



הטכניון – מכון טכנולוגי לישראל
Technion – Israel Institute of Technology

ספריות הטכניון

The Technion Libraries

בית הספר ללימודי מוסמכים ע"ש ארווין וויאן 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, להפיצו 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.

קידוד אדפטיבי של אוטות דיבור באנטropיה מירבית

חבור על מחקר

לשם מלאי חלקו של הדרישות לקבלת התואר

מג'יסטר במדעים

בהתבונת חשמל

n x p

נרי מלחב

833

העכניון-מכון טכנולוגי לישראל
הפקולטה להנדסת חשמל
הפקולטה למדעי המחשב

סבון

חוגש לסנט הטכניון - מכוון טכנולוגית לישראל
חיפה 1985/03/01



000001009880

המחקר נערך בהנחייתו של פרופסור דוד מלאך
בפקולטה להנדסת חשמל.

תודתי העמוקה בתונת לפרופ' דוד מלאך
על הבניהתו ומעורבותו המסורה והמלאת
בכל שלבי מחקר.

ברצוני להודות לציפי פורטנוי על עזרתה
הגדיביה במעבדה לעיבוד אותות מהלך
עובדתי.

תוכן הענייניםעמוד

1	תקציר
2	רשימת הסימונים והקיצורית
6	פרק 1 - מבוא
6	1.1 - הגדרת הבעית
7	1.2 - שיטות מקובלות לקידוד אותות דיבור
16	1.3 - מטרות המחקר
17	1.4 - מבנה החיבור
18	פרק 2 - רקע עיוני לדחיסה באנתרופיה מירבית
18	2.1 - התפישה של אנתרופיה מירבית
20	2.2 - תאור המ מערכת
22	2.3 - אופייני החזוי והשערו.
23	2.3.1 - חזאי והשער עבור טיבית אחת לדגט
31	2.3.2 - חזאי והשער עבור שתי טיביות לדגט.
38	2.4 - תכונות המערכת
45	פרק 3 - ביצועי המערכת עבור מקורות סטציונריים.
45	3.1 - תמלות של יחס האות לרעש והאנתרופיה במקדם הקורלציה.
50	3.2 - רגישות שגיאת השזרה לטיטה במקדם הקורלציה.
51	3.3 - רגישות שגיאת השזרה לטיטה מהפילוג הגאוסי.
53	3.4 - גרסאות של המערכת עבור מקורות שאינם גאוסיים.
56	פרק 4 - ביצועי המערכת עבור מקורות לא סטציונריים - אותות דיבור
56	4.1 - היצורן באדפטציה של המערכת.
56	4.2 - המערכת האדפטיבית המוצעת
56	4.2.1 - אדפטציה של ההגבר.
59	4.2.2 - אדפטציה של מקדם הקורלציה.
65	4.3 - ביצועי המערכת עבור אותות דיבור.
68	פרק 5 - חסינות להפרעות בערוץ.
68	5.1 - דיוון איקוני בהשפעת הפרעות הערוץ
71	5.2 - השפעת ראש הערוץ על שזרור אותות סטציונריים
73	5.3 - השפעת ראש הערוץ על שזרור אותות דיבור.
75	פרק 6 - סכום ומסקנות

תוכן העניינים (המשך)

עמוד

77	נפח א' - הפקת מיליצים אקראיים לא גואטימיט עם צרונן
80	נפח ב' - קידוב מפורש לפונצ'ית הפילוג הגואטיגי
81	נפח ג' - הוכחת הביטוי עבור מונה המשערך
83	מקורות
I	תקציר (אנגלית)

ת ק צ י ר

בעבודה זו נבחנת סכמת קידוד צורת-גל (waveform coding) המתבססת על גישה, שעליה יש לקידד את המקור לקבלת אנטרופיה מירבית במידע המשדר, דהיינוו, לקבלת הסתברויות הופעה שווה לאלמנטים שונים באלפבית של המידע המועבר לעורוץ.

אפיון המערכת דרש הנחות סטטיסטיות אודיות המקור, דבר המגביל במידה מסוימת את משפטאות המקוריות לה היא מתאימה. עם זאת, אבגונה מובייל לטבמה פשוטה ומיורה, ובכך ראיינו את המבנה המשעי לחקרתה. המבנה העיוני נבע בעיקר מהחזוי הלא-lienari הנדרש לימושה של המערכת, ומtoo מטרה לאפיין את קריטריון האנטרופיה המירבית כתחליף לקריטריון המקובל של שגיאה ריבועית ממוצעת מינימלית.

ביצועי המערכת נבחנים בהשוואה למערכות קיימות לגבי תהליכיים סטציונריים מהיבטים שונים: יחס אותן לרעש בשחזר, רגישות לטבויות בתוכנות הסטטיסטיות של המקור ומידת החסינות להפרעות בעורוץ. עבורו אותות דיבור מוצעות גרסאות אדפטיביות של המערכת ב- $k\text{-bit/sec}$ 8 ו- 16 kbit/sec , שעבורן נבחנו בנוסף לניל, גם האיכות של אותן המשוחרר ומידת מובנותו.

הגראטה הלא אדפטיבית של מערכת האנטרופיה המירבית הוצאה זה מכבר בספרות, אך למשתמש ידיעת המחבר, לא נבחנה עבורי מקורות סטציונריים, אלא ישירות לגבי תמונות ודיבור. כמו-כו, לא הוצאה עד כה גראטה אדפטיבית. התוצאות האובייקטיביות של הסימולציות לא הعلاו, למעשה, שקיים יתרון כלשהו להעדרי את מערכת האנטרופיה המירבית בתוכנות הגוכחים על פבי מקדי צורת גל קיימים. אדרבא, עבורי תהליכיים צרי-טרט הסתמנת עדיפות בולטת למערכות התקיימות. מבחינת איכות השזר דיבור הובנה גם כן עדיפות קלה לסטמות מקובלות של קידוד צורת-גל, אך ההבדלים לא היו משמעותיים. אולם, לצורך שמר המבוגנות בשזר לפי אנטרופיה מירבית השתמן צורך באדפטציה מלאה של המערכת, דבר המגדיל את מידת סיובוקיותה לאין ערוך ביחס למערכות אליהן הושוותה.

יחד עם זאת, מועלות מספר הצעות לשיפור המערכת וモוצעים כווננים למחקר נוספים.

רשימת סימונים וקיצורים

A_{n,k}(y) - קבוצת סדרות המדגמת באורוּ k {x_{n-k}, ..., x₁} הגרומות לחזאי לפלוט ערך y ברגע ח.

A_{n,k}ⁱ(y) - האיבר ה- i בסדרת קבוצות המתכנסות ל-(y).

B_n(ρ) - פולינום ממעלה ρ המתאר את מקדם הקורלציה לאחר א-טילינאריות במקרה $x = N(x)$.

C - קיבול עroz.

C_n - משערך המכפלה הפנימית בין דגימות סמכיות.

D - עוזות = תוחלת השגיאה הריבועית בשחזר.

E{·} - אופרטור התוחלת.

e_n - אמצע המשודר בזמן הדיסקרטי ח.

e_n¹ - הסיבית המשמעותית של e_n.

e_n['] - אמצע הנקלט בזמן הדיסקרטי ח.

F(e_n, e_{n-1}, y_{n-1}) - גרסה אי של המשוחרר עבור 2 = R: ניצול שתי הסיביות של e_{n-1}.

$\tilde{F}(e_n^1, e_{n-1}^1, y_{n-1})$ - גרסה ב' של המשוחרר עבור 2 = R: ניצול הסיבית המשמעותית בלבד של e_{n-1}.

f_x(x) - פונקציית ציפיפות הסתברות של x.

F_x(x) - פונקציית פילוג הסתברות של x.

F_o - אינדקס דגימת חיתוך (oversampling index).

g(x) - פונקציית ציפיפות הסתברות של מ.א. גואטי ($\sigma^2(N)$).

g(x,y) - פונקציית ציפיפות הסתברות משותפת לזוג משתנים אקראיים גואטיים במשותף.

g(x|y) - ציפיפות הסתברות של x המותנית ב-y כאשר x ו-y מ.א. גואטיים במשותף.

g(x|y>y_o) - ציפיפות הסתברות של x המותנית במאורע {y > y_o} כאשר x ו-y מ.א. גואטיים במשותף.

G(e_n, e_{n-1}, y_{n-1}) - המשוחרר עבור סיבית אחת לדגם.

H(e_n) - אנטרופיה מדיר ראשון של המידע המשודר.

H(e_n | e_{n-1}) - אנטרופיה מדיר ראשון המותנית בעבר המידע המשודר.

רשימת סימונים וקיצורים (המשך)

- אנטרופיה מסדר ראשון המותנית בעבר המידע המשודר וערך החזוי. $H(e_n | e_{n-1}, y_{n-1})$
- משערך האנטרופיה מסדר ח בעבר סדרת המדגם ה- i -ית. $\hat{H}_n^i(e)$
- משערך האנטרופיה המותנית מסדר ראשון עבור סדרת המדגם ה- i -ית. $\hat{H}_i(e_n | e_{n-1})$
- משערך האנטרופיה המותנית מסדר ראשון. $\hat{H}(e_n | e_{n-1})$
- $H_b(x) = -x \log_2 x - (1-x) \log_2 (1-x)$ - האנטרופיה הבינארית: $(x \log_2 x + (1-x) \log_2 (1-x))$
- K - מטר סדרות המדגם.
- פונקציה חזוי לסדרה כמוני תחומרן עבור $2 = R$. $L(y_n, e_n)$
- קירוב של $L(y_n, e_n)$. $\hat{L}(y_n, e_n)$
- פונקציה חזוי. $M(y_n, e_n)$
- $M_{opt}(y_n, e_n)$ - האיבר ה- n בסדרת פונקציה חזוי המתכנסות ל- $M_i(y_n, e_n)$.
- פונקציה חזוי אופטימלית. $M_{opt}(y_n, e_n)$
- med{•} - חציון (median) של פלוג טוגולי.
- N - אורך סדרת המדגם.
- $N(x)$ - אי ליניאריות (להמראת פלוג טוגולי).
- $N^{-1}(x)$ - הפונקציה ההפוכה ל- $N(x)$.
- p - הסתברות טוענת בודדת בערוץ.
- $p(x)$ - פונקציה פלוג טוגולי (צפיפות הסתברות) של משתנה אקראי x.
- $\hat{p}_i(e_1^n)$ - המופיע היחסי של הווקטור (e_1, e_2, \dots, e_n) בסדרת המדגם ה- i -ית.
- P(x) - אופין חזוי בארגומנט אחד.
- $P^{-1}(x)$ - הפונקציה ההפוכה ל- $P(x)$.
- Q(x) - דוגמאות לפונקציה הפלוג של ח.מ.א. הגאומטרי.
- $Q^{-1}(x)$ - הפונקציה ההפוכה ל- $Q(x)$.
- R - קצב השדור (בטייכות לדגט).
- R(D) - פונקציה קצב-עוזת.
- $R(t, s)$ - פונקציה עזר למונה המשערך המשוחרר.
- S_n - משתנה המצביע של מגננון בקרת ההגבר האדפטיבית (AGC).
- S'_n - גרטה של S_n ללא זהוי מסלול גבולוי.

רשימת סימונים וקיצורים (המשך)

$T(x)$	- פונקציה עזר לקרוב (x).P.
(y_n, e_n)	- פונקציה חזוי לטף כמוני עליון עבור 2 = R.
$\hat{U}(y_n, e_n)$	- קירוב של (y_n, e_n) .
$\{V_n\}$	- רוש לבן גאוסי בעל תוחלת אפס ופזר ייחידה.
w_n	- ערך חזוי לטף כמוני תחתון ברגע n.
x_n	- אותן הכנות (= המקור) ברגע n.
\hat{x}_n	- אותן המשוחזר במקלט ברגע n.
x'_n	- אותן הכנות לאחר הפעלה אי לינאריות.
\hat{x}'_n	- שזרור של x'_n .
\tilde{x}_n	- אותן המשוחזר המנורמל ברגע n (במערכת עם AGC).
x_i^j	- וקטור הכבריות $(x_j \dots x_{i+1}, x_i)$ כאשר $i > j$.
y_n	- ערך החזויה במדיר (טף הכמוני האמצעי) ברגע n.
y'_n	- הגרסתה של y_n במקלט.
z_n	- ערך החזויה לטף הכמוני העליון ברגע n.
α	- קבוע = 0.67448
α_i^n	- מקדמי הפולינום (α_i^n) B.
Δ	- גודל צעד הכמוני.
Δ_{opt}	- גודל צעד הכמוני האופטימלי.
y_n'	- שגיאת החזויה עקב הפרעות בערזץ = $y_n' - y_n$.
λ	- מקדם "שכחתי" של מגנון האדפטציה.
λ_1	- מקדם היישכחתי כאשר המערכת אינה במלול גבולי.
λ_2	- מקדם היישכחתי כאשר המערכת מתונה במלול גבולי.
s	- סטיית התקן של המקור.
\hat{s}_n	- המשערך של s ברגע n.
σ	- מקדם הקורלציה בין דוגמאות סמוכות של מקור טציזונרי.
$\hat{\sigma}_n$	- המשערך של σ ברגע n.
c_m	- מספר אותיות הערזץ לשניה.

רשימת סימונים וקיצוריים (המשך)

s_o - מספר אותיות המקור לשניה.

$\{s_k\}$ - סדרת האוטוקורלציה של התהליין הגaussיאי (לפניהם תילינאריות).

$\{s'_k\}$ - סדרת האוטוקורלציה של התהליין אחרי תילינאריות.

.Adaptive Differential Pulse Code Modulation - ADPCM

.Differential Pulse Code Modulation - DPCM

.Delta Modulator - DM

.Linear Delta Modulator - LDM

.Continuously Variable Slope Delta modulation - CVSD

.Maximum Entropy - ME

פרק 1 - מבוא1.1 הגדרת הבעייה

בבית הדחיטה (קידוד) של אותות דיבור מהוות מזח כבר גושא למחקרים רבים, אך בשל האתגר הכרוך בפתרוניה, והן בשל הישומים הנלוויים לה. השימוש במקדרים לאותות דיבור הוא ויזוני לצורך אחסון עיל של דיבור במגاري מידע או לשט העברתו בעברוצי תקשורת ספרתיים צלי-טרט. לחילופין, דחיטה עילית של מידע הגלוט באות דיבור מאפשרת העברת מספר רב של שייחות בו-זמןית בערוצ אחד. האתגר המרכזי הבינץ בפני החוקרים בתחום זה הוא כיצד למצער את קצב שידור המידע (בטיبيות לשניה) עבור ביצועים רצויים, או לחילופין כיצד לקבל ביצועים אופטימליים לקצב שידור נתון. המונח "ביצועים", בהקשר זה, מכובן לשני ממדים: מוכנות (intelligibility) ותאילכיות (quality) של דיבור המשוחזר במקלט. מוכנות הדיבור היא מדרט טטיטטי כמותי, אותו ניתן להגדיר היטב על פני אוכלוסייה עשרה מטפיך של מאזיניט, והוא קשורה להיבט האינפורטיבי גריידא של אות הדיבור. לעומת זאת, האיכות מהויה מדרט סובייקטיבי המציין באיזו מידת ההאזנה נעימה. שני המדרדים חניל הם על פי רוב בלתי תלויים כלומר, דיבור מובן אינו בהכרח בעל איכות גיבתא, אך גורם לירידת משמעותית באיכות. ככל זאת איכות משפיעה בעקיפין על מוכנות, משות שהאזנה ממושכת לאות באיכות גרוועה עלולה לעילף את המאזין האנושי ולגרום לירידת בתפישה.

הקיים המרכזי הכרוך במחקר ופיתוח של שיטות דחיטת דיבור הוא שלא מוכר מדרט כמותי מובחן אשר מופיע נאמנה את האיכות ו/או מוכנות מחד גיבתא, וכי יכול להיות פונקציית מטרה בתחילת האופטימיזציה של תכנון המקדר מיידך.

מסיבה זו מתקבל לבחור בפונקציות מטרת המשקפות מידות עותות שונות של אות המשוחזר ביחס לזה המקורי בתחום הזמן או בתחום התדר. הפונקציה המקובלת ביותר היא שגיאה ריבועית ממוצעת. למרות שעיל פי רוב אין בכוחה לשקף נאמנה את איכות או מוכנות, צאתה פונקציית מחיר זו לפופולריותבולטת בספרות עקב מספר סיבות:

1. אופטימיזציה של השגיאת הריבועית מובילה במקרים רבים לפתרון מערכת משוואות לינאריות.

2. מוכרים מתרת האינפורטיציה חסמים תאורטיים (פונקציה קצב-عروת) המתאים לשגיאה ריבועית ובאמצעותם ניתן להעריך ביצועים של אלגוריתמים.

3. לשגיאת הריבועית הממוצעת יש מובן פיטקללי והוא הטעק של שגיאת הקידוד (או רעש הקידוד) ועל כן היא קשורה ישירות ליחס אותן לרעש (SNR).

ברם, כפי שנראה בהמשך, בקצבבי שידור נומכית, בהם מתבצע ע"י המקודד כמווי (קונטיניזציה) למספר קטן יחסית של רמות ומידת אי הליניאריות של מערכת השידור היא גדולה, אין עוד הצדקה לפטור את בעיית האופטימיזציה באמצעות מערכת משוואות ליניאריות וממעוררים קשיים בפתרונה המדוייק. ישנו איפוא צורך לחפש קרייטריון אופטימיזציה תחליף לבעית הדחיטה בקצבבי שידור נומכית.

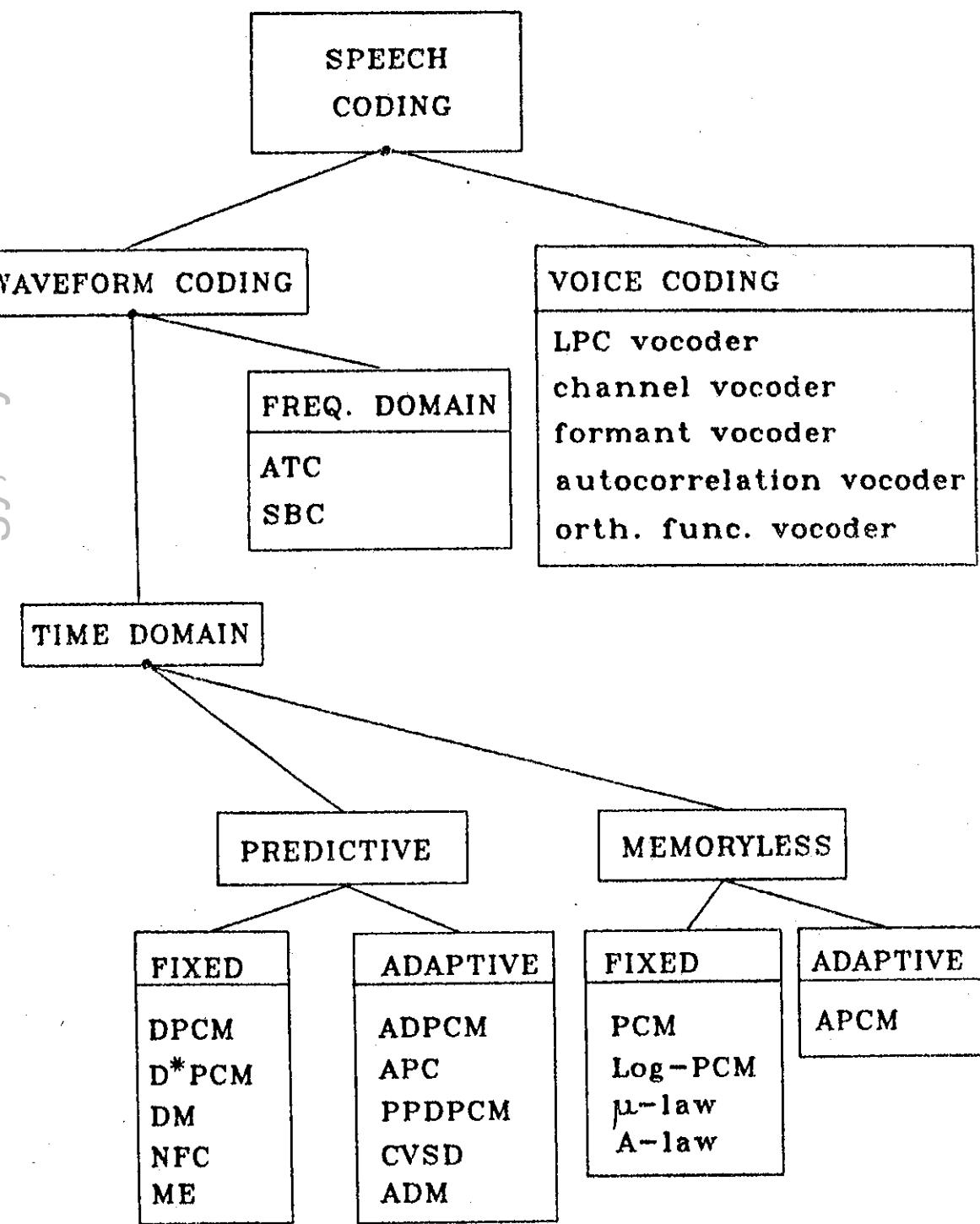
בעובדה זו נבחנת גישה תחליף לבעית הדחיטה אשר מבוססת על מקסימיזציה של האנטropיה במידע המשודר (bit-stream). מקורו של רעיון זה, שהוצע לראשונה ע"י Angel ו-Daigle ([2], [3]), הוא בכך שאנטropיה מיליבית היא תכונה נלוית של מקודד אופטימולי לקצב נמוך.

בהמשך פרק זה נסקור בקצרה גישות ושיטות קיימות לדחיטת דיבור ונעמוד על מעלותיהן ו מגבלותיהן. לבסוף, נציג את מטרות המחקר הנוכחי וננסח בקצרה את מבנה חיבור זה ותרומתו.

1.2. שיטות מקובלות לקידוד דיבור

בטעיף זה נתאר בקצרה מספר מערכות לדחיטת דיבור שהוצעו בספרות. המגוון הרב של סוגי המערכות מייד, למעשה, על אף שלא קיים פתרון אחד לבעית הדחיטה שהוא משביע רצון לאלווטין. לא נעמוד בטעיף זה על כל סוגי המערכות הקיימות אלא נתרכז בעיקר במערכות המושתתות על גישה הדומה לזה של המערכת הנדגונה בחיבור זה. תשומת לב תוקדש, כמו-כו, לתמורה (trade-off) שבין קצב השדרור, האיכות והטיפוליות.

קיימות למעשה שתי תפישות מרכזיות בנושא הקידוד של אותות דיבור: קידוד קול (voice coding) או קידוד צורת-אל (waveform coding). מערכות קידוד-קול (vocoders) מבוססות על מודל פיזיקלי כלשהו של מנגןון ההפקה של דיבור. המידע המשודר במערכות מסווג זה הוא אוסף הפרמטרים לפי המודל המתאים למתקבלים מתוך שערוך על פני קטיעים קווזי-סטאצionarioים של האות. הייצוג הפרטורי יכול להיות מסווגים שונים (ראה ציור 1.1): למשל, רכיבי הספקטרום בזמן קצר (כמו ב-channel vocoder), מוקדי ווזוי ליניארי המתארים את עוטפת הספקטרום (LPC), תדרי מהודה עיקריים (formant vocoder), סדרת האוטוקורלציה (autocorrelation vocoder) וко'. הסיבוכיות החישובית מוקדי פרוק אורטוגונלי (orthogonal function vocoder) וכו'. הסיבוכיות החישובית של מערכות קידוד קול היא יחסית גבוהה ודרושים מנגןוני סyncron בין המרכיבים השונים של המידע המשודר (תדר יסודי, החלטה voiced/unvoiced ופרמטרים). אותן המשוחזר נשמע על פי רוב לא טבעי וחביצועים תלויים בדבר ורגישים לרעש. מאידך, קצב השדרור נומכית למדי ($s/4.8\text{ kbit}$) וזו את בגל ניצול עיל של היתירות (redundancy) הגדולה של המידע הגלום באות דיבור.



ציור מס' 1.1 - משפחות של מקדי דיבור.

Fig. 1.1 - Families of speech encoders.

מקדי צורת-gal

המחלקה השניה של מקדי אוט - מקדי צורת-gal, שאליה משתיכת המערכת הנדרנה בחבור זה, היא עשירה וענפה למדי ובה עוסקת בהמשך סעיף זה. במערכות מסווג זה הקידוד אינו פרטורי ומתרחן, כפי שמשמעותו מהשם, היא שימור צורת gal.

מקדי צורת-gal מתוכנים עקרונית ללא תלות באופי אות המקור ומתחמייט, לפיכך, באותו מידה עבור אותן מסווגים שונות - דיבור, מוסיקה, תמונות, אותן תקשורת, אותן ביולוגיים וכו'. לגבי אותן דיבור, מקדי צורת-gal אינם רגילים באופן ניכר למיפויים של הדובר ולרעש רקע, הם פשוטים יחסית מבחינה חישובית אבל פועלם בקצב גבויים יחסית למקדי קול (בדי'ק מעל 7.2 kbit/s). ניתן גם לאכנו מקדי צורת-gal באופן מתאים יותר לאור האות על מנת לתרוויח ביצועים ו/או ביעילות הקידוד וזאת מתוך ציפויות סטטיסטיות, הכרת אופיו הפיזיקלי של אותן ו cedar.

מאחר שמקדים אלו בניגוד למקדי-קול, כאמור, מתוכנים לשאזר צורת gal, ניתן בקלות עבורם מרד ביצועים אוביקטיבי - יחס אותן לרעש. ברם, כפי שצוין בסעיף 1.1 לא תמיד ישנה התאמאה בין יחס אותן לרעש לבין ההערכה האוביקטיבית, קרי: האיכות. יתרה מזאת, עבור האיכות אין כל מדדים ממשיים. מקובל לפיכך [5], לאפיון איכות על-ידי ארבע רמות:

- .1 - רוחב סרט עד KHz 7 ללא רעש ועוצמת (מעל kbit/s).
- .2 - יחס רעש מעל 30dB, תחום תדרים Hz 200-3200, עות % 2-3, (16-64 kbit/s).
- .3 - מובנות גבואה, איכות בינונית (s/s).
- .4 - צליל בלתי טבעי, קשה לזהות את הדובר (פחות מ-s/4.8 kbit).

שלוש הרמות הראשונות של איכות הן בקצב הפעולה של מקדי צורת-gal וחרביעית מתאימה למקדי קול.

את מקדי צורת gal ניתן לחלק לשני סוגים: מקדים במוחות הזמן ומקדים במוחות המתדר. מקדים במוחות התדר, כמו SBC (Sub-Band Coder) ו-ATC (Adaptive Transform Coder) [5], מבוססים על גישה אשר על פיו מחלקים את רוחב הסרט למספר פסי תדר או רכיבים ספקטראליים ומקדים כל אחד מהם בנפרד. היתרון בכך הוא שמספר הסיביות לכל פס תדר קבוע באופן דינמי על פי רמת האנרגיה היחסית שבוחמו. באופן כזה, הקצאת הסיביות בין תחומי התדר השונים היא מתאימה, ונitin למעשה, לשלוט באופן הספקטראלי של רעש הקוונטיזציה. במקדי SBC בקצב של kbit/s 16 מתקבלת איכות השcoleה זו של מקדי ADPCM בקצב kbit/s 22. זה האחרון הוא מקדי במוחות הזמן, שעליו יורחב הדיבור בהמשך. הטיבוביות של SBC, לעומת זאת, גבוהה יותר: במשדר SBC מופעל בנק מסננים המפרק

את האות לפסי תדר, ובכל פט מדר מבוצעים דצימציה וקידוד בתחום הזמן תוך החלטה שוטפת בדבר הקצת סיביות בין תחומי התדר השוניים. במקיד h-ATC, שהוא בעל סיבוכיות גבוהה מזו של SBC, מבצעים כמוי למקדמי התמרת קרטיס (KLT) Karhunen-Loeve (DCT) של קטעי דיבור קבועים באורכם בזאת אחר זה.

מקדמי צורת-גָל בתחום הזמן

בתחום הזמן קיימים סוגים רבים ומגוונים של מקדים. פשוטים שבתם, אין מנצלים את התלוות הפטיטטיבית בין הדוגמים הסטטיסטיים של אותן הכנסיטה, ומבצעים כמוי לכל דגם בנפרד ולא תלות בעבר, ולפיכך הם נקראים מקדים אטריא זכרון (Memoryless Predictive encoders או instantaneous encoders). בטכניות מתוחמות יותר (quantization) הזכרונו של אותן הכנסיטה מנוצל בשיטות שונות.

המקיד פשוט ביותר ללא זכרונו הוא ה-PCM (Pulse Code Modulation), אשר בו מחבצע כמוי אחד והוא למעשה כמוי ממיר אנלוגי-ספרתי (A/D). עבור מקיד מסווג זה דרושות לפחות 11 סיביות לדגם על מנת להבטיח שהשזר ישמע נאמן ולא עוותים (מעל 3dB).

בגרסת מתקדמת מעט יותר של ה-PCM, המכוי אינו אחד אלא מפולג באופן לוגריאתמי (log-PCM). ע"י כך משליגים כמוי עדין יחסית עבור רמות אותן נמוכות, שכן על פי רוב שכיחות יותר, וכמו גם יותר בرمות הגבוהות. מסתבר שעבור PCM-log דרישות 8-7 סיביות לדגם עבור toll quality האיכות המקובלת במערכות טלפוןניות.

מובן כי ידיעת הפולג הפטיטטיבי של עוצמת אותן הכנסיטה מטייעת למכנוך קוונטיזר מתקאים. ה-PDF optimized quantizer, שהוצע מזה כבר ע"י Max ו-Lloyd, הוא חטיב ביותר שבין המקדים (חטרי הזכרונו) על פי קרייטריוון עותם נקבע כאשר ידוע פולג הכנסיטה (כפי ששთमע משמו). ברט, במרבית הישומים אין כל ידע אפרורי או דוחות פולג הכנסיטה. יתרה מזאת, על-פי-רוב אותן הכנסיטה איננו סטציונרי, אבל על כן יורחוב הדיבור בהמשך. במקרים כאלו טבעי לבצע כמוי אחד, משומם ש מראש לא ברור שכמיי יהיה עדייף. במקרה אחד מטוגן זה יש להabil את המתחום בו מניחים שהאות משתנה. שגיאת המכוי תהיה שניי סוגים: שגיאת רוויה (overload error), הנובעת מחיראה של האות מתחום המכוי, ושגיאת "גרגראית" (granular error), שהיא תוצאה המכוי בתוך המתחום. מובן כי הקטנת המתחום הדינמי של המכוי תקטין את השגיאת ה"גרגראית" על אשבען הגדלת שגיאת הרוויה.

על-מנת למזער את מידת ההשפעה המשולבת של שתי שגיאות אלו יש למצוא את המתחום הדינמי האופטימלי והוא מובן תלוי בפולג הכנסיטה. לדוגמה, עבור פולג כניסה גאוסי כמוי אופטימלי (לא אחד) יבטיח יחסאות לרעש של $B_p(4.3-6R)$ כאשר מספר הסיביות לדגם - R הוא גדול מספיק ($5 \geq R$). לעומת זאת, כמוי אופטימלי אחד יוריד את

יחס ל- $B_p(0.6-5R)$. החטט העליון עברו במקרה היגנטי, שבתון ע"י R 6 (מתוך פונקציית קצב העוזות), איבנו בר-הgasma בטכניות כמו סקלריות. תוצאות נוספות נוספות עברו פלוגטיס מסוג גמא ולפלט מובאות במפורט בספרם של Rabiner & Schaffer [9], וכמו כן בספרם של Jayant ו-Noll [7].

אות השיטות לקבלת כמו אחד היא הפעלת פונקציה לא ליניארית מתאימה (x) N לפניו ה כמו (comanding) והפעלה התופכית לה בשזר.

משמעות שעל-ידי בחירה מתאימה של (x) N ניתן להקטין במידה ניכרת את מידת הרגשות של שגיאת ה כמו לעוצמת האות בתחום רחב מדי.

פונקציות מקובלות עבור (x) N הן ה- law-n וה- A-law. לפונקציות אלו התנחות ליניארית בסביבת הראשית ואופי לוגריטמי עבור ערכיהם גבוהים של x. מערכות law-n ו- A-law ניתן לצפות ליחס אותן לרעש של $B_p(10-6R)$ בקיורוב. לתאור מפורט יותר של מערכות אלו ראה [5], [7] ו-[9].

כמו אדפטיבי (APCM)

גישה אחרת של פיה מניטים להטמוד עט השינויים בעוצמה של האות היא שימוש במנגנון הסתגלות (אדפטציה) היילומדים את מהלך עוצמת האות ומאמילים את ה כמו לפי הצורך.מערכות תקשורת טיפוסיות, בהן התוחם הדינמי של אות דיבור מגע ל- $B_p(40$ עדiphota מיעוט אדפטיבית מסווג זה על פני מערכות A-comanding אשר קבועות בזמן. מנגנון אדפטיבי מאפשר "ניפוי" או "כווץ" של אופי ה כמו בהתאם לעוצמת האות בתחום רחב מאוד.

את מנגנון האדפטציה ניתן להפעיל לפי אות הכנית או לפי האות המשוחזר. במקרה הראשון (AQF - Adaptive Quantization Forward) יש צורך לשדר בנוסף מידע נלווה (side information) בדבר האדפטציה של הקוונטיזר וזאת משום שבמקלט אין כל גישה לאות הכניתה. במקרה השני (AQB - Adaptive Quantization Backward) אין צורך בשידור כל מידע נוספת משום שהאות המשוחזר זמין גם במקלט וביתן על-פיו לחשב באופן מוקומי את משתנה האדפטציה המתאים. מובן כי משדר AQF מכיל גם העתק של המקלט. היתרונו של AQB הוא, כאמור, בכך שאין צורך במידע נלווה, ולפיכך גם לא דרוש סync'רו עם ערך המידע העיקרי. אולי, חסרונו הוא בכך שהוא רגיש יותר מה-AQF להפרעות בערוצ, המשדר שלו מסובך יותר והאדפטציה שלו היא טובת פחות משום שהוא פועל על אות המשוחזר ולא על אות המקורי.

מנגנון האדפטציה מכיל משערר (רכורטיבי) של עוצמת האות (המקור או המשוחזר) ונרמול מתאים (AGC - Automatic Gain Control), או לחילוף משערר של גודל צעד ה כמו (step-size). שתי האפשרויות שקולות להלוטון. לצורך שערור שוטף של עוצמת האות, מקובלות שתי גרסאות:

1. בקרת הגבר עם זכרון אקספוננציאלי:

$$\sigma^2(n) \triangleq (1-\lambda) \sum_{k=0}^{\infty} \lambda^k x^2(n-k) = \lambda \sigma^2(n-1) + (1-\lambda)x^2(n) \quad (1.1)$$

$$(1 < \lambda \leq 0)$$

λ - מקדם היישכחה".

2. בקרת הגבר עם זכרון סופי:

$$\sigma^2(n) \triangleq \frac{1}{M} \sum_{k=0}^{M-1} x^2(n-k) = \sigma^2(n-1) + \frac{1}{M} [x^2(n) - x^2(n-M)] \quad (1.2)$$

עבור המנגנון בעל זכרון אקספוננציאלי, שהוא המקובל יותר, ערכו של λ נוע על פי רוב בין 0.9 ל-0.99 ובעמצעותיו ניתן לשוט על מהירות האדפטציה. תפקיד דומה יש לפרמטר M עבור המשערך בעל הזכרון הסופי ב-(1.2).

לגביה אדפטציה של גודל צעד המכמי ישנה ב-AQB אפשרות לביצה על-פי מילת הקוד המשודרת במקום מתוך האות המשומזר. אלגוריתם עדכונו גודל הצעד - Δ_n הוא באופן עקוריוני מהצורה הבאה:

$$\Delta_{n+1} = \Delta_n \cdot (|\text{밀ות קוד קודמות}| M) \quad (1.3)$$

מנגנון זה מאפשר לקוונטייזר ל"המכווץ" או ל"ייתנפח" בהתאם להתנהגות מילות הקוד המשודרות כוכח שנoilיט בעוצמת האות ועל פי לוגיקה מתאימה.

כדי להטגבר על השגיאות בערזק מקובל להכניס גורם "שכח" $1 < \beta \leq 0$:

$$\Delta_{n+1} = \Delta_n^\beta M \cdot (|\text{밀ות קוד קודמות}|) \quad (1.4)$$

באופן כזה נמנעים מהזכרון האינטימי של (3) לטעויות בערזק.

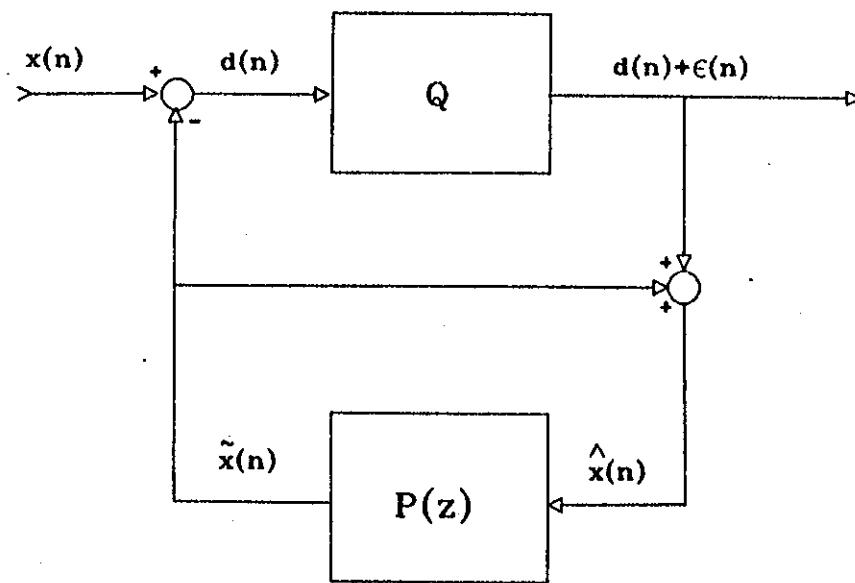
מקדים בעלי זכרון

בתיקות מעטה לטכניות קידוד מתקדמות יותר מ-PCM המנצלות את הזכרון הטבעי של אות הכניסה.

תורת האינפורמציה מבטיחה ניצול מלא של הזכרון הגלום באות הכניסה כאשר מפעלים טכניות של כימוי רב-מיידי (Vector Quantization - VQ) והן עדיפות על פני שיטות סקלריות גם עבור מקורות חסרי זכרון. אולט שיטות אלו כרוכות בסיבוכיות חישובית ובדרישות זכרון גבוהות לאין שעור. במרבית המקרים גם אין אלגוריתמי חכמנו מדוייקים המבטיחים קוונטייזר רב-מיידי אופטימלי, וזאת גם כאשר פילוג

הכניתה ידוע. במערכות פשוטות יותר מונצלים את הזכור הטבוע באמצעות חיזוי (Predictive Quantization). הגישה היא לקדר את שארית החזוי שהיא בעלת אנרגיה נמוכה יחסית ובכך לאפשר הקטנה ניכרת של תחום הכמוני ולכך גם צמצום של גודל הצעד.

במשדר PCM* [6, 9], אותן הרכבת עובר סיגנון "מלביון" לפני הכמוני, היינו, מקדמי המטען נבחרים כך שאות המוצא ממנו יכיל אנרגיה מזערית ותס מוחשיים בטכניקות של חיזוי לינארי. במקלט מתבצע סיגנון הפוך לשארית החזוי (לאחר שעברת כמוני). האתגרון של שיטה זו הוא ששגיאת הכמוני "נצחעת", ובದ"כ גם גדלה עצמאית כתוצאה מפעולות הסיגנון הפוך שבמקלט. מטיבה זו מודיעיפיט לשלב את החזאי בחוג משוב סיבוב הקוונטייזר כמו במשדר DPCM (Differential PCM) [6, 9] המתואר בציור (1.2).



ציור מס' 1.2 - משדר DPCM.

Fig. 1.2 - DPCM transmitter.

יתירון של מבנה זה ביחס ל-PCM הוא, שగיאת השחזר - $(n)\hat{x}(n) = (n)x(n)$ א' שווה בדיקוק לשגיאת הכמוני. א' אם מקבלת מ- $(n)d$ שהוא, כאמור, א' קטן יחסית. השפור ביחס א' לרעש בתשווואה ל-PCM הוא 11dB-4. כשת שב-PCM בחרת גודל צעד הכמוני

מתבצעת משיקולי תמורה (trade-off) בין שגיאת רוויה ושגיאת "גרגירית" ובקבועה על-פי התוצאות הדינמי של אותן, הרי גם ב-DPCM גודל הצעד משפיע על האיזון שבינו לשגיאות אלו. בהיותה של DPCM מערכת דיפרנציאלית, גודל הצעד שלה נקבע גם על פי הקורלציה בין הדגמים הסטטיסטיים של אותן. כאשר הקורלציה גבוהה רצוי לבחור גודל צעד קטן ע"מ להרחיק את השפעת הרעש ה"גרגירית". אם הקורלציה נמוכה, קרי אותן משתבוחות במתירות, אז יש לאפשר גם לפחות המשוחזר להשטנות במתירות ולהמנע משגיאת "רוויה-שיוף" (slope overload) וזאת ע"י בחירת צעד כמו גודל יחסית. דברים דומים אמורים גם לגבי DM (Delta Modulator) המהווה למעשה גרטה של DPCM עם שתי רמות כמו בלבד (סיבית אחת לדגם). החטרון של DPCM לגבי PCM* P הוא שהחזרי מתבצע על פי אותן המשוחזר. גם חישוב המקדמים האופטימליים עבור החזרי $(z)^P$ כרוך עקרונית בהתרת מערכת מסוימות לא לינאריות, וזאת עקב האינטראקציה הגובהת של החזרי עם הקוונטייזר שהוא אלמנט לא לינאר. כאשר מספר הרמות הוא גבוהה מספיק, ניתן לראות את שגיאת הכמוני $(a)^E$ כרעש לבן אידיטיבי שאינו תלוי בכניםה, ואז בעית תכנון החזרי לובשת שוב אופי לינאר. במקרה זה סט המקדמים האופטימלי $T(a_p, \dots, a_1, \underline{a}) = \underline{a}$ של החזרי $\underline{z} = z^P$ מתקבל ע"י התרת מערכת מסוימת (ראה גם [9]).

$$(R + \sigma_e^2 I) \underline{a} = \underline{r} \quad (1.4)$$

כאשר \underline{r} – טור החזרי (שערכו הטיפוסי עבור דיבור אינו עולה על 4), R היא מטריצה קאך שבת האיבר $-T(z_i)$ הוא האיבר ה- $(j-i)$ של סדרת האוטוקורלציה – $(a)^E$ של אותן הכניטה $(a)^x$, ו- $\underline{a} = T(x^1, \dots, x^2, x^1)$. אם מספר רמות הכמוני הוא גבוהה, עוצמת שגיאת הכמוני σ_e^2 היא זניחה בהשוואה ל- R ו- \underline{a} מתקיים מתוך התרת בעית חזוי טנדנטית. ברום, כאשר מספר רמות הכמוני הוא נמוך, לא ניתן עוד להזניח את σ_e^2 והוא בעצם תלוי במקדים \underline{a} , דבר ההופך את הבעיה לבתאי לינארית. יתרה מזאת, מערכת המשוואות (1.4) מתקפת אך ורק תחת התנזה $-T(a)^E$ הוא רעש לבן ובתאי תלוי בכניםה, דבר המתקיים כאמור רק אם מספר הרמות גבוהה מספיק. פירוש הדבר שקיים קושי בתכנון החזרי בקצב שדור נמוכים כמו אלו בהט أنهן עוסקים במתגרת חיבור זה (סיבית אחת או שתי סיביות לדגם).

אותהagiשות לבעה זו היא לוותר על חミニימיזציה של שגיאת השזרור עבור קצב השדרור הנוכחי, ומחליף, למוצר את קצב השדרור ללא עוזות נוספת ע"י הפעלת טכניקות של קידוד אנטרופיה (entropy coding) בטור למשדר הקיטים. הקושי בכך הוא שלאלגוריתמים מסווג זה פועלם בד"כ בקצב משתנה, דבר הדורש חוץ (buffer) ומסרב את פעולת הבדיקה. יתרה מזאת, קידוד אנטרופיה אינו מavailם כאשר נדרש לתקן מערכת אופטימלית לקצב שדרור נתון. במלחיף, הוצאה בשנת 1983 ע"י Daigle ו-Angel ([2], [3]) הגישה שעליה

מתבסס מחקר זה, ולפיה החזאי מתוכנן לשירות ל渴別ת אנטרופיה מירביה במיוזע המשודר בקצב המקורי (Maximum Entropy coding - ME) ועיי' כך הוא הופך לבתאי דחיס ואינו צורך בקידוד אנטרופיה. ההנחה (שנבדקת בין השאר במתגרת מחקר זה), היא שאם רמות הcamoi הן שותה הסתברות אזי הן מנוצלות באופן יעיל ויש בכך תועלת כלשהי למען מצומצם שגיאת השחזר. אנטרופיה מירביה היא גם תוכנה נלווה של מערכת דחיטה אופטימלית המגשימה את פונקציית קצב העות, ועל כך יורח הדיבור בפרק 2. תכונה נוספת של מערכת ME המוצעת היא, שהחזאי שלח איננו ליגנארי. בכלל, טכניקות לא נוטפת של לינאריות לחזוי אופטימלי לא זכו לפופולריותגדולה בין המערכות של Predictive Coding בעיקר בשל הקשי העקרוני הכרוך בימושו: חישוב תוחלת הדגם הנוכחי של אותן, המותנית בדגמי עבר מהאות המשוחרר דרוש ידיעת האופי הסטטיסטי המשותף לאות הכנישה ולשגיאת הcamoi וכרכוך לפיכך בפתרון בעיות בנות "מידה גדולה של אי-לינאריות" יתרה מזאת, האופי הסטטיסטי הבתאי ידוע והלא סטציונרי של דיבור איננו מעודד נקיטת גישה מסווג זה. עבודות שנעשו בתחום זה על אף קשיים אלו מתוארות במאמרו של [6] Gibson.

גרסת גוטפה של קידוד דינריציאלי היא NFC (Noise Feedback Coding). NFC מתקבל ממקור הרוחבה של סכמה שcolaה-L-DPCM על פיו את הכנישה עובר סינון מלביון (בדומה ל-PCM*), ושארית החזוי עוברת כמו תוק הייזון חזזר של שגיאת הcamoi המותנית היותרן של מבנה כזה הוא שהוא מאפשר שליטה בספקטרום שגיאת השחזר ומיסור (masking) מיטאים של אותן הרצוי על פני רעש הקוונטיזציה (למאור מפורט יותר של NFC ראה [5], [6] ו-[7]).

מקדדים אדפטיביים בעלי זכרון

על מנת לנצל את הזכרון של אותן הכנישה גם כאשר הוא אינו סטציונרי דרישות גרסאות אדפטיביות של מקדי חזוי המטוגלות להתמודד עם שנויות בעוצמת אותן (כפי שהסביר נעשה ב-APCM), ולעתים גם עם שנויות באופי הטפקטרלי שלו. מערכת אחת מטוג זה היא ADPCM (Adaptive PCM). במונח "Adaptive" בקשר זה קיים חוסר הילוך (APCM ב-PCM) עקביות מסוימות בספרות - לעומת אדפטיביות מיוhostת לקוונטיזר בלבד (כמו ב-APCM) ולעתים גם חזאי. האדפטציה של הקוונטיזר מטבחה, בדומה לזה ב-APCM, תוק התאמת שוטפה של גודל עצם להתנהגות הדינמית של אותן (ראו [6], [7]). אם בנותם מבעניהם אדפטציה חזאי אפשר לשפר את הביצועים בכ-4dB וזו גם הגדלת סיבוכיות ל-2 או 3, תורמת משמעותית להגדלת חס אוט לרוש: המהיר הוא כموון הגדלת סיבוכיות המערכת. האדפטציה של חזאי, בדומה לזה של הקוונטיזר, יכולה להתבצע לפי מדיניות AQF או AQB. במקרה הראשון מקובל לשדר מקדמי חזוי קבועים למוקטעהן כמוידע גלובינה AQF או AQB. במקרה השני מקובל לשדר מקדי חזוי קבועים (batch) של דוגמים מאות המקור אחת לפרק זמן מסוימים לאחר שערכם על פני אכווה (batch) של דוגמים מאות המקור בטכניקות חזוי לינארית מוכראות [9]. ב-AQB לעומת זאת, מקדי חזוי משתנים מרגע לרגע באופן שוטף ומתקבלים מתוך המידע הנקלט בשיטות רקורסיביות כגון LMS, מסנן קלמן וכיוצא (ראו [6]). האמור לעיל לגבי המגרעות והמעלות של AQB ו-AQF תקף גם כאן.

מציעי גישת האנתרופופיה המירביה כתחליף ל-DPCM ([2],[3]) לא חיעו גרטה אדפטיבית שתקביל ל-ADPCM. במתגרת חبور זה מוצעת גרטה כזו.

הගרטאות האדפטיביות של דלאט-מודולטור, ADM (Adaptive Delta Modulator) ו-CVSD (Continuously Variable Slope Delta modulation) נבדלות זו מזו בלוגיקה של עדכון גודל הצעד על פי מלות הקוד המשודרות. זו האחיזנה משווית בחברת זה לארטה האדפטיבית המוצעת של אנטרופיה מירבית בסביבה אחת לדגס.

במערכות מתקדמות יותר כמו APC (Adaptive Predictive Coding) ו-PPDPCM (Pitch Predictive DPCM) מוצלים גם את ה"יתירות לטוח ארוך" (long term redundancy) של דיבור הנובעת מחזוריותו המוקנית. לצורך פעולה יעילה של מגננון חזוי מתאים דרוש כמובן אלגוריתם מהימן לשערור מדר יסודי (pitch detection) המהווה קושי שלעצמם בעבודאות דיבור, וכורע איפוא, בסיבוכיות לא מבוטלת. לפיכך, בנסיבות חبور זה לא נבחנות מערכות מסווג זה אלא נעשה נסילון להעריך את מערכת האנתרופיה המירבית במתכונת הקבועה ובגרסתה האדפטטיבית בהשוואה ל-DPCM ו-ADPCM (short term redundancy). מרכיבות אלו מוצלות כדי רק את ה"יתירות בזמן קצר" (short term redundancy) של דיבור אך הן יחסית פשוטות למשוח וחסינות יותר לרעש ערוץ.

מטרות המוחלט 1.3

מחקר זה מכובן לבחון ולהעריך את הפוטנציאלי הטמון בגישה האנתרופופיה המירבית בהשוואה ל-DPCM מלביטלים שונים של ביצועים: יחסאות לרוש, איקות השחזור, רגשות לסתיות באופי הסטטיסטי ומידת החסינות להפרעות בערז. בנוסף, מוצעות גרסאות אדפטיביות של קידוד באנתרופופיה מירבית והן נבחנות בהשוואה ל-ADPCM-CVSD לגביהם.

תורת האינפורמציה קובעת כי אנטרופיה מירבית היא תוצאה של דחיסה אופטימלית כאשר הקידוד מתבצע בבלוקים שאורכם שואף לאינסוף. ל מערכת הנדונה, לעומת זאת, יש זכרון של צעד זמן אחד בלבד (חזאי מטר רשות). למרות זאת, ניתן עניין לחזור ולקבל הערכה בדבר הביצועים שניתן לצפות ממנה, הודות למספר מניעים:

1. פשטות המימוש של המערכת.
 2. המשרוש בחזאי לא לינאריאי.
 3. קבלת אפיון של מידת הקשר שבין אנטנורופיה לבין שגיאת השazor.

שאלות אחרות שתחבנה כוגנות למידת השיפור בפיתוח המערכת כאשר מעלים את קצב השידור
שלוח (טיביות אחת לשתי סיביות לדגם), להשפעת קצב הדגימה על הביצועים, ולתפרקיד
הأدפתציה של האזאי (בנוסף לאדפתציה של הגבר) בגרסת האדפטיבית.

1.4 מבנה חברור

בפרק 2 מובא תחילה דיוון בעקרון האנתרופופית המירביה כבסיס לגישה הננקטת במחקר זה. לאחר שמתוארת המערכת שהוצעה ע"י Angel ו-Daigle [2][3], נערך פיתוח עיוני של האופיניות האתאיימיט במשדר ובמקלט עבור כמו 2-2 רמות ו-4 רמות. לבסוף, מטוכמות מספר תכונות לליליות של המערכת ומצוינות מעלהותיה ומגרעוטיה המרכזיות.

בפרק 3 נבחנים ביצועי המערכת (בשילובת ל-DPCM) לגבי מקורות טציזונריאים, מוקדשת תשומת לב למידת רגישותה לטטיות באפיגנים אסתטיטיים של המקור (סוג הפלוג והקורלציה בין הדגימות הטמוכרים), ונעשה נסיוון להסבירם. עבור פלוגים לא אוטיים נבדקות אפשרויות של companding לצורך הסבתם לפלוג גaussi.

גורסאות אדפטיביות של המערכת מוצעות ונבחנות בפרק 4 לגבי אורתות דיבור. יחס האות לרעש ואיכות המשוחרר מושווים ל-ADPCM ול-CVSD, וקיימת התיחסות לשיבוכוthon היחסית של המערכות המוצעות.

פרק 5 מוקדש להערכת מידת רגישותן של המערכות להפרעות בערזץ התקשורתי. תחילה מובא בקצרה דיוון איכותי בדבר מידת ההשפעה של הפרעה (בודדת) בערזץ על אותן המשוחרר במקלט EM בהשוואה לזה במערכת בעלת חזאי לינארי (כמו DPCM). בהמשך מטוכמות תוצאות סימולציה עבור אותן רגשות טציזונריאים. לגבי דיבור מתוארת גם השפעת עצמת ההפרעות על איכות אותן המשוחרר וubahnoto.

בסכום מטוכמות תוצאות העבודה והמקנות המרכזיות ממנה. כמו כן מועלות מספר הציעות לשיפור ולמחקר נוספים.

פרק 2 - רקע עיוני לדחיסה באנטropיה מירבית

בפרק זה יובא נושא עיוני מפורט של הדחיסה באנטרופיה מירבית. תחילת נדוע בעקרון האנטרופיה המירבית כתנאי הכרחי לדחיסה אופטימלית, בהמשך תוצג המערכת המוצעת ב-[2], ו-[3], כמו כן יפורט הפתוח של אופני הדחיסה והשזרור (אשר לא פורסם במקורות הניל) ולבסוף יסוכמו מספר תכונות עקרוניות של המערכת ויצוינו מעלותיה ופגיעותיה העיקריות.

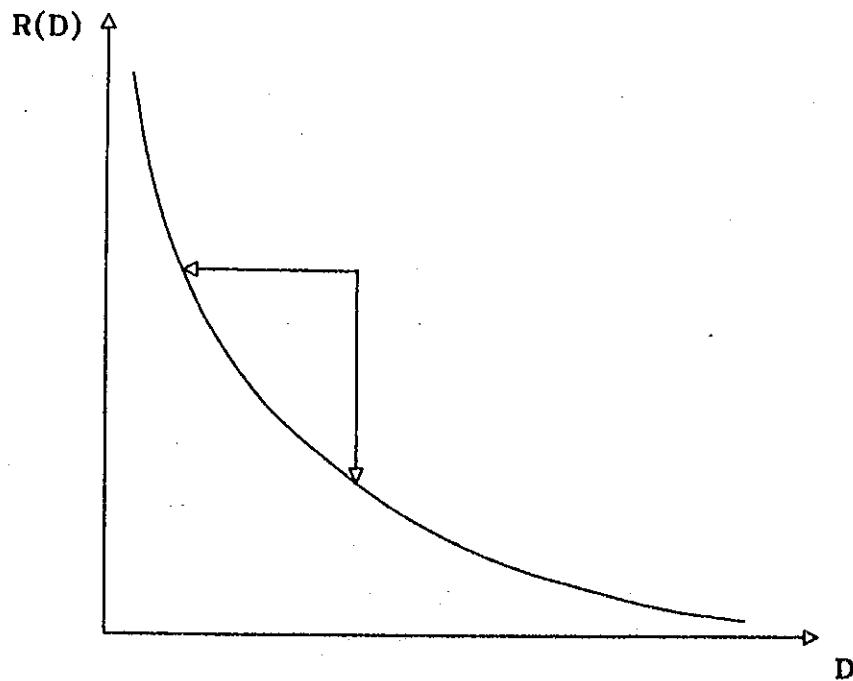
2.1 התפיסה של אנטרופיה מירבית

בסעיף 1.2 תוארה מערכת DPCM, המתוכנת ע"פ עקרון השגיאת הריבועית המוצעת המינימלית, והושם דגש על המגבילות בקבלה מקדמי חזוי מתאימים בקצב שידור נומכית (סיבית אחת או שתיים לדגם) עבורות הכינוי אינו עדין די הצורך.

כפי שצוין במכוא, הוצע בשנת 1983 ע"י Angel ו-Daigle ([2],[3]) קרייטריוון אופטימיזציה תחליפי – אנטרופיה מירבית של המידע המשודר. השkol לבחרת קרייטריוון זה, להערכת המחבר, היה שאנטרופיה מירבית היא תכונה נלווה של מערכת דחיסה שהיא אופטימלית במובן עותת מינימלי בשזרור. ההסבר לכך נובע מהסקול הבא: נניח שלרטותינו מערכת דחיסה אופטימלית כניל. אילו האנטרופיה של המידע המשודר ממנה לא תהיה מירבית, פרוש הדבר שנייה יהיה להמשיך ולדוחסו לא עותת בסוף (entropy coding) וע"י כך לרדת בקצב ומכך שהקצב המקורי אינו מינימלי עבר הערות המוצע הנ透. לחילופין, פרשו של דבר שהעותות אינם מינימלי עבר הקצב הנ透, דבר העומד בסתירה להנחה דלעיל. בציור 2.1 מתוארת באופן איקוני פונקציית קצב עותת טיפוסית. כאשר 'בקודת העבודה' קצב-עות נמצאת מעל העוקם שבעזר, אז היא אינה אופטימלית ונניתן לשפה ע"י הקטנת הקצב ו/או התקנת העותת.

הערות:

1. אלגוריתם דחיסה המבוסס על מקסימיזציה של האנטרופיה מחייב הנחות כלשהן או דומות תכונות הסטטיסטיות של המקור.
2. מכונת האנטרופיה המירבית מהוות תנאי הכרחי בלבד לעותת מינימלי ובמפורש אי-בנה תנאי מטפיך.



ציור מס' 2 - פונקציה קצב-עוות. שיפור אלגוריתם דחיטה ע"י הורדת הקצב או הקטנת העוות.

Fig. 2.1 - Rate-distortion function. Improving the data compression algorithm by reducing the rate or by decreasing the distortion.

דוגמא פשוטה לדחיטה מ-ה-אופטימלית באנטרופיה מירבית היא הכמה הטקלרי של מקור גausי חסר זכרון עם ממוצע אפס ושונות² בסביבית אחת לדגם. כטוצה מכמוני כנייל מתקבל סדרה חרטת זכרון עם אייב ביניاري אשר אותיותיו שוות הסתברות מטעמי סימטריה, ועל-כן האנטרופיה שלה תהיה מירבית (ושווה ליחידה). שנות השגיאת המינימלית האפשרית בשזרור המקור מסדרה זו נמנעה ע"י (ראה גם [7]):

$$D = 0.363 \sigma^2$$

לעומת זאת, שנות השגיאת המינימלית האפשרית עברו אלגוריתם דחיטה אופטימלי בתונה ע"י פונקציית עוות-קצב של מקור גausי חסר זכרון ובעור במקרה $R = 1$ מקבל: ([7])

$$D = \frac{\sigma^2}{2^{2R}} = 0.25 \sigma^2$$

הסיבה להבדל נובעת מכך שהאלגוריתם האופטימלי מתקבל ע"י כמו רב-מימדי (vector quantization).

אף-על-פי-כן קיים קשר בין האנטרופיה לבין העות, וזאת משום שהאנטרופיה מבטאת למעשה באיזו מידת הנצול של כל רמות ה כמו הוא ייעיל. מטעוריות, איפוא, מספר שאלות עקרוניות כגון: מה הפוטנציאל הטמון בשיטות דחיסת המושתנות על אנטרופיה מירבית? האם קרייטריוון האנטרופיה המירבית עדיף, בוגבליות הקיימות בקביבים נמכבים, על פניו קרייטריוון השגיאה הריבועית הממוצעת המינימלית? מהו הקשר שבין שני קרייטריווני אופטימיזציה אלו? וכו'. עם שאלות כגון אלה בנסה להתמודד בהמשך.

2.2 תאור המערך

בטעיף זה נציג את המערכת המוצעת ע"י Daigle ו-Angel, שעליה מוחבסט חיבור זה. התפישה המנחה את עבודתם היא כאמור מקסימיזציה של אנטרופית המידע המשודר. על-מנת שניתן יהיה למצוא פתרון אנליטי נוח לבנית תכנון המערכת, המחברים מגבלים את עצם למקסימיזציה של האנטרופיה המותנית מסדר ראשון בלבד ומגנחים לשטך שהמקורה הוא גausי עם פובקציית קווריאנט ידועה (למשל, תחליך גaus-מרקוב). בציור 2.2 תוצג מזוארית המשדר והמקלט עבור קצב של סיבית אחת לדגם. בתמוך (סעיף 2.3.2) תוצג הרחבה עבור קצב של שתי טיביות לדגם.

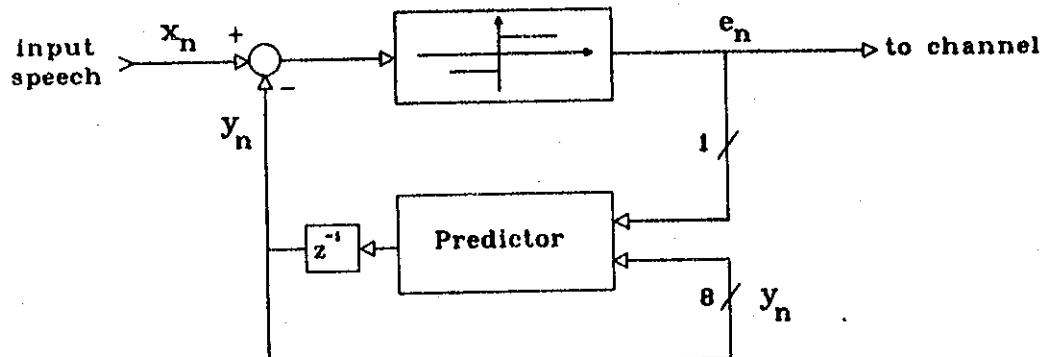
השדר, המתואר בחלק (a) הוא בעל מבנה עקרוני דומה לזה של משדר DPCM בעל חזאי (predictor) מסדר ראשון. אך בנגדוד לחזאי DPCM, המתוכנן למזרע את אנרגיית השארית, החזאי המוצע (משדר ראשון אף הוא) מתוכנן באופן כזה שהסתברויות המופיע של אפס או אחד ביציאת המשדר בצד הזמן הבא (e_{n+1}) תהיה שווה בהנתן המידע העומד לרשות החזאי מצד הזמן הנוכחי, קרי: $e_n = 0$. לעומת זאת, פונקציית החזוי, שננסנה כ-
 $M(y_n, e_n) = 1$, נדרש לקיים את התנאי:

$$P_r\{e_{n+1} = 1 | e_n, y_n\} = P_r\{e_{n+1} = -1 | e_n, y_n\} = \frac{1}{2} \quad (2.1)$$

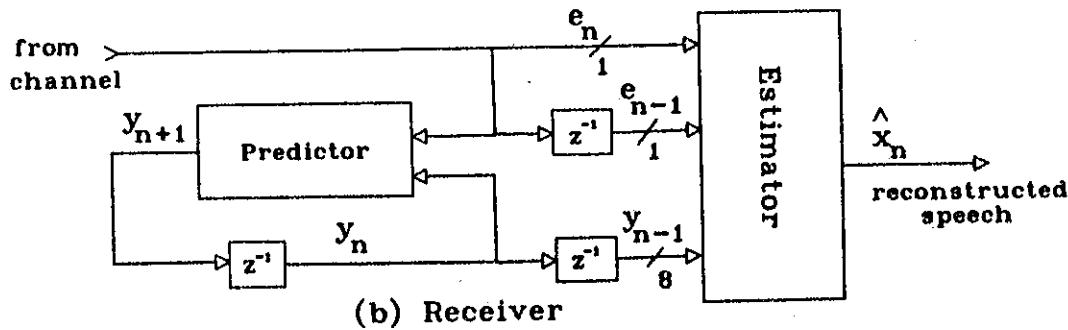
(טעןנו נוחיות נגידר את האיב הבינארי של המשדר מעל המספרים 1 ו-1).

באופן כזה מובטו מקטימים לאנטרופיה המותנית מסדר ראשון (כלומר $P_r\{e_{n+1} | e_n, y_n\} = 1$) במידע המשדר וכתוצאה לכך גם אי תלות סטטיסטי בין סיביותיו הסמכות.

(LAGBI AIIB BINARI: אם מתקיים $P_r\{e_{n+1} | e_n, y_n\} = 1$ אז גם $P_r\{e_{n+1} | e_n\} = 1$).



(a) Transmitter



ציור מס' 2.2 - סכמת הדרישה המוצעת ע"י Angel ו-Daigle.
Fig. 2.2 - The compression scheme proposed by Angel and Daigle.

המכלול המתואר בחלק (ב) של ציור 2.2, מכיל חזאי דומה לצורך שחזור y ובנוסף לכך משערך של אותן הרכיבות הנעזר במידע הזמן הנוכחי מהצעדי הנוכחי או דותה תחווות ומעבר (קרי: $y_{n-1}, e_{n-1}, y_n, e_n$):

$$\hat{x}_n = E\{x_n | e_n, e_{n-1}, y_n, y_{n-1}\} \quad (2.2)$$

מאחר ש- y מוגדר חד-משמעות ע"י y_{n-1}, e_{n-1} ו- y_n מתוך פונקציה חזזוי (\cdot, \cdot, M , אזי ניתן להשミニ את y ולבסוף את (2.2) באורפן הבא:

$$\hat{x}_n = E\{x_n | e_n, e_{n-1}, y_{n-1}\} \quad (2.3)$$

ועל-כן כל האינפורמציה הדרושה למשערך היא השלישייה $\{e_{n-1}, e_n, e_{n+1}\}$ כמפורט בצדורה.

אחר שפונקציית החזוי והשערוך המתקבלות הן בלתי מפורשות (כפי שנראה להלן), המחברים מוציאים להשען באופן נומרי "אות ותמים" ולאחסן בטבלאות קבועות (lookup tables). בדרךazzo המערכת לא נדרש לבצע כל פעולות חישוב אלא גישות לזכרו בלבד ולכן היא מהירה מאוד. מהירות זו חיונית במיוחד כאשר נדרש לדוחות אותן וידאו אשר נדגמים בקצב גובהם מאוד. לעומת זאת, שימוש דחיטה של אותן דיבור, אשר נדגמים בקצב נמוך יחסית, אין למעשה כל הצדקה להשתמש בטבלאות קבועות אשר אינן מאפשרות כל גמישות נוכחה שנויות באופי הטעיטיסטי של המדור (أدפטציה). לשם שימוש בטבלאות קבועות המחברים מוציאים להזנת את הסטרוות החיריגת של המשטנה האקראי הגאולוגי מתחום (ס-4ס) (כאשר ס מציין את טויחת התקן) ולהקצות 8 סיביות בתחום זה בלבד. תפוקת החזאי y תיזג, איפוא, עזרת 8 סיביות (ראו ציור 2.2) וע"כ טבלת החזוי תהיה בגודל 512×8 וטבלת השערוך שבמקלט בגודל 1024×8 , ככלומר דרישות הזכרו הן צנויות למדי ולמערכת 256 מצבים בלבד.

הערה: במחיר שלACPML גודל טבלת השערוך ניתן לשפר את טיב השערוך בעזרת מידע מעוד הזמן (2-ה) חיננו: $E\{x_n | e_{n-2}, e_{n-1}, e_n\} = x_n$.

המחברים מארים את ביצועי האלגוריתם עבור דחיטת דיבור [2] ותמונות [3] בהשוואה לגרסה (לא אדפטיבית) של DPCM עבור קצבים של טיבית אחת ושתי סיביות לדגם. לטענותם, אלגוריתם האנטרופיה המירבית (להלן יקרא בקיצור ME) מגלה כושר עקיבהמצוין לשינויים מהירים בגובה אותן בעוד ש- DPCM לויה בתופעה של slope overload שמקורה במוגביה על הקצב המקסימלי של השטנות באוט המשוחזר. לעומת זאת, בקטעים חלקיים יחסית הביצועים של DPCM עדיפים (מאחר שהוא מזעך את השגיאה הריבועית). מנקודת ראות של איקות הדיבור, השזרור של ME שומר לטענות על הייעורו של הצליל אך מושך מעט רעש רקע בהשוואה ל- DPCM. המחברים לא הצינו גרטאות אופטיביות של האלגוריתם.

2.3 אופני החזוי והשערוך

בטעיף זה נתאר את פתו אופני החזוי והשערוך של אלגוריתם האנטרופיה המירבית שמוarrow בטעיף (2.2) וכאמור, לא פורסם במקורות [2] ו-[3]. בטעיף 2.3.1 יוקדש הדרינו עבור קצב של טיבית אחת לדגם. הביטויים המתקבלים הן עבור החזאי והן עבור המשערך הם בלתי מפורשים וחישובם מחייב, ככל הנראה, שימוש בטכניקות נומריות.

טעיף 2.3.2 תובא הרחבה מתאימה עבור קצב של שתי סיביות לדגם.

2.3.1 מחזאי ותמשער עבור סיבית אחת לדגם (8 kbit/sec)

יהי $\{x_n\}$ תהליך אקראי גaussiano סטציונרי בעל תוחלת אפס. נסמן:

$$\sigma^2 \triangleq E(x_n^2) \quad (2.4.a)$$

$$\rho \triangleq E(x_n x_{n+1}) / \sigma^2 \quad (2.4.b)$$

ונניח שפרמטרים אלו ידועים.

פונקציה צפיפות התסתברות המשותפת לשני דגמים סמוכים נתונה ע"י:

$$g(x_n, x_{n+1}) = (2\pi\sigma^2 \sqrt{1-\rho^2})^{-1} \exp\left\{-\frac{x_n^2 - 2\rho x_n x_{n+1} + x_{n+1}^2}{2\sigma^2(1-\rho^2)}\right\} \quad (2.5)$$

חזרה

תקיד זהה, כאמור בסעיף (2.2), הוא להבטים שהקונטינזר יפיק $e_{n+1} = +1$ או -1 בהסתברויות שווות. לשטך, יש לדרש ש- $y_{n+1} = M(y_n, e_n)$ יהיה החזיוון (median) של צפיפות התסתברות המותנית (x_n) עבור המקרה $+1$, $e_n = +1$ או החזיוון של הצפיפות (y_n) עבור המקרה -1 , $e_n = -1$.

עבור המקרה 1, לפי Bayes:

$$g(x_{n+1} | x_n \geq y_n) = \frac{g(x_{n+1}, x_n \geq y_n)}{P(x_n \geq y_n)} \quad (2.6)$$

המכנה של (2.6) נתון ע"י $\frac{y_n}{\sigma}$ כאשר:

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$$

LAGBI המונה בביטוי (2.6):

$$g(x_{n+1}, x_n \geq y_n) = \int_{y_n}^\infty g(x_{n+1}, \theta) d\theta = \int_{y_n}^\infty g(x_{n+1}) g(\theta | x_{n+1}) d\theta = \quad (2.7)$$

$$= g(x_{n+1}) \int_{y_n}^\infty g(\theta | x_{n+1}) d\theta = g(x_{n+1}) \int_{y_n}^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\theta - \rho x_{n+1})^2}{2\sigma^2(1-\rho^2)}\right\} d\theta$$

הערה: הסימנו g מציין בכל מקרה חוק הסתברות גאוטי אם לפונקציה (\cdot) g מציין ארגומנט אחד, פרשו צפיפות הסתברות נורמלית חד ממדית.

מ-(2.7) נובע:

$$g(x_{n+1}, x_n \geq y_n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x_{n+1}^2/2\sigma^2} Q\left(\frac{y_n - \rho x_{n+1}}{\sigma\sqrt{1-\rho^2}}\right) \quad (2.8)$$

ועיל לפ (2.6):

$$g(x_{n+1} | x_n \geq y_n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x_{n+1}^2/2\sigma^2} \frac{Q\left(\frac{y_n - \rho x_{n+1}}{\sigma\sqrt{1-\rho^2}}\right)}{Q(y_n/\sigma)} \quad (2.9)$$

עד הזרוי עבור צעד הזמן הבא y_{n+1} מקיים איפוא:

$$\begin{aligned} P_r\{e_{n+1}=1 | e_n=1, y_n\} &= P_r\{x_{n+1} \geq y_{n+1} | x_n \geq y_n\} = \\ &= [\sqrt{2\pi}\sigma Q\left(\frac{y_n}{\sigma}\right)]^{-1} \int_{y_{n+1}}^{\infty} e^{-\theta^2/2\sigma^2} Q\left(\frac{y_n - \rho\theta}{\sigma\sqrt{1-\rho^2}}\right) d\theta = \frac{1}{2} \end{aligned} \quad (2.10)$$

נתקבלה, אם כן, משואה אשר בה σ , ρ ו- y הם פרמטרים והגעלה $-y_{n+1}$ היא גבול של אינטגרל אשר לא נמצא עבورو ביטוי מפורש. לפיכך, האפשרות היחידה שמסתמנת לפטור משווה זו היא בטכניות נומריות (גיזטון-רפסון וכיד').

נציין את המקרה $1 = \rho$ עבורי הפונקציה Q , המופיעה באינטגרנד המשווה, הופכת לפונקציה מדרגת:

$$\lim_{\rho \rightarrow 1} Q\left(\frac{y_n - \rho\theta}{\sigma\sqrt{1-\rho^2}}\right) = \begin{cases} 0 & , \theta < y_n \\ 0.5 & , \theta = y_n \\ 1 & , \theta > y_n \end{cases} \quad (2.11)$$

כאשר $1 = \rho$ מקיימים גם $y_{n+1} > x_n$ כיון שבמקרה זה $x_{n+1} = x_n$ בהתברות 1 י- y_{n+1} מקבל למשועה-משמעות של חיזיו הפלוג הסגולי $(g(x_n | x_n > y_n))$, אשר בברור חייב להיות גדול מ- y . עיל אם $1 = \rho$ ניתן לרשום את (2.10) :

$$\left[\sqrt{2\pi} \sigma Q\left(\frac{y_n}{\sigma}\right) \right]^{-1} \int_{y_n}^{y_{n+1}} e^{-\theta^2/2\sigma^2} d\theta = \frac{1}{2} \quad (2.12)$$

מכאן

$$Q\left(\frac{y_n}{\sigma}\right) - Q\left(\frac{y_{n+1}}{\sigma}\right) = \frac{1}{2} Q\left(\frac{y_n}{\sigma}\right)$$

או

$$y_{n+1} = \sigma Q^{-1} \left[\frac{1}{2} Q\left(\frac{y_n}{\sigma}\right) \right] \quad (2.13)$$

עבור המקרה $1 = n$ ניתן לקבל משווה דומה. לחילופין, אפשר לקבל פתרון עבור מקרה זה מתוך הביטוי (2.10) ע"י החלפת סימנים של המשתנים המתאימים. פעולה זו מותרת הודות לסימטריה של הפילוג הגאוסי.

$$g(x_{n+1} | x_n < y_n) = g(x_{n+1} | -x_n > -y_n) = g(x_{n+1} | x_n > -y_n) \quad (2.14)$$

כלומר, ע"מ לקבל את ערך החזוי x_{n+1} עבור $1 = n$, הפעולה תדרישה היא להפוך את הסימן של y_n , לפטור את משווה (2.10) ולבסוף להפוך את טימן הפתרון x_{n+1} . לכן, פונקציית החזוי $(\cdot, \cdot) M$ ניתנת להציג כאותן הבאות:

$$M(y_n, e_n) = e_n P(e_n, y_n) \quad (2.15)$$

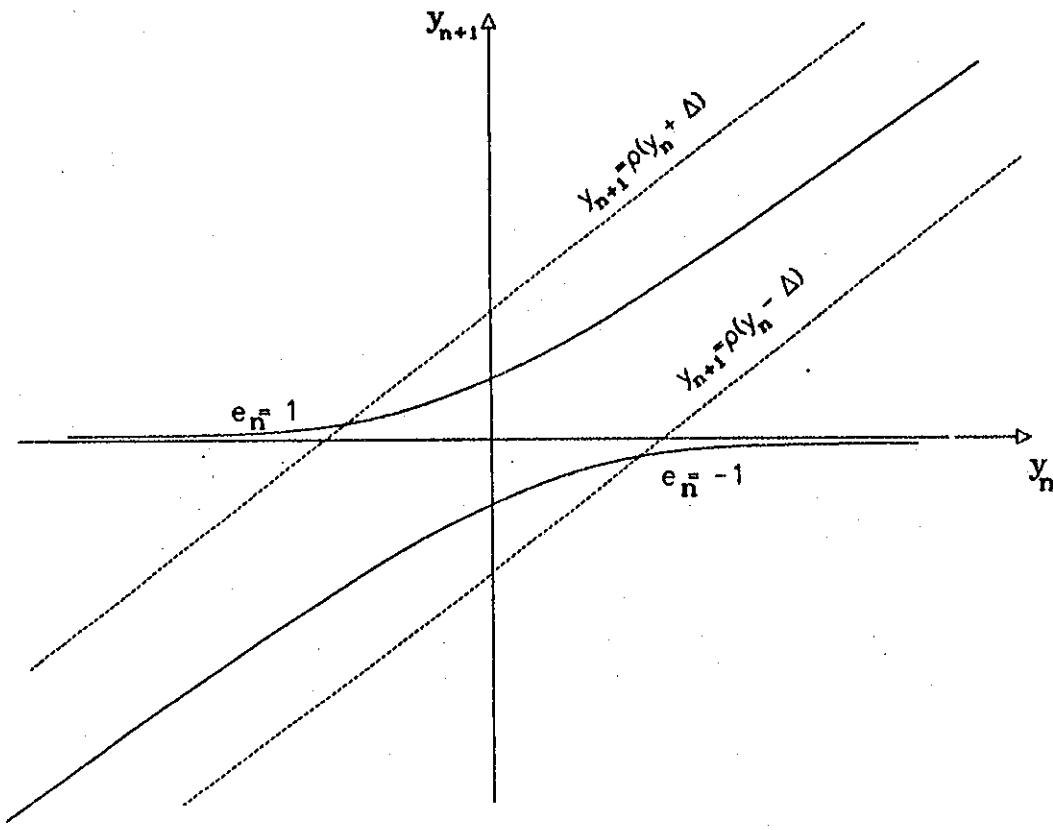
כאשר $(\cdot, \cdot) P$ היא פונקציה של ארגומנט אחד בלבד.

למשל, עבור המקרה $1 = n$ נוכל לרשום את (2.13) בצורה הבאה:

$$y_{n+1} = e_n \sigma Q^{-1} \left[\frac{1}{2} Q\left(\frac{y_n}{\sigma}\right) \cdot \frac{y_n}{\sigma} \right] \quad (2.16)$$

לפתרון עבור $1 = n$ אין ערך מעשי כשלעצמו מאחר שבסייך $1 < n$, אולם בהמשך הדברים נראה שנitin להעזר בביטוי זה ככלי לקירוב עבור מקרים בהם $1 < n$.

נדגיש ש奧פִילִינִי החזוי המתקבלים אינם לינאריים בניגוד לחזאים המקובלים בדיביך. בציור (2.3) מוצגים אופִילִינִי החזוי של אנטרופיה מירבית ושל דלתא מודולטור עבור $\rho = 0.9$.



ציור מס' 2.3 - אופייני חזוי של ME ושל LDM עבור $\rho = 0.9$:
 קו רציף - חזאי אנטרופיה מירבית.
 קו מרוטק - חזאי LDM

Fig. 2.3 - Prediction characteristics of ME and LDM for
 $\rho = 0.9$:
 solid line - Maximum Entropy predictor.
 dashed line- LDM predictor.

המשערך

תפקיד המשערך שבמקלט הוא לשחזר את אותה הרכבה עיינית שפור של תפוקת החזאי. השפור מקבל בזכות העובדה שהמשערך, בוגוד לחזאי, מנצל גם אינפורמציה אודות ההוויה (e_n) ולא רק בעבר. המשערך הוא פונקציה של שלושה משתנים (ראה (2.3)). נסמן פונקציה זו ב- G ונגדיר:

$$\hat{x}_n = E\{x_n | e_n, e_{n-1}, y_{n-1}\} \stackrel{\Delta}{=} G(e_n, e_{n-1}, y_{n-1}) \quad (2.17)$$

התווחלת המותגנית אנויל מתקבלת ע"י איינטגרציה בתחום המתאים עבור x ו- y .
למשל עבור $1 = e_n = e_{n-1} = y_{n-1}$ נוכל להציג $p(y_{n-1})$ ולקבל:

$$\hat{x}_n = G(1, 1, y_{n-1}) = \frac{\int_{-\infty}^{\infty} x dx \int_{y_{n-1}}^{\infty} g(x, y) dy}{\int_{-\infty}^{\infty} dx \int_{y_{n-1}}^{\infty} g(x, y) dy} \quad (2.18)$$

עבור שלושת המקרים הנוגזרים מתבלים בטווים דומים ע"י החלפת תחומי איינטגרציה בהתאם:

$$G(-1, 1, y_{n-1}) = \frac{\int_{-\infty}^{p(y_{n-1})} x dx \int_{y_{n-1}}^{\infty} g(x, y) dy}{\int_{-\infty}^{p(y_{n-1})} dx \int_{y_{n-1}}^{\infty} g(x, y) dy} \quad (2.19)$$

במקרים עבורם $-1 = e_n = -p(-y_{n-1})$ מתקיים

$$G(1, -1, y_{n-1}) = \frac{-p(-y_{n-1}) \int_{-\infty}^{\infty} x dx \int_{-\infty}^{y_{n-1}} g(x, y) dy}{-p(-y_{n-1}) \int_{-\infty}^{\infty} dx \int_{-\infty}^{y_{n-1}} g(x, y) dy} \quad (2.20)$$

$$G(-1, -1, y_{n-1}) = \frac{\int_{-\infty}^{-P(-y_{n-1})} x dx \int_{-\infty}^{y_{n-1}} g(x, y) dy}{\int_{-\infty}^{-P(y_{n-1})} dx \int_{-\infty}^{y_{n-1}} g(x, y) dy} \quad (2.21)$$

קל להראות (מטעמי סימטריה של הפלוג הגאומטרי) שמקיימת:

$$G(e_n, e_{n-1}, y_{n-1}) = -G(-e_n, -e_{n-1}, -y_{n-1}) \quad (2.22)$$

לצורך חישוב המונח של משווהה (2.18) נגידו:

$$R(t, s) \stackrel{\Delta}{=} \int_s^{\infty} x dx \int_t^{\infty} g(x, y) dy \quad (2.23)$$

ניתן להראות (ראה בסוף ג') כי מקיימת:

$$R(t, s) = \frac{\sigma}{\sqrt{2\pi}} \left[e^{-s^2/2\sigma^2} Q\left(\frac{t - \rho s}{\sigma\sqrt{1-\rho^2}}\right) + \rho e^{-t^2/2\sigma^2} Q\left(\frac{s - \rho t}{\sigma\sqrt{1-\rho^2}}\right) \right] \quad (2.24)$$

לצורך חישוב המכנה נקבעו בפיטוי:

$$\int_s^{\infty} dx \int_t^{\infty} g(x, y) dy = P_r\{x_n > s, x_{n-1} > t\} = P_r\{x_n > s | x_{n-1} > t\} \cdot P_r\{x_{n-1} > t\}$$

אם (אזי עיף תחגדרת):

$$P_r\{x_n > P(t) | x_{n-1} > t\} = \frac{1}{2}$$

מאותר ש- $P(x_{n-1} > t) = Q\left(\frac{t}{\sigma}\right)$ קיבל עבור המכנה:

$$\int_{P(t)}^{\infty} dx \int_t^{\infty} g(x, y) dy = \frac{1}{2} Q\left(\frac{t}{\sigma}\right) \quad (2.25)$$

ע"י הצבת (2.23) ו-(2.25) ב-(2.18) נקבל:

$$G(1,1,t) = \frac{2R(t, P(t))}{Q\left(\frac{t}{\sigma}\right)} \quad (2.26)$$

לABI משווה (2.19) המונה נתונה ע"י:

$$\int_{-\infty}^{P(t)} x dx \int_t^{\infty} g(x,y) dy = \int_{-\infty}^{\infty} x dx \int_t^{\infty} g(x,y) dy - \int_{P(t)}^{\infty} x dx \int_t^{\infty} g(x,y) dy = \quad (2.27)$$

$$= R(t, -\infty) - R(t, P(t)) = \frac{\rho\sigma}{\sqrt{2\pi}} e^{-t^2/2\sigma^2} - R(t, P(t))$$

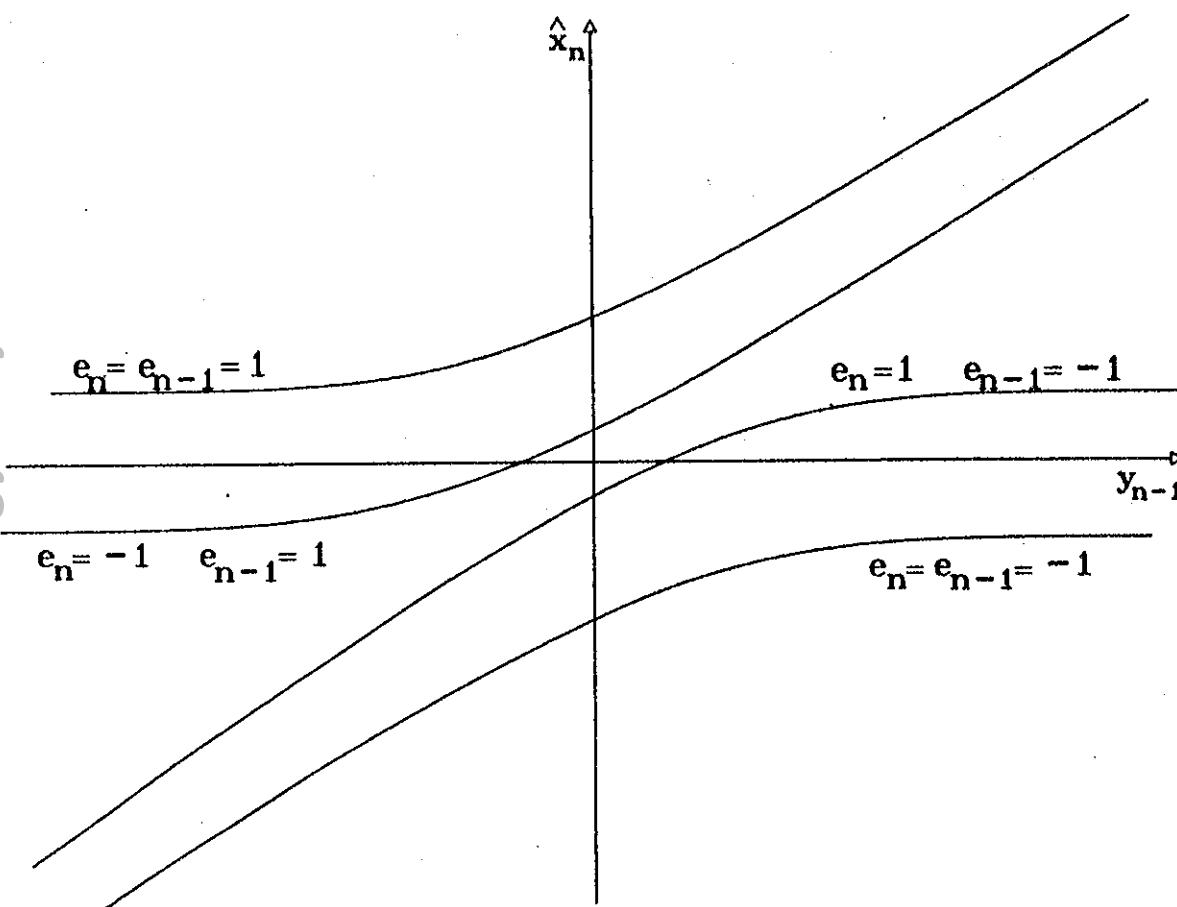
המכנה נוטר כמו בביטוי (2.25) ומקבלים:

$$G(1,-1,t) = \frac{2}{Q\left(\frac{t}{\sigma}\right)} \left[\frac{\rho\sigma}{\sqrt{2\pi}} e^{-t^2/2\sigma^2} - R(t, P(t)) \right] \quad (2.28)$$

את שני המקרים הנדרשים (2.20) ו-(2.21) ניתן לקבל מזור (2.26), (2.28) ומשווא בזיהות (2.22). שנקבלה משקלוי סימטריה.

נתקבלו איפוא גושאות "מפורשות" (עד כדי תפונקציות Q ו-P שאוثرו ניתן לקרב) עבור המשער. חישובות תבוא לידי ביטוי בפרק 4 בו יוצגו גרטאות אדפטיביות של המערכת.

בציור (2.4) מזורים אופייני המשער עבור $0.9 = \rho$ בארכעת המקרים.



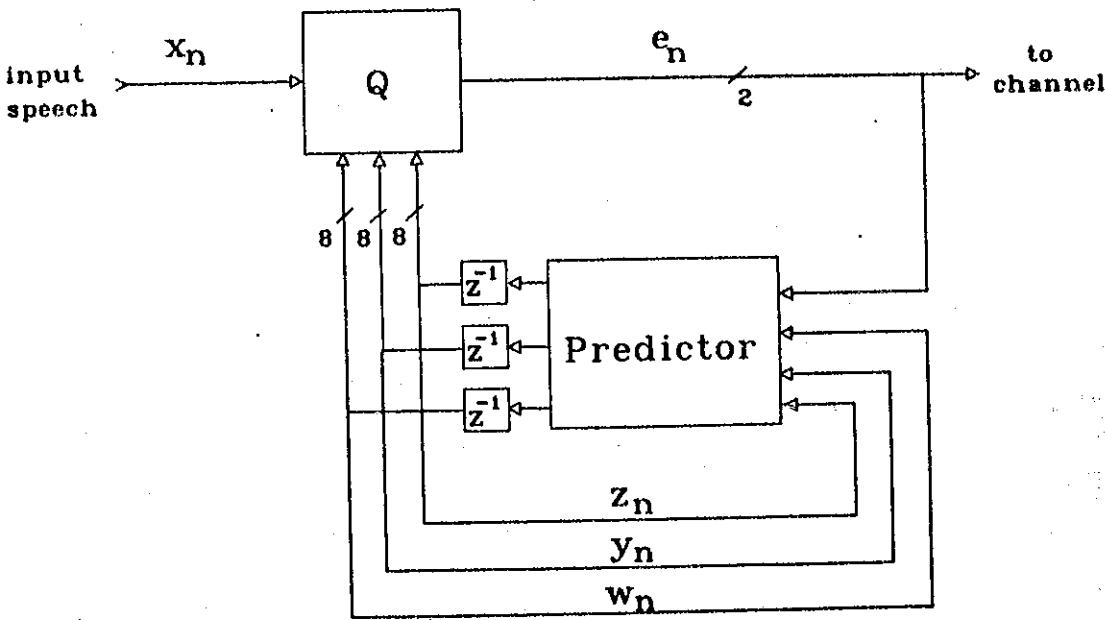
ציר מס' 2.4 - אופייני השערות עבור $\rho = 0.9$
Fig. 2.4 - Estimation characteristics for $\rho = 0.9$

2.3.2 חוץאי והשדר עבור שתי סיביות לדגם (16 kbit/sec)

בטעיף זה נורויב את התוצאות שנטקלו בטעיף 2.3.1 עבור קצב של שתי סיביות לדגם. נראה, שתרחבה ישירה היא מטובכת לאין שoor ומתוער צורך להגביל את מספר המცבים. לשם כך תוצענה שתי אפשרויות שונות.

חצאי

אנטרופיה מירבית בקצב של שתי סיביות לדגם פירושה ארבע רמות כמו שות הטעברות. לשם קבלתו, על חוץאי לספק לקוונטייזר בכל רגע ורגע. לשורה ערכית המחליקת את הישר המשי לאביבה תחומיות שווי הטעברות. הרחכת פונקציית החזוי למקורה הניל תור. שמירה על אסטרטגיית שטוארת בטעיף 2.3.1 פירושה מימוש המשדר המתואר בציור (2.5).

ציור מס' 2.5 - משדר אופטימלי עבור שתי סיביות לדגם.Fig. 2.5 - Optimal transmitter for two bits per sample.

בכל יחידת זמן x עובר כמוני לשתי סיביות (e_n) ע"פ התוצאות המוגדרות ע"י
השלישיה $\{w_n, y_n, z_n\}$

$$e_n = \begin{cases} 00 & , \quad x_n < w_n \\ 01 & , \quad w_n \leq x_n < y_n \\ 10 & , \quad y_n \leq x_n < z_n \\ 11 & , \quad x_n \geq z_n \end{cases} \quad (2.29)$$

בהתהו ש- $\{x_n\}$ תהליכי גאוסי עט צפיפות הסתברות דו-מימדית ידועה (משוואת 2.5),
השלישיה $\{w_{n+1}, y_{n+1}, z_{n+1}\}$ מחושבת לקלט אנטרופיה מותנית מירבית כלהלן:

$$P_r\{e_{n+1} = i | e_n, w_n, y_n, z_n\} = \frac{1}{4} \quad \text{עבור } i = 0, 1, 2, 3 \quad (2.30)$$

סכמה כזו, שהיא הרחבה טבעיות מסיבית אותה לדגט, יוצרת בעיה של גודל זכרון הוחורג
לאין שעור מהמקובל במערכות מעשיות - טבלת החזוי מקבלת ט"כ 26 סיביות בתוכת
ומפיקה 24 סיביות נתוניות. לצורך זה דרוש זכרון של 192 Mbyte.

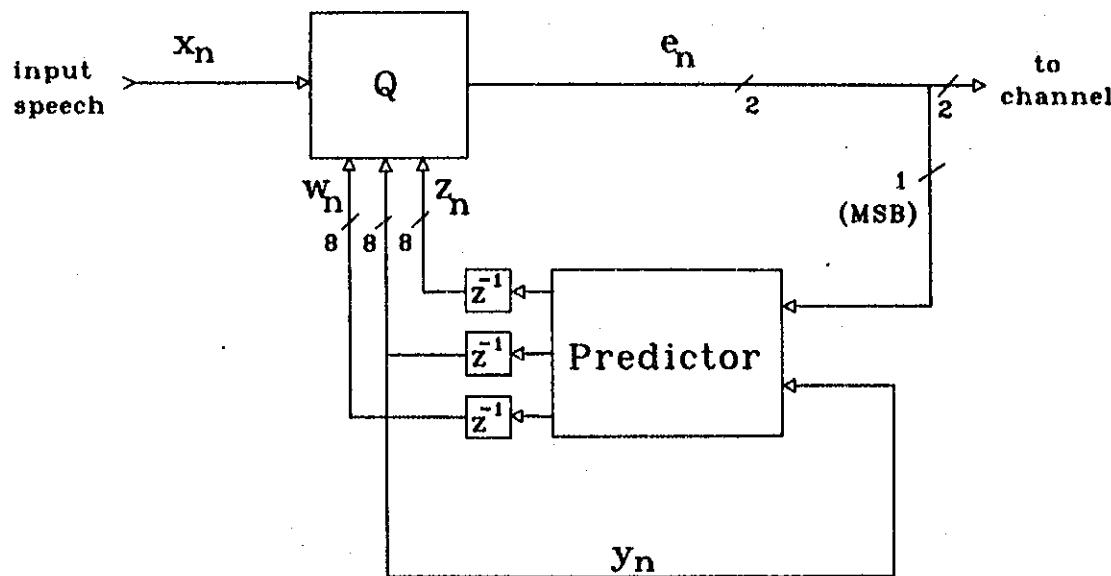
קיט, את כן, הכרתו להabil את מספר המצביים. לשם כך מוצעת סכמה מת-אופטימלית אשר
mbטיחת אנטרופיה מותנית מירבית מסדר ראשון בהנתן המידע העומד לרשות החזאי: הסיבית
המשמעותית של e_n מצעד הזמן הקודם והחיצונית. בציור (2.6) מואר המשדר המוצע.

מערכת זו היא מת-אופטימלית במובן זה שלא כל המידע העומד לרשות המשדר מצעד הזמן
הגוכחי מנוצל לצורך כמו אופטימלי בצעד הזמן הבא. במיליט אורת, האופטימיזציה
מתקיימת על $H(e_{n-1} | e_n)$ במקום על $(H(w_{n-1}, y_{n-1}, z_{n-1}) | e_n)$ ומתורת האינפור-
מציה ידוע כי מתקיימת:

$$H(e_n | e_{n-1}, z_{n-1}, y_{n-1}, w_{n-1}) \leq H(e_n | e_{n-1}, y_{n-1})$$

כעת מובן כי מקסימיזציה של אגף שמאל תגרור מקסימיזציה של אגף ימין (כי שני
האגפים חסומים מלעיל ע"י המספר 2) אך ההפך כמובן אינו נכון.

המחברים Angel ו-Daigle [3] מתייחסים גם כן לביעית ההרחבה לשתי סיביות. הם איבט
מתארים כיצד פתרות אר מציגנים גודל הזכרון הדרוש עבור החזאי והמשער. הפתרון
המושע בסעיף זה מתאים לפתרונות מבחינת דרישות הזכרון.



ציור מס' 2.6 - משדר תת-אופטימלי לקאב שתאי סיביות לדגם.

Fig. 2.6 - Suboptimal transmitter for two bits/sample.

נשובה עליה לתאריך פונקציית החזוי עבור המשדר שבציור 2.6. נטמן ב- e_n^1 את חסיבת המשמעותית של e_n^1 . מתקבל את תחומי ההמצאות הוג של x_n באופן זהה לחולוטין למתואר בטעיף 2.3.1 עבור סיבית אחת לדגם. דמיינו:

$$e_n^1 = \text{sign} (x_n - y_n) \quad (2.31)$$

יעים לקבל ארבעה תחומי שווי הסתברות, דרוש לחלק את הפלוגים $(y_n^1 | x_{n+1} > x_n)$ ו- $(y_n^1 | x_{n+1} < x_n)$ לארבעה שטחים שוויים כל אחד. כלומר עבור $e_n^1 = 1$ נקבל:

$$\int_{-\infty}^{w_{n+1}} g(\theta | x_n > y_n) d\theta = \frac{1}{4} \quad (2.32)$$

$$\int_{-\infty}^{y_{n+1}} g(\theta | x_n \geq y_n) d\theta = \frac{1}{2} \quad (2.33)$$

$$\int_{-\infty}^{z_{n+1}} g(\theta | x_n \geq y_n) d\theta = \frac{3}{4} \quad (2.34)$$

(למקרה $y_n < x_n$ המשוואות דומות אך התגנינה היא לא פי המאורע: $y_n < x_n$).

מtower משווה (2.33) עולה כי החזיוון מחושב באופן זהה לחלווטין לזה המתואר בסעיף (2.3.1). המשוואות (2.32) ו-(2.34) הן משוואות דומות למדי. השוני למשהו הוא רק באגף ימין אך הן מגדירות פונקציות שונות. נסמן:

$$w_{n+1} = L(y_n, e_n^1) \quad (2.35)$$

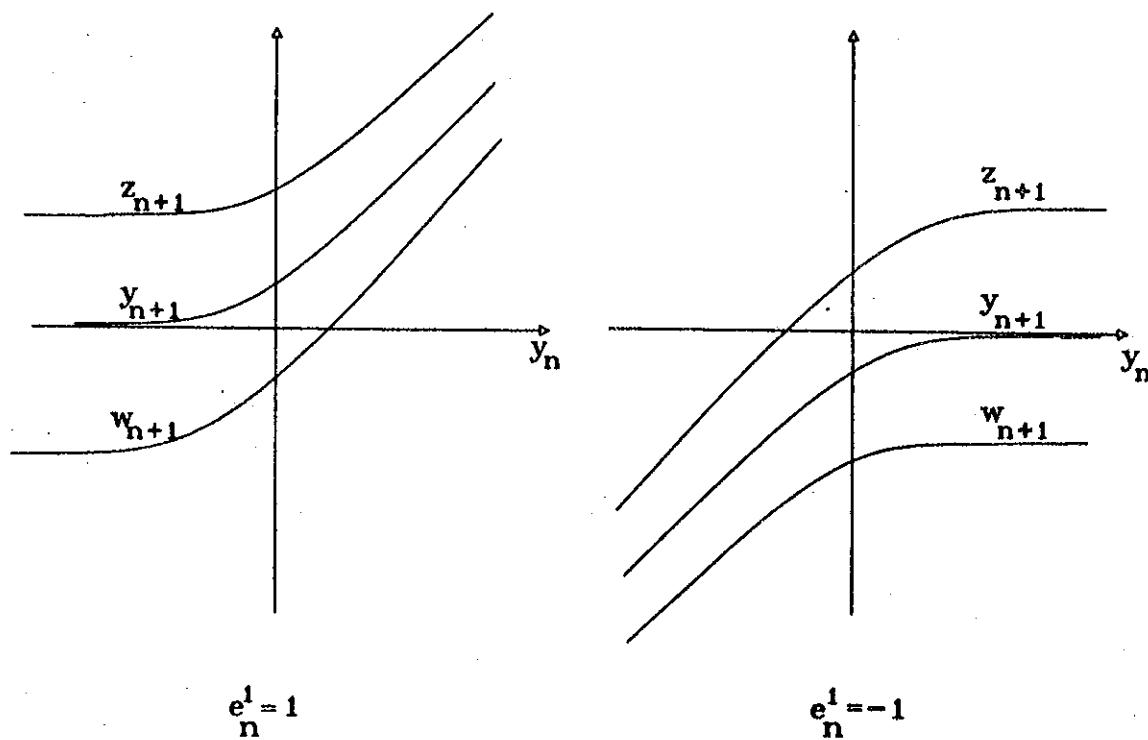
$$z_{n+1} = U(y_n, e_n^1) \quad (2.36)$$

קל להראות כי פונקציות אלו מקיימות "טימטריה הדדית" חילנו: $L(-y, -e) = -U(-y, -e)$. עבור $Q^{-1} = \sigma$ נקבל בדומה ל-(2.16) ביטויים "מפורשיט" (למעט הפונקציות Q ו- U) אשר איננו מפורשוט) עבור הפונקציות $L(0, \cdot)$ ו- $U(\cdot, 0)$ כדלקמן:

$$w_{n+1} = \sigma Q^{-1} \left[\frac{3}{4} Q\left(\frac{y_n}{\sigma}\right) \right] \quad (2.37)$$

$$z_{n+1} = \sigma Q^{-1} \left[\frac{1}{4} Q\left(\frac{y_n}{\sigma}\right) \right] \quad (2.38)$$

בצירור (2.7) מתוארות שלושת רמות החזווי עבור $\sigma = 0.9$.



ציור מס' 2.7 - חזאי עבור שני סיביות לדגם ($\rho = 0.9 = \varphi$).

Fig. 2.7 - Predictor of the two bits-sample system ($\rho = 0.9$).

המשורך

שахזרו את ה captions ע"י המשערך מבחן בדומה למתואר בסעיף 2.3.1. כאשר הקצב הוא שני סיביות לדגם אזי $e_n^1 = e_{n-1}$ מוצגים ע"י שתי סיביות כ"א. לצורך חישוב המשערך יש לבצע כת אינטגרציות כפולות על תחומי המוגדרים ע"י שלושת גבולות הכמוני w_n , y_n ו- z_n . נגידיר עתה את הfonkציה

$$F(e_n, e_{n-1}, y_{n-1}) \stackrel{\Delta}{=} E\{x_n | e_n, e_{n-1}, y_{n-1}\}$$

עבור שתי סיביות לדגם ונבחן את תוכנותיה. בטור דוגמא עבור המקרה $00 = e_{n-1}$ ו- $10 = e_n$ מופיעים הביטוי

$$\hat{x}_n = F(10, 00, y_{n-1}) = \frac{\int_{y_n}^{z_n} \int_{-\infty}^{w_{n-1}} g(x, y) dy dx}{\int_{y_n}^{z_n} \int_{-\infty}^{w_{n-1}} g(x, y) dy dx} \quad (2.39)$$

ביטויים דומים של $F(e_n, e_{n-1}, y_{n-1})$ (עם תחומי האינטגרציה המתאימים) ניתנים לרשום עבור 15 מקライט הנוטרים.

על מנת שבאgal ימינו יהיה רשום ביטוי התלויה אך ורק ב- y_{n-1} ולא בגבולות הכמוני הנוספים (w_{n-1} ו- z_{n-1}) יש צורך להגדיר ולהשא את פונקציות המעבר מ- y_{n-1} אל z_{n-1} ו- w_{n-1} :

לצורך זה נסמן ב- $(\cdot)^{-1}P$ את הפונקציה החופשית ל- $(\cdot)P$ המוגדרת ב-(2.15) (פונקציה P^{-1} קיימת מאחר ש- P מוגוטונית ממש). בעזרה $(\cdot)^{-1}P$ נוכל לקבל את w_{n-1} ו- z_{n-1} כדלקמן:

$$z_{n-1} = \begin{cases} U[P^{-1}(y_{n-1}), 1] & , \quad y_{n-1} > 0 \\ U[-P^{-1}(-y_{n-1}), -1] & , \quad y_{n-1} < 0 \end{cases} \quad (2.40)$$

$$w_{n-1} = \begin{cases} L[P^{-1}(y_{n-1}), 1] & , \quad y_{n-1} > 0 \\ L[-P^{-1}(-y_{n-1}), -1] & , \quad y_{n-1} < 0 \end{cases} \quad (2.41)$$

הארגון הראשון בכל אחד מהביטויים הבניל $(y_{n-1})^{-1}P$ או $(-y_{n-1})^{-1}P$ שווה ל- y והארגון השני - ל- e_{n-2} (אשווות בכל מקרה לסימן של y_{n-1}). פונקציית המעבר z_{n-1} ו- w_{n-1} מחושבות לפי הביטויים (2.35), (2.15) ו-(2.36) בהתאם.

כעת המשער $(\cdot)P$ מוגדר ע"י y_{n-1} , e_{n-1} ו- e_n בלבד. המגרעה של פתרון מסווג זה היא שהפונקציות המתקבלות עבור שעבורו \hat{x} הן בלתי מפורשות, חישובו מסובך למורי וכתוואה מכיוון אי-כן אפשרות אדפטציה כלשהיא נוכח השתנות אפשרית של \hat{x} ו- s עבור מקורות לא סטטיזונריים וזאת בגין לbijoyim שתתקבלו עבור סיבית אותה (2.26) ו-(2.28).

לפיכך מוצעת בנווט גרטה מת-אופטימלית $\tilde{F}(e_n, e_{n-1}^1, y_{n-1})$ אשר מנצלת עבור e_{n-1} רק את הטיבית המשמעותית. (עבור e מנוצלות שתי הטיביות). פתרון מסог זה אפשר להרחב בנקל את התוצאות שנטקלו עבור סיבית אחת דלעיל ולהשתמש בביטויים "מפורשיים" בדומה ל-(2.26) ו-(2.28). כאשר $1 = e_{n-1}^1$ קיבל עבור המבנה של המשער $(\text{כ"א} \text{ מרבעת המקרים האפשרים של } e) \text{ את הביטוי:}$

$$\frac{1}{4} Q\left(\frac{y_{n-1}}{\sigma}\right)$$

וזאת משיקולים דומים לאלה שמהם נמקלה הנוטה (2.25).

את המונה ניתן בכל מקרה לבטא כഫיש של ערכי הפונקציה $(s, t) R(s, t)$ בשתי נקודות מתאימות בדומה ל-(2.27). בדרך זו קיבל:

$$\tilde{F}(00, 1, x) = \frac{R(x, -\infty) - R(x, L(x, 1))}{\frac{1}{4} Q\left(\frac{x}{\sigma}\right)} = \frac{\frac{\sigma}{\sqrt{2\pi}} e^{-x^2/2\sigma^2} - R(x, L(x, 1))}{\frac{1}{4} Q\left(\frac{x}{\sigma}\right)} \quad (2.42)$$

$$\tilde{F}(01, 1, x) = \frac{R(x, L(x, 1)) - R(x, M(x, 1))}{\frac{1}{4} Q\left(\frac{x}{\sigma}\right)} \quad (2.43)$$

$$\tilde{F}(10, 1, x) = \frac{R(x, M(x, 1)) - R(x, U(x, 1))}{\frac{1}{4} Q\left(\frac{x}{\sigma}\right)} \quad (2.44)$$

$$\tilde{F}(11, 1, x) = \frac{R(x, U(x, 1))}{\frac{1}{4} Q\left(\frac{x}{\sigma}\right)} \quad (2.45)$$

עבור $0 = e_{n-1}^1$ ניתן להעזר בביטויים הניל ובשקלוי סימטריה בדומה ל-(2.22) (כאשר את תפליך הפוך הטימן עבור e מחליף ההפור אלוגי של הטיביות). ע"י שהשימוש בביטויים אלו יהיה עשוי לצרכי אדפטציה של s ו- t יש צורך למצוא פונקציות מפורשות המקרבות נאמנה את הfonקציות L , M ו- U . לנושא זה תוקדש תשומת לב בפרק 4. לפני כן בפרק 3 נבחן מהו המפטור הכרוך בויתור על השימוש בfonקציה F והחלפתה ב- \tilde{F} עבור מקורות סטציונריים.

2.4 תכונות המערכת

בטעיף זה נדון במספר תכונות עקרוניות של אופייני החזוי והשערור שפותחו לעיל. חלון יסיעו בהמשך לקבלת קרובים מפורטים לפונקציות החזוי והשערור עבור גרסאות אדרטיביות של המערכת. לבסוף נציג מספר מעלות ומגרעות של סכמת האנתרופיה המירבית מנוקדת ראות הנדסית.

תכונות אופייני החזוי והשערור

1. סימטריה - תכונת הטימטריה של אופייני החזוי והשערור שצויינה מספר פעמים במהלך הדיון בטיעוף 2.3 (נותאות 2.15, 2.22, 2.35 ו-2.36). במהלך הדיון בטיעוף 2.3, מקורה בסימטריה הזוגית של הפלוג הגאוטי. היא שוויה להועיל לצורך הקטנת הדרכן הנדרש עבור הטבלאות ב-50% במקרים בהם הדבר דרוש.

2. אי לינאריות - פונקציית החזוי והשערור אינן לינאריות כפי שראינו בטיעוף 2.3. מקורה של אי הלינאריות בעובדה שפעולות חמי או חלא אינה לינארית. כפי שראינו בפרק 1, במערכות מקובלות כמו DPCM ו-DM, לעומת זאת, מקובל להעטלם מי הליינאריות הכרוכה בדמיון גם בקבצים הנומוכים בהם אנו עוסקים ולהשתמש בחזאים לינאריים.

3. התנהגות אטימפטוטית

המונח "התנהגות אטימפטוטית" בהקשר זה מכון לאופי פונקציות החזוי והשערור כאשר הארגומנט y שואף ∞ או $-\infty$.

ניתן להראות שפונקציית החזוי עבור $y = \lim_{n=1}^{\infty} e$ מקיימת:

$$\lim_{y \rightarrow \infty} [P(y) - \rho y] = 0 \quad (2.46)$$

$$\lim_{y \rightarrow -\infty} P(y) = 0 \quad (2.47)$$

נושה (2.46) נובעת מהעובדה שאם $y \rightarrow \infty$ אז צפיפות הסתברות המותנית $(y > x_n) g$ מתקrbת ל- $(y > x_{n+1}) g$ על-פי משפט לופיטל:

$$\lim_{y \rightarrow \infty} [g(x_{n+1} | x_n > y) - g(x_{n+1} | x_n = y)] = \lim_{y \rightarrow \infty} \left[\frac{\int_y^{\infty} g(x_{n+1}, \theta) d\theta}{\int_y^{\infty} g(\theta) d\theta} - \frac{g(x_{n+1}, y)}{g(y)} \right]$$

$$= \lim_{y \rightarrow \infty} \left[\frac{\frac{\partial}{\partial y} \int_y^{\infty} g(x_{n+1}, \theta) d\theta}{\frac{\partial}{\partial y} \int_y^{\infty} g(\theta) d\theta} - \frac{g(x_{n+1}, y)}{g(y)} \right] = 0$$

הפלוג הטעולי ($y = x_{n+1} | x_n > y$) g הוא פלוג גאוסי עם חציון (= תוחלת) השווה ל- y ומפוק נובע (2.46).

בתו (2.47) נובע מכך שם $\infty \rightarrow y$ אזי המאורע $\{y > x_n\}$ מתרחש בתסתברות השוואפת ל-1 ולכן הפלוג הטעولي המתאים מקרב לפלוג הטעולי הגaussiano הבתוי מותנה ($x_{n+1} | x$) אשר עבورو החציון הוא אפס.

משיקולים דומים לאלו נקבל עבור הפונקציות U ו- L את האטימפטוטות:

$$\lim_{y \rightarrow \infty} [L(y, 1) - \rho y] = -\sigma \sqrt{1-\rho^2} \alpha \quad (2.48)$$

$$\lim_{y \rightarrow -\infty} L(y, 1) = -\sigma \alpha \quad (2.49)$$

$$\lim_{y \rightarrow \infty} [U(y, 1) - \rho y] = \sigma \sqrt{1-\rho^2} \alpha \quad (2.50)$$

$$\lim_{y \rightarrow -\infty} U(y, 1) = \sigma \alpha \quad (2.51)$$

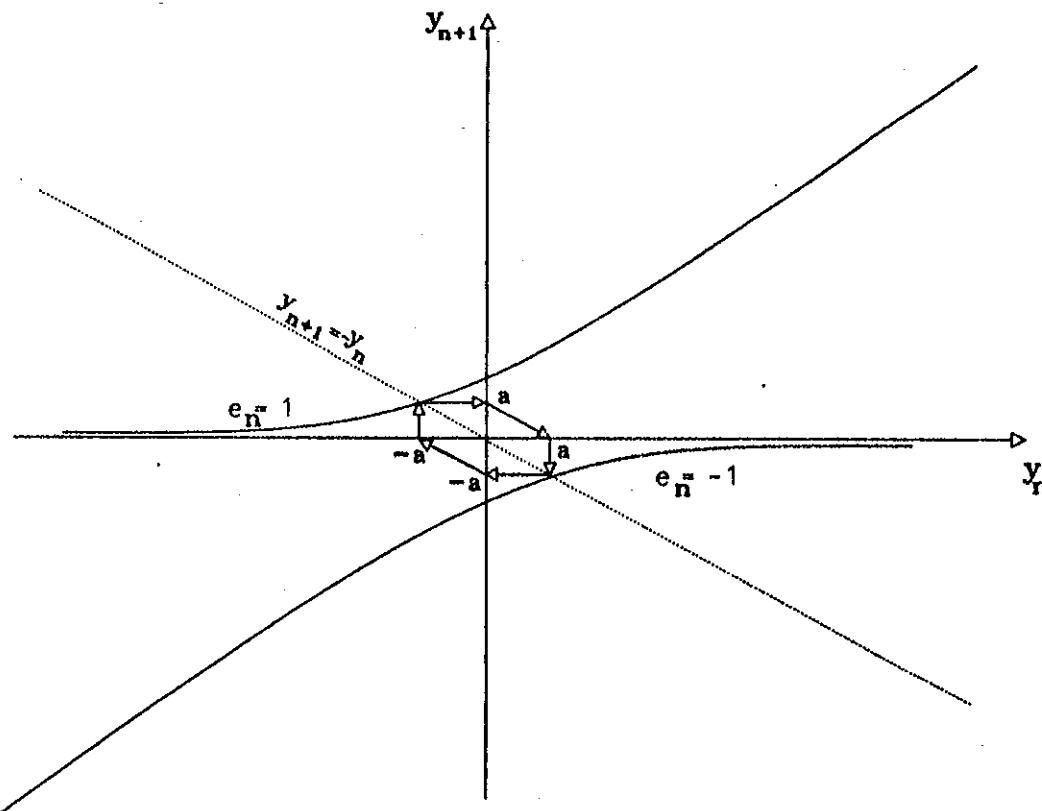
$$\text{כאשר } \Delta \stackrel{\Delta}{=} Q^{-1} \left(\frac{1}{4} \right) = 0.67448 \alpha$$

בדרכ דומה, ניתן להראות שגם המשער (עבור שני קבועי השדרור) הוא בעל אטימפטוטה לינארית (בעלת שפוע ρ) כאשר $\infty \rightarrow y$ ואטימפטוטה אופקית כאשר $-\infty \rightarrow y$ (עבור המקרה $1 = e^{1/n-1}$).

כאמור, תוכנות אטימפטוטיות אלו יתגלו בהמשך לקבלת קרובים עבור גרסאות אדפטיביות של המערכת.

4. מטוללים גבוליים (limit cycles)

בנich כי בקטע מסוים את הכניטה $\{x_n\}$ הוא אפס לכל n (Idling). במקרה זה המשדר, בהיותו מערכת לא לינארית בעלת משוב שלילי, נכנס למסלול גבולי וחוואני מפיק תנודות $y_n = a(-1)^n$ (הינו $a > 0$ ו- a - לטרוגין) כמפורט ציור 2.8.



ציור מס' 2.8 - תופעת המסלול האבולוי.

Fig. 2.8 - The limit cycle effect.

משרעת התנודות a עבור מקרה 1 = σ מתקבלת מתוך התרת המשוואות:

$$P(-a) = \sigma Q^{-1} \left[\frac{1}{2} Q\left(-\frac{a}{\sigma}\right) \right] = a$$

$$a = \sigma Q^{-1} \left(\frac{1}{3}\right) = 0.431 \sigma$$

אשר פתרונגה:

עבור $1 < \alpha$ מתקבל מושעת נמוכה יותר.

התוצאות המתקבלות לא משתנה אם אם n נניח $\alpha = -\frac{1}{n}$ להשתנות בתחום $(-\alpha, +\alpha)$. בambilits אחרות, התופעה מתרחש בקטועים בהם $\alpha = -\frac{1}{n}$ לש אנרגיה נמוכה יחסית. באופן דומה אם α הוא קבוע סטטי וגדול בערכו המוחלט $\alpha > 1$, מתקבל גם כן מיטול גבולי אך בעל מחזור ארוך יותר ואmplיפיטודה α .

המשמעות של תופעה זו היא שבקטועים חלקים יחסית או נמוכי אנרגיה ישמעו באוט המשוחזר צפופים בתדר גבוהה. בעית דומה (channel idle noise) מוכרת גם במערכות DPCM ודلتא מודולטור בהן היא נפרת ע"י התקנת גודל צעד הCOMMON בהתאם. בפרק 4 יוצע מבגרון להרחקת השפעת המיטול הגבולי עבור מערכת האנטרופיה המירבית.

5. שחזור עבר סדרת החזוי ("אובטראוביליות")

נניח כי ברגע נתון ידוע למקלט (טיביתאות לדגט) ערך החזו y_n . מתחבר שניינו עקרונית לשחזר את עבר הסדרה y ללא הגבלה באמצעות נוסחת הנסיגה הבאה:

$$y_{n-1} = \begin{cases} P^{-1}(y_n) & , \quad y_n > 0 \\ -P^{-1}(-y_n) & , \quad y_n < 0 \end{cases} \quad (2.52)$$

$$e_{n-1} = \text{sign}(y_n) \quad (2.53)$$

דבר זה אפשרי באופן תאורטי הודות לשני גורמים:

א. הפונקציה P מוגנות (ועיל חפיקה).

ב. לפונקציות $(x)^{-1}$ ו- $(-x)^{-1}$ טווחים זרים.

גורם ב' לא קיים ב-DPCM שעבורו חזאי לינארי (ראה ציור 2.3) כיוון שהאופיגנים של החזאי אינם בטוחחים זרים ולא ניתן לבצע נסיגה בזמן באופן חד-משמעי.

טענה זו עומדת לכואורה בסתירה עם תורת האינפורמציה: די כביכול בשדור ערך אחד של y (במוקום סדרה של y) כדי לשחזר ללא גידול בעוות, בлок ארוך כרצובנו של המקור. סתייה זו מתইישבת בכך, שדרוש ליצא את y בבדיקה אינטואטיבית. אלמלא כן מצטברות שגיאות גדולות וחולכות מתלין הנסיגה המתוואר וזאת משום שהפונקציה $(x)^{-1}$ היא בעלת נגזרת גדולה מיחידה בכל מקומות (מגדילה הפרשיות). ברם, ניתן בכלל זאת ניתן לנצל תכונה זו ע"מ להקטין את קצב השדור ללא עות נסף, אך אפשרות כזו לא נבדקה במטרת חיבור זה.

נסכם איפוא מספר מעילות ומגרעות של המערכת המוצעת:

מעלות

1. **אנתרופופיה מירבית** - תכונה רצiosa עבור דחיטה אופטימלית: כאמור בסעיף 2.1, מתקיים התנאי ההכרחי לקבלת עותם מינימלי בשחוור בגלל הניצול של כל רמות חכמי במידה שווה.
 2. **שימוש בחזווי לא לינארי** - בוגוד למערכות קיימות המוגבלות בד"כ לחזווי לינארי אשר בד"כ אינו אופטימלי.
 3. **פשטת הממוש** - בדוחסה ובשחוור לא נדרש כל פעולות חשוב אלא גישות לזכרון בלבד. תכונה זו, כאמור, מסייעת בעיקר בישומי דחיטה של תМОנות בעות אשר עברון נדרש קצב דגימה גבוהה מאוד וקיימיט, לפיכך, אילוצים חמורים בדבר הסיבוכיות החישובית.
- על שלושת גורמים אלו, למעשה, מושתת המוטיבציה העיקרית למחקר זה: לבחון את **האנתרופופיה המירבית** כקריטריון תחלIFI לשגיאה לריבועית ממוצעת, לבדוק האם קיימת עדיפות לחזווי לא לינארי על פני לחזווי לינארי, ולהשוות ביצועים עם מערכות קיימות על רקע עלילונות המערכת המוצעת מבחינת **סיבוכיות חישובית**.

מגבלות

1. **מגבליות על הסטטיסטיקה של המקור** - המערכת מתוכננת למקור גאוסי. בנוסף ידוע אפריאורי של הפליג טגולי (ידיעת μ ו- σ). יתרה מזאת, ההנחה של פלוג גאוסי אמונה מפשטה את האנליזה הדרישה למכנו מחד גיטא, אך מיידך, ידוע כי מבין כל המקורות הסטטיזונרים, המקור הגaussi הוא הקשה ביותר לדחיטה (ambilן המשתנים האקראיים בעלי אותה שוננות, למשתנה האקראי הגaussi האנתרופופיה הגובלת ביותר). לפיכך קיים הפעלה מסוימת בהגחת הפלוג הגaussi אם המקור העומד לרשותינו אינו צזה.
 2. **חוסר גמישות לשינויים בתכנון** - המערכת המוצעת אינה גמישה להגדלת קצב השדר או להגדלת סדר החזווי. ראיינו כי המעבר מטביה אחת לשתי סיביות לדגם הצריך חישוב חדש של כל אופייני החזווי והשעור. יתרה מזאת, התעורר צורך לוותר על מסדר אופטימלי מטעמים מעשיים ולהסתפק בפתרון מת-אופטימלי (סעיף 2.3.2).
- המעבר לחזאי מסדר שני עבר סיבית אחת לדגם אף הוא בעייתי משום ששיטם כך דרוש זכרון של 128 Mbyte (8 סיביות כתובות של $1-n$, 8 סיביות של $2-n$ וסיבית אחת של n). גם באדרטציה של המערכת קיימים קשיים מסוים שפונקציות החזווי

והשעורה, כאמור, איןנו פונקציות מפורשות של הפרמטרים הטעיטיטיים α ו- β . לצורך האדפטציה יש הכרח להסתפק בקרובים לפונקציות אלו אשר יהוו פונקציות מפורשות כפי שנראה להלן בפרק 4.

3. תת-אופטימליות - חזאי וחישור המתוכנים כמתואר בסעיף 2.3 לעיל אינו אופטימליים מאחר שאינפומציה בה הם נעדרים ברגע נתון אינה מנוצלת באופן עיל. נtabונן למשל בחזאי לטיבית אחת עבורו חישבנו את החיזון של הפלוג הטגולי ($y_{n+1}^n | x$). לצורך פתרונו בעיה זו תהייחסנו של x^n ו- x^{n+1} כאל שני משתנים אקראיים ולמעשא לא הובאה בחשבון העובדה שהט ל Kohim מזור תוליך. ברט, כאמור $y_n < x^n$ או $y_n > x^n$ הנתו לחזאי ברגע n כלשהו גם אינפומציה מסוימת אודות העבר של x^n שאיא רלוונטי לגבוי חזוי ושערוך בזמן עתידי, וזאת גם אם המקור הוא מוקבי מסדר ראשון. לצורך הגדרת הפתרון האופטימלי לביעית החזוי נגידר את (y, k) קבוצה כל סדרות המדגט באורך k $\{x_{n-1}, \dots, x_{n-k}, x_{n-k+1}, \dots, x_{n-1}\} \triangleq \{x_{n-k}^{n-1} | x\}$ של המקור שעבורן החזאי מפיק את הערך y ברגע n ($y = x_n$). החזאי האופטימלי עבר $1 = e$ יהיה איפוא נתון עיינ:

$$M_{opt}(y, 1) = \lim_{k \rightarrow \infty} \text{med}\{x_{n+1} | x_n \geq y, x_{n-k}^{n-1} \in A_{n,k}(y)\} \quad (2.54)$$

כאשר $\text{med}\{\cdot\}$ מציין חיזון של פלוג. באופן דומה ניתן להגיד את המשערך האופטימלי בעזרת אופרטור התוחלת המותנית.

הчисוב ב-(2.54) הוא בעייתי, ראשית, משום שיש קושי בחישוב הקבוצה (y, k) ו שנית, משום שיש צורך לחשב חיזונים ותוחלות של פלוגים רב-מיידים ולעבור לאבול (בהנחה שהוא קיים).

נשאלת השאלה האם ניתן למתקרב ל-(2.54) באופן איטרטיבי? לשם כך מוצעת דרך הבאה: נגידר את $(e, M_k(y))$ כפונקציית החזוי המתקבלת בצעד ה- k -י של האיטרציה. נבחר כי ניחושי המחלמי (y, e) את החזאי M כפי שהוא בסעיף 2.3. נעזר עתה בחוכנה מס' 5 דלעיל על פיה ניתן באופן אורתוי לשזר את עבר סדרת החזוי $\{y_1^{n-1} | y\}$ ואת עבר סדרת הקמי $\{e_1^{n-1} | e\}$ מזור ידיעת y . ונגידר:

$$M_1(y_n, e) = \lim_{k \rightarrow \infty} \text{med} \left\{ \int_{(y_n, e_n)} dx_n \int_{(y_{n-1}, e_{n-1})} dx_{n-1} \dots \int_{(y_{n-k}, e_{n-k})} dx_{n-k} g(x_{n-k}^{n-1}) \right\}$$

כאשר הסימון (e_i, y_i) מציין את תחום האינטגרציה $[y_i, \infty)$ אם $i = 1$ ואת $(-\infty, y_i]$ אם $i = -1$ ($e_i = e$ ובכך מוגדר ה"ניחוש" החחלמי עבר (y, e) $A_{n,k}(y)$). $(x_{n-k}^{n-1} | y)$ מציין את חפלוג הטגולי המשותף לרכיבי הוקטור $(x_{n-k}, \dots, x_{n+1})$.

בצד האיטרצייה הבא נשים לmouseup על פניו כל אוסף סדרות המדגם $\{x_{n-k}^{n-1}\}$ ה"ימותרות"
שמושרות ע"י הפונקציה $(e, y)_1^M$. נסמן אוסף זה ב- $(y_n, A_{n,k}^1)$ ונקבל:

$$M_2(y_n, e) = \lim_{k \rightarrow \infty} \text{med} \left\{ \int_{(y_1, e_n)} dx_n \int_{A_{n,k}^1(y_n)} \dots \int_{i=n-k}^{n-1} dx_i g(x_{n-k}^{n+1}) \right\}$$

באופן כזה ממשיכים את תהליך עבור M_3 וכו' עד לתנאי עצירה מסוים. שאלות
המעורבות מעבר לקשיים החישוביים הכרוכים בתהליך זה הן הבאות:

1. האם גם עבור $0 > i$ הפונקציות $(e, y)_i^M$ מקיימות את תכונה מס' 5, תיינו, שניתן
לשזר חד ערכית העבר ולקבוע בקלות את הקבוצות $(y_n, A_{n,k}^i)$?
2. האם $(e, y)_k^M$ מתכנס ל- $(y, e)_{\text{opt}}^M$?

התהליך האיטרטיבי המוצע ושאלות הניל לא נבחנו, אלא הטעפנו במטרת חיבור זה
בפתרונות שהוצעו בסעיף 2.3 עקב פשوطם היחסית ומתוך הנחתה שמירב האינפורמציה
החשובה לחזוי ושערו מילא גלומה בהוויה ותרומת העבר היא משנית. ברם, כפי שנראה
בפרק הבא, ככל שקדם קורלציה ק גביה יותר, כך השפעת העבר משמעותית יותר ועל
כך החזאי והמשמעות המוצעים בסעיף 2.3 מתרחקים מלאו האופטימליים.

פרק 3 – ביצועי המערכת עבור מקורות סטציונריים

בפרק זה נבחן את ביצועי מערכת האנטרופיה המירבית (ME) למידוד מקורות סטציונריים בקצבים של סיבית אחת ושתי סיביות לדגס. מתעניין בשבר שבין יחס האות לרעש באוט המשוחזר והאנטרופיה של המידע המשודר לבין מקדם הקורלציה ק' וنبחנו את רגישות יחס האות לרעש לטטייה במקדם הקורלציה של המקור (הינו באיזו מידת נפגעים הביצועים כאשר אין תאמת מדיקת בין המקור למערכת). בנוסף, תבדק התנהגות הביצועים כאשר המקור אינו גאוטי וביצד ניתן לשפרם. עבור מקורות גאוט-מרקוב בערכת השוואת עם מערכות קיימות (LDM עבור סיבית אחת ו-DPCM עבור שתי סיביות לדגס) ועם החפט המאורת של יחס האות לרעש הנקבע מתוך פונקציה קצב העוות של מקורות גאוט-מרקוב מסדר ראשון. המחברים Daigle ו-Angel ([2], [3]) לעומת זאת, השוו את ביצועי המערכות הנילישירות על אותן דבors ותחמונות.

מדידת הביצועים עבור מקורות סטציונריים מאפשרת לקבל הערכה מבוקרת ובתאי מלאה אודוות מהירות האדפטציה שדרש למקורות לא סטציונריאים וההפסד הצפוי ביציעים נוכח אופיו הלא-גאוטי של אותן דיבורים.

חישובו היישר של יחס האות לרעש מתוך תוחלת הספק שגיאת השחזר באופן עיוני (וללא סימולציה) כירור בקשיט הנובעים בעיקר מהעובדה שלא נמצא בטוי מפורש עבור הפלוג הסגול של ערך החזוי \hat{x}_n , שהוא כמובן חיווני לצורך זה. מסתיבת זו בחרנו לאמור גדלים אלו באמצעות סימולציות.

3.1 התלות של שגיאת השחזר והאנטרופיה במקדם הקורלציה

לצורך חישוב של יחס האות לרעש והאנטרופיה כמלות ב- \hat{x}_n נעלכה סדרת סימולציות בהן הוזנתה המערכת עבור ערכים שונים של ק', בסדרת תהליכי גאוט-מרקוב סינטטיים עם ערכי ק המתאימים.

ערכים ק שנבדקו היו 0.98, 0.95, 0.9, 0.8, 0.5, 0.2 = ק. עבור כל אחד מערכי ק הניל הופקו $50 = K$ סדרות מדגם של תהליך גאוט-מרקוב מסדר ראשון באורך של $4096 = N$ דגימות כ"א באמצעות משוואת ההפרש:

$$x_n = \rho x_{n-1} + \sigma \sqrt{1-\rho^2} v_n \quad (3.1)$$

כאשר $\{v_n\}$ הוא רעש לבן גאוטי בעל תוחלת אפס ושותיות יחידה.

בנוסף, עבור כל ערך של e חושבו אופיליני החזווי והשערוני המתאימים. על פניו סדרות המדגמת נסדו לכל e הגדלים הבאים:

1. אנטרופיה מותנית משוערת מסדר ראשון. לצורך הגדרת נגדי תחילה משערר אנטרופיה מסדר n עבור סדרת המדגם ה- n -ית:

$$\hat{H}_n^i(e) \stackrel{\Delta}{=} - \sum_i \hat{p}_i(e_1^n) \log \hat{p}_i(e_1^n)$$

כאשר $(e_1^n)_k$ מצינו לגבי סדרת המדגם ה- n את המופיע היחסי של כל אחד מ- n^2 הבלוקים האפשריים באורך n - $\frac{n}{1}$ מעל האיבר הבינארי.
האנטרופיה המותנית המשוערת מסדר ראשון לגבי סדרת המדגם ה- n -ית מוגדרת ע"י:

$$\hat{H}_i(e_n | e_{n-1}) \stackrel{\Delta}{=} \hat{H}_2^i(e) - \hat{H}_1^i(e)$$

והאנטרופיה המשוערת הרצוייה נתונה ע"י:

$$\hat{H}(e_n | e_{n-1}) = \frac{1}{K} \sum_{i=1}^K \hat{H}_i(e_n | e_{n-1})$$

2. יחס אותן לרעש:

$$SNR \stackrel{\Delta}{=} 10 \log_{10} \left[\frac{\frac{1}{NK} \sum_{n=1}^{NK} x_n^2}{\frac{1}{NK} \sum_{n=1}^{NK} (x_n - \hat{x}_n)^2} \right]$$

כאשר \hat{x}_n הוא השזoor של x במקלט.

ובכן שליחס אותן לרעש יש חשיבות מעשית גדולה יותר מאשר לאנטרופיה כאשר מדובר בביטויים. למדידת האנטרופיה יש בעיקר ערך לימודי ומטרתה למעשה לבחון את מידת הצימוד בין שגיאת השזoor לאנטרופיה ולהסביר על ידי כך על השאלה באיזו מידה מוצדקת גישת המקסימיזציה של האנטרופיה כתחליף למינימיזציה של השגיאה הריבועית הממוצעת (כפי שראינו בסעיף 1.2, בעית המינימיזציה הישרה של השגיאה הריבועית הממוצעת היא קשה לפתרון בקביניט נומכית).

החותם העליון על יחס אותן לרעש שנייתן לצפות בשזoor תחlixir אוט-מרקוב מסדר ראשון נוצר מפונקציה קצב העורות שלו הנתונה ע"י (ראה גם [4]):

$$R(D) = \frac{1}{2} \log_2 \left[\frac{\sigma^2(1-\rho^2)}{D} \right] \quad (3.2)$$

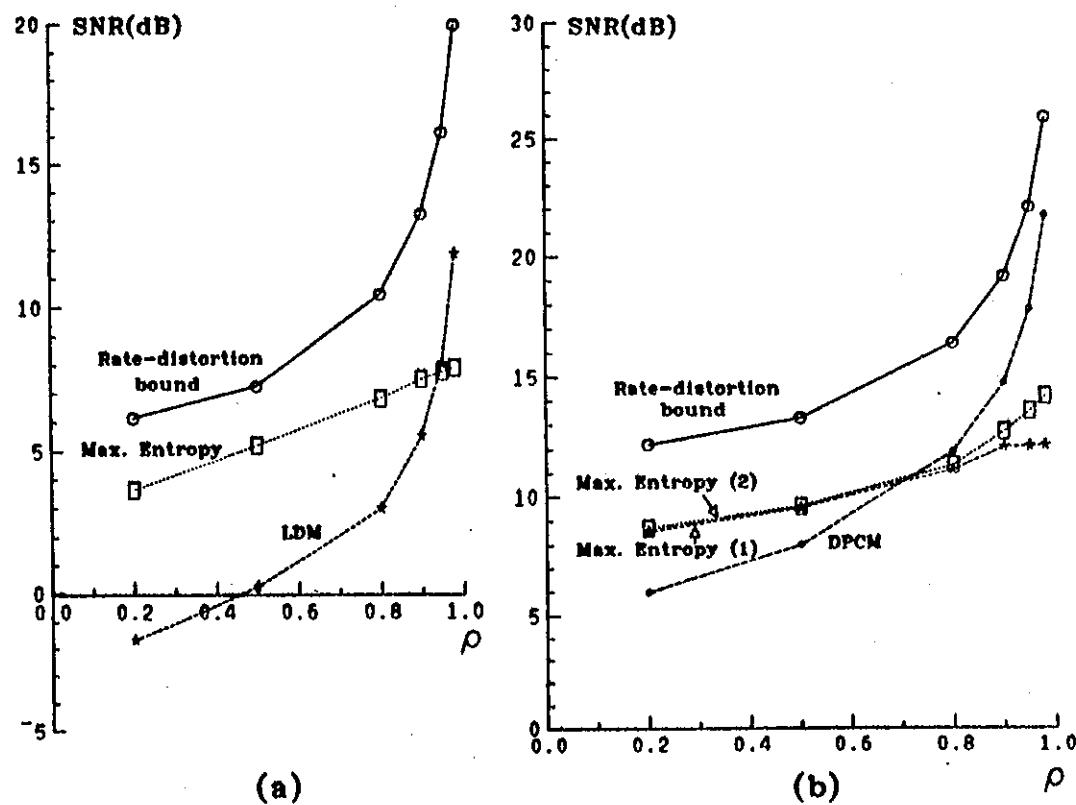
כאשר:

D - השגיאה הריבועית הממוצעת בשחזר $(\hat{x}_n - x_n)^2$
 R - מספר הסיביות לאות מקור (= מס' סיביות לדגם).

מתוך (3.2) ניתן לחוץ את יחס האות לרעש המרבי:

$$\text{SNR(dB)} = 10 \log_{10} \left(\frac{\sigma^2}{D} \right) = 6.02R - 10 \log_{10}(1-\rho^2) \quad (3.3)$$

עבור המקרים בהם אנו עוסקים ערכו של R הוא 1 או 2. בציור 3.1 מוצגים גרפים של יחס זה לרעש המתkeletal ע"י מערכת האנטרופיה המרבית (ME), ע"י DPCM וכן ע"י החסם תעליוון לפי משווה (3.3). עבור דגם/סיבית 1 = R ועבור דגם סיביות 2 = R.



ציור מס' 3.1 - יחסאות לרעש כמויות ב- ρ עבור מערכות שונות:

(a) דגם/סיבית 1 = R (b) R = 2 דגם/סיבית 2 = R.

Fig. 3.1 - Signal-to-noise ratio versus ρ for the various systems:

(a) $R = 1$ bit/sample, (b) $R = 2$ bits/sample.

בציוור (3.1a) נארכט השוואה עם LDM (Linear Delta Modulator) שהוא למעשה גרטה של DPCM בסיבית אחת לדגם ובציוור (3.1b) מוצגים באורה דומה אופייני לחס אותן עבור עבור DPCM בקצב של שתי סיביות לדגם (16kbit/sec). לארכט LDM הותאם עבור כל ערך של ρ גודל צעד אופטימלי על פי נוסחה אמפרית מקובלת (ראה [9]):

$$\Delta_{\text{opt}} = E^{\frac{1}{2}}(x_n - x_{n-1})^2 \ln(2F_o) = \sigma \sqrt{2(1-\rho)} \ln(2F_o) \quad (3.4)$$

כאשר $F_o = F_s / 2F_N$ - הוא מקדם "דיגימת יתר" (oversampling index). F_s - תדר ניקויסט, F_N - קצב הדיגמה.

אם הדיגמה היא בקצב ניקויסט אזי F משוקף תפקיד של מספר הסיביות לדגם ($=1$). מאחר ש- $\{x_n\}$ היא סדרה טיננטית המיצרת במחשב עבור המקרים בהם געסוק בפרק זה, אזי כנich לשם נוחיות כי $\{x_n\}$ "נדגמה" בקצב ניקויסט ($=1$) וע"י נוכל לרשום:

$$\Delta_{\text{opt}} = \sigma \sqrt{2(1-\rho)} \ln 2 \quad (3.5)$$

עבור DPCM ($R=2$) גודל הצעד האופטימלי נתබל באופן נטיוני לכל ערך של ρ . עבור המערכת המוצעת ME מוצגים חלק (b) של הציגור הביצועים עבור שתי גרסאות שונות במקצת אשר תוארו בסעיף 2.3.2 לעיל בעזרת הפונקציות $(\cdot, \cdot, \cdot, \cdot, \cdot, \cdot)$ ו- $(\cdot, \cdot, \cdot, \cdot, \cdot, \cdot)$ (בציוור: (1) Max Entropy ו-(2) Max Entropy בהתאם). ניתן לראות כי התפיש אצפוי ביצועים כתוצאה משימוש ב- \tilde{F} כתחליף ל- F הוא קטן למדי.

בשני חלקי הציגור ניתן להבחין במספר גורמים מאפיינים:

א. כל המערכות משלבות ביצועים הנוגלים באורה שימושוותי מהחסם העליון. ההפרש מבטא למעשה את שעור התפיש ביצועים כתוצאה מהעובדת שכמוני הינו סקלרי (אל החסם ניתן להקרב בטכניות של כמו וקטורי).

ב. במעבר מקצב שדרור של סיבית 1 = R לקצב סיביות 2 = R עבור DPCM יש שיפור ביחס אותן לרעש שאינו נופל מ-6dB (בדומה לביטוי (3.3)). לעומת זאת, עבור ME השיפור timid קטן יותר. הסבר אפשרי לכך הוא היתור על חזאי אופטימלי עבור $2 = R$ כתמואר בסעיף 2.3.2.

ג. בנגוד למערכות DPCM ו-LDM עבורן מתקיים:

$$\lim_{\rho \rightarrow 1} \text{SNR}_{\text{opt}} = \infty$$

הרי עבור מערכת ME אופיין לחס אותן לדוש הוא בעל שיפור מתון גם כאשר ρ מתקרב ל-1. מקור התגבר לכך הוא בתופעת המסלול האבולטי (limit cycle) אשר צוינה בפרק הקולדט. תופעה זו קיימת אמג נגט ב-DPCM וב-LDM אך השפעתה קטנה הוודאות

לעובידה שגודל הצעד הנבחר על פי ביתוי (3.5) קטן כאשר δ מתקרב ל-1. המסקנה במערכות ME, לעומת זאת, אין אפשרות לשולות על 'גודל הצעד' במישרין. המסקנה המעשית מכאן היא שאין לצפות לשיפור שימושי בפיתוחים של ME אס מגדיילים את קצב הדגימה.

עבור ערכיט נומוכים של ק, המערכת ME מגלה ביצועים עדיפים. לעומת זאת, עבור ערכיט גבוחית של ק, כפי שצוין לעשרה בסעיף ג', הממערכות DPCM ו-LDM משיגות יחסאות לרעש האגדל באורך משמעותי מזיה המתקבל ע"י ME למראות שזו האחידות מוכננה למחליכי גאוס-מרקוב מלכתחילה. החיבור העיוני לכך נובע מחתה-האופטימליות בתכנון אופיני החזווי והשדרוך כפי שתוארה בפרק הקודם וצוויננה כאחת המוגדרות של המערכת. הביצועים של שתי המערכות משתווים עבור $0.95 \leq k$ כאשר $1 = R$ ועבור $0.8 \leq k$ כאשר $2 = R$. ערכיט אלו של ק הם בתחום האופיני עבור אותן דיבור קולי (Voiced Speech) הדגום בקצב kHz 8. לפיכך, יש עניין לבדוק איזו מן המערכות עדיפה עבור אותן דיבור אמייתי. עם זאת, יש לזכור כי אותן דיבור איננו בעל פולוג גאוסי ב嚷נות להנחות על פיהם תוכנבה מערכת ME.

האנטropיות מסדר ראשון ($\hat{H}|e_n|e_{n-1}$) שנתקבלו עבור כ"א מהמערכות הביל מטוכמות בטבלה (3.1). עבור $ME_1 - 2 = R$ מופיעים ערכי ($\hat{H}|e_n|e_{n-1}$).

Table 3.1 - The conditional first order entropy for the various systems.

מערכת system		ρ	0.2	0.5	0.8	0.90	0.95	0.98
R = 1	ME	0.99975	0.99983	0.99991	0.99995	0.99993	0.99983	
	LDM	0.7805	0.94939	0.986756	0.977164	0.969844	0.95133	
R = 2	ME	1.99810	1.99939	1.99824	1.99860	1.99925	1.99760	
	PDCM	1.41794	1.50128	1.5511	1.64974	1.61837	1.61104	

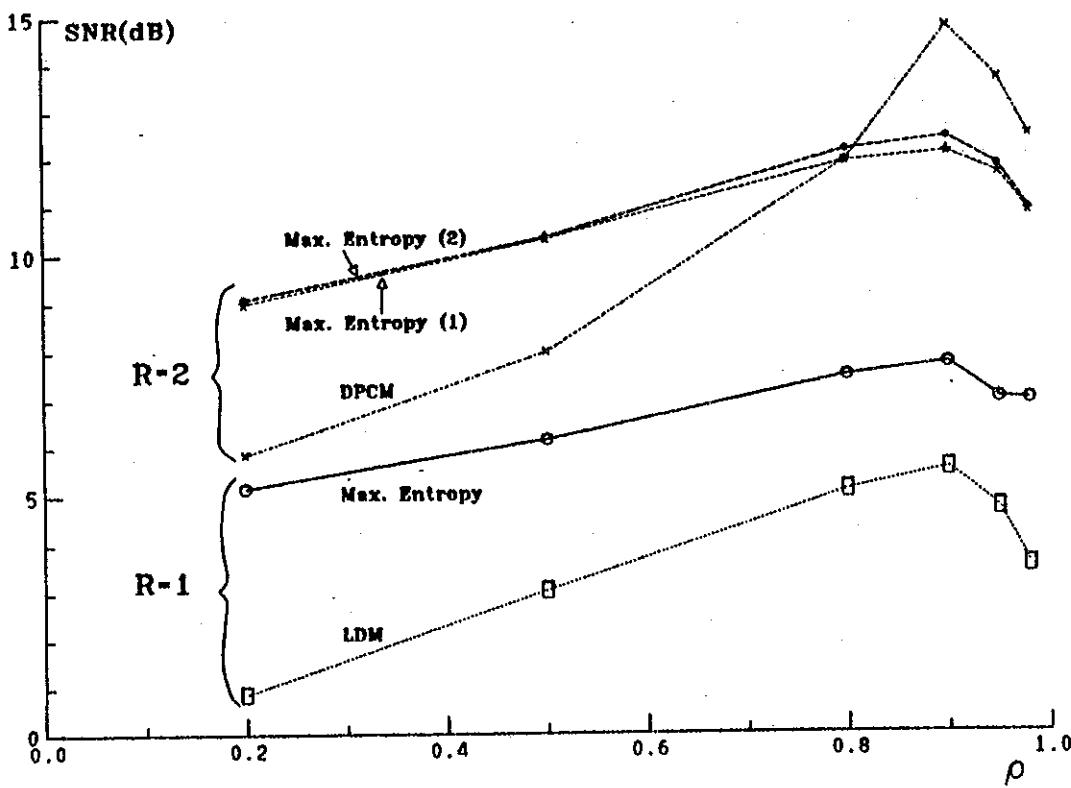
כפוי, האנטרופיה המותנית מסדר ראשון עבורי מערכת ME למעשה מגשימה את קצב השדרור לכל ערך של k הינו: $R = \frac{1}{e^{H(k)}}$. תוצאה זו מתיישבת עם האמור לעיל בפרק 1. עבורי DPCM מתקבלות אנטרופיות הנמוכות באופן משמעותי מallow של ME בעיקר כאשר $2 = R$. תוצאה זו צפואה משוט שמי רמות הקונטינזיה עבורי $2 = R$, שהן נוספות על המוגנות לעיל לא ניתן להציגו במובhawk על קשר כלשהו בין האנטרופיה המירבית. מן התוצאות ישנו עבורי $1 = R$ ב-DPCM, גורמות לחיריה נוספת בין אנטרופיה המירבית. מנו מרונות שהאנטרופיה בשזר. עבורי ערכיו k גבוהים ל-DPCM יחס אותן לרעש עדיף על ME למרונות שהאנטרופיה שלו נמוכה יותר. פרושו של דבר של ידי קידוד אנטרופיה (entropy coding) של יציאות מסדר DPCM ניתן לרזרת בקצב השדרור (המוציא) ללא פגיעה נוספת בשזר ובעקבות להישג בו זמן קצב שדרור נמוך יותר ויחס אותן לרעש גבוה יותר מאשר ב-ME.

3.2 רגישות הביצועים לטיטה במקדם הקורלציה

בטעיף זה נבחן באיזו מידת נפגעים הביצועים של ME כאשר אין התאמה מדוייקת בין מקדם הקורלציה k על פיו מתוכננת המערכת לבין מקדם הקורלציה של המקור. בדומה לטעיף 3.1, מערך השוואת בין מערכת ME למערכות המקובלות DPCM ו-LDM.

במסגרת נסוי זה הוזן מקור גaus-מראקוב בעל פרמטר $0.9 = k$ למערכות המתוכננות עבורי הערכיהם השונים של k (בדומה לטעיף 3.1). בציור (3.2) מוארת התנהגות יחס אותן לרעש עבורי קבועים של טיבית אחת ושתי טיביות לדגמ.

כפוי, בכל המערכות מתבל יחס אותן לרעש מירבי כאשר יש התאמה בין המקור למערכת, היינו עבורי $0.9 = k$. עבורי שני קבועי השדרור יש למערכת ME רגישות נמוכה לטיטה בערך של k (בציור ניתן לראות כי השיאים בייחס אותן לרעש רחבים למדי). פרושו של דבר שעבורו מקורות לא סטציונריים מהירות האדפטציה הדרישה אינה גדולה. לעומת זאת, למערכות DPCM ו-LDM רגישות יותר גבוהה במקצת. עבורי דגם/טיבית $1 = R$ בנסוי זה, מערכת האנטרופיה המירבית עדיפה על פני LDM בכל מוחם הערכיהם של k , אך עבורי דגם/טיביות $2 = R$ אין הדבר כך. כאשר k מתרחק מ-1 הביצועים של DPCM עדיפים על אלו של ME. בדומה להנחה ש叙述ה בטעיף 3.1, גם כאן אנו עדים לשיפור גדול יותר של יחסunas עבורי לרעש ב-DPCM כתוצאה מהטotta טיבית בשדרור למרונות שכעת אין התאמה בין המקור למערכת. כמו כן, יש לזכור כי גפקיד הפורט k ב-DPCM מرتبط גם בקיימות גודל הצעד (ראה נסחה 3.5) בנוסך להיותו פרמטר חזוי. דבר זה מאפשר לכובן באופן רצוני את עצומות רעש "ארוותית" (slope-overload noise) ורעש ה"גרגירית" (granular noise) זה על חשבונו זה עד לקבלת גודל צעד אופטימלי. לעומת זאת במערכת ME בגרסת הנוכחות אין אפשרויות כזו כאמור לעיל.



aczior מס' 3.2 - רגישות יחס הזרע לרעש לטטיה במקדם הקורלציה.

עבור $\rho = 0.9$ יש התאמה בין המקור לבין המערכת.

Fig. 3.2 - The sensitivity of the SNR to deviations of the correlation coefficient. For $\rho = 0.9$ the system matches the source.

בדומה לטעיף 3.1, עבור $R = \hat{F}$ המשערק כמתואר בטעיף 2.3.2 (בציוויל):
 (Max. Entropy (2) נופל ביצועים אך כמעט מהשערק F (בציוויל: (1)

3.3 רגישות שגיאת השחזר לטטיה מהפילוג הגאוטי

בטעיף זה נבחנת שגיאת השחזר (יחס אוזן לרעש) כאשר המקור אינו בעל פלוג גאוטי. נבחנים מקורות בעלי פלוג גמא ופלוג לפלט מחדר גיטא ומוקור בעלי פלוג אחד בלבד. הפלוג האחד שונה באורך שימושי באופיו מהשניים הראשונים (אשר מקובלים בספרות [9] כמודלים הסטטיסטיים של אותן דיבורים) ועל כן ניתן להניח שבדיקה הביצועים עבור שלושת הפלוגים הבנויים יש כדי "ילכוט" משפחת של פילוגים עשויה די הצורך.

לצורך חקמת תהליכיים אקראיים לא גausסיאנים בעלי זכרון ניתן למשל, להפעיל פונקציה לא ליניארית מתאימה (חטף זכרון) על תהליך גausס-מרקוב המתקבל בעזרת ביטוי 3.1 לעיל (ליתר פרוט ראה [10] וכן בטפוח א'). עבור כל אחד מהמקורות הנ"ל השונות נבחרה כך שתוצאות השתנות האפשרי יכטט את הקטע [ס+4, ס-4] בדומה לאופן בו הדבר נעשה עבור פלוג גausס.

בטבלה 3.2 מצוינו יחסאות אלו לשימוש שנקבלו לכל אחד מהמקורות הנ"ל עבור $\rho = 0.9$ (ערך אופייני לדבר קולי הדגום ב-KHz 8).

טבלה מס' 3.2 - יחס אותן לרעש (dB) עבור $\rho = 0.9$ ופלוגי בגיטה שונות.

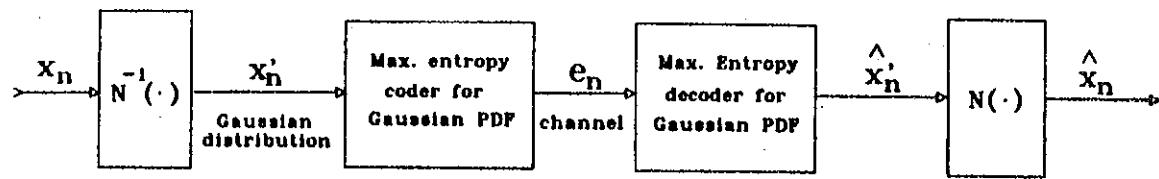
Table 3.2 - Signal-to-noise ratio (dB) for $\rho = 0.9$ and various input probability density functions.

חוק הסתגלות Distribution law	צפיפות הסתברות حد מילדי Scalar PDF	R = 1	R = 2
gauss Gauss	$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$	7.76	12.45
גמא Gamma	$p(x) = \frac{4\sqrt{3}}{\sqrt{8\pi\sigma x }} e^{-\frac{\sqrt{3} x }{2\sigma}}$	5.69	7.31
לפלס Laplace	$p(x) = \frac{1}{\sqrt{2\sigma}} e^{-\frac{\sqrt{2} x }{\sigma}}$	6.94	10.18
אחד Uniform	$p(x) = \begin{cases} \frac{1}{\sqrt{3}\sigma}, & x < \frac{\sqrt{3}}{2}\sigma \\ 0, & \text{אחרת} \end{cases}$	7.66	12.84

מהטבלה עולה כי הביצועים עבור פלוג אחיד דומים לאלו עבור פלוג גausס. לעומת זאת, עבור הפלוגים גמא ולפלס (בעיקר פלוג גמא) יש עליה שימושותית בשגיאת השחזור. דבר זה נובע ככל הנראה מתוਪעת המסלול האבולטי אשר מוצאת את ביטוייה ביתר שאת בפלוגים אלה שעבורם יש הסתברות גבוהה מאוד בסביבה הקרובה של הראשית. המסקנה לכך היא שיש לצפות לירידה לא מבוטלת ביחסותן לרעש אם מפעילים את המערכת על אותן דיבור בגל אופיו הלא-gausס.

3.4 גרטאות של המערכת עבור מקורות שאינם גaussiani

מהתוואות שהוצעו בסעיף 3.3 עולה כי יש להתאים את המערכת לפולוג הכניטה. בדרך המקובלות לבצע זאת היא להפעיל פונקציה לא לינארית חסרת זכרון בכניטה, שתפקידה לשנות את הפולוג הטוגולי בהתאם, בדומה לפעולת ה-companding במערכות בהן חמיוי אינו אחיד כמו law-law, A-law וכדי (ראה [9]). תיינוagiשה היא להפוך את פולוג הכניטה (גמא, לפט או אחיד) לפולוג גaussiי באמצעות פונקציה ליל' מתאימה (\cdot^{-1}) (ראה נושא אי'), לדוחות את האות המתקבל באמצעות ME ולהפעיל את הפונקציה הליל' הפוכה (\cdot) על האות המשוחזר, כמוראה בציור 3.3.



ציור מס' 3.3 - מערכת ME עבור פולוג לא גaussiי.

Fig. 3.3 - ME system for non-Gaussian distribution.

האות המשוחזר נתון אם כן ע"י:

$$\hat{x}_n = N(\hat{x}'_n) = N(E\{x'_n | e_n, e_{n-1}, y_{n-1}\}) \quad (3.6)$$

במקרה שהשערוד מתרצע בלאו-הcli בעזרת טבלאות קבועות, כמוואר בפרק 2, אז עדיף להפעיל את התוחלת על הפונקציה הלא לינארית, תיינו:

$$\hat{x}_n = E\{N(x'_n) | e_n, e_{n-1}, y_{n-1}\} = \frac{\int N(x) dx \int g(x,y) dy}{\int dx \int g(x,y) dy} \quad (3.7)$$

כאשר האינטגרציות מבוצעות על התוחמים המתאימים כמוואר בסעיף 2.3.

למקרה $R = 2$ מצאנו בפרק 2 עבור המשערוד \tilde{F} ביטויים "מפורשייט" של התוחלת המותנית של x_n . בפרט, ביטויים כאלה לא נוכל לבדוק למצוא עבור התוחלת המותנית של $(x'_n)N$

מאותר ש-(0) A היא פונקציה לא לינארית. לכן לשערך \tilde{F} יתאים הבטו (3.6) ולשערך F (המואחסן בלאו-הכי בטבלאות קבועות) נתאים את נוסחה (3.7).

במילים אחרות, נתבעין במשתנים הבאים:

גרסת א'

$$\hat{x}_n = N[\tilde{F}(e_n, e_{n-1}, y_{n-1})] = N(E\{x'_n | e_n, e_{n-1}^1, y_{n-1}\}) \quad (3.8)$$

גרסת ב'

$$\hat{x}_n = E\{N(x'_n) | e_n, e_{n-1}, y_{n-1}\} \quad (3.9)$$

הערה: עבור גרסת ב' יש להשRITE ביצור 3.3 את הבלוק הימני ביותר.

העדיפות של גרסת ב' מקורה בשני גורמים:

א. שתי הסיביות של e_{n-1} מנוצלות.

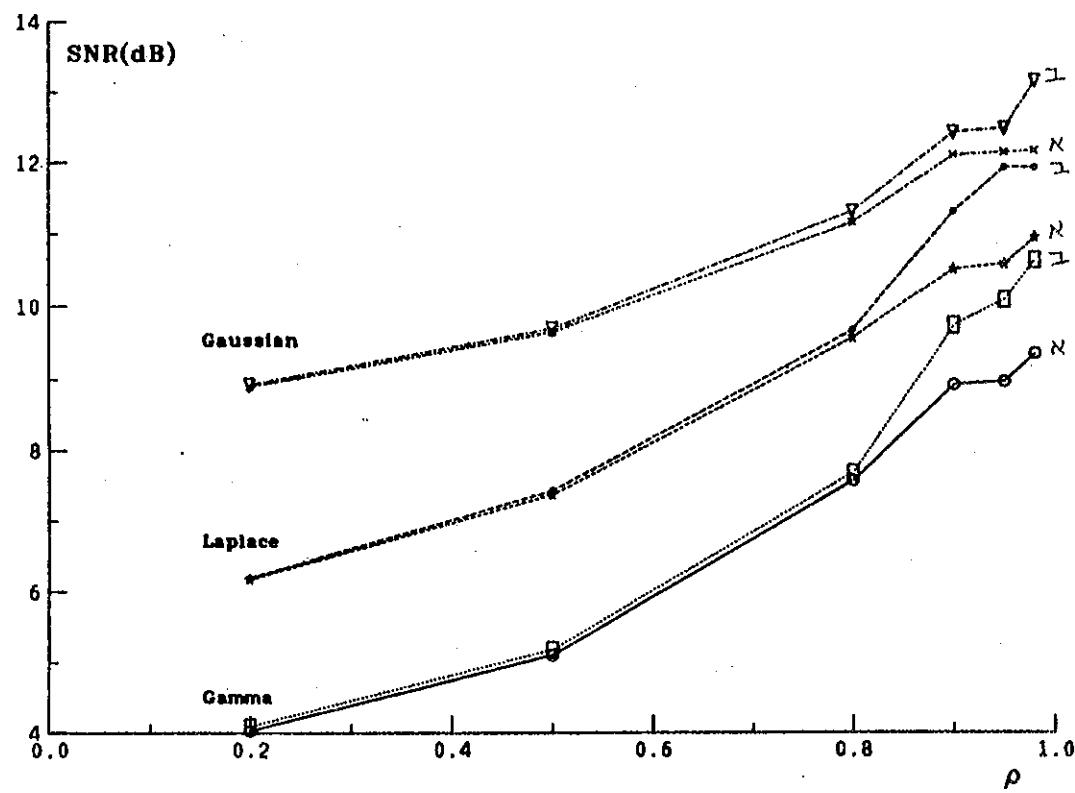
ב. בגרסת ב' מזעריט את ארגיגית ההפרש הרצוי ($\hat{x}_n - x$) בניגוד לגרסת א' בה מזעריט את ($x_n - \hat{x}_n$) כזכור המטריקה שונת.

邏輯, גרסת א' מאפשרת אדפטציה מאוחר שהיא מתאפשרת מפונקציה מפורשת של הפרמטרים ק ו-ס ועל כן היא מתאימה גם עבור $1 = R$.

ביצור 3.4 מトואר יחס אותן לרוש עבור $2 = R$ בשתי הגרסאות למקורות המפוגלים לפי גאות, גמא ולפלס כאשר כל אחד מהם מוזן ל מערכת המתאימה. (אות הוכנחת הווק בכל מקרה ע"י הפעלה פונקציה לא לינארית מתאימה על תחליק גאוס-מרקוב מתאים).

邏輯ו מטכבר שעבור הפלוגים בהם אנו מטענינו: גמא ולפלס, העדיפות של גרסת ב' על-פבי גרסת א' היא קטנה למדי עבור ערכיהם נומכיהם של ק אך גדרה שימושותית כאשר $1 + \kappa$. זאת מושט שהמידע שנושא הסיבית הנוספת בגרסת ב', כאשר ק מתרך ליחידה, גשיטת שימושותית יותר לABI שערוך והוויה מטור העבר. ברם, עבור הפלוג הגאוסי שבו מתקבלו יחסיהם לרוש האגובהים ביותר, ההבדל בין הגרסאות אינו גדול גם עבור ערכי ק המתאים לדיבור ($0.9 \approx \kappa$). משמעות תוצאות אלו היא שגרסת א' תモחתמת לפולוג גאוסי, עדיפה למעשה לגבי אמות דיבור משום שהיא מאפשרת אדפטציה ללא הפסד שימושותי ביחס אותן לרוש.

שתי הגרסאות של המערכת (א' ו-ב') עבור פלוגים לא גאוסיים מtbodyות למעשה על ההנחה שאות הוכנחת נוצר מטור הפעלה פונקציה לא לינארית חסרת זכרון על תחליק גאוסי. מובן כי זהו מודל שאינו בהכרח נאמן למהליך התפקיד של אותן דיבור. בפרק 4 נבחן, בין היתר, אם אכן מודל כזה משפר את הביצועים לגבי אותן דיבור.



ציור מס' 3.4 - ביצועי הממערכות המתוכננות לפולוגי כביטה שונות.

Fig. 3.4 - Performances of the systems designed for various distribution laws.

פרק 4 – ביצועי המערכת עבור מקורות לא טציזונרליים – אותן דיבור

בפרק זה תוצענה שתי גרסאות אדפטיביות של מערכת האנטרופיה המירבית ויבחנו ביצועיהן לאותות דיבור בהשוואה לגרסאות אדפטיביות של DPCM ו-LDM, הינו: CVSD ו-ADPCM בהתאם.

4.1 הצורך באדפטציה של המערכת

כפי שצוין בפרק 3, Daigle ו-Angel ([2], [3]) לא הצליחו גרסאות אדפטיביות עבור מערכת האנטרופיה המירבית אלא הפעילו את המערכת הקבועה על אותן דיבור ותמונות. ברם, במרבית היישומים המקוריים המקור שעומד לרשותינו אינו טציזונרי וקיים צורך לאפשר הסtagלות של המערכת לשינויים (אייטיליט) בפרמטרים הסטטיסטיים של המקור, כשזהה הדבר נעשה בגרסאות אדפטיביות של טכניקות קידוד צורת-גל מקובלות כמו ADPCM, ADM, CVSD וכיוצאין. לצורך זה נדרש מגנון אשר יעקוב אחר מהלך ההשתנות של ס-היחסים לעוצמת אותן אכניתה ושל ס – מקדמת הקורלציה בין הדגמים הסטטיסטיים שלו.

הבדל שמעותי במיוחד בין ערכי ס-היחסים לבין מוצאים בין דבר קולי (voiced) לבין דבר גטול-קול (unvoiced). עבור קצב דגימה של kHz 8 בדבר קולי ערכו של ס-היחס נע בערך בין 0.8 ל-0.95 בעוד שבדבר גטול-קול מתקבלים לעיתים אפלו ערכים שליליים. גם העוצמה – ס של אותן שונא באופן ניכר בשני המקדים. יתר על כן, המבדלים בין שני סוגי הדבר (כמו למשל מעברים מעוצר למונע ולהפך) עשויים להיות מהירים מאד ועל כן נדרשה אדפטציה מהירה מספיק שתחמכו אפקטים של טשטוש מחרומים ביניהם.

4.2 המערכת אדפטטיבית המוצעת

לצורך אדפטציה של המערכת יש לשערך את ס-היחס (או פון רקורסיבי) בכל יחידת זמן ולעדכן את פונקציות החזוי והשערך בהתאם.

4.2.1 אדפטציה של ההגבר

את הפרמטר ס מוצע לשערך באמצעות מגנון בקרת הגבר אדפטיבית (AGC) כאמור:

$$S_n = \lambda S_{n-1} + (1-\lambda) \tilde{x}_n^2 \quad (0 < \lambda \leq 1) \quad (4.1)$$

$$\hat{\sigma}_n = \sqrt{s_n} \quad (4.2)$$

כאשר:

\hat{x}_n - האות המשוחזר באינדקס הזמן n (ראה גם ציור מס' 1).

λ - מקדם "שכחיה" המציין את מהירות האדפטציה (ערכים מקובלים: $0.995 \leq \lambda \leq 0.95$).

\hat{s}_n - המשערך של s באינדקס הזמן n .

באופן כזה ניתן לנצל את אות הכניסה x_n ולשמור על תחום דינמי מתאים. את האות המשוחזר - \hat{x}_n יש לכפול שוב ב- \hat{s}_n . במשוואות (4.1), (4.2) מתוארת אדפטציה על פי האות המשוחזר (feedback adaptation) אשר יתרונה בכך שאין צורך בשדר של מידע נלווה (side information) ועל כן גם טగרווון לא נדרש. אם מבצעים, לעומת זאת, את האדפטציה לפי אות הכניסה (feed-forward adaptation) כאמור, \hat{x}_n מחליף את \hat{x}_n במשוואת (4.1), אז נדרש לשדר גם את \hat{s}_n במקביל כדי מידע נלווה, אך הריגשות לטעויות בערוצן קטנה יותר. מאחר שאנו עוסקים בדוחט בקצב שדור נתון, נתרכז באדפטציה על-פי האות המשוחזר. לצורך זה המשדר חייב להכיל העתק של המקלט לשם קבלת \hat{x}_n כמפורט בציור 4.1.

הकושי העיקרי של בקרת הגבר לפי האות המשוחזר מוצא את ביטויו בקטעי השקט בתמונת הפעת המסלול/global אשר אינה אפשרית ל- s_n לדעוך במידה מסוימת וגורמת להופעת צפופים באות המשוחזר. יש>Ifao צורך לאפשר מגננון אשר יזהה מסלול/global ויאלץ את s_n לקטונו בקטעים בהם האנרגיה נמוכה במיוחד. לצורך זה נדריך את מגננון בקרת הגבר באופן הבא:

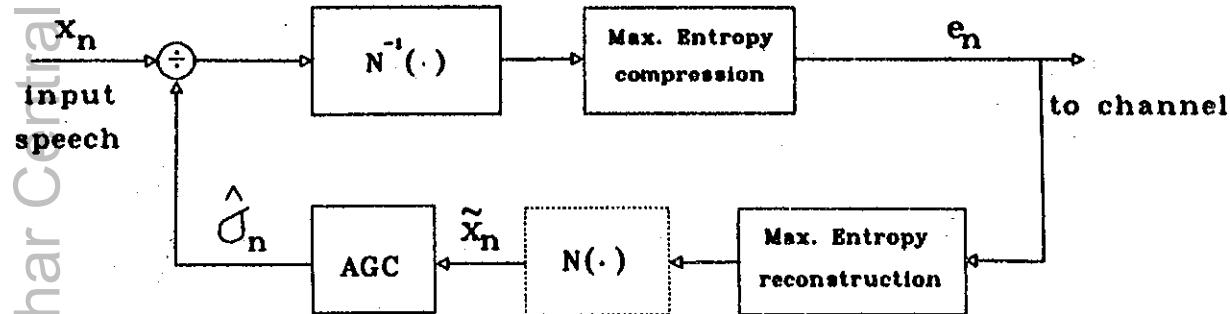
$$s_n = \begin{cases} \lambda_1 s_{n-1} + (1-\lambda_1) \hat{x}_n^2, & \text{המערכת אינה בתונה בمسلול-global,} \\ \lambda_2 s_{n-1}, & \text{המערכת בתונה בمسلול-global,} \end{cases} \quad (4.3)$$

כאשר λ_1 ו- λ_2 הם מספריים חיוביים הקטנים מ-1.

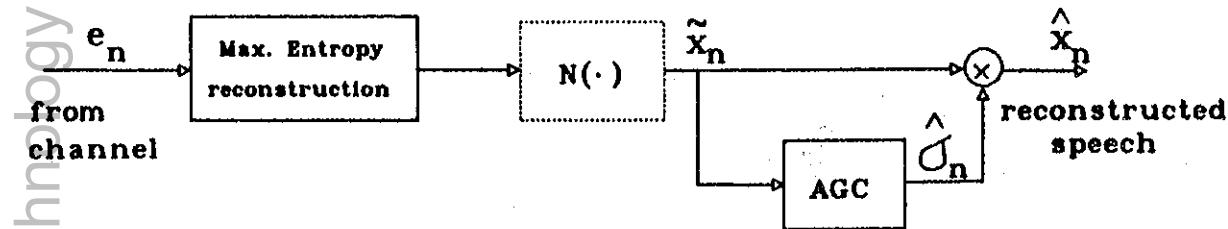
זהוי מסלול-global ניתן להציג ע"י התמונה המאפיינת שלו והוא הסימן המתחלף של (ראה סעיף 2.4) או לחילופין הקשר:

$$e_n = -\text{sign}(y_n) \quad (4.4)$$

אשר מקיימים גם עבור ערכים שליליים של y_n . מרגע, Ifao, לאבירץ על מסלול-global אם (4.4) מתקיים עבור שלושה ערכים עוקבים של y_n . גישת זו דומה לזה המקובל ב-ADM אשר עברו המכפיל (באנלוגיה - λ_2) קטן מ-1 אם שלושת מילוט חקוד האחרוןות לא כולם שוות.



(a) Transmitter



(b) Receiver

ציוויל מס' 4.1 - מושדר ומקלט עם בקרת הגבר מסתגלת. המלבן המקיף את הפעונציה "(·)N" משורטט בכו מרווח מסוים שקיים מותנה בגרסת השערוך (ראה פרק 3).

Fig. 4.1 - Transmitter and receiver with AGC. The rectangle that surrounds the function "N(·)" is dashed because its existence is dependent upon the estimation version (see chapter 3).

נעיר של אלגוריתם כזה מאפשר לקבל באופן טבעי קצב עלייה (λ_1) וקצב ירידה (λ_2) של ההגבר השונים זה מזה. גמישות כזו חיונית לנוכח משכי העלייה בעוצמת הקיצרים יחסית ומשci הירידה הארוכים אשר אופיינניים לדיבור אמייתי.

4.2.2 אדרפטציה של מקדם הקורלציה

המשערד המוצע \hat{c}_n של מקדם הקורלציה c ברגע n מתקבל כלהלן:

$$c_n = \lambda c_{n-1} + \tilde{x}_n \tilde{x}_{n-1} \quad (4.5)$$

$$s_n' = \lambda s_{n-1}' + \tilde{x}_n^2 \quad (4.6)$$

$$\hat{\rho}_n = \frac{c_n}{s_n'} \quad (4.7)$$

כאשר יש להגביל את ערכי \hat{c}_n לבב יחרגו מתחום $(-1, 1)$ ו- s_n' היא הגרסה של AGC כדומה ל-(4.1) – ללא זהוי מסלול גבולי.

עבור $2 = R$ אדרפטציה של c היא רלוונטיית אך ורק לגבי גרטה א' של השعروר אשר לפיה נעצרים בפונקציה F שתיא פונקציה מפורשת של c ועל כן ניתן להציב בה את \hat{c}_n העדכני לאותו רגע. (עבור $1 = R$ המשערד הוא גם כן פונקציה מפורשת של c). לעומת זאת, אופייני החזווי אינם פונקציות מפורשות של c לפיכך יש למצוא עבורם קירובים מתאימים.

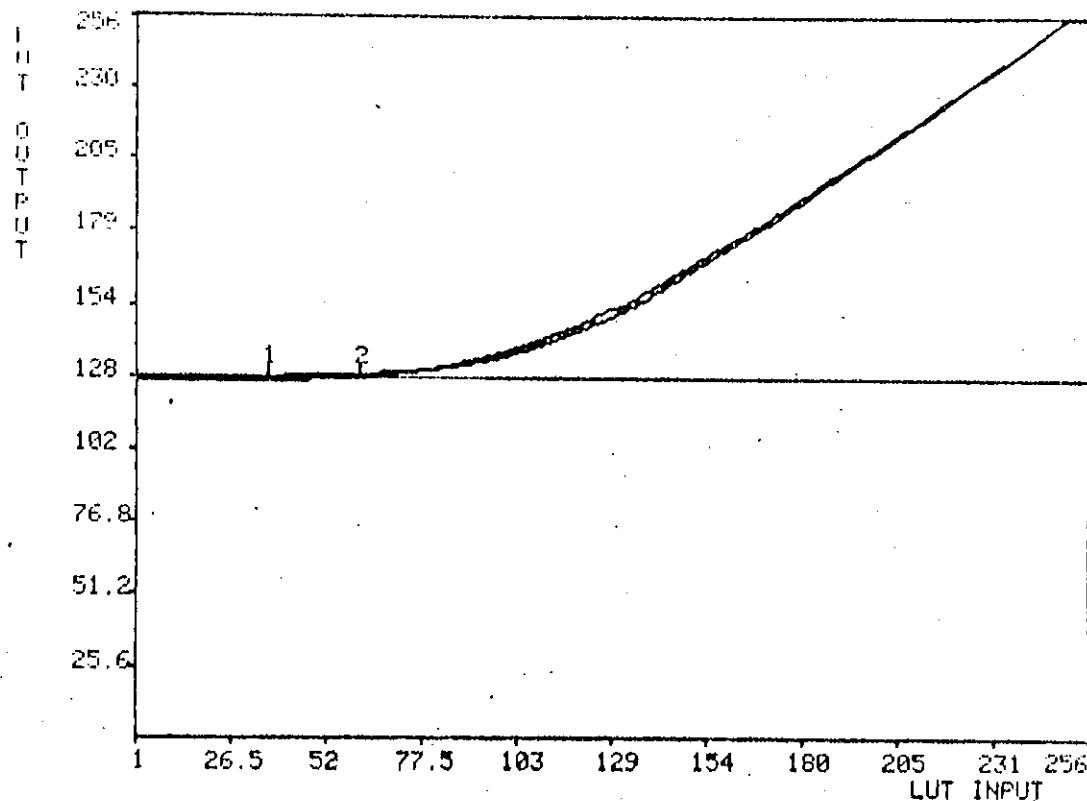
עבור החזיוון y , נזר בתוכנה מט' 3 (סעיף 2.4), בדבר האורי האטימפטוטי של אופייני החזווי, על פי עבור $\infty \rightarrow |y|$ לפונקציה $(y) P$ התנהגות לינארית במשתנה c . מכאן ומנוסחה (2.13) המבנת את הפונקציה P עבור המקרה $1 = c$ מוצע הקירוב הבא:

$$\tilde{P}(y) = \rho \sigma Q^{-1} \left[\frac{1}{2} Q\left(\frac{y}{\sigma}\right) \right] \quad (4.8)$$

MOVED CI L- (y) \tilde{P} אטימפטוטות זהות לאלו של $(y) P$ ועל-כן לפחות עבור ערכי $|y|$ גדולים הקירוב הוא נאמן. המגראעת של קירוב זה היא השימוש בפונקציה $(0) Q^{-1}$ שתהיא רגישה לשגיאות בארגומנט שלת בפרט כאשר הוא מתקרב לאפס או ליחידה, ועל-כן הטעלאות בהן יש להשתמש לצורך קבלת חיבות לאכיל מספר גדול של טיביות. קירוב אפשרי אחר לפונקציה P , אף הוא בעל התכונות האטימפטוטיות הרצויות, מתקבל מתוך החלפת החזיוון בתוחלת המותגנית המתאימה:

$$\hat{P}(y) = E\{x_{n+1} | x_n > y_n\} = \rho E\{x_n | x_n > y_n\} = \rho \sigma \frac{\exp(-y^2/2\sigma^2)}{\sqrt{2\pi} Q(y/\sigma)} \quad (4.9)$$

כעת מופיעה במכנה הפונקציה (4.0) שubahrah קיים קירוב מפורש שהוא נאמן למדי ולפיו שהגיאה אינה עולה על $7.5 \times 10^{-8} \pm 15$ (ראה נספח ב'). בציור (4.2) מתוארת הפונקציה P בהשוואה לפונקציה \hat{P} עבור $\rho = 1$.



ציור מס' 4.2 - הערך האמתי P והמקורב \hat{P} עבור $\rho = 1$.

Fig. 4.2 - The true predictor P and its approximation \hat{P} for $\rho = 1$.

להלן נסמן:

$$T(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi} Q(x)} \quad (4.10)$$

ונוכל לרשום את (4.9) בצורה:

$$\hat{P}(y) = \rho\sigma T\left(\frac{y}{\sigma}\right) \quad (4.11)$$

עבור $2 = R$ יש צורך בקירובים דומים גם עבור הfonקציית L ו- \hat{U} המוגדרות ב-(2.35) ו-(2.36) בהתאם. כעת דרישות אסימפטוטיות בהתאם ל-(2.48) – (2.51). הקירובייט המוצעים לצורך זה בעזרת הfonקציה (\cdot) T נתוניים ע"י:

$$\hat{L}(y,1) = -\hat{U}(-y,-1) = -\alpha\sigma + \rho\sigma T\left(\frac{y}{\sigma} - \alpha \frac{\sqrt{1-\rho^2} - 1}{\rho}\right) \quad (4.12)$$

$$\hat{U}(y,1) = -\hat{L}(-y,-1) = \alpha\sigma + \rho\sigma T\left[\frac{y}{\sigma} + \alpha \frac{\sqrt{1-\rho^2} - 1}{\rho}\right] \quad (4.13)$$

(4.12) ו-(4.13) מתקבלים ע"י מתאמת האסימפטוטיות של הfonקציה (\cdot) T לאסימפטוטות חדישות ע"פ (2.48) – (2.51) כאמור.

לצורך קבלת הערכה בדבר טיב הקירובים הנ"ל עבור $2 = R$, נבדקו האנטרופיות המתקבלות מהזנת תחלייך גאוס-מרקוב סטציונרי למערכת בעלת אופייני חזוי מקרובים נגד אלו. המתקבלות עבור אופייני חזוי המדוייקת. התוצאות מסומנות בטבלה מס' 4.1.

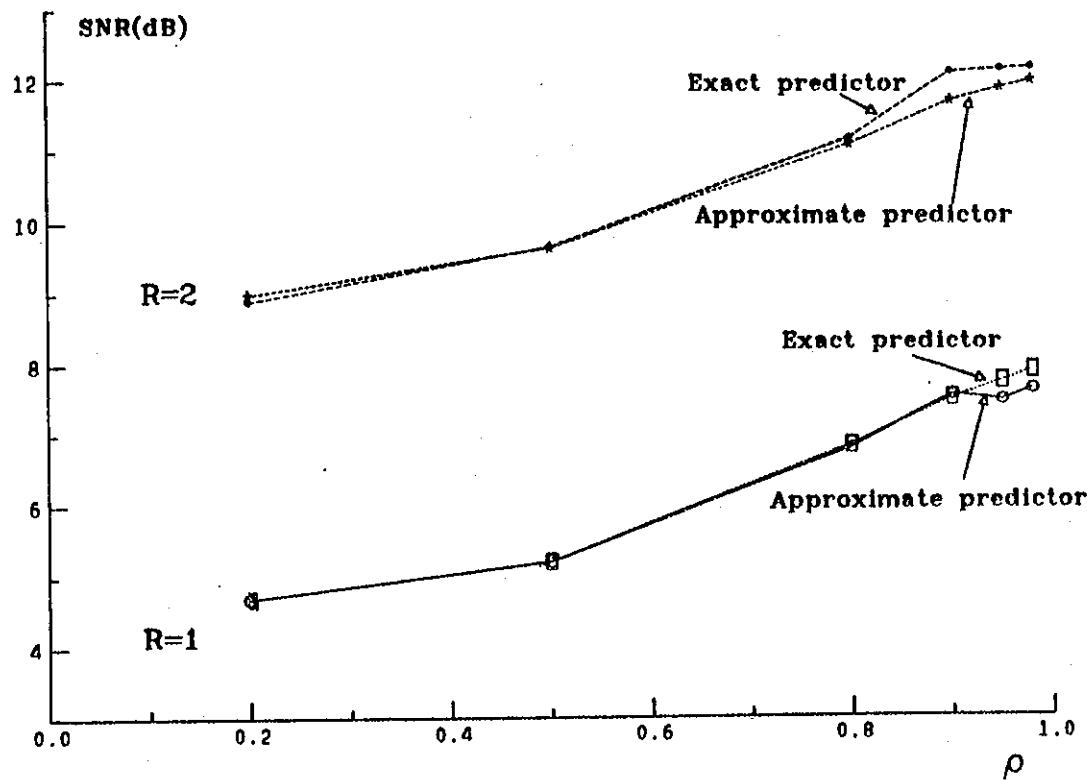
טבלה מס' 4.1 – האנטרופיות המתקבלות עבור החזאי המדוייק ועבור החזאי המקורב ($R = 2$).

Table 4.1 - The resulting entropies for the exact predictor and the approximated predictor (for $R = 2$).

ρ	חזאי מדוייק	חזאי מקרוב
0.2	1.99785	1.99815
0.5	1.99962	1.99944
0.8	1.99664	1.99622
0.9	1.99861	1.99380
0.95	1.99636	1.96873
0.98	1.99948	1.91009

מבחן ניתן לראות כי ככל ש- ρ גדול כן טיב הקירוב יוריד מאחר שההבדלים בין תוחלת מותגנית לבין חציון גודליהם יותר. אולם עבור ערכי ρ האופיליניות לדיבור ($0.9 - 0.8$) הביצועים של החזאי המקורב בהשוואה למדוייק הם מש比יעי רצון.

בציור מס' 4.3 מתואר יחס האות לרעש המתקבל (באותם תנאי סימולציה) בשחזרו של המערכת המקורבת (תיינו, זו המבוססת על חזאי מקורב) לעומת יחס האות לרעש של המערכת המדוייקת לפי גרטה שעורך א' (כאשר $x = (x)N$).



ציור מס' 4.3 - ביצועי המערכת עבור חזאי מקורב לעומת אלו עבור החזאי המדוייק.

Fig. 4.3 - System performance for the approximated predictor compared to that for the true predictor.

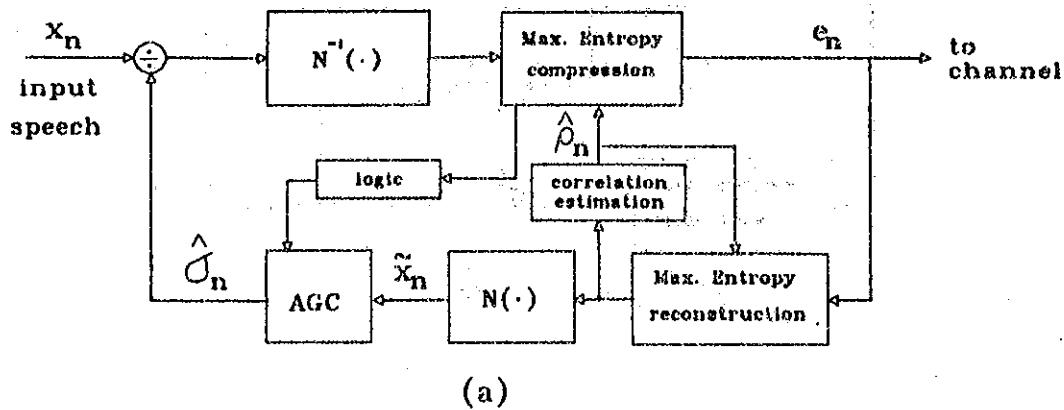
בדומה לנכפה מטבלאות (4.3) גם כאן ההפדר בביטויים גדול והורק ככל ש- σ גדול בעיקר עבור $2 = R$ - בתחום ערכי κ הרלוונטי לדיבור מתќבל הפדר של $C-BP1$.

קושי נוסף הכרוך באדפטציה של κ נוגע לטיבוכיות החישובית הגדולה יחסית ל-ADPCM ו-CVSD. מטיבה זו נבחן בטיעף הבא שתי גרטאות אדפטיביות של מערכת האנתרופיה המירבית:

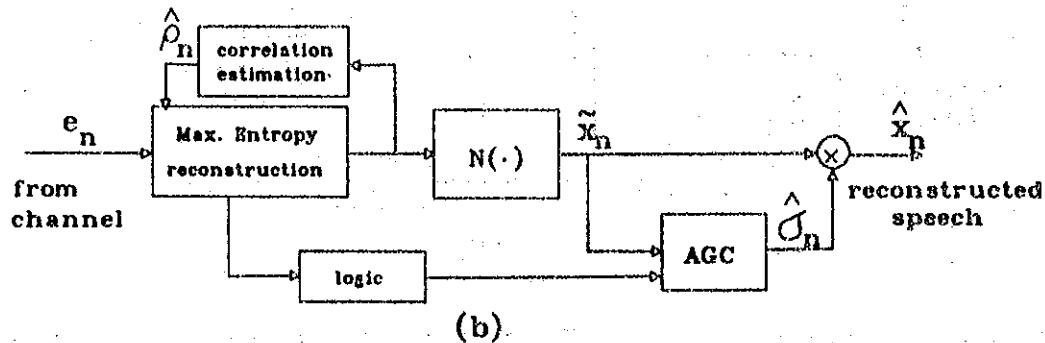
גרטה מס' 1 - אדפטציה של κ ושל σ על פי גרטת שעורך א' כמתואר לעיל.

גרטה מס' 2 - אדפטציה של σ בלבד תוך שימוש בטבלאות חזוי וشعורך מדוייקות ל- κ נומינלי לפי גרטת שעורך ב'.

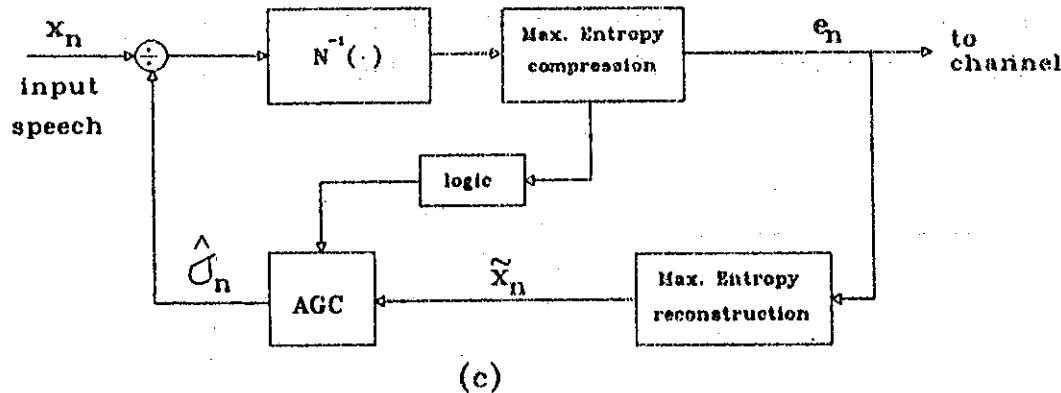
עבור גרטה מס' 2 דרוש AGC בלבד ועל-כן היא בעלת סיבוכיות חישובית דומה ל-ADPCM (ללא חזאי אדפטיבי). הטבלאות הנבחרות הן עבור ערך "אופייני" של κ (למשל 0.9). בציור מס' 4.4 מתוארים המקלט והמשדר עבור שתי הגרטאות האדפטיביות של ME.



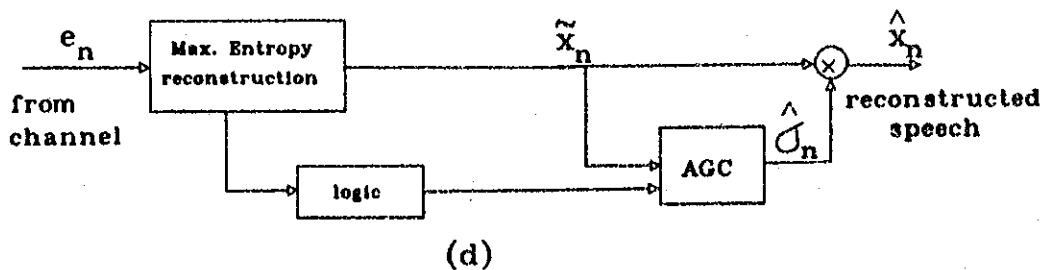
(a)



(b)



(c)



(d)

ציור מס' 4.4 - גרסאות אדפטטיביות של ME. גרטה מס' 1 : (a) - מועד, (b) - מקלט.

גרטה מס' 2 : (c) - מועד, (d) - מקלט.

Fig. 4.4 - Adaptive versions of ME. Version no. 1: (a) - transmitter, (b) - receiver. Version no. 2: (c) - transmitter, (d) - receiver.

4.3 ביצועי המערכת עבור אותות דיבור

בטעיף זה יתוארו ביצועי המערכות האדפטיביות של האנתרופופיה המירבית עבור אותות דיבור הדגומים בקצב של KHz 8 בקצב שדר של 16 kbit/sec ($R = 1$) ו- $R = 2$ (R = 2). הדיוון ייערך בלוווי משווה עם CVSD (עבור $R = 1$ ועם ADPCM בעל חזאי מסדר ראשון ($R = 2$)). לגבי המקהלה $R = 2$ נשווה את הביצועים של שתי הגרסאות האדפטיביות של ME אשר הוצעו בטיעוף 4.2 לעיל. עבור כל אחת מן הגרסאות האדפטיביות הניל נבחן את האפשרויות הבאות לבחירת הfonkcija הלא-לינארית ($\cdot(x)$ N :

$$.1. x = \cdot(x)N - \text{עבור פלוג לבנייה גאוטי.}$$

$$.2. \frac{1}{\sqrt{2}} \text{sign}(x) \ln(1 - |2Q(x) - 1|) = \cdot(x)N - \text{המתאים לפלוג לפולס.}$$

$$.3. \frac{1}{\sqrt{3}} x^2 \text{ sign}(x) = \cdot(x)N - \text{עבור פלוג גמא.}$$

כל אחת מהfonkcije הnil מעתקה לשטנה אקראי בעל פלוג גאוטי/פלפס/גמא (בהתאם) עם תוחלת אפס ושורנות יחידה, לשטנה אקראי גאוטי בעל אותם מומנטים (ראה גם נספח אי).

עבור קצב שדר של kbit/sec 8 בוצעה סימולציה על אותן דיבור הדגומים מקולו של גבר הקורא משפט אשר כולל בעל אופי קולי ("We were away a year ago"). הוצאות האובייקטיביות הטובות ביותר בשזרור של האנתרופופיה המירבית נתקבלו עבור $x = \cdot(x)N$ (המתאים למקור גאוטי) ומקדמי בקרת הגבר $\lambda_1 = 0.95$ ו- $\lambda_2 = 0.99$. לחס אותן לרעש הכלול (overall SNR) שבתקבל הוא 5.96dB וערך דומה - 5.9dB נמדד עבור יחס אותן לרעש הסגמנטלי (Segmental SNR) כאשר כל קטע (Segment) כולל 256 דגימות (32 msec). האנתרופופיה הממוצעת שנתקבל היא 0.99079.

מציאות דומות למדי נתקבלו עבור CVSD - יחס אותן לרעש: 6.23dB ו- 5.9dB בהתאם להאנתרופופיה - 0.99035. מנוקודת ראות סובייקטיבית, שני השחזרורים נשמעים דומים ולmedi והאנתרופופיה - 0.99035. מעשה, להציג על אחד מהם כעדיף. את שניהם מאפיינת התוחשה, שהקהל המשוחרר אמ衲 איננו מעוזת, אך טבול ברעש וקע רחב טרט בעצמה גבולה. האיכות בשני המקרים אינה משביעה רצון, מסיבה זו, אכן, לא מתקבל עלי-פי-רווב להפעיל את מערכת CVSD בקצב דגימה של KHz 8 אלא בקצבים גבוהים יותר בתמם מקובל札יה בין הדגמים הסטטיסטיות גדולות די חשוב.

עבור $R = 2$ נבחנו הביצועים, בנוטף לדבר הניל, גם לגבי אותן דיבור אחרים הדגומים (גם כנ-ב- KHz 8) מתקבליהם של גבר ושל אשה בקורסום שלושה משפטים כל אחד:

1. A lathe is a big tool.
2. Joe brought a young girl.
3. An icy wind raked the beach.

בנגוד למשפט הבודדים, משפטים אלו כוללים גם יחידות דיבור נטולות קול (unvoiced).

לABI גרטה מס' 1 של מערכת האנטרופיה המירבית (חכלה כאמור אדפטציה מלאה) ערכיו יחסאות לרעש הגבויים ביותר, בסימולציה לABI דבר שהזכר לעיל עברו $1 = R$, נתקבלו עברו הבחירה: $0.95 = \lambda_1$, $0.98 = \lambda_2$ ו- $x = (x)N$. יחסאות לרעש הכלול שנמדד עברו מקרה זה היה 9.83dB ואטגנטלי - 9.44dB . האנטרופיה הממוצעת שנתקבלה הייתה 1.981 .

לABI גרטה מס' 2 (המכילה בקרת ABI בלבד) נתקבלו יחסיותאות לרעש של 10.1dB ו- 10.2dB בהתאם לאנטרופיה של 1.971 סיביות, כאשר ערכו של κ היה 0.9 (ערכיו λ_1, λ_2 והפונקציה $(\cdot)N - \text{כ腻}$). עברו ADPCM, לעומת זאת, נתקבל יחסאות לרעש כולל של 12.9dB ויחס טגנטלי של 12.3dB . האנטרופיה הממוצעת שנמדדה עברו המידיע המשודר - 1.742 .

הערכתה הסובייקטיבית של אייקות השائز, למעשה הבדלים בין שתי הגרסאות של ME. בדומה כאמור לעיל לABI $1 = R$, גם כאן לשائز מלאוה רעש רקע ארבעצמתה הנמוכה באופן ניכר מזו שנתקבלה עברו $1 = R$. המובנות (intelligibility) של הדבר או נפגעת והצליל אינו מעוז. למרות העדריפות ב-SNR, השائز של ADPCM לא נמצא עדיף משמעותית על השائزים של ME - הוא מכיל אמנם מידת פחוותה במקצת של רעש רקע אך ההבדל בין לבינו ME זעום למדי.

כאשר נבדקו הביצועיםABI סדרת משפטי השניה אשר מכילה גם קטעי unvoiced העברו גרטה מס' 1 תוצאות עדירות במקצת - יחסאות לרעש טגנטלי של 11.24dB לעומת 10.35dB בגרטה מס' 2 (ModelPropertyים דומים לאלו נתקבלו גם עברו יחסאות גראטה מס' 2 אינה עברו המאזין ההבדלים בין שני השائزים אין גודלים אר נמצאו כי גראטה מס' 2 אינה משוחרת נסנה קטעי unvoiced, דבר הפוגע במובנות. כך למשל, העיצור /ch/ (צ') בamilah "beach" הנשמעת כמו "bee". תופעה זו מתרחשת הן בקולו של ABI והן בקולת השאה. בגרטה מס' 1, לעומת זאת, בעית זו אינה קיימת. עברו ADPCM התקבל משائز משפטיים אלו יחסאות לרעש טגנטלי של 14dB ויחס כולל של 13dB , אך בדומה לעיל, העדריפות הסובייקטיבית של ADPCM היא קטנה למדי.

התוצאות הנ"ל ABIABI ("... We were away") עומדות בנגוד לצפוי - עברו גרטה מס' 1 התקבלו יחסיותאות לרעש הנמוכים כמעט מלוא של גראטה מס' 2 למרות שזו חרושנה מכילה אדפטציה מלאה. יש לכך מספר הסברים אפשריים:

1. גרטה מס' 1 עדיפה על גרטה מס' 2 של ME עבור מקורות גאוט-מרקוב. לא מובטח כי כך הדבר גם לגבי אותות דיבור.

2. המשפט הביניל מכיל יחידות דיבור קוליות (voiced) בלבד, עבورو ערך קבוע של ρ כבגרטה מס' 2 (כמו $0.9 = \rho$), שהוא ערך אופייני, עשוי להיות עדיף על המשערך $\hat{\chi}$ כבגרטה מס' 1, המתוות תחיליך רועש (גם אם המקור סטציונרי) אשר אינו משקף נאמנה את התוצאות האיטיות לאחסית של ρ בתחליך יצירתיו של דיבור אמיתי. לגבי סדרת המשפטים השנייה, המכילה גם קטעי unvoiced, תוצאות גרטה מס' 1 נמצאו, כאמור, צפוי, עדיפות על אלו של גרטה מס' 2.

מהוצאות שהוצעו עבור ME עולה בין היתר, כי אין יתרון להשתמש בפונקציות חלא לינאריות (α) A עבור פילוגי גמא ולפלס - מהוצאות הטובות ביותר שנתקבלו היו בכל המקרים עבור $\alpha = (\alpha)A$. פרושו של דבר שהמודל של הפעלה פונקצייה לא לינארית חסרה צורך לשט קבלת פלוג טוגלי דמי דיבור על תחיליך גאוט-מרקוב אינו מתאים למנגנון ההפקה האמיתית של דיבור.

בנוסף, נצין את האנתרופיות הגבוהות המתקבלות מערכת ME לאות דיבור למרות שאינו תחיליך גאוטי. יחד עם זאת מתקבל מהוצאות שעיל שהאנתרופיה אינה תמיד גבוהה עם יחס התאות לרועש - בדומה לתוצאות שנתקבלו עבור תחליכים סטציונריים - האנתרופיה עבור ADPCM נמוכה יותר מאשר עבור ME, אבל יחס התאות לרועש גבוה יותר.

פרק 5 – חסיניות להפרעות בערוץ

בפרק זה נדוען בביצועי הגרסאות השונות של מערכת האנטרופיה המירבית מן ההיבט של מידת חסינות השחזרה להפרעות בערוץ התקשורתי כאשר אין משתמשים בזופני תקון שגיאות (צט"ש). מידת חסינות החישובית ומשאבי הצלרונו הנדרשים ע"י אלגוריתמי צפינת ערוץ וצט"שים מקובלים (זופני בלוקים, זופני קונבולוציה, פענוח ויטרבי, פענוח סדרתי וכו') עומדים לאין ערוך על אלו של צפינת המקור עצמו ושאזרו ולפיכך אין טעם רב לשלב צפינת ערוץ במערכות מסווג זה.

תחילת ניטוק בהיבט העיוני של השפעת טוות בודדת בערוץ על עצמת ההפרעה הנמстраה למקלט ונאמוד את משך התאוששותו ממנה. בהמשך, תוצגה תוצאות נסיגיות בדבר השפעת הסתברות הטוות בערוץ ביןארី חסר צלרוון על יחסאותיו לרעש שהזור מקורות טציאונריים ודיבור. בדומה לפרקים הקודמים, תערוך השוואה עם אלגוריתמים המקובלים לקידוד צורת-גל – CVSD ו-DPCM.

5.1 דיוון איקוני בהשפעת הפרעות בערוץ

ניתן דעתנו תחילת למערכת האנטרופיה המירבית הקבועה עבור $\frac{\text{טיבית}}{\text{dagm}} = R$. המשדר שלו מהווה מערכת לא לינארית בעלת משוב המRCTב ע"י אופין החזוי. לפיכך המכוניות העיקריות הכרוכות במעבר טוות בודדת בערוץ אל יציאת המקלט קבועות ע"י החזאי. נסמן ב- y_n את תפקת החזאי אשר במשדר וב- e_n את זו המתבלט במקלט ונניח כי ברגע $n = 0$ מתרחשת טוות בודדת בערוץ, כלומר:

$$\begin{aligned} y_0 &= y'_0 \quad ; \quad e_0 = -e'_0 \\ e_n &= e'_n \quad (\text{לכל } n \geq 0) \end{aligned}$$

כאשר $\{e_n\}$ מצין את המידע המשודר ו- $\{e'_n\}$ את זה התקלט. שגיאת החזאי – $y_n - y'_n$ – הנטראת למשערך עזרת "כווץ" עבור $n \geq 2$, כיוון שאופיני השערוך הוא פונקציות בעלות נגזרות אשר אינן עלות על מ- $(1 \leq |m|)$ – פונקציות "מכווצות" (כפי שהוזכר בסעיף 2.4). בגלל אופייה הרקורסיבי של פועלם החזוי, שגיאת החזאי y_n הngerמת מטעות ברגע $n = 0$ מוצאת את ביטוייה גם עבור m מניט עמידים $\dots, 1, 2, \dots = n$, אך היא דועכת בנסיבות כיוון שאופין החזאי (y_n גס הוא "מכווץ" שגיאות ב- n (שוב, כיוון שנגזרתו אינה עולה על m).

הודות לתכונת "כווץ" זו של אופין החזווי, מסוגל המקלט ל"התואוש" מטעות בודדת בערזץ ועל-כן גם מסדרת טעויות אשר תמרוחחים בינהן גדולים מטפิก, ככלומר, שכיחותן קטנה.

כאמור, לאחר שהנזרת המירבית של אופין החזווי אינה עולה על ρ , אז השגיאה ϵ_n מקיימת:

$$|\epsilon_{n+1}| \leq |\rho \epsilon_n| \quad (5.1)$$

על כן דעיכת שגיאת החזווי מהירה מדיעיכה אקספוננציאלית בעלת קבוע ρ .

$$|\epsilon_n| \leq |\epsilon_1|^{\rho^{n-1}} \quad (5.2)$$

ערך המירבי של ϵ_1 הוא בקירוב $\sigma \rho$ כיון שהתחום הדיינמי המירבי הוא $(-\sigma, \sigma)$ ולכן ההפרש המירבי של ערכי החזווי נתון ע"י:

$$M(4\sigma, 1) - M(4\sigma, -1) \approx 4\sigma\rho - 0 = 4\sigma\rho$$

לפיכך נוכל לחסום את שגיאת החזווי ϵ_n ע"י:

$$|\epsilon_n| \leq |4\sigma\rho|^n \quad (5.3)$$

שגיאת השערוך הנובעת מטעות הערזץ היא כאמור לכל היותר ϵ_n . מ-(5.3) נובע, שככל ש- ρ גדול יותר בערכו חnochלט, כן יש לצפות ממשר הפרעה ארוך יותר.

על פי (5.3) נוכל לחסום מלעיל את משך התהאוששות מהפרעה בודדת עבר $\rho = 0.9$. אם בנו כי $n = 128$ מתקדים כאשר ההפרש ביןיהם קטן מחייב סיבית, אז משך התהאוששות אינו עולה על:

$$\frac{\log(\frac{0.5}{128})}{\log 0.9} \approx 53$$

עבור קצב דגימה של 8 KHz תוצאה זו מתאימה ל- 6 msec בקירוב. מאידך, האורך המתוחלתי המירבי של השגיאה הוא גדול למדי - כמחצית המתחום הדיינמי. יש לצפות, איפוא, שטויות הערזץ ימסרו למקלט כפוליטים צרים וגבויים ויגרמו לאפקט חרומה באופיו לרעש "ברד", הנתרפם כמתוח ריריות (shot noise) ברקע האות המשוחזר.

עבור דלתא-מודולטור, לעומת זאת, אופין החזוי הוא לינארי - גורם ת"י כווז" שווה (בדיוון) ל- ρ ומשמעותה התחוליתית שווה לכפלים גודל הצעד - Δ , ועל-כן ההפרעה הנ מסרת במקצת זה בתונה ע"י:

$$\epsilon_n = 2\Delta\rho^n = 2\sigma \sqrt{2(1-\rho)} \rho^n \quad (5.4)$$

כעת משי' התחוששות מההפרעה רב יותר (כי קבוע הזמן שווה ל- ρ ולא חסום ע"י ρ כאמור) אך עצמה התחוליתית נמוכה יותר ועל כן השפעתו על האזנה של אותן המשוחזרות קטינה יותר.

הערכות:

1. עבור חזאי האנטרופיה המירבית ניתן לאופן לקבל חסם מלעיל דורך יותר של קצב התחוששות מהפרעה בודדת, בעזרת הנגזרת של הפונקציה $(y)^{\hat{P}}$ (המודרנת ב-(4.8)) כגבול העליון של התחום הדינמי $-\sigma = y$. במקרה זה מתבלט הנגזרת המירבית של אופין החזוי המקורב $(y)^{\hat{P}}$:

$$\hat{P}'(4\sigma) = 0.9284 \rho$$

ועל-כן נוכל לרשום מחדש את (5.3) באופן הבא:

$$|\epsilon_n| \leq |4\sigma| \cdot (0.9284 \rho)^n \quad (5.5)$$

2. עבור המקרה $2 = R$ הדיוון לעיל תקף אם הטעות היא בסביבה המשמעותית של תמלול המשודרת. אם הטעות היא רק בסביבה לשנית אז היא נמסרת למקלט אך ורק ברגע הופעתה ונעלמת מיד לאחר מכן כיוון שהיא לא מועברת למשוב החזוי.

3. עבור הגרטאות האדפטיביות ההפרעות הנ מסרות לשזהר גורמות גם לשגיאות בשערור σ ו- ρ . קצב הדיעיכה של שגיאות אלו תלויות ב- ρ ובקבועי הזמן של האדפטציה (λ_1 ו- λ_2). אילו היינו מאמצים גרטה של אדפטציה על פי אותן הנסיבות אז לא הייתה השפעה הדדית בין שגיאות האדפטציה לשגיאות השזהר וארגון היתה קטנה יותר.

4. קובל ערך בינהרי סימטרי חסר זכרון בעל הסתברות טעות - κ נתון ע"י:

$$C = 1 - H_b(p)$$

כאשר:

$$H_b(p) \triangleq -p \log p - (1-p) \log(1-p)$$

פונקציית קצב העוות של מקור גאוט-מרקוב מסדר ראשון נתונה ע"י:

$$R(D) = \frac{1}{2} \log_2 \left[\frac{\sigma^2}{D} (1-\rho^2) \right]$$

נסמן ב- ρ_s את מספר אותיות המקור לשניה וב- c את מספר אותיות הערוץ לשניה.
לפי משפט עיבוד הנתונים (ראה [4]) מתקיים:

$$R(D) \leq \frac{\rho_c}{\rho_s} \cdot c$$

כלומר:

$$\frac{1}{2} \log_2 \left[\frac{\sigma^2 (1-\rho^2)}{D} \right] \leq \frac{\rho_c}{\rho_s} [1 - H_b(p)] \quad (D \leq \sigma^2 \frac{1-\rho}{1+\rho})$$

הוא מספר אותיות הערוץ לאות המקור = R . מכאן נוכל לקבל חסם אוניברסלי עבור יחס האות לרעש בתנאים חרלוונטיים $\frac{\rho_c}{\rho_s}$

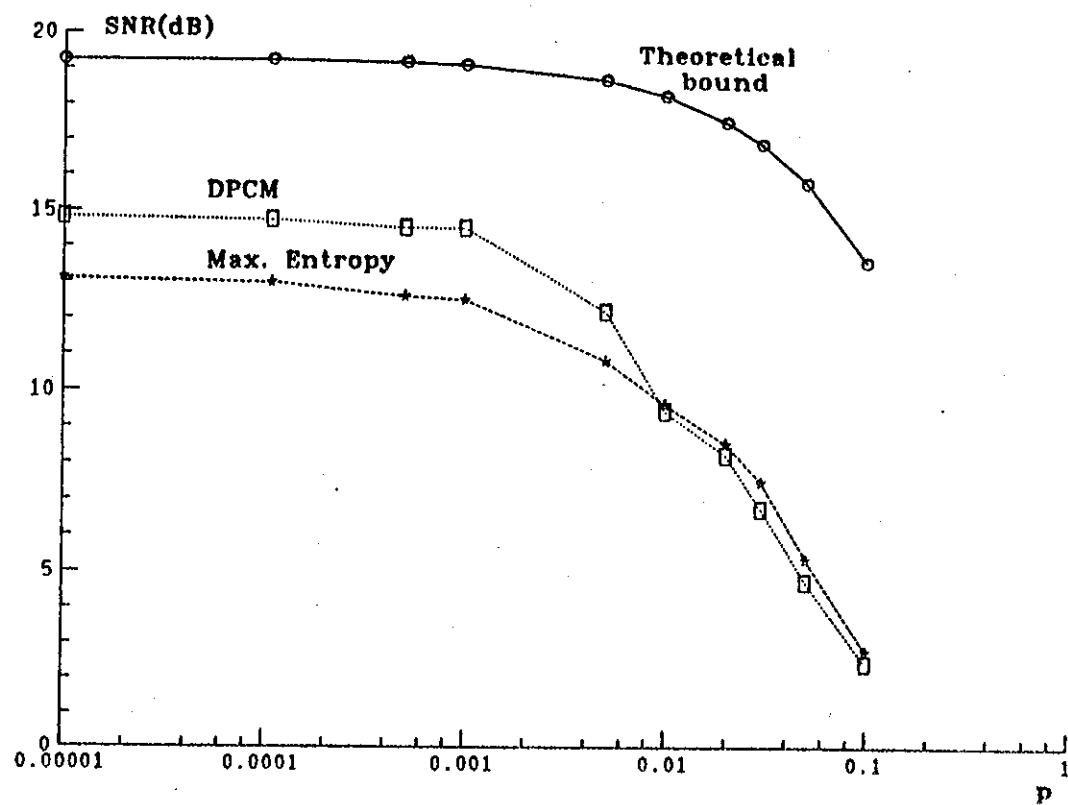
$$\text{SNR(dB)} \leq 6.02 R [1 - H_b(p)] - 10 \log_{10} (1 - \rho^2) \quad (5.6)$$

לחסם זה ניתן להתקרב את מבצעים צפינת המקור וצפינת הערוץ מתאימים. בסעיף הבא
נשווה בין השאר את התוצאות המתקבלות לחסם זה.

5.2 השפעת רעש הערוץ על שחזור אותות סטציונריאליים

בסעיף זה נתאר תוצאות גטווי הדמיה של הגרסת הלא-אדפטיבית של אנטרופיה מירביה
עבור $2 = R$, הפעלה בערוץ בינארי סימטרי חסר-זיכרון, בהשוואה ל-DPCM באותו תנאי.
לצורך הנגסו הופק תחילה גאוט-מרקוב מסדר ראשון של 4096 דגימות בעל מקדם קורלציה
 $0.9 = k$ והוזן למשדר. סיביות המידע המשודר הועברו דרך ערוץ בינארי סימטרי חסר
זיכרון בעל הסתברות טעות k . יחס האות לרעש בשחזר נמדד עבור ערכיהם השונים של k ,
בשתי המערכיות. כמו כן חושב החסם התיאורטי עבור יחס האות לרעש לכל ערך מתאים
של k על-פי משפט עיבוד הנתונים (5.6).

התוצאות שהתקבלו מוצגות בציור (5.1).



ציור 5.1 - יחס אות לרעש כמחלות בהסתברות הטעות עבור מחליכים סטציונריים.

Fig. 5.1 - SNR versus channel error probability for stationary processes.

מחזקאות ניתן בראש וראשונה לראות את הפרש הניכר בין החסם חמיורי לבין ביצועי המערכת, הנובע מכך שצפינית המקור בתוך אינה אופטימלית וצפינת ערוץ אינה מתבצעת כלל. בנוסף לכך, אנו עדים לתופעה שעבור ערכיהם קטנים של k , אמנים הביצועים של DPCM עדיפים (כפי שראינו בפרק 3 עבור $0 = k$), אך מעבר לערך מסוים מתחפכות היצירות ומתהמכת עדיפות למערכת ME. הסבר אפשרי לכך, נובע מן האמור בסעיף 5.1 לגבי זמני "התאוששות" טעויות ערוץ בודדות: ראיינו כי עבור DPCM זמן התאוששות ארוכים יותר מאשר עבור ME. מעבר לערך מסוים של k הזמן המוצע בין טעויות סמוכות קטן ממש התאוששות טעויות בודדות ושגיאות השזר הנגרמות מצטברות ומביאות לירידה תלולה יותר ביחס להוות לרעש.

המשמעות המוצע בין שתי טוויות סוכנות בעורץ נתון ע"י:

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} p \equiv \frac{1}{p}$$

משך התאוששות מטעות בודדת עבור DPCM ניתן לחישוב בנקל מתוך (5.4) וערכו עבור $0.9 = k$ הוא 38 דגמים (מתוך השוואת $\frac{1}{k}$ למחצית הסיבית). הערך המתאים של k הוא, איפוא: $0.026 = \frac{1}{38} = k$. המוצאות שביור (5.1) מצבויות על $0.01 = k$ בעל הנקודה בה לראשונה יחס האות לרעש ב-DPCM נופל מזה ב-ME. לעומת זאת הגדיל מתאים לחישוב (האגט) לעיל.

5.3 השפעת רעש העורץ על שחזור דיבור

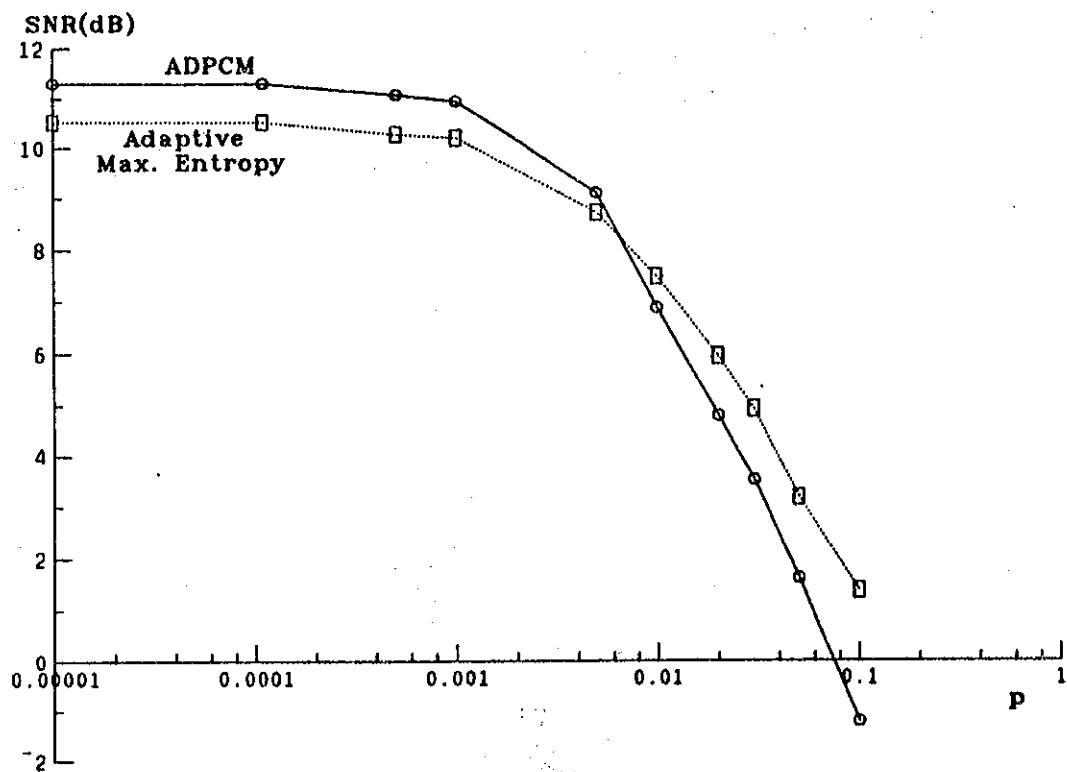
בטעיף זה נתאר תוצאות השוואת בין הגישה האדפטיבית של ME (AGC בלבד) לבין ADPCM בשתי סיביות לדגם עבור אותן דיבורים.

בדומה למקרה בטעיף 5.2, נערcta הדמיה של המערכות הניל בעורץ רועש לגבי דחיטה ושהזר של המשפט "We were away a year ago". עבור ערכיהם שונים של k בשתי המערכות חושב יחס אות לרעש טגמנטלי (SNRSEG), לפי קטעים בני 256 דגמים כ"א. התוצאות מוצגות בציור 5.2.

בתוצאות של עורץ רועש אלגוריתם עדכו גודל הצעד זוקן למנגנון "שכחיה" של טוויות כדי שמתואר במبدأ (ראה ביטוי (1.4)). בחירות ערד קטן מיחידה עבור β (במקרה זה - $0.98 = \beta$) פוגעת ביצועים כאשר העורץ אינו רועש. מטיבה זו נתקבל עבור ADPCM ערד הנמור ב- $2dB$ מזה שנתקבל בפרק 4 עבור $1 = \beta$. בדומה לתוצאות בטעיף 5.2, גם כאן מתרחשת תופעה, שבה החל מערך מסוים של k , ביצועי ADPCM נופלים мало של ME, והאטבר לכך דומה לזה שבטעיף הקודם. אולם, כתף ההשפעה של טוויות בעורץ היא מרכיבת יותר מאשר במערכות הקבועות מושם שגמרת, כאמור, הפרעה גם במנגנון עדכו גודל הצעד (או בקצת ההגבר) בנוסף לשגיאה בתפקיד החזאי.

הרעש הנמור מהעורץ למקלט ME נשמע כמעט יריות חדות (פולטים גבוהים וצריכים), ועבור ערכיהם גבוהים של k ($0.1, 0.05 = k$) עוצמת הרעש פוגעת במוגבלות. במקלט הרעש נשמע דומה, אך עמוס יותר באופיו ומפריע במידה פחותה כמעט לאוזנה.

למרות שהמדד האובייקטיבי (יחס אות לרעש) מצבע על רגישות גבוהה יותר ל-DPCM, הבדיקה הסובייקטיבית מגלת כי שחזור ADPCM עדיף במקצת ונוח יותר להאזנה.



ציור 5.2 - יחס אות לרעש כתלות בהסתברות טעות עבור דיבור.

Fig. 5.2 - SNR versus channel error probability for speech.

פרק 6 - סיכום ומסקנות

בעבודה זו נבחנת מערכת קידוד באנטרופיה מירבית שהוצאה בספרות, כנגד גרטאות של מערכת DPCM (הדומה לה במבנה העקרוני), מהיבטים שונים של ביצועים: חת אות לרעש, רגישות לטטיות שונות בפרמטרים ועמידות בפני הפרעות ערוץ. עבור אותן לא סטציונריאליים, הוצעו הרחבות אדפטיביות למערכת הקימת והן נבחנו לגבי דיבור ב- s/kb ו- $s/8$ ביחסוואה לגרטאות אדפטיביות של מערכות המכמי הדיפרנ- ציאליות הקימות באותו קצבי שידור. תשומת לב הוקשה, כמו-כן, לאיכות ולモבגנות של אותן המשוחזר ואופי העותאים והreaslers הנגרמים ע"י שגיאת השחזר במקלט.

התוצאות שנתקבלו מנוטילים של הדמית המערכות לא הצבעו, למעשה, על עדיפות כלשהי למערכת האנטרופיה המירבית במתכוונת בא תואר בחיבור זה. עבור מקורות גאטס-מרקוב סטציונריאליים מסדר ראשון, מתאימים להנחות המכנו של מערכת האנטרופיה המירבית, הטמינה אמנת עדיפות על פני DPCM עבור ערכיהם נוכחים של הקורלציה בין דגימות עוקבים. ברם, עבור ערכיהם גבוהים של ק, היוצרות נתפסו, ול-ADPCM נמצאה עדיפות ביחס אותן לרעש. הערכיהם האופלניים של מקדם הקורלציה עבור אותן דיבור הדגומות ב-KHz 8 מתאימים בקירוב לערך הגבול של מקדם הקורלציה אשר בו ביצועי שתי המערכות משתווים. לפיכך, ראיינו בכל זאת עניין להשוות בין המערכות עבור דיבור. מיידך, אותן דיבור איננו מלהיליך גאטס-מרקוב, ולכן התוצאות של ADPCM נמצאו טובות מעט יותר (3dB-2) מאשר אלו של המקדדים האדפטיביים באנטרופיה מירבית. מבחינת האיזות, נמצאו הבדלים קטנים למדי, אבל על-מנת להבטיח השזזור מתימן של עצורים מסוימים בדיבור, דבר חיוני לשימור המובגנות, ראיינו כי יש לאפשר אדפטציה של החזאי ואין להסתפק בברית הגבר (AGC) בלבד. אולם, האדפטציה של חזאי אינה תורמת באופן משמעותי להגדלת חת אותן לרעש, והיא מצריכה סיבוכיות חישובית גדולה בהשוואה ל-ADPCM.

אופיו הלא-لينארי של החזאי במערכת הנבחנת, תורם לעקבות טוביה באות המשוחזר כאשר יש שניים מתייררים באות אך הוא גורם לתנודות גדולות יחסית (idle channel noise) את אותן נמור או משתנה באיכות. נוצרים כמו-כן דפקים קצר-זמן אך גבוהים במשמעות כאשר מתרחשות טעויות בערזץ התקשורות. ההפרעות הנמרצות מרווח ADPCM, לעומת זאת, הן מושכות אך נוכחות יותר בעצמתן ולפיכך הן מושכות פחות על התזוזה במקלט.

מගענות נוספות של מערכת האנטרופיה המירבית שהוזכרו בפרק 2, הן שדרושות הנחות בדבר האופי הلطיטטי של המקור, וחסר גמישותה של המערכת לשינויים המכנו.

המקנה העיונית באשר לגישת של אנטרופיה מירבית, מתוך התוצאות עבור תהליכיים סטציונריים, היא שאין לצפות להתבדות בין אכونة האנטרופיה המירבית לבין מינימום שגיאה ריבועית מושעת, עבור מערכות בעלות צרכן המוגבל לצעד זמן אחד בלבד. תורת האינפורמציה מבטיחה את קיומן המשוחף של שתי תכונות אלו רק אם מבצעים כמו וקטורי אופטימלי בימד השווים לאינסוף, אך לא עבור מלמד סופי.

ידוע גם שכמוני סקלרי לפי Max ו-Lloyd, המבטיח שגיאת ריבועית מושעת מינימלית, איןבו בעל אנטרופיה מירבית. מכאן מתקשת הצעה פשוטה לשיפור המערכת באופן הבא: עבור כל אחד מהמצבים האפשריים π של החזאי (אשר מופיע סופי), יש לתקן את רמות הטעוי לצעד הזמן הבא $(t+1)$ על פי Max ו-Lloyd מתוך הפלוג הסגול של π^{t+1} המותנה במידע הזמן לחזאי מצעד הזמן הנוכחי - t , קרי: $\pi^t \rightarrow \pi^{t+1}$. כמובן מושעת מערכת לשיפור קידוד דיפרנציאלי על פי Max ו-Lloyd.

ובן כי גם המערכת זו יהיה אופינית לא לינאריות ותכונתיה יהיה כפוף להנחות טטיטיות לגבי המקור, אך יש לצפות שלפחות עבור תהליכיים סטציונריים, אשר להם היא מותאמת ת尽情 שגיאת שחזור קטנה יותר מאשר במערכת האנטרופיה המירבית. דרישות הזכור של מערכת זו דומות לאלו של מערכת האנטרופיה המירבית.

הצעות נוספות לשיפור מערכת האנטרופיה המירבית הועלו במהלך הדברים ונזכרים שוב בקצרה:

1. ニיצול הייאובטרוביולוגיי של משתנה המ痴 π לצורך הקטנת הקצב (סעיף 2.4).
2. שימוש במידע מצעד הזמן $(t-2)$ לצורך שיפור השערוך במחיר הכפלת גודל הזכרונו הדרוש למשערך (סעיף 2.2).
3. שיפור איטרטיבי של החזאי והתקינות לחזאי אופטימלי (סעיף 2.4).

אפשרויות אלו, שלא נבחנו במסגרת חבור זה, אך עשויות להביא לשיפורים בביצועי המערכת, יכולות להיות בסיס למחקר נוספת.

גופנו Ai – הפקת תהליכיים אקראיים לא אוטומטים בעלי זכרון

על מנת להפיק תהליך אקראי בעל זכרון ופלוג לא גausי (לפלט, גמא, אחיד וכיו"ב) מוצע ע"י Sondhi (ראה גם [10]) להפעיל אי ליניאריות חסרת זכרון (Ai) על תהליך המתקבל מפיגנו מתאים ("צביעה") של רוש לבן גausי. אי ליניאריות (Ai) נבחרת בהתאם לפלוג הרצוי על פי טענת העזר הבאה ([10]):

יהי x משתנה אקראי ותחי (x) פונקציה של פלוג שלו. אם (x) f_x גזירה אזי המשתנה האקראי (x) $= y$ מפולג אחיד על $[0,1]$. דבר זה נובע מהשקל הבא:

$$f_y(y) = \frac{f_x(x)}{|F'_x(x)|} = \frac{f_x(x)}{f_x(x)} = 1$$

ומשתנה בתחום $[0,1]$ ועל כן מפולג אחיד בתחום זה.

ע"י שימוש בעקרון זה ניתן לבנות פונקציות לא ליניאריות המunikות משתנה אקראי x בעל פלוג כלשהו F_1 למשתנה אקראי y בעל פלוג כלשהו F_2 דרך הפילוג האחד: $y = F_2^{-1}[F_1(x)]$. בהדרך זו מקבלים עבור פלוג אחיד בתחום $[-\sqrt{3}, +\sqrt{3}]$

$$N(x) = 2\sqrt{3} (Q(x) - \frac{1}{2}) \quad (A.1)$$

ועבור פלוג לפלאט

$$N(x) = -\frac{1}{\sqrt{2}} \operatorname{sign}(x) \log[1 - |2Q(x) - 1|] \quad (A.2)$$

פלוג גמא מתקבל ע"י אי ליניאריות:

$$N(x) = \frac{1}{\sqrt{2}} x^2 \operatorname{sign}(x) \quad (A.3)$$

(A.1), (A.2) ו-(A.3) מגדירים משתנה אקראי גausי בעל תוחלת אפס ושונות ייחודית למשתנים אקראיים, אף הם בממוצע אפס ופזרור יחידה, אך בפולוגים המתאימים.

על מנת לתכנן את המנגנון דרשו לדעת את הקשר בין מקדם תקורלציה k לפני האיליניאריות:

$$\rho_k \triangleq E \{x_n \cdot x_{n-k}\}$$

לבין מkładם הקורולציה ρ_k' אחדי האיליניאריות:

$$\rho_k' \triangleq E\{N(x_n) N(x_{n-k})\}$$

עבור פלוג גמא ופלוג אחד נמצאו בטויים מפורשים ([10]):

פלוג אחד:

$$\rho_k' = \frac{6}{\pi} \sin^{-1} \frac{\rho_k}{2} \quad (A.4)$$

פלוג גמא:

$$\rho_k' = \frac{2}{3\pi} [(1+2\rho_k^2) \sin^{-1} \rho_k + 3\rho_k \sqrt{1-\rho_k^2}] \quad (A.5)$$

במקרה של פלוג לפט לא נמצא בטוי מפורש לקשר הג"ל.

המגבלה של שיטה זו היא שלא כל סדרת אוטוקורולציה $\{\rho_k\}$ הרצiosa עבור התהיליך מתקבל, משליפה (לפי קשיים בדימה ל-A.4 ו-A.5) דרישת חוקיות לגבי הסדרה $\{\rho_k\}$ שאמורה אף היא לקיים תכונות של סדרת אוטוקורולציה (סדרה המגדירה מטריצת טופלייז חיובית חזי מוגדרת). ברם, מאחר שאנו עוסקים במערכות בעליות זכרון של יחידת זמן אחת, הדרישת מצטמת לקביעת האיבר הראשון ρ_1 בלבד. דרישת זו תמיד ניתן למלא ע"י חישוב הערך המתאים של ρ_1 והפקת מהליך גאוס-מרקוב מתאים לפניהם האיליניאריות.

נציין משפה נוספת של פונקציות לא-lienariות עבורן קיימות קשרים מפורשים בין ρ_k לבין ρ_1 , והוא משפחת החזקות האיליניאריות של x .

$$N(x) = x^n \quad (p \text{ טבעי}, 1 - p = n)$$

במקרה זה ρ_k הוא פולינום ממעלה n ב- ρ .

$$\rho_k' = B_n(\rho_k) = \sum_{i=1}^n \alpha_i^n \rho_k^i \quad (A.6)$$

ρ_k^n היא פונקציה אי-זוגית של ρ_k אם (α) N אף היא אי-זוגית. וע"י מתקדים $\alpha_{2j}^n = 0$ על-ידי שימוש בזהות ([10]):

$$\frac{\partial^2 g(x,y)}{\partial x \partial y} = \frac{\partial g(x,y)}{\partial y} \quad (A.7)$$

(כאשר g מוגדרת כמו ב-(2.5) עבור $1 = \sigma$) ו מביצוע אינטגרציה בחלקים מתקבל הקשר הבא:

$$\frac{\partial^2}{\partial \rho^2} \iint_{-\infty}^{\infty} x^{n+2} y^{n+2} g(x,y) dx dy = (n^2 + 3n + 2)^2 \iint_{-\infty}^{\infty} x^n y^n g(x,y) dx dy$$

הMOVIL לנוסחה הרקורסיבית:

$$\frac{\partial^2}{\partial \rho^2} B_{n+2}(\rho) = (n^2 + 3n + 2)^2 B_n(\rho)$$

מתוך תנאי הגבול: $0 = B_n(0)$ ו- $B_n(1) = 1$ מתקבלות משוואות המעבר הבאות:

$$\alpha_1^{n+2} = 1 - \frac{(n^2 + 3n + 2)^2}{(2n+1)(2n+3)} \sum_{k=1}^n \frac{\alpha_k^n}{(k+1)(k+2)} \quad (\text{A.8})$$

עבור $k = 3, 5, \dots, (n+2)$

$$\alpha_k^{n+2} = \frac{(n^2 + 3n + 2)^2}{(2n+1)(2n+3)} \cdot \frac{\alpha_{k-2}^n}{k(k-1)} \quad (\text{A.9})$$

תבאי תחילת: $\alpha_1^1 = 1$

$$\rho'_k = B_3(\rho_k) = 0.6\rho_k + 0.4\rho_k^3 \quad \text{דוגמא: עבור } x^3 = N(x) \text{ מקבלים}$$

הערה: במקרים בהם לא נמצאו ביטויים אנליטיים עבור הקשר בין ρ_1 ו- ρ_{-1} ניתן לחפיק את תנאי הרצוי ע"י מספר איטרציות בבתרת ρ .

נספח ב' - פונקציה מפורשת לקירוב הפלוג הנגוטי

פונקציית הפלוג - (x) Q ניתנת לקירוב ע"י פונקציה מפורשת באמצעות פיתוח לטורCDFקמן (ראה [1]) :

$$Q(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sum_{i=1}^5 b_i t^i + \epsilon(x) \quad (B.1)$$

כאשר:

$$t = \frac{1}{1 + px} \quad (B.2)$$

$$|\epsilon(x)| \leq 7.5 \cdot 10^{-8}$$

$$p = 0.2316419$$

$$b_1 = 0.319381530$$

$$b_2 = -0.356563782$$

$$b_3 = 1.781477937$$

$$b_4 = -1.821255978$$

$$b_5 = 1.330274429$$

נספח ג' ~ חוכחת לבטווי עבור מונת המשערך

מונת המשערך הוגדר ב-(2.23) כפונקציה הבאה:

$$R(t,s) = \int_s^\infty x dx \int_t^\infty g(x,y) dy \quad (C.1)$$

או

$$\begin{aligned} R(t,s) &= \int_s^\infty x g(x) dx \int_t^\infty g(y|x) dy = \int_s^\infty x g(x) Q\left(\frac{t - \rho x}{\sigma\sqrt{1-\rho^2}}\right) dx = \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_s^\infty x e^{-x^2/2\sigma^2} Q\left(\frac{t - \rho x}{\sigma\sqrt{1-\rho^2}}\right) dx \end{aligned} \quad (C.2)$$

נגידו:

$$U'(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \quad (C.3)$$

$$V(x) = Q\left(\frac{t - \rho x}{\sigma\sqrt{1-\rho^2}}\right) \quad (C.4)$$

אזי עיין אינטגרציה בחלוקת נקבל:

$$R(t,s) = \frac{\sigma}{\sqrt{2\pi}} \left[-e^{-x^2/2\sigma^2} Q\left(\frac{t - \rho x}{\sigma\sqrt{1-\rho^2}}\right) \Big|_s^\infty + \right. \quad (C.5)$$

$$\left. + \frac{\rho}{\sqrt{2\pi}\sigma^4(1-\rho^2)} \int_s^\infty e^{-x^2/2\sigma^2} \exp\left\{-\frac{(t - \rho x)^2}{2\sigma^2(1-\rho^2)}\right\} dx \right] =$$

$$= \frac{\sigma}{\sqrt{2\pi}} e^{-s^2/2\sigma^2} Q\left(\frac{t - \rho s}{\sigma\sqrt{1-\rho^2}}\right) + \rho \int_s^\infty g(x,t) dx =$$

$$= \frac{\sigma}{\sqrt{2\pi}} e^{-s^2/2\sigma^2} Q\left(\frac{t - \rho s}{\sigma\sqrt{1-\rho^2}}\right) + \rho \frac{\sigma}{\sqrt{2\pi}} g(t) \int_s^\infty g(x|t) dx =$$
$$= \frac{\sigma}{\sqrt{2\pi}} \left[e^{-s^2/2\sigma^2} Q\left(\frac{t - \rho s}{\sigma\sqrt{1-\rho^2}}\right) + \rho e^{-t^2/2\sigma^2} Q\left(\frac{s - \rho t}{\sigma\sqrt{1-\rho^2}}\right) \right]$$

REFERENCESרשימת מקורות

- [1] Abramowitz and Stegun, Handbook of Mathematical Functions. National Bureau of Standards, 1964.
- [2] E. Angel, L. Daigle and M. Rodriguez, A Maximum Entropy Encoder for Speech. IEEE ICASSP 1983 pp. 1292-1295.
- [3] E. Angel and L. Daigle, A High Speed Maximum Entropy Encoder for Images. ICASSP 1983, pp. 1236-1239.
- [4] T. Berger, Rate Distortion Theory. Prentice-Hall, Cliffs N.J. 1971.
- [5] J.L. Flanagan et al., Speech Coding. IEEE Transactions on Communications, Vol. COM-27 No. 4, April 1979.
- [6] J.D. Gibson, Adaptive Prediction in Speech Differential Encoding Systems. Proc. IEEE Vol. 68 No. 4, April 1980, pp. 488-525.
- [7] N.S. Jayant and P. Noll, Digital Coding of Waveforms. Englewood Cliffs N.J. Prentice-Hall 1984.
- [8] R.J. McEliece, The Theory of Information and Coding. Addison-Wesley, 1977.
- [9] L.R. Rabiner and R.W. Schaffer, Digital Speech Processing. Prentice-Hall, Inc. Englewood Cliffs N.J. 1978.
- [10] M.M. Sondhi, Random Processes with Specified Spectral Density and First Order Probability Density. Bell Systems Technical Journal Vol. 62 No. 3, March 1983, pp. 679-701.

ADAPTIVE MAXIMUM ENTROPY CODING
OF SPEECH SIGNALS

Research Thesis

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In Electrical Engineering

by

Neri Merhav

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

Heshvan 5746

HAIFA

November 1985

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

My sincere expression of gratitude to Professor D. Malah, for his devoted guidance and full involvement in all stages of this work.

Many thanks to Zippy Portnoy for the generous help at the Signal Processing Laboratory.

CONTENTS

	<u>Page</u>
ABSTRACT	1
LIST OF SYMBOLS AND ABBREVIATIONS.	2
CHAPTER 1 - INTRODUCTION	6
1.1 - Problem description.	6
1.2 - Conventional coding schemes.	7
1.3 - Research topics.	16
1.4 - Structure of the thesis.	17
CHAPTER 2 - THEORETICAL BACKGROUND FOR MAXIMUM ENTROPY CODING.	18
2.1 - The concept of Maximum Entropy Coding.	18
2.2 - System description	20
2.3 - Prediction and estimation characteristics.	22
2.3.1 - The predictor and the estimator for 1 bit/sample .	23
2.3.2 - The predictor and the estimator for 2 bits/sample.	31
2.4 - The properties of the system	38
CHAPTER 3 - THE PERFORMANCE OF THE SYSTEM FOR STATIONARY SOURCES . . .	45
3.1 - The dependence of the SNR and the entropy upon the correlation coefficient.	45
3.2 - The sensitivity of the reconstruction error to deviations of the correlation coefficient	50
3.3 - The sensitivity of the reconstruction error to deviations from the Gaussian distribution law	51
3.4 - Versions of the system for non-Gaussian sources.	53
CHAPTER 4 - THE PERFORMANCE OF THE SYSTEM FOR NONSTATIONARY SOURCES - SPEECH SIGNALS	56
4.1 - The necessity of adaptation.	56
4.2 - The proposed adaptive scheme	56
4.2.1 - Adaptation of the gain	56
4.2.2 - Adaptation of the correlation coefficient.	59
4.3 - The performance of the system for speech signals	65

CONTENTS (Cont.)

	<u>Page</u>
CHAPTER 5 - THE IMMUNITY TO CHANNEL ERRORS	68
5.1 - The influence of channel errors - a general discussion . .	68
5.2 - The affect of channel noise on the reconstruction of stationary signals	71
5.3 - The affect of channel noise on the reconstruction of speech signals	73
CHAPTER 6 - SUMMARY AND CONCLUSIONS.	75
Appendix A : The production of non-Gaussian correlated random processes	77
Appendix B : An explicit function that approximates the error function	80
Appendix C : A proof for the expression of the numerator of the estimator	81
REFERENCES	83
ABSTRACT (English)	I

A B S T R A C T

In this research we examine a recently proposed waveform coding scheme, which is based upon maximizing the entropy of the transmitted bit-stream. The theoretical motive for this research is driven by the non-linear prediction required to realize the system, and the investigation of the maximum entropy criterion as a substitute for the widely used minimum mean square error criterion. The simple implementation of the fast scheme proposed, provides the practical motive for this work. However, system design requires some prior statistical assumptions, which might limit the class of sources for which the system is suitable.

The performance of the system is compared to that of existing systems for stationary processes in several aspects: Signal-to-noise ratio, sensitivity to deviations in the statistical features of the source, and the degree of immunity to transmission errors. Adaptive versions of the scheme of 8 kbit/sec and 16 kbit/sec are suggested for speech signals. The quality and intelligibility of the resulting reconstructed signals are examined in addition to the above measures.

The non-adaptive maximum entropy scheme is already known in the literature. Yet, to the best knowledge of the author, has not been tested for stationary sources, rather directly for speech signals and images, for which no adaptive maximum entropy schemes have been proposed so far.

In fact, simulation results using objective performance measures (e.g. SNR), reveal no advantage in applying the maximum entropy system in its present configuration, over conventional waveform coders (such as ADM and ADPCM). On the contrary, prominent preference to the latter is noticed for highly correlated signals. However, considering the quality of the reproduced speech, only a slight superiority of the conventional schemes is observed. To preserve the intelligibility, a full adaptation of the maximum entropy system is needed, and consequently, an extreme increase in computational complexity compared to existing systems, is observed.

Nevertheless, several improved alternatives are recommended for future research.