Block Code

*GF(2) είναι το σύνολο δυαδικών λέξεων… Έτσι π.χ $GF(2)^{n-k}$ είναι το σύνολο των δυαδικών λέξεων μήκους n-k

- Ένας δυαδικός block code $\mathcal{C}$ (n,n-k)
  - Έχει n-k εγγραφές
  - Τα στοιχεία του $\mathcal{C}$ ονομάζονται κωδικές λέξεις
  - Τα μηνύματα προκύπτουν από το $\mathrm{GF}(2)^{n-k}$ και αντιστοιχίζονται με τις κωδικές λέξεις.
  - Ο ρυθμός του κώδικα είναι: $R \triangleq (n-k)/n$

  - Ένας γραμμικός κώδικας περιγράφεται από τον γεννήτορα πίνακα $\mathbf{G} \in \mathrm{GF}(2)^{n-k \times n}$ που αποτελεί την αλγεβρική βάση του $\mathcal{C}$
  - Η κωδικοποίηση περιγράφεται από:

$$\mathbf{m} \mapsto \mathbf{G}^{\mathsf{T}}\mathbf{m}.$$

Block Code

- Ο δυικός του γραμμικού κώδικα $\mathcal{C}$ είναι ο (n,n-k) ορθογώνιος, δηλ:

$$\mathcal{C}^{\perp} \triangleq \left\{ \mathbf{c} \in \mathrm{GF}(2)^n : \forall \mathbf{x} \in \mathcal{C} \quad \sum_{i=1}^{n} c_i x_i = 0 \right\}.$$

- Με άλλα λόγια περιέχει όλα τα $\mathrm{GF}(2)^n$ που είναι ορθογώνια του $\mathcal{C}$
- Ο $\mathcal{C}^{\perp}$ αποδεικνύεται επίσης γραμμικός κώδικα.
- Ο γεννήτορας του $\mathcal{C}^{\perp}$ ονομάζεται partity-check matrix (πίνακας ελέγχου ισοτιμίας) για το $\mathcal{C}$ ($\mathbf{x} \in \mathcal{C}$)

$$\mathbf{H} \in \mathrm{GF}(2)^{k \times n} \qquad \mathbf{G}\mathbf{H}^{\mathsf{T}} = \mathbf{0} \qquad \mathbf{H}\mathbf{x} = \mathbf{0}$$

Block Code

- Για ένα γραμμικό κώδικα $\mathcal{C}$ με πίνακα ισοτιμίας **H** και για $\mathbf{s} \in \mathrm{GF}(2)^k$ το σύνολο

$$\mathcal{C}(\mathbf{s}) \triangleq \{\mathbf{x} \in \mathrm{GF}(2)^n : \mathbf{H}\mathbf{x} = \mathbf{s}\}$$

- Καλείται coset (ομοσύνολο). Ισχύει: $\mathcal{C} = \mathcal{C}(\mathbf{0})$
- Το coset είναι μια “μετάφραση” του αρχικού κώδικα.
- Αν $\mathbf{H}\mathbf{x}' = \mathbf{s}$, τότε $\mathcal{C}(\mathbf{s}) = \{\mathbf{x}' \oplus \mathbf{x} : \mathbf{x} \in \mathcal{C}\}$
- Το $\mathbf{x}' \in \mathcal{C}(\mathbf{s})$ με το μικρότερο βάρος, ονομάζεται “αρχηγός του coset”
- Το **s** ονομάζεται σύνδρομο.

Γραμμικοί Block Codes (Απλούστερη περίπτωση)

- Στη συγκεκριμένη περίπτωση ο Γεννήτορας έχει δομή ώστε στα πρώτα bits της κωδικής λέξης να υπάρχει το μήνυμα.
- Κακή περίπτωση για τη μυστικότητα, αλλά το αναφέρουμε λόγω απλότητας. (standard array format)

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \cdots & 0 & p_{11} & p_{21} & \cdots & p_{m1} \\ 0 & 1 & \cdots & 0 & p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{1k} & p_{2k} & \cdots & p_{mk} \end{bmatrix} = [\mathbf{I}_k \; \mathbf{P}]$$

$$\mathbf{H} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} & 1 & 0 & \cdots & 0 \\ p_{21} & p_{22} & \cdots & p_{2k} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mk} & 0 & 0 & \cdots & 1 \end{bmatrix} = [\mathbf{P}^{\mathsf{T}} \; \mathbf{I}_m]$$

$\mathbf{P}^{\mathsf{T}}$ $(m \times k)$ $\quad$ $\mathbf{I}_m$ $(m \times m)$

$$\mathbf{H}\,\mathbf{G}^{\mathsf{T}} = [\mathbf{P}^{\mathsf{T}} \; \mathbf{I}_m]\,[\mathbf{I}_k \; \mathbf{P}]^{\mathsf{T}} = [\mathbf{P}^{\mathsf{T}} \; \mathbf{I}_m] \begin{bmatrix} \mathbf{I}_k \\ \mathbf{P}^{\mathsf{T}} \end{bmatrix} = \mathbf{P}^{\mathsf{T}}\,\mathbf{I}_k \oplus \mathbf{I}_m\,\mathbf{P}^{\mathsf{T}} = \mathbf{P}^{\mathsf{T}} + \mathbf{P}^{\mathsf{T}} = \mathbf{0}$$

Αποκωδικοποίηση και ο ρόλος του Συνδρόμου

- Έστω ότι $\mathbf{r} = \mathbf{x} \oplus \mathbf{e}$ είναι η λαμβανόμενη κωδικολέξη που προκύπτει από την άθροιση μιας έγκυρης και ενός διανύσματος σφάλματος **e**
- Ισχύει: $\mathbf{Hr} = \mathbf{H}\left(\mathbf{x} \oplus \mathbf{e}\right) = \mathbf{Hx} + \mathbf{He} = 0 + \mathbf{s}$
- Αν δεν έχει γίνει σφάλμα, τότε το Σύνδρομο σύμφωνα με τα παραπάνω είναι μηδέν.
- Αποκωδικοποίηση:
  - Αν **s** = **0**, τότε δεν έχει συμβεί σφάλμα, δηλαδή **r** = **x**
  - Αν **s** ≠ **0**, τότε έχει συμβεί κάποιο σφάλμα και η έγκυρη κωδικολέξη βρίσκεται βάσει της ελάχιστης απόσταση Hamming dmin{**r**, **x**$_i$} (μοιάζει και είναι επίπονο)

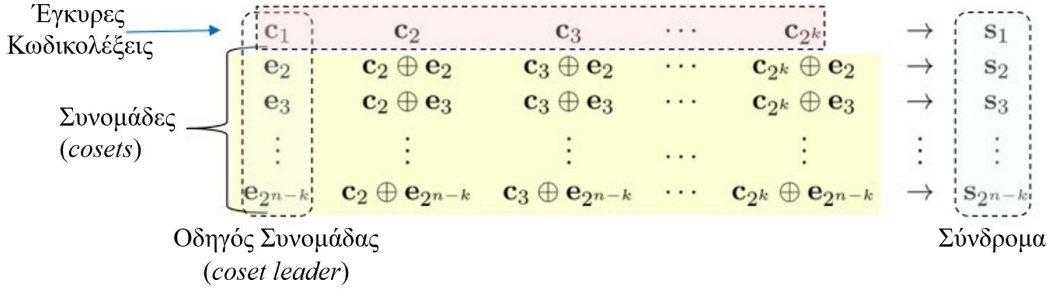
Αποκωδικοποίηση και ο ρόλος του Συνδρόμου

$$\mathbf{s} = \mathbf{e}\,\mathbf{H}^{\mathrm{T}} = \mathbf{e}\,[\mathbf{P}^{\mathrm{T}}\ \mathbf{I}_m]^{\mathrm{T}} = \mathbf{e}\left[\left(\mathbf{P}^{\mathrm{T}}\right)^{\mathrm{T}}\right] = \mathbf{e}\begin{bmatrix}\mathbf{P}\\ \mathbf{I}_m\end{bmatrix}$$

$$[s_1,\ s_2,\ \ldots,\ s_m] = [e_1,\ e_2,\ \ldots,\ e_n]\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{km} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \mathbf{P} \\ \\ \\ \\ \mathbf{I}_m \end{matrix}$$

- Πρόκειται για σύστημα m εξισώσεων με n αγνώστους και άρα έχει πολλές λύσεις ως προς **e**
- Ειδικότερα, υπάρχουν 2^{n-k} δυνατές διαφορετικές λύσεις
- Πίνακας συνδρόμων για πιθανά σφάλματα. Στην $1_η$ στήλη βρίσκεται η κωδικολέξη $\mathbf{c}_1 = \mathbf{0}$ και οπωσδήποτε $\mathbf{s}_1 = \mathbf{0}$
  - το σφάλμα προκύπτει από τη γραμμή που αντιστοιχίζει το σύνδρομο με το σφάλμα **e**
  - Η κωδική λέξη υπολογίζεται:

$$\mathbf{c} = \mathbf{e} \oplus \mathbf{r}$$



LDPC

- Οι δυαδικοί κώδικες LDPC είναι μια ειδική κατηγορία δυαδικών γραμμικών κωδίκων, που χαρακτηρίζονται από έναν αραιό πίνακα ελέγχου ισοτιμίας H, ο οποίος περιέχει πολύ μικρότερο αριθμό μονάδων από μηδενικά. Με άλλα λόγια, οι εξισώσεις ελέγχου ισοτιμίας που ορίζουν τον κώδικα περιλαμβάνουν μόνο ένα μικρό αριθμό bit. Αντί να προσδιορίζετε τον κωδικό LDPC με βάση τον πίνακα ελέγχου ισοτιμίας, είναι βολικό να χρησιμοποιήσετε μια γραφική αναπαράσταση του **H** που ονομάζεται γράφημα Tanner

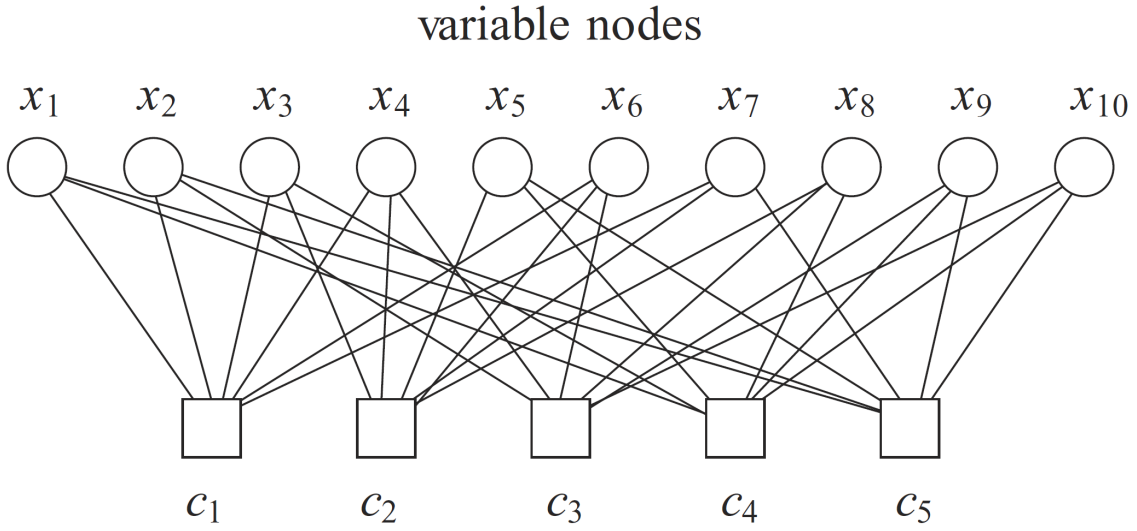
Tanner Graph

- είναι ένα διμερές γράφημα με n κόμβους μεταβλητές και k κόμβους ελέγχου.
- Ο j-οστός κόμβος ελέγχου αναπαριστά το

$$c_j = \bigoplus_{i=1}^{n} x_i h_{ji}$$

- Μια γραμμή ενώνει το x_i με το c_j αν $h_{ji} = 1$

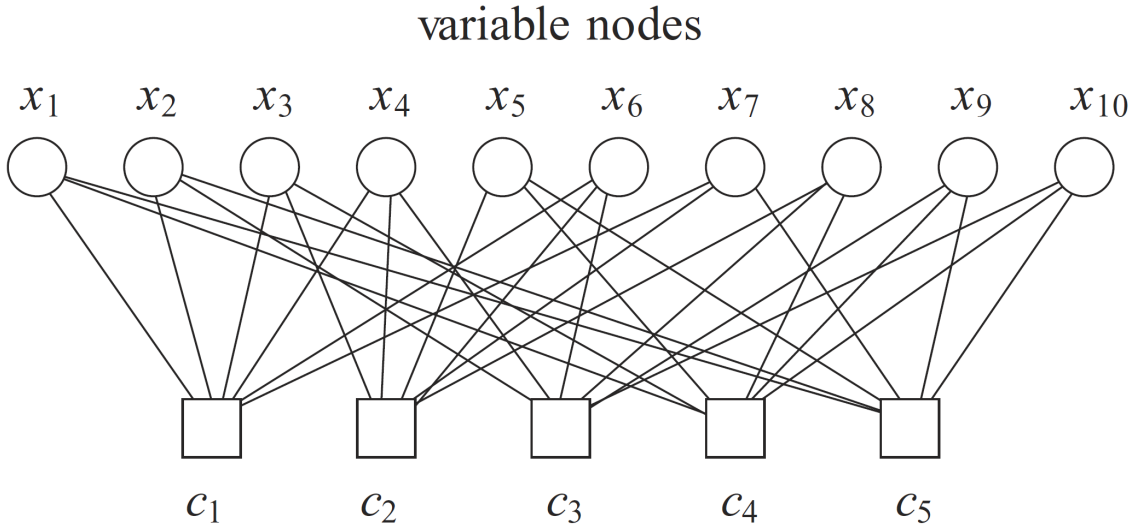
$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



Tanner Graph

- Τάξη ενός κόμβου είναι ο αριθμός των ακμών που καταλήγουν σε αυτόν
- Για ένα Tanner γράφο μπορεί να υπολογιστεί η κόμβου-μεταβλητής κατανομή ακμής-τάξης

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



Tanner Graph

- Για ένα Tanner γράφο μπορεί να υπολογιστεί η κόμβου-μεταβλητής κατανομή ακμής-τάξης όπου λ_i είναι το κλάσμα των ακμών που προσπίπτουν σε ένα κόμβο μεταβλητής τάξης i.
- Αντίστοιχα ορίζεται η κατανομή για τον κόμβο ελέγχου με τάξη j ρ_j
- Οι κατανομές ακμής – τάξης εκφράζονται πολλές φορές σαν πολυώνυμα

$$\lambda(x) = \sum_{i \geqslant 1} \lambda_i x^{i-1} \qquad \rho(x) = \sum_{j \geqslant 1} \rho_j x^{j-1}$$

- Ρυθμός:

$$R = 1 - \frac{\int_0^1 \rho(x)\mathrm{d}x}{\int_0^1 \lambda(x)\mathrm{d}x}$$

Tanner Graph

- Σημειώστε ότι ένας πίνακας ελέγχου ισοτιμίας **H** καθορίζει ένα μοναδικό γράφημα Tanner και συνεπώς μια μοναδική κατανομή τάξεων, ενώ μια κατανομή τάξεων καθορίζει ένα σύνολο κωδικών με τον ίδιο ρυθμό R (για παράδειγμα, όλες οι μεταθέσεις των κόμβων σε ένα γράφημα έχουν τον ίδιο βαθμό διανομή).
- Για μεγάλα μήκη μπλοκ, όλοι οι κωδικοί σε ένα δεδομένο σύνολο έχουν περίπου την ίδια απόδοση αποκωδικοποίησης. Ως εκ τούτου, οι κωδικοί LDPC συχνά καθορίζονται από τις κατανομές βαθμών τους (λ(x), ρ(x)) μόνο.
- Ένας κωδικός LDPC ονομάζεται κανονικός (regular) εάν όλοι οι μεταβλητοί κόμβοι έχουν τον ίδιο βαθμό και όλοι οι κόμβοι ελέγχου έχουν τον ίδιο βαθμό. Διαφορετικά, ονομάζεται ακανόνιστος (irregular).
- Πρακτικά αυτό σημαίνει ότι αν το **H** έχει ίσο αριθμό άσσων σε όλες τις γραμμές και το ίδιο ισχύει και για τις στήλες, τότε είναι regular.

Tanner Graph

Example 6.1. A rate-$\frac{1}{2}$ regular $(3, 6)$ LDPC code is such that all variable nodes have degree 3, and all check nodes have degree 6. Its degree distributions are simply

$$\lambda(x) = x^2 \quad \text{and} \quad \rho(x) = x^5.$$

Example 6.2. The following irregular degree distributions correspond to another rate-$\frac{1}{2}$ LDPC code:

$$\lambda(x) = 0.106\,257x + 0.486\,659x^2 + 0.010\,390x^{10} + 0.396\,694x^{19},$$
$$\rho(x) = 0.5x^7 + 0.5x^8.$$

Αποκωδικοποίηση μετάδοσης μηνύματος LDPC

- Εκπεμπόμενη ακολουθία $\mathbf{x} = (x_1, \ldots, x_n)^\intercal \in \mathcal{C}$
- Διάδικο κανάλι χωρίς μνήμη $(\{0, 1\}, p_{Y|X}(y|x), \mathcal{Y})$
- Λαμβανόμενη ακολουθία $\mathbf{y} = (y_1, \ldots, y_n)^\intercal$
- Αποκωδικοποίηση με log-likelihood ratio

$$\lambda_i = \log\left(\frac{\mathbb{P}[X_i = 0|\mathbf{y}]}{\mathbb{P}[X_i = 1|\mathbf{y}]}\right) \quad \text{for } i \in [\![1, n]\!]$$

- Το πρόσημο δείχνει το 0 ή το 1 και το πλάτος είναι μέτρο αξιοπιστίας της εκτίμησης.

Αποκωδικοποίηση μετάδοσης μηνύματος LDPC

- Για κάθε $i \in [\![1, n]\!]$ έστω ότι $\mathcal{N}(i)$ δηλώνει τους δείκτες των κόμβων ελέγχου που συνδέονται με τον κόμβο μεταβλητής x_i στο γράφημα Tanner. Το σύνολο μπορεί να ληφθεί από τον πίνακα ελέγχου ισοτιμίας

$$\mathcal{N}(i) \triangleq \{j : h_{ji} = 1\}$$

- Αντίστοιχα για κάθε $j \in [\![1, k]\!]$ ορίζονται τα σύνολα:

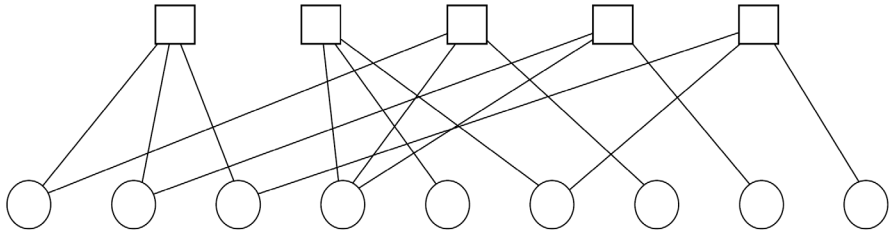
$$\mathcal{M}(j) \triangleq \{i : h_{ji} = 1\}$$

- Το LLR μπορεί να προσεγγιστεί με τον αλγόριθμο belief-propagation.

Αποκωδικοποίηση μέσα από το Tanner graph

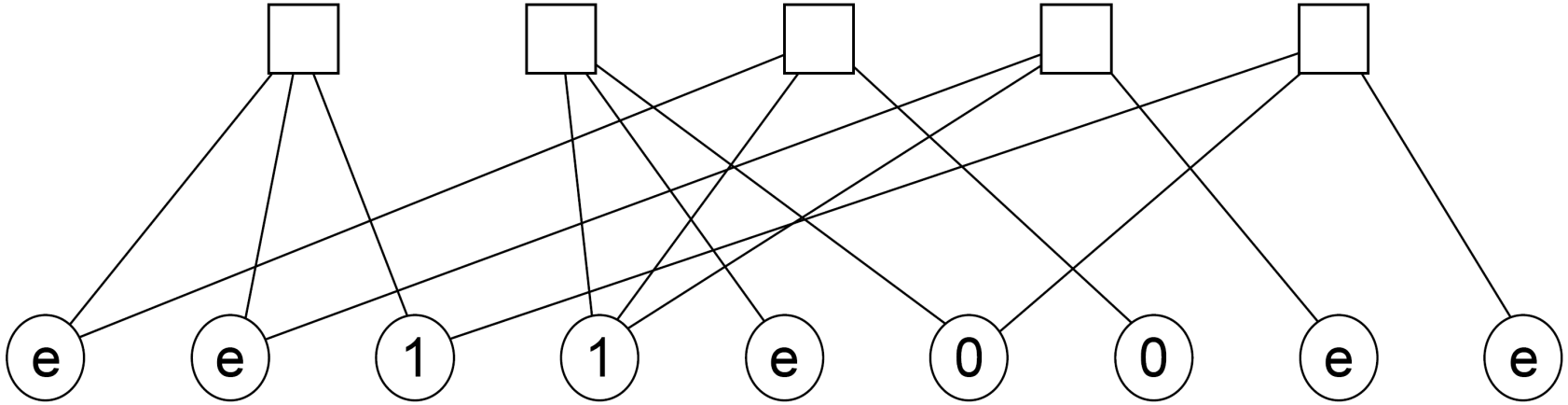
- Φτιάχνουμε το Tanner Graph, π.χ.

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



- Τοποθετούμε στους κόμβους μεταβλητών το λαμβανόμενο (με τα πιθανά erasures)

Αποκωδικοποίηση μέσα από το Tanner graph



- Επαναληπτικά και στη σειρά τσεκάρω αν υπάρχει κόμβος ελέγχου που βλέπει ένα απλό erasure.
- Διορθώνω το λάθος μέσω άρτιας ισοτιμίας.
- Συνεχίζω να κοιτάζω για κόμβο ελέγχου με ένα απλό erasure.
- Αν κάνω έστω και μια διόρθωση, θα πρέπει να ξανά-σκανάρω από την αρχή μήπως κόμβος με δυο erasures έχει διορθωθεί σε κόμβο με ένα erasure.
- Αν σε σκανάρισμα δεν κάνω διόρθωση ο αλγόριθμος τέλειωσε.
- Αν δεν έχει μείνει erasure, τα κατάφερα.

Αποκωδικοποίηση μέσα από το Tanner graph