אוקטובר 2011 – פברואר 2012

מיכאל מולוצ'ניקוב

תורת האינפורמציה 1
סיכום הרצאות

מבוסס על הרצאותיו של פרופ' רמי זמיר

חוברת זו הינה סיכום אישי של המחבר שנכח בהרצאות.

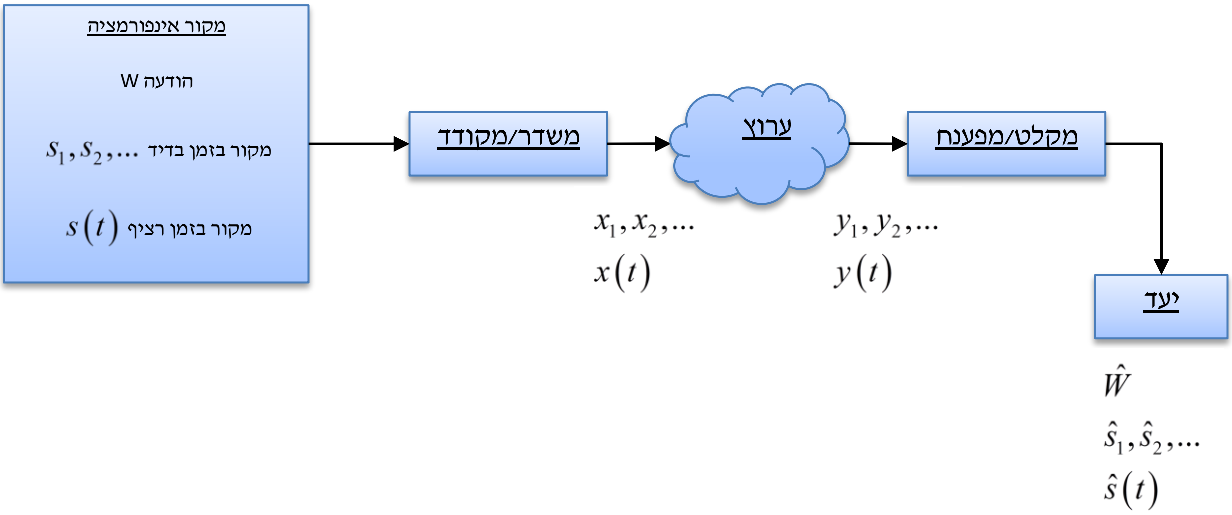
אין המחבר מתחייב לנכונות המידע בחוברת זו.

חוברת זו מהווה חוברת עזר להרצאות ואין היא מהווה תחליף אליהן!

תוכן העניינים

1. הקדמה

1.1. תרחיש התקשורת הבסיסי



1.2. דוגמאות

- ✓ מקורות:
  "קובץ" – מידע דיגיטלי, דיבור או מוסיקה, תמונה או וידאו
- ✓ ערוצים:
  קו טלפון, כבלים, אלחוטי (Wi-Fi), סלולארי, לוויינים), התקני זיכרון – DOK, HD, Tape.
- ✓ מגבלות וקריטריונים:
  - הגבלת הספק על הכניסה לערוץ: $Ex^2(t) \le P_{\max}$
  - הגבלה על ביצועי תקינות המידע (SNR ו-BER): $P\{\hat{s}_n \ne s_n\}, P\{\hat{W} \ne W\}, d(\hat{s}_n, s_n)$

  כאשר אלה מסמנים קריטריונים עבור ביצועי ערוץ התקשורת.

1.3. מה זו אינפורמציה?

אינפורמציה היא מידה של אי-ודאות – של הפתעה. כלומר, ככל שערך מסויים של משתנה הוא לא צפוי (מפתיע), כך ערך זה מכיל יותר אינפורמציה. לצורך המחשה, נביט בדוגמא הבאה: נגיד ואנו שואלים מישהו 2 שאלות לגבי הילד שלו: "האם זה בן או בת?" ו"מה שמו?" ומקבלים את התשובה "דניאל". "דניאל" הוא שם סביר גם לבן וגם לבת באותה המידה. כלומר, האי-ודאות במקרה זה יותר גבוהה מהתשובה לשאלה "בן או בת".

$$Info\{\text{son or daughter}\} < Info\{\text{What's the name? Daniel}\}$$

בנוסף, מתקיים:

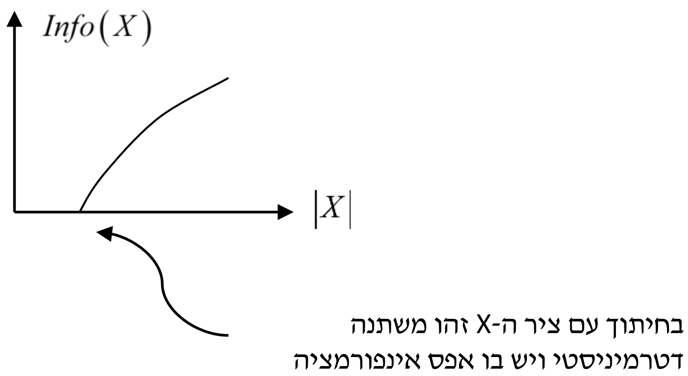
$$Info\{\text{s/d, name}\} \underset{\text{statistically independent}}{=} Info\{\text{s/d}\} + Info\{\text{name}\}$$

1.4. תכונות רצויות ממדד אינפורמציה

א. מתייחס למשתנים אקראיים: משתנים בדידים, כלומר א"ב סופי -

$$\left(X,Y\right)\sim P\left(x,y\right)\ ,\ X\sim P\left(x\right);x\in X$$

ב. עבור מ"א שמתפלג אחיד $X\sim unif\left(x\right)$, האינפורמציה עולה מונוטונית עם גודל ה-א"ב



ג. <u>אדיטיביות</u>:

האינפורמציה הכוללת של מספר מ"א בת"ס היא סכום האינפורמציות:

$$Info\left\{X,Y\right\}\underset{X\perp Y}{=}Info\left\{X\right\}+Info\left\{Y\right\}$$

ד. <u>עידון (refinement, grouping)</u>:

אם Y נותן פירוט ביחס ל- X (כמובן שהוא תלוי בו) אזי:

$$Info\left\{X,Y\right\}=Info\left\{X\right\}+\sum_{x}P\left(x\right)\cdot Info\left\{Y\mid X\right\}$$

מהחוקים הנ"ל ניתן להראות:

$$H\triangleq entropy\triangleq H\left(X\right)\triangleq H\left(P\left(x\right)\right)\triangleq\sum_{x\in X}P\left(x\right)\log\left(\frac{1}{P\left(x\right)}\right)$$

זהו גודל המקיים את כל התכונות הרצויות ממדד של אינפורמציה.

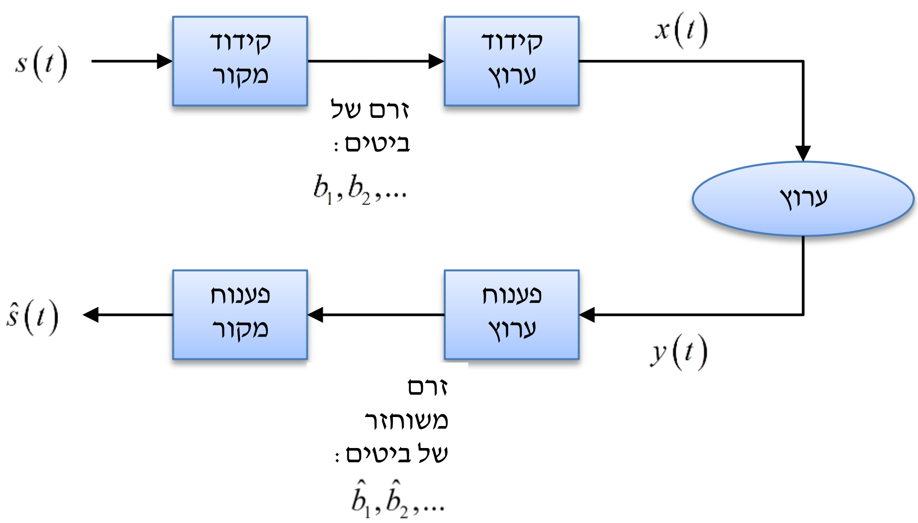
1.5. עיקרון ההפרדה

באופן כללי, המשדר (או המקודד) צריך להמיר את אות המקור $s(t)$ לאות בכניסת הערוץ $x(t)$. קיימות מספר סיבות להמרה זו:

א. התאמת אות המקור לתווך השידור.
ב. קיום אילוצים (הספק, תחום תדרים וכו'...)
ג. שיפור החסינות לרעש.

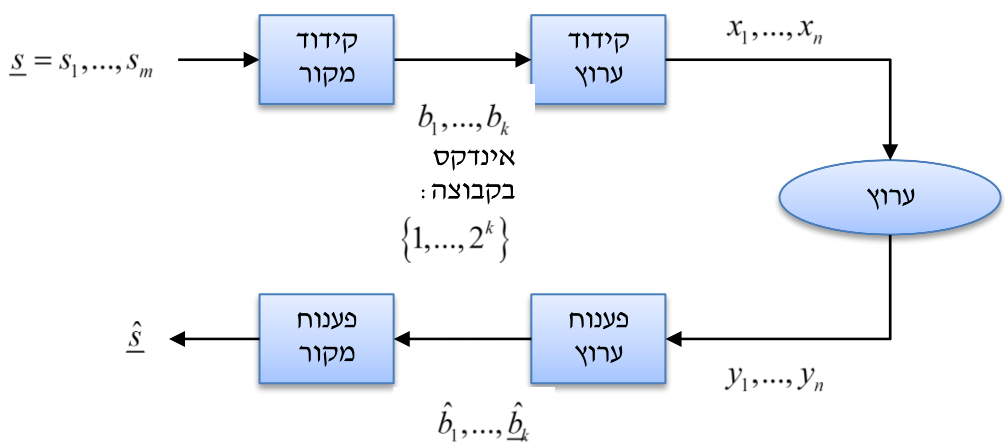
שאנון הראה ש:

"ניתן לפרק את בעיית התקשורת הכללית לקידוד מקור וקידוד ערוץ"

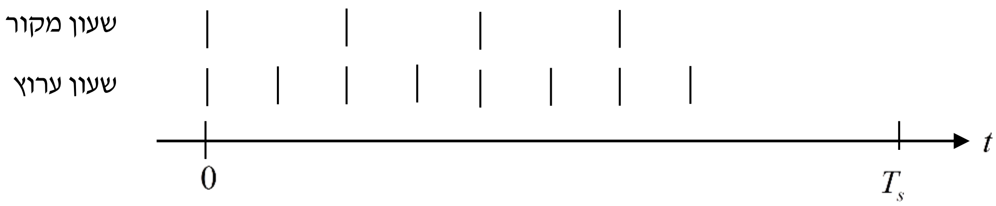


- מקודד/מפענח המקור לא תלוי במאפייני הערוץ.
- מקודד/מפענח הערוץ לא תלוי במאפייני המקור.
- מטרת מקודד המקור היא "לדחוס" אותו לקצב הביטים <u>המינימלי</u> האפשרי שעדיין מקיים את הקריטריון הרצוי (עיוות או הסתברות שגיאה).
- מטרת מקודד הערוץ היא להעביר ביטים דרך הערוץ בקצב <u>המקסימלי</u> האפשרי שעדיין מאפשר פענוח תקין שלהם.

✓ בד"כ נעבוד עם <u>מודל בזמן בדיד</u> (ביחס לקידוד של בלוק באורך סופי):



אנו יכולים להבחין כי קידוד המקור מתאים k ביטי קוד ל- m דגימות מקור, וקידוד הערוץ מתאים וקטור x^n באורך n ל- k ביטי מקור. אם נסתכל על זה מבחינת קצב השעונים של המקור והערוץ, נראה כי הקצב של שעון הערוץ שונה משעון המקור (אלא אם $m=n$):



1.6. משפט הקידוד של שאנון

עבור מקור נתון $(s_1, s_2, \ldots)$ קיימת מידת אינפורמציה H שהינה פונקציה של המאפיינים הסטטיסטיים של המקור. עבור ערוץ נתון (מעבר מכניסה X למוצא Y) קיים קיבול C שהינו פונקצייה של המאפיינים הסטטיסטיים של הערוץ.

- ✓ אם $H \le C$ אזי ניתן להעביר את המקור דרך הערוץ עם הסתברות שגיאה קטנה כרצונינו.
- ✓ אם $H > C$ אזי לא ניתן לבצע את הנאמר לעיל.

כלומר, H הינו קצב הדחיסה המינימלי האפשרי של המקור, ו- C הינו קצב השידור המקסימלי האפשרי דרך הערוץ.

- ✓ עבור מקור רציף (בערכים של דגימותיו ולא בזמן) מחליפים את H בפונקציית קצב עיוות $R(D)$.

1.7. דוגמאות לקידוד מקור

א. קידוד בינארי: קיימות 2 מילים ב-א״ב: $\{u_0, u_1\}$ והמיפוי הוא: $u_0 \to \text{'0'},\quad u_1 \to \text{'1'}$.

ב. קידוד עשרוני: הא״ב הוא: $X = \{0, \ldots, 9\}$. ניתן לייצג כל מילה ע״י 4 ביטים (כיוון ש- $|X| = 10$), וניתן לעשות זאת בצורה הבאה:

$$0 \to 0000$$
$$1 \to 0001$$
$$\ldots$$
$$9 \to 1001$$

ע״פ הגדרה שתוגדר בהמשך, קצב הקוד הוא $R = 4\,bit/sample$.

ג. ננסה לשפר את הקידוד הקודם. נניח כי המקור פולט את המילים בצורה רציפה, למשל ...,3,7,0,9,2,5. נחלק את המילים הנפלטות לקבוצות של 3: ... ,5 ,2 ,9, 0 ,7 ,3. מספר המילים שניתנות לייצוג בצורה זו הוא $|X^3| = 1000$ ולצורך כך נדרשים 10 ביטים. כך במקום להשתמש ב-12 ביטים השתמשנו רק ב-10.

- ✓ קצב הקידוד:

$$R = \frac{\text{amount of bits to be decoded}}{\text{length of the source vector}}$$

במקרה של דוגמא ג׳:

$$R = \frac{10}{3} = 3.33\,bits/sample$$

✓ <u>אסימפטוטות באורך בלוק המקור</u>:

עבור וקטור מקור באורך m, ניתן לייצגו ע"י $\left\lceil \log_2 \left| X^m \right| \right\rceil$ ביטים. ולכן הקצב:

$$R=\frac{\left\lceil m\cdot\log_2|X|\right\rceil}{m}\le\frac{\left(m\cdot\log_2|X|+1\right)}{m}\xrightarrow[m\to\infty]{}\log_2|X|$$

כלומר, עבור בלוקים גדולים דיים הקצב יהיה:

$$R=\log_2(10)=3.32\,bits/sample$$

ד. <u>מירוץ סוסים</u>:

לסוס 1 קיים סיכוי לזכות $P=1/2$, סוס 2: $P=1/4$, סוס 3: $P=1/8$, סוס 4: $P=1/16$ וסוסים 5-8: $P=1/64$. צריך לקבוע קוד יעיל להודעה איזה סוס יזכה במירוץ.

למשל, אם נבחר באורך הודעה קבוע, הקצב יהיה: $R=3bits/message$ (זאת כיוון שקיימים 8 סוסים). שיטה יותר טובה היא לתת להודעה עם ההסתברות הגבוהה פחות ביטים (יותר צפוי שהסוס הזה יזכה אז מאורע זה מכיל פחות אינפורמציה). נבצע את הקידוד בדרך הבאה:

$1\to'0'$

$2\to'10'$

$3\to'110'$

$4\to'1110'$

$5-8\to'1111**'$

$$R=E\left[length(X)\right]=\frac{1}{2}\cdot1+\frac{1}{4}\cdot2+\frac{1}{8}\cdot3+\frac{1}{16}\cdot4+\frac{1}{16}\cdot6=2\,{bits}/{message}$$

זהו קידוד באורך משתנה ואנו רואים כי הוא יותר טוב מקידוד באורך קבוע. ניתן לוודא זאת עם חישוב האנטרופיה:

$$H\left(\frac{1}{2},\frac{1}{4},\frac{1}{8},\frac{1}{16},...,\frac{1}{64}\right)=...=2bit$$

ה. <u>מקור ברנולי</u>:

כאשר ת"א X_n מפולג ברנולי: $X_n\sim Bernoulli(p)$, הכוונה היא שזהו מקור בינארי iid, ומתקיים: $P\{X_n=1\}=p,\ P\{X_n=0\}=1-p$.

ע"פ חוק המספרים הגדולים (LLN), עבור m גדול מספיק יהיו ברצף שפלט המקור בקירוב $p\cdot m$ '1'-ים ו- $(1-p)\cdot m$ '0'-ים.

נשאלת השאלה כמה קומבינציות באורך m קיימות אשר מקיימות את ה-LLN במדויק?

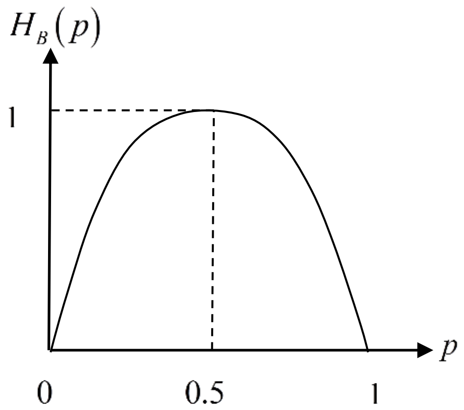
התשובה היא (מטעמי קומבינטוריקה): $\binom{m}{p\cdot m}$.

$$\binom{m}{p\cdot m}=\frac{m!}{(p\cdot m)!\cdot((1-p)\cdot m)!}\underset{\substack{Striling\\Approximation}}{\approx}2^{m\cdot H_B(p)}$$

$$H_B(p)\triangleq p\cdot\log\left(\frac{1}{p}\right)+(1-p)\cdot\log\left(\frac{1}{1-p}\right)$$

המקודד ימפה את הוקטור $x_1,\dots,x_m$ (תחת ההנחה ש- LLN מתקיים) לאינדקס בקבוצה בגודל $2^{m\cdot H_B(p)}$. לכן הקוד יהיה באורך: $\log_2\left(2^{m\cdot H_B(p)}\right)=m\cdot H_B\left(p\right)$. ולכן קצב הקוד הינו:

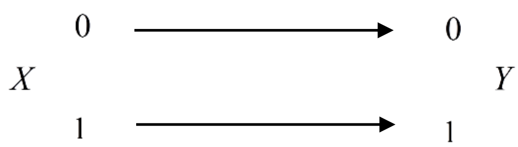
$$R=\frac{\text{code length}}{\text{amount of source bits}}=\frac{m\cdot H_B\left(p\right)}{m}=H_B\left(p\right)$$



כאשר $p=1/2$ קיימת ההפתעה הרבה ביותר בתוצאה. ולכן נראה כי שם קיימת הכי הרבה אינפורמציה, ולא ניתן לבצע דחיסה של המקור.

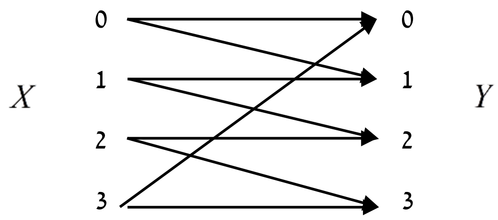
1.8. דוגמאות לקידוד ערוץ

א. <u>ערוץ בינארי "שקט"</u>:
זהו בעצם ערוץ דטרמיניסטי:



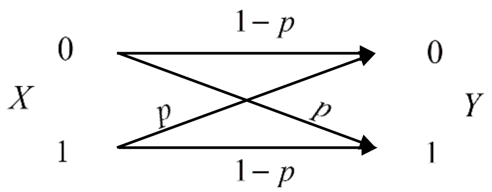
הערוץ יכול להעביר רק ביט אחד בזמן נתון ולכן קיבול הערוץ הוא $C=1{}^{\text{bit}}\!/_{\text{channel use}}$.

ב. הערוץ מ- X ל- Y הוא מהצורה הבאה:



ההסתברות של כל חץ היא 0.5. קיבול הערוץ הוא בעצם כמות המידע שאפשר להעביר בשימוש ערוץ יחיד ללא שגיאת פענוח. בסכימה הנ"ל קיבול הערוץ הוא גם כן 1 ביט לשימוש ערוץ. המיפוי מתבצע בצורה הבאה: ביט '0' ממופה לכניסה "0" וביט '1' ממופה לכניסה "2". בדרך זו לא תיפול שום שגיאה בפענוח כי אם מתקבל "0" או "1" אז ברור כי שודר '0' ואם מתקבל "2" או "3" אז שודר '1'.

ג. <u>ערוץ בינארי סימטרי (BSC)</u>:



ניתן לייצג את הערוץ בצורה הבאה: $Y = X \overset{\text{xor}}{\oplus} Z$, כאשר Z הינו רעש ברנולי עם הסתברות p. אם מכניסים את הבלוק $x_1,\ldots,x_n$, ע"פ חוק המספרים הגדולים יהיו $n\cdot p$ חילופי סימן. ולכן, עבור וקטור $\underline{Y}$ שנקלט תהיה הסתברות גבוהה יותר לוקטורי $\underline{X}$ מסוימים מאשר לאחרים. נסתכל על בעיית הקידוד בצורה הבאה: לכל כניסה $\underline{X}$ קיים "כדור" של ערכי $\underline{Y}$ מסוימים אשר עשויים להתקבל. רדיוס הכדור הוא $n\cdot p$ ולכן גודל הכדור הוא: $\binom{n}{n\cdot p} \approx 2^{n\cdot H_B(p)}$. נרצה לבצע קידוד כך שהכדורים שנוצרים לא "ייכנסו" אחד בתוך השני וכך לא יהיו טעויות בפענות. עבור וקטור בגודל n, גודל מרחב המוצא הוא 2^n וגודל של כל "כדור אי-ודאות" הוא $2^{n\cdot H_B(p)}$. נסמן ב- M את כמות וקטורי ה- $\underline{X}$ שלאחר מעבר בערוץ הכדורים שלהם לא "יחפפו" והם יהיו ברי-הבחנה. מתקיים:

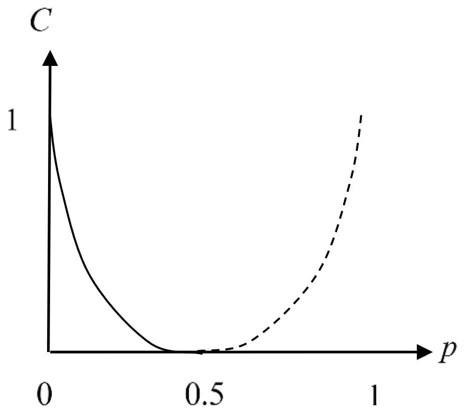
$$M \le \frac{2^n}{2^{n\cdot H_B(p)}} = 2^{n\cdot(1-H_B(p))}$$

מכאן נובע חסם על קיבול הערוץ (שנמדד בביט לשימוש ערוץ):

$$C \le \frac{\log_2(\text{amount of possible input vectors})}{n} = \frac{\log_2\left(2^{n\cdot(1-H_B(p))}\right)}{n} =$$

$$= \frac{n\cdot(1-H_B(p))}{n} = 1-H_B(p)$$

בהמשך הקורס נראה כי קצב זה הוא בר-השגה, כלומר: $\underline{C = 1-H_B(p)}$.



1.9. מאפייני הגישה של תורת האינפורמציה

א. גישה הסתברותית

ב. מדברים על אורך בלוק קידוד שהולך ל- ∞ (שימוש בחוק המספרים הגדולים).

ג. מתעלמים מסיבוכיות מימוש.

2. מדדי אינפורמציה

2.1 אנטרופיה

נתון מ"א $X \sim P(x)$ עם א"ב בדיד $X=\{1,2,...,M\}$ עם הסתברויות $\underline{p}=\{p_1,p_2,...,p_M\}$ בהתאמה. <u>האנטרופיה</u> מוגדרת בצורה הבאה:

$$Entropy: H(X)=H(p_1,...,p_m) \triangleq \sum_{i=1}^{m} p_i \cdot \log\left(\frac{1}{p_i}\right) = -\sum_{i=1}^{m} p_i \cdot \log(p_i)$$

<u>הערות</u>:

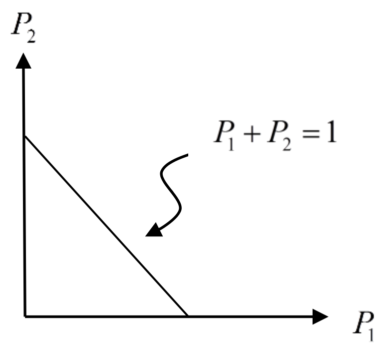
- ✓ עבור $p_i=0$ מתקבל הביטוי $0\cdot\infty$. ביטוי הנותן תוצאה 0 כיוון שהאפס הוא מוחלט והאינסוף הוא בשאיפה.
- ✓ אם נציב $p_i=0$ בהגדרה השניה של האנטרופיה, יתקבל הערך $0\cdot\log 0$. ערך זה אינו מוגדר. כיוון שמתקיים: $\lim_{x\to 0} x\cdot\log(x)=0$, מגדירים: $0\cdot\log 0 \triangleq 0$.
- ✓ אם Log הוא בבסיס 2 אז האנטרופיה נמדדת בביטים. אם עובדים בבסיס הטבעי (ln) אז מודדים ב-[nat].
- ✓ ניתן לבחון את האנטרופיה בצורה הבאה:

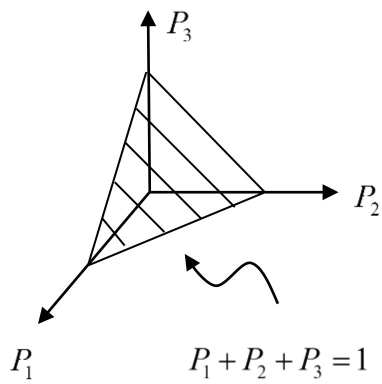
$$H(X)=\sum_{x\in X} p(x)\cdot\log\left(\frac{1}{p(x)}\right)=E\left(\log\frac{1}{p(X)}\right)$$

בעצם זוהי תוחלת על מידת ההפתעה או לחלופין תוחלת על האינפורמציה העצמית.

2.2 תכונות יסודיות של האנטרופיה

א. פונקציה רציפה וגזירה של הוקטור $\underline{p}$ בתוך הסימפלקס (תחום ערכים שמקיימים ערכי הסתברות). בנוסף, מתקיים $H=0$ על שפת הסימפלקס.
<u>סימפלקסים בדו- מימד ובתלת-מימד</u>:





ב. תמיד מתקיים: $H\ge 0$, ושוויון מתקיים אמ"מ אחד ה- p_i הוא 1 וכל השאר 0 – כלומר, במקרה של משתנה דטרמיניסטי.

ג. תמיד מתקיים: $\log(m) \underset{p_i=\frac{1}{m}}{=} H(uniform) \ge H$.

<u>הוכחה</u>:

יש למצוא את הוקטור $\{p_1,...,p_m\}$ שהצבתו בביטוי האנטרופיה ייתן ערך מקסימלי. על וקטור זה להיות וקטור הסתברות. זוהי בעצם בעייה של מציאת נקודת קיצון תחת אילוץ:

$$\max_{\substack{\{p_1,\dots,p_m\} \\ p_i \ge 0 \\ \sum_{i=1}^{m} p_i = 1}} \sum_{i=1}^{m} p_i \cdot \log\left(\frac{1}{p_i}\right)$$

נגדיר לגרנז'יאן:

$$L(p_1,\dots,p_m) = \sum_{i=1}^{m} p_i \cdot \log\left(\frac{1}{p_i}\right) + \lambda\left(\sum_{i=1}^{m} p_i\right)$$

$$\frac{\partial L}{\partial p_i} = 0; i = 1,\dots,m$$

$$\Rightarrow p_i = \exp\{-1-\lambda\} = const$$

כלומר, בפילוג אחיד מתקבלת נקודת מקסימום. ניתן לוודא זאת ע"י ביצוע נגזרת שניה.

ד. תכונת העידון/קיבוץ (השווה עם תכונה ד' בסעיף 1.4 לעיל):

$$H(p_1,\dots,p_m) = H(p_1+p_2, p_3,\dots,p_m) + (p_1+p_2)\cdot H_B\left(\frac{p_1}{p_1+p_2}\right)$$

ה. תכונת הקמירות של האנטרופיה:

אם $\underline{p}$ ו- $\underline{q}$ וקטורי הסתברות, אזי:

$$H\left(\lambda\cdot\underline{p} + (1-\lambda)\cdot\underline{q}\right) \ge \lambda\cdot H(\underline{p}) + (1-\lambda)\cdot H(\underline{q})$$

2.3. הגדרות נוספות הקשורות לאנטרופיה

יהיו X,Y מ"א עם פונקציית צפיפות הסתברות משותפת $P(X,Y)$.

א. אנטרופיה משותפת:

$$H(X,Y) = H(P(X,Y)) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\cdot\log\left(\frac{1}{p(x,y)}\right)$$

ב. אנטרופיה המותנית בערך מסוים:

$$H(X|Y=y) = \sum_{x\in X} p(x|y)\cdot\log\left(\frac{1}{p(x|y)}\right)$$

ג. אנטרופיה מותנית (ממוצעת):

$$H(X|Y) = \sum_{x\in X}\sum_{y\in Y} p(y)\cdot p(x|y)\log\left(\frac{1}{p(x|y)}\right)$$

ד. אינפורמציה הדדית:

$$I(X;Y) = \text{"}I(P(X),P(Y|X))\text{"} = H(X) - H(X|Y)$$

$$\underset{\text{easy to prove}}{=} H(Y) - H(Y|X) \underset{\text{easy to prove}}{=} H(X) + H(Y) - H(X,Y)$$

תכונות:

✓ $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$. לכלל זה קוראים גם "כלל השרשרת לאנטרופיה".

✓ $H(X|Y) \le H(X)$ ושוויון מתקיים אמ"מ $X \perp Y$ (X ו- Y הם בת"ס).

✓ מתקיים $I(X;Y) \ge 0$, ושוויון אמ"מ $X \perp Y$.

הגדרה:

האינפורמציה העצמית של הערך x היא:

$$I_x = \log\left(\frac{1}{p(x)}\right)$$

וכפי שראינו בתחילת הפרק:

$$H(X) = EI_x = E\left[\log\frac{1}{p(X)}\right] = E\left[-\log\left(p(X)\right)\right]$$

ובאופן דומה ניתן לראות כי:

$$H(X|Y=y) = E\left[-\log\left(p(X|y)\right)\right]$$
$$H(X,Y) = E\left[-\log\left(p(X,Y)\right)\right]$$

דוגמא:

במקרה של מ"א בינארי, מתקיים כזכור:

$$H(X) = H_B(p) = -p\cdot\log p - (1-p)\cdot\log(1-p)$$

במקרה האקראי ביותר $(p=0.5)$, האינפורמציה העצמית:

$$I_x = \log\frac{1}{0.5} = 1bit$$
$$\Rightarrow H(X) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 1 = 1bit$$

ובמקרה הדטרמיניסטי שבו מתקבל הערך 0 בצורה דטרמיניסטית:

$$I_0 = \log\frac{1}{1} = 0 \rightarrow \text{no surprise!}$$
$$\mathrm{I}_1 = \log\frac{1}{0} = \infty \rightarrow \text{very surprising!}$$
$$H(X) = 1\cdot 0 + 0\cdot\infty = 0$$

2.4. יחסי התניה ואנטרופיה

2.4.1. משפט ההתניה והאנטרופיה

משפט: התניה מורידה (לא מעלה) את האנטרופיה. כלומר, $H(X|Y) \le H(X)$ ושוויון מתקיים אמ"מ $X \perp\!\!\!\perp Y$.

דוגמאות:

א. X,Y מ"א בינאריים. Y מתפלג בינארית עם פרמטר 0.5. הקשר בין X ל- Y מתואר בטבלה הבאה:

Y\X	0	1
0	1	0
1	0	1

נשאלת השאלה מהם $H(X|Y)$ ו- $H(X)$. קל להבחין כי כאשר Y ידוע אז X הוא דטרמיניסטי ולכן - $H(X|Y)=0$.

$$H(X) = \frac{1}{2}\cdot\log(2) + \frac{1}{2}\cdot\log(2) = 1bit$$

ב. <u>ערוץ BSC</u>:

נסמן את הסתברות החילוף בערוץ ב- q (Cross-Over Probability). כניסת הערוץ היא X ומוצא הערוץ הוא Y. $X \sim unif\{0,1\}$. ניתן לראות מטעמי סימטריה או הוכחה קלה כי $p(Y=0)=p(Y=1)=0.5$. כזכור, האנטרופיה של משתנה המפולג באחידות היא $\log M$ ולכן:

$$H(Y)=\log 2=1$$

$$H(Y|X)=\sum_{x=0}^{1}p(x)\cdot H(Y|X=x)=\sum_{x=0}^{1}p(x)\cdot\sum_{y=0}^{1}p(y|x)\cdot\log\left(\frac{1}{p(y|x)}\right).$$

$$=\sum_{x=0}^{1}p(x)\cdot H_B(q)=H_B(q)\cdot\sum_{x=0}^{1}p(x)=H_B(q)\le 1$$

וגם בדוגמא זו רואים כי התנייה לא מעלה את האנטרופיה.

על-מנת להוכיח את המשפט הנ"ל יש צורך בכמה כלים. נגדיר אותם כעת:

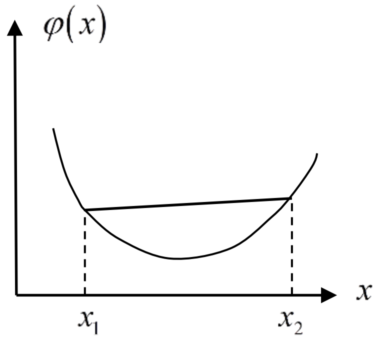
2.4.2. קמירות של פונקציות

<u>הגדרה</u>:

פונקציה $\varphi(x):\mathbb{R}^1\to\mathbb{R}^1$ היא <u>*קמורה*</u> (Convex) בתחום D אם לכל $x_1,x_2\in D$ הקומבינציה הקונבקסית של הפונקציה בנקודות x_1,x_2 היא גדולה או שווה מהפונקציה של הקומבינציה הקונבקסית של x_1,x_2. כלומר, לכל $0\le\lambda\le 1$, מתקיים:

$$\lambda\cdot\varphi(x_1)+(1-\lambda)\cdot\varphi(x_2)\ge\varphi\left(\lambda\cdot x_1+(1-\lambda)\cdot x_2\right).$$

במילים אחרות, לכל $x_1,x_2\in D$, אם נמתח מיתר מ- x_1 ל- x_2, המיתר יהיה גבוה יותר מהפונקציה.



<u>הערות</u>:

- ✓ בניגוד לפונקציה הנ"ל שהיא "שמחה", הפונקציה "העצובה" נקראת פונקציה קעורה (Concave). אולם, מעתה נקרא לכל הפונקציות (גם לשמחה וגם לעצובה) קמורות ונציין במפורש האם הכוונה היא לשמחה או לעצובה.
- ✓ ההגדרה המתטית לפונקציה קמורה $\cap$ היא כנ"ל רק שהמיתר עובר מתחת לפונקציה והכיוון של הסימן של האי-שוויון מתחלף.
- ✓ ניתן להראות את שקילות ההגדרה הנ"ל ל- $\frac{\partial^2\varphi}{\partial x^2}\ge 0$ בתחום (בתנאי שהפונקציה גזירה בתחום).
- ✓ ניתן להרחיב את ההגדרה הנ"ל גם למקרה של פונקצייה וקטורית: $\varphi:\mathbb{R}^n\to\mathbb{R}^1$.
- ✓ נאמר ש- $\varphi(x)$ היא <u>קמורה ממש</u> (במובן החזק) אם האי-שוויון הוא חזק.

2.4.3. אי-שוויון ינסן (Jansen)

אם פונקציה קמורה (בתחום מסוים) ו- X מ"א (או ו"א) בתחום הנ"ל. אזי:

$$E\left(\varphi(X)\right) \geq \varphi\left(E(X)\right)$$

<u>דוגמא</u>:

נניח כי X הינו מ"א בינארי המקבל את הערכים x_1, x_2 בהסתברויות $\lambda, (1-\lambda)$ בהתאמה.

$$\varphi\left(E(X)\right) = \varphi\left(\lambda \cdot x_1 + (1-\lambda) \cdot x_2\right)$$
$$E\left(\varphi(X)\right) = \lambda \cdot \varphi(x_1) + (1-\lambda) \cdot \varphi(x_2)$$

וכיוון ש- $\varphi(X)$ הינה קמורה אז המשפט מתקיים.

- ✓ ניתן להוכיח את משפט ינסן ל-א"ב כללי וסופי ע"י שימוש באינדוקציה.
- ✓ שוויון באי-שוויון ינסן עבור פונקציה קמורה ממש מתקבל אמ"מ X הוא דטרמיניסטי.
- ✓ בנוסף, יתקבל שוויון באי-שוויון עבור פונקציה שהיא בעצם קו ישר.

2.4.4. קמירות האנטרופיה בפילוג

<u>טענת עזר</u>: פונקציית האנטרופיה היא תמיד קמורה $\cap$ ממש ביחס לוקטור הפילוג.

<u>דוגמא</u>:

אם ניזכר ב- $H_B(p)$ שהיא פונקציית האנטרופיה במקרה הבינארי, מתקיים בה:

$$H_B\left(\lambda \cdot p_1 + (1-\lambda) \cdot p_2\right) \geq \lambda \cdot H_B(p_1) + (1-\lambda) \cdot H_B(p_2)$$

ושוויון יתקיים במקרה בו: $p_1 = p_2$.

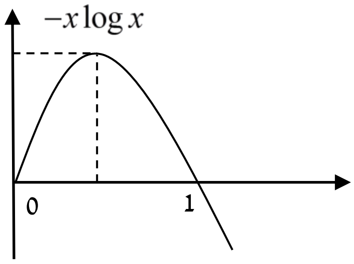
<u>הערה</u>:

$H_B(p)$ הינה Concave ולכן הסימן באי-השוויון התחלף ממה שהוגדר עבור פונקציה קמורה.

באופן כללי, האנטרופיה $H(\underline{p})$ קמורה ביחס לפילוג $\underline{p}$. כלומר, עבור k פילוגים $\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_k$ ו- $\lambda_1, \lambda_2, \ldots, \lambda_k$ כאשר $\lambda_i \geq 0$ ו- $\sum_{i=1}^{k} \lambda_i = 1$ מתקיים:

$$H\left(\sum_{i=1}^{k} \lambda_i \cdot \underline{p}_i\right) \geq \sum_{i=1}^{k} \lambda_i \cdot H\left(\underline{p}_i\right)$$

שוויון יתקבל כאשר כל ה- $\underline{p}_i$ שווים.



<u>הוכחת הטענה:</u>

הוכחת הטענה נובעת מקמירות ∩ של הפונקציה $-x\cdot\log(x)$ בתחום $[0,1]$.

2.4.5. הוכחת משפט ההתניה והאנטרופיה

ניזכר בהגדרה:

$$H(X|Y)\triangleq\sum_{y\in Y}p(Y=y)\cdot H(X|Y=y)=-\sum_{y\in Y}p(y)\cdot\sum_{x\in X}p(x|y)\cdot\log\left(p(x|y)\right)$$

נתייחס ל- $p(x|y)$ (כאשר y הוא קבוע) כאל וקטור הסתברויות $\underline{p}_y$ ולהסתברויות כאל λ_y. לכן:

$$H(X|Y)=\sum_{y\in Y}\lambda_y\cdot H\left(\underline{p}_y\right)\underset{\text{entropy is concave}}{\leq}H\left(\sum_{y\in Y}\lambda_y\cdot\underline{p}_y\right)=H\left(\sum_{y\in Y}p(y)\cdot p(x|y)\right)$$
$$=H\left(p(x)\right)=H(X)$$

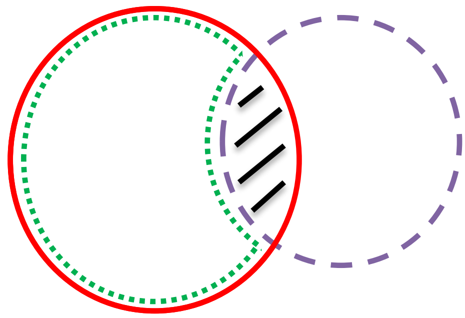
מש״ל.

2.4.6. השלכת משפט ההתניה והאנטרופיה לגבי האינפורמציה ההדדית

ניתן לשים לב כי כיוון שההתניה לא מעלה את האנטרופיה, מתקיים כי <u>האינפורמציה ההדדית היא גודל אי-שלילי</u>:

$$I(X;Y)\triangleq H(X)-H(X|Y)\triangleq H(Y)-H(Y|X)\geq 0$$

ניתן להסתכל על התובנה הנ״ל בעזרת דיאגרמת Venn:



<u>אדום (ללא קווקו)</u> - $H(X)$, <u>סגול</u> (קווקו קווים) - $H(Y)$, <u>ירוק (קווקו נקודות)</u> - $H(X|Y)$, <u>השטח המקווקו בין העיגולים</u> - $I(X;Y)$

כיוון ששטח הוא גודל אי-שלילי, כך גם $I(X;Y)$ וכל גדלי האנטרופיה.

<u>הגדרה</u>:

<u>*דיברגנס*</u> זהו מדד מרחק בין פילוגי הסתברות. עבור פילוגים $p(X)$ ו- $q(X)$, הדיברגנס מוגדר כך:

$$D(p\|q) \triangleq \sum_{x\in X} p(x)\cdot\log\frac{p(x)}{q(x)}$$

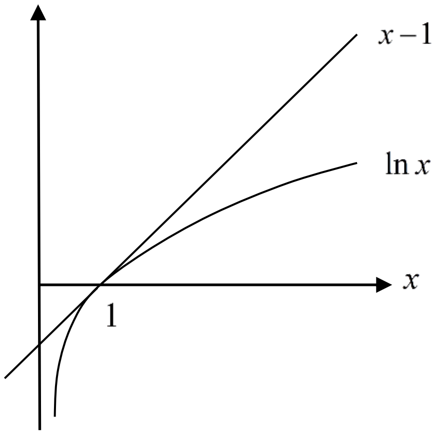
<u>הערה</u>: יש לשים לב כי הנוסחא לא סימטרית ביחס ל- p ו- q.

<u>טענה</u>:

$D(p\|q)\geq 0$ ושוויון מתקיים אמ"מ $p(X)=q(X)$.

<u>הוכחה</u>:

נשתמש בעובדה כי $\ln x \leq x-1$ ושוויון מתרחש כאשר $x=1$:



נבצע את הפיתוח הבא:

$$-\ln x \geq 1-x$$
$$\ln\frac{1}{x}\geq 1-x$$
$$\ln t \geq 1-\frac{1}{t}$$

נשתמש בתכונה זו עבור הדיברגנס:

$$\begin{aligned} D(p\|q) &= \sum_{x\in X} p(x)\cdot\log\frac{p(x)}{q(x)} \geq \sum_{x\in X} p(x)\cdot\left(1-\frac{q(x)}{p(x)}\right) \\ &= \sum_{x\in X} p(x) - \sum_{x\in X} q(x) = 1-1=0 \end{aligned}$$

אם נראה כי האינפורמציה ההדדית היא מקרה פרטי של הדיברגנס נוכל להוכיח את הטענה כי האינפורמציה ההדדית היא אי- שלילית.

<u>טענה</u>:

האינפורמציה ההדדית היא מקרה פרטי של דיברגנס.

<u>הוכחה</u>:

$$\begin{aligned} &I(X;Y)=H(X)-H(X|Y) \\ &H(X,Y)=H(X|Y)+H(Y) \\ &\Rightarrow I(X;Y)=H(X)+H(Y)-H(X,Y) \\ &\qquad = -\sum_{x\in X} \underbrace{p(x)}_{=\sum_{y\in Y} p(x,y)} \cdot\log\left(p(x)\right) - \sum_{y\in Y} \underbrace{p(y)}_{=\sum_{x\in X} p(x,y)} \cdot\log\left(p(y)\right) + \sum_{x\in X}\sum_{y\in Y} p(x,y)\cdot\log\left(p(x,y)\right) \end{aligned}$$

$$=\sum_{x,y} p(x,y)\cdot\left(-\log(p(x))-\log(p(y))+\log(p(x,y))\right)$$

$$=\sum_{x,y} p(x,y)\cdot\log\left(\frac{p(x,y)}{p(x)\cdot p(y)}\right)=D\left(p(x,y)\,\|\,p(x)\cdot p(y)\right)$$

כלומר, האינפורמציה ההדדית היא הדיברגנס בין הפילוג המשותף $p(x,y)$ לבין הפילוג הבת״ס $p(x)\cdot p(y)$ עם אותם הפילוגים השוליים.

ניתן גם להראות כי:

$$I(X;Y)=\sum_{x\in X} p(x)\sum_{y\in Y} p(y\,|\,x)\log\frac{p(y\,|\,x)}{\sum_{x'\in X} p(y\,|\,x')\cdot p(x')}=\sum_{x\in X} p(x)\sum_{y\in Y} p(y\,|\,x)\log\frac{p(y\,|\,x)}{p(y)}$$

$$=\sum_{x\in X} p(x)\cdot D\left(p(y\,|\,x)\,\|\,p(y)\right)\triangleq\sum_{x\in X} p(x)\cdot I(X=x;Y)$$

באופן זה האינפורמציה ההדדית מבוטאת כפונקציה של פילוג הכניסה ושל פילוג המעבר בערוץ.

2.5. סיכום ביניים ומשמעויות נוספות

נזכיר את המושגים שנלמדו בפרק זה:

- ✓ $H(X)$ - אי-הוודאות הממוצעת לגבי המ״א X.
- ✓ $H(Y\,|\,X)$ - אי-הוודאות הממוצעת הנותרת ב- Y לאחר ש- X ידוע.
- ✓ $H(X,Y)$ - אי-הוודאות המשותפת הממוצעת לגבי הזוג X,Y. או לחלופין, אי-הודאות ב- X ועוד אי-הודאות הנותרת ב- Y לאחר ש- X ידוע. מתקיים:

$$H(X,Y)=H(X)+H(Y\,|\,X)=H(Y)+H(X\,|\,Y)$$

- ✓ $I(X;Y)$ - מידת הירידה באי-ודאות לגבי X לאחר קבלת Y. מתקיים:

$$I(X;Y)=H(X)-H(X\,|\,Y)=H(Y)-H(Y\,|\,X)=H(X)+H(Y)-H(X,Y)$$

- ✓ $D(p\,\|\,q)$ - לערך זה קוראים דיברגנס או מרחק אינפורמציוני או מרחק קולבאק-לייבלר. ערך זה הוא בעצם ״אנטרופיה יחסית״ – המרחק האינפורמציוני בין הפילוג $\underline{p}$ לפילוג $\underline{q}$. כזכור, ערך זה אינו סימטרי ביחס ל- $\underline{p}$ ו- $\underline{q}$.

  משמעויות:

  א. אקספוננט הסיכוי שמקור עם פילוג $\underline{q}$ ייצור סדרה עם פילוג אמפירי $\underline{p}$. כלומר - $2^{-nD(p\|q)}$.

  ב. המחיר האינפורמציוני על שימוש בפילוג ״לא מתואם״ (המקור מתנהג לפי $\underline{p}$ כאשר הקוד מתואם ל- $\underline{q}$).

  ניתן להראות כי מתקיים:

$$H(X)=\log|X|-D\left(p(X)\,\|\,unif\,(|X|)\right)$$

- ✓ $0\le H(X)\le\log|X|$ - אי-ודאות מינימלית (אפס) אם X דטרמיניסטי ומקסימלית אם הוא משתנה אחיד.
- ✓ מתקיימים אי-השוויונים הבאים:

$$0\le H(X\,|\,Y)\le H(X)$$

$$0\le I(X;Y)\le H(X),\;\; 0\le I(X;Y)\le H(Y)$$

$$0\le D(p\,\|\,q)\le\infty$$

2.6. אינפורמציה מותנית

✓ $I(X;Y|Z)$ - מידת הירידה באי-ודאות של X לאחר קבלת Y, כאשר Z ידוע מראש.

מתקיים:

$$I(X;Y|Z) \triangleq \sum_{z \in Z} p(z) \cdot I(X;Y|Z=z) = E_Z\left[I(X;Y|Z=z)\right]$$

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

מתקיים כי $I(X;Y|Z)$ לא בהכרח גדול מ- $I(X;Y)$ ויכול להיות אף קטן יותר.

אינפורמציית הצד יכולה:

א. להגדיל את האינפורמציה במקרה בו למשל Z נותן מידע לגבי הערוץ מ- X ל- Y או ההיפך.

ב. להקטין את האינפורמציה ההדדית במקרה בו למשל קיימת שרשרת מרקוב $Y \leftrightarrow X \leftrightarrow Z$. כלומר, Z נותן מידע על X ולכן Y נותן פחות אינפורמציה על X במקרה זה.

2.7. אי-שוויון עיבוד הנתונים

הגדרה:

שלישייה מרקובית מסומנת כך: $Z \leftrightarrow Y \leftrightarrow X$ תמיד מתקיים:

$$p(x,y,z) = p(x) \cdot p(y|x) \cdot p(z|x,y)$$

ובשלישייה מרקובית מתקיים:

$$p(x,y,z) = p(x) \cdot p(y|x) \cdot p(z|y)$$
$$\Rightarrow p(z|x,y) \equiv p(z|y)$$

או לחלופין, X ו- Z הם בת"ס בהנתן Y.

אי-שוויון עיבוד הנתונים (Data Processing Inequality):

אם X,Y,Z הם שלישייה מרקובית $X \leftrightarrow Y \leftrightarrow Z$, אזי:

$$I(X;Z) \le I(X;Y) \text{ and } I(Y;Z)$$

ושוויון מתקיים אמ"מ $I(X;Y|Z)=0$ (או מטעמי סימטריה: $I(Z;Y|X)=0$).

הוכחה:

$$I(X;Y,Z) = I(X;Y) + \underbrace{I(X;Z|Y)}_{\substack{=0 \\ \text{derived from the definition} \\ \text{of Markov chain}}}$$

$$I(X;Y,Z) = I(X;Z) + \underbrace{I(X;Y|Z)}_{\ge 0}$$

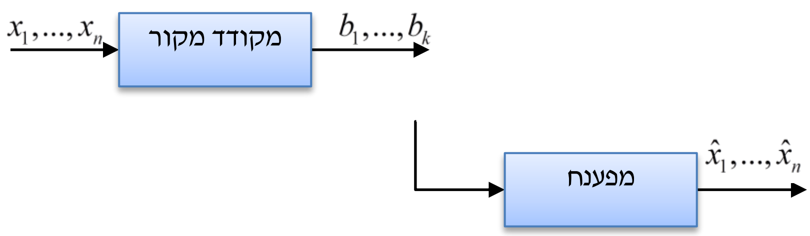
$$\Rightarrow I(X;Z) = I(X;Y) - I(X;Y|Z)$$

3. תכונות של סדרות מקור ארוכות

3.1. AEP – Asymptotic Equi-Partition Property

נושא זה מוגדר ע"י רמי זמיר בתור: "שוויון באחרית הימים"

<u>בעיית קידוד המקור</u>:



<u>דוגמאות</u>:

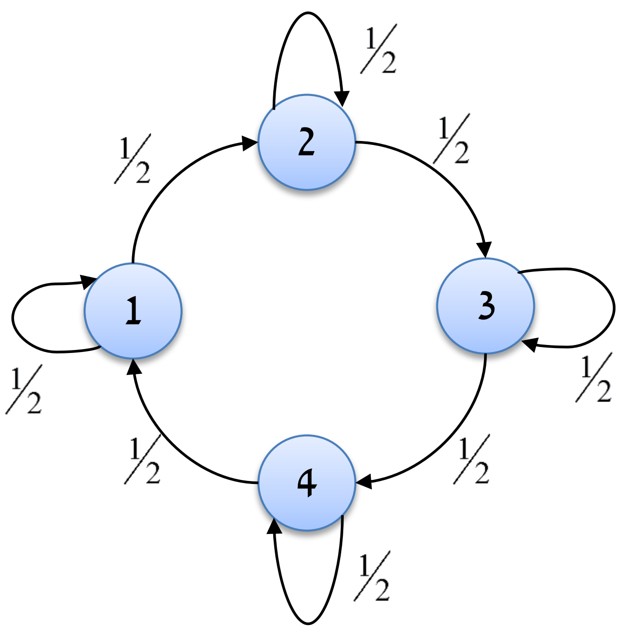
א. $bernoully(1/2)$. זהו מקור לא דחיס. זאת משום שקיימות 2^n אפשרויות להודעת כניסה וכולן שוות הסתברות.

ב. מקור קוואנטרי $(1,2,3,4)$, iid, אשר מקיים:

$$p(x=1)=p(x=3)=0.5$$
$$p(x=2)=p(x=4)=0$$

בדוגמא זו קיימות 2^n אפשרויות שוות הסתברות ולכן לקידוד זה דרושים n ביטים. אנו יכולים לחסוך בקידוד n ביטים!

ג. <u>מקור קוואנטרי מרקובי</u>:



$$X_1 \sim unif\,(1,2,3,4)$$
$$p(X_{n+1}\,|\,X_n)=0.5$$

נספור כמה סדרות קיימות בכל אורך:

- $n=1$: 4 סדרות
- $n=2$: 2x4 סדרות
- $n=3$: 2x2x4 סדרות

.
.
.

✓ $2^{n-1}\times 4$ סדרות

לכן, כדי לקודד n דגימות צריך: $\log(2^{n-1}\times 4)=n+1\approx n$ דגימות וזה שווה בערך ל- 1bit לדגימה. גם פה ניכר החיסכון אם יימצא קידוד מתאים.

משפט:

$x_1, x_2, \ldots$ הוא תהליך iid (בתורת האינפורמציה מקור iid שקול לערוץ חסר זיכרון) שמתפלג בפילוג $p(X)$. נסמן ב- $p(X_1,\ldots,X_n)$ את הפילוג המשותף של הוקטור $X_1,\ldots,X_n$. אזי מתקיים:

$$-\frac{1}{n}\log P(X_1,\ldots,X_n)\xrightarrow[n\to\infty]{} H(X) \text{ in probability}$$

תזכורת:

התכנסות בהסתברות:

$$p\left(\left|-\frac{1}{n}\log\left(p(X_1,\ldots,X_n)\right)-H(X)\right|<\varepsilon\right)\xrightarrow[\substack{n\to\infty \\ \forall\varepsilon>0}]{}1$$

נבחן מספר מקרים:

- ✓ עבור $n=1$, הביטוי הינו $\log\frac{1}{p(X_1)}$ והאנטרופיה שלו היא:

$$E\left[\log\frac{1}{P(X_1)}\right]=H(X)$$

- ✓ עבור $n=2$, הביטוי הינו $\frac{1}{2}\log\left(\frac{1}{p(X_1,X_2)}\right)$ והאנטרופיה שלו היא:

$$E\left[\frac{1}{2}\log\left(\frac{1}{p(X_1,X_2)}\right)\right]\underset{iid}{=}E\left[\frac{1}{2}\log\left(\frac{1}{p(X_1)\cdot p(X_2)}\right)\right]$$
$$=E\left[\frac{1}{2}\log\frac{1}{p(X_1)}+\frac{1}{2}\log\frac{1}{p(X_2)}\right]=E\left[\log\frac{1}{p(X_1)}\right]=H(X)$$

הוכחה:

עבור מקור חסר זיכרון:

$$p(X_1,\ldots,X_n)=\prod_{i=1}^{n}p(X_i)$$
$$\Rightarrow -\frac{1}{n}\log\left(p(X_1,\ldots,X_n)\right)=-\frac{1}{n}\sum_{i=1}^{n}\log\left(p(X_i)\right)$$

אם $X_1, X_2, \ldots$ הם iid אז לכל פונקציה $g(x)$, אז $Y_i=g(X_i)$ הם גם iid.

לפי חוק המספרים הגדולים, כש- $n\to\infty$, הממוצע מתכנס לתוחלת:

$$-\frac{1}{n}\sum_{i=1}^{n}\log p(X_i)\xrightarrow[n\to\infty]{}-E\left[\log p(X_i)\right]\underset{\text{in probability}}{=}H(X)$$

__הערה__: המשפט ניתן להרחבה למקורות ארגודיים כלליים – כפי שנראה בהמשך.

3.2. הקבוצה האופיינית

<u>הגדרה</u>:

<u>הקבוצה האופיינית</u> (Typical Set) במובן החלש:

$$A_\varepsilon^{(n)} \triangleq \left\{ \underline{X} \in X^n : \underbrace{2^{-n[H+\varepsilon]}}_{p_{\min}} \le p\left(\underline{X}\right) \le \underbrace{2^{-n[H-\varepsilon]}}_{p_{\max}} \right\}$$

<u>תכונות הקבוצה האופיינית</u>:

א. לכל $\underline{X} \in A_\varepsilon^{(n)}$ מתקיים: $p\left(\underline{X}\right) \le 2^{-n[H\pm\varepsilon]}$. כלומר: $H-\varepsilon \le -\frac{1}{n}\log p\left(\underline{X}\right) \le H+\varepsilon$.

ב. עבור וקטור אקראי $\underline{X} \sim p\left(\underline{X}\right)$ ועבור $\delta > 0$ ו- n מספיק גדול מתקיים: $p\left(\underline{X} \in A_\varepsilon^{(n)}\right) \ge 1-\delta$. תכונה זו נובעת ממשפט ה-AEP וניתן לנסח אותה כך:

$$p\left\{\left|-\frac{1}{n}\log p\left(\underline{X}\right) - H\left(X\right)\right| < \varepsilon\right\} \underset{\text{for n that's large enough}}{\ge} 1-\delta$$

ג. לכל n מתקיים:

$$\left|A_\varepsilon^{(n)}\right| \le 2^{n\cdot[H+\varepsilon]} = \frac{1}{p_{\min}}$$

מסי האיברים בקבוצה הוא לכל היותר 1 מחולק ב- $p_{\min}$.

ד. אם n מספיק גדול, אזי:

$$\left|A_\varepsilon^{(n)}\right| \ge (1-\delta)\cdot 2^{n\cdot[H-\varepsilon]} = \frac{1-\delta}{p_{\max}}$$

סעיף זה נובע מסעיף ב'.

<u>הגדרה</u>:

<u>סדרה מטיפוס q</u> הינה סדרה בינארית המכילה q 1-ים ו- $(1-q)$ 0-ים.

<u>הגדרה</u>:

נגדיר *<u>שוויון במובן האקספוננציאלי</u>*. נאמר שסדרה $a_1, a_2, \ldots$ עולה (יורדת) עם אקספוננט E

ומסמנים: $a_n \doteq 2^{n\cdot E}$ אם מתקיים: $\frac{1}{n}\log a_n \xrightarrow[n\to\infty]{} E$

<u>הערה</u>:

שוויון אקספוננציאלי מתעלם מגדלים תת-אקספוננציאליים. לדוגמא:

$$2^{10n} + 2^{5n} \doteq 2^{10n},\ n^{10}\cdot 2^{nE} \doteq 2^{nE},\ 2^{nE}\cdot 100 \doteq 2^{nE},\ 2^{-10n} + 2^{-5n} = 2^{-5n}$$

<u>דוגמא</u>:

נבחן סדרה שמפולגת ברנולי עם פרמטר p. ההסתברות שהמקור יפלוט סדרה מסוימת מטיפוס q היא:

$$p\left(X_1,\ldots,X_n\right) = p^{q\cdot n}\cdot(1-p)^{(1-q)\cdot n} = 2^{n\left[q\log p + (1-q)\log(1-p)\right]} \underset{q=p}{=} 2^{-nH}$$

למשל, עבור $p=1/10$:

$$p(101)=p(011)=p(110)=\left(\frac{1}{10}\right)^2\cdot\left(\frac{9}{10}\right)^1$$

מסי הסדרות מטיפוס q :

$$\binom{n}{q\cdot n}\doteq 2^{n\cdot H_B(q)}$$

מכאן נובע שההסתברות לקבל סדרה מטיפוס q ממקור שמתפלג לפי p היא :

$$2^{n\left[q\log p+(1-q)\log(1-p)\right]}\cdot\binom{n}{q\cdot n}\doteq 2^{n\left[q\log p+(1-q)\log(1-p)\right]}\cdot 2^{nH_B(q)}=2^{-n\left[q\log\frac{q}{p}+(1-q)\log\frac{1-q}{1-p}\right]}=2^{-n\cdot D(q\|p)}$$

<u>AEP עבור מקור ברנולי עם פרמטר P</u>:

$$A_\varepsilon^{(n)}=\left\{\underline{X}:2^{-n\left[H_B(p)+\varepsilon\right]}\le p\left(\underline{X}\right)\le 2^{-n\left[H_B(p)-\varepsilon\right]}\right\}$$

בעצם, זוהי קבוצת כל הסדרות שמסי האחדים בהן הוא $n\cdot\left(p\pm\varepsilon\right)$, כלומר בקירוב $n\cdot p$ אחדים.

<u>דוגמא לאופייניות במובן החלש שאיננה במובן החזק</u> :

נניח מקור המפולג בצורה הבאה : $\left(A,B,C,D\right)\sim\left(\frac{1}{2},\frac{1}{4},\frac{1}{8},\frac{1}{8}\right)$.

הסדרה AAAABBCD - <u>אופיינית במובן החזק</u> כי כל סימבול מופיע בדיוק לפי ההסתברות.

הסדרה AAAABBCC - <u>אופיינית במובן בחלש</u> כי אמנם C,D הם בעלי אותה ההסתברות אך הסדרה לא מקיימת את הסתברות המקור באופן הדוק.

בקבוצה $A_\varepsilon^{(n)}$ מופיעות גם סדרות האופייניות במובן וחזק וגם במובן החלש כיוון שסדרות אלה מקיימות :

$$2^{-n[H+\varepsilon]}\le p\left(\underline{X}\right)\le 2^{-n[H-\varepsilon]}$$

3.3. משפט ה-AEP ההפוך

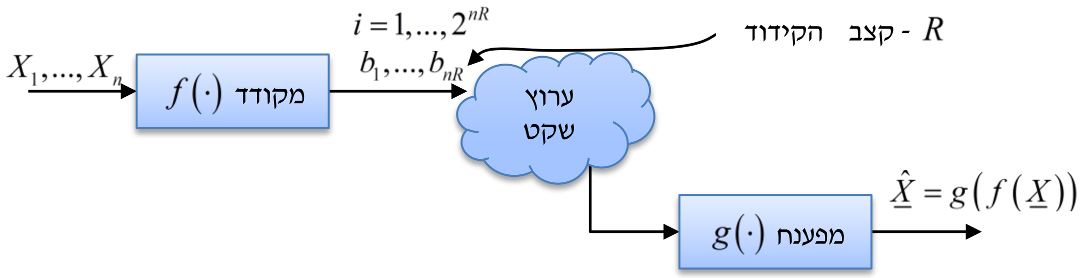
כשאומרים "הפוך" בתורת האינפורמציה הכוונה היא למה ש-"בלתי-אפשרי".

א. <u>הגדרה איכותית</u> : עבור n מספיק גדול, הגודל של קבוצת הסדרות עם הסתברות "משמעותית" הוא לפחות $\approx 2^{nH}$.

ב. <u>גרסא חלשה</u> : לכל $\varepsilon>0$, אם n מספיק גדול ואם $B^{(n)}$ היא קבוצת סדרות באורך n המקיימת $p\left(B^{(n)}\right)>1-\varepsilon$, אזי בהכרח : $\left|B^{(n)}\right|>2^{n[H-\varepsilon]}\cdot\left(1-2\varepsilon\right)$.

ג. <u>גרסא חזקה</u> : לכל $\varepsilon>0$ ומימד n מספיק גדול, אם $\left|B^{(n)}\right|<2^{n[H-\varepsilon]}$ אזי $p\left(B^{(n)}\right)<\varepsilon$.

4. קידוד מקור בקצב קבוע (קידוד בלוק)



<u>הסתברות השגיאה</u> מוגדרת כך: $p\left(\hat{\underline{X}}\neq\underline{X}\right)$

<u>הגדרה</u>: <u>קצב בר השגה בקידוד ללא עיוות (Lossless Compression)</u>. נאמר ש- R הוא קצב בר השגה אם לכל $\varepsilon>0$ קיימת מערכת קידוד (פונקציות f ו- g) במימד n (מספיק גדול) כך ש: $p\left(\hat{\underline{X}}\neq\underline{X}\right)<\varepsilon$.

<u>משפט הקידוד הראשון של שאנון</u>:

קצב בר ההשגה המינימלי האפשרי למקור ארגודי הוא קצב האנטרופיה $\bar{H}$ של המקור.

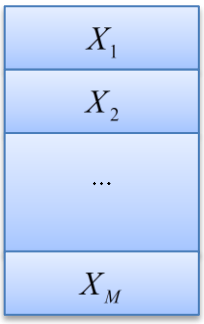
במקרה הפרטי של מקור חסר זיכרון:

$$\bar{H}=H=-\sum_i p_i\log p_i$$

<u>פרשנות</u>:

- ✓ <u>הצד הישר</u>: לכל $R>\bar{H}$ ניתן למצוא מערכת קידוד במימד מספיק גדול להשגת הסתברות שגיאה קטנה כרצונינו.
- ✓ <u>הצד ההפוך</u>: אם נתונה מערכת בקצב $R<\bar{H}$, אזי בהכרח הסתברות השגיאה לא "קטנה כרצונינו".

<u>הוכחת הצד הישר</u>:



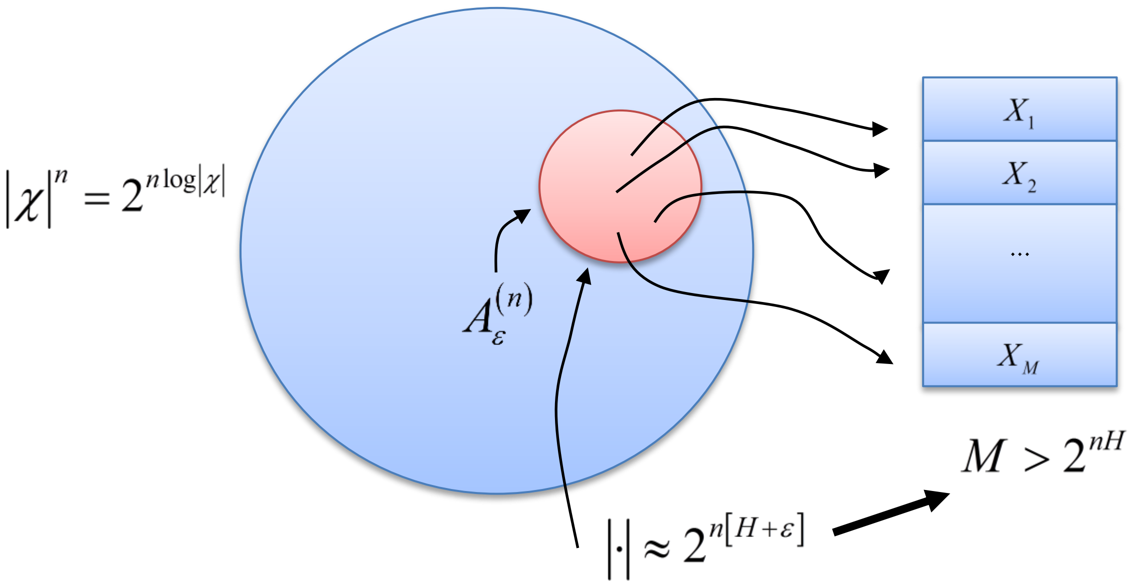
א. נבנה טבלה עם M סדרות.

ב. המקודד שולח את האינדקס i של המקום בטבלה של סדרת המקור שהתקבלה.

ג. אם המקור פלט סדרה $\underline{X}$ שלא נמצאת בטבלה, ישודר אינדקס כלשהו i (אקראי או קבוע – לא באמת משנה).

ד. המפענח מכיר את הטבלה וישחזר את $\hat{\underline{X}}$ לפי המקום ה- $\hat{i}$ בטבלה.

בנייה חכמה של הטבלה תאכלס אותה בסדרות מהקבוצה האופיינית $A_\varepsilon^{(n)}$ בתוספת (אם יש מקום) של סדרות לא אופייניות עם הסתברות משמעותית.



משפט ה- AEP מבטיח שעבור n מספיק גדול:

$$p\left(\underline{X}\in Table\right)>1-\varepsilon \Rightarrow p\left(error\right)<\varepsilon$$

<u>הערות</u>:

- ✓ הגודל של $A_\varepsilon^{(n)}$ קטן משמעותית מגודלה של χ^n.
- ✓ וקטור $\underline{X}\in A_\varepsilon^{(n)}$ יתקבל בהסתברות גבוהה.
- ✓ הסתברות האלמנטים ב- $A_\varepsilon^{(n)}$ אחידה בקירוב והיא שווה ל- $\doteq 2^{-n[H\pm\varepsilon]}$.

<u>הוכחת הצד ההפוך</u>:

נתונה מערכת בקצב הנמוך מהאנטרופיה $R<\bar{H}$, באורך בלוק m. נניח כי $\bar{H}-R=2\varepsilon$. נבחר n שהוא כפולה שלמה של m שעבורו: $p\left(x_1,\dots,x_n\in A_\varepsilon^{(n)}\right)>1-\varepsilon$ - דבר שהוא אפשרי ע״פ ה-AEP. נפעיל את מערכת הקידוד n/m פעמים כדי לקודד את הסדרה $x_1,\dots,x_n$.

$$p\left(\underbrace{x_1,\dots,x_n}_{\underline{X}}\in Table\right)=\underbrace{p\left(\underline{X}\in Table\cap A_\varepsilon^{(n)}\right)}_{(1)}+\underbrace{p\left(\underline{X}\in Table\cap \bar{A}_\varepsilon^{(n)}\right)}_{(2)}$$

$$(1)\le p_{\max}\cdot\left(\text{Table Size}\right)\le 2^{-n[H-\varepsilon]}\cdot 2^{nR}=2^{-n\underbrace{[H-R-\varepsilon]}_{\ge 0}}\xrightarrow[n\to\infty]{}0$$

$$(2)\le\varepsilon\leftarrow\text{Result of AEP}$$

הראינו כי ההסתברות של וקטור להיות בספר הקוד שואפת לאפס כאשר $R<\bar{H}$.

5. מקורות בעלי זיכרון

בניגוד למקורות שראינו עד עכשיו שהיו חסרי זיכרון, במקורות בעלי זיכרון מתקיים:

$$p\left(x_1,\ldots,x_n\right)\neq\prod_{i=1}^{n}p\left(x_i\right)$$

ולכן, ע"פ כלל השרשרת לאנטרופיה:

$$H\left(X_1,\ldots,X_n\right)=H\left(X_1\right)+H\left(X_2\,|\,X_1\right)+H\left(X_3\,|\,X_2,X_1\right)+\ldots+H\left(X_n\,|\,X_{n-1,}X_{n-2},\ldots,X_1\right)\neq\underbrace{n\cdot H\left(X_1\right)}_{\text{entropy of a memoryless source}}$$

ולמעשה מתקיים: $H\left(\underline{X}\right)<n\cdot H\left(x_1\right)$.

<u>**טענה**</u>:

עבור מקור סטציונרי, הסדרה $H\left(X_n\,|\,X_{n-1},\ldots,X_1\right)$ יורדת עם n.

<u>הוכחה</u>:

נביט בביטוי - $H\left(X_0\,|\,X_{-1},\ldots,X_{-n}\right)$. כאשר נגדיל את n ביטוי זה מקבל התניות נוספות ועקב כך ערכו יורד עם הגדלת n. עקב סטציונריות מתקיים - $H\left(X_{n+1}\,|\,X_1,\ldots,X_n\right)=H\left(X_0\,|\,X_{-n},\ldots,X_0\right)$ ולכן גם הסדרה $H\left(X_n\,|\,X_{n-1},\ldots,X_1\right)$ יורדת עם הגדלת ערכו של n.

<u>מסקנה</u>: קיים גבול לסדרה.

<u>הגדרה</u>:

נגדיר עבור מקור סטציונרי את שני הגדלים הבאים:

$$H_\infty^{(I)}\triangleq\lim_{n\to\infty}H\left(X_n\,|\,X_{n-1},\ldots,X_1\right)$$

$$H_\infty^{(II)}\triangleq\lim_{n\to\infty}\frac{1}{n}H\left(X_1,\ldots,X_n\right)$$

<u>דוגמא</u>: עבור מקור מרקובי סטציונרי מסדר ראשון: $H_\infty^{(I)}=H\left(X_2\,|\,X_1\right)$.

נשים לב שעבור מקור חסר זיכרון: $H_\infty^{(I)}=H_\infty^{(II)}=H\left(X_1\right)$.

<u>**טענה**</u>:

עבור מקור סטציונרי מתקיים: $H_\infty^{(I)}=H_\infty^{(II)}=\bar{H}$ כאשר $\bar{H}$ הינו קצב האנטרופיה.

<u>**טענה**</u>:

אם המקור הינו סטציונרי וארגודי, אזי מתקיים בו AEP, כלומר:

$$\lim_{n\to\infty}-\frac{1}{n}\log p\left(X_1,\ldots,X_n\right)=\bar{H}\ \text{ in prob.}$$

<u>הערה</u>: למעשה, ה- AEP (גם למקורות חסרי זיכרון וגם מקורות סטציונרים כלליים) מתקיים במובן החזק יותר של התכנסות בהסתברות 1.

<u>מסקנה</u>: ניתן להגדיר את הקבוצה האופיינית כפי שהיא הוגדרה בפרקים הקודמים והיא תקיים את כל תכונותיה, למרות שהמקור הוא לא חסר זיכרון.

6. קידוד מקור באורך משתנה

עד כה, לקחנו דגימות מקור ומיפינו כל וקטור של דגימות לוקטור של ביטים, כלומר, אורכי הסדרות n ו-k היו קבועים לאורך הקידוד. אפשר גם לדבר גם על קידודים בעלי אורכים לא קבועים. קיימים 3 סוגים של קידוד באורך משתנה:

- ✓ Fixed to variable – המילים במוצא המקודד הן באורך משתנה.
- ✓ Variable to fixed – המילים בכניסה למקודד הן באורך משתנה.
- ✓ Variable to variable – גם בכניסה וגם במוצא המקודד – המילים הן באורך משתנה.

בקורס זה נתמקד באפשרות הראשונה – המילים בכניסה למקור הן באותו האורך, והמקודד משדר מילים בעלות אורך משתנה.

<u>דוגמא</u>:

נבחן מספר דוגמאות לקידוד באורך משתנה. נניח $n=1$. בנוסף:

$$\chi=\{1,2,3,4\}$$
$$f(x)=\underline{b}=b_1,\dots,b_{k(x)}$$
$$f(x):\chi\to\{0,1\}^*$$

משמעותה של הכוכבית היא שהאורך של הקידוד עשוי להשתנות.

X	קוד 1	קוד 2	קוד 3	קוד 4
1	0	0	10	**0**
2	0	010	00	**10**
3	0	01	11	**110**
4	0	10	110	**111**

- ✓ <u>קוד 1</u>: זהו קוד סינגולרי. לא ניתן לפענח ממנו איזה סימבול שודר.

הקודים הבאים הם <u>לא</u> סינגולריים, אבל נבחן האם הם ברי-פענוח כאשר משרשרים בזו אחר זו מילות קוד של דגימות מקור עוקבות:

- ✓ <u>קוד 2</u>: זהו קוד לא סינגולרי אך איננו בר פענוח יחיד. למשל בקבלת הסדרה 010 לא ברור אם שודר 010 או 01,0.
- ✓ <u>קוד 3</u>: זהו קוד בר-פענוח יחיד אך הוא לא בר-פענוח מיידית. למשל, נוכל להבדיל בין הסדרות ...11,00,00,00 ו- ...110,00,00,00 רק כשיגיע הסימבול שייסים את סדרת האפסים והדבר לא בהכרח מיידי.
- ✓ <u>קוד 4</u>: זה קוד שהוא בר פענוח ומיידי (Instantaneous).

<u>**התכונה של קוד בר-פענוח מיידי**</u>:

קוד בר-פענוח מיידי מקיים את תנאי הקידומת (prefix free code) – <u>אף מילת קוד איננה תחילית במילת קוד אחרת</u>.

<u>הגדרה</u>:
<u>הקוד האופטימלי</u> מוגדר בצורה הבאה:

$$\bar{L}^{opt}\left(\underline{p}\right)\triangleq\min_{\{l_1,\dots,l_n\}}\sum_{i=1}^{n}p_i\cdot l_i=\min_{\{l_1,\dots,l_n\}}El$$

כאשר $l_1,\dots,l_n$ הם אורכים אפשריים לקוד בר-פענוח. כאשר p_i היא הסתברות אות המקור ה- i ו- l_i הוא אורך מילת הקוד (מספר הביטים) שמתאימה לאות המקור ה- i. בעצם, הקוד האופטימלי הוא קוד בו האורך הממוצע של המילים בקוד הוא הקצר ביותר.

באופן דומה ניתן להגדיר את ספר הקוד האופטימלי:

$$\mathbb{C}^{opt}\left(\underline{p}\right)\triangleq\arg\min_{\{l_1,\dots,l_n\}}\sum_{i=1}^{n}p_i\cdot l_i=\arg\min_{\{l_1,\dots,l_n\}}El$$

<u>שאלות מתבקשות</u>:

א. אילו אורכים $l_1,\dots,l_n$ אפשריים?
ב. האם גם עבור קודים באורך משתנה מתקיים $\bar{L}\geq H$? (כלומר, האם גם במקרה זה האנטרופיה היא האורך המינימלי האפשרי?)
ג. איך מוצאים את הקוד האופטימלי?

6.1. אי-שוויון קראפט (Kraft)

נניח קידוד בקוד בינארי המקיים את תנאי הקידומת למילות קוד באורכים $l_1,\dots,l_M$.

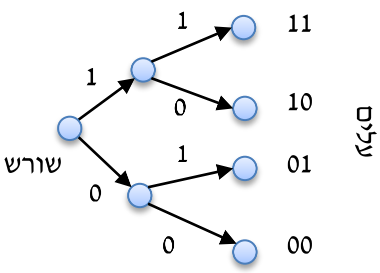
<u>טענה</u>: הקוד מקיים את תנאי הקידומת אמ"מ $\sum_{i=1}^{M}2^{-l_i}\leq 1$.

בפרט, אם $\sum_{i=1}^{M}2^{-l_i}\leq 1$, אזי ניתן למצוא קוד קידומת באורכים $l_1,\dots,l_M$.

<u>הערה</u>: אי-שוויון קראפט הוא תנאי שאינו קשור למקור מסוים או להסתברויות מסוימות – כפי שניתן לראות – הוא אינו מדבר על הסתברות כלל!

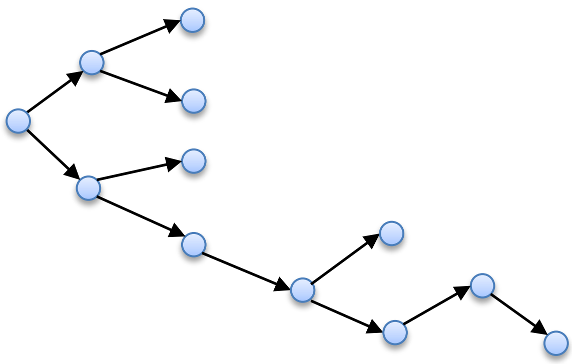
6.1.1. שקילות קוד קידומת לעץ בינארי

ניתן להסתכל על קוד קידומת כעל ענפים בעץ בינארי. למשל:



זהו עץ בינארי המתאר קוד בעל אורך קבוע. ה-MSB יוצא מהשורש של העץ ומסתיים בעלה שהוא ה- LSB של המילה. תנאי הקידומת שקול לכך שצמתים פנימיים לא מייצגים מילת קוד.

נביט בעץ הבא:

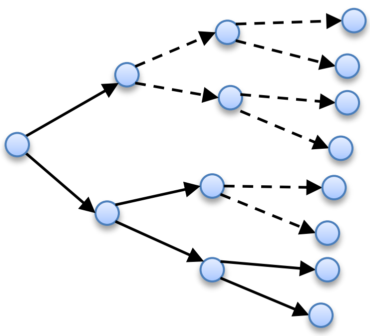


עץ זה מכיל צמתים בהם קיים פיצול יחיד. מבחינה אינפורמציונית צמתים אלה הם מיותרים וניתן להסתפק בצומת האחרונה שנבעה מ-2 פיצולים. נעדיף לעבוד עם עץ שהוא "<u>גזום למשעי</u>" (לא מונח מקצועי – אלא כינוי שהמציא רמי זמיר). עץ גזום למשעי זהו עץ בינארי שמכל צומת (פרט לעלים) יוצאים 2 ענפים בדיוק.

<u>הגדרה</u>:

$l_{\max}$ - עומק העץ. זהו בעצם אורך המסלול (מן השורש עד העלה) הארוך ביותר בעץ.

6.1.2. הוכחת אי-שוויון קראפט



נבנה על כל עלה בעץ שנמצא באורך הקטן מ- $l_{\max}$ "תת-צמרת" שתשלים אותו כך שייצא לנו עץ בינארי מלא (כפי שמודגם בציור). גודל תת הצמרת (מספר העלים) שבנויה על כל מילה באורך l_i הוא: $2^{l_{\max}-l_i}$. סה"כ עלים שנוספו לעץ אחרי בנייתן של תת הצמרות קטן או שווה ל- $2^{l_{\max}}$. או בשפה מתמטית:

$$\sum_{i=1}^{M} 2^{l_{\max}-l_i} \le 2^{l_{\max}} \Rightarrow \sum_{i=1}^{M} 2^{-l_i} \le 1$$

מההוכחה ניתן לראות שעבור עץ גזום למשעי, אי- השוויון מתקיים בשוויון.

6.2. קוד שאנון-פאנו (Shannon-Fano)

כזכור, לפי שאנון - $H = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$. נשווה גודל זה עם - $\bar{L} = \sum_{i=1}^{n} p_i \cdot l_i$. השאלה המתבקשת היא האם $\log\left({}^{1}\!/_{p_i}\right)$ מקיים את אי-שוויון קראפט? התשובה היא שכן (משום ש- $\sum_{i=1}^{n} 2^{-l_i} = \sum_{i=1}^{n} p_i = 1$), אך $\log\left({}^{1}\!/_{p_i}\right)$ איננו מס' שלם כמו l_i עבור נוסח האי-שוויון. לכן, שאנון ופאנו הציעו את האורכים הבאים עבור הקידוד:

$$l_i \triangleq \left\lceil \log \frac{1}{p_i} \right\rceil$$

מצד אחד, אחרי העיגול הערכים עדיין מקיימים את אי-שוויון קראפט. מצד שני, מתקיים - $l_i \le \log\frac{1}{p_i}+1$. ומכאן נובע כי: $H \le \sum_{i=1}^{n} p_i \cdot l_i \le H+1$.

מאי-שוויון קראפט ניתן להגדיר שוב את הקוד האופטימלי:

$$\bar{L}^{opt}\left(\underline{p}\right) \triangleq \min_{\left\{l_1,\ldots,l_n;\sum_{i=1}^{n} 2^{-l_i}\le 1\right\}} \sum_{i=1}^{n} p_i \cdot l_i$$

<u>הערה</u>: היות והקוד האופטימלי הוא טוב לפחות כמו קוד שאנון-פאנו, אז גם לגביו מתקיים:

$$H\left(\underline{p}\right) \le \bar{L}_{opt}\left(\underline{p}\right) \le H\left(\underline{p}\right)+1$$

עבור קוד בלוק באורך n מתקיים כי האורך הממוצע של כל מילות הקוד ביחד מקיים:

$$nH \le L_{total} \le nH+1$$

ולכן האורך הממוצע למילת קוד מקיים:

$$H \le \bar{L} \le H+\frac{1}{n}$$

6.3. השלכה של אי-שוויון קראפט לגבי עודף אורך הקוד הממוצע מעל האנטרופיה

נניח כי $l_1,\ldots,l_M$ מקיימים את אי-שוויון קראפט. נגדיר:

$$q_i \triangleq \frac{2^{-l_i}}{\underbrace{\sum_{j=1}^{M} 2^{-l_j}}_{\triangleq K}}$$

מתקיים כי: $K \le 1$ ושוויון מתקיים אמ"ם הקוד מתאים לעץ גזום למשעי. נשים לב כי $\underline{q}=q_1,\ldots,q_M$ הינו וקטור הסתברות חוקי.

<u>**טענה**</u>:

בהנתן מקור עם פילוג $\underline{p}=p_1,\ldots,p_M$, מתקיים:

$$\bar{L}=\sum_{i=1}^{M} p_i \cdot l_i = H\left(\underline{p}\right)+\underbrace{\underbrace{D\left(\underline{p}\,\|\,\underline{q}\right)}_{\ge 0}+\underbrace{\log\frac{1}{K}}_{\ge 0}}_{\text{Redundency of the code over the entropy}}$$

<u>מסקנה</u>: מתקיים כי $\bar{L} \ge H$ (אי-אפשר לרדת נמוך יותר מהאנטרופיה) ושוויון מתקיים אמ"ם הקוד הינו גזום למשעי ובנוסף $p_i = 2^{-l_i}$.

<u>**טענה**</u>:

קוד באורך משתנה יכול להשיג את חסם האנטרופיה כאשר מקודדים וקטורים ארוכים של דגימות מקור.

<u>הוכחה</u>:

קוד שאנון-פאנו מקיים: $\bar{L} \le H\left(X\right)+1$. עבור $n>1$, מתקיים הקשר הבא עבור קצב האינפורמציה:

$$R \triangleq \frac{1}{n}\cdot\bar{L} \le \frac{1}{n}H\left(X_1,\ldots,X_n\right)+\frac{1}{n}\xrightarrow[n\to\infty]{}\bar{H}$$

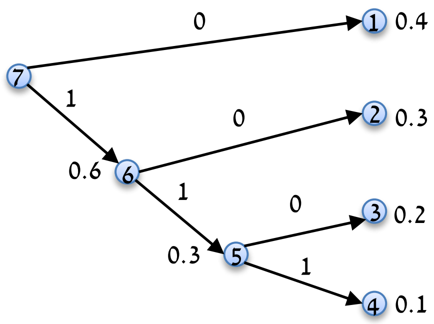
6.4. קוד האפמן (Huffman)

קוד האפמן הינה שיטה לבנייה של קוד prefix באורך משתנה שהוא הקוד האופטימלי, כלומר משיג את המינימום של האורך הממוצע. שיטה זו הומצאה ע״י האפמן ב-1952. נדגים שיטה זו בדוגמא הבאה:

נניח כי עלינו לקודד מקור הפולט 4 דגימות בהסתברויות הבאות: $\{0.4, 0.3, 0.2, 0.1\}$. השלבים בבניית הקוד הם כדלהלן:

1. מסדרים את ההסתברויות הנ״ל בעמודה – מהנמוכה עד הגבוהה. אלה הם העלים של העץ הבינארי.
2. מאחדים את 2 האירועים בעלי ההסתברויות הנמוכות ביותר לאירוע בודד. בעץ הבינארי מאחדים את 2 הצמתים הנ״ל לצומת יחיד המהווה את סכומם.
3. ממשיכים בשלב 2 עד אשר העץ נבנה במלואו.
4. כאשר העץ בנוי, מקצים את הביט '0' לכל ענף שעולה למעלה מצומת ואת הביט '1' לכל ענף שיורד מטה (בפועל, אפשר לוותר על סעיף זה ולבצע את ההקצאה הנ״ל בכל דרך שרוצים). הולכים משמאל לימין ומרכיבים את מילות הקוד.

למשל עבור הדוגמא הנ״ל:



מילות הקוד עבור העץ הנ״ל הן: $\{0,10,110,111\}$. האנטרופיה של המקור הנ״ל היא: $H = 1.846\,\text{bit}$.

האורך הממוצע של הקוד הנ״ל הוא: $\bar{L} = 1.9\,\text{bit}$. העץ הנ״ל הוא גזום למשעי ולכן - $\sum_{i=1}^{n} 2^{-l_i} = 1$.

<u>הערות</u>:

א. העץ הבינארי המתאים למקור מסוים אינו יחיד! אם קיים מקרה בו קיימים מס' איחודים אפשריים שסכומם שווה, אז לקוד זה יהיו מספר עצים בינאריים אפשריים. יש לציין שלכולם יהיה אותו אורך קוד ממוצע.

ב. ניתן להשיג את האנטרופיה ע״י קידוד האפמן אם הסתברויות המקור הן חזקות של 0.5. למשל, עבור המקור $\{0.5, 0.25, 0.125, 0.125\}$ ניתן לבנות עץ בינארי כמו בדוגמא הנ״ל ולקבל: $\bar{L} = 1.75\,\text{bit} = H$. (במקרה זה, קוד האפמן וקוד שאנון-פאנו מתלכדים!)

ג. עבור קוד בו $p_j > p_k$ מתקיים כי $l_j \le l_k$. כלומר, לא יתכן שדגימת מקור שהסתברותה נמוכה יותר מדגימת מקור אחרת תקבל מילת קוד קצרה יותר ממנה.

ד. שתי מילות הקוד המייצגות את דגימות המקור בעלות ההסתברויות הנמוכות ביותר הן באותו האורך, והן נבדלות ביניהן רק בביט האחרון.

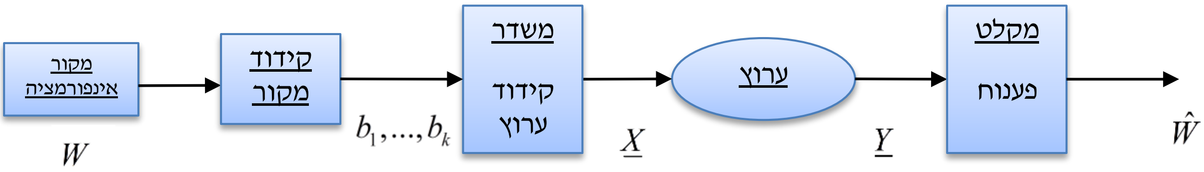
ה. כאשר בונים קוד האפמן מסדר r (לאו דווקא בינארי), התהליך הוא זהה למעט הנקודות הבאות:

i. במקרה בו מס' דגימות המקור שיש לקודד אינו $r + n\cdot(r-1)$ כאשר n הינו מספר טבעי כלשהו, יש להוסיף למקור זה מספר מילות מקור ״דמה״ על-מנת להשלים לערך הנ״ל ולתת לכל דגימת מקור שכזו הסתברות 0.

ii. קוד האפמן במקרה ה- r -י ישיג את האנטרופיה כאשר הסתברויות דגימות המקור הן חזקות של $1/r$.

7. קיבול של ערוץ רועש

ניזכר בסכימה של ערוץ התקשורת:



בפרק הקודם התמקדנו בקידוד מקור האינפורמציה. בפרק זה נעסוק בקידוד הערוץ.

<u>מטרות המשדר/מקודד</u>:

א. התאמת א"ב מהאות המקור לכניסת הערוץ.
ב. התאמה לאילוצי הכניסה של הערוץ (לדוגמא: אילוץ הספק או אילוץ רצפים חוקיים).
ג. יצירת אותות (מילות קוד) חסינות לרעש (ברות-פענוח).

<u>דוגמא 1</u>:

נביט בערוץ BSC עם הסתברות להחלפת ביט ε. נקרא לרצף בכניסה לערוץ b_n ולרצף המוצא $\hat{b}_n$. כזכור, בערוץ זה ניתן להעביר ביט יחיד לשימוש ערוץ, כלומר:

$$R = 1\frac{bit}{channel\ use}$$

נגדיר <u>הסתברות שגיאה לביט אינפורמציה</u>:

$$p_{e,bit} \triangleq p\left(\hat{b}_n \neq b_n\right)$$

ובמקרה של הערוץ שלנו:

$$p_{e,bit} = \varepsilon$$

כמו כן, <u>הסתברות השגיאה לכל בלוק האינפורמציה</u>:

$$p_{e,block} \triangleq p\left(\underline{b} \neq \hat{\underline{b}}\right) = 1 - p_{correct} = 1-(1-\varepsilon)^m \xrightarrow[m\to\infty]{} 1$$

<u>הערה</u>: נניח כי $\varepsilon \le 0.5$, כי אם אחרת אזי עדיף עבור כל מוצא לבחור את ההיפך ממה שהתקבל.

על-מנת להקטין את הסתברות השגיאה ניתן לשדר כל ביט מספר פעמים. כלומר, נשכפל כל ביט מוצא m פעמים ונשדר ביטים אלה בערוץ. כמובן שהגיוני במקרה זה להחליט איזה ביט שודר ע"פ רוב הביטים שהתקבלו. כלומר, אם התקבלה סדרה שרוב הביטים בה הם 0 – נחליט ששודר 0. הסתברות השגיאה במקרה זה תהיה ההסתברות שהתחלפו יותר ממחצית הביטים, כלומר:

$$p_e = \sum_{k>\frac{m}{2}} \binom{m}{k} \varepsilon^k \cdot (1-\varepsilon)^{m-k} \le \frac{m}{2} \cdot \underbrace{\binom{m}{m/2}}_{\approx 2^{m\cdot H_B\left(\frac{1}{2}\right)} = 2^m} \varepsilon^{\frac{m}{2}} \cdot (1-\varepsilon)^{\frac{m}{2}}$$

$$\le m \cdot \underbrace{\left(2\sqrt{\varepsilon(1-\varepsilon)}\right)}_{<1}^{m} \xrightarrow[m\to\infty]{} 0$$

הקצב במערכת הנ"ל:

$$R = \frac{1\text{information bit}}{m \text{ channel uses}} = \frac{1}{m} \xrightarrow[m\to\infty]{} 0$$

אמנם הסתברות השגיאה קטנה כרצונינו, אך ככל ש- m עולה כך גם קצב הקוד יורד ל-0. נרצה קוד שהסתברות השגיאה תהיה קטנה כרצונינו ושקצב הקוד לא ישאף לאפס.

<u>דוגמא</u>: שידור BPSK על ערוץ AWGN
כזכור, בערוץ זה מתקיים כי הסתברות השגיאה לביט היא:

$$p_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

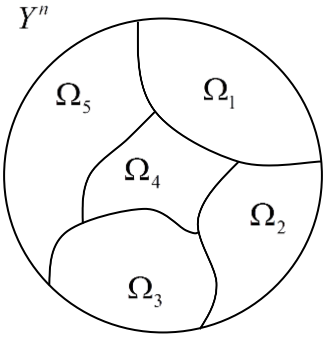
שימוש בקוד חזרות כפי שהוצג בדוגמא הקודמת שקול להארכת זמן הביט T_b. ולכן גם האנרגיה לביט E_b גדלה. ולפי הנוסחא שהוצגה, ניתן להיווכח כי הסתברות השגיא יורדת. גם כאן ככל שנגדיל את T_b הסתברות השגיאה תרד אך כך גם קצב השידור.

<u>הגדרת התרחיש</u>:
ע"פ הדיאגרמה שהוצגה בתחילת הפרק, נגדיר פורמלית את המשתתפים בתרחיש:

- ✓ <u>מקור האינפורמציה</u> פולט הודעה W - $W \in \{1,\ldots,M\}$.
- ✓ <u>המקודד</u> מתאים את המילה W לאחת ממילות ספר הקוד $\underline{X} \in X^n$ ע"י פונקציה כלשהי $f(\cdot)$.

$$\underline{X} = f(W)$$
$$\mathbb{C} \triangleq \{\underline{X}_1,\ldots,\underline{X}_M\} \leftarrow \text{ Code Book}$$
$$\underline{X}_i = f(i)$$

- ✓ <u>הערוץ</u> הינו התאמה הסתברותית כלשהי בין כניסתו $\underline{X}$ ליציאתו $\underline{Y} \in Y^n$ - $p(\underline{Y} \mid \underline{X})$.
- ✓ <u>המפענח</u> מבצע מה שנקרא "many to one mapping". הוא פונקציה כלשהי $g(\cdot)$ הממפה קבוצות של וקטורי $\underline{Y}$ להחלטה על המילה ששודרה: $\hat{W} = g(\underline{Y})$.



תחומי החלטת המפענח הם:

$$\Omega_1,\ldots,\Omega_M ; \Omega_i \subset Y^n$$

המפענח פועל בצורה הבאה:

$$\underline{Y} \in \Omega_i \Rightarrow \hat{W} = g(\underline{Y}) = i$$

<u>מאפייני המערכת</u>:

- ✓ <u>קצב</u>:

$$R \triangleq \frac{\log M}{n}\left[\frac{\text{bit}}{\text{channel use}}\right]$$

- ✓ <u>הסתברות שגיאה להודעה</u>:

$$p_e(i) = p\left\{\hat{W} \neq W \mid W = i\right\} = p\left(\underline{Y} \notin \Omega_i \mid W = i\right)$$
$$P_{e,\max} = \max_{i=1,\ldots M} p_e(i) \leftarrow \text{Maximal error probability}$$
$$\overline{p}_e = \frac{1}{M}\sum_{i=1}^{M} p_e(i) \leftarrow \text{Average error probability}$$

הערה: השימוש במיצוע אחיד עבור $\bar{p}_e$ הוא שרירותי – ללא קשר לפילוג של ההודעה W בפועל.

מטרתינו היא למצוא קשר מיטבי בין n, R והסתברות השגיאה המקסימלית או הממוצעת.

7.1. משפט הקידוד השני של שאנון (משפט הקידוד לערוץ רועש)

הגדרה:

קצב בר-השגה – נאמר ש- R הינו קצב בר-השגה בערוץ $p\left(\underline{Y} \mid \underline{X}\right)$ נתון, אם קיימת סדרה של מערכות $\left\{n, R, f\left(\cdot\right), g\left(\cdot\right)\right\}$ עבור $n \to \infty$ כך ש:

$$p_{e,\max} \xrightarrow[n \to \infty]{} 0$$

במילים אחרות, לכל $\varepsilon > 0$ ניתן למצוא מערכת עם אורך בלוק n מספיק גדול כך ש- $p_{e,\max} < \varepsilon$.

הגדרה:

קיבול אופרטיבי הינו הקצב בר-ההשגה המקסימלי:

$$C^{op} = \sup\left\{R : R \text{ is achievable}\right\}$$

הגדרה:

קיבול אפס שגיאה - C_0, מוגדר כמו קיבול אופרטיבי רק שכאן דורשים: $p_e \equiv 0$.

הגדרה:

קיבול אינפורמציוני הוא גודל שהגדיר שאנון במטרה לתת נוסחא פשוטה לקיבול האופרטיבי.

עבור ערוץ בדיד חסר זיכרון (DMC), הקיבול האינפורמציוני מוגדר כך:

$$C^{\inf} = \max_{p(X)} I\left(p\left(X\right), p\left(Y \mid X\right)\right) = \max_{p(X)} I\left(X;Y\right)$$

משפט הקידוד השני של שאנון:

$$C^{op} = C^{\inf}$$

הערות:

א. הקיבול האופרטיבי מתייחס לקוד בלוק באורך n $\left(n \to \infty\right)$. למרות זאת, עבור DMC, מספיק לחשב את הקיבול האינפורמציוני עבור $n = 1$ (החישוב הוא חד-מימדי, כאשר המערכת המשיגה אותו היא $n = \infty$ מימדית).

ב. המשפט הנ"ל מדבר על הסתברות שגיאה לבלוק! כלומר, אם ההודעה היא $W \in \left\{1, 2, \ldots, 2^{nR}\right\}$, כלומר בלוקים של nR ביטים, אזי אם נפלה שגיאה בביט אחד אז זוהי טעות בכל הבלוק.

ג. האינפורמציה ההדדית הינה פונקציה קמורה $\cap$ בפילוג המבוא. לכן קיים פילוג $p^*\left(X\right)$ שמביא אותה למקסימום (בד"כ $p^*\left(X\right)$ הינו יחיד ואינו נמצא על השפה של הסימפלקס). לפילוג זה קוראים "*הפילוג המגשים*".

ד. מתקיים כי $I\left(X;Y\right) \le \log\left|X\right|, \log\left|Y\right|$. ולכן מתקיים: $C^{\inf} \le \log\left|X\right|, \log\left|Y\right|$.

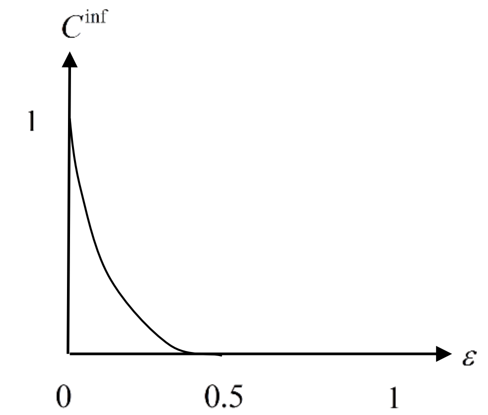
דוגמאות:

א. ערוץ BSC:

$$C^{\inf} = \max_{p(X)} I(X;Y) = \max_{p(X)}\left[H(Y) - H(Y\mid X)\right] = \max_{p(X)}\left[H(Y) - H_B(\varepsilon)\right]$$
$$= \max_{p(X)}\left[H(Y)\right] - H_B(\varepsilon) \le 1 - H_B(\varepsilon)$$

האי-שוויון בסוף הוא בר השגה עבור:

$$p^*(X) = \left(\frac{1}{2},\frac{1}{2}\right)$$
$$\Rightarrow C^{\inf} = 1 - H_B(p)$$



ב. ערוץ מודולו אדיטיבי:

$$Y = (X + Z) \bmod q$$
$$X, Z \in \chi$$
$$|\chi| = q$$
$$C^{\inf} = \max\left[H(Y)\right] - H(Z) \le \log q - H(Z)$$

ועבור הפילוג האחיד: $p^*(Y) = unif[1,q]$ מתקיים: $H(Y) = \log q$. ולכן:

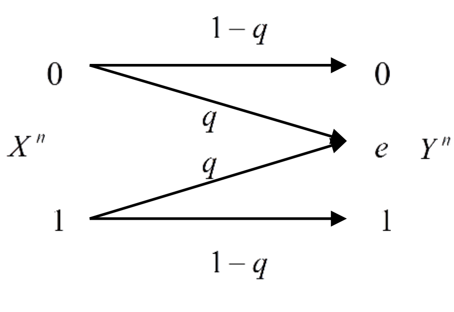
$$C^{\inf} = \log q - H(Z)$$

ג. ערוץ עם מחיקות:

זהו ערוץ שבו הביט מגיע בהצלחה או מגיע בערך לא ברור הנקרא "מחיקה".

מתקיים:

$$C^{\inf} \le \log 2 = 1bit$$



המקסימום מושג ע"י פילוג אחיד על ביטי הכניסה. ואז:

$H(X)=1$ ו- $H(X\mid Y) = (1-q)\cdot 0 + q\cdot 1 = q$.

ומכאן נובע:

$$C^{\inf} = H(X) - H(X\mid Y) = 1-q$$

זה הגיוני כיוון שתוך n ביטים המועברים בערוץ רק $n(1-q)$ "יגיעו בשלום".

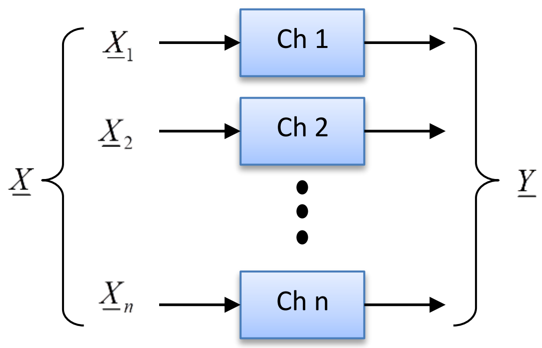
ולכן:

$$R \le \underbrace{\frac{n(1-q)}{n}}_{\substack{\text{the case in which the decoder}\\ \text{knows where the erasures occur}}} = 1-q\,\frac{\text{bits}}{\text{channel use}}$$

אם קיים משוב, בכדי להעביר n ביטים במלואם, נצטרך לשדר בערך $\frac{n}{1-q}$ ביטים (כולל retransmissions). הנוסחא של $C^{\inf}$ מוכיחה כי גם <u>ללא משוב</u> ניתן להשיג אסימפטוטית קצב של $1-q$.

ד. <u>ערוץ סכום</u>:
אם ניתן לחלק את הערוץ הנתון למספר תתי ערוצים בלתי תלויים אחד בשני, אזי הקיבול הכולל נתון כפונקציה של הקיבולים של תתי- הערוצים:

$$C = \log \sum_{i=1}^{n} 2^{C_i}$$



ניתן לרשום את הפילוג המגשים הכולל כפונקציה של הקיבול הכולל, הקיבולים של תתי-הערוצים והפילוג המגשים בתתי-הערוצים:

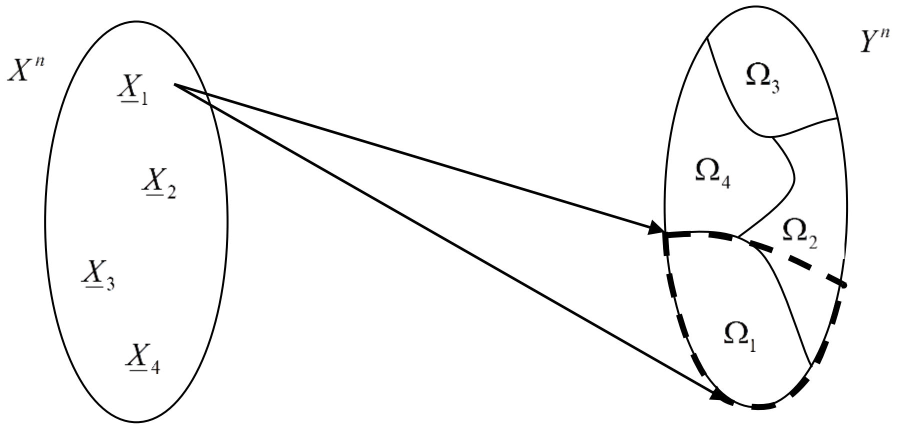
$$p^*\left(X\right) = 2^{-(C-C_i)} \cdot p^*\left(X_i\right)$$

7.2. "המניפות של שאנון"

ההוכחה מבוססת על עיקרון "המניפות של שאנון". קיימות 2 דרכים להסתכל על הבעיה – "הערוץ הישר" "והערוץ ההפוך".

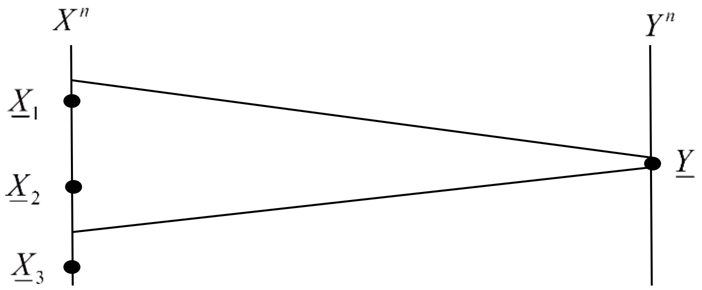
<u>הערוץ הישר</u>:

הערוץ הינו העתקה מ- X^n ל- Y^n. כל $\underline{X}_i \in X^n$ מועתק לאיזור כלשהו $\Omega \subset Y^n$.



באילוסטרציה הנ"ל $\underline{X}_1$ הועתק לאיזור Ω שכולל בתוכו גם חלק מ- Ω_2. החלק שהועתק ל- Ω_2 הינו הסתברות השגיאה.

הערוץ ההפוך:



בניגוד לשיטה הישרה שבה המניפה הייתה קדמית (מ- X^n לכיוון Y^n), במקרה זה המניפה היא אחורית. לכל $\underline{Y}$ מותאמים $\underline{X}_i$ שהיו עשויים לגרום לו.

7.2.1. אקספוננט סף-ההצלחה

אם מבצעים 2^{nR} ניסויים בת"ס (iid) עם סיכוי הצלחה 2^{-nr} בכל ניסוי. אז ההסתברות P להצלחה באיזשהו ניסוי תהיה:

$$P \xrightarrow[n\to\infty]{} \begin{cases} 1 & R>r \\ 0 & R<r \end{cases}$$

הוכחה:

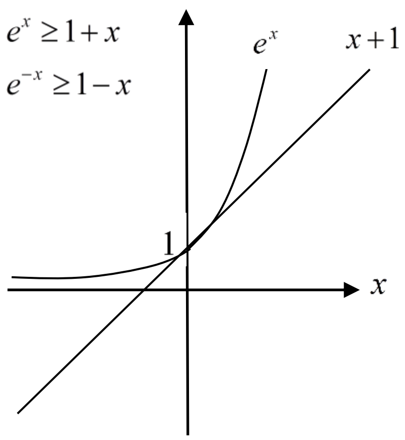
נגדיר E_i - המאורע של ההצלחה בניסוי ה- i, ו- $P_i = p\left(E_i\right)$ היא ההסתברות של מאורע זה.

$$P = p\left\{\bigcup_{i=1}^{2^{nR}} P_i\right\} \underset{\substack{\text{union}\\ \text{bound}}}{\leq} \sum_{i=1}^{2^{nR}} P_i = 2^{nR}\cdot 2^{-nr} = 2^{-n(r-R)} \xrightarrow[\substack{n\to\infty\\ R<r}]{} 0$$

הוכחנו את השוויון כאשר $R<r$. כעת נבחן את המאורע שהיה כישלון בכל הניסויים - $\overline{P}$:

$$\overline{P} = \prod_{i=1}^{2^{nR}} \overline{P}_i = \prod_{i=1}^{2^{nR}} \left(1-P_i\right) = \left(1-2^{-nr}\right)^{2^{nR}} \underset{e^{-x}\geq 1-x}{\leq} \left(e^{-2^{-nr}}\right)^{2^{nR}} = e^{-2^{n(R-r)}} \xrightarrow[\substack{n\to\infty\\ R>r}]{} 0$$

ובמקרה זה הסתברות P של המאורע המשלים שואפת ל-1.



7.3. הוכחת המשפט "בנפנוף ידיים"

נביט עתה על תמונות "המניפות" בערוץ ההפוך מ- Y ל- X :

א. גודל תחום ה- $\underline{X}$ -ים האפשריים ביחס ל- $\underline{Y}$ נתון בהנחת התנהגות "אופיינית" (AEP) הוא : $\approx 2^{nH(X|Y)}$. תחום זה מסומן ע"י 🍎.

ב. גודל תחום ה- $\underline{X}$ -ים האופיניים הכולל הוא : $\approx 2^{nH(X)}$.

ג. הסיכוי שמילת קוד "מתחרה" (שלא באמת שודרה) תיפול בתחום 🍎 היא :

$$\approx \frac{2^{nH(X|Y)}}{2^{nH(X)}} = 2^{-nI(X;Y)}$$

ד. בהנחה של התנהגות "אופיינית", המילה האמיתית ששודרה נמצאת בוודאות בתחום 🍎.

ה. מסי המילים המתחרות הוא $M-1$:

$$M-1=2^{nR}-1\le 2^{nR}$$

ו. ע"פ אקספוננט סף ההצלחה :

$$p_e \xrightarrow[n\to\infty]{} \begin{cases} 1 & R>I(X;Y) \quad \leftarrow \text{error event} \\ 0 & R<I(X;Y) \quad \leftarrow \text{correct decoding} \end{cases}$$

ז. היות ו- $I(X;Y)$ הוא לכל היותר C (ומושג ע"י $p^*(X)$), אזי ניתן להשיג $p_e\to 0$ עבור כל קצב המקיים : $R<C$.

ח. הוכחה זו היא "בנפנוף ידיים" משום שהיא מתעלמת מה- ε במשפט ה-AEP (כאילו שחוק המספרים הגדוליים מתקיים "בדיוק").

ננסח מחדש את <u>משפט הקידוד של שאנון</u> :

<u>המשפט הישר</u> : אם $R<C$, אזי ניתן למצוא מערכת קידוד $(n,R,f(\cdot),g(\cdot))$ עם p_e קטן כרצונינו.

<u>המשפט ההפוך</u> : אם $R>C$, אזי הסתברות השגיאה היא בהכרח גדולה ממש מ-0.

7.4. הוכחה למשפט ההפוך

נניח כי $p_e\equiv 0$. נתונה מערכת בקצב R . נראה שאם $p_e\equiv 0$ אזי בהכרח : $R\le C$.

נגדיר פילוג אחיד על ההודעה W :

$$H(W)=\log M=nR$$

מצד שני :

$$H(W)\overset{(a)}{=}H(W|Y^n)+I(W;Y^n)\overset{(b)}{=}I(W;Y^n)\overset{(c)}{\le}I(X^n;Y^n)\overset{(d)}{\le}\sum_{i=1}^{n}I(X_i;Y_i)\overset{(e)}{\le}n\cdot C$$

ולכן : $R\le C$.

(a) – מהגדרת האינפורמציה ההדדית.

(b) $p_e=0$ ולכן W ידוע דטרמיניסטית מתוך ידיעת Y^n, ולכן : $H(W|Y^n)=0$.

(c) - אי שוויון עיבוד הנתונים : $W\leftrightarrow X^n\leftrightarrow Y^n$.

(d) -

$$I(X^n;Y^n)=H(Y^n)-H(Y^n|X^n)\underset{\substack{\text{DMC}\\ \text{without feedback}}}{=}H(Y^n)-\sum_{i=1}^{n}H(Y_i|X_i)=$$

$$\underset{\substack{\text{mutual entropy is}\\ \text{smaller or equal to the}\\ \text{sum of the marginal entropies}}}{\leq} \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i \mid X_i) = \sum_{i=1}^{n} I(X_i ; Y_i)$$

(e) - כל אחד מאיברי הסכום חסום מלעיל ע"י $C^{\inf}$ (ע"פ ההגדרה).

כעת, נגדיר את <u>אי-הודאות הנותרת</u> כ- $H(W \mid Y^n)$. מההוכחה הקודמת נובע כי:

$$H(W \mid Y^n) \geq n \cdot (R - C)$$

7.4.1. <u>אי-שוויון פאנו</u> (הקשר בין אי-הודאות הנותרת לבין הסתברות שגיאת הגילוי)

נתונים זוג משתנים אקראיים A, B. נתייחס ל- A בתור האות הרצוי, ל- B בתור המדידה ול- $\hat{A}(B)$ כגלאי של A מתוך B. נגדיר הסתברות שגיאה: $p_e = p(\hat{A} \neq A)$. לכל גלאי $\hat{A}(B)$ מתקיים:

$$\underline{H(A \mid B) \leq H_B(p_e) + p_e \cdot \log(|A| - 1)}$$

נציב אי-שוויון זה בביטוי לאי-הודאות הנותרת:

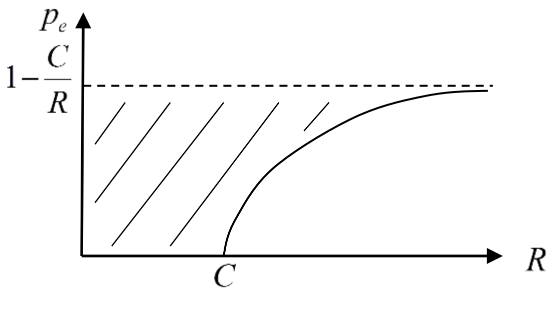
$$H(W \mid Y^n) \leq H_B(p_e) + p_e \log(2^{nR} - 1) \leq 1 + p_e \log(2^{nR}) = 1 + n \cdot R \cdot p_e$$

וביחד עם העובדה כי: $H(W \mid Y^n) \geq n \cdot (R - C)$ מתקבל:

$$nR \leq 1 + p_e \cdot nR + nC$$

$$\Rightarrow p_e \geq 1 - \frac{C}{R} - \frac{1}{nR} \xrightarrow[n \to \infty]{} 1 - \frac{C}{R} \xrightarrow[R \to \infty]{} 1$$

קיבלנו חסם תחתון עבור הסתברות השגיאה להודעה כפונקציה של הקצב והראינו שכאשר $R > C$ הסתברות השגיאה בהכרח חיובית!



<u>הערות</u>:

קיימות מספר דרכים לכמת "אמינות" של מערכת:

א. BER – Bit Error Rate. אחוז הביטים השגויים. דרך שמאוד פופולרית בתעשייה.

ב. $p_{e,\text{block}}$ - מחלקים את מידע הקלט והפלט לחבילות (פאקטות), וערך זה סופר את אחוז החבילות שהגיעו שגויות.

ג. $p_{e,\text{erasures}}$ - אחוז החבילות שאבדו או הגיעו לא תקינות. או לחלופין, אחוז ה- retransmissions.

ד. T_{delay} - זמן מרגע שידור ביט/חבילה עד אשר היא פוענחה באופן תקין.

המשפט של שאנון מדבר על הסתברות שגיאה לבלוק. ההודעה המשודרת W היא בלוק של nR ביטים, כאשר - $W \in \{1,\ldots,2^{nR}\}$. אם ההודעה פוענחה אפילו בביט אחד שגוי, אז מוגדר כי הייתה טעות.

מההוכחה הקודמת ראינו כי אם n גדול מאוד אז מתקיים - $p_{e,n-\text{block}} \geq 1-\frac{C}{R}$. נשאלת השאלה האם כאשר n הוא סופי, אולי ניתן למצוא ספר קוד שיניב $R>C$ והסתברות שגיאה 0. טענה זו נסתרת ע"י שרשור של M קידודים באורך n לבלוק ארוך המכונה "סופר-בלוק" שגודלו מקיים - $N=M\cdot n$. הסופר בלוק מקיים את אי-שוויון פאנו ולכן מתקיים:

$$p_{e,N\text{-block}} = \left\{ \begin{array}{c} \text{error probability somewhere} \\ \text{in the big block} \end{array} \right\} \geq 1-\frac{C}{R} \underset{R>C}{>} 0$$

7.5. Joint – AEP – וקבוצה אופיינית במשותף

ההגדרות בסעיף זה מאוד דומות למה שכבר הוגדר עבור המקרה החד-מימדי. נבצע את ההגדרות בסעיף זה עבור פילוגים מהצורה $p(\underline{X},\underline{Y})=\prod_{i=1}^{n} p(X_i,Y_i)$, כלומר, הזוגות $(X_i,Y_i),(X_j,Y_j)$ הם iid לכל $i \neq j$, אך קיימת תלות בין X_i ל- Y_i.

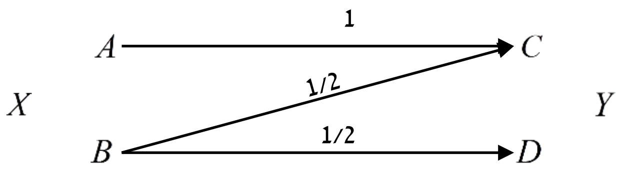
<u>הגדרה</u>:

קבוצה אופיינית ε - במשותף (במובן החלש):

$$A_{\varepsilon}^{(n)}(X,Y)=\left\{ (\underline{X},\underline{Y}) \in X^n \times Y^n : \begin{array}{c} \left| -\frac{1}{n}\log p(\underline{X})-H(X) \right| < \varepsilon \\ \left| -\frac{1}{n}\log p(\underline{Y})-H(Y) \right| < \varepsilon \\ \left| -\frac{1}{n}\log p(\underline{X},\underline{Y})-H(X,Y) \right| < \varepsilon \end{array} \right\}$$

<u>דוגמה</u>: פילוג משותף $p(X,Y)$

נביט בערוץ הבא:



דוגמה לסדרות אופייניות:

$$\begin{array}{lccccccccccccc} n= & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & \ldots \\ \underline{X}= & A & B & A & A & B & A & B & B & B & A & B & \ldots \\ \underline{Y}= & C & D & C & C & C & C & D & C & D & C & C & \ldots \end{array}$$

הפילוג המשותף הוא ההסתברות לאירוע $p(X,Y)$, כאשר - $X \in \{A,B\}, Y \in \{C,D\}$

משפט (J-AEP) :

א.

$$p\left(\left(\underline{X},\underline{Y}\right)\in A_\varepsilon^{(n)}\left(X,Y\right)\right)\xrightarrow[n\to\infty]{}1,\forall\varepsilon>0$$

ב.

$$\left(\underline{X},\underline{Y}\right)\in A_\varepsilon^{(n)}\left(X,Y\right)\Rightarrow p\left(\underline{X},\underline{Y}\right)\doteq 2^{-nH(X,Y)}$$

ג.

$$\left|A_\varepsilon^{(n)}\left(X,Y\right)\right|\doteq 2^{nH(X,Y)}$$

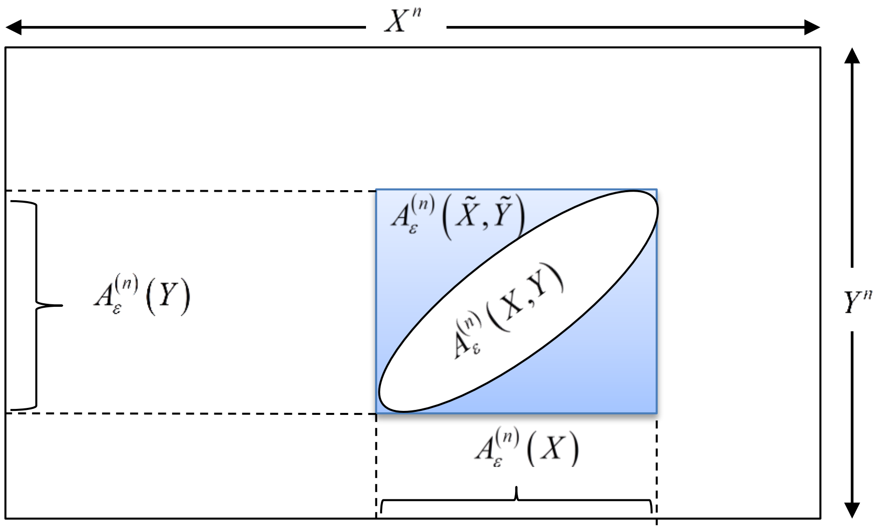
ד. נגדיר $\left(\tilde{\underline{X}},\tilde{\underline{Y}}\right)\sim iid\ p\left(\underline{X}\right)\cdot p\left(\underline{Y}\right)$ כלומר,הפילוג <u>השולי</u> הוא כמו של $\left(\underline{X},\underline{Y}\right)$ אך הפילוג <u>המשותף</u> הוא חסר תלות. נקרא למאורע $\left\{\left(\tilde{\underline{X}},\tilde{\underline{Y}}\right)\in A_\varepsilon^{(n)}\left(X,Y\right)\right\}$ "התחזות", משום שזהו המקרה שבו זוג הוקטורים $\tilde{\underline{X}}$ ו- $\tilde{\underline{Y}}$ שהם בת"ס, ניראים כאילו הם אופייניים ביחס לפילוג <u>התלוי</u> $p\left(X,Y\right)$. עבור אירוע התחזות מתקיים :

$$p\left\{\left(\tilde{\underline{X}},\tilde{\underline{Y}}\right)\in A_\varepsilon^{(n)}\left(X,Y\right)\right\}\doteq 2^{-nI(X;Y)}$$

או ליתר דיוק :

$$\left(1-\varepsilon\right)2^{-n\left[I(X;Y)-3\varepsilon\right]}\le p\left\{\left(\tilde{\underline{X}},\tilde{\underline{Y}}\right)\in A_\varepsilon^{(n)}\left(X,Y\right)\right\}\le 2^{-n\left[I(X;Y)+3\varepsilon\right]}$$

נמחיש בצורה גראפית :



מתקיים :

- $\left|A_\varepsilon^{(n)}\left(X\right)\right|\doteq 2^{nH(X)}$, $\left|A_\varepsilon^{(n)}\left(Y\right)\right|\doteq 2^{nH(Y)}$, $\left|A_\varepsilon^{(n)}\left(X,Y\right)\right|\doteq 2^{nH(X,Y)}$.
- בנוסף, $\left|A_\varepsilon^{(n)}\left(\tilde{X},\tilde{Y}\right)\right|=2^{n(H(X)+H(Y))}$ וזאת כי עבור הפילוג $p\left(\tilde{\underline{X}},\tilde{\underline{Y}}\right)=p\left(\underline{X}\right)\cdot p\left(\underline{Y}\right)$ מתקיים :

$$H\left(\tilde{X},\tilde{Y}\right)=H\left(\tilde{X}\right)+H\left(\tilde{Y}\right)=H\left(X\right)+H\left(Y\right)$$

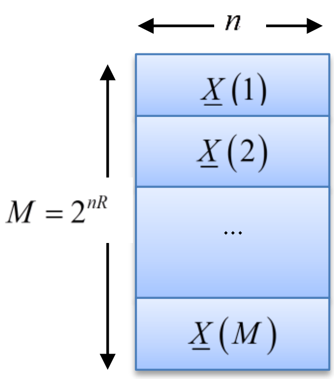
הוכחת סעיף ד׳:

עד כדי ε באקספוננט, המאורע הוא של זוג $\left(\underline{X},\underline{Y}\right)$ בקבוצה הקטנה $A_\varepsilon^{(n)}\left(X,Y\right)$ ביחס לפילוג אחיד (בקירוב) על-פני הקבוצה הגדולה $A_\varepsilon^{(n)}\left(\tilde{X},\tilde{Y}\right)$:

$$p\begin{pmatrix}\text{event of}\\ \text{pretending}\end{pmatrix}=\frac{\left|A_\varepsilon^{(n)}\left(X,Y\right)\right|}{\left|A_\varepsilon^{(n)}\left(\tilde{X},\tilde{Y}\right)\right|}\doteq\frac{2^{nH(X,Y)}}{2^{n\left[H(X)+H(Y)\right]}}=2^{-nI(X;Y)}$$

כאשר השוויון האחרון נובע מהזהות - $I\left(X;Y\right)=H\left(X\right)+H\left(Y\right)-H\left(X;Y\right)$.

7.6. הוכחת המשפט הישר



ההוכחה של שאנון מסוססת על שני עקרונות:

א. קוד אקראי.

ב. פענוח "אופייניות משותפת" היחס לפילוג הגרלת הקוד.

נגריל מטריצה $n\times 2^{nR}$ של סימבולים מהא"ב של X ע"פ פילוג $p\left(X\right)$ בת"ס בין רכיבי המטריצה. נעביר את הטבלה לידיעת המקודד $f\left(\cdot\right)$ והמפענח $g\left(\cdot\right)$. המפענח יודע את פילוג המעבר בערוץ $p\left(Y\mid X\right)$, ונניח כי פילוג ההודעות W אינו ידוע מראש.

<u>פעולת המקודד</u>:

להעברת ההודעה $1\le i\le 2^{nR}$, $W=i$, המקודד משדר את המילה $\underline{X}\left(i\right)$ (כלומר, את שורה i בטבלת הקוד).

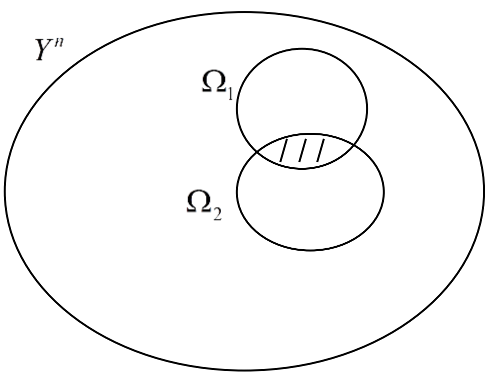
<u>פעולת המפענח</u>:

איזורי ההחלטה $\Omega_1,\dots,\Omega_n$ במפענח נקבעים ע"פ העיקרון של "האופייניות המשותפת":

$$\hat{W}=\begin{cases} i & \text{if } \left(\underline{X}\left(i\right),\underline{Y}\right)\in A_\varepsilon^{(n)}\left(X,Y\right) \quad \text{for some unique i}\\ ? & \text{otherwise}\Rightarrow \begin{array}{c}\text{all the codewords } \underline{X}\left(i\right) \text{ are not typical with } \underline{Y}\\ \text{or there is more then one codeword that is typical with } \underline{Y}\end{array}\end{cases}$$

כאשר $A_\varepsilon^{(n)}\left(X,Y\right)$ מחושב לפי הפילוג המשותף - $p\left(X,Y\right)=p\left(X\right)\cdot p\left(Y\mid X\right)$.

נסתכל על פעולת המפענח בצורה גראפית:



Ω_1 הוא קבוצת הוקטורים - $\left\{\underline{Y}:\left(\underline{X}(1),\underline{Y}\right)\in A_{\varepsilon}^{(n)}(X,Y)\right\}$. האיזור בו קיימת חפיפה בין Ω_1 ל-
Ω_2 קיימת אי-ודאות הוא איזור בו לא ניתן להכריע כי אכן שודר $\underline{X}(1)$. לכן, המפענח יחליט באיזור
זה: $\hat{W}=?$. אם החלטות המפענח הן $\hat{W}=?$ או $\hat{W}\neq W$, הדבר מוגדר "כמאורע שגיאה".

<u>ניתוח הסתברות השגיאה:</u>

נגדיר את המאורע E_i כמקרה בו $\underline{X}(i)$ ו- $\underline{Y}$ אופייניים במשותף ביחס לפילוג $p(X,Y)$:

$$E_i \triangleq \left(\underline{X}(i),Y\right)\in A_{\varepsilon}^{(n)}(X,Y),\ i=1,\dots,M$$

אם $W=1$ (כלומר מילת הקוד ששודרה היא $\underline{X}(1)$) אזי:

$$\text{Error event} = \underbrace{E_1^C}_{\text{complement}} \cup E_2 \cup E_3 \dots \cup E_M$$

כאשר $E_2,\dots,E_M$ הם מאורעות של התחזות.

נחשב את הסתברות השגיאה הממוצעת הן ביחס להתנהגות הערוץ, ביחס להגרלת ספר הקוד $\mathbb{C}$ והן ביחס לאיזו מילת קוד $\underline{X}(i)$ שודרה:

$$\bar{\bar{P}}_e=\sum_{\mathbb{C}} p(\mathbb{C})\cdot\bar{p}_e(\mathbb{C})=\sum_{\mathbb{C}} p(\mathbb{C})\frac{1}{M}\sum_{i=1}^{M}\underbrace{p_{e,i}(\mathbb{C})}_{\substack{\text{error probability of code }\mathbb{C}\\ \text{if the codeword i was transmitted}}}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\underbrace{\sum_{\mathbb{C}} p(\mathbb{C})\cdot p_{e,i}(\mathbb{C})}_{\substack{\text{independent of i because}\\ \text{the code is chosen randomly}}}=\sum_{\mathbb{C}} p(\mathbb{C})\cdot p_{e,1}(\mathbb{C})\triangleq\bar{p}_{e,1}$$

<u>הערה</u>: הסבר על על הסימון - $\bar{\bar{p}}_e$:

- ✓ p_e - פרמטר המלמד על אקראיות הערוץ.
- ✓ $\bar{p}_e$ - כנ"ל, ובנוסף מתבצע מיצוע על מילות הקוד $1,\dots,M$.
- ✓ $\bar{\bar{p}}_e$ - כנ"ל, ובנוסף מתבצע מיצוע על כל ספרי הקוד האפשריים $\mathbb{C}$.

נחזור לחשבון. עתה נרשום את הביטוי עבור השגיאה כאשר שודר $\underline{X}(1)$ (כפי שהראינו – זה שקול להסתברות השגיאה הכוללת):

$$\bar{\bar{p}}_e=p\left(E_1^C\cup\bigcup_{i=2}^{M}E_i\right)\underset{\text{union bound}}{\leq}p\left(E_1^c\right)+\sum_{i=1}^{M}\underbrace{p\left(E_i\right)}_{\substack{\text{codeword i is typical with }\underline{Y}\text{ despite}\\ \text{that codeword 1 was transmitted}}}$$

$$=p\left(E_1^c\right)+(M-1)\,p\left(E_2\right)$$

מה- Joint-AEP נובע:

$$p\left(E_1^C\right)=p\left\{\left(\underline{X},\underline{Y}\right)\notin A_\varepsilon^{(n)}\left(X,Y\right)\Big|\begin{matrix}\underline{X}=\underline{X}(1)\\ \underline{Y}\text{ is generated from }\underline{X}\text{ by channel }p\left(Y\mid X\right)\end{matrix}\right\}$$

$$\le\varepsilon\ \text{ if n is large enough}$$

$$\mathrm{p}\left(E_2\right)=p\left\{\left(\underline{X}(2),\underline{Y}\right)\notin A_\varepsilon^{(n)}\left(X,Y\right)\Big|\begin{matrix}\underline{X}=\underline{X}(1)\\ \underline{Y}\text{ is generated from }\underline{X}\text{ by channel }p\left(Y\mid X\right)\end{matrix}\right\}$$

$$=\text{probability of pretending}\le 2^{-n\left[I(X;Y)-3\varepsilon\right]}$$

$$\Rightarrow\bar{\bar{p}}_e\le\varepsilon+\underbrace{\left(2^{nR}-1\right)}_{\substack{\text{amount of}\\ \text{pretending codewords}}}\cdot 2^{-n\left[I(X;Y)-3\varepsilon\right]}<\varepsilon+\underbrace{2^{-n\left[I(X;Y)-R-3\varepsilon\right]}}_{\substack{\le\varepsilon\\ n\text{ is large enough}\\ \text{if }R<I-3\varepsilon}}\le 2\varepsilon$$

מכאן, שבממוצע על כל הקודים האפשריים הסתברות השגיאה שואפת ל-0 כאשר n שואף לאינסוף, בתנאי ש- $R<I\left(X;Y\right)$.

<u>נקודות חשובות לציון</u>:

א. חשוב לציין שוב כי בפיתוח הנ"ל דובר על הסתברות שגיאה להודעה או לחלופין לבלוק.

ב. כאשר נבחר בפילוג ייצור הקוד האקראי $p\left(X\right)=p^*\left(X\right)$ נקבל כי $C=I\left(X;Y\right)$.

ג. אם נבחר $\varepsilon\to 0$, נוכל להשיג $\bar{\bar{p}}_e\xrightarrow[n\to\infty]{}0$ לכל R קרוב כרצונינו ל- C.

ד. <u>מציאת קוד מסויים</u>: אם $\bar{\bar{p}}_e=E\left[\bar{p}_e\right]\le 2\varepsilon$ (תוחלת על פני הקודים האפשריים $\mathbb{C}$), אזי בהכרח קיים לפחות קוד אחד בקבוצה שעבורו $\bar{p}_e<2\varepsilon$.

ה. <u>הסתברות שגיאה מקסימלית</u>: נסדר את M מילות הקוד לפי הסתברויות השגיאה שלהן באופן הבא: $p_{e_1}\le p_{e_2}\le\ldots\le p_{e_M}$. נגדיר ספר קוד על המילים בעלות הסתברות השגיאה הנמוכות ביותר:

$$\mathbb{C}'\triangleq\left\{\underline{X}(1),\ldots,\underline{X}\left(\frac{M}{2}\right)\right\}$$

$$2\varepsilon>\bar{p}_e=\frac{1}{M}\sum_{i=1}^{M}p_{e_i}\ge\frac{1}{M}\sum_{i=\frac{M}{2}+1}^{M}p_{e_i}\ge\frac{1}{M}\cdot\frac{M}{2}\cdot p_{e_{\frac{M}{2}}}=\frac{p_{e_{\frac{M}{2}}}}{2}$$

ולכן, כל מילות הקוד ב- $\mathbb{C}'$ בעלות הסתברות שגיאה נמוכה מ- 4ε. כלומר - $p_{e_{\max}}<4\varepsilon$ עבור קוד $\mathbb{C}'$. בבניה של $\mathbb{C}'$ התשמשנו בחצי ממילות הקוד אשר היו קיימות בקוד המקורי $\mathbb{C}$. נבדוק כמה הפסדנו בקצב הקידוד:

$$R=\frac{1}{n}\log\left\{\begin{matrix}\text{amount of}\\ \text{codewords in }\mathbb{C}'\end{matrix}\right\}=\frac{1}{n}\log\frac{M}{2}=\frac{1}{n}\log M-\frac{1}{n}\log 2=R-\frac{1}{n}\xrightarrow[n\to\infty]{}R$$

כלומר, עבור n גדול מספיק ההפסד הוא זניח.

ו. <u>"רוב הקודים הם טובים"</u>: מ"א אי-שלילי מקיים ע"פ אי-שוויון מרקוב כי לכל קבוע ממשי $t>0$ מתקיים:

$$P\left(Z\ge t\cdot EZ\right)\le\frac{1}{t}$$

מכאן נובע:

$$p\left\{\begin{array}{c}\text{random code}\\ \text{has error probability}\geq 100\cdot E\left[\bar{p}_e\right]\\ \bar{p}_e\left(\mathbb{C}\right)\end{array}\right\}\leq\frac{1}{100}$$

כלומר, ההסתברות לקוד אקראי עם הסתברות שגיאה 200ε קטנה מ-0.01. ובאופן כללי, ככול שאורך הבלוק גדל, חלק גדול יותר של הקודים באנסמבל הם טובים כרצונינו.

ז. בפיתוח של החסם (חסם האיחוד) דרשנו "רק" אי-תלות בזוגות.

7.7. משפט הפוך עבור הסתברות השגיאה לביט – BER

כאמור, במשפט של שאנון דיברנו על הסתברות השגיאה להודעה. לכל הודעה ניתן להתייחס ע"פ הביטים המרכיבים אותה. כאשר $W\in\left\{1,\ldots,2^{nR}\right\}$ היא ההודעה אותה משדרים, ניתן על הביטים היוצרים אותה - $W=\left\{B_1,\ldots,B_L\right\}$ כאשר $L=nR$. באופן שקול, המילה המפוענחת היא - $\hat{W}=\left\{\hat{B}_1,\ldots,\hat{B}_L\right\}$. נגדיר הסתברות שגיאה לביט אינפורמציה:

$$BER\triangleq\frac{1}{L}\sum_{i=1}^{L}p\left(\hat{B}_i\neq B_i\right)$$

<u>טענה</u>:

$$H_B\left(BER\right)\geq 1-\frac{C}{R}$$

כלומר, כאשר $R>C$, אזי בהכרח $BER>0$.

<u>הוכחה</u>:

מהוכחת המשפט ההפוך להודעה (פרק 7.4):

$$H\left(W|Y^n\right)=\left\{\begin{array}{c}\text{the remaining}\\ \text{uncertainty}\end{array}\right\}\geq n\left(R-C\right)$$

ומכאן נובע:

$$H\left(\underbrace{B_1,\ldots,B_L}_{W}|Y^n\right)\geq n\left(R-C\right)$$

נגדיר:

$$E_i\triangleq\begin{cases}1 & \hat{B}_i\neq B_i\\ 0 & \hat{B}_i=B_i\end{cases}$$

$$H\left(B_1,\ldots,B_L|Y^n\right)=H\left(E_1,\ldots,E_L|Y^n\right)\leq H\left(E_1,\ldots,E_L\right)\leq\sum_{i=1}^{L}H\left(E_i\right)$$

$$=\sum_{i=1}^{L}H_B\underbrace{\left(\tilde{p}_{e_i}\right)}_{\substack{\text{error probability}\\ \text{in the } i\text{-th bit}}}=L\cdot\frac{1}{L}\sum_{i=1}^{L}H_B\left(\tilde{p}_{e_i}\right)\underset{\substack{\text{convexity of}\\ H_B(\varepsilon)}}{\leq}L\cdot H\underbrace{\left(\frac{1}{L}\sum_{i=1}^{L}\tilde{p}_{e_i}\right)}_{=BER}$$

$$\Rightarrow L\cdot H_B\left(BER\right)\geq n\left(R-C\right)$$

$$\Rightarrow H_B\left(BER\right)\geq 1-\frac{C}{R}$$

7.8. תנאי KKT לחישוב קיבול בערוץ DMC

<u>הערה</u>: פרק זה מבוסס על מסמך שנכתב ע"י ניר וינברגר. הוא בא לסכם את עיקרי הדברים. פרק זה משלים את הדוגמאות שניתנו בפרק 7.1.

כזכור, עבור ערוץ DMC שכניסתו היא X ומוצאו הוא Y ניתן לבטא את האינפורמציה ההדדית בצורה הבאה:

$$I(X;Y)=\sum_{x\in X}p(x)\cdot I(X=x;Y)=\sum_{x\in X}p(x)\cdot D\left(p(y|x)\|p(y)\right)=$$

$$=\sum_{x\in X}p(x)\sum_{y\in Y}p(y|x)\log\frac{p(y|x)}{p(y)}=\sum_{x\in X}p(x)\sum_{y\in Y}p(y|x)\log\frac{p(y|x)}{\sum_{x'\in X}p(y|x')\cdot p(x')}$$

ולפי שאנון:

$$C=I(X;Y)=\sum_{x\in X}p(x)I(X=x;Y)$$

<u>משפט</u>: (תנאי KKT)

$p^*(X)$ הינו הפילוג המגשים אמ"מ לכל x עבורו $p^*(x)>0$ מתקיים כי - $I(X=x:Y)=C$ ולכל x עבורו $p^*(x)=0$ מתקיים - $I(X=x:Y)\le C$.

<u>אלגוריתם לחישוב קיבול ערוץ</u>:

- ✓ במקרים פשוטים אפשר לחסום את $I(X,Y)$ מלמעלה (משיקולים של מקסימום אנטרופיה) ואח"כ להראות שעבור פילוג כניסה "הגיוני" (למשל פילוג אחיד) החסם הוא בר-השגה, וזהו בהכרח הפילוג המגשים.
- ✓ במקרים מורכבים יותר, ניתן "לנחש" פילוג כניסה מגשים, ולבדוק באמצעות תנאי ה-KKT כי $I(X=x:Y)$ מקבל את אותו הערך לכל x עבורו $p(x)>0$ וערך הנמוך מכך עבור x-ים שעבורם $p(x)=0$. במקרה זה - $C=I(X=x;Y)$ אשר מחושב עבור x כלשהו שמקיים $p(x)>0$ ו-$p(x)$ הוא בהכרח $p^*(x)$.

7.9. קיבול אפס שגיאה ואקספוננט שגיאה

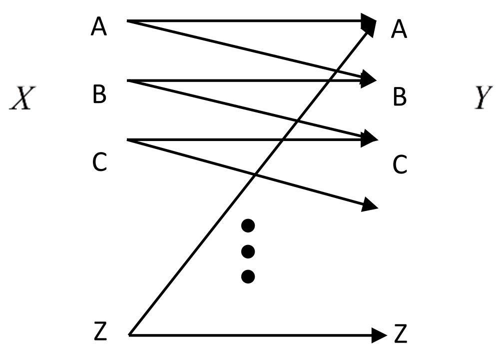
<u>הגדרה</u>:

<u>*קיבול אפס שגיאה*</u> הינו הקצב המקסימלי הניתן להשגה עם $p_e\equiv 0$ (להבדיל מההגדרה עבור הקיבול הרגיל בה $p_e\to 0$). מסמנים את קיבול אפס השגיאה ב-C_0.

לעולם מתקיים כי - $C_0\le C$.

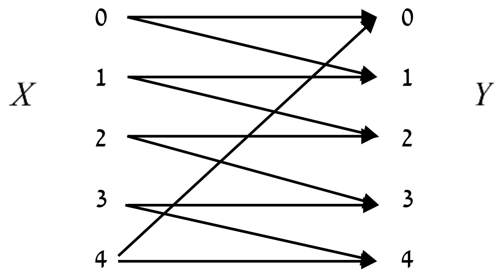
<u>דוגמא</u>: Noisy Typewriter

נתון הערוץ הבא: (בכל הקשתות הסתברות המעבר היא 0.5)



בערוץ זה - $L=26$, ובנוסף מתקיים - $C=\log 26-1=\log 13$.
אם נבחר כל כניסה שניה ונמפה אליה מילת כניסה, נוכל לפענח בערוץ זה ללא שגיאה. כלומר - $C_0=\log 13$, כלומר - $C_0=C$.

לעומת זאת, נתבונן בערוץ פנטגון : (גם כאן, בכל הקשתות הסתברות המעבר היא 0.5)



בערוץ זה - $C=\log 5-1$. ניתן לראות כי עבור המיפוי הבא של הכניסה בזוגות ניתן להעביר מידע בערוץ ללא שגיאה עבור מילות הקוד הבאות - $\{11,23,35,54,42\}$. ולכן :

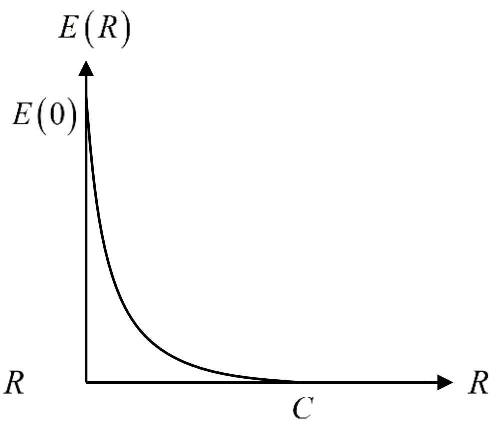
$$C_0 \geq \frac{1}{n}\log M=\frac{1}{2}\log 5<C$$

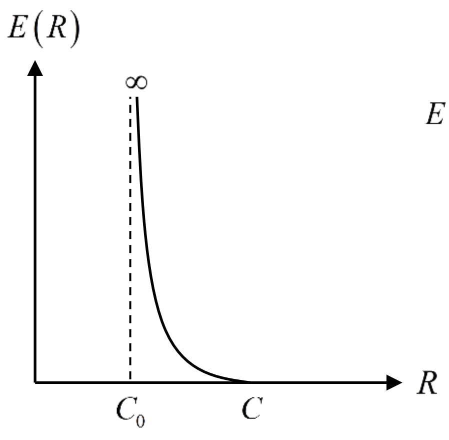
<u>הערות חשובות</u> :

א. אם $R\leq C_0$, אזי ניתן להשיג $p_e\equiv 0$ עבור איזשהו אורך בלוק סופי n .

ב. אם $C_0<R<C$, מתקיים כי הסתברות השגיאה הנמוכה ביותר הניתנת להשגה מקיימת - $p_e^{opt}\sim 2^{-nE(R)}$, כאשר $E(R)$ נקרא "<u>אקספוננט השגיאה</u>". זהו גודל חיובי. עבור $R>C$, לפי ה- AEP, $E(R)=0$. להלן גרפים של $E(R)$ עבור 2 מקרים :

ערוץ ללא קיבול אפס שגיאה	ערוץ עם קיבול אפס שגיאה





8. קידוד משותף מקור-ערוץ (J.S.C.C)

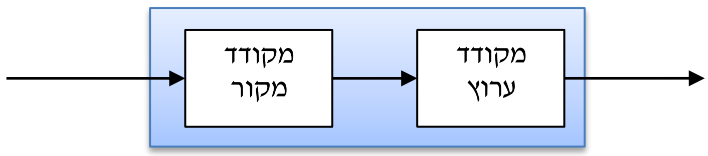
בסכימה הנ״ל מקודד המקור ומקודד הערוץ אוחדו לבלוק בודד – כך גם מפענח הערוץ ומפענח המקור.

8.1 המשפט הישר לקידוד משותף מקור-ערוץ

אם $\bar{H}(u) < C\cdot r$, כאשר $r \triangleq n/L$ שימושי ערוץ לדגימת מקור - אזי ניתן לשחזר את u במפענח עם הסתברות שגיאה קטנה כרצוננו.

<u>הוכחה</u>:

ע״י עיקרון ההפרדה של מקודד מקור ומקודד ערוץ, נפריד את המקודד המשותף ל-2 מקודדים:



המקודד ממיר L דגימות מקור ל- m ביטים. לכן, ע״פ משפט קידוד מקור מתקיים - $L\cdot\bar{H}(u) \le m$.

מצד שני, ע״פ משפט קידוד ערוץ $m \le nC$:

$$L\cdot\bar{H}(u) \le m \le n\cdot C$$

$$\Rightarrow \bar{H}(u) < \frac{n}{L}\cdot C = r\cdot C$$

בעצם, הסכימה הנ״ל מקיימת הן משפט קידוד המקור והן את משפט קיבול הערוץ.

8.2 המשפט ההפוך לקידוד משותף מקור-ערוץ

אם $\bar{H}(\underline{u}) > r\cdot C$, אזי לא ניתן לקיים תקשורת אמינה. בפרט מתקיים:

$$H_B(SER) + SER\cdot\log(|u|-1) \ge \bar{H}(u) - r\cdot C$$

כאשר:

$$SER \triangleq \frac{1}{L}\sum_{i=1}^{L} p(\hat{u}_i \ne u_i)$$

<u>בדיקה</u>:

עבור מקור שהוא ״ביטי אינפורמציה״ (מקור ברנולי 0.5) מתקיים - $|u| = 2$, $\bar{H}(u) = 1$ ולכן:

$$r = \frac{n}{L} = \frac{n}{nR} = \frac{1}{R}$$

9. ערוצים מיוחדים

9.1. ערוצים עם זיכרון

עד עכשיו דיברנו על ערוצי DMC בהם מתקיים:

$$p\left(\underline{Y}\mid\underline{X}\right)=\prod_{i=1}^{n}p\left(Y_i\mid X_i\right)$$

עבור ערוץ כללי, אין הנוסחא הנ"ל בהכרח נכונה. על כן נגדיר – באופן אנאלוגי לפרק קידוד מקור עם זיכרון – את הקיבול לערוץ עם זיכרון:

$$C^{(n)^{\inf}}\triangleq\frac{1}{n}\max_{p\left(X^n\right)}I\left(X^n;Y^n\right)$$

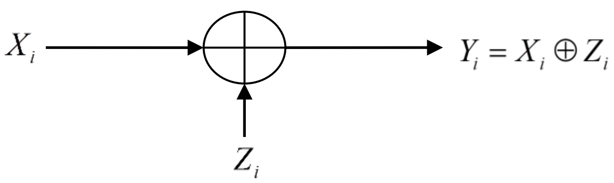
נסמן:

$$C^{(n)}\xrightarrow[n\to\infty]{}C^{(\infty)}$$

במקרה הכללי לא בהכרח מתקיים כי - $C^{op}=C^{(\infty)}$.

<u>מקרה פרטי</u>:

ערוץ רעש אדיטיבי מודולו L, כאשר הרעש איננו iid.



בערוץ זה מתקיים:

$$C^{(\infty)}=\log L-\bar{H}\left(Z\right)\underset{\substack{\text{if the noise}\\ \text{is stationary}\\ \text{and ergodic}}}{=}C^{op}$$

כאשר $\bar{H}\left(z\right)$ הוא קצב האנטרופיה של הרעש Z.

<u>**המשפט ההפוך החזק**</u>:

אם $R>C^{(\infty)}$ אזי בהכרח - $p_{e,block}^{\max}\xrightarrow[n\to\infty]{}1$.

משפט זה מוכח בעזרת משפט ה- AEP ההפוך (ראה פרק 3.3).

<u>הסבר</u>:

כזכור, איחוד איזורי ההחלטה של המפענח Ω_i כאשר $i=1,\ldots,M$ שווה לתחום וקטורי מוצא הערוץ Y^n.

מכאן שגודל איזור ההחלטה הקטן ביותר מקיים:

$$\left|\Omega_{\min}\right|\leq\frac{\left|Y^n\right|}{M}=\frac{L^n}{2^{nR}}=2^{n(\log L-R)}$$

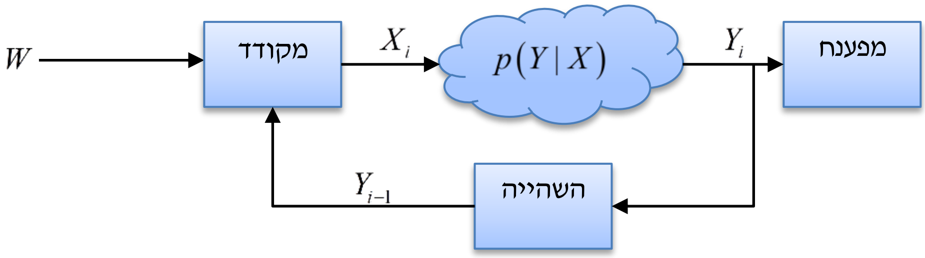
מכאן שאם $R>C^{(\infty)}$ אזי בהכרח:

$$\left|\Omega_{\min}\right|<2^{n\left(\bar{H}(Z)-\delta\right)}$$

עבור איזשהו $\delta>0$. הסתברות השגיאה המקסימלית היא הסיכוי ש- $\underline{Z}$ לא נופל בתוך תחום בגודל $\left|\Omega_{\min}\right|$. ע"פ ה-AEP ההפוך, $\delta>0$ גורר כי הסתברות השגיאה הולכת ל-1.

9.2. ערוצים עם משוב

משוב הינה שיטה בה המפענח משיב למקודד מידע על מה שהוא קלט, ומתקבלת החלטה ע״פ הפרוטוקול האם לשדר שוב במקרה של טעות. משוב יכול לשפר את הסתברות השגיאה ואת אקספוננט השגיאה (כולל הגדלה של C_0), אך <u>אינו</u> יכול להגדיל את הקיבול C של ערוץ <u>חסר זיכרון</u>.

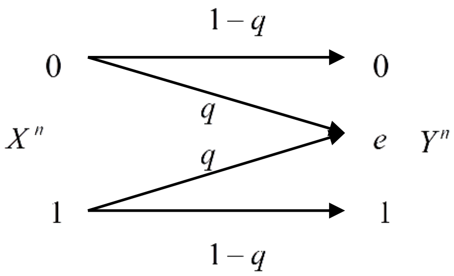


עבור מקודד ללא משוב: $X_i = f_i(W) \Leftrightarrow X^n = f(W)$ $i=1,\dots,n$.

עבור מקודד עם משוב: $X_i = f(W, Y_1, \dots, Y_{i-1})$.

<u>דוגמא</u>: משוב מוריד הסתברות שגיאה.

נבחן ערוץ עם מחיקות:



כזכור, בערוץ הנ״ל - $C = 1-q \frac{bit}{c.u.}$. על-מנת לשפר את הביצועים, נשדר k ביטי אינפורמציה, ובכל שלב נשדר מחדש את הביטים שנמחקו. כלומר סדר השידור יהיה כדלהלן:

$$X^n = [K \text{ information bits}], [q_1 K \text{ erased bits - 1st iteration}],$$
$$[q_2 q_1 K \text{ erased bits - 2nd iteration}], \dots$$

כאשר q_i הוא מס׳ הביטים שנמחקו בסיבוב ה- i.

ע״פ חוק המספרים הגדולים - $q_i \xrightarrow[k\to\infty]{} q$.

לכן, מס׳ הפעמים שיש לשדר הוא:

$$n \approx K \cdot \sum_{i=0}^{\infty} q^i = \frac{K}{1-q} \Rightarrow R = \frac{K}{K/(1-q)} = 1-q$$

אנו משיגים את הקיבול עם הסתברות שגיאה 0, וקיבול אפס שגיאה מתמזג עם קיבול הערוץ.

נשים לב כי בערוץ חסר-זיכרון עם משוב לא מתקיים - $p(\underline{Y}|\underline{X}) = \prod_{i=1}^{n} p(Y_i|X_i)$ אלא - $p(Y_i|X_1,\dots,X_i) = p(Y_i|X_i)$. כלומר, ניתן לבטא את הערוץ כשרשרת מרקוב:

$$Y_i \leftrightarrow X_i \leftrightarrow (X_1^{i-1}, Y_1^{i-1}, W)$$

בערוץ עם משוב, X_i תלוי ישירות ב- Y_1^{i-1}, ולכן Y_i תלוי ב- X_{i+k} (כאשר $k \geq 0$) ישירות ולא רק דרך X_i.

כזכור, עבור ערוצי DMC ללא משוב התקיים:

$$H\left(Y^n \mid X^n\right)=\sum_{i=1}^{n} H\left(Y_i \mid X_i\right)$$

ומכאן הוכחנו כי:

$$I\left(X^n;Y^n\right)\le\sum_{i=1}^{n} I\left(X_i;Y_i\right)$$

עבור ערוץ DMC עם משוב המשוואות הנ"ל אינן מתקיימות, ולכן נציג את אי- השוויון בדרך חלופית:

$$I\left(W;Y^n\right)\underset{(*)}{\le}\sum_{i=1}^{n} I\left(X_i;Y_i\right)\le nC$$

<u>הוכחה של (*)</u>:

$$I\left(W;Y^n\right)\underset{\substack{\text{chain rule for}\\ \text{mutual entropy}}}{=}\sum_{i=1}^{n} I\left(W;Y_i \mid Y_1^{i-1}\right)$$

$$\begin{aligned}I\left(W;Y_i \mid Y_1^{i-1}\right)&\le I\left(W,Y_1^{i-1};Y_i\right)\le I\left(W,Y_1^{i-1},X_i;Y_i\right)=\\&=I\left(X_i;Y_i\right)+\underbrace{I\left(W,Y_1^{i-1};Y_1 \mid X_i\right)}_{\substack{=0\\ \text{due to Markovity}}}\end{aligned}$$

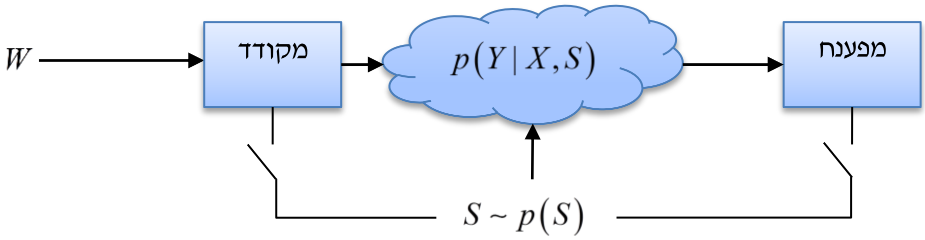
עבור W אחיד - $H\left(W\right)=nR$. ולכן, קיבלנו כי אי-הודאות הנותרת מקיימת:

$$H\left(W \mid Y^n\right)\ge n\left(R-C\right)$$

ויחד עם אי-שוויון פאנו הוכחנו שגם עם משוב לא ניתן לשדר מקצב הגדול מקיבול הערוץ.

<u>הערה חשובה</u>:
הטענה הנ"ל לא תופסת בהכרח לערוצים עם <u>זיכרון</u>. כלומר, בערוץ עם זיכרון המשוב כן יכול להגדיל את הקיבול.

9.3. ערוצים עם אינפורמציית צד



בסכימה הנ״ל, הערוץ מ- X ל- Y תלוי בין היתר גם במשתנה מצב S. משתנה זה הינו משתנה אקראי היכול לבטא פרמטר משתנה בערוץ כדוגמת דעיכות. נבחן את הקיבול במצבים שונים כאשר המקודד או המפענח יודעים את ערכו של S:

<u>מצב הערוץ ידוע למפענח ולא למקודד</u>:

$$C=\max_{p(X)} I(X;Y,S)=\max_{p(X)}\left\{\underbrace{I(X;S)}_{=0 \text{ because } X\perp S}+I(X;Y|S)\right\}=\max_{p(X)} I(X;Y|S)$$

<u>מצב הערוץ ידוע גם למקודד וגם למפענח</u>:

$$C=\max_{p(X|S)} I(X;Y|S)=\sum_{s\in S} p(S=s)\cdot I(X;Y|S=s)$$

<u>מצב הערוץ לא ידוע במקודד וגם במפענח</u>:

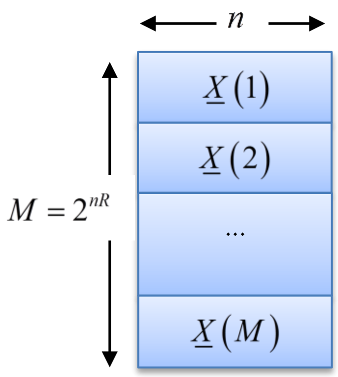
$$C=\max_{p(X)} I(X;Y)$$

כאשר פילוג המעבר השקול מתקבל ע״י:

$$p(Y|X)=\sum_{s\in S} p(s)\cdot p(Y|X,S=s)$$

10. קודים אקראיים לעומת קודים ליניאריים

כזכור, עבור קוד אקראי הגרלנו ספר קוד $\mathbb{C}$:



כל $k=nR$ ביטים (כאשר $M=2^{nR}$) בוחרים את המילה באורך n שתשודר בערוץ. שאנון הציע גילוי בשיטת האופייניות המשותפת, כלומר לבדוק עבור וקטור שנקלט $\underline{Y}$ האם קיימת מילת קוד i יחידה עבורה מתקיים - $\left(\underline{X}(i),\underline{Y}\right)\in A_{\varepsilon}^{(n)}(X,Y)$, אם כן אז המפענח מחליט $\hat{X}=i$.

נבחן את ערוץ $Bern(0.5)$. כזכור, עבור ערוץ זה הפילוג המגשים הוא אחיד. בערוץ זה, על גודל המילון לקיים:

$$M=2^{nR}\approx 2^{nC}=2^{n(1-H_B(p))}$$

הערה:

גילוי אופייניות משותפת בערוץ BSC שקולה ל- $\left(\underline{Y}\oplus\underline{X}(i)\right)\in A_{\varepsilon}^{(n)}(Z)$ כאשר Z הינו רעש הברנולי של הערוץ. נשים לב כי $A_{\varepsilon}^{(n)}$ הוא כדור המינג הכולל משקל המינג $\approx np$.

עתה נבחן קוד בלוק ליניארי. קוד שכזה מאופיין בד"כ ע"י 3 פרמטרים - (n,k,d):

- ✓ n - אורך סדרת המידע לאחר קידוד.
- ✓ k - אורך סדרת המידע לפני הקידוד.
- ✓ d - המרחק המינימלי של הקוד.

פעולת קידוד המידע מבוצעת באופן הבא - כל k סימבולי מידע מבוטאים כוקטור $\underline{x}$ ומוכפלים במטריצה. בצורה זו מתקבל וקטור באורך n:

$$\underset{n\times 1}{\underline{X}}=\underset{n\times 1}{\underline{C}}=\underset{n\times k}{\underline{\underline{G}}}\cdot\underset{k\times 1}{\underline{x}}$$

כאשר המטריצה $\underline{\underline{G}}$ נקראת "המטריצה היוצרת" של הקוד.

הערה: כיוון שאנו עובדים על מרחב בינארי, כל פעולות הכפל והחיבור הן מודולו 2.

התיאור הנ"ל של הקוד $\mathbb{C}$ גורר כי כל מילות הקוד $\underline{C}_i$ יוצרות מרחב ליניארי, כלומר:

$$\underline{C}_i,\underline{C}_j\in\mathbb{C}\ \Rightarrow\ \underline{C}_i\oplus\underline{C}_j\in\mathbb{C}$$

יותר מכך, k וקטורי העמודה של $\underline{\underline{G}}$ פורשים את המרחב הוקטורי שמהווה את ספר הקוד $\mathbb{C}$.

מקרה פרטי: קוד נקרא "קוד סיסטמתי" אם $\underline{\underline{G}}$ שלו היא מורכבת מבלוקים בצורה הבאה:

$$\underset{n\times k}{\underline{\underline{G}}}=\left[\begin{array}{c}\underset{k\times k}{I}\\ \hline \underset{(n-k)\times k}{R}\end{array}\right]$$

כאשר $\underline{\underline{I}}$ הינה מטריצת יחידה בגודל $k \times k$ ו- $\underline{\underline{R}}$ היא מטריצה כלשהי המתארת את תוספת היתירות של הקידוד. באופן שקול ניתן לתאר את מילת הקוד כך - $\underline{C} = \left[\underline{x}, \underline{r}\right]$, כלומר מילת הקוד היא בעצם מילת המידע המקורית באורך k בתוספת $\left(n-k\right)$ ביטי יתירות.

נשלים את המטריצה $\underline{\underline{G}}$ עם מטריצה $\underline{\underline{H}}^T$ לבסיס שפורש את כל המרחב $\left\{0,1\right\}^n$ בצורה הבאה:

$$n \left[\underbrace{G}_{k} \,\middle|\, \underbrace{H^T}_{n-k} \right]$$

למטריצה $\underline{\underline{H}}$ קוראים "מטריצה בדיקת זוגיות". בעזרתה מגדירים את הקוד באופן שקול:

$$\underset{(n-k)\times n}{\underline{\underline{H}}} \cdot \underset{n\times 1}{\underline{C}} = \underset{(n-k)\times 1}{\underline{0}} \;, \quad C \in \mathbb{C}$$

<u>הגדרה:</u>

<u>מרחק מינימלי</u> של קוד d מוגדר בצורה הבאה:

$$d = d_{\min} = \min_{\underline{C}_i \neq \underline{C}_j \in \mathbb{C}} d_H\left(\underline{C}_i, \underline{C}_j\right) = \min_{\underline{C} \in \mathbb{C}} W_H\left(\underline{C}\right)$$

בהגדרה הנ"ל:

- ✓ $d_H\left(\underline{a}, \underline{b}\right)$ - <u>מרחק האמינג</u> בין הוקטורים $\underline{a}$ ו- $\underline{b}$ (שהם באותו האורך). זהו מס' הביטים השונים בין 2 הוקטורים.
- ✓ $W_H\left(\underline{a}\right)$ - <u>משקל האמינג</u>. שהו מס' ה-1-ים בוקטור $\underline{a}$.

אם המרחק המינימלי בין 2 מילות קוד הוא d, הדבר אומר כי אם היו פחות מ- $\left\lfloor \frac{d-1}{2} \right\rfloor$ חילופי ביט במילת קוד, נוכל לתקן אותם ולשייך את המילה הנקלטת למילה שקרובה אליה ביותר ע"פ מרחק האמינג. לחלופין, ניתן להסתכל על המרחב הוקטורי כאוסף של כדורים שהמרחק בין מרכזם הוא d ואין חפיפה ביניהם. כל מרכז מתאר את אחת ממילות הקוד בספר $\mathbb{C}$. המילה שמתקבלת משויכת למילת הקוד שמייצגת את הכדור בו היא נקלטה.

עבור ערוץ BSC – עבור n מספיק גדול, מס' חילופי הביט בערוץ הוא $n \cdot p$, ואם מתקיים $n \cdot p \leq \frac{d}{2}$, אזי המרחק המינימלי של הקוד מקיים - $d = 2 \cdot n \cdot p$.

<u>דוגמא</u>: קוד האמינג $\left(7,4,3\right)$

בקוד זה מטריצת בדיקת הזוגיות היא:

$$\underline{\underline{H}} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

מילות הקוד שמתוארות ע"י המטריצה הנ"ל מקיימות:

$$\underline{C} = \left\{ \underline{C} : \underset{3\times 7}{\underline{\underline{H}}} \cdot \underset{7\times 1}{\underline{C}} = \underset{3\times 1}{\underline{0}} \right\}$$

המכפלה $\underline{\underline{H}} \cdot \underline{C}$ מתארת את סכום העמודות במטריצה $\underline{\underline{H}}$ שבאינדקס שלהן $\underline{C}=1$. ע"פ המבנה של $\underline{\underline{H}}$, כל הוקטורים $\underline{C}$ צריכים להכיל לפחות שלושה 1-ים על-מנת לקיים את ההגדרה שלהם (ניתן ליצור את הוקטור $\underline{0}$ ע"י מינימום 3 עמודות מ- $\underline{\underline{H}}$). כלומר, המרחק המינימלי של הקו הוא - $d=3$ וקוד זה יוכל לתקן חילוף אחד לכל היותר.

<u>הגדרה</u>:

<u>הסינדרום</u> של וקטור $\underline{Y}$ הוא - $\underline{s}=\underline{\underline{H}} \cdot \underline{Y}$.

נביט איך מפענחים שגיאה בעזרת קוד האמינג והסינדרום. נניח ונקלט וקטור $\underline{Y}$ בערוץ עם רעש ברנולי - $\underline{Y}=\underline{X} \oplus \underline{Z}$. נרצה לפענח מ- $\underline{Y}$ את $\underline{X}$ כאשר $\underline{X}$ מקודד לפי מטריצה $\underline{\underline{H}}$. נחשב את הסינדרום של $\underline{Y}$:

$$\underline{s}=\underline{\underline{H}} \cdot \underline{Y}=\underline{\underline{H}} \cdot (\underline{X} \oplus \underline{Z})=\underbrace{\underline{\underline{H}} \cdot \underline{X}}_{=0} \oplus \underline{\underline{H}} \cdot \underline{Z}=\underline{\underline{H}} \cdot \underline{Z}$$

אם מתקיים כי $\underline{s}=0$, הדבר אומר כי נקלטה המילה ששודרה ללא חילופי ביט. אם היה חילוף ביט יחיד, הדבר אומר כי בוקטור $\underline{Z}$ היה קיים רק איבר יחיד ששווה ל-1. בעצם, תחת ההנחה שהייתה רק חילוף אחד באינדקס i, הסינדרום שיתקבל בעצם יהיה העמודה ה- i של $\underline{\underline{H}}$. ולכן, נוכל לשערך כי וקטור הרעש הוא אפסים ויש רק 1 באינדקס i, והשחזור יתבצע באמצעות:

$$\hat{X}=\underline{Y} \oplus \hat{\underline{Z}}$$

כאשר $\hat{\underline{Z}}$ הינו שערוך וקטור הרעש.

<u>חסם האמינג</u>:

מס' מילות הקוד M מקיים:

$$M \leq \frac{2^n}{\left|\text{Hamming ball}\left(r=\frac{d}{2}\right)\right|} \underset{\text{for large } n}{\approx} 2^{n\left(1-H_B\left(\frac{d}{2n}\right)\right)}=2^{nC}$$

וכפי שניתן לראות, הדבר מסתדר עם התיאוריה של שאנון.

<u>חסם ורשמוב-גילברט</u>:

<u>קיים</u> קוד בינארי המקיים:

$$\left|\text{Hamming ball } (r=d)\right| \geq \frac{2^n}{M}$$

כלומר, קיים קוד שעבורו מתקיים - $R>1-H_B(2p)$.

על-מנת לחבר את התיאוריה של שאנון עם הקודים הלינאריים, נגריל את $\underline{\underline{G}}$ באופן אקראי ע"י פילוג אחיד עם הסתברות 0.5. זהו בעצם קוד ליניארי אקראי. באופן זה, כל מה שהוכח ע"י שאנון עבור קודים אקראיים תקף גם לגבי קודים ליניאריים. ניתן להוכיח שביחס למפענח ML או חיפוש בכדור BSC כי $p_e \to 0$ עבור $n \to \infty$ בתנאי ש- $R<1-H_B(p)$.

11. תורת האינפורמציה לאותות רציפים

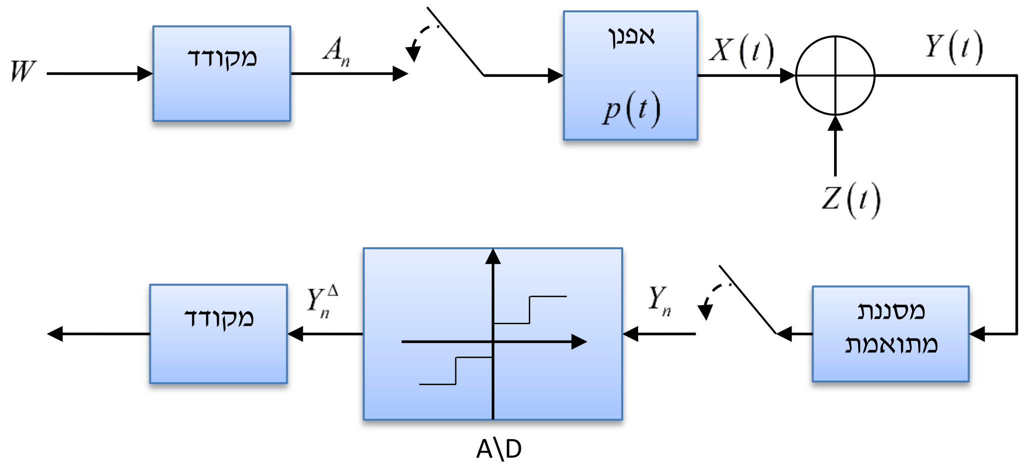
11.1. הקדמה

כאשר אנטרופיית המקור גדולה מקיבול הערוץ ($H>C$), קיימת בעייה בהעברת מידע דרך הערוץ בהסתברות שגיאה קטנה כרצוננו. מה בכל זאת ניתן לעשות?

- ✓ לא לשדר חלק מדגימות המקור.
- ✓ לשדר לאט יותר – כלומר להקטין את הקצב R.
- ✓ קוונטיזציה של האות – דחיסה עם עיוות.

עבור אות רציף מתקיים $H>C$, ויותר מכך - $H\to\infty$ כי קיימים אינסוף ערכים אותם משתנה רציף יכול לקבל.

11.1.1. מודל טיפוסי של ערוץ



המידע W מקודד לביטי המידע A_n שמאופננים ע"י האפנן $p(t)$ והופכים לאות רציף שמשודר בתווך הרציף. התווך ממודל בד"כ כערוץ אדיטיבי גאוסי (AWGN). אות זה עובר בתווך, נקלט ועובר גילוי ע"י מסננת מתואמת. אח"כ נדגם ומומר חזרה לאות דיגיטלי ע"י רכיב A\D. האות במוצא הוא קוונטיזציה לאות הנקלט. ניתן לשים לב כי הכניסה והמוצא של הסכימה הם דיגיטלים ולא רציפים. לכן, ניתן למצוא ערוץ דיגיטלי שקול לכל המרכיבים הנמצאים בין אות הכניסה לאות המוצא. נרצה לפתח תורה המאפיינת ערוץ רציף עם כלים שלמדנו עבור הערוץ הבדיד.

11.2. מדדי אינפורמציה לאותות רציפים

נניח כי X,Y הם מ"א רציפים שמקבלים אינסוף (בר או לא בר מניה) של ערכים. באופן כללי משתנים אקראיים רציפים מאופיינים ע"י פונקציית התפלגות:

$$F_X(x)=p(X\le x)$$

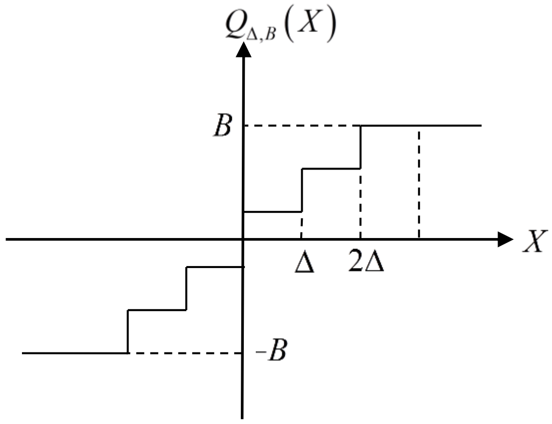
אם פונקציית ההתפלגות היא גזירה, אזי קיימת למ"א פונקציית צפיפות הסתברות:

$$f_X(x)=\frac{dF}{dx}\approx p(X=x)\cdot dx$$

עבור מ"א המקבלים אינסוף ערכים מתקיים כי $H\to\infty$. מצד שני, האינפורמציה ההדדית היא סכום של שני איברים השואפים לאינסוף - $I(X;Y)=H(X)-H(X|Y)$, ולכן ערך זו יכול להיות סופי. נרצה איפוא להגדיר את האינפורמציה ההדדית בדרך "עוקפת אנטרופיה".

הגדרה:

קוונטייזר הוא פונקציה המקבלת משתנה רציף וממירה אותו למס' ערכים. לדוגמא, קוונטייזר עם 6 רמות נראה כך:



מגדירים $X^{\Delta} \triangleq Q_{\Delta,B}(X)$. X^{Δ} הוא בעצם קירוב של X ע"י מספר סופי של ערכים (המדרגות שבגרף). המרחק בין הערכים הוא B/Δ כאשר B ו- $-B$ הם ערכי הרוויה של הקוונטייזר – הוא לא יכול להוציא ערך מחוץ לתחום זה.

בניגוד ל- $H(X)$, $H(X^{\Delta})<\infty$ לכל $B<\infty, \Delta>0$. לכן, ניתן להגדיר:

$$I(X;Y) \triangleq \sup_{\Delta,B} I(X^{\Delta};Y^{\Delta}) = \sup_{\Delta,B}\left\{H(X^{\Delta}) - H(X^{\Delta} \mid Y^{\Delta})\right\}$$

ניתן להראות כי:

$$I(X;Y) = \lim_{\substack{\Delta\to 0\\ B\to\infty}} I(X^{\Delta};Y^{\Delta})$$

הדבר נובע מתכונת העידון של האנטרופיה. בצורה דומה ניתן להגדיר את $D(p_X \| p_Y)$.

נצמצם את הדיון למ"א בעלי פונקיית צפיפות הסתברות (כלומר, משתנים אקראיים בעלי פונקציית התפלגות $F(X)$ גזירה). נניח המשתנים האקראיים X,Y הם בעלי פונקציית צפיפות משותפת $f(X,Y)$. ניזכר כי מתקיים:

$$f(X) = \int_{-\infty}^{\infty} f(X,y)\,dy$$

$$f(X \mid Y) = \frac{f(X,Y)}{f(Y)}$$

הגדרה:

"אנטרופיה דיפרנציאלית" עבור משתנה X המפולג לפי $f(X)$ מוגדרת כך:

$$h(X) \triangleq h(f(\cdot)) \triangleq -\int_{-\infty}^{\infty} f(x)\cdot \log f(x)\,dx = E\left\{-\log\frac{1}{f(X)}\right\}$$

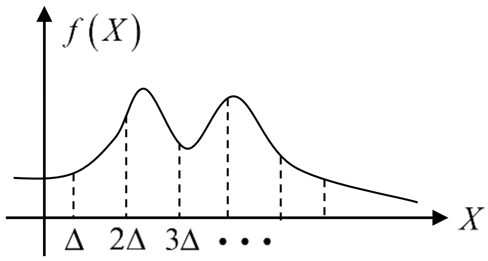
ניתן להגדיר באופן דומה את: $h(X \mid Y), h(X,Y), h(Y), h(X)$.

משפט:

אם $f(X)$ היא אינטגרבילית רימן אזי:

$$H(X^{\Delta}) + \log\Delta \xrightarrow[\Delta\to 0]{} h(X)$$

הוכחה:



$$p_i \triangleq p(X\in \text{cell } i)=\int_{\text{cell } i} f(x)dx=\Delta\cdot f(x_i)$$

כאשר x_i הוא אחד מהערכים בתא ה- i, כלומר - $x_i\in \text{cell } i$. נמשיך:

$$H(X^\Delta)=-\sum_i p_i\log p_i=-\sum_i p_i\log(\Delta\cdot f(x_i))=-\sum_i p_i\log\Delta-\sum_i p_i\log f(x_i)$$
$$=-\log\Delta-\sum_i\left[f(x_i)\cdot\log f(x_i)\right]\cdot\Delta\xrightarrow[\Delta\to 0]{}-\log\Delta+h(X)$$

<u>הערה</u>: אנטרופיה דיפרנציאלית יכולה לקבל ערכים שליליים. לכן, היא אין לה משמעות של אינפורמציה.

11.3. השלכות של אנטרופיה לצפיפות שהיא אינטגרבילית רימן

$$I(X^\Delta;Y)=H(X^\Delta)-H(X^\Delta\,|\,Y)=\left[H(X^\Delta)+\log\Delta\right]-\left[H(X^\Delta\,|\,Y)+\log\Delta\right]$$
$$\xrightarrow[\Delta\to 0]{}h(X)-h(X\,|\,Y)$$

כזכור, מתקיים:

$$I(X;Y)=\lim_{\substack{\Delta\to 0\\ B\to\infty}} I(X^\Delta;Y^\Delta)$$

ולכן נובע כי ניתן לבטא את האינפורמציה ההדדית בצורה הבאה (אם קיימות צפיפויות הפילוג):

$$I(X;Y)=h(X)-h(X\,|\,Y)=$$
$$=h(Y)-h(Y\,|\,X)=$$
$$=h(X)+h(Y)-h(X,Y)$$

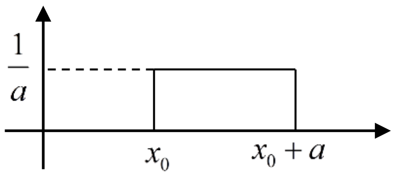
נשים לב כי $I(X;Y)$ הינו גודל חיובי:

$$\underbrace{I(X;Y)\triangleq\sup_{\Delta,B}\underbrace{I(X^\Delta;Y^\Delta)}_{\geq 0}}_{\geq 0}$$

בנוסף, ניתן להגדיר אינפורמציה מותנית כמו במקרה הבדיד ומתקיימים גם כלל השרשרת ואי-שוויון עיבוד הנתונים.

<u>דוגמאות</u>:

א. מ"א אקראי אחיד - $X\sim unif(x_0,x_0+a)$

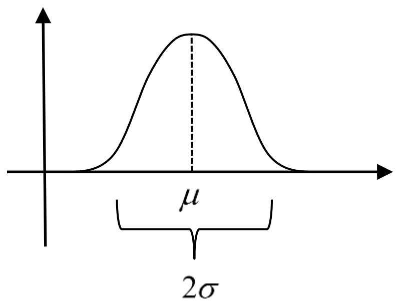


במקרה זה מתקיים:

$$h(X)=\log a$$

נשים לב שאם $a=1$ אז $h(X)=0$ ואם $a<1$ אז $h(X)<0$.

ב. פילוג גאוסי - $X \sim N(\mu,\sigma^2)$.



$$f(X)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

במקרה זה מתקיים:

$$h(X)=\frac{1}{2}\log\left(2\pi e\sigma^2\right)$$

נשים לב כי בשתי הדוגמאות מתקיים כי $h(X)$ איננה תלויה בתוחלת הפילוג!

11.4. תכונות מקסימום אנטרופיה

א. אם X מוגבל באינטרוול בגודל a, אזי $-\infty \le h(X) \le \log(a)$ ושוויון מתקיים אמ"מ X הוא משתנה אחיד על-פני האינטרוול הנ"ל.

ב. אם $\mathrm{var}(X) \le \sigma^2$, כלומר השונות היא בעלת ערך סופי, אזי $-\infty \le h(X) \le \frac{1}{2}\log\left(2\pi e\sigma^2\right)$, ושוויון מתקיים אמ"מ X הינו גאוסי - $X \sim N(\mu,\sigma^2)$.

<u>הערה</u>: ניתן להוכיח כי אם $EX=\mu$ ו- $\mathrm{var}(X)=\sigma^2$, אזי:

$$0 \le D\left(f(X)\,\|\,N(\mu,\sigma^2)\right)=h\left(N(\mu,\sigma^2)\right)-h(X)$$

ומכאן נובע סעיף ב'.

11.5. משפט ה-AEP במקרה הרציף

<u>הגדרה</u>:

"*<u>קבוצה אופיינית</u>*" למקור iid עם צפיפות פילוג $f(X)$ מוגדרת בצורה הבאה:

$$A_\varepsilon^{(n)} \triangleq \left\{X^n : \left|-\frac{1}{n}\log f\left(X^n\right)-h(X)\right|<\varepsilon\right\}$$

<u>משפט</u>:

א. לכל $X^n \in A_\varepsilon^{(n)}$ מתקיים: $f\left(X^n\right)\approx 2^{-nh(X)}$.

ב. מחוק המספרים הגדולים נובע כי עבור $\varepsilon>0$ ו- n גדול מספיק מתקיים:

$$p\left\{X^n \in A_\varepsilon^{(n)}\right\}>1-\varepsilon$$

ג. הנפח של הקבוצה האופיינית:

$$(1-\varepsilon)2^{n[h(X)-\varepsilon]}\underset{\substack{\text{for } n \text{ that's}\\ \text{large enough}}}{\le} vol\left(A_\varepsilon^{(n)}\right)\underset{\forall n}{\le} 2^{n[h(X)+\varepsilon]}$$

כלומר, לאנטרופיה דיפרנציאלית יש משמעות של "אקספוננט הנפח" של סדרות ארוכות של אותו המ"א. ניתן גם להגדיר בצורה דומה קבוצה אופיינית במשותף של זוג מ"א רציפים.

11.6. בעיית הקיבול לערוץ רציף כללי

עבור הערוץ הרציף הקיבול האינפורמציוני מוגדר בצורה דומה:

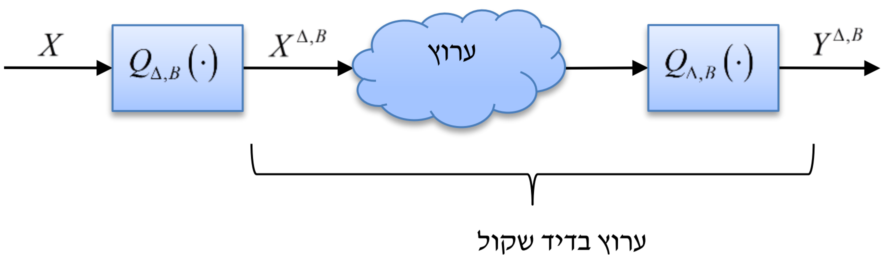
$$C^{\inf} \triangleq \max_{p(X)} I(X;Y)$$

עבור ערוץ חסר-זיכרון מתקיים כי הקיבול האופרטיבי שווה לקיבול האינפורמציוני: $C^{\inf} = C^{op}$.

<u>הצדקה למשפט</u>:

יש להזכר בהוכחת המשפט עבור המקרה הבדיד ורק אז לקרוא חלק זה.

א. <u>הכיוון הישר</u>:



ניתן להשיג את C של הערוץ הבדיד ע"י משפט הקידוד לערוצים בדידים ואז להשאיף לגבול: $\Delta \to 0, B \to \infty$.

ב. <u>הכיוון ההפוך</u>:

ניתן להפעיל את המשפט ההפוך כרגיל:

$$I(W:Y^n) \le n \cdot C$$

כאשר ההודעה W הינה בדידה והמוצא Y^n הוא רציף.

האנטרפויה של הודעות שוות הסתברות היא כזכור:

$$H(W) = nR$$

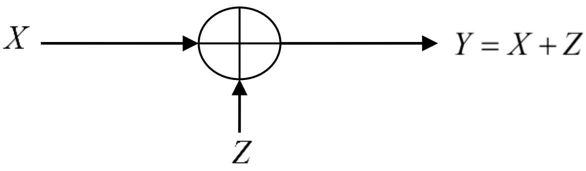
ולכן, גם פה מתקיים:

$$H(W|Y^n) \ge n(R-C)$$

החסם התחתון מושג ע"י שימוש באי-שוויון פאנו.

12. הערוץ הגאוסי הלבן (AWGN Channel)

ניתן לתאר את הערוץ האדיטיבי הגאוסי בצורה הבאה:

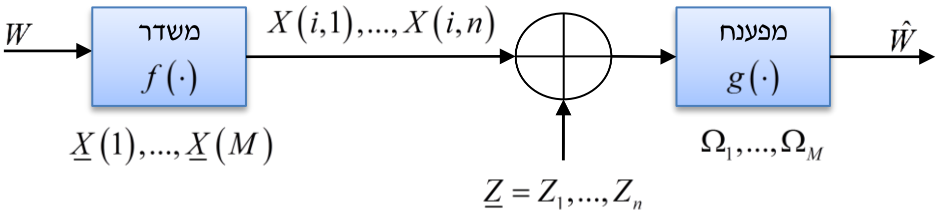


Z הינו רעש לבן גאוסי המפולג $Z \sim N(0, N)$, כלומר:

$$f_Z(z) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{z^2}{2N}}$$

Z הינו בת"ס באות הכניסה X.

סכימת ערוץ השידור עליה נדון היא:



12.1. אילוץ הספק בערוץ רציף

כשהערוץ הוא אנאלוגי אזי קיים <u>אילוץ הספק</u> כניסה על האות אותו משדרים בערוץ:

א. <u>גרסא מחמירה</u>:

כל מילת קוד מקיימת את האילוץ:

$$\frac{1}{n}\left\|\underline{X}(i)\right\|^2 = \frac{1}{n}\sum_{j=1}^{n}\left|\underline{X}(i,j)\right|^2 \le P \quad ; \quad \forall i = 1,\dots,M$$

כלומר, ההספק של כל מילות הקוד קטן מערך מסוים.

ב. <u>גרסא מקילה (ממוצעת)</u>:

בגרסא זו ההספק הממוצע של כל מילות הקוד קטן מערך מסוים:

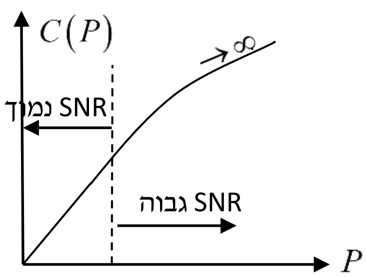
$$\frac{1}{M}\sum_{i=1}^{M}\frac{1}{n}\left\|\underline{X}(i)\right\|^2 \le P$$

12.2. הקיבול האינפורמציוני של ערוץ גאוסי

הקיבול האינפורמציוני של ערוץ גאוסי נתון ע"י:

$$C^{\inf} = \max_{\{X : EX^2 \le P\}} I(X; X+Z) = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

הערך של P/N מכונה "יחס האות לרעש" (SNR).



<u>הוכחה</u>:

$$I(X;Y)=h(Y)-h(Y|X)=h(X+Z)-h(Z)\le h_{\max}-h(Z)$$

האנטרופיה הדיפרנציאלית המקסימלית מתקבלת כאשר הפילוג של המ״א X הוא גאוסי, ואז גם $X+Z$ הוא גאוסי ולכן:

$$C^{\inf}=\max_{\substack{p(X)\\ EX^2\le P}} I(X;Y)=\frac{1}{2}\log\left(2\pi e(P+N)\right)-\frac{1}{2}\log(2\pi eN)=\frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

<u>תכונות</u>:

א. הקיבול ביחס אות לרעש גבוה הוא:

$$C_{HSNR}=\frac{1}{2}\log\left(\frac{P}{N}\right)$$

כלומר, הקשר בין יחס האות לרעש לקיבול הינו לוגריתמי.

ב. הקיבול ביחס אות לרעש נמוך הוא:

$$C_{LSNR}=\frac{\log e}{2}\cdot\frac{P}{N}$$

כלומר, הקשר בין היחס אות לרעש לקיבול הערוץ הוא קשר ליניארי.

ג. הקיבול הוא פונקציה קמורה $\cap$ של אילוץ ההספק.

ד. הקיבול הוא פונקציה מונוטונית לא יורדת של אילוץ ההספק.

12.3. הוכחת המשפט ההפוך עם אילוץ הספק

נמצא חסם על האינפורמציה הנותרת עבור הגרסא הממוצעת של אילוץ ההספק:

$$I(W;Y^n)\le I(X^n;Y^n)\underset{(1)}{\le}\sum_{j=1}^{n}I(X_j,Y_j)\underset{(2)}{\le}\sum_{j=1}^{n}C(P_j)\underset{(3)}{\le}n\cdot C\left(\frac{1}{n}\sum_{j=1}^{n}P_j\right)\underset{(4)}{\le}n\cdot C(P)$$

כאשר P_j הינו ההספק הממוצע בזמן j:

$$P_j\triangleq\frac{1}{M}\sum_{i=1}^{M}\left|X(i,j)\right|^2$$

הסברים למעברים:

(1) סכום אינפורמציות הדדיות שוליות גדול שווה לאינפורמציה ההדדית המשותפת.

(2) הקיבול הוא חסם עליון על האינפורמציה ההדדית.

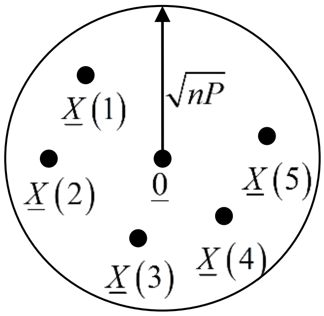
(3) קונבקסיות הקיבול כפונקציה של ההספק (כי הקיבול הוא בעצם אינפורמציה הדדית והיא קונבקסית).

(4) אילוץ ההספק.

המשך ההוכחה משתמש באי-שוויון פאנו – בדיוק באותו אופן כפי שעשינו בעבר. נשים לב כי המשפט ההפוך מתייחס לאילוץ ההספק בגירסא <u>המקילה</u>.

12.4. כדורי אינפורמציה במרחב האותות

ניתן לבטא את מרחב האותות ב- $\mathbb{R}^n$ בצורה גראפית. נבחן תחילה את מילות הקוד היוצאות מהמשדר.
עקב אילוץ הספק <u>המחמיר</u> הן כולן מקיימות: $\left\|X(i)\right\| \le \sqrt{nP}$. כלומר, "המרחק" של כל מילות הנקודות מנקודת האפס מוגבל ע"י ערך מסוים. נבטא זאת בעזרת כדור:



כדור זה מכונה "כדור השידור". בכדור זה כל מילות הקוד מוגבלות ברדיוס $\sqrt{nP}$ מ- $\underline{0}$.
נגדיר עתה כדור n - מימדי:

$$B(\underline{c},r) \triangleq \left\{\underline{X} : \left\|\underline{X}-\underline{c}\right\| \le r\right\}$$

הנפח של כדור n-מימדי הוא:

$$vol\left(B(\underline{c},r)\right) = V_n \cdot r^n$$

כאשר V_n הוא נפח של כדור n-מימדי עם רדיוס 1.
מתקיים:

$$V_1 = 2$$
$$V_2 = \pi$$
$$V_3 = \frac{4}{3}\pi$$
$$\vdots$$

$$V_n = \frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} \;;\; n \text{ even} \quad ; \quad V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{\frac{n}{2}}}{\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{5}{2}\cdot\ldots\cdot\frac{n}{2}\cdot\sqrt{\pi}} \;;\; n \text{ odd}$$

<u>אסימפטוטיקה במימד</u>:

א. קירוב אקספוננציאלי כאשר $n \to \infty$:

$$V_n \doteq \left(\frac{2\pi e}{n}\right)^{\frac{n}{2}}$$

ב. היחס בין נפח כדור לבין קליפת כדור.

נחשב את היחס בין כדור ברדיוס r לבין כדור ברדיוס "קצת" קטן יותר $r(1-\varepsilon)$:

$$\frac{\text{Ball volume of radious } r(1-\varepsilon)}{\text{Ball volume } r} = \frac{V_n \cdot \left(r(1-\varepsilon)\right)^n}{V_n \cdot r^n} = (1-\varepsilon)^n \xrightarrow[n\to\infty]{} 0$$

כלומר, עבור $n \to \infty$ רוב הנפח של הכדור מרוכז בקליפתו.

12.5. מקלט אופייניות משותפת

כזכור, מקלט אופייניות משותפת הוא:

$$\hat{W}=\begin{cases} i & if\ \left(\underline{X}(i),\underline{Y}\right)\in A_{\varepsilon}^{(n)}(X,Y) \quad \text{for some unique i} \\ ? & \text{otherwise} \Rightarrow \begin{array}{c}\text{all the codewords } \underline{X}(i) \text{ are not typical with } \underline{Y} \\ \text{or there is more then one codeword that is typical with } \underline{Y}\end{array}\end{cases}$$

הקבוצה האופיינית של מ"א X היא:

$$A_{\varepsilon}^{(n)}(X)=\left\{X^n \mid -\frac{1}{n}\log f\left(x^n\right)=h(X)\pm\varepsilon\right\}$$

עבור מ"א גאוסי $X\sim N\left(0,\sigma_x^2\right)$:

$$f\left(X^n\right)=\left(\frac{1}{\sqrt{2\pi\sigma_x^2}}\right)^n e^{-\frac{\|X\|^2}{2\sigma_x^2}}$$

ולכן:

$$A_{\varepsilon}^{(n)}(X)=\left\{X^n \mid \frac{1}{n}\left\|x^n\right\|^2=\sigma_x\pm\varepsilon'\right\}$$

התיאור הנ"ל הוא בעצם תיאור של קליפה כדורית ברדיוס $\sqrt{n\sigma_x^2}$. עבור $n\gg 1$, נקרב את הקליפה הכדורית לכדור, ולכן מתקיים:

$$A_{\varepsilon}^{(n)}(X)=B\left(\underline{0},\sqrt{n\sigma_x^2}\right)$$

הקבוצה האופיינית במשותף של X ו- Y היא:

$$A_{\varepsilon}^{(n)}(X,Y)=\left\{\left(x^n,y^n\right): \begin{array}{c} -\frac{1}{n}\log f_X\left(x^n\right)=h(X)\pm\varepsilon \\ -\frac{1}{n}\log f_Y\left(y^n\right)=h(Y)\pm\varepsilon \\ -\frac{1}{n}\log f_{X,Y}\left(x^n,y^n\right)=h(X,Y)\pm\varepsilon \end{array}\right\}$$

ובמקרה בו $Y=X+Z$, $X\sim N\left(0,\sigma_x^2\right), Z\sim N\left(0,\sigma_z^2\right)$ ובנוסף $X \perp\!\!\!\perp Z$ מתקיים:

$$A_{\varepsilon}^{(n)}(X,Y)=\left\{\left(x^n,y^n\right): \begin{array}{c} \frac{1}{n}\left\|x^n\right\|^2=\sigma_x^2+\varepsilon' \\ \frac{1}{n}\left\|y^n\right\|^2=\sigma_y^2+\varepsilon' \\ \frac{1}{n}\left\|y^n-x^n\right\|^2=\sigma_z^2+\varepsilon' \end{array}\right\}$$

נבחן עתה את מודל הערוץ ההפוך:

ניתן לבטא את הקבוצה האופיינית כך:

$$A_{\varepsilon}^{(n)} = \left\{ \left(x^n, y^n \right) : \begin{array}{c} \text{same marginal constraints as before} \\ x^n \in B\left(\alpha^* y^n, \sqrt{n \cdot LMMSE} \right) \end{array} \right\}$$

$$\alpha^* = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} \leftarrow \text{Wiener constant}$$

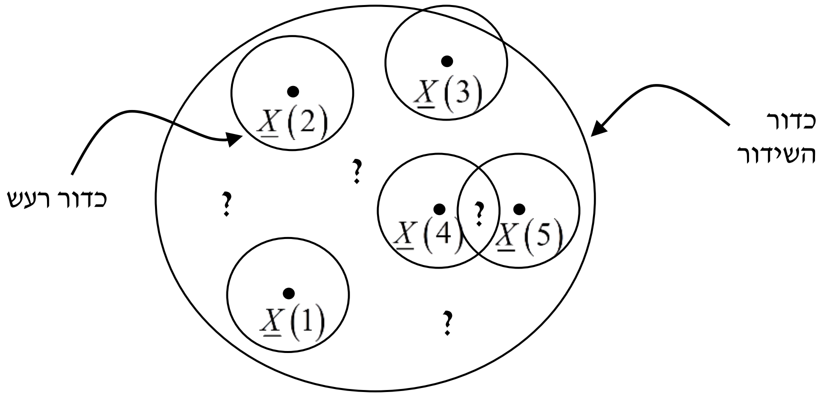
$$LMMSE = \frac{\sigma_x^2 \cdot \sigma_z^2}{\sigma_x^2 + \sigma_z^2}$$

עבור HSNR ($\sigma_x^2 \gg \sigma_z^2$) מתקיים:

$$\alpha^* = 1$$

$$LMMSE = \sigma_z^2$$

בעצם אנו יכולים להבחין שאזורי הגילוי הם בעצם "כדורי רעש" סביב $\underline{Y}$ ברדיוס $\sqrt{n\sigma_z^2}$. לחלופין, ניתן לצייר את הכדורים סביב מילות הקוד ולזהות את ה-$\underline{y}$-ים שעבורם יכריז המפענח "?". נתאר את תמונת הקליטה בצורה גראפית:



12.6. מפענח סבירות מירבית (Maximum Likelihood)

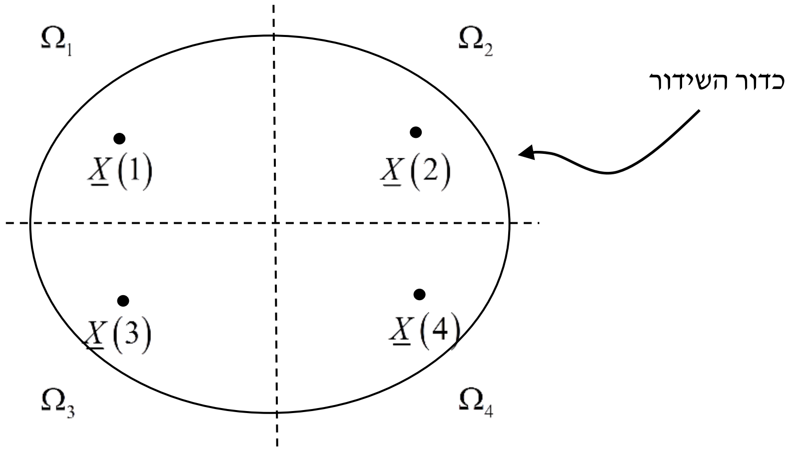
משערך ה-ML מוגדר בצורה הבאה:

$$\hat{W}_{ML}=\arg\max_i p\left(\underline{Y}\,|\,\underline{X}(i)\right)$$

עבור הערוץ האדיטיבי הגאוסי:

$$\hat{W}_{ML}=\arg\max_i f_{Y|X(i)}\left(\underline{y}\,|\,\underline{x}(i)\right)=\arg\max_i \underbrace{f_Z\left(\underline{Y}-\underline{X}(i)\right)}_{\propto\left\|\underline{Y}-\underline{X}(i)\right\|}=\arg\min_i\left\|\underline{Y}-\underline{X}(i)\right\|$$

כלומר, משערך ML מקבל קלט $\underline{Y}$ ובוחר את ה- $\underline{X}(i)$ שהכי "קרוב" אליו. ובצורה גרפית:



בהמשך הניתוח נניח הנחת HSNR - כלומר שכדורי הרעש המופיעים סביב ה-$\underline{X}(i)$ נמצאים כולם (או רובם) בתוך כדור השידור. במקרה זה איזור ההחלטה הקטן ביותר מקיים:

$$vol\left(\Omega_{\min}\right)\le\frac{vol\left(\text{Tx ball}\right)}{M}=\frac{V_n\left(\sqrt{nP}\right)^n}{2^{nR}}$$

<u>הגדרה</u>:

<u>*רדיוס אפקטיבי*</u> של איזור ההחלטה $\Omega_{\min}$ מוגדר להיות הרדיוס של הכדור בעל אותו הנפח כמו $\Omega_{\min}$. כלומר:

$$V_n\cdot r_{eff}^n=\frac{V_n\cdot\left(\sqrt{nP}\right)^n}{2^{nR}}\Rightarrow r_{eff}=\frac{\sqrt{nP}}{2^R}$$

נשים לב כי בד"כ תחום ההחלטה עם הנפח המינימלי יגרום להסתברות השגיאה המקסימלית:

$$p_e^{\max}=\max_i p_e(i)$$

12.7. חסם תחתון על הסתברות השגיאה ואקספוננט אריזת הכדורים

<u>אי השוויון "האיזו פרימטרי" – Iso-perimeteric</u>:

$$p\left(\underline{Z}\notin\Omega_i\right)\ge p\left(\underline{Z}\notin\tilde{\Omega}_i\right)$$

כאשר $\tilde{\Omega}_i$ הינו כדור שנפחו זהה לנפח של Ω_i. כלומר, כדור היא צורה עבורה S הוא הכי נמוך עבור נפח נתון, כאשר S יכול להיות קוטר, מומנט שני, שטח פנים,...וגם הסתברות שגיאה בנוכחות רעש גאוסי!

ולכן הסתברות השגיאה מקיימת:

$$p_e^{\max}\ge p\left(\underline{Z}\notin\tilde{\Omega}_{\min}\right)=p\left(\underline{Z}\notin B\left(\underline{0},r_{eff}\right)\right)=p\left(\left\|\underline{Z}\right\|>r_{eff}\right)$$

בעצם, הטענה שקולה לכך שתתרחש שגיאה כאשר הרעש גדול מסף מסוים. ניתן לכתוב :

$$\sum_{i=1}^{n} Z_i^2 > r^2$$

הביטוי המדויק לשגיאה הוא :

$$\int_{r}^{\infty} f_{\|Z\|}\left(\tilde{r}\right) d\tilde{r}$$

אבל במקום לחשב במדויק, ננסה לחסום את הביטוי הנ"ל.

<u>חסם צ'רנוף למעבר סף</u> :

עבור סכום iid מתקיים:

$$p\left(\sum_{i=1}^{n} Z_i^2 \geq r^2\right) \leq \left(\frac{E\left\{e^{sZ^2}\right\}}{e^{s\frac{r^2}{n}}}\right)^n ; \forall s > 0$$

עבור $Z \sim N\left(0, N\right)$, מחשבים את הפונקציה היוצרת מומנטים $E\left\{e^{sZ^2}\right\}$, מבצעים אופטימיזציה על $s > 0$ ומקבלים :

$$p\left\{\|\underline{Z}\|^2 \geq r_{eff}^2\right\} \leq e^{-nE_{sp}\left(\frac{r_{eff}^2}{r_{noise}^2}\right)}$$

כאשר:

$$E_{sp}\left(x\right) = \begin{cases} \frac{1}{2}\left[x - 1 - \ln x\right] & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

$$r_{noise} = \sqrt{nN}$$

בעזרת "שיטת האינטגרציה של לפלס" ניתן להוכיח כי חסם צ'רנוף הדוק במובן אקספוננציאלי עבור Z גאוסי, כלומר :

$$p\left(\|\underline{Z}\| > r_{eff}\right) \doteq e^{-nE_{sp}\left(\frac{r_{eff}^2}{r_{noise}^2}\right)}$$

<u>אקספוננט אריזת הכדורים</u> :

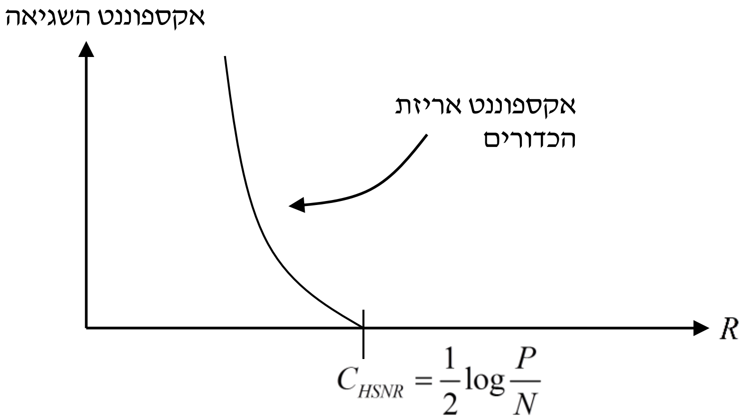
זהו בעצם מצב כאשר כל המרחב דחוס באופן מלא בכדורים ברדיוס r_{eff}.

$$p_e^{\max} \overset{\bullet}{\geq} e^{-nE_{sp}\left(\frac{r_{eff}^2}{r_{noise}^2}\right)}$$

כאשר אי-השוויון האחרון הוא <u>במובן אקספוננציאלי</u> גם כן.

נשים לב כי מתקיים :

$$\frac{r_{eff}^2}{r_{noise}^2} = \frac{\left(\frac{\sqrt{nP}}{2^R}\right)^2}{\left(\sqrt{nN}\right)^2} = \frac{P/N}{2^{2R}} = \frac{2^{2C_{HSNR}}}{2^{2R}} = 2^{2\left[C_{HSNR} - R\right]}$$



נשים לב כי אקספוננט אריזת הכדורים חיובי לכל $R<C$.

<u>משפט הפוך חזק</u>:

אם $R>C$ אזי מתקיים כי $r_{noise}>r_{eff}$. ובנוסף, לפי ה-AEP:

$$p\left(\|\underline{Z}\|>r_{eff}\right)\xrightarrow[n\to\infty]{}1$$

12.8. השגת הקיבול ע"י קוד אקראי אחיד וגלאי הסף

<u>תזכורת</u>:

עבור DMC, הסתברות השגיאה של קוד המוגרל לפי $p(X)$ מקיימת:

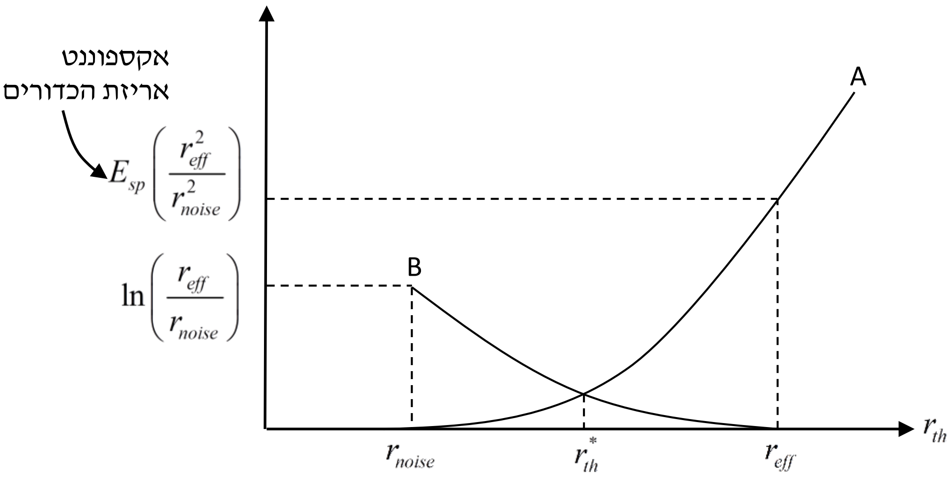
$$\bar{\bar{p}}_e\le \underbrace{p\begin{Bmatrix}\text{non-typical}\\ \text{behaviour}\end{Bmatrix}}_{\le\varepsilon}+\underbrace{(M-1)\cdot\underbrace{2^{-n[I(X;Y)-3\varepsilon]}}_{\text{probability of pretending}}}_{\text{average amount of pretending codewords}}$$

בהקבלה, בערוץ AWGN וקוד אקראי המתפלג באחידות על כדור השידור מתקיים:

$$\bar{\bar{p}}_e\le p\left(\|\underline{Z}\|>r_{th}\right)+(M-1)\frac{vol\left(B\left(\underline{0},r_{th}\right)\right)}{vol\left(B\left(\underline{0},\sqrt{nP}\right)\right)}\le p\left(\|\underline{Z}\|>r_{th}\right)+2^{nR}\frac{V_n\left(\sqrt{n\cdot r_{th}}\right)^n}{V_n\left(\sqrt{n\cdot P}\right)^n}=$$

$$=p\left(\|\underline{Z}\|>r_{th}\right)+\left(\frac{r_{th}}{r_{eff}}\right)^n\doteq\underbrace{e^{-nE_{sp}\left(\frac{r_{th}^2}{r_{noise}^2}\right)}}_{A}+\underbrace{e^{-n\ln\left(\frac{r_{eff}}{r_{th}}\right)}}_{B}$$

כאשר r_{th} הינו ערך סף שעל פיו המשערך מחליט. במקרה של <u>אופייניות משותפת</u>: $r_{th}\approx r_{noise}$



הערה: כאשר מסכמים את e^{-nA} ו-e^{-nB}, הדומיננטי ביניהם (עבור $n\to\infty$) הוא בעל האקספוננט המינימלי.

מסקנה מהגרף: כל עוד $r_{eff} > r_{noise}$ (כלומר, $R<C$) אזי רדיוס החיפוש r_{th}^{*} גדול מרדיוס הרעש r_{noise}, ולכן אקספוננט השגיאה של גלאי הסף הוא חיובי, ומכך נובע שניתן להשיג את הקיבול.

מצד שני, האקספוננט שהתקבל פחות טוב מאקספוננט אריזת הכדורים (אשר חוסם את הסתברות השגיאה מלעיל).

12.9. ניתוח גילוי מרחק מינימלי (Nearest Neighbor) לקוד אקראי

האלגוריתם בניתוח יהיה כדלהלן: נתנה ב- $\|Z\| = r$, נחשב את הסיכוי למילה מתחרה בכדור החיפוש $B(\underline{Y},R)$, נכפיל בפונקציית הצפיפות $f_{\|Z\|}(r)$ ונבצע אינטגרל לפי dr:

$$\bar{\bar{p}}_{e,NN} \le \int_0^\infty f_{\|Z\|}(r)\cdot \min\left\{\underbrace{(M-1)\cdot\frac{vol(\text{Search Ball})}{vol(\text{Tx Ball})}}_{=\left(\frac{r}{r_{eff}}\right)^n},1\right\}dr = \underbrace{\int_0^{r_{eff}} f_{\|Z\|}(r)\left(\frac{r}{r_{eff}}\right)^n dr}_{(I)} + \underbrace{\underbrace{\int_{r_{eff}}^\infty f_{\|Z\|}(r)dr}_{=p(\|Z\|>r_{eff})}}_{(II)}$$

שיטת לפלס לאינטגרציה:
מתקיים השוויון הבא:

$$\int_a^b e^{-nE(x)}dx \doteq e^{-n\cdot E_{\min}}$$

כאשר:

$$E_{\min} \triangleq \min_{a\le x\le b} E(x)$$

ואומרים כי "האינטגרל נשלט ע"י $E_{\min}$ בגבול $n\to\infty$.

משיטת לפלס נובע כי האקספוננט של $f_{\|Z\|}(r)$ זהה לאקספוננט של $\int_r^\infty f_{\|Z\|}(\tilde{r})d\tilde{r}$, כלומר, הוא שווה ל- $e^{-nE_{SP}\left(\frac{r^2}{r_{noise}^2}\right)}$. מכאן שבמובן אקספוננציאלי:

$$\bar{\bar{p}}_{e,NN} \overset{\bullet}{\le} \underbrace{\int_0^{r_{eff}} e^{-n\cdot E_{sp}\left(\frac{r^2}{r_{noise}^2}\right)}\cdot e^{-n\cdot\ln\left(\frac{r_{eff}}{r}\right)}dr}_{(I)} + \underbrace{e^{-nE_{sp}\left(\frac{r_{eff}^2}{r_{noise}^2}\right)}}_{(II)}$$

האיבר (I) נשלט ע"י:

$$r^* = \begin{cases} r_{eff} & if\ \frac{r_{eff}}{r_{noise}} < \sqrt{2} \\ \sqrt{2}\cdot r_{noise} & else \end{cases}$$

מכאן נובע שעבור $r_{noise} \le r_{eff} \le \sqrt{2}r_{noise}$, האקספוננטים (I) ו- (II) שווים. לכן, האקספוננט מתלכד עם אריזת הכדורים, כלומר:

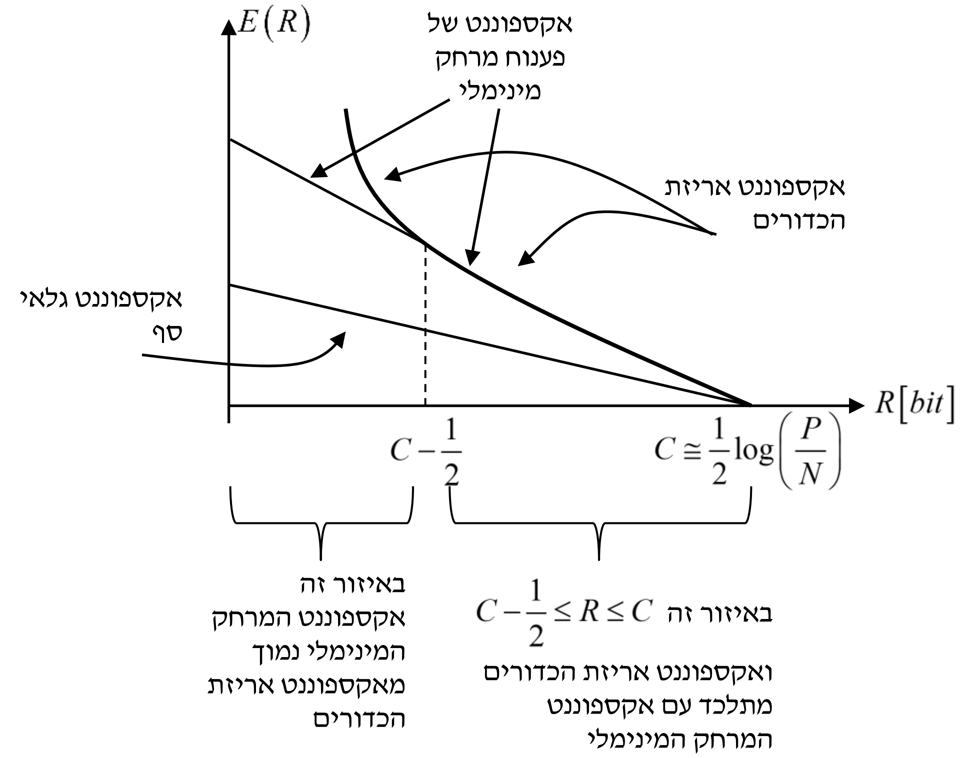
$$\bar{\bar{p}}_{e,NN} \overset{\bullet}{\le} e^{-nE_{sp}\left(\frac{r_{eff}^2}{r_{noise}^2}\right)}$$

בעוד שעבור $r_{eff} > \sqrt{2}r_{noise}$, האיבר הראשון (I) דומיננטי, ולכן האקספוננט נמוך מאקספוננט אריזת הכדורים, ושווה ל:

$$E_{sp}\left(\frac{r^2}{r_{noise}^2}\right)+\ln\left(\frac{r_{eff}}{r}\right)\Bigg|_{r=\sqrt{2}r_{noise}} = E_{sp}\left(2\right)+\ln\left(\frac{r_{eff}}{\sqrt{2}r_{noise}}\right)$$

נסכם את את נושא אקספוננטי שגיאה בערוץ גאוסי לבן ע"י השוואה בין האקספוננטים שונים.

ניזכר ש- $r_{eff} = r_{noise}$ מתאים לקצב $R = C$, והנקודה של $r_{eff} = \sqrt{2}r_{noise}$ מתאימה ל- $R = C - \frac{1}{2}[bit]$.



13. אקספוננט שגיאה לערוץ בדיד DMC

נתונות שתי מילות קוד $\underline{X}(1)$ ו- $\underline{X}(2)$. ע"פ משערך ML:

$$p\left(\underline{Y} \mid \underline{X}(2)\right) \underset{\hat{W}=1}{\overset{\hat{W}=2}{\gtrless}} p\left(\underline{Y} \mid \underline{X}(1)\right)$$

והסתברות השגיאה (עבור ערוץ חסר זיכרון):

$$p_{e,1} = p\left\{p\left(\underline{Y} \mid \underline{X}(2)\right) > p\left(\underline{Y} \mid \underline{X}(1)\right) | W=1\right\} =$$

$$= p\left\{\sum_{i=1}^{n} \underbrace{\ln \frac{p\left(Y_i \mid X_i(2)\right)}{p\left(Y_i \mid X_i(1)\right)}}_{\triangleq Z_i} > 0 \mid W=1\right\} = p\left\{\sum_{i=1}^{n} Z_i > 0 \mid W=1\right\}$$

אם מילות הקוד מוגרלות $\underline{X}(j) \sim iid\, Q(X)$, אזי Z_i הם iid, ואפשר להפעיל עליהם את חסם צ'רנוף:

$$p\left\{\sum_{i=1}^{n} Z_i > 0\right\} \le \left(E\left\{e^{sZ}\right\}\right)^n, \forall s > 0$$

לאחר חישוב הפונקציה היוצרת מומנטים $E\left\{e^{sZ}\right\}$ מקבלים:

$$\bar{p}_{e,1} \le \left(\sum_y \left[\sum_x Q(x)\sqrt{p(y \mid x)}\right]^2\right)^n, \; s = \frac{1}{2}$$

$$\triangleq e^{-n \log R_0}$$

כאשר:

$$R_0 \triangleq -\log\left(\sum_y \left[\sum_x Q(x)\sqrt{p(y \mid x)}\right]^2\right)$$

עבור $M = 2^{nR}$ מילים, נפעיל את חסם האיחוד:

$$\bar{P}_e \le (M-1)e^{-nR_0} = e^{-n[R_0 - R]}$$

<u>הערה</u>: R_0 קטן מהקיבול, ולכן חסם זה איננו שימושי בתחום $R_0 < R < C$

<u>חסם גלאגר (1965)</u>:

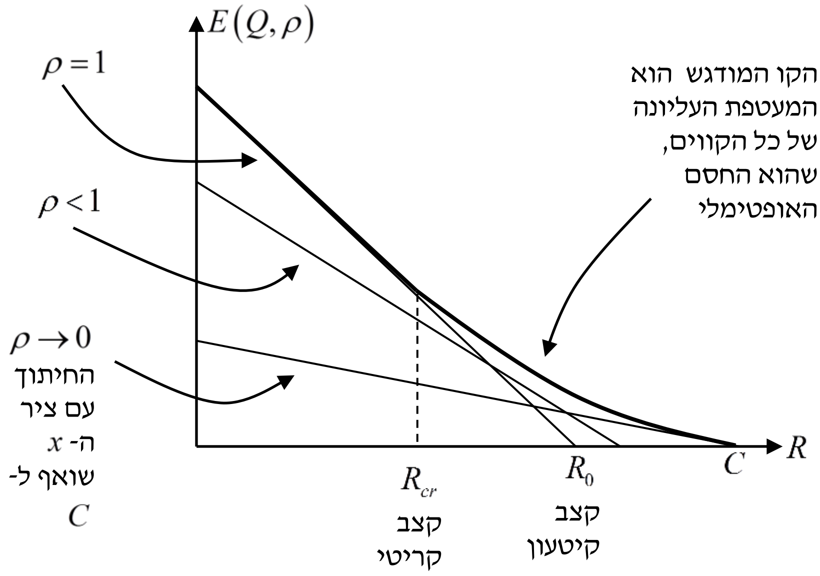
$$\bar{\bar{p}}_e \le e^{-nE(R)}$$

$$E(R) = \max_{0 \le \rho \le 1}\left[\max_Q E_0(Q,\rho) - \rho R\right]$$

$$E_0(Q,\rho) \triangleq -\log\left(\sum_y \left[\sum_x Q(x) \cdot \left(p(y \mid x)\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right)$$

נשים לב כי הביטוי עבור $E_0(Q,\rho)$ מתלכד עם R_0 עבור $\rho = 1$.

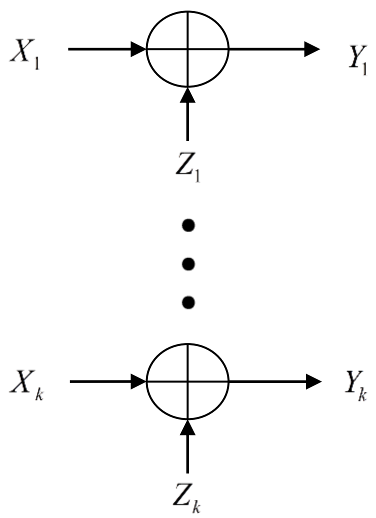
להלן ציור של חסם גלאגר עם ρ -ים שונים :



<u>הערה</u> : הגרף הנ״ל צויר עבור הפילוג המגשים $Q^*(x)$. באופן כללי, נקודת המפגש הימנית ביותר עם ציר R היא המקסימום המושג על האינפורמציה ההדדית ע״י ρ ו- $Q(x)$ כלשהם.

14. ערוצים צבעוניים

14.1. ערוצים גאוסיים במקביל עם אילוץ הספק משותף



בסכימה הנ״ל קיימים k ערוצים במקביל ומתקיים :

$$Z_j \sim N(0,N_j),\ j=1,\dots,N$$

$$Z_i \perp\!\!\!\perp Z_j,\ i\neq j$$

$$\sum_{j=1}^{k} EX_j^2 < P$$

נרצה לחשב את קיבול הערוץ הנ"ל:

$$C\left[\frac{\text{bit}}{\text{vector use}}\right]=\max_{\{p(X)|E\|X\|^2<P\}} I\left(\underbrace{X_1,\dots,X_k}_{\underline{X}};\underbrace{Y_1,\dots,Y_k}_{\underline{Y}}\right)\underset{(1)}{=}\max_{\{p(X)|E\|X\|^2<P\}}\sum_{j=1}^{k} I\left(X_j;Y_j\right)=$$

$$\underset{(2)}{=}\max_{\left\{p_1,\dots,p_k|\sum_{j=1}^{k}p_j<p\right\}}\sum_{j=1}^{k}\frac{1}{2}\log\left(1+\frac{p_j}{N_j}\right)=\max_{\left\{p_1,\dots,p_k|\sum_{j=1}^{k}p_j<p\right\}}\sum_{j=1}^{k}\frac{1}{2}\left(\log\left(N_j+p_j\right)-\log N_j\right)$$

כאשר מעבר 1 נובע מכך שקיבול ערוצים בת"ס מושג ע"י כניסות בת"ס, ומעבר 2 נובע מכך שמקסימום אינפורמציה הדדית בערוץ גאוסי מושגת ע"י כניסה גאוסית. הביטוי הנ"ל הוא בעצם בעיית אופטימיזציה של "הקצאת הספקים" (Power Allocation).

נמצא חסם עליון על הסכום בביטוי הנ"ל:

$$\sum_{j=1}^{k}\frac{1}{2}\left(\log\left(N_j+p_j\right)-\log\left(N_j\right)\right)\underset{\text{convexity}}{\le}\frac{1}{2}k\left(\log\left(\frac{1}{k}\sum_{j=1}^{k}\left(N_j+p_j\right)\right)-\underbrace{\frac{1}{k}\sum_{j=1}^{k}\log N_j}_{=\log\bar{\bar{N}}}\right)=$$

$$=\frac{1}{2}k\log\left(\frac{\bar{N}+\bar{p}}{\bar{\bar{N}}}\right)\le\frac{1}{2}k\log\left(\frac{\bar{N}+P/k}{\bar{\bar{N}}}\right)$$

$$\bar{\bar{N}}=\sqrt[k]{N_1\cdot\ldots\cdot N_k}$$

זהו החסם העליון של שאנון לקיבול ערוץ אדיטיבי כלשהו. שוויון בחסם יתקיים כאשר:

$$p_j+N_j=const=\bar{N}+\bar{p}=\bar{N}+\frac{P}{k}$$

נחזור לאופטימיזציה, ונבצע אותה בעזרת כופלי לגרנז':

$$L\left(p_1,\dots,p_k,\lambda\right)=\sum_{j=1}^{k}\frac{1}{2}\log\left(1+\frac{p_j}{N_j}\right)-\lambda\sum_{j=1}^{k}p_j$$

$$\frac{\partial L}{\partial p_j}=\frac{1}{2}\frac{1}{p_j+N_j}-\lambda=0,\ j=1,..,k$$

$$\underset{\nu=\frac{1}{2\lambda}}{\Rightarrow} p_j=\nu-N_j$$

כאשר ν נבחר כאשר: $\sum_j p_j=P$.

<u>תנאי קון-תאקר</u>:

מקסימיזציה של פונקציה קמורה על-פני תחום קמור מקיימת במקסימום:

- הגרדיאנט שווה ל-0 אם הוא מושג בתוך התחום.

או:

- גרדיאנט פונה בניצב לשפה (כלפי חוץ), אם הוא מושג בשפה.

במקרה שלנו, השפה היא: $p_1\ge 0,\dots,p_k\ge 0$ וצריך להתקיים:

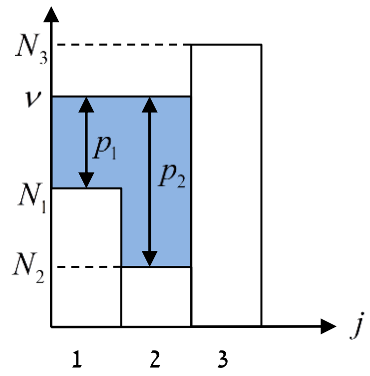
$$\frac{\partial L}{\partial p_j}=0\ \text{ if }\ p_j^{opt}>0\quad;\quad\frac{\partial L}{\partial p_j}\le 0\ \text{ if }\ p_j^{opt}=0$$

ולכן:

$$p_j=\left[\nu-N_j\right]^+$$

$$\left[x\right]^+=\begin{cases}x & x\geq 0\\ 0 & x\leq 0\end{cases}$$

ניתן לבטא את הפיתרון הנ"ל ע"פ כלל מזיגת המים:



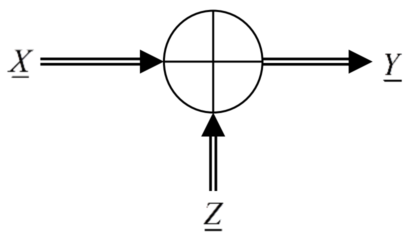
ν הינו גובה פני המים, וההספק שיש להקצות לכל תת-ערוץ הוא ההפרש בין רמת הרעש ל- ν.

ולכן הקיבול הוא:

$$C=\sum_j\frac{1}{2}\log\left(1+\frac{p_j^{opt}}{N_j}\right)=\sum_{\left\{j|p_j^{opt}>0\right\}}\frac{1}{2}\log\frac{\nu}{N_j}$$

14.2. ערוץ וקטורי עם רעש צבעוני

נתון הערוץ הבא:



כאשר $\underline{Z}$ הינו רעש צבעוני המתפלג לפי: $\underline{Z}\sim N\left(\underline{0},K_Z\right)$. נבצע המרה של הבעייה הנ"ל לבעייה של ערוצים גאוסיים במקביל. נעשה זאת ב-2 שלבים:

<u>שלב 1</u>:

ליכסון מטריצת הקווריאנס של הרעש:

$$K_z=Q\cdot\Lambda\cdot Q^{-1}$$

כאשר:

$$\Lambda=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_k\end{pmatrix}$$

λ_i הינם הערכים העצמיים של K_Z ו- Q הינה מטריצה אורתונורמלית המורכבת מהוקטורים העצמיים של K_Z: $\left[\underline{v}_1\ \vdots\ \underline{v}_2\ \vdots\ \cdots\ \vdots\ \underline{v}_k\right]$.

נציין כי K_Z הינה מטריצה <u>אי-שלילית מוגדרת</u>:

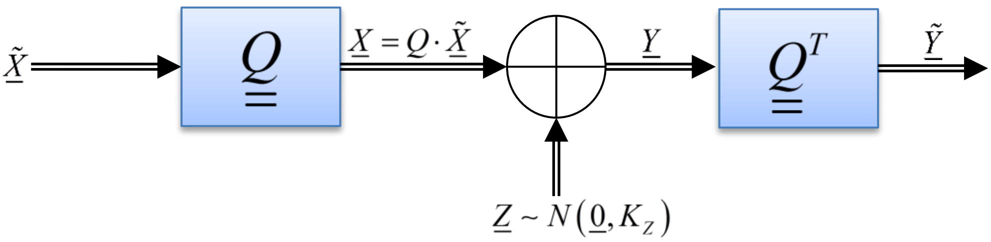
- $\lambda_i \geq 0$ והם ממשיים.
- $\underline{v}^T \cdot K_Z \cdot \underline{v} \geq 0$.
- ו"ע אורתונורמליים - $\underline{v}_i \perp \underline{v}_j$

מהתנאים הנ"ל נובע כי:

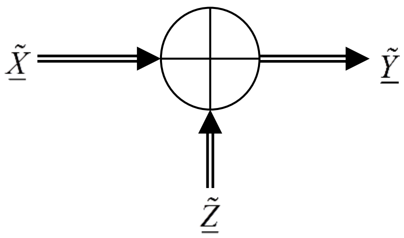
$$Q^{-1} = Q^T \Leftrightarrow Q \cdot Q^T = I$$

<u>שלב 2</u>:

המרה למרחב מלוכסן (מרחב הטרנספורם). נכפיל את הסכימה הנ"ל במטריצות בצורה הבאה:



נשים לב שהסכימה הנ"ל שקולה לסכימה הבאה:



כאשר מתקיים:

$$\underline{\tilde{X}} = Q^T \cdot \underline{X}$$
$$\underline{\tilde{Z}} = Q^T \cdot \underline{Z}$$
$$\underline{\tilde{Y}} = Q^T \cdot \underline{Y}$$

<u>הערה</u>:

נשים לב שהטרנספורם הנ"ל לא פגע באילוץ ההספק על הכניסה בבעיה המקורית:

$$\left\|\underline{\tilde{X}}\right\|^2 = \left\|Q^T \underline{X}\right\|^2 = \left(Q^T \underline{X}\right)^T \cdot \left(Q^T \underline{X}\right) = \underline{X}^T \cdot \underbrace{Q \cdot Q^T}_{=I} \cdot \underline{X} = \underline{X}^T \cdot \underline{X} = \left\|\underline{X}\right\|^2$$

$$\Rightarrow E\left\|\underline{\tilde{X}}\right\|^2 = E\left\|\underline{X}\right\|^2 \leq P$$

נבחן את הקווריאנס של $\underline{\tilde{Z}}$:

$$\operatorname{cov}\left(\underline{\tilde{Z}}\right) = \operatorname{cov}\left(Q^T \cdot \underline{Z}\right) = Q^T \operatorname{cov}\left(\underline{Z}\right) Q = Q^T \left(Q \cdot \Lambda \cdot Q^T\right) Q = \Lambda$$

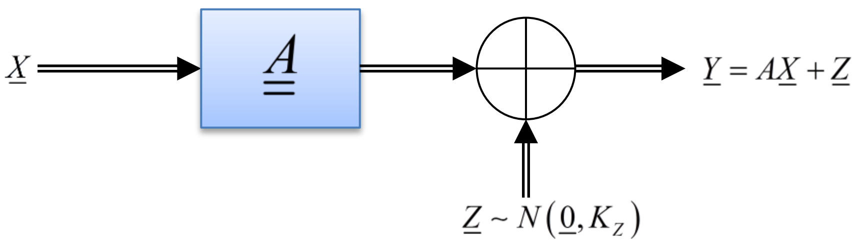
כלומר, רכיבי $\underline{\tilde{Z}}_j$ הם בת"ס!

לכן:

$$C\left(\underline{Y} = \underline{X} + \underline{Z}\right) = C\left(\underline{\tilde{Y}} = \underline{\tilde{X}} + \underline{\tilde{Z}}\right)$$

והקיבול הנ"ל מחושב בעצם ע"י טכניקת "מזיגת המים" על הערוץ השקול (בו מטריצת הקווריאנס היא אלכסונית).

14.3. ערוץ מטריצי



בעצם, ניתן לבטא כל רכיב של וקטור המוצא $\underline{Y}$ בצורה הבאה:

$$Y_j = A_{jj} \cdot X_j + \sum_{i \neq j} A_{ji} \cdot X_i + Z_j$$

<u>הערות</u>:

א. אפשר לראות שאם A לא ריבועית אז ניתן לבצע רדוקציה למטריצה ריבועית עם מימד $rank(A)$, ללא הפסד אינפורמציה הדדית.

ב. לאחר שהבעיה בצורתה המוקטנת כוללת מטריצה ריבועית מדרגה מלאה, ניתן תמיד (ללא הפסד אינפורמציה) לעבור לצורה קאנונית (ערוץ בו הכניסה מוכפלת במטריצת היחידה):

$$\underline{Y} = \underline{X} + \underline{\tilde{Z}}$$

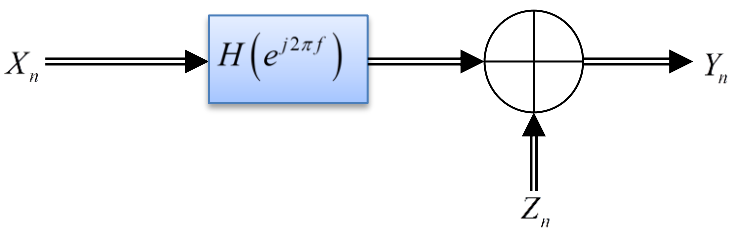
כאשר $\underline{\tilde{Z}}$ הינו "רעש צבעוני" ונתון ע"י:

$$\underline{\tilde{Z}} = A^{-1} \cdot \underline{Z}$$

ולפי הפרק הקודם, ערוץ עם רעש צבעוני ניתן להמיר לערוצים גאוסיים במקביל.

14.4. ערוץ פילטר

לפי מה שהראינו בפרקים הקודמים, עתה נוכל לפתור את הבעיה הבאה:



כאשר המסנן הוא מסנן LTI והרעש הוא תהליך גאוסי סטציונרי.ע"י הפעלת המסנן ההפוך במוצא הערוץ, נמיר את הערוץ עם המסנן לערוץ שקול קאנוני עם רעש צבעוני וללא המסנן. נרשום את הקיבול של הערוץ בתור גבול של קיבול על בלוקים:

$$C = \lim_{k \to \infty} C^{(k)}$$

כאשר:

$$C^{(k)} \triangleq \frac{1}{k} \max I\left(X_1^k ; X_1^k + Z_1^k\right)$$

ניתן להראות כי הסדרה $C^{(k)}$ הינה מונוטונית עולה, ולכן הגבול שהוא הקיבול בהכרח קיים.

הקיבול של הערוץ $X^k \to X^k + Z^k$ תלוי בקווריאנס $K_Z^{(k)}$ של וקטור הרעש Z^k. כפי שראינו בפרקים הקודמים, הקיבול בערוץ זה מתקבל ע״י ״מזיגת מים״ על פני הע״ע של $K_Z^{(k)}$. נוכיח כי עבור $k \to \infty$, הע״ע שואפים להיות דגימות של הספקטרום של התהליך האינסופי הסטציונרי $Z_1, Z_2, \ldots$.

Toeplitz Limit Distribution Theorem .14.4.1

יהא Z_n תהליך סטציונרי עם ספקטרום הספק $S\left(e^{j2\pi f}\right)$. נגדיר: $K_Z^{(k)} = \text{cov}\left(Z_1, \ldots, Z_k\right)$. ויהיו $\lambda_1^{(k)}, \ldots, \lambda_k^{(k)}$ הע״ע של המטריצה הנ״ל. לכל פונקציה $g(\cdot)$ מתקיים:

$$\frac{1}{k}\sum_{j=1}^{k} g\left(\lambda_j^{(k)}\right) \xrightarrow[k\to\infty]{} \int_{-\frac{1}{2}}^{\frac{1}{2}} g\left(S\left(e^{j2\pi f}\right)\right) df$$

כלומר, הע״ע שאופים להיות הדגימות של הספקטרום והו״ע שואפים לפונקציות הרמוניות:

$$K_Z^{(k)} = Q^{(k)} \cdot \Lambda^{(k)} \cdot Q^{(k)^T}$$

נשים לב כי $Q^{(k)}$ הינו פירוק פורייה ו- $\Lambda^{(k)}$ הינו צפיפות ההספק הספקטרלית.

$$C = \lim_{k\to\infty} \frac{1}{k} \max_{\{\}} I\left(X_1, \ldots, X_k; X_1 + Z_1, \ldots, X_k + Z_k\right) = \lim_{k\to\infty} \begin{pmatrix} \text{water pouring} \\ \text{regarding } K_Z^{(k)} \end{pmatrix} = \begin{matrix} \text{water pouring} \\ \text{over the spectrum} \end{matrix}$$

מכאן נובע כי <u>קיבול ערוץ גאוסי עם רעש צבעוני</u> הוא:

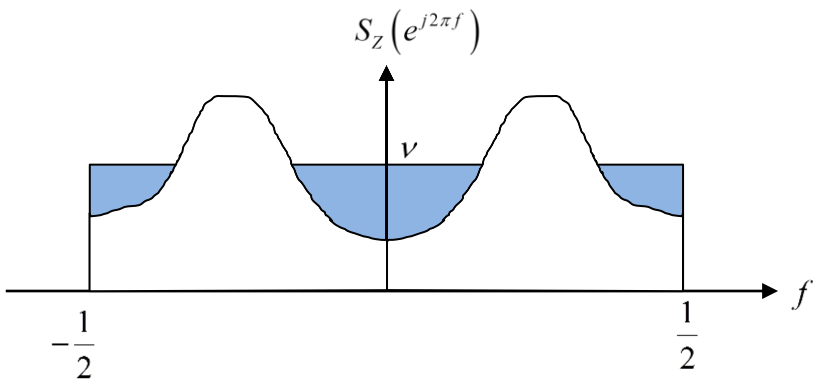
$$C^{(k)} = \frac{1}{k} \sum_{\left\{j | p_j^{opt} > 0\right\}} \frac{1}{2} \log\left(\frac{\nu^{(k)}}{\lambda_j^{(k)}}\right) \xrightarrow[k\to\infty]{} \int_{\left\{f | \nu > S_Z\left(e^{j2\pi f}\right)\right\}} \frac{1}{2} \log\left(\frac{\nu}{S_Z\left(e^{j2\pi f}\right)}\right) df$$

כאשר ν הוא ״גובה המים על פני הספקטרום״. נבחר כך שיתקיים:

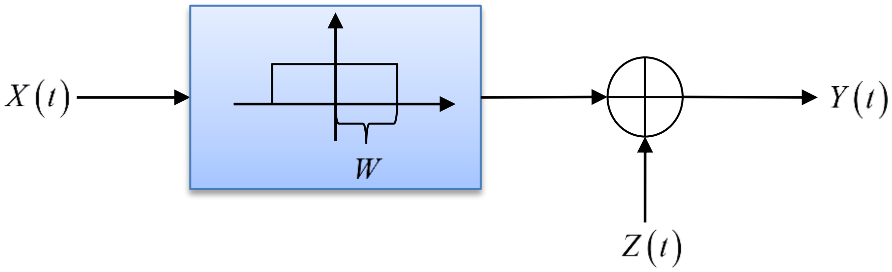
$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\nu - S_Z\left(e^{j2\pi f}\right)\right]^+ df = P$$

נשים לב שהפילוג המגשים הוא תהליך גאוסי סטציונרי עם ספקטרום הספק:

$$S_X\left(e^{j2\pi f}\right) = \left[\nu - S_Z\left(e^{j2\pi f}\right)\right]^+$$



14.5. ערוץ ה-AWGN הרציף בזמן (מוגבל הסרט)



הרעש $Z(t)$ הוא רעש לבן גאוסי בעל צפיפות הספק ספקטרלית $\frac{N_0}{2}$. W הינו רוחב הסרט החד-צדדי של השידור ו- T הינו משך זמן התשדורת. ספר הקוד (אוסף התשדורות האפשריות) הוא:

$$\{X_i(t)\},\ 0 \le t \le T,\ i = 1,\dots,M$$

הקצב R הוא ביחידות ביט לשניה ומוגדר ע"י:

$$R = \frac{\log M}{T}\left[\frac{bit}{\sec}\right]$$

אילוץ הספק/אנרגיה עבור התשדורת ה- i:

$$E_i = \int_0^T X_i(t)\,dt \le P \cdot T$$

14.5.1. פירוק אורתונורמלי באינטרוול הזמן [0,T]

אם $\int_0^T X^2(t)\,dt < \infty$, אזי קיים פירוק מהצורה:

$$X(t) = \sum_{i=0}^{\infty} X_i \cdot \varphi_i(t)$$

כאשר X_i הם מקדמי פירוק כלשהם ו- $\varphi_i(t)$ הן פונקציות בסיס אורתונורמליות, כלומר:

$$\langle \varphi_i(t), \varphi_j(t) \rangle \triangleq \int_0^T \varphi_i(t) \cdot \varphi_j(t)\,dt = \begin{cases} 0 & i \ne j \\ 1 & i = j \end{cases}$$

המקדמים X_i מקיימים:

$$X_i = \int_0^T X(t) \cdot \varphi_i(t)\,dt = \langle X(t), \varphi_i(t) \rangle$$

<u>משפט פרסבל</u>:

$$E = \int_0^T X^2(t)\,dt = \sum_{i=0}^{\infty} X_i^2$$

<u>דוגמאות לבסיסים</u>:

א. <u>פורייה</u>:

$$\varphi_i(t) = \frac{1}{\sqrt{T}} \exp\left\{-j2\pi \frac{i}{T} t\right\}$$

אם $X(t)$ מוגבל סרט $[-W, W]$, אזי מספיקות $n = 2WT + 1$ פונקציות בסיס לפירוק שלו.

ב. <u>דגימה בקצב נייקויסט</u>:

$$X_i \triangleq X\left(i \frac{t}{T}\right),\ \ \varphi_i(t) = \sin c(\ \ )$$

עבור $Z(t)$ שהוא כאמור רעש לבן גאוסי עם צה"ס $\frac{N_0}{2}$ ההיטלים שלו על כל בסיס אורתונורמלי הם iid גאוסיים: $Z_i \sim N\left(0,\frac{N_0}{2}\right)$.

<u>טענה</u>:
אין הפסד אינפורמציה במעבר לפירוק אורתונורמלי, ולכן קיבול הערוץ המוגבל סרט נתון ע"י:

$$C=\frac{1}{T}\max_{\substack{\sum_{i=1}^{n} EX_i^2\le P\cdot T\\ n=2WT+1}} I\left(X_1,...,X_n;Y_1,...,Y_n\right)=\frac{1}{T}n\frac{1}{2}\log\left(1+\frac{P\frac{T}{n}}{\frac{N_0}{2}}\right)$$

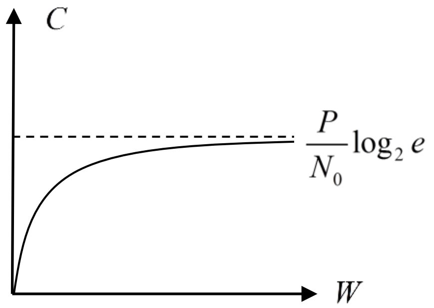
ואם לוקחים בקירוב $n\approx 2WT$ מקבלים:

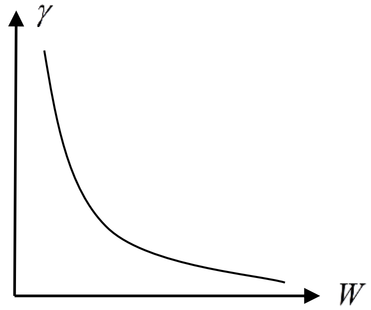
$$\boxed{C=W\log\left(1+\frac{P}{W\cdot N_0}\right)}$$

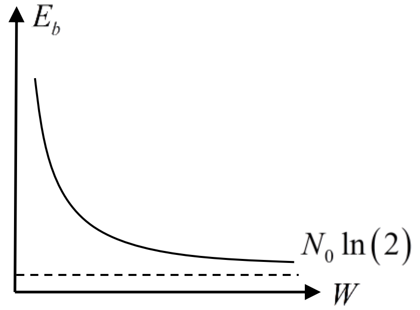
וזהו קיבול ערוץ גאוסי רציף בזמן ומוגבל סרט!
נגדיר 2 גדלים:

γ - <u>נצילות ספקטרלית</u> – מוגדרת ע"י C/W והיחידות הן $\left[bit/\sec\times Hz\right]$.

E_b - <u>נצילות אינפורמציונית</u> – או לחלופין אנרגיה לביט אינפורמציה. מוגדרת ע"י P/C והיחידות הן $\left[\frac{Joul}{bit}\right]$. להלן הגראפים של הגדלים שמצאנו כפונקציה של רוחב הסרט W:







מהגרף האחרון ניתן לראות כי הכמות המינימלית של אנרגיה הדרושה להעביר ביט אינפורמציה בערוץ AWGN רציף בזמן היא $N_0\ln 2$ ולכן:

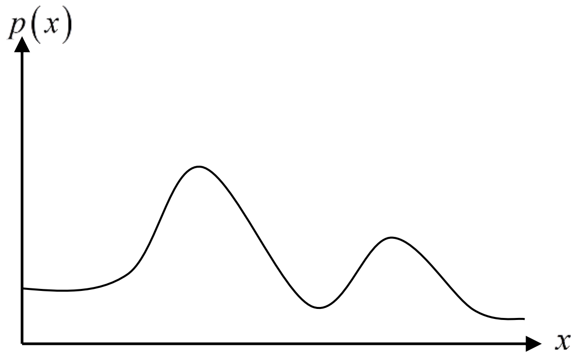
$$\frac{E_b}{N_0}\ge\ln 2=-1.6db$$

החסם הנ"ל ניקרא "חסם שאנון".

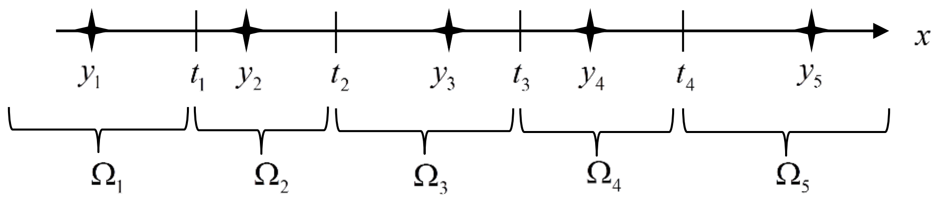
15. דחיסה עם עיוות

15.1. קוונטיזציה סקלרית

נתון מ"א רציף X עם פילוג כלשהו:



נרצה לבצע דיסקרטיזציה של המקור הרציף למקור בדיד כך שהעיוות (כפי שיוגדר בהמשך) יהיה המינימלי לפי מדד כלשהו. ניתן להסתכל על בעיית הדיסקרטיזציה בצורה הבאה - נחלק את ציר ה- x לחלקים:



בעצם, אנו מחלקים את הציר ל- M תחומים זרים $\Omega_1,\dots,\Omega_M$, כאשר הגבולות של התחומים הם ה- t_i וכל ערך של המ"א X ממופה ל- y_i אשר נמצא בתחום של אותו הערך. קוונטייזר טוב הוא קוונטייזר שה- t_i וה- y_i שלו משיגים עיוות מינימלי לפי מדד כלשהו.

<u>הערה</u>: קצב השידור במקרה הנ"ל הוא $R=\log M$.

ננסח את פעולת הקוונטייזר $Q(x)$ מתמטית:

$$\left.\begin{array}{ll} f(x)=i & if\ \ x\in\Omega_i \\ g(i)=y_i=\hat{x}_i & i=1,\dots,M \end{array}\right\}\Rightarrow Q(x)=g\left(f(x)\right)$$

מסמנים את מדד העיוות ב- $d(x,y)$.

דוגמאות ל- $d(x,y)$:

א. MSE: $(y-x)^2$

ב. Rth-power: $|y-x|^r$

מגדירים גודל D המקיים:

$$D\triangleq E\left[d\left(x,Q(x)\right)\right]=\int_{-\infty}^{\infty}p(x)\cdot d\left(x,Q(x)\right)dx$$

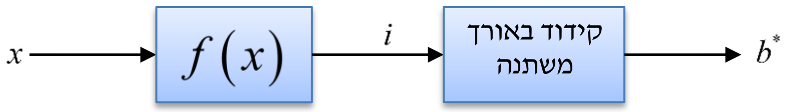
בעיית הקוונטיזציה הקלאסית היא למצוא $y_1,\dots,y_M$ ו- $t_1,\dots,t_{M-1}$ שמביאים למינימום את D עבור M (או R) נתונים.

נציג אלגוריתם למימוש הקוונטייזרים:

אלגוריתם לויד:
זהו אלגוריתם איטרטיבי למציאת הקוונטייזר. הוא מתבצע בצורה הבאה:

א. נקבע t-ים.
ב. נבצע אופטימיזציה על ה- y-ים.
ג. נקבע y-ים.
ד. נבצע אופטימיזציה על ה- t-ים

וחוזר חלילה...
האלגוריתם נעצר כאשר העיוות באינטרציה כלשהי הוא טוב מספיק.
קוונטיזציה עם קידוד אנטרופיה:



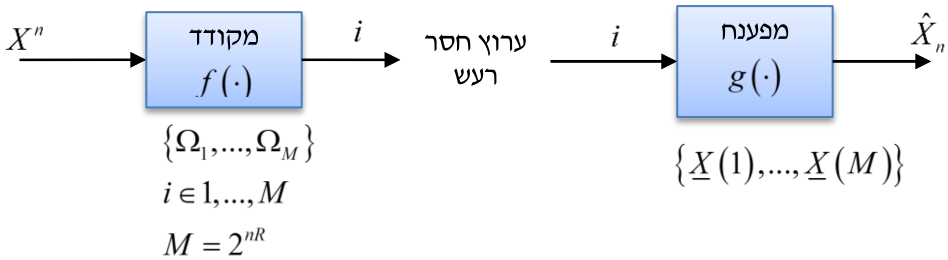
$$R \cong H(index) = H(f(x)) \equiv H(Q(x))$$

הערה: הקוונטייזר האופטימלי בהנחה של קידוד אנטרופיה יהיה בד"כ שונה מקוונטייזר לויד הרגיל.
קוונטיזציה אחידה:
מחלקים את התחום לחלקים שווים. אם גודל כל תא הוא Δ, ניתן להוכיח כי עבור MSE, D הוא המומנט השני של פילוג אחיד בתא Δ, כלומר:

$$D_{MSE} = \frac{\Delta^2}{12}$$

15.2. קוונטיזציה וקטורית

הפעם לא לוקחים סקלר אלא וקטור באורך n כלשהו ועליו מבצעים קוונטיזציה. נתאר את הבעייה כמערכת קידוד מקור. אם נדייק אז כבעיית קידוד בלוק בקצב קבוע:



גם במקרה הזה $\Omega_1,...,\Omega_M$ הינם תחומים זרים. הקצב במקרה הוקטורי הוא: $R = \frac{1}{n}\log M$.

D עבור וקטורים מוגדר בצורה הבאה:

$$D \triangleq E\left[d\left(\underline{X}, Q(\underline{X})\right)\right] \triangleq \frac{1}{n} E\left[\sum_{i=1}^{n} d\left(X_i, Q(X_i)\right)\right]$$

הגדרה:
הזוג (R,D) הוא זוג קצב עיוות בר-השגה אם קיימת סדרה של מקודדים-מפענחים בקצב R - (f_n, g_n) כך ש:

$$\lim_{n\to\infty} E\left\{d\left(\underline{X}, Q_n(\underline{X})\right)\right\} \le D$$

כאשר:

$$Q_n(\underline{X}) = g_n\left(f_n(\underline{X})\right)$$

הקצב הנמוך ביותר שהוא בר-השגה ביחס לעיוות רצוי D יקרא הקצב האופרטיבי. מגדירים פונקציית קצב עיוות אופרטיבית בצורה הבאה:

$$R^{oper}(D) \triangleq \inf\left\{R : (R,D) \text{ - Achievable}\right\}$$

15.3. משפט קידוד המקור עם עיוות של שאנון

נתון מקור $X\sim p(X)$ חסר זיכרון iid. מגדירים קצב אינפורמציוני:

$$R^{\inf}(D)\triangleq \min_{\left\{P(\hat{X}|X):E\left[d(\hat{X},X)\le D\right]\right\}} I(X;\hat{X})$$

כאשר את האילוץ ניתן לכתוב בצורה הבאה:

$$\left\{p(\hat{X}\mid X):\int p(x)p(y\mid x)d(x,y)dxdy\le D\right\}$$

טענה:

$$R^{oper}(D)=R^{\inf}(D)$$

הערוץ (בד"כ יחיד) שמשיג את המינימום $p^{*}(\hat{X}\mid X)$ נקרא "הערוץ המגשים"

דוגמאות:

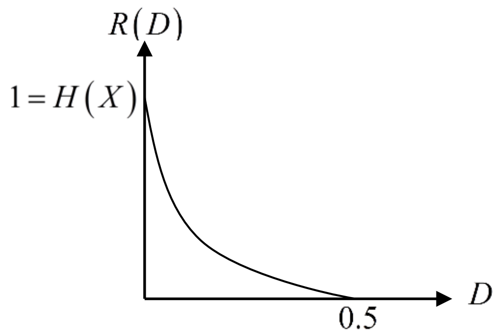
א. מקור ברנולי 0.5:

תחת מדד עיוות המינג:

$$d(x,y)=\begin{cases}0 & x=y\\ 1 & x\neq y\end{cases}\quad,\quad x,y\in\{0,1\}$$

מתקיים:

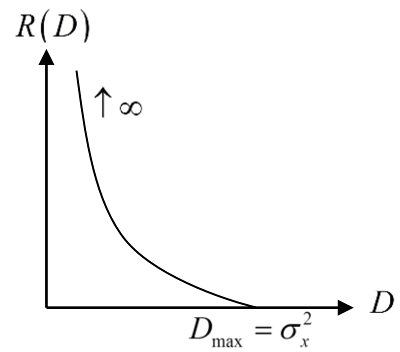
$$R(D)=1-H_B(D),\quad 0\le D\le\frac{1}{2}=D_{\max}$$



ב. מקור גאוסי:

המקור X מפולג $N(0,\sigma_x^2)$. תחת הנחת מדד עיוות ריבועי $d(x,y)=(y-x)^2$ מתקיים:

$$R(D)=\frac{1}{2}\log\left(\frac{\sigma_x^2}{D}\right)$$



הערה: אם מדד העיוות הוא כזה ש- $d(x,y)=0$ מחייב ש- $y=x$, אזי מתקיים $R(D=0)=H(X)$.

15.4. תכונות פונקציית קצב העיוות

א. $R(D)\equiv 0$ עבור $D\geq D_{\max}$.

$D_{\max}$ הוא העיוות הממוצע המתקבל ממיפוי של X לנקודה קבועה y שהיא האופטימלית ביחס למדד העיוות:

$$D_{\max}\triangleq \min_y E\left\{d\left(X,y\right)\right\}$$

דוגמא: במקרה של ממד עיוות ריבועי, ה- y האופטימלי הוא התוחלת ו- $D_{\max}$ הוא השונות.

ב. $R(D)$ מונוטונית לא עולה בתחום $0\leq D\leq D_{\max}$.

הוכחה: התחום עליו עושים מינימיזציה גדל עם D ולכן המינימום בהכרח לא עולה.

ג. $R(D)$ היא פונקציה קמורה ביחס ל- D:

$$R\left(\lambda D_1+(1-\lambda)D_2\right)\leq \lambda R\left(D_1\right)+(1-\lambda)R\left(D_2\right)$$

הוכחה:

נגדיר:

$$p_\lambda\left(\hat{X}\mid X\right)=\lambda p_1^*\left(\hat{X}\mid X\right)+(1-\lambda)\,p_2^*\left(\hat{X}\mid X\right)$$

כאשר $p_1^*\left(\hat{X}\mid X\right)$ ו- $p_2^*\left(\hat{X}\mid X\right)$ הם ערוצים המגשימים עבור D_1 ו- D_2 בהתאמה. קל לוודא כי:

$$\int p(x)\,p_\lambda\left(\hat{x}\mid x\right)d\left(x,\hat{x}\right)dxd\hat{x}=\lambda D_1+(1-\lambda)D_2\triangleq D$$

מצד שני, כיוון שהאינפורמציה ההדדית קמורה ביחס לפילוג המעבר:

$$I\left(p(X),p_\lambda\left(\hat{X}\mid X\right)\right)\leq \lambda\cdot\underbrace{I\left(X;\hat{X}_1\right)}_{=R(D_1)}+(1-\lambda)\cdot\underbrace{I\left(X;\hat{X}_2\right)}_{=R(D_2)}$$

מהגדרת פונקציית קצב העיוות נובע כי:

$$R(D)\leq I\left(p(X),p_\lambda\left(\hat{X}\mid X\right)\right)$$

ולכן:

$$R(D)\leq \lambda R\left(D_1\right)+(1-\lambda)R\left(D_2\right)$$

15.5. החסם התחתון של שאנון

נניח שמדד העיוות הוא הפרשי, כלומר: $d(x,y)=func(y-x)$. נגדיר משתנה "מקסימום אנטרופיה" ביחס ל- $d(\cdot)$:

$$H_{\max}(D)=\max_{\{E[d(u)\leq D]\}}H(u)\qquad \text{discrete}$$
$$h_{\max}(D)=\max_{\{E[d(u)\leq D]\}}h(u)\qquad \text{conitinious}$$

מתקיים:

$$R(D)\geq R_{SLB}(D)\triangleq\begin{cases}H(X)-H_{\max}(D) & \text{discrete}\\ h(X)-h_{\max}(D) & \text{continious}\end{cases}$$

<u>הוכחה למקרה של MSE</u>:

$$E\left(\hat{X}-X\right)^2 \le D$$

$$I\left(X;\hat{X}\right)=h\left(X\right)-h\left(X\mid\hat{X}\right)\underset{(1)}{=}h\left(X\right)-h\left(X-\hat{X}\mid\hat{X}\right)\ge$$

$$\underset{(2)}{\ge}h\left(X\right)-h\left(X-\hat{X}\right)\underset{(3)}{\ge}h\left(X\right)-h_{\max}\left(D\right)\underset{(4)}{=}h\left(X\right)-\frac{1}{2}\log\left(2\pi eD\right)$$

הסברים למעברים:

(1) הזזה בקבוע לא משנה אנטרופיה.

(2) התניה מורידה אנטרופיה.

(3) הגדרה של $h_{\max}$.

(4) משתנה גאוסי הוא בעל האנטרופיה המקסימלי עבור מומנט שני נתון.

שוויון ב-SLB מתקיים אמ"מ ניתן לפרק את X בת"ס הבא: $X=\hat{X}+U^*$ כאשר U^* הוא משתנה מקסימום אנטרופיה (גאוסי עבור המקרה של MSE). בצורה זו מתקבל "ערוץ גאוסי הפוך" (X מופק מ-$\hat{X}$).

<u>הגדרה</u>:

<u>הספק אנטרופיה</u> זוהי שונות של מ"א גאוסי (ו"א או ת"א גאוסי לבן) עם אותה האנטרופיה כמו של המשתנה (ו"א, ת"א) הנתון. מסמנים את הספק האנטרופיה ב-$p_E\left(X\right)$.

עבור סקלר:

$$p_E\left(X\right)=\frac{2^{2h(X)}}{2\pi e}$$

עבור וקטור:

$$p_E\left(\underline{X}\right)=\frac{2^{\frac{2}{n}h(\underline{X})}}{2\pi e}$$

מההגדרה הנ"ל ניתן לרשום:

$$R_{SLB}\left(D\right)=\frac{1}{2}\log\left(\frac{p_E\left(X\right)}{D}\right)$$

הביטוי הנ"ל נראה כמו פונקציית קצב העיוות הגאוסית רק שבמקום σ_x^2 מופיע $p_E\left(X\right)$. נשאלת השאלה האם/מתי ניתן להשיג את ה-SLB:

א. עבור <u>מדד עיוות ריבועי</u>, אם $X\sim N\left(0,\sigma_x^2\right)$ ו- $0\le D\le\sigma_x^2$, אזי הערוץ ההפוך קיים:

$$\underset{N\left(0,\sigma_x^2\right)}{X}=\underset{N\left(0,\sigma_x^2-D\right)}{\hat{X}}+\underset{N(0,D)}{U^*}$$

כאשר U^* בת"ס ב- $\hat{X}$, ובמקרה זה:

$$R\left(D\right)=R_{SLB}\left(D\right)=\frac{1}{2}\log\left(\frac{\sigma_x^2}{D}\right)$$

ב. עבור <u>מדד האמינג</u>: קל להוכיח שעבור מקור ברנולי p מתקיים:

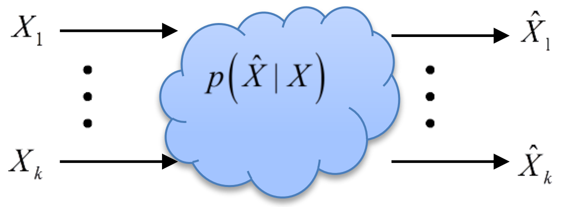
$$R\left(D\right)=R_{SLB}\left(D\right)=H_B\left(p\right)-H_B\left(D\right)$$

$$0\le D\le p=D_{\max}$$

גם במקרה זה מתקיים הערוץ ההפוך.

15.6. כלל מזיגת המים למקורות

נבחן בעיה של מקורות גאוסיים במקביל:



בסכימה הנ"ל מתקיים:

$$X_i \sim N\left(0,\sigma_i^2\right)$$

$$\sum_{i=1}^{k} E\left(\hat{X}_i - X_i\right)^2 \le D$$

נבחן את פונקציית קצב העיוות:

$$R(D) = \min_{\sum_{i=1}^{k} E\left(\hat{X}_i - X_i\right)^2 \le D} I\left(X_1,...,X_k;\hat{X}_1,...,\hat{X}_k\right) \underset{(1)}{\ge} \min_{\sum_{i=1}^{k} E\left(\hat{X}_i - X_i\right)^2 \le D} \sum_{i=1}^{k} I\left(X_i;\hat{X}_i\right) \ge$$

$$\underset{(2)}{\ge} \min_{\left\{D_1,...,D_k \mid \sum D_i \le D\right\}} \sum_{i=1}^{k} R\left(D_i\right) \underset{(3)}{=} \min_{\left\{D_1,...,D_k \mid \sum D_i \le D\right\}} \sum_{i=1}^{k} \left[\frac{1}{2}\log\left(\frac{\sigma_i^2}{D_i}\right)\right]^+$$

הסברים למעברים:

(1) עבור כניסות בת"ס, זיכרון בערוץ יכול רק להגדיל את האינפורמציה ההדדית.

(2) הגדרת $R\left(D_i\right)$

(3) פונקציית קצב העיוות הגאוסית.

לבעייה שהגענו אליה בפיתוח קוראים "בעיית הקצאת העיוותים"

15.6.1. הקצאת עיוותים אופטימלית

<u>הערה</u>: חסם שאנון לבעייה של מקורות גאוסיים במקביל הוא:

$$R(D) \ge R_{SLB}(D) = \frac{1}{2}\log\frac{p_E\left(\underline{X}\right)}{D/k} = \frac{1}{2}\log\frac{\sqrt[k]{\sigma_1^2 \cdot ... \cdot \sigma_k^2}}{D/k}$$

והוא מושג בשוויון כאשר: $D/k \le \min \sigma_i^2$. במקרה זה קיים הערוץ ההפוך ואז מתקיים:

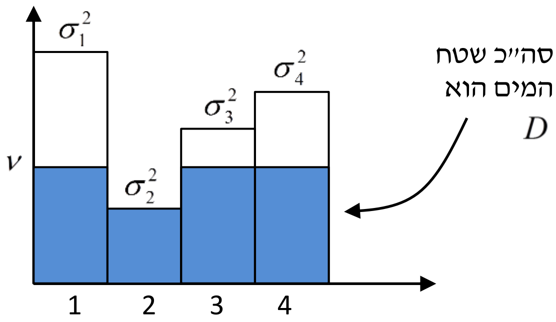
רעש גאוסי לבן+"משהו"=מקור וקטורי

<u>הסבר</u>: במקרה שה-SLB מושג בשוויון, אפשר לקבל את הוקטור $X_1,...,X_k$ כסכום של "משהו" ועוד רעש גאוסי לבן עם הספק D/k. מכאן נובעת הקצאת עיוותים אחידה $D_i = D/k$.

במקרה הכללי (שחסם ה-SLB איננו הדוק) אז ע"י כופלי לגרנז', גזירה ותנאי קון-תאקר מקבלים כלל מזיגת מים:

$$D = \sum_{i=1}^{k} D_i = \sum_{i=1}^{k} \min\left\{\nu,\sigma_i^2\right\} = \sum_{\substack{\text{unencoded}\\\text{sources}}} \sigma_i^2 + \nu \cdot \left(\begin{matrix}\text{amount of}\\\text{encoded sources}\end{matrix}\right)$$

$$R = \sum_{i=1}^{k}\left[\frac{1}{2}\log\frac{\sigma_i^2}{\nu}\right]^+ = \sum_{\substack{\text{encoded}\\\text{sources}}} \frac{1}{2}\log\frac{\sigma_i^2}{\nu}$$



נשים לב שאם גובה המים ν הוא <u>מתחת</u> המקור החלש ביותר (כלומר, כל המקורות הם מקודדים) אזי מתקיים התנאי לכך שחסם ה-SLB הוא הדוק.

15.7. מקור גאוסי סטציונרי בדיד בזמן

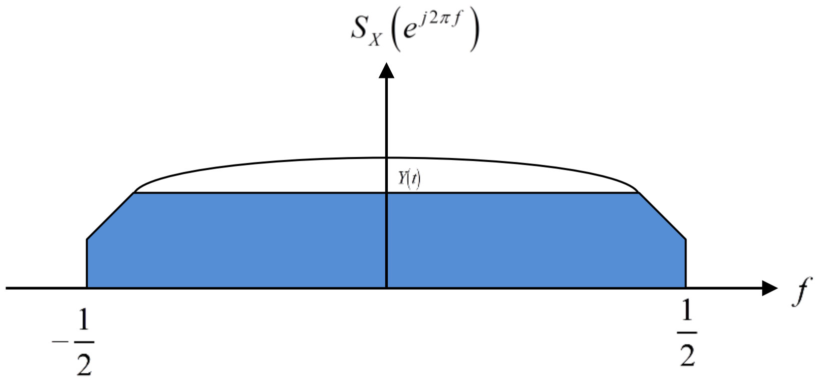
נבחן את פונקציית קצב העיוות ע״י מדד עיוות ריבועי (MSE):

$$\bar{R}(D) \triangleq \lim_{n\to\infty} \frac{1}{n} \min_{\|\underline{X}-\hat{\underline{X}}\|^2 \le nD} I\left(\underline{X};\hat{\underline{X}}\right)$$

הגבול הנ״ל קיים הודות לסטציונריות. בנוסף התהליך $X_1, X_2, \ldots$ הינו גאוסי עם פונקציית אוטוקורלציה $R_X(k)$ וספקטרום $S_X\left(e^{j2\pi f}\right) = \mathcal{F}\left\{R_X(k)\right\}$. פונקציית קצב העיוות נתונה ע״י פיתרון פרמטרי (מזיגת מים ביחס לספקטרום):

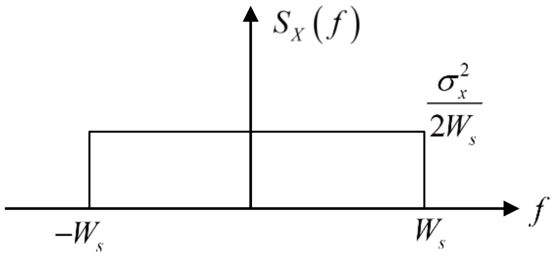
$$D = \int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\nu, S_X\left(e^{j2\pi f}\right)\right\} df$$

$$R = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\frac{1}{2}\log\frac{S_X\left(e^{j2\pi f}\right)}{\nu}\right]^+ df$$



15.8. מקור גאוסי רציף בזמן, לבן ומוגבל סרט

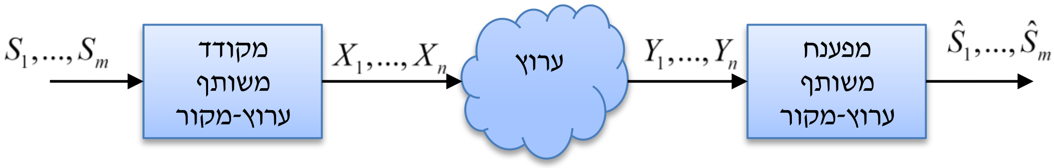
נתון מקור גאוסי X שהינו כאמור רציף בזמן, לבן ומוגבל סרט. הספקטרום של התהליך נראה כך:



פונקציית קצב העיוות למקור הנ"ל היא:

$$R(D)=\frac{2W_s\cdot T}{T}\cdot\frac{1}{2}\log\frac{\sigma_x^2/2W_s}{D/2W_s}=W_s\log\frac{\sigma_x^2}{D}$$

16. קידוד משותף מקור-ערוץ עם עיוות



נניח שזמן השידור הוא T (בזמן זה ישנן m דגימות מקור ו- n שימושי ערוץ).

נסמן ב- $D_{\min}$ את העיוות המינימלי האפשרי עבור כל הקידודים האפשריים:

$$D_{\min}\triangleq\min_{\substack{\text{all possible}\\ \text{systems}}}\{\text{Distortion}\}$$

עבור פתרון ספרתי לבעיה (שימוש בעיקרון ההפרדה: קידוד מקור וקידוד ערוץ) מתקיים:

$$R(D)\left[\frac{bit}{sample}\right]\cdot\frac{m}{T}\le R\left[\frac{bit}{\sec}\right]\le C\left[\frac{bit}{channel\,use}\right]\cdot\frac{n}{T}$$

הדבר אפשרי רק אם: $R(D)\le\frac{n}{m}\cdot C$.

כיוון ש- $R(D)$ היא פונקציה מונוטונית יורדת, מתקיים כי:

$$D_{\min}=R^{-1}\left(\frac{n}{m}\cdot C\right)$$

ניתן להגדיר את הבעייה הנ"ל כשרשרת מרקוב:

$$\underline{S}\leftrightarrow\underline{X}\leftrightarrow\underline{Y}\leftrightarrow\underline{\hat{S}}$$

עבורה מתקיים אי-שוויון עיבוד הנתונים:

$$I\left(\underline{S};\underline{\hat{S}}\right)\le I\left(\underline{X};\underline{Y}\right)$$

מצד שני:

$$I\left(\underline{X};\underline{Y}\right)\le n\cdot C$$

$$m\cdot R(D)\le I\left(\underline{S};\underline{\hat{S}}\right)$$

מכאן נובע כי לכל מערכת שידור-קליטה (ספרתית או אנאלוגית) מתקיים:

$$R(D)\le \frac{n}{m}\cdot C$$

ולכן $D_{\min}$ של המערכת הספרתית הוא חסם לכל מערכת תקשורת!

16.1. קידוד משותף של מקור גאוסי דרך ערוץ גאוסי

הצבה במשפט ה-J.S.C.C של מקור גאוסי לבן ברוחב סרט W_s משודר בערוץ ברוחב סרט W_c:

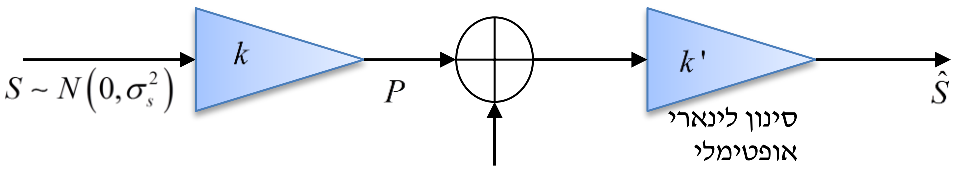
$$r\triangleq \frac{W_c}{W_s}\leftarrow \text{uses to samples ratio}$$

$$C=W_c\cdot\log\left(1+\underbrace{\frac{P}{W_C\cdot N_0}}_{=SNR}\right)\left[\frac{bit}{\sec}\right]$$

$$R(D)=W_s\cdot\log\left(\frac{\sigma_s^2}{D}\right)\left[\frac{bit}{\sec}\right]$$

$$\Rightarrow D_{\min}=\frac{\sigma_s^2}{(1+SNR)^{\frac{W_c}{W_s}}}$$

במקרה בו $W_c=W_s$ אזי ניתן לשדר את האות גם בצורה אנאלוגית ולהשיג את אותו $D_{\min}$ כמו הפתרון הספרתי (בעיקרון ההפרדה).



17. מקורות רציפים בעלי זיכרון

נתון ת"א $\ldots, X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots$ נגדיר שני גדלים:

$$\overline{h}^{(I)} \triangleq \lim_{n\to\infty} \frac{1}{n} h\left(X_1,\ldots,X_n\right)$$

$$\overline{h}^{(II)} \triangleq \lim_{n\to\infty} h\left(X_n \mid X_{n-1},\ldots,X_1\right)$$

<u>**טענה**</u>: אם X_n הוא תהליך סטציונרי אזי הגבולות קיימים ומתקיים:

$$\overline{h}^{(I)} = \overline{h}^{(II)} \triangleq \overline{h}$$

נבחן את המקרה הגאוסי הסטציונרי:

$$EX_n = \mu,\ R_x\left(k\right) = EX_nX_{n+k},\ S_X\left(e^{j2\pi f}\right) = \mathcal{F}\left\{R_X\left(k\right)\right\}$$

<u>תזכורת</u>: אם $\mu = 0$ אזי מתקיים:

$$\sigma_x^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_X\left(e^{j2\pi f}\right) df$$

<u>**טענה**</u>: עבור ת"א גאוסי סטציונרי, נפעיל את ה-Toeplitz limit distribution theorem על $\overline{h}^{(I)}$ ונקבל:

$$\overline{h} = \exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{2}\log\left(2\pi e \cdot S_X\left(e^{j2\pi f}\right)\right) df\right\}$$

17.1.1. קישור להספק אנטרופיה

($p_E\left(X_1^\infty\right)$ היא השונות של ת"א גאוסי <u>לבן</u> עם אותו קצב אנטרופיה כמו של X_1^∞)

מתקיים:

$$\overline{h} = \frac{1}{2}\log\left\{2\pi e \cdot p_E\left(X_1^\infty\right)\right\}$$

נשים לב כי:

$$p_E\left(X_1^\infty\right) = \exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \log\left(S_X\left(e^{j2\pi f}\right)\right) df\right\}$$

לגודל זה קוראים <u>"הממוצע הגיאומטרי של הספקטרום"</u>.

17.1.2. קישור עם חיזוי ליניארי

חזאי ליניארי מסדר k שמשיג את ה- LMMSE הוא בעל הצורה הבאה:

$$\hat{X}_n = \sum_{i=1}^{k} a_i \cdot X_{n-i}$$

עבור ת"א גאוסי:

$$X_n \mid_{X_{n-1},\ldots,X_{n-k}} \sim N\left(\hat{X}_n, LMMSE_k\right)$$

עבור חזאי מסדר ∞:

$$h\left(X_0 \mid X_{-1}, X_{-2},\ldots\right)=\frac{1}{2}\log\left(2\pi e\cdot LMMSE_{\infty}\right)$$
$$\Rightarrow LMMSE_{\infty} = p_E\left(X_1^{\infty}\right)$$

שזהו בעצם הממוצע הגיאומטרי של הספקטרום.

17.1.3. קישור למשפט הגבול המרכזי והספק אנטרופיה

<u>אי-שוויון הספק אנטרופיה (EPI)</u>:

אם X ו- Y הם בת"ס, אזי מתקיים:

$$p_E\left(X+Y\right)\geq p_E\left(X\right)+p_E\left(Y\right)$$

ושוויון אמ"מ X ו- Y הם גאוסיים.

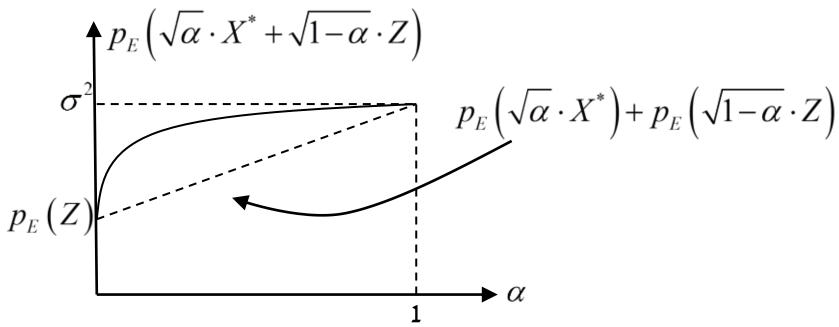
<u>פרשנות</u>: סכום של מ"א לא גאוסיים הופך להיות "יותר גאוסי" (נשים לב כי מבחינת שונות מתקיים:

$$\mathrm{var}\left(X+Y\right)=\mathrm{var}\left(X\right)+\mathrm{var}\left(Y\right)$$

<u>דוגמא</u>:

נבחן את הסכום: $\sqrt{\alpha}\cdot X^{*}+\sqrt{1-\alpha}\cdot Z$, כאשר X^{*} הינו גאוסי עם שונות σ^2, Z מ"א לא גאוסי עם שונות σ^2 ו- α הינו סקלר המקיים: $0\leq\alpha\leq 1$. מתקיים:

$$\mathrm{var}\left(\sqrt{\alpha}\cdot X^{*}+\sqrt{1-\alpha}\cdot Z\right)=\sigma^2 \quad \forall\alpha$$



מקווה שהסיכום היה ברור ומועיל.

בהצלחה במבחן!