Codes LDPC Non-binaires Hybrides

Lucile Sassatelli, David Declercq

{sassatelli,declercq}@ensea.fr

ETIS ENSEA/UCP/CNRS UMR 8051,Cergy, France

18 mars 2008

Funded by DGA

Plan

Plan

Objectifs

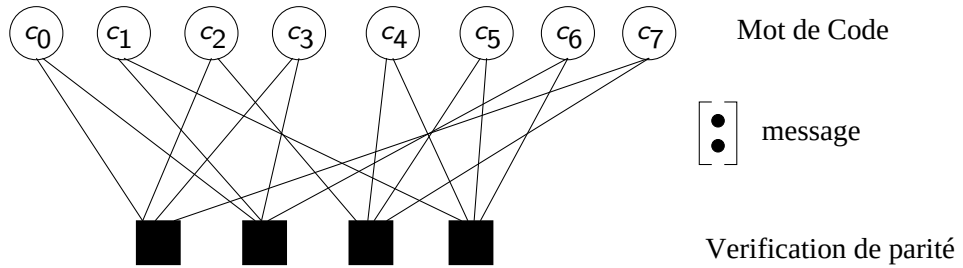
- Les codes LDPC hybrides sont une classe très générale de codes LDPC, généralisant les familles existantes de codes LDPC (LDPC reguliers ou irreguliers, binaires or non-binaires, codes IRA, multiedge LDPC).

- Généraliser les familles existantes de LDPC en préservant les avantages:
  - Encodage simple (linear time)
  - Décodage peu complexe (BP-like decoder)
  - Caractérisation asymptotique de la classe hybride avec techniques DE ou EXITcharts
    - ★ Prédiction des performances de décodage asymptotiques
    - ★ Preuve du comportement à seuil, en développant une condition de stabilité.
    - ★ Optimisation des paramètres d'une classe

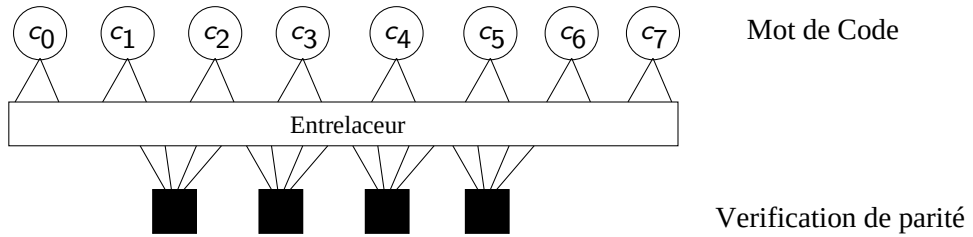
Codes LDPC sur $GF(2)$: Cas Binaire

- Code LDPC binaire: Application linaire de $GF(2)^K$ vers $GF(2)^N$.
- graphe de Tanner = graphe bayésien non orienté

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

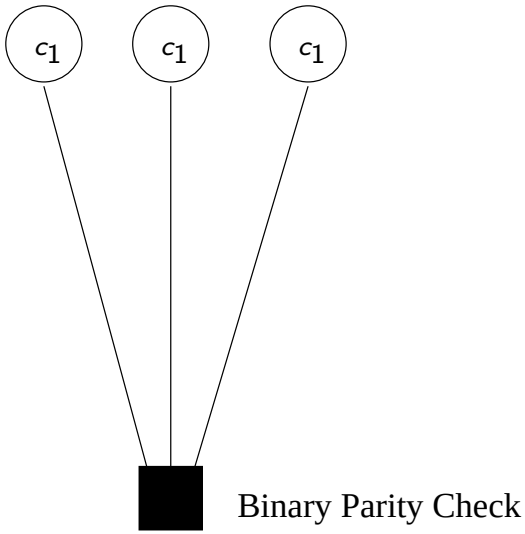


Famille $(d_v = 2, d_c = 4)$



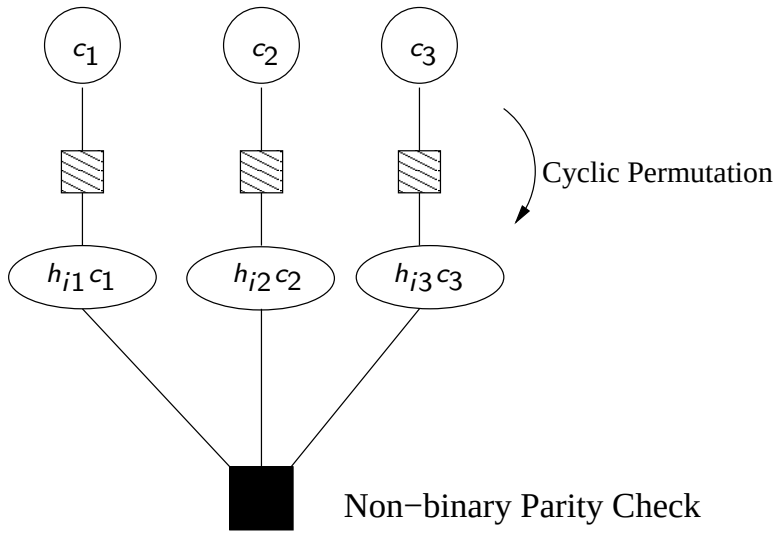
Graphes de Tanner de Noeuds de Vérification de Parité

Binary LDPC



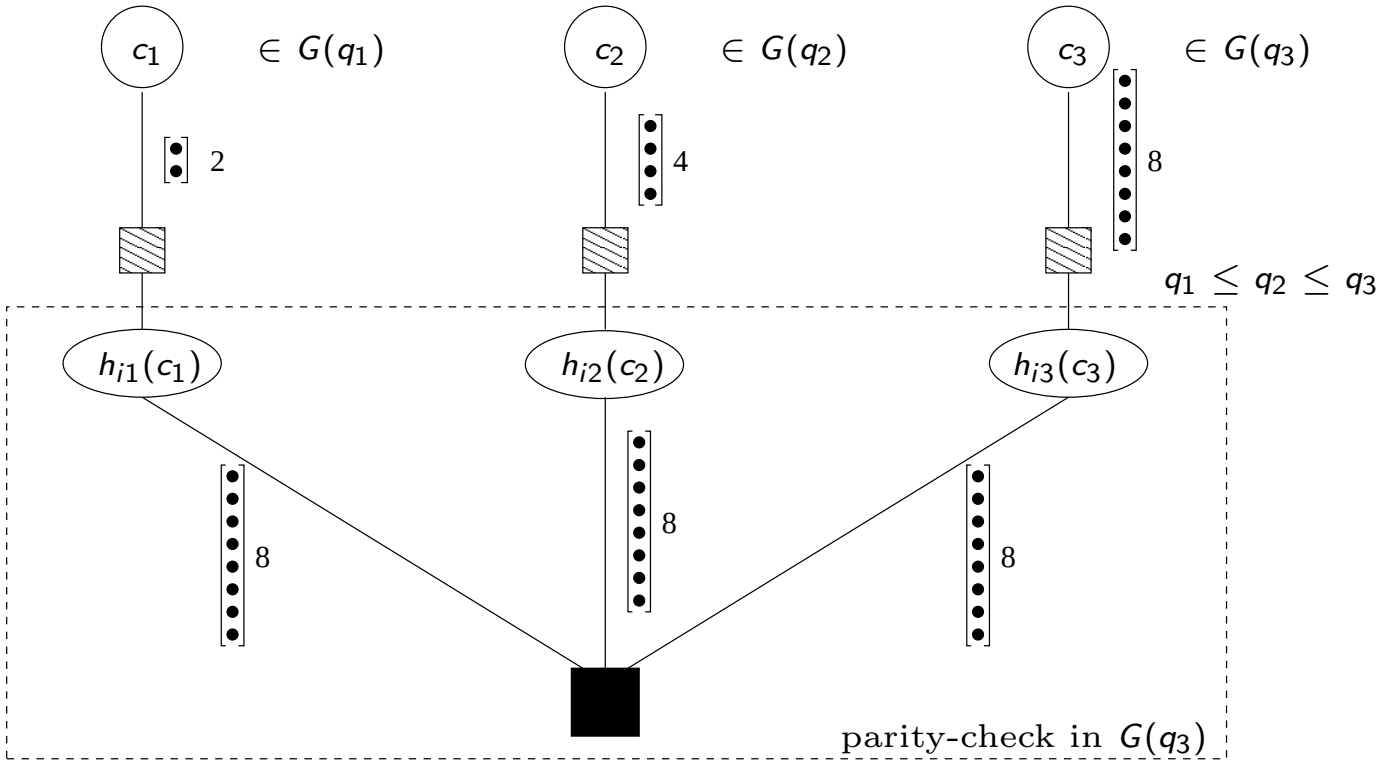
$$c_1 + c_2 + c_3 = 0, \qquad \{c_i\} \in GF(2)$$

Non−binary LDPC



$$h_{i1}c_1 + h_{i2}c_2 + h_{i3}c_3 = 0, \qquad (h_{ij}, c_j) \in GF(q)$$

Vérification de Parité d'un Code LDPC Hybride



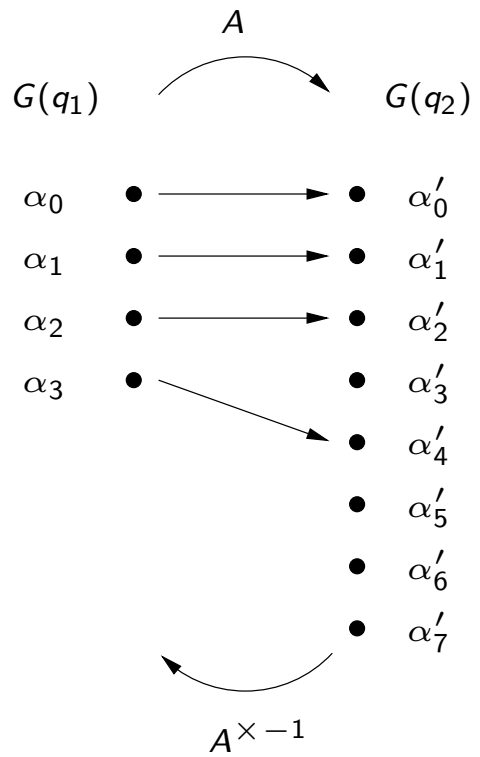
$$h_{i1}(c_1) + h_{i2}(c_2) + h_{i3}(c_3) = 0, \qquad h_{ij}(c_j) \in G(q_3)$$

définit un code composant sur le groupe $G = G(q_1) \times G(q_2) \times G(q_3)$

Transformation des Messages par Applications Linéaires

$G(q_1) = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$

$G(q_2) = \{\alpha'_0, \alpha'_1, \alpha'_2, \alpha'_3, \alpha'_4, \alpha'_5, \alpha'_6, \alpha'_7\}$



- Extension :
  $\mathbf{y} = \mathbf{x}^{\times A}$ étendu de $\mathbf{x}$ par A est défini par :
  $\forall i = 0, \ldots, q_2 - 1$

  si $i \notin \mathrm{Im}(A)$, $\quad y_i = 0$

  si $i \in \mathrm{Im}(A)$, $\quad y_i = x_j$ avec j tel que $\mathbf{i} = A\mathbf{j}$

- Troncature :
  $\mathbf{x} = \mathbf{y}^{\times A^{-1}}$ tronqué de $\mathbf{y}$ par A^{-1} est défini par :

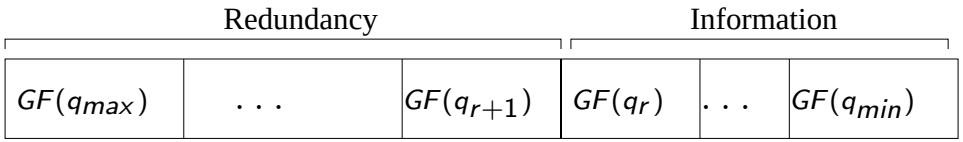
  $j = 0, \ldots, q_1 - 1$, $x_j = y_i$ avec j tel que $\mathbf{j} = A^{-1}\mathbf{i}$

Mot de Code et Matrice de Vérification de Parité Hybrides

On considère un code LDPC Hybride sur $G = G(q_{min}) \times \ldots \times G(q_{max})$

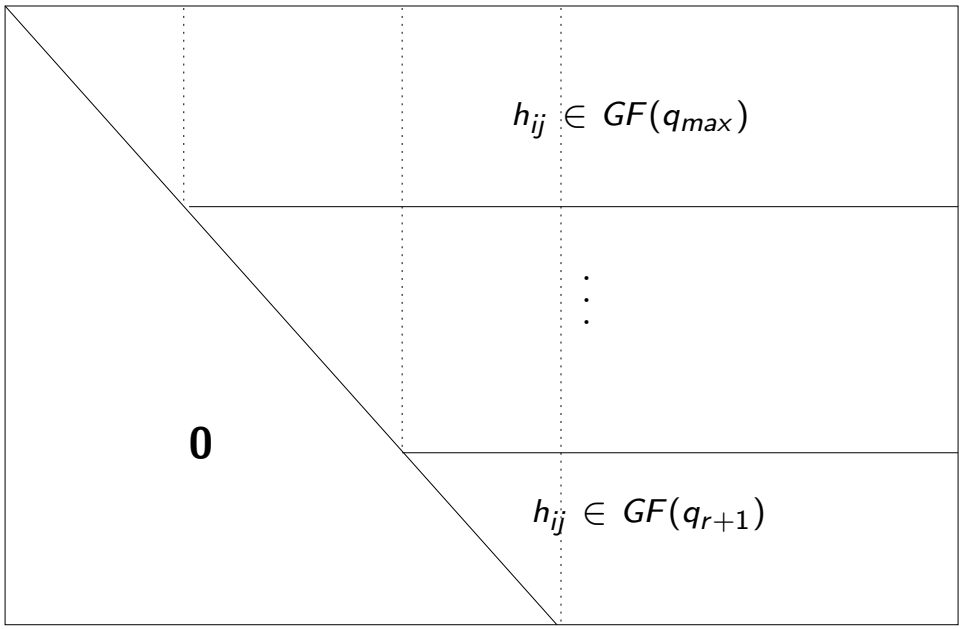
1st constraint: Increasing sorting of the field orders

$\mathbf{c} =$



2nd constraint: Upper−triangular matrix

$\mathbf{H} =$



Décodeur Belief Propagation

[Goupil06] *Goupil et al., FFT-based BP Decoding of General LDPC Codes over Abelian Groups, IEEE Trans. on Comm., to appear 2006*

Une itération de décodage dans le domaine probabilité:

Step 1 **Mise à jour noeuds de variable** dans $G(q_k)$: produit des messages entrant

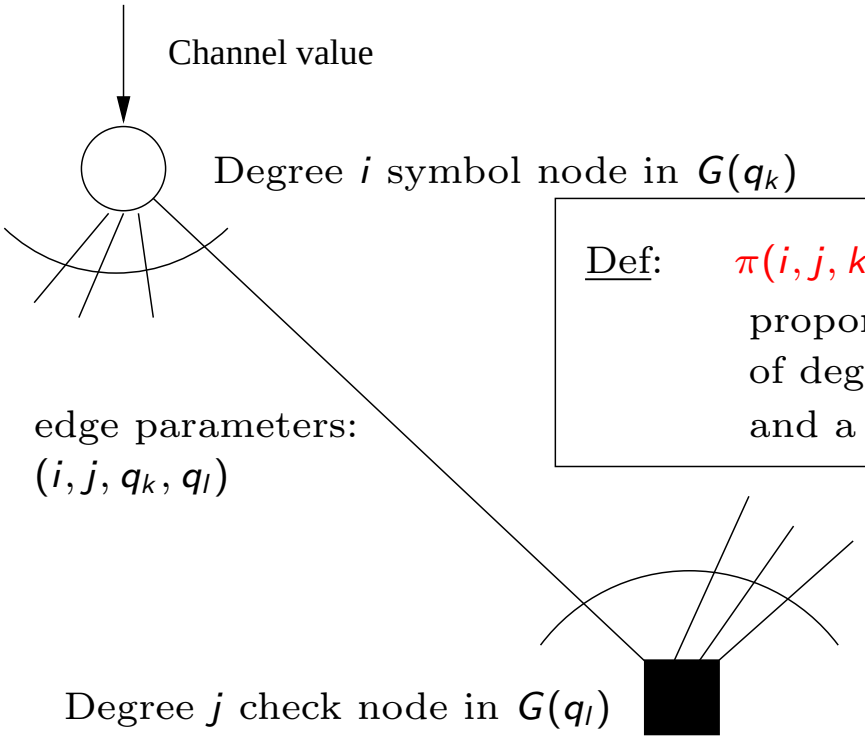
Step 2 **Extension des messages par A** $G(q_k) \rightarrow G(q_l)$: dans le vecteur message de taille q_l à **0**, remplir les q_k composantes de Im(A) par le message entrant

Step 3 **Mise à jour noeuds de parité** dans $G(q_l)$ dans le domaine de Fourier

- $\rightarrow$ FFT de taille q_l
- $\rightarrow$ Produit des vecteurs FFT terme-à-terme
- $\rightarrow$ IFFT de taille q_l

Step 4 **Troncation des messages par A** de $G(q_l) \rightarrow G(q_k)$: les q_k composantes sont remplies avec les composantes dans Im(A) du message entrant

Paramétrisation des codes LDPC hybrides



Def: $\pi(i, j, k, l) \in [0, 1]$:
proportion of edges that link a data node of degree i in $G(q_k)$ and a check node of degree j in $G(q_l)$

Les codes LDPC hybrides ont une paramétrisation riche puisque l'espace des paramètres a 4 dimensions.

Preuve du comportement à seuil: contraintes et hypothèses

- On veut montrer le **comportement à seuil** des codes LDPC hybrides:

  si la probabilité d'erreur a déjà décru en dessous d'un certain point, alors elle doit converger vers zéro.

  $\longrightarrow$ Le but est de dériver une **condition de stabilité** pour les codes LDPC hybrides.

- On a besoin de:
  - probabilité d'erreur indépendante du mot de code envoyé
    $\longrightarrow$ La **symétrie** du canal permet de ne travailler qu'avec le mot de code nul.
  - **LA-invariance** pour la condition de stabilité.

Plan

Symétrie (1)

[Bennatan06] *Bennatan et al., Design and Analysis of Non-binary LDPC Codes for Arbitrary Discrete-Memoryless Channels, Trans. on Inform. Theory, Feb. 2006*

- $c \in G(q)$ est un symbole du mot de code
- $\mathbf{y}$ est le vecteur proba correspondant à l'observation du canal
- $\mathbf{y}^{+a} = (y_a, \ldots, y_{a+(q-1)})$
- $\mathbf{w} = LDR(\mathbf{y})$ and $\mathbf{w}^{+a} = LDR(\mathbf{y}^{+a})$: $w_i^a = w_{a+i} - w_a, \quad \forall i = 0 \ldots q-1$

Definition (1)

Un canal est symétrique ssi le vecteur proba observation $\mathbf{Y}$ *vérifie:*

$$\forall a \in G(q), \quad P(\mathbf{Y} = \mathbf{y} | c = a) = P(\mathbf{Y} = \mathbf{y}^{+a} | c = 0)$$

$$\iff \forall a \in G(q), \quad P(\mathbf{W} = \mathbf{w} | c = 0) = e^{w_a} P(\mathbf{W} = \mathbf{w}^{+a} | c = 0)$$

Symétrie (2)

- Le décodeur hybride préserve la symétrie des messages.

Lemma (2)

La probabilité d'erreur d'un code LDPC hybride, utilisé sur un canal symétrique et décodé avec BP, est indépendante du mot de code transmis.

Invariance par applications linéaires (LA-invariance)

- But: Prouver que la LA-invariance des messages est induite par le choix uniforme des extensions.
- La LA-invariance permet de caractériser simplement les densités des messages.

Definition (2)

Un vecteur aléatoire $\mathbf{Y}$ *de taille* q_l *est LA-invariant ssi pour tout* k *et* $(A^{-1}, B^{-1}) \in T_{k,l} \times T_{k,l}$, *les vecteurs aléatoires* $\mathbf{Y}^{\times A^{-1}}$ *et* $\mathbf{Y}^{\times B^{-1}}$ *sont identiquement distribués.*

Lemma (3)

Si un vecteur aléatoire $\mathbf{Y}$ *de taille* q_l *est LA-invariant, alors toutes ses composantes sont identiquement distribuées.*

LA-invariance (2)

Definition (3)

Soit $\mathbf{X}$ *un vecteur aléatoire de taille* q_k*, on définit l'extension aléatoire de taille* q_l *de* $\mathbf{X}$*, notée* $\tilde{\mathbf{X}}$*, comme le vecteur aléatoire* $\mathbf{X}^{\times A}$*, où* A *est choisi uniformément dans* $E_{k,l}$ *et indépendant de* $\mathbf{X}$*.*

Lemma (4)

Un vecteur aléatoire $\mathbf{Y}$ *de taille* q_l *est LA-invariant ssi il existe* q_k *et un vecteur aléatoire* $\mathbf{X}$ *de taille* q_k *tel que* $\mathbf{Y} = \tilde{\mathbf{X}}$*.*

$\longrightarrow$ On étudie la famille de codes LDPC hybrides, paramétrisée par $\pi(i, j, k, l)$, avec des applications linéaires choisies uniformément.

Plan

La condition de stabilité

- $\Omega = \sum_{j,k,l} \pi(i=2,k,j,l) \frac{q_k - 1}{q_l - 1}(j-1)$
- $\Delta = \sum_{k,l} \pi(k,l) \frac{1}{q_l - 1} \sum_{i=1}^{q_k - 1} \int \sqrt{p(y|i)p(y|0)}dy$

Theorem

- *Si* $\Omega\Delta \geq 1$, *la probabilité d'erreur sera bornée au dessus de zéro.*
- *Si* $\Omega\Delta < 1$, *la probabilité d'erreur converge vers zéro (stable).*

1^{re} remarque: stabilité des codes LDPC sur $GF(q)$

[Bennatan06] *Bennatan et al., Design and Analysis of Non-binary LDPC Codes for Arbitrary Discrete-Memoryless Channels, Trans. on Inform. Theory, Feb. 2006*

- Pour un code LDPC sur $GF(q)$ classique sur BI-AWGNC, la condition de stabilité hybride est réduite à:

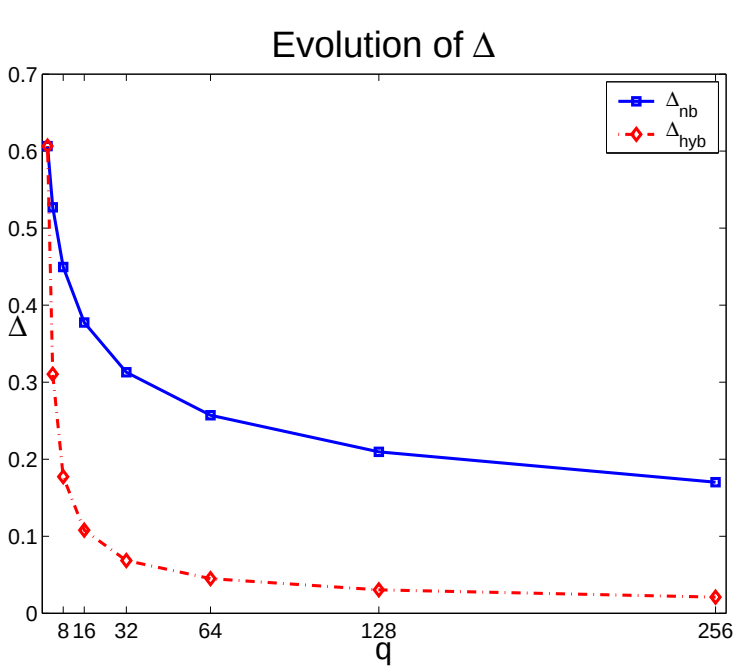
$$\begin{aligned} \Omega_{nb} &= \rho'(1)\lambda'(0) \\ \Delta_{nb} &= \frac{1}{q-1}\sum_{i=1}^{q-1} exp(-\frac{1}{2\sigma^2}n_i) \end{aligned}$$

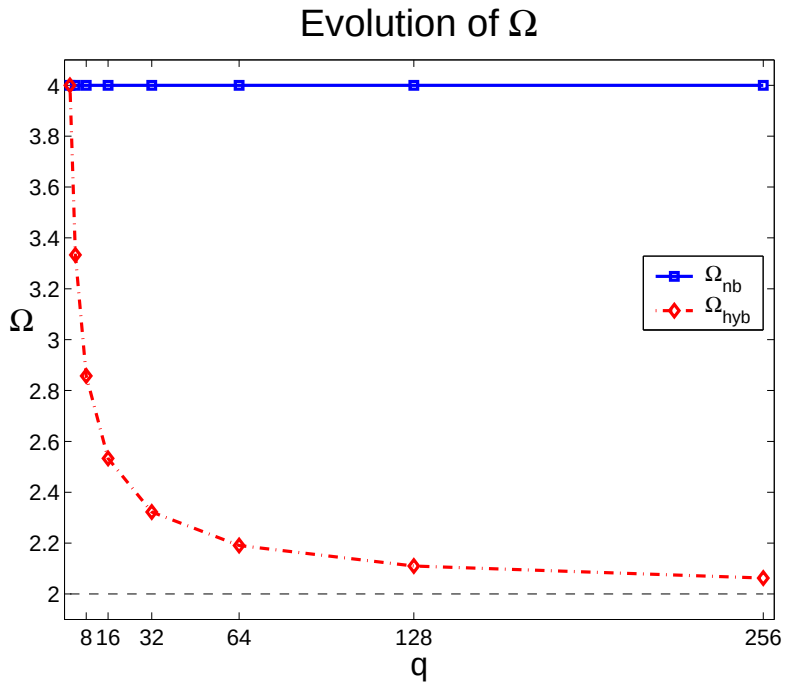
  avec n_i, le nombre de 1 dans la représentation binaire de $\alpha_i \in G(q)$.

- $\lim_{q\to\infty} \Delta = 0$
- Sur canal BI-AWGN, ça signifie que tout code LDPC sur $GF(q)$ est stable dès que q est assez grand.
- Les cycle codes LDPC non-binaires, avec $d_v = 2$, sont stables si q est assez grand.

2^e remarque: la condition hybride est moins contraignante

- Codes LDPC hybrides sur $G(q) - G(2)$ et non-binaires sur $GF(q)$.
- $R = \frac{1}{2}$, $d_v = 2$, $q = 2 \ldots 256$
- Canal BI-AWGN
- $\Omega_{hyb} \leq \Omega_{nb}$ and $\Delta_{hyb} \leq \Delta_{nb}$
- Un code LDPC hybride peut être stable à SNR plus bas qu'un code LDPC non-binaire classique.





Plan

Analyse de l'évolution de l'information mutuelle

- Analyse asymptotique dans le cas du canal BI-AWGN, pour une famille de codes LDPC hybrides paramétrée par $\pi(i, j, k, l)$ et applications linéaires choisies uniformément.
- Décodage réussi $\Leftrightarrow \lim_{t\to\infty} x_{APP}^{(t)} = 1$
- Pour ça, on suit l'évolution de l'IM des messages au cours des itérations
- Messages symétriques et LA-invariants.
- Chaque composante de message a une distribution avec un Δ_∞ en raison des changements de groupes. $\rightarrow$ On ne peut pas supposer que tous les messages sur le graphe sont gaussiens, pour utiliser leur moyenne (infinie).
- Pour parvenir à une projection des densités sur un seul paramètre, on considère les moyennes de messages gaussiens qui ont la même information mutuelle que les messages circulant sur le graphe.

Modèle gaussien des messages pour le canal BI-AWGN

- Modèle général pour un message rapport log de densité (LDR) $\mathbf{W}$ symétrique et suivant une loi gaussienne.

$$\begin{aligned} \mathbf{W} &\sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma}) \\ \mathbf{\Sigma}_{i,j} &= \mathbf{m}_i + \mathbf{m}_j - \mathbf{m}_{i \oplus j}, \qquad \{i,j\} \in GF(q_k) \end{aligned}$$

  - Cas de l'observation rapport log de vraisemblance (LLR) venant du canal

$$\mathbf{m} = \frac{2}{\sigma^2}\mathbf{n}, \qquad \mathbf{n} = [n_0 \dots n_{q_k-1}]$$

  - Cas des messages LA-invariants rentrant et sortant des noeuds de parité dans $G(q_l)$, de même information mutuelle que messages gaussiens LA-invariants de moyenne

$$\mathbf{m} = m\mathbf{1}_{q_l}$$

$\rightarrow$ Ces vecteurs gaussiens, de même IM que les messages du graphe, étant contraints, la connaissance de $\mathbf{m}$ est suffisante pour caractériser la densité $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$

Information mutuelle de vecteurs gaussiens contraints

[Bennatan06] *Bennatan et al., Design and Analysis of Non-binary LDPC Codes for Arbitrary Discrete-Memoryless Channels, Trans. on Inform. Theory, Feb. 2006*

- IM d'un message LDR $\mathbf{w}$ de taille q

$$x_{\mathbf{w}} = 1 - \mathbb{E}_{\mathbf{w}} \left(1 + \sum_{i=0}^{q-1} e^{-w_i} \right)$$

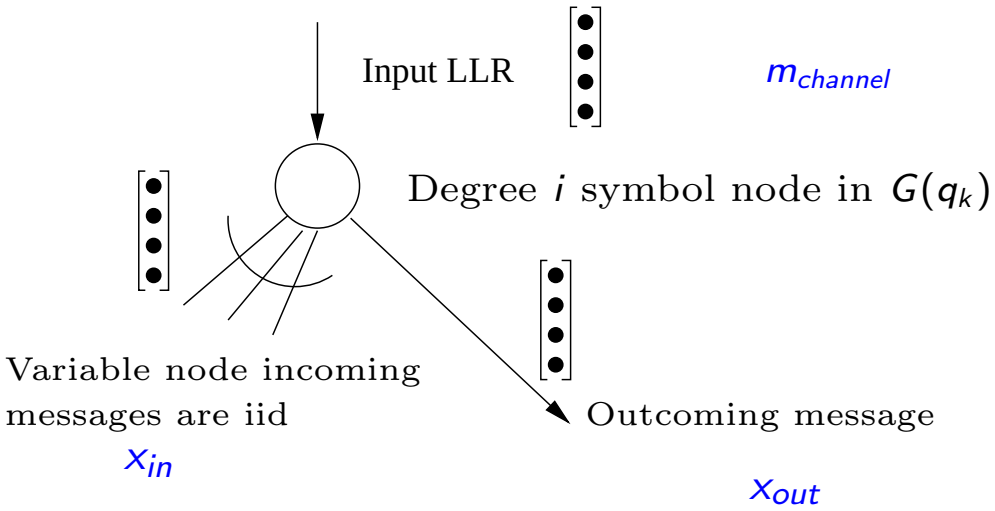
  - Messages non-LA-invariants (sorties des noeuds de variable)

$$x_{\mathbf{w}} = J_v(\mathbf{m}, q_k), \qquad p(\mathbf{w}) = \text{fct}(\mathbf{m} = \frac{2}{\sigma^2}\mathbf{n} + (i-1)m\mathbf{1}_{q_k-1})\}$$

  - Messages LA-invariants (liés aux check nodes)

$$x_{\mathbf{w}} = J_c(m, q_l), \qquad p(\mathbf{w}) = \text{fct}(\mathbf{m} = m\mathbf{1}_{q_l-1})$$

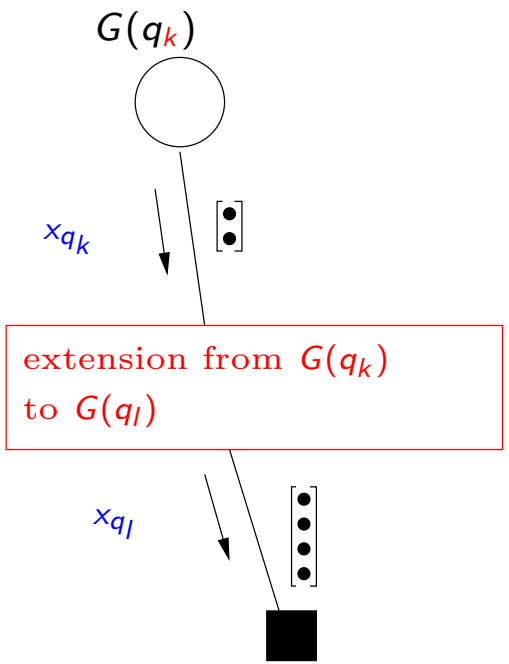
Evolution de l'IM à travers les noeuds de variable

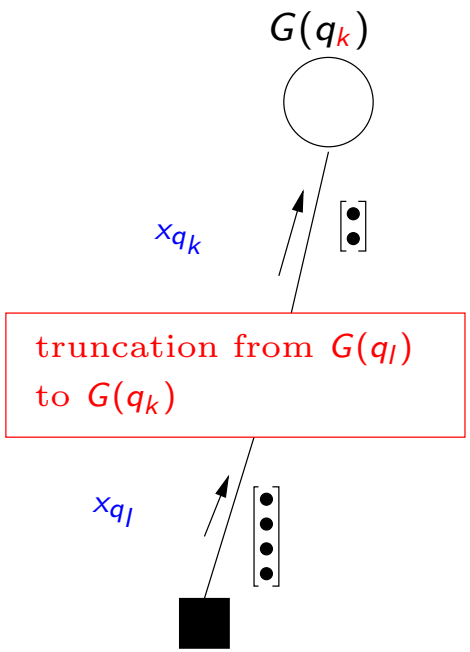


$$x_{out} = J_v\left(m_{channel} + (i-1)J_c^{-1}(x_{in}, q_k), q_k\right)$$

La symétrie des messages LDR est conservee par l'addition de la mise à jour du noeud de variable

Evolution de l'IM à travers les noeuds d'extension/troncation



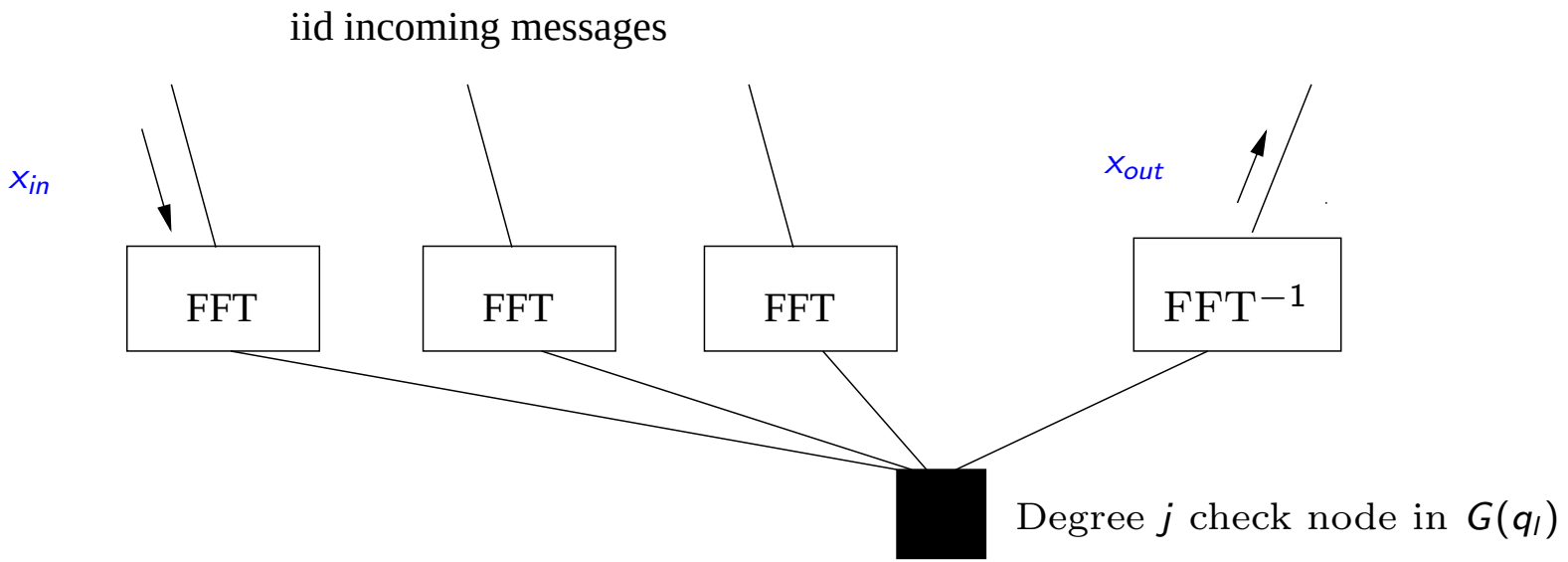


$$(1 - x_{q_l}) log_2(q_l) = (1 - x_{q_k}) log_2(q_k)$$

$$x_{q_k} = J_c\left(J_c^{-1}(x_{q_l}, q_l), q_k\right)$$

Les messages étendus et tronqués sont symétriques si les messages entrant le sont.

Evolution de l'IM à travers les noeuds de parité



$$x_{out} = 1 - J_c\left((j-1)J_c^{-1}(1-x_{in}, q_l), q_l\right)$$

Si les messages d'entree sont symetriques, les messages de sortie obtenus par convolution le sont aussi

Fonction EXIT des codes LDPC hybrides (1)

- Décodage réussi $\Leftrightarrow \lim_{t\to\infty} x_{APP}^{(t)} = 1$
  Or on n'a pas d'expression $x_{APP}^{(t+1)} = fonction(x_{APP}^{(t)})$
  $\rightarrow$ pas de condition de la forme $\lim_{t\to\infty} x_{APP}^{(t)} = 1$
  $\Leftrightarrow x_{APP}^{(t+1)} > x_{APP}^{(t)}$

- Or $\lim_{t\to\infty} x_{APP}^{(t)} = 1 \Leftrightarrow \lim_{t\to\infty} x_{ext}^{(t)} = 1$ où x_{ext} est l'IM moyenne des messages après extensions

- On a avec x_{ext}: $x_{ext}^{(t+1)} = fonction(x_{ext}^{(t)})$

La condition d'évolution vers 1 de l'IM est donc: $x_{ext}^{(t+1)} > x_{ext}^{(t)}, \quad \forall t.$
[1]

[1] Dans la suite: $x_{ext} = x_{vc}$

Fonction EXIT des codes LDPC hybrides (2)

$$x_{cv,k}^{(j,l')^{(t)}} = J_c\left(J_c^{-1}\left(1 - J_c\left((j-1)J_c^{-1}(1 - x_{vc}^{(t)}, q_{l'}), q_{l'}\right), q_{l'}\right), q_k\right)$$

$$x_{vc}^{(t+1)} = \sum_{i,k} \pi(i,k) \sum_{l} 1 - \pi(l|i,k) \frac{log(q_k)}{log(q_l)} \left(1 - J_v\left(\mathbf{m}_{sc}{}^{q_k} + (i-1)J_c^{-1}(\sum_{j,l'} \pi(j,l'|i,k) x_{cv,k}^{(j,l')^{(t)}}, q_k)\mathbf{1}_{q_k-1}, q_k\right)\right)$$

- Paramétrisation de la famille de codes LDPC hybrides
- Observation venant du canal
- Suivi de l'information mutuelle

$\longrightarrow$ Les propriétés de symétrie et de LA-invariance permettent le suivi de l'information mutuelle

Plan

2 exemples de bons codes LDPC hybrides

- Stratégie d'optimisation: pour un rendement donné, trouver la famille hybride avec le plus bas seuil de convergence

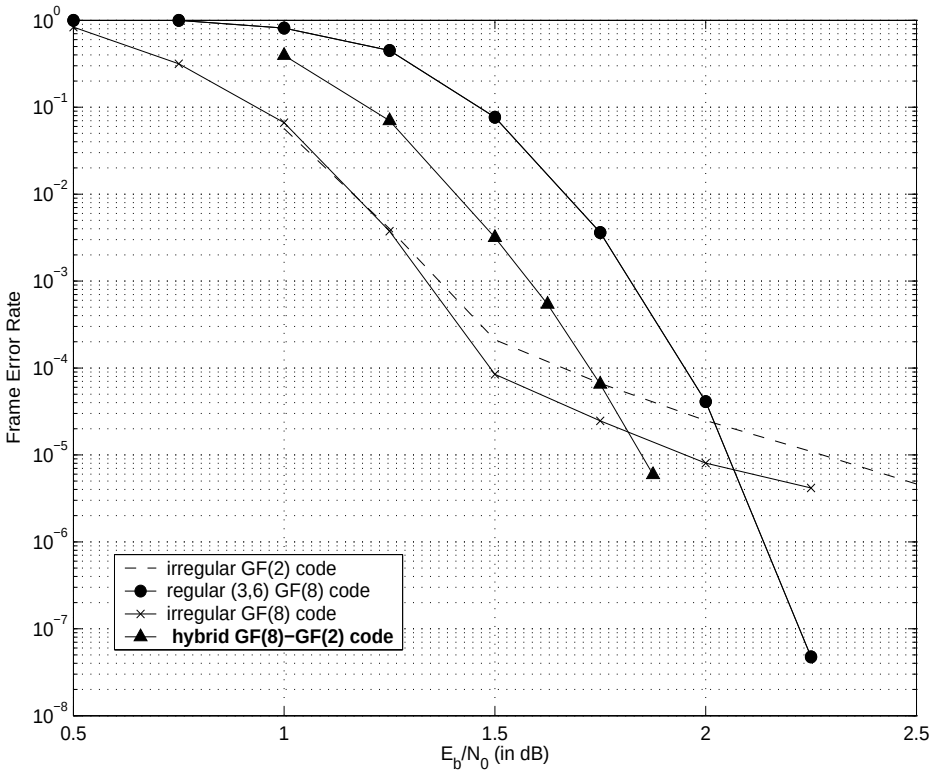
- Pour éviter une optimisation jointe et rendre le problème linéaire, les paramètres des noeuds de parité sont fixés $\rightarrow$ Optimisation de $\pi(i, k|j, l)$

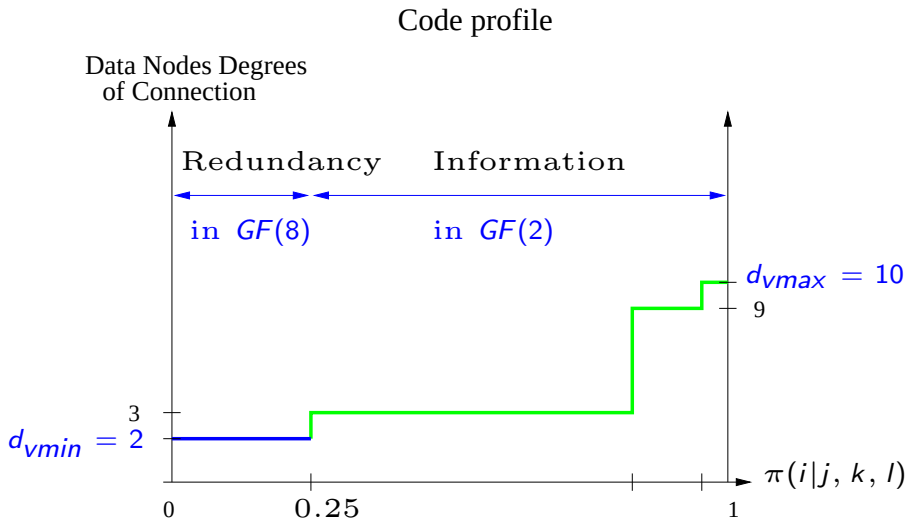
  - 1^{er} exemple : optimisation des degrés des noeuds de variable $\pi(i|k, j, l)$ pour un profil de groupes donné

  - 2^{e} exemple : optimisation du profil de groupes des noeuds de variable $\pi(k|i, j, l)$ pour un profil de connection donné.

1^{er} Exemple: Optimisation des connections des noeuds de variable

$R = \frac{1}{2}$, $N_{bit} = 3008$, $Nitermax = 500$



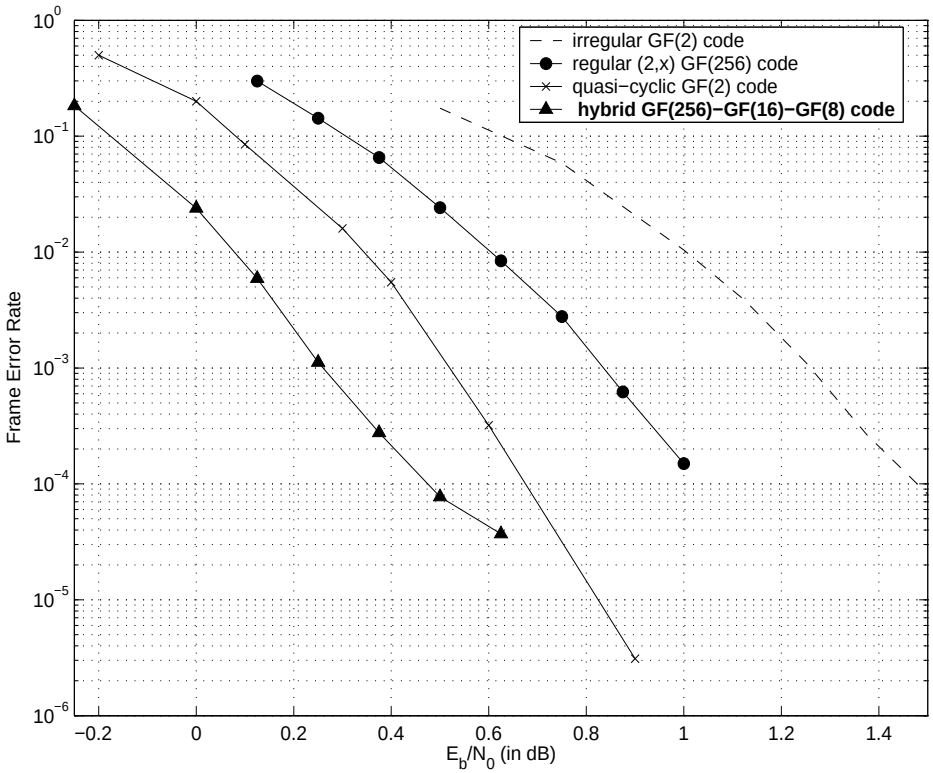


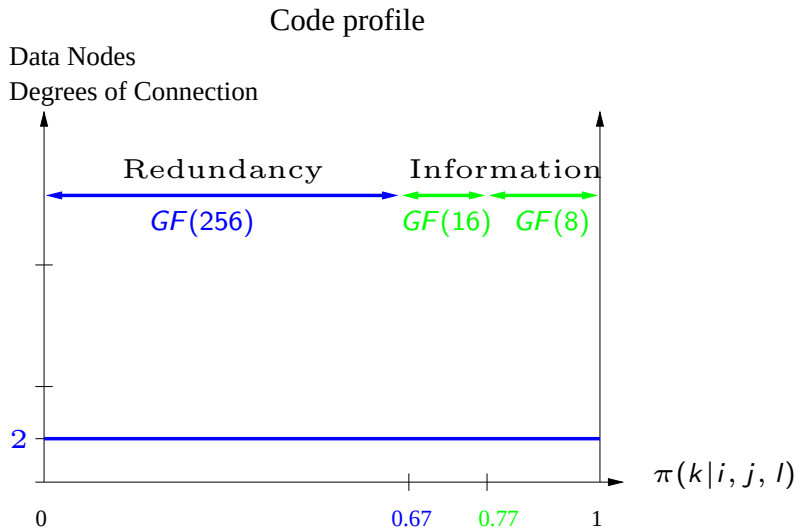
With average check node density $\overline{\rho} = 14.3$

- : A priori assumptions
- : Result of LP optimization

2^e Exemple: Optimisation des groupes des noeuds de variable

$R = \frac{1}{6}$, $N_{bit} = 6144$, $Nitermax = 500$





With connectivity profile $d_v = 2, d_c = 3$

- : A priori assumptions
- : Result of LP optimization

Codes LDPC hybrides pour les faibles rendements (1)

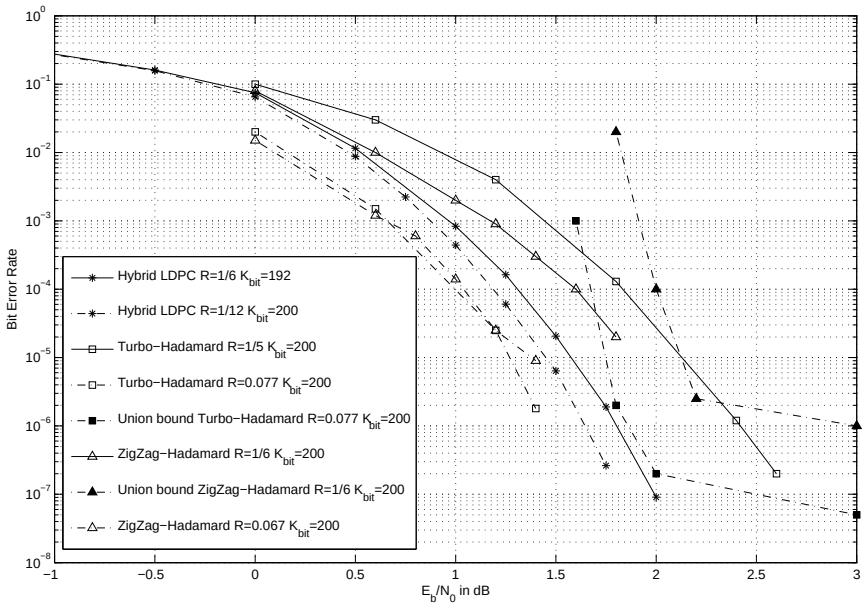
[Poulliat06] *Poulliat et al., Design of regular (2,dc)-LDPC codes over GF(q) using their binary images, IEEE Trans. Commun., 2007*

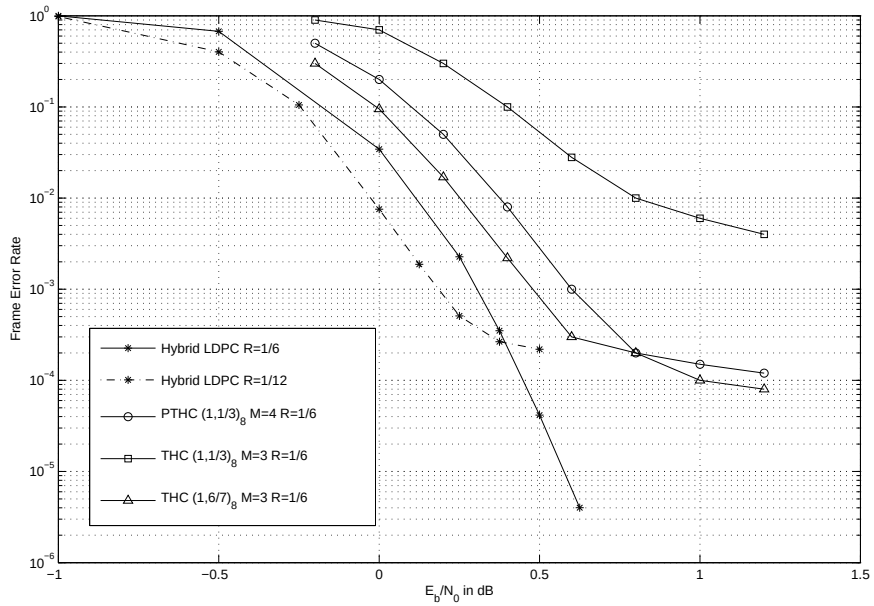
- Les codes bas rendements sont importants pour les communications à faible SNR, ou large bande utilisant le CDMA pour remplacer le code d'étalement.
- Les codes hybrides sont particulièrement bien adaptés aux bas rendements:

$$R = \frac{R_{graph} \sum_{k \in \mathcal{I}} \pi(k) \log_2(q_k)}{R_{graph} \sum_{k \in \mathcal{I}} \pi(k) \log_2(q_k) + (1 - R_{graph}) \log_2(q_{red})}$$

- L'optimisation taille finie [Poulliat06] pour annuler les cycles dans le graphe convient particulièrement aux codes hybrides en raison de la structure des applications linéaires.

Codes LDPC hybrides pour les faibles rendements (2)



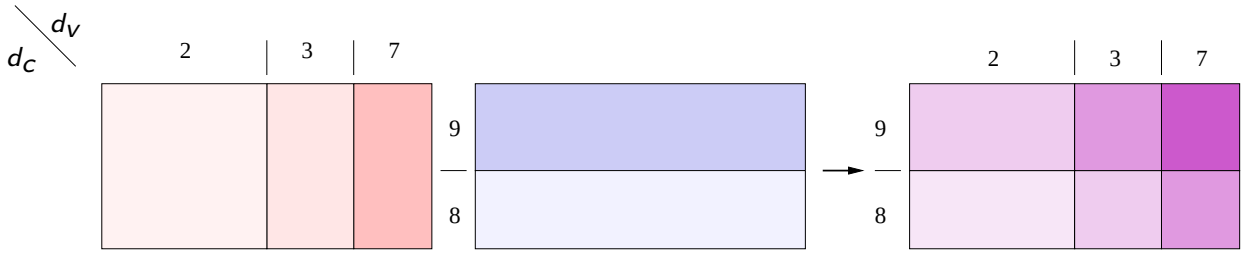


Conclusion

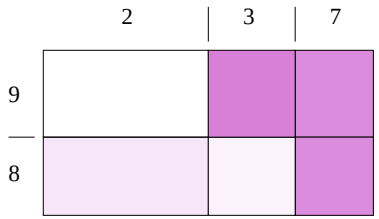
- Nouvelle classe de codes LDPC complètement irréguliers: Codes LDPC non-binaires hybrides.
- Analyse asymptotique des performances.
- Des propriétés spécifiques, comme la **LA-invariance**, ont été étudiées pour aboutir à une condition de stabilité pour les codes LDPC hybrides.
- Optimisations asymptotique et à taille finie pour le canal BI-AWGN.
- Codes particulièrement intéressants pour les bas rendements.

Discussion

- Binary LDPC codes: $\{\lambda, \rho\} \rightarrow$ assumes isotropic repartition of the **H** matrix density



- Detailed representation $\Pi_{i,j} \rightarrow$ allows to control localized density of **H**



- Hybrid LDPC Codes: $\pi(i, j, k, l) \rightarrow$ allows to jointly control the localized densities of **H** and the size of the symbols in each area