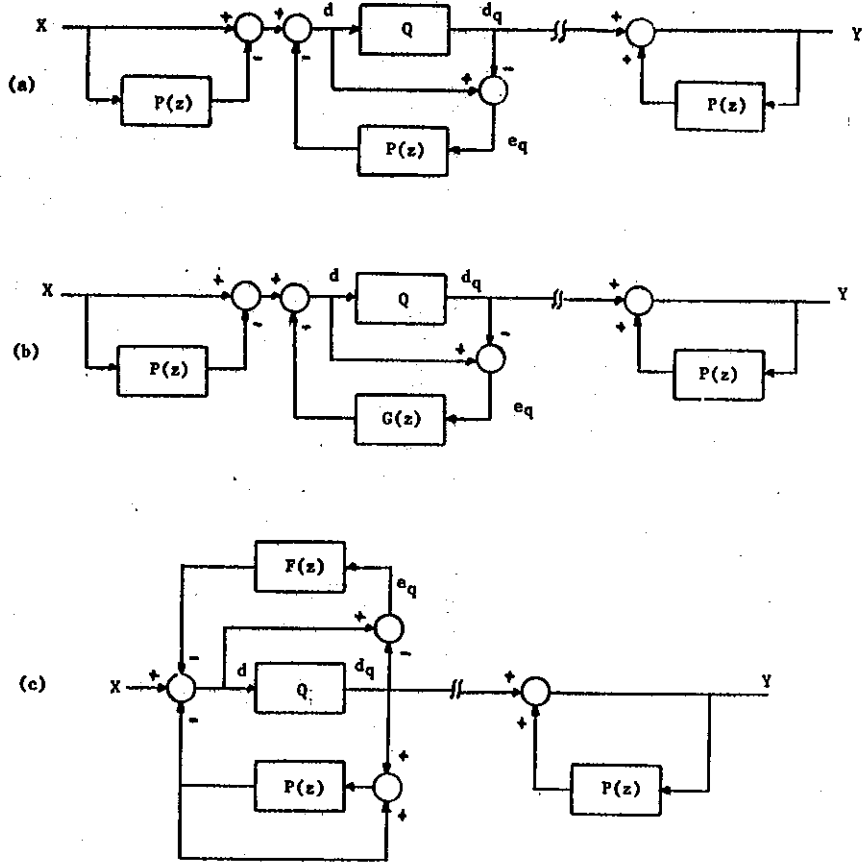
$$(5.2) \quad Y(z) = X(z) - e_q(z) \frac{1 - G(z)}{1 - P(z)}$$

מערכת קדוד זו מכונה ב-[Jayant84] - NFC.

עבור שגיאת קוונטיזציה לבנה נקבע ספקטרום שגיאת השחזור על פי המסנן $(1-G)/(1-P)$. בעזרת מסנן זה ניתן לעצב כרצוננו את ספקטרום שגיאת השחזור. עיצוב ספקטרום השגיאה נועד לתחום אותה לאיזור הממוסך על-ידי אות הדיבור, או במקרה שאין זה אפשרי (כאשר ה-SNR נמוך), לעצב את ספקטרום השגיאה על-פי קו הגבול של איזור זה. עבור אותות דיבור קוליים ספקטרום שגיאה שטוח מידי פירושו SNR נמוך באיזור התדרים הגבוהים, תופעה המתבטאת ברעש "hiss". ספקטרום שגיאה הקרוב מידי לצורת ספקטרום האות פירושו SNR נמוך באיזור התדרים הנמוכים, תופעה המתבטאת ב"חיספוס" האות המשוחזר [Makhoul79].

למערכת הקידוד המתוארת באיור 5.2b יש חסרון כאשר מדובר במערכת אדפטיבית אחורית. במערכת זו תחליך האדפטיציה מבוצע הן במקדד והן במפענח, והאות המשוחזר ביציאת המפענח אינו קיים גם במקדד. באיור 5.2c מתוארת מערכת קידוד NFC בה האות המשוחזר קיים גם במקדד. הקוונטייזר הוכנס חזרה לחוג החיזוי בתוספת מסנן משוב לשגיאת הקוונטיזציה - $F(z)$. כתוצאה מכך מתקבל האות המשוחזר הבא:

$$(5.3) \quad Y(z) = X(z) - e_q(z)[1 + F(z)]$$



איור 5.2 - תרשימי מלבנים של מערכות קידוד ADPCM (a) ומערכות NFC (b), (c).
 Figure 5.2 - Block diagram of ADPCM coder (a), and NFC coders (b), (c).

Postfiltering 5.2

שיטה נוספת לשיפור סוביקטיבי של איכות האות המשוחזר במערכות קידוד ADPCM מוצעת ב-[Jayant86] על-ידי תוספת מסנן שיחזור נוסף Postfilter (PF) במוצא המפענח (נסמן מקודד ADPCM בתוספת PF ב-ADPCM-PF):

$$(5.4) \quad Y_p(z) = [X(z) - e_q(z)]PF(z)$$

כאשר Y_p הוא האות המשוחזר במוצא ה-ADPCM-PF ו- $PF(z)$ היא פונקציית התמסורת של מסנן ה-PF במישור Z .

כפי שניתן לראות על-פי משוואה (5.4), מסנן ה-PF פועל הן על שגיאת הקוונטיזציה והן על האות המקודד. מכאן, שמסנן ה-PF אינו משנה את צורת ה-SNR כפונקציה של התדר - $SNR(f)$. כלומר, תדרים בהם התקבל SNR נמוך על-ידי מערכת ה-ADPCM ישארו באותו חיחס גם במערכת ה-ADPCM-PF. השיפור הסוביקטיבי באיכות האות המשוחזר מתקבל כתוצאה ממיסוך התדרים בעלי העוצמות הנמוכות (בהם מתקבל ה-SNR הנמוך) על-ידי תדירויות בעלי עוצמות גבוהות יותר (בהם מתקבל ה-SNR הגבוה).

לצורך ביצוע משימה זו דרוש כי למסנן ה-PF תהיה צורת ספקטרום הדומה לצורת ספקטרום האות המקודד. צורת הספקטרום של מסנן ה-PF, המתקבלת כפשרה בין המיסוך המבוקש לבין עיוות האות המתקבל, היא גרסה "מעודנת" של צורת ספקטרום אות הדיבור (בדומה לצורת הספקטרום של מסנן ה-NFC).

ב-[Jayant86] מוצע מסנן PF אדפטיבי מסוג pole-zero הנתון ב-(5.5), אשר תגובת התדר שלו נגזרת מתגובת התדר של מסנן השחזור במפענת.

$$(5.5) \quad PF(z) = \frac{1+B_P(z)}{1-A_P(z)}, \quad A_P(z) = \sum_{i=1}^2 a_i \alpha^i(t) z^{-i} \quad B_P = \sum_{i=1}^6 b_i \beta^i z^{-i}$$

$$0 \leq \alpha(t) \leq 1 \quad 0 \leq \beta \leq 1$$

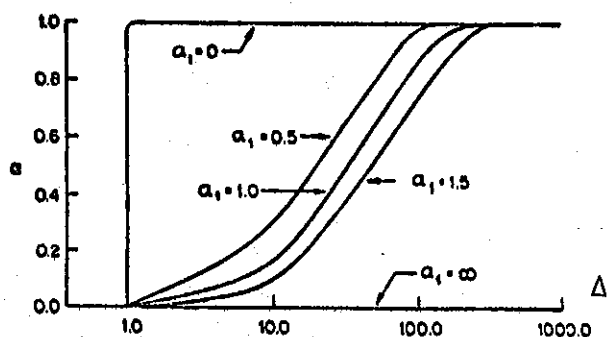
כאשר $\{a_i\}$ ו- $\{b_i\}$ הם אוסף מקדמי מסנן השחזור מסוג pole-zero ו- α, β הם גדלים הקובעים את מידת הסינון.

עבור $\alpha = \beta = 0$ מקבלים $FP(z) = 1$, כלומר מערכת קידוד ADPCM רגילה, עבור $\alpha = \beta = 1$ מקבלים מסנן PF זהה למסנן השחזור. בערכי ביניים של α, β מתקבל מסנן PF עם קטבים ואפסים בתדירויות זוויתיות כשל מסנן השחזור אך הקרובים יותר לראשית הצירים במישור Z.

מידת הסינון של מסנן ה-PF נקבעה ב-[Jayant86] באופן נסיוני: עבור מסנן האפסים - מידת סינון מכסימלית, קרי $\beta = 1$. עבור מסנן הקטבים - מידת סינון מועטה יותר המשתנה בזמן. זאת מאחר שמסנן הקטבים גורם לעיוות משמעותי יותר של האות המקודד. בקטעי זמן בהם האות משוחזר על-ידי מערכת ה-ADPCM באופן "מוצלח" לא דרוש סינון משמעותי, וערכי α נמוכים מספיקים. לעומת זאת, בקטעי זמן בהם מערכת ה-ADPCM נכשלת, דרוש סינון משמעותי ואז משתמשים בערכי α גבוהים יותר. Jayant מציע אלגוריתם פשוט, המנסה לעקוב אחר ביצועי ה-ADPCM, ומסתמך על פרמטרים המצביעים על רווח החיזוי - המקדם הראשון של מסנן השחזור מסוג all-pole, ועל עוצמת שגיאת הקוונטיזציה - צעד הקוונטיזציה:

$$(5.6) \quad \alpha(t) = 0.98\alpha(t-1) + 0.02e^{-20\alpha_1(t)(\Delta(t)-\Delta_{\min})}$$

בנוסף למשוואה (5.6) חסום $\alpha(t)$ מלמעלה על-ידי הערך 0.8. באיור 5.3 מתוארים ערכי α המתקבלים כתלות במקדם a_1 וצעד הקוונטיזציה Δ (עבור $\Delta_{\min} = 1$).



איור 5.3 - α כפונקציה של המקדם הראשון של מסגן הקטבים - a_1 וצעד הקוונטיזציה Δ (עבור $\Delta_{\min}=1$) [Jayant86]

Figure 5.3 - α as a function of the first all-pole predictor coefficient - a_1 and the quantizer step size Δ (for $\Delta_{\min}=1$) [Jayant86].

5.3 השוואת ביצועי המערכות השונות

בוצעה השוואה סוביקטיבית בלבד בין מערכות קידוד השונות שהוצגו לבין מקודד ADPCM רגיל. לצורך השוואה זו מומשו בסימולציות מחשב מערכת ה-ADPCM-PF שהוצגה ומערכת ה-NFC המתוארת באיור 5.3c עם מסגן $F(z)$ עבורו מתקבל האות המשוחזר הבא:

$$(5.7) \quad Y(z) = X(z) - e_q(z) \frac{1+B_F(z)}{1-A_F(z)}, \quad A_F(z) = \sum_{i=1}^2 a_i \alpha_F^i(z) z^{-i}, \quad B_F(z) = \sum_{i=1}^6 b_i \beta_F^i z^{-i}$$

$$0 \leq \alpha_F \leq 1 \qquad 0 \leq \beta_F \leq 1$$

נבדקו ערכים שונים של צמידים α_F, β_F קבועים בזמן בתחום $[0,1]$. מתוך החשוואה התקבל כי השיפור המתקבל באיכות האות המשוחזר על-ידי מערכת ה-ADPCM-PF ביחס למערכת קידוד ADPCM הוא משמעותי, ועולה במידה ניכרת על זה המתקבל על-ידי מערכת ה-NFC בכל ערכי α_F, β_F . שנוס, זאת למרות העיוות הנגרם לאות המקודד במערכת זו על-ידי מסגן ה-PF.

אנו מניחים כי הפער בין ביצועי המערכות השונות נעוץ בהבדלים הבאים שביניהם:

- (1) **בתחום התדר** - בהנחת שגיאת קוונטיזציה לבנה מתקבלת במערכת קידוד ADPCM שגיאת שחזור לבנה, ומכאן שצורת ה- $SNR(f)$ היא כצורת הספקטרום של האות המקודד, במערכת קידוד NFC מעצבים את צורת ה- $SNR(f)$ לצורה שטוחה יותר. המוטיבציה לכך היא המודל של אוזן האדם על-ידי מערך מסננים בהם עובר האות הנקלט טיפול נפרד בכל פס תדר. לעומת זאת, במערכת קידוד ADPCM-PF מגבירים את האות בתדרים בהם ה- $SNR(f)$ גבוה, על חשבון הנחתת תדרים בהם ה- $SNR(f)$ נמוך. בשיטה זו מנצלים את החשפעות ההדדיות שבין פסי התדר השונים באוזן האדם, לצורך מיסוך אותם התדרים בהם מתקבל $SNR(f)$ נמוך.
- (2) **בתחום הזמן** - המיקום הפיזי של מסגן העיצוב במערכת ה-ADPCM-PF הוא ביציאת המפענת. פעולת הסינון לצורך עיצוב הספקטרום מבוצע על-ידי "החלקת" האות המשוחזר.

לעומת זאת, במערכת ה-NFC מסנן העיצוב ממוקם בצד המקודד והקוונטייזר נמצא בתוך חוג הסינון שלו. באופן זה צביעת שגיאת הקוונטיזציה מתקבלת על-ידי תוספת מעברים בין רמות הייצוג של הקוונטייזר.

פרק 6 – אלגוריתם הפרדה בין אותות דיבור ונתונים

עד כה עסקנו במקודדי אותות תוך הדגשה על התאמתם לאותות דיבור. בכלל זה, סקרנו בפרק הקודם שתי דרכים לשיפור סוביקטיבי של האות המקודד ע"י צביעת שגיאת הקוונטיזציה. מקודדים אלו חושוו, במשך ביצוע העבודה, עם מקודד ADPCM רגיל וכן עם מקודד PCM עם קוונטיזר מתואם בקידוד אותות נתונים. בסימולציות שבוצעו, התקבל כי במקודדים בהם נצבעת שגיאת הקוונטיזציה מתקבל BER גבוה באופן משמעותי ביחס לשאר המקודדים. כמו כן התברר יתרונו של מקודד ה-PCM על מקודד ה-ADPCM בעמידותו לשגיאות ערוץ.

תוצאות אלו הביאו אותנו לרעיון של שימוש במקודד אשר ימותג בהתאם לסוג האות בכניסה (דיבור/נתונים) בין שני המצבים הבאים: מקודד ADPCM-PF עבור אותות דיבור / מקודד APCM (Adaptive PCM) עבור אותות נתונים. מיתוג זה בין שני המצבים יבוצע באופן אוטומטי, על-סמך אלגוריתם סיווג האות לאחת משתי קבוצות אלו.

עקב המחקר הרב המבוצע בתחום זיהוי אותות דיבור וכן כתוצאה מפיתוח אלגוריתמים לדחיסת אותות דיבור על-ידי ניצול איזורי השקט, פותחו אלגוריתמים לאיבחון אותות דיבור מתוך רעש הרקע. בשל הדמיון לבעיית הסיווג שלנו, נציג דוגמא אחת הניתנת למימוש בזמן אמת.

אבחנה טובה של האות מתוך רעש הרקע מתקבלת בעזרת מספר פרמטרים המשלימים זה את זה כמדד לעוצמת האות, מידת סטציונריות האות לזמן קצר וכן קצב חציות האפס כמפורט לחלן [Cameron83]:

- (i) עוצמה - עוצמת האות מושווית עם סף עוצמה נתון. קיום אות מאובחן כאשר עוצמתו גבוהה מרמת הסף. סף זה נועד לשם איבחון של אותות דיבור קוליים. מתוך סימולציות לרעש המתקבל בקצה של ערוץ טלפון באורך מכסימלי, נקבע סף זה בגובה -30 dBmO.
- (ii) סטציונריות - סטציונריות האות נבחנת על-פי השוואת עוצמת חבלוק הנוכחי עם ממוצע עוצמת חבלוקים האחרונים. עבור אותות דיבור מתקבל ערך מוחלט של ההפרש בין עוצמות אלו הגדול מהמתקבל על-ידי רעש הרקע ברוב המקרים.
- (iii) חציות אפס -אנרגיה רבה של אותות אי-קוליים (בהם נכללים גם לחישות ועיצורים) מרוכזות בתדירות גבוהות יחסית. עובדה זו משתקפת בקצב חציות אפס גבוה מזה שיש לרעש הרקע (הקשר שבין צורת ספקטרום האות לקצב חציות האפס יוסבר בהמשך). בנוכחות רעש רקע בעוצמה גבוהה, האבחנה מבוצעת על-סמך מדידת אורך סדרות זילופי הסימן בדגמים עוקבים.

סיווג בין אותות דיבור ונתונים מוצע ב- [Adoul77] על סמך פרמטרים דומים לאלו שהוזכרו כעוצמת האות, קצב חציות האפס ונקודות האקסטרמום, מדד לסטציונריות האות וכן היחס שבין מכסימום העוצמה לערכה הממוצע כמדד לצורת האות.

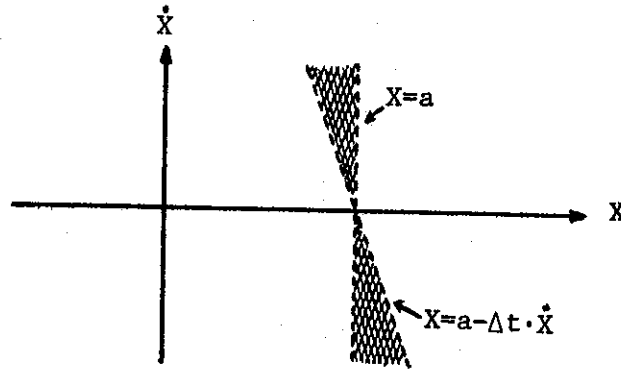
בעבודתנו, אנו מוותרים על מדידת קצב נקודות האקסטרמום ומשלבים יחדיו את שתי המדידות האחרונות המוצעות ב-[Adoul77].

6.1 חציות אפס ונקודות אקסטרמום

קצב חציות אפס ליחידת זמן של תהליך אקראי רציף $\{X_t\}$, ארגודי, סטציונרי ומוגבל סרט ניתן לחישוב מתוך פונקציית הפילוג המשותפת לתהליך ולניגזרתו $f_{x\dot{x}}(x, \dot{x})$. מאחר ו- X מוגבל סרט, ניתן ליצגו בקטע הזמן $(t, t+\Delta t)$, עבור Δt מספיק קטן, על-ידי קטע ישר המתחיל בנקודה X_t בשיפוע \dot{X}_t . לפיכך חציית הרמה $X=a$ יכולה להתרחש פעם אחת בלבד, אם בכלל, בקטע זמן זה. חתנאי לחציית הרמה הוא [Crandall63]

$$(6.1) \quad [X_t < a \wedge \dot{X}_t > \frac{a-X_t}{\Delta t}] \vee [X_t > a \wedge \dot{X}_t < \frac{a-X_t}{\Delta t}]$$

כאשר הסמנים \wedge ו- \vee מייצגים את הפעולות הלוגיות "and" ו-"or" בהתאמה. תנאי זה לחציית הרמה $X=a$ מתואר גם באיור 6.1.



איור 6.1 - הצמד x, \dot{x} ברגע t עבורו מתקבלת חציית הרמה $X=a$ בקטע הזמן $(t, t+\Delta t)$ נמצא באיזור המוצל [Crandall63].

Figure 6.1 - Combination of x and \dot{x} at time t may result in crossing of $X=a$ during the time interval $(t, t+\Delta t)$ lie in the shaded area [Crandall63].

הסתברות לחציית רמה זו בקטע הזמן $(t, t+\Delta t)$ היא ההסתברות לחמצא באיזור המוצל, המתואר באיור 6.1, השווה ל-

$$(6.2) \int_0^{\infty} dx \int_{a-x\Delta t}^a dx f_{XX}(x, \dot{x}) + \int_{-\infty}^0 dx \int_a^{a-x\Delta t} dx f_{XX}(x, \dot{x})$$

בגבול $\Delta t \rightarrow 0$ ועבור הרמה $a=0$ מקבלים את המוצע קצב חציות האפס למידת זמן

$$(6.3) E[Z_0] = \int_{-\infty}^{\infty} |\dot{x}| f_{XX}(0, \dot{x}) d\dot{x}$$

כאשר Z_0 הוא מונה מספר נקודות האקסטרומום ליחידת זמן. קל להראות כי עבור תהליך אקראי גאוסי סטציונרי, התהליך ונגזרתו בלתי תלויים סטטיסטית -

$$E[XX'] = E\left[\frac{1}{2} \frac{d}{dt} X^2\right] = \frac{1}{2} \frac{d}{dt} E[X^2] = 0$$

$$f_{XX'}(x, \dot{x}) = f_X(x) \cdot f_{\dot{X}}(\dot{x})$$

ומכאן

עם הצבת $f_{XX'}(x, \dot{x})$ ב-(6.3) נקבל [Rice45], [Kak69]

$$(6.4a) E[Z_0] = \frac{1}{\pi} \sqrt{-\frac{R''(0)}{R(0)}} = 2 \frac{\sqrt{\int_{-\infty}^{\infty} f^2 S(f) df}}{\sqrt{\int_{-\infty}^{\infty} S(f) df}}$$

כאשר $R(\tau)$ ו- $S(f)$ הן פונקציות האוטוקורלציה וספקטרום ההספק של התהליך בהתאמה, ו- $R_{(\tau)}^{(n)}$ היא הנגזרת ה- n -ית של $R(\tau)$. באופן דומה מחושב קצב חציות האפס של נגזרת התהליך (הוא קצב נקודות האקסטרומום), מתוך פונקצית הפילוג המשותפת של הנגזרת הראשונה והשניה של התהליך [Kak69]

$$(6.4b) E[Z_1] = \frac{1}{\pi} \sqrt{-\frac{R^{(4)}(0)}{R(0)}} = 2 \frac{\sqrt{\int_{-\infty}^{\infty} f^4 S(f) df}}{\sqrt{\int_{-\infty}^{\infty} f^2 S(f) df}}$$

כאשר Z_1 הוא מונה מספר נקודות האקסטרומום ליחידת זמן.

עבור המקרה הגאוסי הבדיד קצב חציות האפס הממוצע לקטע דגימה מתקבל באופן מיידי מתוך פונקצית האוטוקורלציה או ספקטרום ההספק של התהליך [Kedem86].

$$(6.5a) \quad \frac{E[Z_0]}{L-1} = \frac{1}{\pi} \cos^{-1} \rho_1 = \frac{1}{\pi} \cos^{-1} \left\{ \frac{\int_0^{\pi} \cos \theta S(\theta) d\theta}{\int_0^{\pi} S(\theta) d\theta} \right\}$$

כאשר ρ_1 הוא מקדם הקורלציה המנורמל בין שני דגמים סמוכים, $-\pi \leq \theta \leq \pi$ הוא ספקטרום החספק ו- Z_0 הוא מונה מספר הציות האפס בבלוק באורך L דגמים. קצב הופעת נקודות האקסטרומום הוא קצב הציות האפס של תהליך ההפרש $X(t) - X(t-1)$ (הנגזרת הדיסקרטית). קצב זה מחושב כמו ב- (6.5a) תוך החלפת ספקטרום החספק של התהליך בספקטרום החספק של נגזרתו.

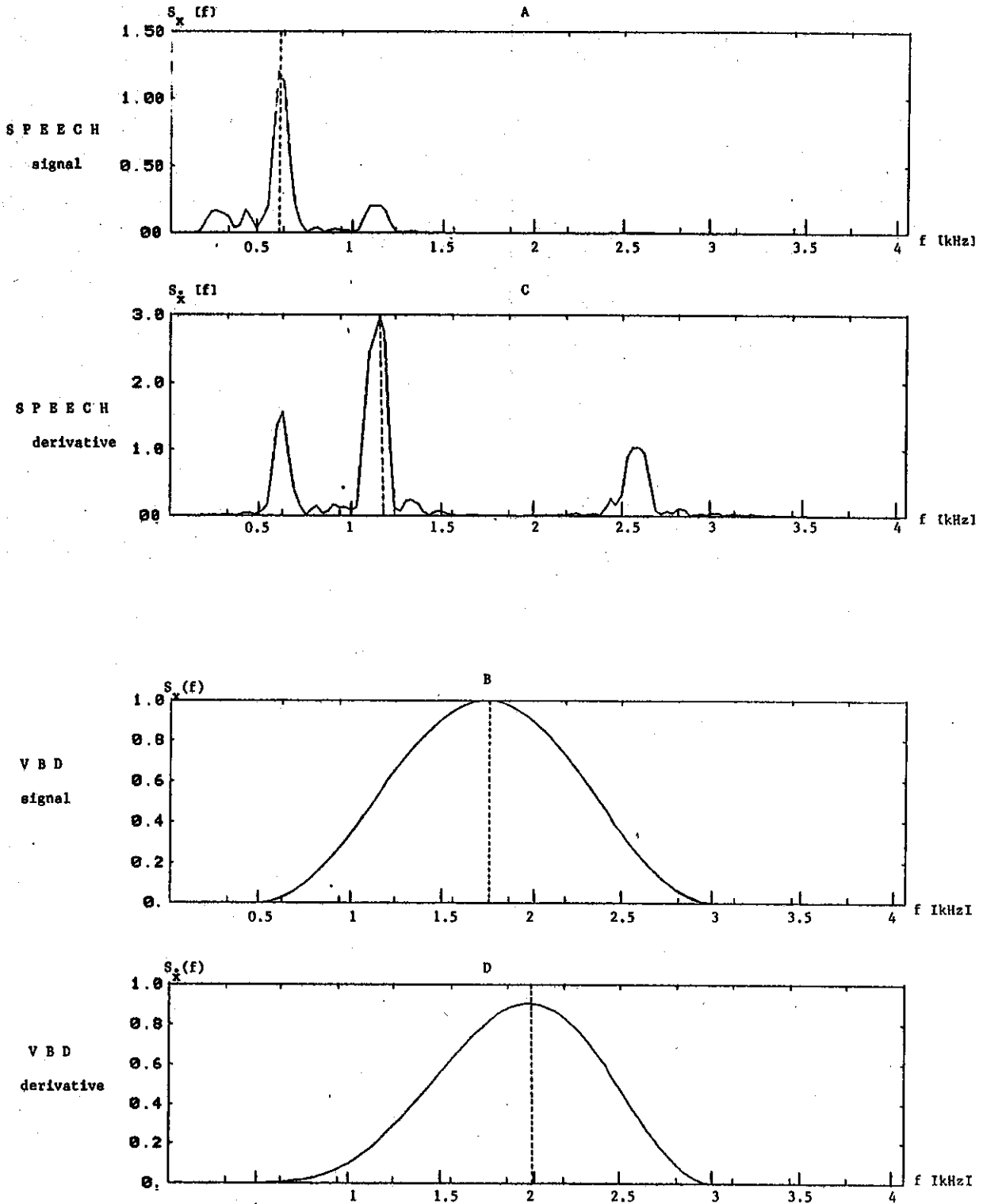
$$(6.5b) \quad \frac{E[Z_1]}{L-1} = \frac{1}{\pi} \cos^{-1} \left\{ \frac{\int_0^{\pi} \cos \theta |H(\theta)|^2 S(\theta) d\theta}{\int_0^{\pi} |H(\theta)|^2 S(\theta) d\theta} \right\}$$

כאשר Z_1 הוא מונה מספר נקודות האקסטרומום בבלוק באורך L דגמים, ו- $H(\theta)$ הוא תגובת התדר של מסנן גוזר ספרהלי.

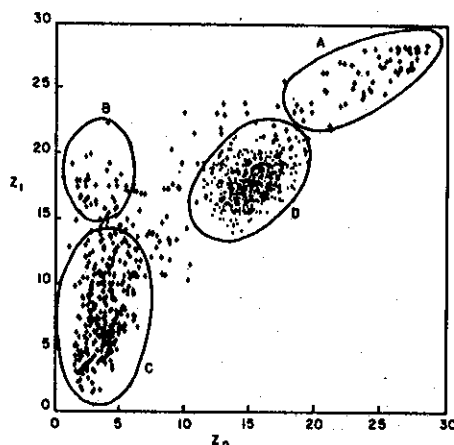
ניתן לראות מתוך (6.5) כי $E[Z_0]$ הוא ממוצע משוקלל של $S(\theta)$ במלים אחרות: "קצב הציות האפס של התהליך נוטה להתיישב בקרבת התדירות הדומיננטית של התהליך" [Kedem86]. באופן דומה, קצב האקסטרומום נוטה להתיישב בקרבת התדירות הדומיננטית של נגזרת התהליך.

6.1.1 קצב הציות האפס והאקסטרומום של אותות דיבור ונתונים

ספקטרום החספק של אותות הנתונים מרוכזים סביב התדירויות 1700-1800 Hz. זהו תחום תדירויות הגבוה באופן משמעותי מתחום התדירויות בו נמצא הפורמנט הראשון של אותות דיבור קוליים. הבדל נוסף הקיים בין האותות הנדונים הוא ביחס שבין ספקטרום האות לספקטרום אות ההפרש שלהם. בעוד שבאותות נתונים אין שינוי קיצוני בצורת ספקטרום האות, הרי שבאותות דיבור קוליים עובר מרכז הכובד מהפורמנט הראשון לפורמנט השני [Christian86], [Eqing69]. תופעות אלו משתקפות כפי שהוסבר קודם לכן, גם בקצב הציות האפס ונקודות האקסטרומום כמתואר באיורים 6.2, 6.3.



איור 6.2 – ספקטרום החספק של אות דיבור קולי, אות נתונים ואותות ההפרש שלהם מתוארים באיורים A, B, C, D בהתאמה. הקו המרוסק מציין את התדר הדומננטי של התהליכים או נגזרותיהם בהתאמה. Figure 6.2 – Power spectrum of speech and VBD signals and their differentiated signals are drawn in the sub figures A, B, C, D respectively. The dashed lines indicate the dominant frequency of the processes or their derivatives, respectively.



איור 6.3 - דיאגרמת פיזור פרמטרים Z_0, Z_1 של אותות דיבור ונתונים בקצב 9600 bps עבור מסגרת זמן באורך $L=32$ דגמים. איזור A אופייני לאותות אי-קוליים, איזורים B ו-C אופייניים לאותות קוליים. איזור D מאוכלס ע"י אותות הנתונים" [Christian86]

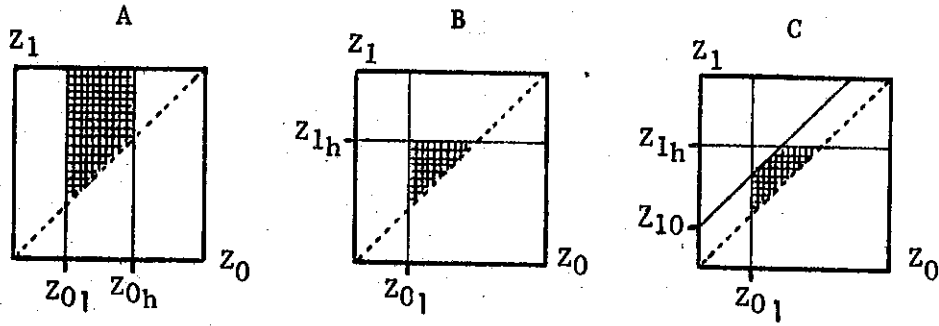
Figure 6.3 - Scatter diagram of the parameters Z_0, Z_1 for speech and VBD signals at a rate of 9600 bps with $L=32$ sample observation windows. Region A is characteristic of unvoiced sounds, regions B and C are characteristic of voiced sounds. Region D is occupied by the VBD signals [Christian86].

מתוך התוצאות המוצגות באיור 6.3, ניתן להבחין בהפרדה בין אותות דיבור ונתונים במישור $Z_0 \times Z_1$ המאפשרת סיווג אותות אלו על-סמך חלוקת מישור זה. יחד עם זאת לא קיימת חלוקה שעבורה לא תתקבלנה שגיאות סיווג כלל. זאת בשל פיזור ערכי Z_0, Z_1 של אותות דיבור, המתפרש גם על התחום המאוכלס על-ידי אותות הנתונים. הקטנת מספר שגיאות הסיווג מבוצעת בשני מישורים:

1. קביעת חלוקת המישור $Z_0 \times Z_1$ כך שהתחום המגדיר אותות נתונים יחרוג מעל ככל הניתן מעבר לתחום הערכים המתקבל על-ידי אותות אלו. באיור 6.4 מתוארות שלוש חלוקות לינאריות של המישור $Z_0 \times Z_1$ אשר נבחנו לצורך הסיווג. הערכים המגדירים את התחומים השונים מתאימים לחישוב Z_0 ו- Z_1 על-פי הביטוי (6.6) ומצורפים לאיור.
2. המישור השני, בו אנו פועלים לצורך הקטנת מספר שגיאות הסיווג של אותות דיבור, הוא השלמת אלגוריתם הסיווג בפרמטרים נוספים הקשורים לעוצמת האות כפי שמפורט ב: 6.2

$$(6.6) \begin{cases} Z_0(t) = (1-\alpha_0)Z_0(t-1) + \frac{1}{2} |\text{sign}[Y(t)] - \text{sign}[Y(t-1)]| \\ Z_1(t) = (1-\alpha_1)Z_1(t-1) + \frac{1}{2} |\text{sign}[Y(t)-Y(t-1)] - \text{sign}[Y(t-1)-Y(t-2)]| \\ \alpha_0 = \alpha_1 = 2^{-6} - \text{effective window} = 128 \end{cases}$$

(1) אותות הנתונים בהם אנו עוסקים מאכלסים במישור $Z_0 \times Z_1$ איזור המכיל את איזור D המסומן באיור 6.3



$$Z_{01} = 18.5 \text{ [2312.5 zero crossing/sec]}$$

$$Z_{0h} = 39 \text{ [4875 zero crossing/sec]}$$

$$Z_{1h} = 39 \text{ [4875 extremum/sec]}$$

$$Z_{10} = 10 \text{ [difference between } Z_1 \text{ and } Z_0]$$

איור 6.4 - שלוש חלוקות לינאריות של המישור $Z_0 \times Z_1$. אותות נתונים נמצאים באיזורים המוצלים.
 Figure 6.4 - Three linear splitting of the $Z_0 \times Z_1$ plane. VBD signals lie inside the shaded areas.

6.2 עוצמת האות

אלגוריתם סיווג המבוסס על-סמך חלוקת המישור $Z_0 \times Z_1$ אשר תואר עד כה, סובל לעתים מטעויות סיווג של אותות דיבור אי-קוליים. הקטנת מספר השגיאות מבוצעת על-ידי השלמת אלגוריתם חסיווג במדידת עוצמת האות, היחס שבין מכסימום העוצמה לערכה הממוצע וכן מדידה חלקית של סטציונריות האות. אנו מבצעים פעולות אלו על-ידי השוואת ממוצע הערך המוחלט של האות עם רמות סף מתאימות.

6.2.1 רמת סף קבועה - T_c

אותות דיבור אי-קוליים מאופיינים בדרך כלל גם בעוצמתם הנמוכה. ב- [Adoul77] מוצעת רמת סף לצורך ניפוי אותות אלו, הדומה בערכה לרמת הסף המוצעת גם ב- [Cameron83] לצורך הפרדה בין אותות דיבור קוליים לרעש הערוץ בגובה $T_c = -30 \text{ dBm0}$. אות שנוצמתו נמוכה מרמת סף זו (כולל רעש ערוץ ומירב אותות הדיבור האי-קוליים) יסווג כאות דיבור. כשאר פרמטרי המקודד, עוצמת האות משוערכת מידי כל דגם. לצורך חסכון בפעולות חישוב, מחושב ממוצע ערכו המוחלט של האות (אשר יסומן לצורך פשטות ב-AVE), במקום ערכו האפקטיבי (RMS), משום כך, יש צורך בגורם תיקון לרמת הסף T_c , השווה עבור פילוג כניסה גאוסי ל- $\sqrt{\pi/2}$.

$$(6.7) \quad AVE(t) = (1 - \alpha_{AVE}) AVE(t-1) + \alpha_{AVE} |Y(t)|, \quad \alpha_{AVE} = 2^{-5}$$

$$(6.8) \quad T_c \text{ (dBm)} = -30 - 10 \log \pi/2 = -32 \text{ dBm0}$$

6.2.2 רמת סף אדפטיבית - T_A

צורת האות: היחס שבין מכסימם העוצמה לערכה הממוצע מנוצל ב-[Adoul77] כמדד לצורת האות. עבור גל סינוסידלי יחס זה שווה ל- $\pi/2$. ע"פ קריטריון מדויק זה, אותות נתונים, המורכבים מגלים סינוסידליים, "קרובים" יותר לגלים סינוסידליים מאשר אותות דיבור.

סטציונריות: התחלה וסיום קטע דיבור (Utterance) מאופינים בין השאר בשינוי מהיר של עוצמת האות. זאת בניגוד לעוצמה קבועה יחסית של אותות נתונים.

אנו מתייחסים לשני מאפיינים אלה בעזרת רמת סף אדפטיבית - T_A הנגזרת מתוך משעריך מכסימם האות - MX כמפורט להלן:

$$(6.9) \quad T_A(t) = MX(t) / C, \quad C = 3$$

ומשעריך המכסימום מחושב באופן הבא:

$$(6.10) \quad MX(t) = \begin{cases} MX(t-1) & , \quad AVE(t) < T_A \\ (1-\alpha_{MD})MX(t-1) + \alpha_{MR} [\max\{MX(t-1), Y(t)\} - M(t-1)] & , \quad \text{else} \\ \alpha_{MD} = 2^{-11} & , \quad \alpha_{MR} = 1/4 \end{cases}$$

אות עבורו ממוצע הערך המוחלט - AVE קטן מרמת הסף T_A מסווג כאות דיבור.

ל- MX קבוע זמן דעיכה איטי, שנקבע לצורך זיהוי סוף קטע דיבור. כמו-כן MX מוקפא לערכו האחרון כאשר הערוץ "נכנס" למצב idle (רמת עוצמה הקטנה מ- T_c). באופן זה מתקבל T_A גבוה יחסית גם בתחילת קטע הדיבור הבא, ומאפשר סיווג נכון.

6.3 סיכום אלגוריתם הסיווג

באיור 6.5 מתוארות דיאגרמות פיזור ערכי הפרמטרים השונים שהוזכרו עבור אותות דיבור ביחס לאזורי ההחלטה - דיבור ונתונים, ובאיור 6.6 מוצגות החלטות הסיווג בציד הזמן. מתוצאות אלה עולה, כי קצב חציות האפס הוא המדד החשוב ביותר. רוב שגיאות הסיווג המתקבלות ע"פ מדד זה מאופיינות בעוצמתן הנמוכה, או שהן מתרחשות בתחילת/סיום Utterance. קטעים אלה של אות הדיבור מסווגים נכונה בעזרת רמות הסף (הקבועה והאדפטיביות) המשלימות את האלגוריתם.

נסמן ב-ID את ההחלטות מסווג האות, אזי -

$$(6.11) \quad ID(t) = \begin{cases} 0 \text{ (VBD)}, [Z_{o_1} \leq Z_o(t) \leq Z_{o_n}] \wedge [AVE(t) \geq \max\{T_A(t), T_C\}] \\ 1 \text{ (Speech)}, \text{ else} \end{cases}$$

בטבלה 6.2 מפורטות תוצאות הסיווג עבור אותות דיבור, וכן עבור אותות נתונים וטונים בתדרים 1.2 KHz ו- 2.4 KHz.

נציין, כי שגיאות הסיווג שהתקבלו, לא השפיעו באופן מהותי על פעולת המקודד, בזכות המעבר ה"רד" בין אופני העבודה השונים כפי שיוסבר בפרק הבא. משום כך גם הסתפקנו בבדיקת מונה חציות האפס, וייתרנו על מונה נקודות האקסטרמום.

Input Signal	σ [dBm0]	הסתברות שגיאות הסיווג			
		0, 0 [10 ⁻³]	40 dB, 0 [10 ⁻³]	20 dB, 10 ⁻³ [10 ⁻³]	10 dB, 10 ⁻³ [10 ⁻³]
Speech	-10	1.89	1.89	20.18	-
Speech	-20	1.71	1.50	3.54	-
VBD	-15	0	0	0	0
1.2 KHz	-15	88	86.5	79.7	71.4
2.4 KHz	-15	41.5	40.1	40.0	55.34

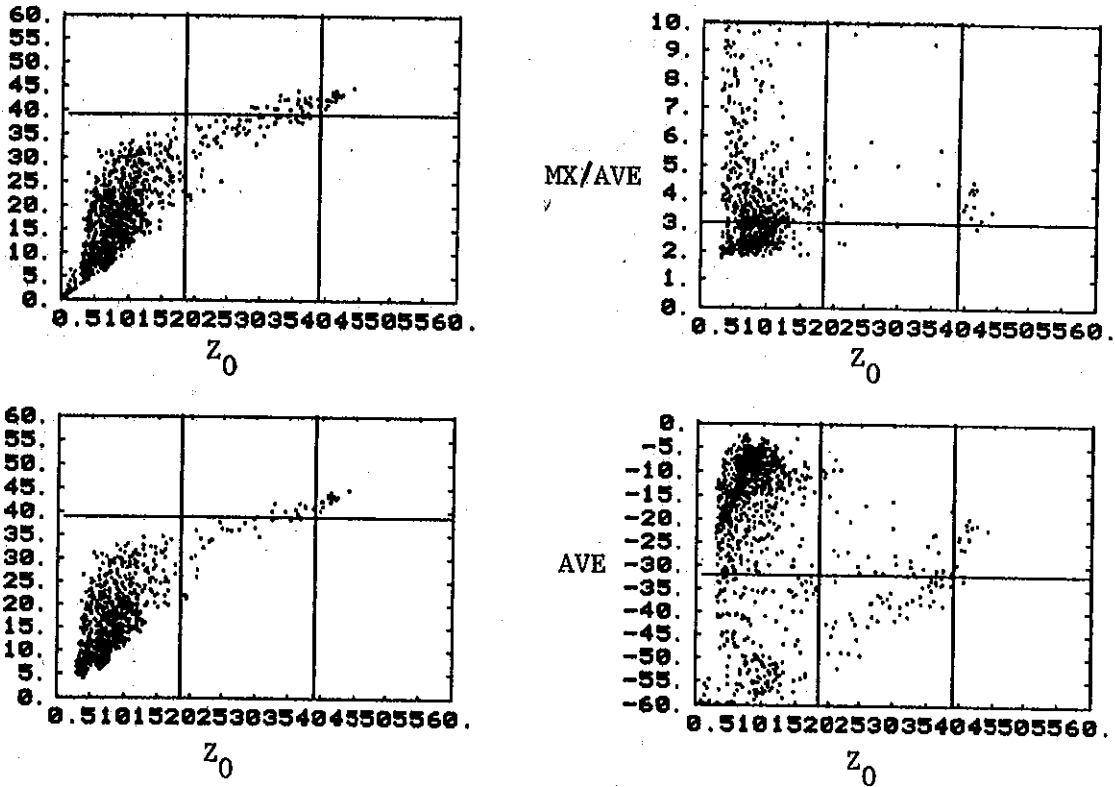
טבלה 6.2 - פירוט שגיאות הסיווג עבור אותות הכניסה השונים, בערוץ שקט ורועש. זוג הספרות N, ρ בראש העמודות 3+6 מתארות את תנאי הערוץ: לאות נוסף רעש גאוסי לבן ב-SNR השווה ל-N dB ושגיאות ערוץ בהסתברות - ρ .

Table 6.2 - Detailed identification for the various input signals, in a clean and noisy channel. The couple of numbers N, ρ at the head of columns 2:4 describes the channel conditions: A white gaussian noise is added to the signal on SNR equal to N dB. The channel has errors with probability - ρ .

במקרים בהם יש צורך במסווג מדוייק יותר ניתן לשפרו בדרכים הבאות:

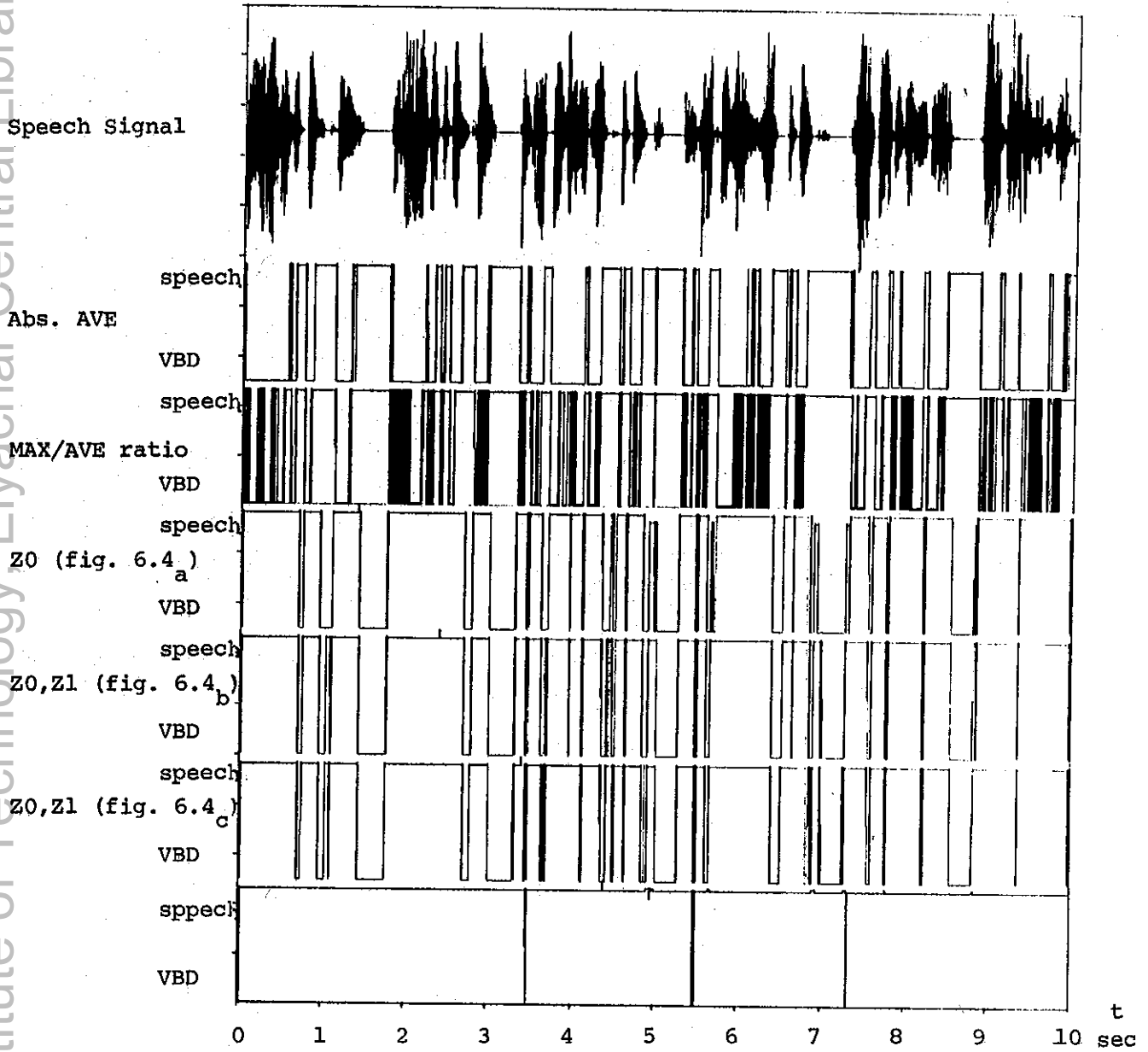
- (1) התיחסות למונה נקודות האקסטרמום.
- (2) הקפאת החלטת המסווג במשך קיום אות: לאחר איבחון קיום אות וקבלת החלטת סיווג יציבה למשך זמן מספיק ארוך (הקבוע מראש), תוקפא החלטה זו עד לרגע בו יאובחן סיום האות או עד לרגע בו תאובחן הופעת האות הבא.

3) ב- [Christian86] מוצע אלגוריתם מדויק לסיווג בין אותות דיבור לאותות Modem בקצב נתונים של 9600 bps, המסתמך על מדידת קצב חציות האפס ונקודות האקסטרמום של האות ושל האות במוצא מסננים שונים.



איור 6.5 - פיזור פרמטרי אות דיבור במישורי ההחלטה הבאים: A - $Z_1 \times Z_0$; B - $Z_1 \times Z_0$ כאשר מוצגים רק ערכים המתאימים לאות שעוצמתו גבוהה מרמת חסף הקבועה T_c . C - $MX/AVE \times Z_0$; D - $AVE \times Z_0$. ההגבלות של איור B;

Figure 6.5 - Scatter diagram of speech signal in the following decision planes: A - $Z_1 \times Z_0$; B - $Z_1 \times Z_0$ while only the values of signal with power greater than the threshold level T_c are presented; C - $MX/AVE \times Z_0$ with the restrictions of figure B; D - $AVE \times Z_0$.



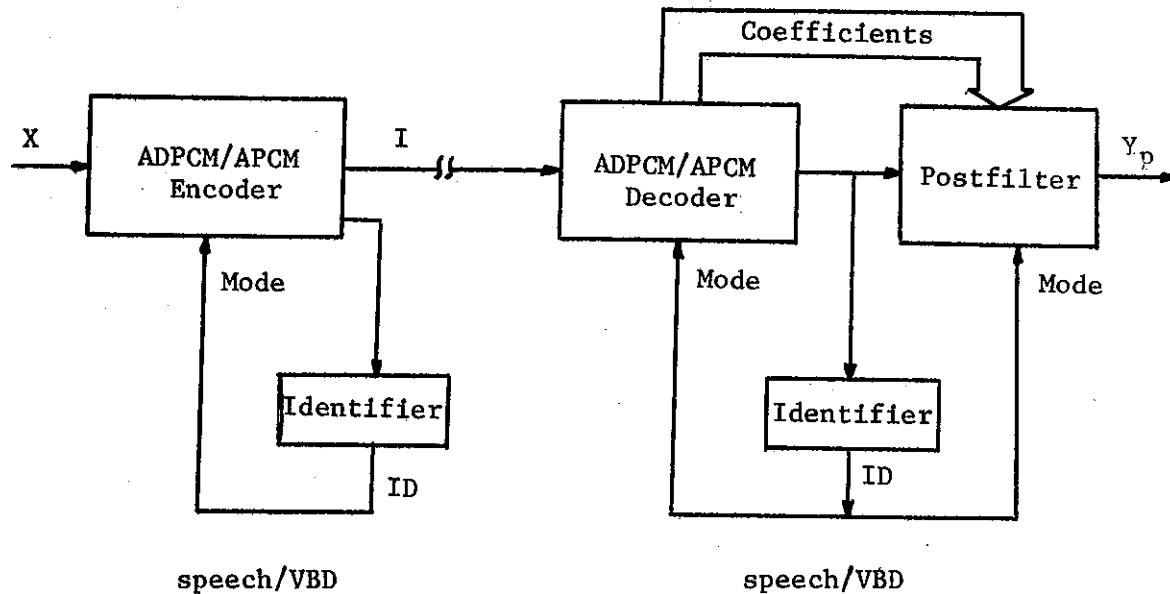
איור 6.6 - אות דיבור והחלטות הסיווג המתקבלות על-ידי הפרמטרים השונים שהתייחסנו אליהם.
Figure 6.6 - The accepted decisions of the identification's parameters considered for speech signal.

פרק 7 – מבנה מערכת הקידוד המוצעת ובצועיה

בעבודה זו מוצעת מערכת קידוד לקצב שידור 16 kbps עבור אותות דיבור ו-VBD (אותות נתונים המאופנים בשיטת PSK או FSK בקצבי נתונים של עד 2400 bps).

מערכת הקידוד ממותגת לאחד משני אופני העבודה הבאים:
עבור אותות דיבור – ADPCM with a Postfilter (ADPCM-PF)
עבור אותות VBD – APCM

אופן העבודה של מערכת הקידוד נקבע באופן אוטומטי בהתאם לסוג אות הכניסה בעזרת מנגנון סיווג האות לאחת משתי קטגוריות: דיבור ו-VBD (המתואר בפרק 6). אלגוריתם הסיווג מסתמך על האות המשוחזר (מסומן ב-Y), ועל-כן אין צורך בשידור החלטת המסווג למפענח כאינפורמציה צד. תרשים מלבנים של מערכת הקידוד המוצעת מתואר באיור 7.1.



איור 7.1 – סכמת מלבנים של מערכת הקידוד המוצעת.
Figure 7.1 – Block diagram of the proposed coder.

אנו מציגים בפרק זה את מבנה מערכת הקידוד המוצעת. תחילה נתאר את מסנן החיזוי מסוג zero-pole, נמשיך בתיאור מסנן ה-PF האדפטיבי וכן נתאר את מבנה הקוונטיזר.

7.1 מבנה החזאי

החזאי מורכב ממסנן קטבים מסדר שני ומסנן אפסים מסדר שני. אדפטציה המקדמים מבוצעת על-פי קירוב ה- $\text{sign}(\cdot)$ לאלגוריתם הגרדיאנט כפי שמבוצע ב-ADPCM32. נתאר בביטוי כללי אלגוריתם אדפטציה זה:

$$(7.1) \quad P_i(t+1) = (1 - \alpha_i)P_i(t) + \mu_i \Delta_i(t)$$

כאשר: P_i הוא המקדם ה- i -י של מסנני החזוי
 α_i קובע את זמן הדעיכה ה- i -י
 μ_i הוא מקדם התיקון ה- i -י
 Δ_i הוא גורם התיקון ה- i -י

במערכת הקידוד המוצעת אנו מעוניינים להשתמש באלגוריתם אדפטציה זה עבור אותות דיבור וחזאי מנוון, קרי $P_i=0 \forall i$ עבור אותות VBD.

נזכור כי יציאת המסווג ID נמצאת ברמה לוגית של '0'/'1' בעת זיהוי אותות דיבור/VBD בהתאמה. יציאה זו מאופיינת בהיותה לוגית במעבר חד ממצב אחד לשני. מעבר איטי יותר בין אופני העבודה של החזאי מתקבל על-ידי הכפלת גורם התיקון Δ_i בתוצאת המסווג:

$$(7.2) \quad P_i(t+1) = (1 - \alpha_i)P_i(t) + \mu_i \Delta_i(t) \cdot ID(t)$$

כאשר מזוהה אות דיבור $ID(t)=1$, לקבלת אלגוריתם אדפטציה זהה לזה הקיים ב-ADPCM32. כאשר המערכת מזוהה אותות VBD $ID(t)=0$. כתוצאה מכך מאופס גורם התיקון, ומקדמי המסננים דועכים לאפס. קצב המעבר בין אופני העבודה של מסנני החיזוי הוא כקצב האדפטציה של מקדמיהם בזמן עבודה נומינלית.

7.2 שיפור סובייקטיבי של איכות האות המשוחרר

בפרק 5 סקרנו שתי שיטות לשיפור סובייקטיבי של איכות האות המשוחרר. שיטת ה-NFC בה מעוצב ספקטרום האות ושיטת ה-Postfiltering בה מועבר האות המשוחרר דרך מסנן נוסף. מהשוואה סובייקטיבית של שתי השיטות התברר יתרון שיטת ה-Postfiltering. נעיר כאן כי מסנן ה-PF גרם להגבר האות. הגבר זה מקוזז על-ידי הכפלה בגורם נרמול קבוע שערכו 1/3 אשר יסומן - N_s (גודל קבוע זה נמצא כמתאים עבור רוב אותות הדיבור). [Jayant86]

בעת קידוד אותות VBD הביאה ה-PF להגדלה משמעותית של השגיאות אשר התקבלו במוצא המפענת. בעיה נוספת מחוזה גורם הנרמול הקבוע, אשר אינו מתאים לאותות אלה. מוצע לכן לבטל את מסנן ה-PF ולקבוע את גורם הנרמול ל-1 בנוכחות אותות VBD.

ביטול המסנן היא משימה קלה, שכן מקדמיו נגזרים ממקדמי מסנן השחזור, ואלה האחרונים מאולצים לאפס בנוכחות אותות VBD. מהירות המעבר של מסנן ה-PF מאופן עבודה אחד לשני (אקטיבי בנוכחות אותות דיבור ומנוון-שקוף - בנוכחות אותות VBD) היא כמהירות המעבר של מסנן השחזור. כתוצאה מניוון מסנן ה-PF בנוכחות אותות VBD, אין צורך עוד בגורם הנרמול. לכן נדרוש כי לגורם הנרמול של המסנן - N יהיו שני המצבים הבאים:

$$(7.3) \quad N(t) = \begin{cases} N_s & , \text{ Input Speech signal} \\ 1 & , \text{ Input VBD signal} \end{cases}$$

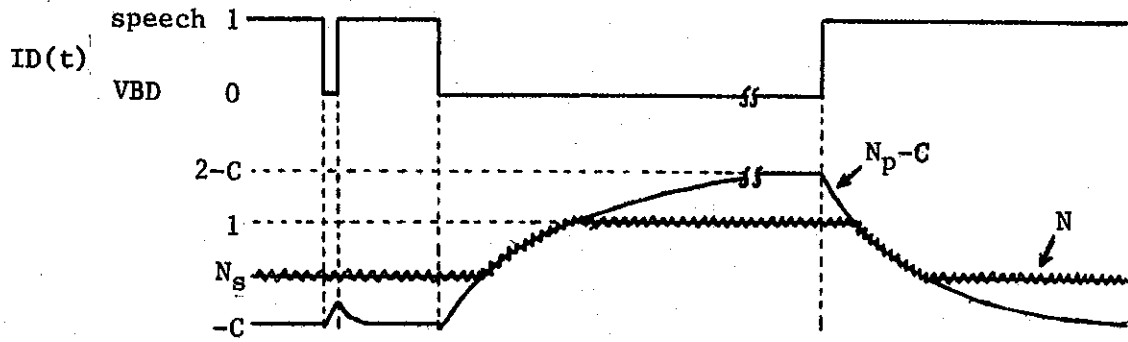
מעבר חד של גורם הנרמול בין מצביו השונים אינו רצוי עקב הקפיצה שתיגרם לעוצמת האות המשוחזר ועקב הקושי שיווצר בסינכרון המפענת למקדד. מעבר ממותן של גורם הנרמול מתקבל על-ידי העברת המשלים ל-1 של ID דרך LPF. מאחר ושנויים בגורם הנרמול משפיעים ישירות על עוצמת האות המשוחזר, דרשנו כי הוא יהיה חסין יותר משאר פרמטרי המערכת הממותגים לשגיאות ערוץ. חסינות זו מתקבלת (ללא הגדלת קבוע הזמן של ה-LPF) על-ידי השחיית תחילת המעבר של גורם הנרמול ממצב יציב אחד לשני כפי שיוסבר להלן. תחום ערכי יציאת ה-LPF יכול את תחום ערכי N(t) האפשריים, כאשר N(t) יתקבל על-ידי קצוץ יציאת ה-LPF לתחום הערכים הדרוש:

$$(7.4a) \quad N_p(t+1) = (1-\alpha_F)N_p(t) + 2\alpha_F[1 - ID(t)], \quad \alpha_F = 2^{-7}$$

$$(7.4b) \quad N(t) = \begin{cases} 1 & , N_p(t) - C \geq 1 \quad (\text{VBD}) \\ N_s & , N_p(t) - C \leq N_s \quad (\text{Speech}) \\ N_p(t) - C & , \text{ else} \quad (\text{Transition state}) \end{cases}$$

$$-N_s \leq C < 1$$

כאשר $N_p(t) - C$ היא יציאת ה-LPF ו-C הוא גודל קבוע המבקר על החשיות המתקבלות מרגע מתוג ID(t) לתחילת השינוי ב-N(t). עבור $C=1$ מתקבלת השחייה מירבית בין מיתוג ID(t) מ-1 ל-0 (דיבור ל-VBD) לתחילת השינוי ב-N(t). עבור $C=-N_s$ מתקבלת השחייה מירבית בין מיתוג ID(t) מ-0 ל-1 (VBD לדיבור) לתחילת השינוי ב-N(t). אנו קבענו את C לערך $(1-N_s)/2$ לקבלת זמני השחייה שווים בשני המעברים, ראה איור 7.2.



איור 7.2 - תגובת גורם הנרמול N - ליציאת המסווג.

Figure 7.2 - Normalization factor N response to the discriminator output.

7.3 הקוונטיזציה

מערכת הקידוד המוצעת כוללת קוונטיזציה אדפטיבית המסתמך על שערך אחורי של וריאנס האות בעזרת אלגוריתם חישוב רקורסיבי עמיד לשגיאות ערוץ [Goodman75] (ראה משוואה 2.8a). התאמת הקוונטיזציה לאותות דיבור (אי-סטציונריים) מחד, ולאותות VBD (שחם בעלי עוצמה קבועה יחסית) מאידך, מבוצעת על-סמך הצעת Petr [Petr82] בה מותאמת מהירות האדפטציה למידת סטציונריות האות. בהצעתו, Petr מממש זאת על-ידי בניית שני צעדי קוונטיזציה: האחד מוושב על-פי קבוע זמן מהיר (המתאים לאותות דיבור), והשני מתקבל על-ידי העברת צעד הקוונטיזציה הראשון דרך LPF (לקבלת צעד קוונטיזציה בעל קבוע זמן איטי המתאים לאותות VBD). צעד הקוונטיזציה הסופי מתקבל מתוך שקלולם הלינארי של השניים בהתאם למדד סטציונריות האות $\alpha(t)$ (ראה סעיף 4.2.1).

במקום מימוש זה אנו מתאימים את "קבוע" זמן האדפטציה למידת סטציונריות האות באופן ישיר, כפי שמוצג להלן:

$$(7.5a) \quad \nabla(t+1) = [1 - \beta(t)] \nabla(t) + \beta(t)W_i(t)$$

$$(7.5b) \quad \beta(t) = \alpha(t)\beta_u + [1 - \alpha(t)]\beta_l, \quad \beta_u = 2^{-5} \quad \beta_l = 2^{-11}$$

כאשר ∇ הוא לוגריתם צעד הקוונטיזציה, β קובע את קצב האדפטציה של הקוונטיזציה ו- W_i הוא לוגריתם של הכופל ה- M_i של הקוונטיזציה. מדד סטציונריות האות $\alpha(t)$ מחושב בדומה לאלגוריתם ה-ADPCM32. עבור אותות אי-סטציונריים $\beta(t) = \beta_u \leftarrow \alpha = 1$ לקבלת קצב אדפטציה מהיר, ועבור $\beta(t) = \beta_l \leftarrow \alpha(t) = 0$ לקבלת קצב אדפטציה איטי.

7.3.1 התחום הדינמי של הקוונטיזציה

התחום הדינמי של הקוונטיזציה נפרש על-ידי צעדי הקוונטיזציה המינימלי והמכסימלי. צעד הקוונטיזציה המינימלי נבחר כבאלגוריתם ה-ADPCM32 לערך $\Delta_{\min}=2$. צעד הקוונטיזציה המכסימלי נקבע לערך $\Delta_{\max} = 2700$. ערך זה מגביל את התחום הדינמי של הקוונטיזציה מלמעלה ל- -3dBm0 ⁽¹⁾. עבור אות כניסה גאוסי חסר זכרון (הגבלה זו זהה להגבלת התחום הדינמי המתקבלת על-ידי מקודד PCM μ -Law בקצב שידור 64kbps).

7.3.2 פרמטרי קוונטיזציה

בפרק 2 תראינו מספר דרכים לחישוב אוסף כופלים לפילוג וקוונטיזציה נתונים (ראה משוואה 2.19a וצמד המשוואות 2.19b ו-2.13b). נעיר, כי ממשוואות אלה מתקבל תחום ערכים מותר ולא דווקא אוסף כופלים יחיד. במקרה של קוונטיזציה בעל 4 רמות ($R = 2 \text{ bits/sample}$) מתקבל פתרון עם דרגת חופש אחת. דרגת החופש גדלה באחד כאשר מתחשבים בתיקון המוצע ב-[Jayant73] לכופל המתאים לרמת הקוונטיזציה הגבוה ביותר. כלומר, מספר דרגות חופש כמספר הנעלמים. מסיבה זו קבענו את אוסף הכופלים באופן נסיוני תוך חיפוש סדרתי.

באיורים 7.3 ו-7.4 מתואר ה-SEG-SNR המתקבל עבור אותות דיבור ו-VBD בהתאמה כפונקציה של הלוגריתם לפי בסיס e של כופלי הקוונטיזציה - $\{\beta_{\mu} W_i\}$ וקוטטזורים האופטימליים לפילוגים: Gaussian, Laplacian, Uniform, ו-Gamma. מתוך אוספי הכופלים שנבחנו הגבלנו עצמנו לאלה המאפשרים את קבלת התחום הדינמי המבוקש. תחום זה מתקבל עבור הקוונטיזציה המתואר במשוואה: $(7.5a) \quad W_0 \leq \varphi(t) \leq W_1$. מכאן הגבלנו עצמנו לערכי W_i הבאים:

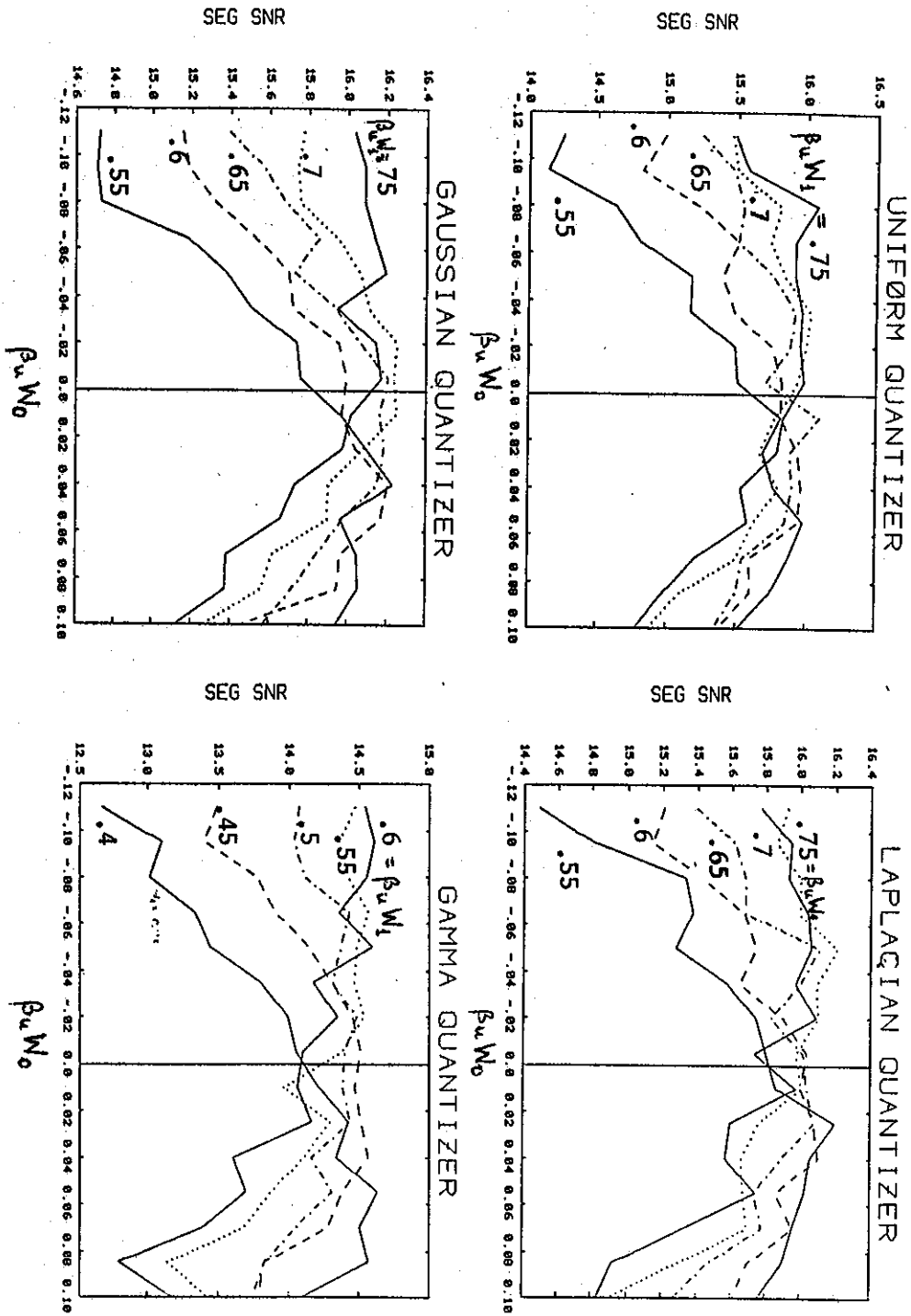
$$(7.6) \quad W_0 \leq \nabla_{\min} , \quad W_1 \geq \nabla_{\max}$$

את אוסף הכופלים בחרנו כפשרה בין האוספים המתאימים לאותות דיבור ונתונים. מתוך האפשרויות שהתקבלו בחרנו בזו הגורמת לקצב האדפטציה האיטי יותר, זאת כדי לקבל רגישות קטנה ככל האפשר לשגיאות ערוץ. משיקולים אלה בחרנו בקוונטיזציה גאוסי עם זוג הכופלים:

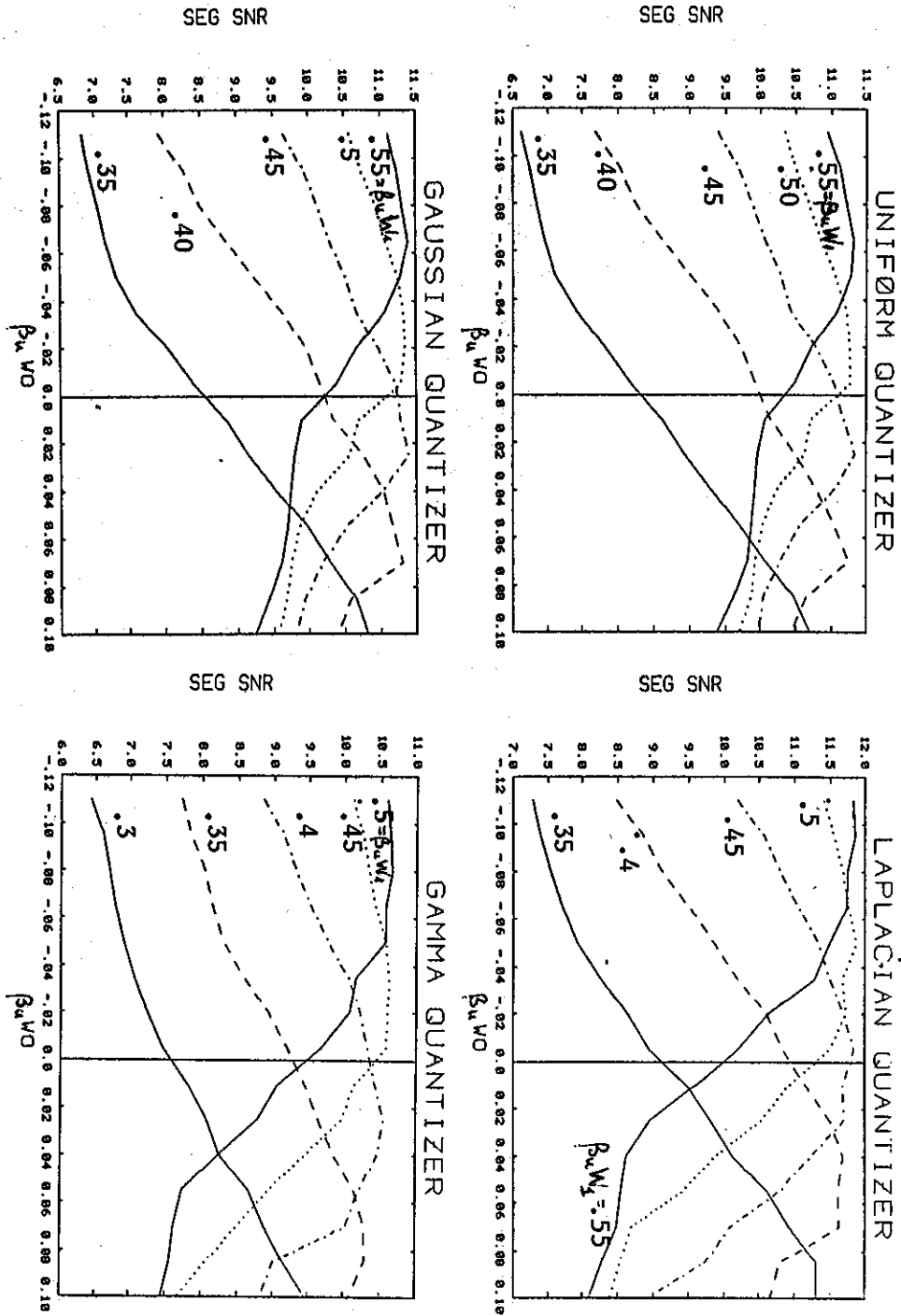
$$M_0 = 1.0202 , \quad M_1 = 1.6487$$

$$\beta_{\mu} W_0 = 0.02 \quad \beta_{\mu} W_1 = 0.50 \quad \text{ולחילופין}$$

(1) 0dBm0 מיוצג על-ידי הערך 4000.



איור 7.3 - SEG-SNR עבור אות דיבור כפונקציה של $\{\beta_u W_i\}$ וסוג הקוונטיזציה.
Figure 7.3 - SEG-SNR for speech signal versus $\{\beta_u W_i\}$ and type of the quantizer.



איור 7.4 - SEG-SNR עבור אותות VBD כפונקציה של $\{\beta_u W_i\}$ וסוג הקוונטיזציה.
 Figure 7.4 - SEG-SNR for VBD signal versus $\{\beta_u W_i\}$ and type of the quantizer.

7.4 ביצועי מערכת המוצעת

ביצועי מערכת הקידוד המוצעת נבחנו בעזרת סימולציות מחשב בהן קודדו אותות דיבור ו-VBD בתנאי ערוץ שקט, בתוספת רעש גאוסי לבן וכן בתנאי ערוץ ספרתי רועש. בגלל הקושי הכרוך בכיול המדידות כפי שיוסבר בהמשך, ביצועי המערכת המוצעת נבחנו תוך בצוע מדידות השוואתיות ביחס לאלגוריתמי קידוד שונים המפורטים להלן:

64 kbps A-Law PCM {A-PCM64}⁽²⁾

32-64 kbps μ -Law PCM { μ -PCM32-64}⁽²⁾

32 kbps ADPCM (CCITT standard) {ADPCM32}⁽²⁾

16 kbps ADPCM {ADPCM16}⁽²⁾, אלגוריתם זה דומה לאלגוריתם המוצע, אך בעל אופן עבודה אחד בלבד וללא מסנן ה-PF.

7.4.1 אותות דיבור

אותות הדיבור אשר קודדו לצורך בחינת איכות האות המשוחזר הם שלשה משפטים בשפה האנגלית הנאמרים על-ידי דובר ודוברת, מוקלטים בעזרת מקרופון בעל איכות גבוהה ומוגבלי סרט לתחום $0.2 \div 3.4$ KHz.

אותות דיבור אלה קודדו בתנאי הערוץ הבאים:

- (1) ערוץ נקי.
- (2) תוספת של רעש גאוסי לבן חסר קורלציה לאות המקודד ביחס אות לרעש $SNR = 20$ dB.
- (3) הכנסת שגיאות ערוץ בהסתברות $Channel BER = 10^{-3}$.

להלן מוגש סיכום בדיקות השמיעה הסוביקטיביות אשר נערכו:

- (1) ביצועי מערכת הקידוד המוצעת עלו תמיד (בכל תנאי הערוץ) על ביצועי אלגוריתם ה-ADPCM16.
- (2) מערכות קידוד ADPCM, בכללם האלגוריתם המוצע, רגישות פחות ממקודדי PCM לשגיאות ערוץ. עוצמת שגיאות השחזור קטנות יותר, והן מטרידות פחות את האוזן האנושית כמוסבר להלן:

א. עוצמת השגיאה - במערכות ADPCM האות המשודר הוא יצוג של אות הפרש, אות בעל וריאנס הקטן מוריאנס אות הכניסה. עוצמת השגיאה הנגרמת לאות זה וכן לאות המשוחזר קטנה מזו הנגרמת לאות המשוחזר במוצא מערכת PCM.

-
- (2) לשם נוחיות נשתמש בכינויים המופיעים בסוגריים { } לאלגוריתמי הקידוד אליהם מיוחסים ביצועי המערכת המוצעת.

ב. למערכות ADPCM זכרון הגורם להתפשטות השגיאה לאורך זמן. כתוצאה מכך מתקבלת שגיאת שחזור המרוכזת באיזור תדרים נמוך יותר המטריד פחות מבחינה סוביקטיבית. השמוש במסנן ה-PF גורם עוד לצביעת השגיאה על-פי צורת ספקטרום מוחלקת של אות הדיבור. עקב כך ההפרעה הסוביקטיבית קטנה במידה נוספת, עד כדי קושי באיבחון של השגיאות.

(3) המערכת המוצעת רגישה פחות מכל שאר המערכות לרעש המתווסף לאות בטרם קידודו. כי למסנן ה-PF ניתן להתייחס כאל מסננת מתואמת לאות הדיבור הגורמת לשפור ה-SNR במוצא המפענת. באיור 7.5 מתוארת השפעת תוספת הרעש בכניסה על ה-SNR במוצא המערכת המוצעת ובמוצא מערכת ה-ADPCM16. זאת ועוד, כפי שצויין קודם, ספקטרום השגיאה מעוצב על-פי ספקטרום אות הדיבור לקבלת רעש נלווה המטריד פחות (סוביקטיבית).

נזכיר שוב כי השיפור הסוביקטיבי מתקבל במערכת המוצעת על חשבון הנחתת תדרים גבוהים באופן משמעותי. לצורך הפרדת התרומה המתווספת לשיפור איכות האות המשוחזר מהנחתת הגבוהים, בוצעה השוואה סוביקטיבית בין האותות המתקבלים במוצא המערכת המוצעת וזה המתקבל במוצא שאר המערכות כאשר ביציאתם האות הועבר דרך LPF (השוואה זו בוצעה בעזרת 8 מאזינים לא מנוסים).

תאור מחלך המדידה:

קביעת תדר הברך של ה-LPF

לכל מאזין הושמעו האות המשוחזר במוצא המערכת המוצעת והאות המקורי המועבר דרך LPF. תדר הברך של ה-LPF הורד בהדרגה מתדר 3.4 KHz (היא תדירות הברך של מסנן ה-LPF אשר דרכו הועבר האות בטרם דגימתו) במטרה למצוא את התדירות בה האותות נשמעו "דומים". תחליך דומה לזה בוצע שנית אלא שבפעם זו העלינו בהדרגה את תדירות הברך מתדר 1 KHz ומעלה. בשתי הפעמים התקבלו תדירויות קרובות בערךן ולפיכך נקבע תדר בברך לכל מאזין כממוצע בין שתי התדירויות שנבחרו.

השיפור הסוביקטיבי הנתרם על-ידי ה-PF

בוצעה השוואה סוביקטיבית בין האות המתקבל במוצא המערכת המוצעת לבין האות המתקבל על-ידי ה-ADPCM16 כאשר ביציאתו מחובר ה-LPF אשר נבחר בסעיף קודם (על-ידי כל מאזין).

דירוג סוביקטיבי של איכות האות במוצא המערכת המוצעת

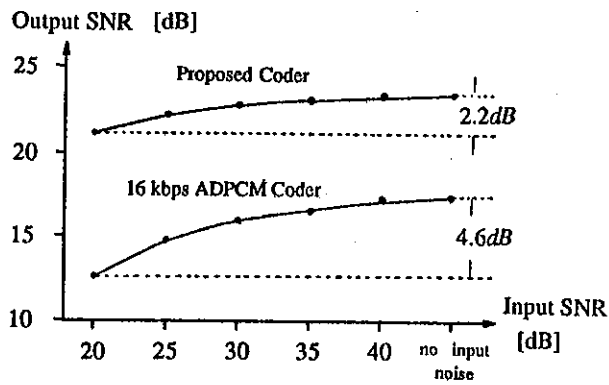
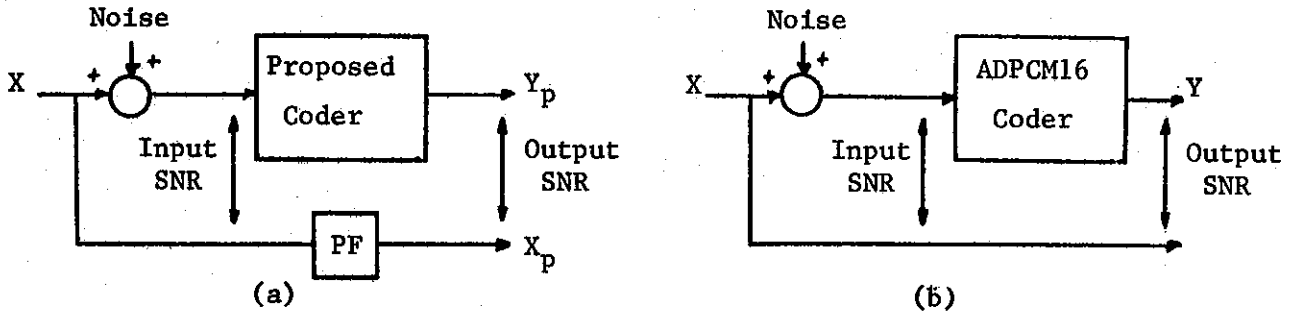
בוצעה השוואה סוביקטיבית בין האות המתקבל במוצא המערכת המוצעת לבין האות המתקבל במוצא מקודד PCM- μ , בקצבי שידור שונים, כאשר ביציאתו חובר ה-LPF הנדון.

סכום התוצאות:

(1) הנחתת הגבוהים המתקבלת על-ידי מסנן ה-PF "דומה" למעבר האות דרך LPF עם תדר ברך בתדר 1.7 KHz

(2) שיפור איכות אות הדיבור הנתרם על-ידי מסנן ה-PF גדול מזה המתקבל על-ידי LPF "שקול" וקבוע בזמן (קביעה זו התקבלה פה אחד על ידי כל המאזינים).

6) איכות אות הדיבור המשוחזר במוצא המערכת המוצעת זומה לזו המתקבלת על-ידי μ -Law 48 (3)
 אשר ביציאתו מחובר LPF "שקול".



(c)

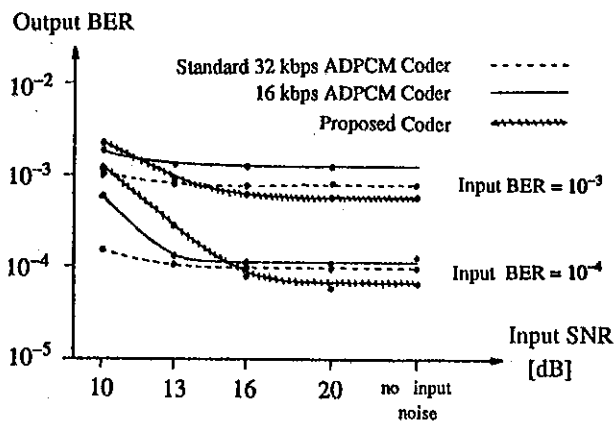
איור 7.5 - מערך בדיקת השפעת רעש כניסה על הרעש במוצא המערכת המוצעת (a) ובמוצא מערכת ADPCM16 (b). התוצאות מוצגות במונחים של SNR (c).

Figure 7.5 - Input noise projection on the output noise at the proposed system (a) and at the ADPCM16 system (b). The results are given in terms of SNR (c).

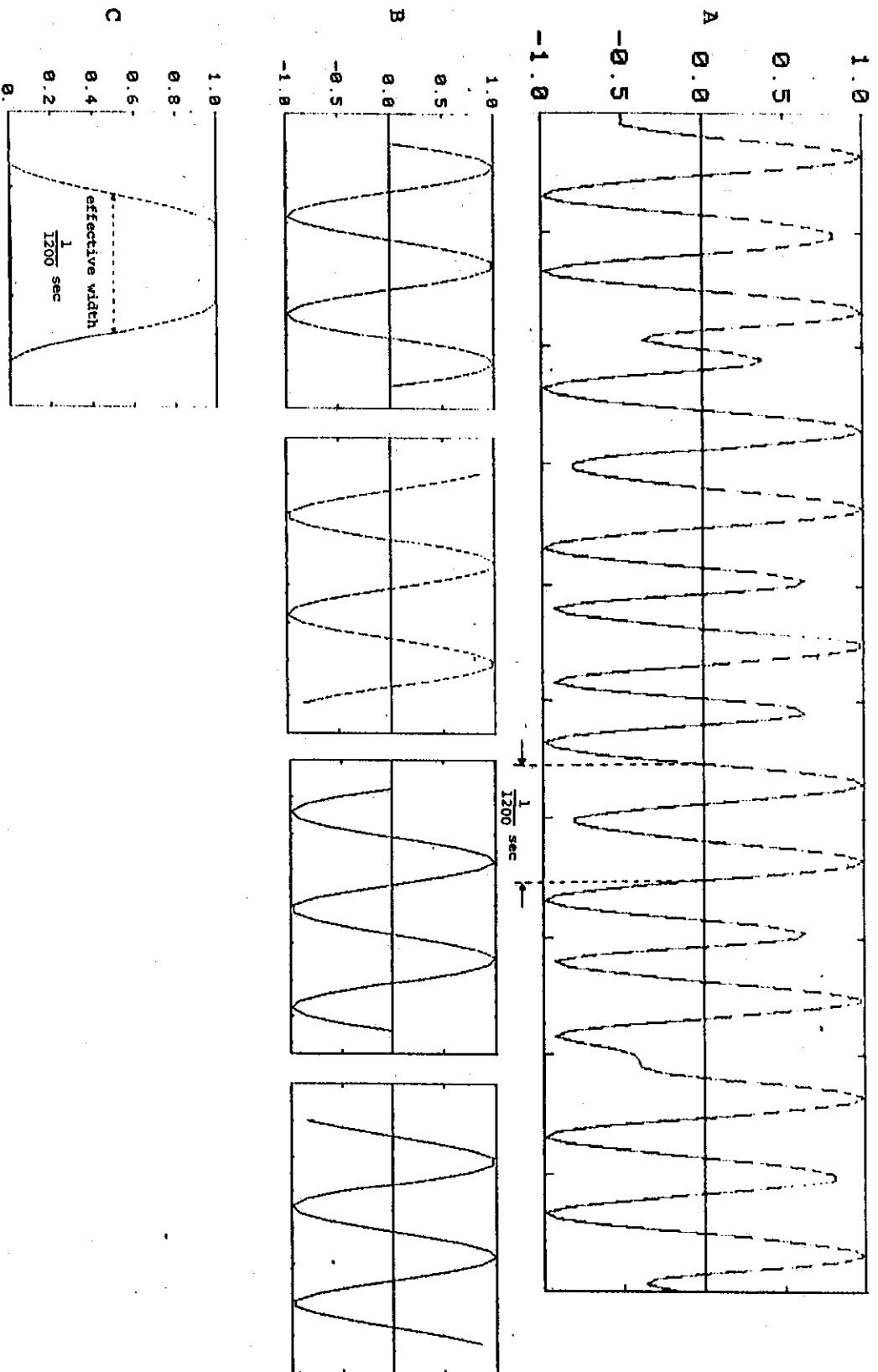
7.4.2 אותות VBD

בדיקת ביצועי המערכת המוצעת בנוכחות אותות VBD בוצעה על-ידי קידוד אותות אשר הופקו ממדמה אשר נכתב למטרה זו. המדמה מדמה אותות Modem באפנון QPSK בקצב אותות נתונים 2400 bps (זה הקצב הגבוה ביותר אליו המערכת מתאימה). לשם הערכת ביצועי המערכת היה צורך בביצוע מדידות השוואתיות ביחס למערכות קידוד אחרות, שכן אותות ה-VBD המדומים אינם זהים לאותות Modem אמיתיים. האות המדומה מורכב מקטעי גל סנדואיכלי במופעים שונים, הנקבעים בהתאם לאות הנתונים המאופנו, מוכפלים בחלון Hamming קטום ללא סינון נוסף המגדיל את רוחב הסרט של האות (ראה איור 7.6).

אות ה-VBD קודד בתנאי ערוץ ספרתי רועש בהסתברויות 10^{-4} ו- 10^{-3} , וכן בתוספת רעש גאוס לכן ב-SNR נמוך עד כדי 10 dB. סיכום תוצאות המדידות החשוואתיות בין מערכות הקידוד ADPCM16, ADPCM32 והמערכת המוצעת הפועלת כ-APCM מופיעות באיור 7.7. ממדידות אלה נובע כי ב-SNR גבוה ביצועי המקודד המוצע טובים יותר, וכי המערכת המוצעת רגישה יותר לתוספת רעש, אך רגישה פחות לשגיאות ערוץ משאר המערכות שנבדקו.



איור 7.7 - BER במוצא מערכות הקידוד השונות כפונקציה של ה-SNR בכניסה ושגיאות הערוץ.
 Figure 7.7 - Output BER versus input SNR and channel error.



אין 7.6 - אות ה-VBD המדומה (a) המתקבל מקטעי גל סינוסואידלי באתר מארביץ אפשרות (b)

המוכפלים בחלון המינג קוסים (c)

Figure 7.6 - Simulated VBD signal (a) composite of sine wave at one of 4 phases (b)

multiplied by a clipped Hamming window (c).

פרק 8 – סכום ומסקנות

בעבודה זו הצגנו מערכת קידוד לקצב שידור 16 kbps הממותגת בין שני אופני עבודה המתאימים לאותות דיבור ונתונים (VBD).

בעת קידוד אותות דיבור המערכת פועלת כמקודד ADPCM בתוספת Postfilter (PF). בעזרת מסנן ה-PF את הדיבור המשוחזר מוגבר בתדרים בהם ה-SNR גבוה ומונחת בתדרים בהם ה-SNR נמוך. באופן זה ממוסכים התדרים האחרונים לקבלת אות הנשמע סוביקטיבית נעים יותר לאוזן. השיפור הסוביקטיבי באיכות האות בא על חשבון הנחתת תדרים גבוהים (בהם התקבל ה-SNR הנמוך). הנחתה זו דומה (סוביקטיבית) לזו המתקבלת על-ידי מסנן LPF בעל תדר ברוך ב-1.7 KHz. בוצעה השוואת המערכת המוצגת עם מקודד PCM לוגריתמי (μ -Law) בקצבי שידור שונים, אשר ביציאתו הוצב מסנן ה-LPF המוזכר לעיל. מהשוואה זו התקבל, כי איכות האות המתקבלת במוצא המערכת המוצגת דומה לאיכות המתקבלת על-ידי מקודד PCM זה הפועל בקצב 48 kbps.

בעת קידוד אותות VBD המערכת פועלת כמקודד APCM (חוג החיזוי וכן מסנן ה-PF מנוונים). התוצאות שהתקבלו על-ידי המערכת המוצגת, בעת קידוד אותות VBD סינטטיים, בקצב נתונים של 32 kbps היו דומות לאלו המתקבלות על-ידי אלגוריתם ה-ADPCM הסטנדרטי לקצב שידור של 32 kbps. המערכת המוצגת התגלתה כרגישה יותר לרעש המתלווה לאות ה-VBD, אך רגישה פחות לשגיאות ערוץ. בתנאי ערוץ ספרתי רועש, בהסתברות שגיאות השווה ל- 10^{-4} , התקבל BER נמוך יותר על-ידי המקודד המוצג מאשר על-ידי אלגוריתם ה-ADPCM הסטנדרטי. מצב זה התהפך כאשר עוצמת הרעש היתה גבוהה, עד כדי SNR השווה לכ -15 dB.

המיתוג בין שני אופני העבודה שהוצגו מבוצע באופן אוטומטי בעזרת אלגוריתם המסווג את אות הכניסה לאחת משתי קטגוריות - דיבור/VBD. אלגוריתם הסיווג מסתמך על פרמטרי האות הבאים: קצב חציות האפס, ממוצע הערך המוחלט (עוצמה), היחס שבין מכסימום הערך המוחלט לעוצמת האות וכן על קצב ירידת עוצמת האות.

עבור אותות VBD התקבל סיווג נכון בכל תנאי הבדיקה שבוצעו. עבור אותות דיבור התקבלו החלטות סיווג לא נכונות בהסתברות של כ- 10^{-3} , אולם בזכות "המעבר חרך" של מערכת הקידוד מאופן אחד לשני, שגיאות אלה לא פגעו כלל באיכות האות המשוחזר.

8.1 כיוונים למחקר המשך

נפריד דיון זה לטיפול באותות VBD ובאותות דיבור:

אותות VBD - התוצאות המוצגות בעבודה זו לגבי אותות VBD מסתמכות על סימולציות שבוצעו על אותות סינטטיים, אשר יוצרו לצורך מטרה זו. אותות אלה מדמים אותות QPSK בקצב נתונים של 2400 bps. האותות הסינטטיים רחבי סרט מאותות VBD אמיתיים בקצב זה ובקצבי שידור נמוכים יותר. מערכת קידוד ADPCM מתפקדת ביעילות רבה יותר ככל שאות הכניסה צר סרט יותר. לפיכך אנו צופים שיפור ביצועי מקודד ה-ADPCM ביחס למערכת המוצגת בתנאי שדה. אנו ממליצים לבחון את ביצועי המערכת בקידוד אותות אלה כנגד מקודד ADPCM רגיל. אם יסתבר, כי ביצועי ה-ADPCM יהיו טובים יותר, אזי ניתן לנוון במערכת המוצגת את מסנן ה-PF בלבד, ולחשאי על כנו את מסנן החיזור.

אותות דיבור - השיפור המתקבל בשיטת ה-Postfiltering מלווה במספר תופעות לוואי לא

רצויות ואלו הן:

- (1) הנחתת תדרים גבוהים הפוגעת באיכות האות.
- (2) הגברת תדרים נמוכים המקשה על חיבור טורי של מספר דרגות קידוד, זאת משום חריגת האות מעבר לתחום הדינמי.
- (3) עקב האדפטיביות של מסנן ה-PF קטן רוחב הסרט של האות מדרגה לדרגה וכבר במוצא שתי דרגות בלבד מתקבל אות משוחזר באיכות גרועה למדי.

אנו מציגים לחלן דרך לשיפור עמידות מסנן ה-PF לחיבור טורי של מספר דרגות קידוד ובן מסנן עיצוב ספקטרום שגיאת הקוונטיזציה מסוג שונה מזה אשר תואר בעבודה.

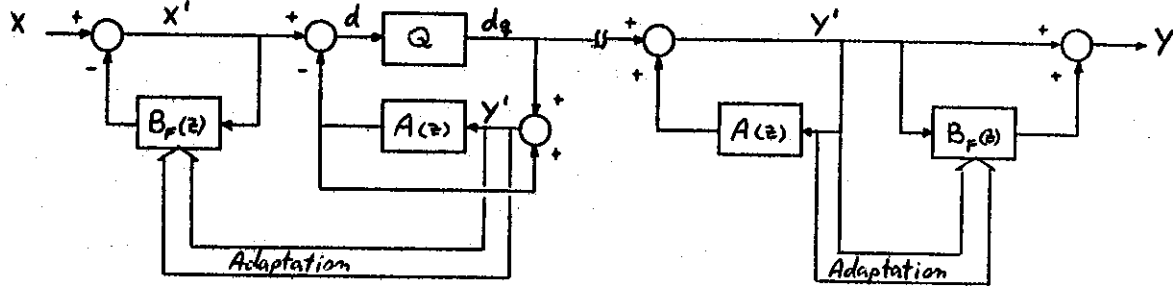
שיפור עמידות לחיבור טורי

כאמור, רוחב הסרט של אות הדיבור קטן מדרגה לדרגה, בעיה החריפה במיוחד באותות דיבור קוליים. במקרים אלה, ככל שגדל מספר הדרגות הטוריות, כך מתרכז ספקטרום האות המשוחזר בתחום התדירות הנמוך יותר. כתוצאה מכך גדל ערכו של מקדם הקורלציה הראשון ויחד איתו גדל גם מקדם מסנן הקטבים הראשון - a_1 . ניתן לנצל תופעה זו לצורך הקטנת מידת הסינון במקרים אלה. לדוגמא, על-ידי הקטנת ערכי α, β הקובעים את מידת הסינון (ראה ביטוי 5.5), עבור ערכי a_1 גבוהים. יש לבדוק באיזו מידה תפגע עקב כך יעילות המסנן בדרגה בודדת.

מבנה שונה של מסנן עיצוב ספקטרום שגיאת הקוונטיזציה

במשך העבודה בדקנו מערכת קידוד ADPCM בה מבוצעת צביעת ספקטרום שגיאת הקוונטיזציה בעזרת מסנן משוב - NFC. בשיטה זו הקוונטיזר נמצא בתוך חוג המשוב (ראה איור 5.3c), ועיצוב הספקטרום מבוצע על-ידי תוספת מעברים בין רמות הקוונטיזציה, ולא על-ידי החלקת האות כפי שזה מבוצע בשיטת ה-Postfiltering. בספרות מוצעת דרך נוספת לעיצוב ספקטרום שגיאת הקוונטיזציה בעזרת מסנן קדם (Pre-filter) ומסנן מוצא (Post-filter) הופכי לו, "החובקים" את מערכת ה-ADPCM [Atal82]. עד עתה מומשו מערכות מסוג זה שהן בעלות אדפטיביות קידמית בלבד, בהן פרמטרי המקודד משודרים אל המפענח כאינפורמציה צד. באיור 8.1 מוצג תרשים

מלבנים של מערכת קידוד כזו עם מסנן עיצוב מסוג "all-zero" בעל אדפטציה אחורית. את מקדמי מסנן העיצוב (Pre ו-Post) ניתן לחשב מתוך האות המשוחזר Y' על-פי אלגוריתם הגרדיאנט או קירוב ה- $\text{sign}(\cdot)$ שלו (ראה ביטוי 3.23).



איור 8.1 - תרשים מלבנים של מערכת קידוד ADPCM עם עיצוב ספקטרום שגיאת הקוונטיזציה בעזרת מסנני Pre ו-Post.

Figure 8.1 - Block diagram of an ADPCM coder with a noise spectral shaping by a Pre and Post filters.

מקורות

- [Adoul77] Adoul J.P. and Pradelles D., "On Line Speech / Data-Modem Identifier for Telephone Network," Proc. IEEE ICASSP, pp. 332-335, May 1977.
- [Atal82] Atal B.S., "Predictive Coding of Speech at Low Bit Rates," IEEE Trans. on Comm., pp. 600-614, April 1982.
- [Belfield84] Belfield W.R., Nishitani T., Raulin J.M., "Objective Test Results for the 32 kb/s ADPCM Coder," IEEE GLOBECOM 84, Conf. Proc., 1984.
- [Christian86] Christian R. and Adoul J.P., "Fast On-Line Speech / Voicedband-Data Discrimination for Statistical Multiplexing of Data with Telephone Conversations," IEEE Trans. on Comm., pp. 744-751, Aug. 1986.
- [Cameron83] Cameron F.L., Chow D.Q., "A Complementary Speech Detection Algorithm", IEEE GLOBECOM 83 Conf. Proc., 1983.
- [Crandall63] Crandall S.H., "Zero Crossing, Peaks, and Other Statistical Measures of Random Responses," the Journal of the ASA, pp. 1693-1699, Nov. 1963.
- [Crochiere82] Crochiere R.E. and Flanagan J.L., "Current Perspectives in Digital Speech," IEEE Comm. Magazine, pp. 32-40, Jan. 1982.
- [Daumer84] Daumer W.R., Mermelstein P., Maitre X. and Tokizawa I., "Overview of the ADPCM Coding Algorithm," IEEE GLOBECOM 84, Conf. Proc. pp. 774-777, 1984.
- [Debasis79] Debasis M., "An Almost Linear Relation Between the Step Size Behavior and the Input Signal Intensity in Robust Adaptive Quantization," IEEE Trans. on Comm., pp.623-629, March 1979.
- [Einarsson81] Einarsson G., "A Robust Adaptive Quantizer With Extended Dynamic Range," IEEE Trans. on Comm., pp. 830-836, June 1981.
- [Ewing69] Ewing G.D. and Taylor J.F., "Computer Recognition of Speech Using Zero-Crossing Information," IEEE Trans. on Audio and Electroacoustics, pp. 37-40, Mar. 1969.
- [Goodman75] Goodman D.J. and Wilkinson R.M., "A Robust Adaptive Quantizer," IEEE Trans. on Comm., pp. 1362-1365, Nov. 1975.
- [Jayant73] Jayant N.S., "Adaptive Quantization With a One-Word Memory," The BSTS pp. 1119-1144, Sep. 1973.
- [Jayant74] Jayant N.S., "Digital Coding of Speech Waveforms: PCM, DPCM and DM Quantizers," Proc. of the IEEE, pp. 611-632, May 1974.

- [Jayant84] Jayant N.S. and Noll P., "Digital Coding of Waveforms – Principles and Applications to Speech and Video," Englewood Cliffs, NJ: Prentice Hall, 1984.
- [Jayant86] Jayant N.S. and Ramamoorthy V., "Adaptive Postfiltering of 16 kb/s – ADPCM Speech,"
- [Kak69] Kak S.C., "Zero Crossing Information for Signal Reconstruction," Electronic Letters, 11th Dec. 1969.
- [Kedem86] Kedem B., "Search for Periodicities By Axis-Crossing of Filtered Time Series," Signal Processing, pp. 129–144, 1986.
- [Makhoul77] Makhoul J., "Stable and Efficient Lattice Methods for Linear Prediction", IEEE Trans. on ASSP Proc., pp. 423-429, October 1977.
- [Makhoul79] Makhoul J. and Berouti M., "Adaptive Noise Spectral Shaping an Entropy Coding in Predictive Coding of Speech," IEEE Trans. on ASSP, pp. 63–73, Feb. 1979.
- [Millar84] Millar D. and Mermelstein P., "Prevention of Predictor Mistracking in ADPCM Coders," IEEE ICC 84 Conf. Proc., pp. 1508–1512, June 1984.
- [Nakamura84] Nakamura M., Sugiyama F., Sue H. and Takada T., "Adaptive Lattice Filter LSI for Speech Signal Processing," IEEE ICC 84, pp. 1492–1495, 1984.
- [Nishitani82] Nishitani T., Aikoha S., Araseki T., Ozawa K. and Maruta R., "A 32 kb/s Toll Quality ADPCM Codec Using a Single Chip Signal Processor," in Proc. Int. Conf. ASSP, pp. 960–963, May 3–5 1982.
- [Petr82] Petr D.W., "32 kb/s ADPCM-DLQ Coding for Network Applications," IEEE GLOBECOM 82, Conf. Proc., Vol. 1, pp. 239–243, Nov. 29 – Dec. 2, 1982.
- [Porat86] פורת ב., קורט "עיבוד אותות אקראיים בזמן בדיד", הנדסת חשמל, הטכניון, חורף 1986.
- [Rabiner78] Rabiner L.R. and Schafer R.W., "Digital Processing of Speech Signals," Englewood Cliffs, NJ: Prentice Hall, 1978.
- [Rice44] Rice S.O., "Mathematical Analysis of Random Noise," BSTJ pp. 282–332, 1944; pp. 46–156, 1945.
- [Suyderhoud84] Suyderhoud H. and Williams G., "Subjective Performance Evaluation of the 32 kb/s ADPCM Algorithm," IEEE GLOBECOM 84, Conf. Proc., pp., 1984.

**A SWITCHED MODE ADPCM CODER FOR SPEECH
AND VOICEBAND-DATA TRANSMISSION AT 16 KBPS**

Research Thesis

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In Electrical Engineering

By

AMIR ZINATY

Submitted to the Senate of the Technion -- Israel Institute of Technology

Tamuz 5747

Haifa

July 1987

**This research was carried out in the Faculty
of Electrical Engineering under the Supervision
of Professor David Malah.**

I wish to express my deep gratitude to Prof. David Malah, for proposing the subject, for his helpful and instructive guidance, and for his patience throughout this work.

I would like to thank TADIRAN - Communication Division which supported this research.

I also wish to thank Mrs. Ziva Avni, System Manager, to Mr. Yoram Or-Chen, Lab's Engineer, and to my friends for helping me.

CONTENTS

	<u>Page</u>
Abstract	1
Symbol List	2
Chapter 1 – Introduction	5
1.1 Research Aims	5
1.2 Quality Criteria	6
1.2.1 Speech Quality Evaluation	6
1.2.2 Voiceband-Data Quality Evaluation	7
1.3 Structure of the Thesis	7
Chapter 2 – Quantization	9
2.1 Optimal Quantizer	10
2.2 Logarithmic Quantizer	11
2.3 Adaptive Quantizer	13
2.3.1 Forward Adaptation	13
2.3.2 Backward Adaptation	13
Chapter 3 – Differential Coders (ADPCM)	19
3.1 An all-pole Predictor	20
3.1.1 Forward Estimation	21
3.1.2 Backward Estimation	24
3.2 An all-zero Predictor	26
Chapter 4 – The Standard 32 kbps ADPCM Coder	27
4.1 Specifications	27
4.2 Algorithm Details	27
4.2.1 Adaptive Quantizer	28
4.2.2 Adaptive Predictor	30
4.2.3 Synchronization Unit	32
4.3 Simulation Results	32
4.3.1 Speech Coding	32
4.3.2 Voiceband-Data Coding	33

CONTENTS (CONT.)

	<u>Page</u>
Chapter 5 – Quantization Noise Shaping	35
5.1 Noise Feedback Coding	36
5.2 Postfiltering	37
5.3 Comparison of the Methods	39
Chapter 6 – Speech / Voiceband-Data Discrimination Algorithm ..	41
6.1 Zero-Crossing and Extremum Points	42
6.1.1 Zero-Crossing and Extremum Points of Speech and Voiceband-Data Signals	44
6.2 Signal Power	47
6.2.1 A Constant Threshold Level – T_c	47
6.2.2 An Adaptive Threshold Level – T_A	48
6.3 Summary of the Discrimination Algorithm	48
Chapter 7 – The Proposed Coding System and its Performance ...	52
7.1 The Switched Predictor	53
7.2 Subjective Enhancement of the Reconstructed Signal	53
7.3 The Quantizer	55
7.3.1 The Quantizer's Dynamic Range	56
7.3.2 The Quantizer's Parameters	56
7.4 Performance of the Proposed System	59
7.4.1 Speech Coding	59
7.4.2 Voiceband-Data Coding	62
Chapter 8 – Summary and Conclusion	64
8.1 Topics for Further Research	66
References	67
Abstract (English)	I

ABSTRACT

In this work we present an improved ADPCM coding system for Speech on Voiceband-Data signals transmission at 16 kbps (Voiceband-Data – VBD signals are binary data, modulated to achieve signals limited to telephone bandwidth. We deal with FSK and PSK modems for data rates up to 2400 bps).

Not like other coding systems, working at this rate, which are dedicated to speech signals only, the proposed system is switched between two modes of operation:

The default state of the coder is the Speech coding mode. At this mode the system is based on the standard 32 kbps ADPCM coder but is modified to operate at 16 kbps. When reducing the transmission rate of an ADPCM coder to 16 kbps, the reconstructed signal is degraded by the quantization error. Several ways for reducing the subjective effect of this degradation are well known, e.g. noise shaping using a quantization error feedback filter (Noise Feedback Coding – NCF), and filtering the reconstructed signal at the decoder (Postfiltering). In the presented system we use an adaptive Postfilter which results in a significant subjective quality improvement of the reconstructed signal.

The combination of an ADPCM coder with the adaptive Postfilter is, however, not suited for VBD as it causes an increase in the output Bit Error Rate (BER). Furthermore, we obtained in our simulations a better performance with APCM than with ADPCM (i.e. no prediction loop). The improvement is even more significant when operation with a noisy digital channel.

The switching between these operating modes is relied on a discrimination algorithm which classifies the signal to one of the two categories: Speech or VBD. The discrimination algorithm is based on the following signal's characteristics: Zero crossing rate, short time power, stationarity, and signal shape.

Based on computer simulations results, it is found that when coding speech signals, the subjective performance of the presented system is significantly superior on the conventional ADPCM coder at the same transmission rate. This is inspite of the low pass filtering effect of the postfilter. When coding VBD signals at a data rate of 2400 bps, the performance of the presented system is similar to this of the standard 32 kbps ADPCM coder.