# Problème II

Laurent Massoulié

March 28, 2007

## Graphes "expansifs" et codes correcteurs d'erreurs

Le graphe aléatoire $g_{N,M,k}$ est défini de la manière suivante. C'est un graphe dont les sommets sont partitionnés en deux ensembles $V_1$ et $V_2$, de tailles respectives $N$ et $M$. Les arêtes sont spécifiées au moyen de variables aléatoires $Z_i(j)$, $i = 1\ldots, N$, $j = 1\ldots, k$, qui sont indépendantes et identiquement distribuées, uniformément sur $\{1, \ldots, M\}$. Le $i$ème sommet de $V_1$ est incident à $k$ sommets de $V_2$, déterminés par les valeurs de $Z_i(1), \ldots, Z_i(k)$ (pour le moment on n'exclut pas la possibilité d'arêtes multiples).

Finalement, pour tout ensemble $C \subset V_1$, on note $\Gamma(C)$ le sous-ensemble des sommets de $V_2$ qui sont incidents à un sommet $i$ appartenant à $C$:

$$\Gamma(C) = \{j \in V_2 : \exists i \in C \text{ tel que } (i,j) \text{ arête de } g_{N,M,k}\}.$$

1. Etablir que pour tout $C \subset V_1$, tout $\ell \in [0, k|C|]$, on a:

$$\mathbf{P}(|\Gamma(C)| \leq k|C| - \ell) \leq \mathbf{P}\left(\mathrm{Bin}(k|C|, 1 - k|C|/M) \leq k|C| - \ell\right). \tag{1}$$

2. On dit qu'un graphe bi-partite tel que $g_{N,M,k}$ est "$(\epsilon, m)$-expansif" s'il vérifie la propriété suivante:

$$\forall C \subset V_1,\ |C| \leq \epsilon|V_1| \Rightarrow |\Gamma(C)| \geq m|C|. \tag{2}$$

   On note $\pi$ la probabilité que $g_{N,M,k}$ soit $(\epsilon, (1-\delta)k)$-expansif, pour deux paramètres $\epsilon$, $\delta \in ]0,1[$ donnés. Etablir l'inégalité:

$$1 - \pi \leq \sum_{a=1}^{\epsilon N} \binom{N}{a} \binom{ka}{\lceil ka\delta \rceil} \left(\frac{ka}{M}\right)^{\lceil ka\delta \rceil}. \tag{3}$$

3. On suppose maintenant $N$ grand, $M \sim \alpha N$ pour un $\alpha > 0$ fixé, et $k$ fixé. Montrer que, pour tout $\delta > 0$ tel que $k\delta > 1$, il existe $\epsilon > 0$ tel que, avec probabilité tendant vers 1 lorsque $N$ tend vers $+\infty$, $g_{N,M,k}$ est $(\epsilon, (1-\delta)k)$-expansif. On pourra par exemple majorer le terme binomial $\binom{N}{a}$ par $(Ne/a)^a$ dans l'inégalité (3) ci-dessus, où $e = \exp(1)$.

4. Montrer que, lorsque $N$ tend vers $+\infty$, pour $k$ fixé et $M \sim \alpha N$, la probabilité que $g_{N,M,k}$ n'admette pas d'arêtes multiples tend vers $\exp(-k(k-1)/2\alpha)$. En déduire, pour $k > 1$ et $\delta > 0$ tel que $\delta k > 1$, l'existence de graphes bi-partites qui sont $(\epsilon, (1-\delta)k)$-expansifs, dont tous les sommets de $V_1$ ont un degré $k$, et qui n'ont pas d'arêtes multiples.

5. Soit un graphe bi-partite comme dans la question précédente, en particulier tel que chaque sommet de $V_1$ est de degré $k$. On l'utilise pour former un code correcteur d'erreurs de la

manière suivante. On considère qu'un mot code $x$ de $N$ bits $x_1, \ldots, x_N$, appartenant à $\{0,1\}^N$, est valide si et seulement si pour tout $j = 1, \ldots, M$, on a:

$$\sum_{i=1}^{N} A_{ij} x_i = 0 \bmod(2)$$

où l'addition est modulo 2, et $A_{ij}$ vaut 1 si le $i$ème sommet de $V_1$ est connecté au $j$ème sommet de $V_2$, et 0 sinon. Les mots codes sont donc les vecteurs du *noyau* d'une application linéaire de $\{0,1\}^N$ dans $\{0,1\}^M$, et sont au moins au nombre de $2^{N-M}$.

Etablir que, lorsque le graphe est $(\epsilon, (1-\delta)k)$-expansif, avec $\delta < 1/2$, alors deux mots codes distincts $x$, $x'$ sont tels que leur distance de Hamming $\sum_{i=1}^{N} |x_i - x'_i|$ est strictement supérieure à $\epsilon N$. Ceci garantit que des corruptions de moins de $\epsilon N/2$ bits lors de la transmission d'un mot code peuvent être corrigées sans ambiguité.

Indice: on se ramènera au cas où l'un des mots codes est le vecteur nul.