EE728
Προχωρημένα Θέματα Θεωρίας Πληροφορίας
7η διάλεξη
(2η έκδοση, 1/4/2013)

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

26 Μαρτίου 2013

Περιεχόμενα 7ης διάλεξης

$I(X^n; Y^n) \leq nC$

Θα αποδείξουμε, κατ' αρχάς, ότι, για Διακριτά Κανάλια Χωρίς Μνήμη, η πληροφοριακή χωρητικότητα ανά χρήση του καναλιού δεν αυξάνει εάν το κανάλι χρησιμοποιηθεί ως κανάλι γινομένου. Δηλαδή, $I(X^n; Y^n) \leq nC$ για οποιαδήποτε $p(x)$, όπου $C = \max_{p(x)} I(X; Y)$.

$$I(X^n; Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1, \ldots, Y_{i-1}, X^n) =$$

$$\overset{(a)}{=} H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \overset{(b)}{\leq} \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i)$$

$$= \sum_{i=1}^{n} I(X_i; Y_i) \leq nC.$$

(a) Το κανάλι δεν έχει μνήμη και δε χρησιμοποιείται ανάδραση. (b) Η από κοινού εντροπία δεν υπερβαίνει το άθροισμα των εντροπιών.

Ανισότητα Fano

- Για την απόδειξη του αντιστρόφου του Θεωρήματος Κωδικοποίησης Καναλιού θα χρησιμοποιήσουμε την Ανισότητα Fano.
- Είδαμε ότι, για κάθε εκτιμητή $\hat{X} = g(Y)$,

$$H(X|Y) \leq H(X|\hat{X}) \leq H(P_e) + P_e \log |\mathcal{X}| \Rightarrow H(X|\hat{X}) \leq 1 + P_e \log |\mathcal{X}|,$$

όπου $P_e = \Pr\{\hat{X} \neq X\}$.

- Εάν θεωρήσουμε Διακριτό Κανάλι Χωρίς Μνήμη με βιβλίο κωδίκων $\mathcal{C}$ και ομοιόμορφα κατανεμημένα μηνύματα M,

$$H(M|\hat{M}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{M \neq \hat{M}\}.$$

Θεώρημα Κωδικοποίησης Καναλιού – Απόδειξη ασθενούς αντιστρόφου

- Θα δείξουμε ότι, για κάθε κώδικα $(2^{nR}, n)$ με $\lambda^{(n)} \to 0$, πρέπει να ισχύει $R \leq C$. Δεδομένου ότι $\lambda^{(n)} \to 0$ και η μέση πιθανότητα σφάλματος $P_e^{(n)} \to 0$.
- Έστω ότι ο δέκτης αποφασίζει ποια ακολουθία μεταδόθηκε με βάση κάποια συνάρτηση αποκωδικοποίησης $\hat{M} = g(Y^n)$. Ισχύει $M \to X^n(M) \to Y^n \to \hat{M}$.
- Έστω, επίσης, ότι το μήνυμα που στέλνεται στο κανάλι επιλέγεται με βάση ομοιόμορφη κατανομή στο σύνολο των πιθανών μηνυμάτων $\{1, 2, \ldots, 2^{nR}\}$. Επομένως, $\Pr\{\hat{M} \neq M\} = P_e^{(n)} = \frac{1}{2^{nR}} \sum_i \lambda_i$.

Θεώρημα Κωδικοποίησης Καναλιού – Απόδειξη ασθενούς αντιστρόφου (2)

- Συνεπώς,

$$nR \overset{(a)}{=} H(M) \overset{(b)}{=} I(M;\hat{M}) + H(M|\hat{M}) \overset{(c)}{\leq} I(M;\hat{M}) + 1 + P_e^{(n)} nR$$
$$\overset{(d)}{\leq} I(X^n;Y^n) + 1 + P_e^{(n)} nR \overset{(e)}{\leq} nC + 1 + P_e^{(n)} nR.$$

(a) M ομοιόμορφη τ.μ., (b) σχέση αμοιβαίας πληροφορίας – εντροπίας, (c) ανισότητα Fano, (d) ανισότητα επεξεργασίας δεδομένων, (e) $I(X^n;Y^n) \leq nC$.

Θεώρημα Κωδικοποίησης Καναλιού
Απόδειξη ασθενούς αντιστρόφου (3)

$$nR \leq 1 + P_e^{(n)} nR + nC \Rightarrow R \leq P_e^{(n)} R + \frac{1}{n} + C.$$

- Από την υπόθεση ότι $\lambda^{(n)} \to 0, P_e^{(n)} R \to 0$ για $n \to \infty$. Επομένως, για $n \to \infty$,
$$R < C.$$
- Λύνοντας ως προς $P_e^{(n)}$, $P_e^{(n)} \geq 1 - \frac{C}{R} - \frac{1}{nR}$. Συνεπώς, εάν $R > C$, $P_e^{(n)} > 0$ για $n \to \infty$.

Θεώρημα Κωδικοποίησης Καναλιού
Απόδειξη ασθενούς αντιστρόφου (4)

- Το αποτέλεσμα αυτό ονομάζεται ασθενές αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού. Αποδεικνύεται (ισχυρό αντίστροφο) ότι, εάν $R > C$, $P_e^{(n)} \to 1$ εκθετικά.
- Συνεπώς, η χωρητικότητα καναλιού C αποτελεί μια πολύ σαφή διαχωριστική γραμμή: Όταν $R < C$ η πιθανότητα σφάλματος τείνει εκθετικά στο 0. Αντίθετα, όταν $R > C$, η πιθανότητα σφάλματος τείνει εκθετικά στο 1.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου

- Θα αποδείξουμε ξανά το αντίστροφο με μία μικρή παραλλαγή στη χρήση της ανισότητας Fano.
- Η απόδειξη αυτή είναι πιο γενική. Όπως θα δούμε σύντομα, μπορεί να εφαρμοστεί και στην περίπτωση που χρησιμοποιείται ανάδραση.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (2)

- Επειδή $M \to X^n(M) \to Y^n \to \hat{M}$,

$$H(M|Y^n) \leq H(M|\hat{M}).$$

- Επίσης, από την ανισότητα Fano,

$$H(M|\hat{M}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{M \neq \hat{M}\}.$$

- Υποθέτοντας, και πάλι, ότι τα μηνύματα M ακολουθούν ομοιόμορφη κατανομή,

$$\begin{aligned} nR = H(M) &\overset{(a)}{=} I(M; Y^n) + H(M|Y^n) \\ &\overset{(b)}{\leq} I(M; Y^n) + 1 + P_e^{(n)} nR \end{aligned}$$

(a) Σχέση εντροπίας-αμοιβαίας πληροφορίας, (b) ανισότητα Fano.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (3)

$$
\begin{aligned}
nR &\leq I(M; Y^n) + 1 + P_e^{(n)} nR \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(M; Y_i | Y^{i-1}) + 1 + P_e^{(n)} nR \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} I(M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} I(X_i, M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR
\end{aligned}
$$

(c) κανόνας αλυσίδας, (d)
$I(M, Y^{i-1}; Y_i) = I(Y^{i-1}; Y_i) + I(M; Y_i | Y^{i-1})$, (e) $X_i = f(M, Y^{i-1})$
(ισχύει ακόμα και όταν χρησιμοποιείται ανάδραση).

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (4)

$$
\begin{aligned}
nR &\leq \sum_{i=1}^{n} I(X_i, M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_i) + 1 + P_e^{(n)} nR \\
&\leq nC + 1 + P_e^{(n)} nR = n\left(C + \frac{1}{n} + P_e^{(n)}\right)
\end{aligned}
$$

(f) $X_i = f(M, Y^{i-1})$ και το κανάλι δεν έχει μνήμη, οπότε $(M, Y^{i-1}) \to X_i \to Y_i$.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (5)

- Επομένως, για $n \to \infty$ και $P_e^{(n)} \to 0$,

$$nR \leq n\left(C + \epsilon_n\right) \Rightarrow R < C.$$

- Παρατηρήστε ότι δεν απαγορέψαμε τη χρήση ανάδρασης στον κώδικα (περισσότερα σύντομα). Αυτό σημαίνει ότι, σε ένα κανάλι χωρίς μνήμη, ακόμα και αν μπορούμε να χρησιμοποιήσουμε ανάδραση, η χωρητικότητα δεν αυξάνει (περισσότερα σύντομα).

Γιατί χρησιμοποιούμε $M \sim \text{Unif}\left[0, 2^{nR} - 1\right]$;

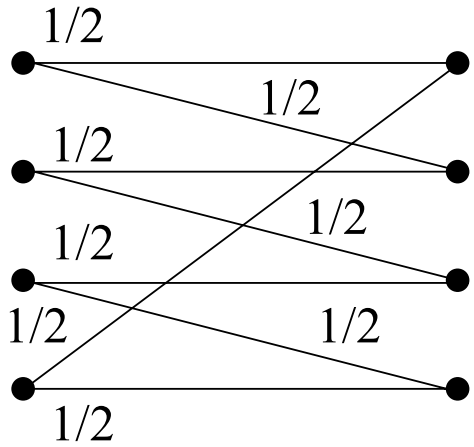
- Στην απόδειξη του ασθενούς αντιστρόφου με χρήση της ανισότητας Fano υποθέσαμε ότι $M \sim \text{Unif}\left[0, 2^{nR} - 1\right]$.
- Μήπως αυτό σημαίνει ότι το αντίστροφο ισχύει μόνο για κώδικες όπου όλες οι κωδικές λέξεις είναι ισοπίθανες;
- Όχι (αυτό είναι ένα λεπτό σημείο). Θυμηθείτε ότι η χωρητικότητα ισούται με το μέγιστο εφικτό ρυθμό μετάδοσης. Ο ρυθμός μετάδοσης ισούται με $\log M/n$.
- Δηλαδή, για να δείξουμε ότι ο ρυθμός μετάδοσης ενός κώδικα είναι R αρκεί να αποδείξουμε ότι υπάρχει κώδικας με 2^{nR} κωδικές λέξεις τον οποίο μπορούμε να χρησιμοποιήσουμε και να επιτύχουμε $P_e^{(n)} \to 0$.
- Για να δείξουμε ότι η χωρητικότητα ενός καναλιού είναι C πρέπει να δείξουμε ότι δεν υπάρχει κώδικας μήκους n με περισσότερες από 2^{nC} κωδικές λέξεις.

Γιατί χρησιμοποιούμε $M \sim \text{Unif}\left[0, 2^{nR} - 1\right]$; (2)

- Η πιθανότητα με την οποία ο χρήστης του κώδικα στέλνει κάθε κωδική λέξη δε μας αφορά (τουλάχιστον όσον αφορά την απόδειξη του αντιστρόφου). Εμείς θέλουμε μόνο να κατασκευάσουμε 2^{nR} κωδικές λέξεις τις οποίες να μπορεί να διακρίνει ο δέκτης με αυθαίρετα μικρή πιθανότητα σφάλματος.
- Η επιλογή $M \sim \text{Unif}\left[0, 2^{nR} - 1\right]$ γίνεται απλώς και μόνο γιατί μας βολεύει στην απόδειξη του αντιστρόφου.

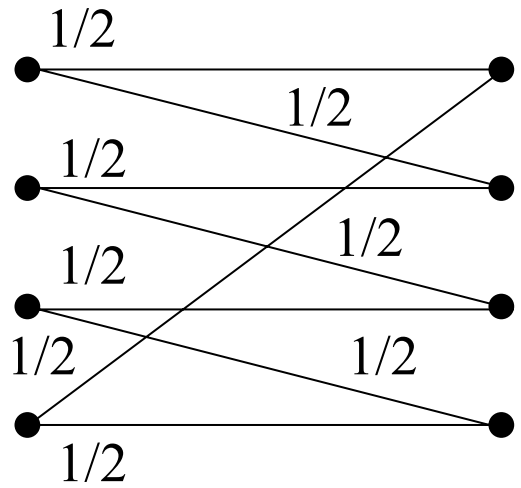
Χωρητικότητα καναλιών με ανάδραση

Παράδειγμα 7.1



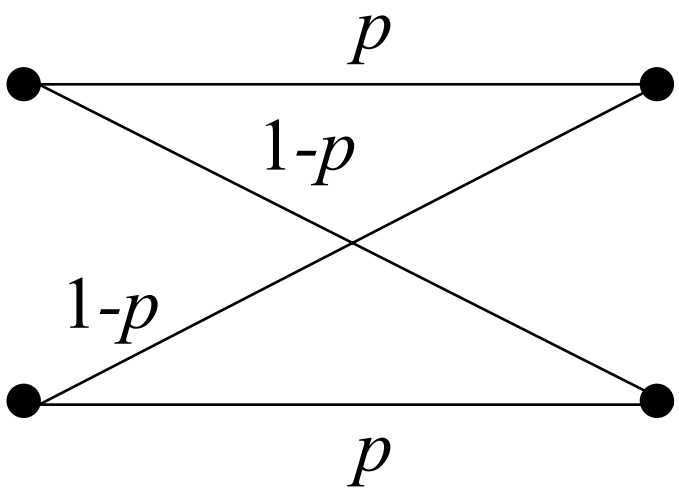
- Θεωρούμε το διακριτό κανάλι χωρίς μνήμη του σχήματος ("ενθόρυβη γραφομηχανή").
- Η χωρητικότητα του καναλιού ισούται με $C = \max I(X;Y) = \max\{H(Y) - H(Y|X)\} = 2 - 1 = 1$ bit.
- Μπορούμε να επιτύχουμε μετάδοση με ρυθμό ίσο με τη χωρητικότητα και με μηδενική πιθανότητα σφάλματος χρησιμοποιώντας π.χ. τις εισόδους 0 και 2. Προφανώς, $R = 1$ bit $= C$.

Παράδειγμα 7.1 (συνέχεια)



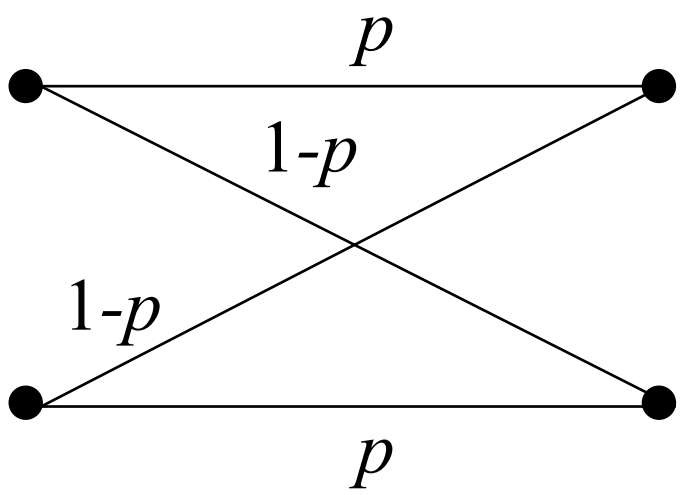
- Ό,τι και να συμβεί στο κανάλι είμαστε βέβαιοι ότι δε θα εμφανιστεί σφάλμα αποκωδικοποίησης.
- Εάν μπορούσαμε να χρησιμοποιήσουμε ανάδραση (feedback), η χωρητικότητα θα παρέμενε η ίδια;

Παράδειγμα 7.2



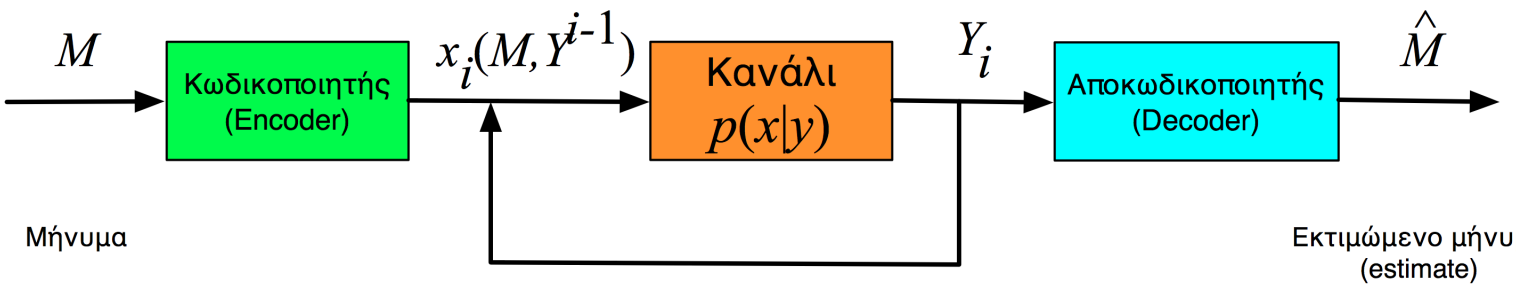
- Ας θεωρήσουμε, τώρα, το δυαδικό συμμετρικό κανάλι.
- Γνωρίζουμε ότι $C = 1 - H(p)$ και ότι η χωρητικότητα επιτυγχάνεται χρησιμοποιώντας και τα δύο μηνύματα με ίση πιθανότητα. Επομένως, κάθε φορά που στέλνουμε ένα από τα δύο μηνύματα στο κανάλι δε γνωρίζουμε εάν το μήνυμα μεταδόθηκε επιτυχώς. Η πιθανότητα σφάλματος ανά μετάδοση είναι μη μηδενική.

Παράδειγμα 7.2 (συνέχεια)



- Τι συμβαίνει εάν χρησιμοποιήσουμε ανάδραση; (όπου γνωρίζουμε εάν έχει εμφανιστεί σφάλμα στο δέκτη;)
- Σημείωση: Όταν χρησιμοποιούμε ανάδραση στο BSC, ο πομπός γνωρίζει ότι συνέβη σφάλμα, όχι, όμως, ο δέκτης!
- Παρόλο που κανείς θα περίμενε, ίσως, το αντίθετο, θα αποδείξουμε ότι, σε διακριτά κανάλια χωρίς μνήμη, η χρήση ανάδρασης δεν αυξάνει τη χωρητικότητα!

Χωρητικότητα καναλιού με ανάδραση – Μοντέλο



- Στο μοντέλο του σχήματος θεωρούμε ότι ο δέκτης στέλνει όλα τα ληφθέντα σύμβολα Y_i στον πομπό άμεσα και χωρίς σφάλματα. Ο πομπός χρησιμοποιεί την πληροφορία που λαμβάνει από το δέκτη προκειμένου να αποφασίσει πώς θα μεταδώσει.

Χωρητικότητα καναλιού με ανάδραση – Ορισμοί

- Ορισμός 7.1. Κώδικας ανάδρασης (feedback code) $(2^{nR}, n)$:
  - Ένας κωδικοποιητής που παράγει ακολουθία $x_i(M, Y^{i-1})$, όπου κάθε x_i είναι συνάρτηση του τρέχοντος μηνύματος M, καθώς και των σημάτων που ελήφθησαν στο δέκτη έως και τη χρονική στιγμή $i-1$: $Y_1, Y_2, \ldots, Y_{i-1}$ και
  - Ένας αποκωδικοποιητής $g : \mathcal{Y}^n \to \{1, 2, \ldots, 2^{nR}\}$.
- Θεωρούμε ότι τα μηνύματα M είναι ομοιόμορφα κατανεμημένα. Επομένως, $P_e^{(n)} = \Pr\{g(Y^n) \neq M\}$, όπου $X^n = x^n(M)$.
- Ορισμός 7.2. Η (λειτουργική) χωρητικότητα με ανάδραση (feedback capacity), C_{FB}, του διακριτού καναλιού χωρίς μνήμη ισούται με το μέγιστο ρυθμό που είναι εφικτός με χρήση κωδίκων ανάδρασης.

Χωρητικότητα καναλιού με ανάδραση

- Θεώρημα 7.3. (Cover 7.12.1): $C_{FB} = C = \max_{p(x)} I(X;Y)$.
- Απόδειξη Είναι, κατ' αρχάς, προφανές ότι $C_{FB} \geq C$ (ευθύ), δεδομένου ότι το κανάλι χωρίς ανάδραση μπορεί να θεωρηθεί ως ειδική περίπτωση του καναλιού με ανάδραση.
- Θα αποδείξουμε ότι $C \geq C_{FB}$ και, επομένως, $C = C_{FB}$.
- Θα χρησιμοποιήσουμε και πάλι την ανισότητα Fano, όπως και στο αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού. Ωστόσο, στην απόδειξη πρέπει να ληφθεί υπόψη ότι στο κανάλι με ανάδραση δεν ισχύει η σχέση $I(X^n;Y^n) \leq nC$.

Χωρητικότητα καναλιού με ανάδραση (2)

- Αρκεί να χρησιμοποιήσουμε την εναλλακτική απόδειξη του αντιστρόφου της προηγούμενης εβδομάδας.
- Η απόδειξη επαναλαμβάνεται αυτούσια για διευκόλυνση και για να τονιστεί ότι δεν επηρεάζεται από την ύπαρξη ή μη ανάδρασης.
- Επειδή $M \to Y^n \to \hat{M} = g(Y^n)$,

$$H(M|Y^n) \leq H(M|\hat{M}).$$

- Επίσης, από την ανισότητα Fano,

$$H(M|\hat{M}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{M \neq \hat{M}\}.$$

Χωρητικότητα καναλιού με ανάδραση (3)

- Υποθέτοντας, και πάλι, ότι η τ.μ. M ακολουθεί ομοιόμορφη κατανομή,

$$nR = H(M) \overset{(a)}{=} I(M; Y^n) + H(M|Y^n)$$
$$\overset{(b)}{\leq} I(M; Y^n) + 1 + P_e^{(n)} nR$$

(a) Σχέση εντροπίας-αμοιβαίας πληροφορίας, (b) ανισότητα Fano.

Χωρητικότητα καναλιού με ανάδραση (4)

$$
\begin{aligned}
nR &\leq I(M; Y^n) + 1 + P_e^{(n)} nR \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(M; Y_i | Y^{i-1}) + 1 + P_e^{(n)} nR \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} I(M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} I(X_i, M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR
\end{aligned}
$$

(c) κανόνας αλυσίδας, (d)
$I(M, Y^{i-1}; Y_i) = I(Y^{i-1}; Y_i) + I(M; Y_i | Y^{i-1})$, (e) $X_i = f(M, Y^{i-1})$.

Χωρητικότητα καναλιού με ανάδραση (5)

$$
\begin{aligned}
nR &\leq \sum_{i=1}^{n} I(X_i, M, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_i) + 1 + P_e^{(n)} nR \\
&\leq nC + 1 + P_e^{(n)} nR = n\left(C + \frac{1}{n} + P_e^{(n)}\right)
\end{aligned}
$$

(f) $X_i = x_i(M, Y^{i-1})$ και το κανάλι δεν έχει μνήμη, οπότε $(M, Y^{i-1}) \to X_i \to Y_i$.

Χωρητικότητα καναλιού με ανάδραση (6)

- Επομένως, $I(M; Y^n) \leq nC$, και

$$nR \leq 1 + P_e^{(n)} nR + I(M; Y^n) \leq P_e^{(n)} nR + 1 + nC.$$

- Διαιρώντας με n, και για $n \to \infty$,

$$R \leq C, \text{ και, επομένως, } C_{FB} \leq C.$$

- Παρόλο που η χρήση ανάδρασης σε διακριτά κανάλια χωρίς μνήμη δεν αυξάνει τη χωρητικότητα, ενδέχεται να διευκολύνει τη μετάδοση. Για παράδειγμα, στο κανάλι διαγραφής, η μετάδοση απλουστεύεται εάν γνωρίζουμε πότε το σήμα εισόδου διαγράφεται.
- Φυσικά, στην πράξη, μπορεί να μην υπάρχει αξιόπιστος δίαυλος ανάδρασης ή να έχει κόστος (π.χ. σε εύρος ζώνης ή καθυστέρηση).