Để chứng minh định lý của Bregman, ta sử dụng 3 mệnh đề dưới đây:

(A) $H(X) \leq \log_2(|suppX|)$. Dấu "="xảy ra khi và chỉ khi X là phân phối đều trên $supp\ X$ (tức $Prob(X = a) = \frac{1}{n}$ với $a \in supp\ X, n = |supp\ X|$.

Chứng minh: Không mất tính tổng quát, giả sử $p_i > 0\ \forall i$. Áp dụng bất đẳng thức **AM-GM** $a_1^{p_1} .....a_n^{p_n} \leq p_1 a_1.....p_n a_n$: Đặt $a_i = \frac{1}{p_i}$ và lấy log 2 vế, ta được:

$$H(X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \leq \log_2 \left(\sum_{i=1}^{n} p_i \frac{1}{p_i} \right) = \log_2 n.$$

(B) $H(X, Y) = H(X) + H(Y|X)$, và tổng quát ta có $H(X_1, ..., X_n) = H(X_1) + H(X_2|X1) + ... + H(X_n|X_1, ..., X_{n-1})$

Chứng minh:

$$\begin{aligned}
H(X, Y) &= -\sum_{i,j} p(a_i, b_j) \log_2 p(a_i, b_j) \\
&= -\sum_{i,j} p(a_i, b_j) \log_2 p(a_i) p(b_j|a_i) \\
&= -\sum_{i,j} p(a_i, b_j) [\log_2 p(a_i) + \log_2 p(b_j|a_i)] \\
&= -\sum_{i,j} p(a_i, b_j) \log_2 p(a_i) - \sum_{i,j} p(a_i, b_j) \log_2 p(b_j|a_i) \\
&= -\sum_{i,j} p(a_i, b_j) \log_2 p(a_i) - \sum_{i,j} p(a_i).p(b_j|a_i) \log_2 p(b_j|a_i) \\
&= -\sum_{i=1}^{m} p(a_i) \log_2 p(a_i) \sum_{j=1}^{n} p(b_j|a_i) + H(Y|X) \\
&= -\sum_{i=1}^{m} p(a_i) \log_2 p(a_i) + H(Y|X) = H(X) + H(Y|X)
\end{aligned}$$

(Do $\sum_{j=1}^{n} p(b_j|a_i) = \sum_{j=1}^{n} \frac{P(b_j.a_i)}{P(a_i)} = \frac{1}{P(a_i)} \sum_{j=1}^{n} P(b_j.a_i) = \frac{1}{P(a_i)} P(a_i) = 1$)

(C) Nếu $supp\ X$ được chia thành d tập $E_1, .., E_d$ sao cho $E_j := \{a \in suppX : |supp(Y|a)| = j\}$ thì:

$$H(Y|X) \leq \sum_{j=1}^{d} Prob(X \in E_j) \log_2 j.$$

Chứng minh: Ta có $H(Y|X) = \sum_{i=1}^{m} p(a_i) H(Y|a_i)$. Tiến hành chia tập $a_1, ..., a_m$ thành các tập con E_j theo giả thuyết và sử dụng kết quả từ **(A)**, ta có:

$$\begin{aligned}
H(X, Y) &= \sum_{j=1}^{d} \sum_{a \in E_j} p(a) H(Y|a) \\
&\leq \sum_{j=1}^{d} \sum_{a \in E_j} p(a) \log_2 j \\
&= \sum_{j=1}^{d} Prob(X \in E_j) \log_2 j.
\end{aligned}$$

Định lý 1. Đặt $M = (m_{ij})$ là ma trận $n \times n$ chỉ chứa hai giá trị 0,1 và đặt $d_1, ..., d_n$ là tổng các hàng của ma trận M, hay $d_i = \sum_{j=1}^{n} m_{ij}$. Khi đó:

$$perM \leq \prod_{i=1}^{n} (d_i!)^{1/d_i}.$$

Chứng minh: Xét $G = (U \cup V, E)$ là đồ thị hai phía tương ứng với ma trận M, trong đó d_i là bậc tương ứng của các đỉnh u_i, và kí hiệu Σ là tập các *perfect matching* của G. Vì $perM = m(G) = |\Sigma|$ nên thay vì tìm cận trên cho $perM$ như định lý 1, ta sẽ tìm cận trên cho $|\Sigma|$. Giả sử $|\Sigma| \neq 0$ và mỗi $\sigma \in \Sigma$ là một hoán vị tương ứng $\sigma(1)\sigma(2)....\sigma(n)$ của các chỉ số. Vì vậy, tương ứng với mỗi giá trị $u_i \in U$ là một giá trị $v_{\sigma(i)} \in U$ theo phép song ánh σ

Ý tưởng ban đầu là chọn σ một cách ngẫu nhiên và xét biến ngẫu nhiên $X = (X_1, X_2, .., X_n) = (\sigma(1), \sigma(2), ..., \sigma(n))$.

Theo mệnh đề **(A)**,

$$H(\sigma(1), \sigma(2), ..., \sigma(n)) = \log_2(|\Sigma|)$$

Do đó chỉ cần chỉ ra

$$H(\sigma(1), ..., \sigma(n)) \leq \log_2(\prod_{i=1}^{n} (d_i!)^{1/d_i}) = \sum_{i=1}^{n} \frac{1}{d_i} \log_2(d_i!). \tag{1}$$

Sử dụng mệnh đề **(B)**, ta có

$$H(\sigma(1), \sigma(2), ..., \sigma(n)) = \sum_{i=1}^{n} H(\sigma(\sigma_i)|\sigma(1), \sigma(2), ..., \sigma(i-1)) \tag{2}$$

Ý tưởng của Radhakrishnan là xét các đỉnh $u_1, u_2, ..., u_n$ theo một *thứ tự ngẫu nhiên* τ , với xác suất là như nhau và bằng $\frac{1}{n!}$, và lấy giá trị kì vọng của các entropy. Nói cách khác, ta xét các cặp *matching* theo thứ tự $\sigma(\tau(1)), \sigma(\tau(2)), ..., \sigma(\tau(n))$. Xét τ cố định, khi đó $k_i = \tau^{-1}(i)$ được hiểu là vị trí của u_i theo thứ tự ngẫu nhiên τ là k_i. Khi đó, biểu thức (2) trở thành:

$$H(\sigma(1), ..., \sigma(n)) = \sum_{i=1}^{n} H(\sigma(i)|\sigma(\tau(1)), ..., \sigma(\tau(k_i - 1)))$$

Khi đó

$$H(\sigma(1), ..., \sigma(n)) = \frac{1}{n!} \sum_{\tau} \left(\sum_{i=1}^{n} H(\sigma(i)|\sigma(\tau(1)), ..., \sigma(\tau(k_i - 1))) \right)$$

Xét biểu thức

$$H(\sigma(i)|\sigma(\tau(1)), ..., \sigma(\tau(k_i - 1))) \tag{3}$$

Để tìm cận trên cho , ta sẽ sử dụng mệnh đề **(C)**, áp dụng với biến ngẫu nhiên $X = \Big(\sigma(\tau(1)), ..., \sigma(\tau(k_i - 1))\Big)$ và $Y = \sigma(i)$. Với τ cố đinh $\in S_n$ và $\sigma \in \Sigma$ đặt $N_i(\sigma, \tau)$ là số các giá trị $k \in [n]$ sao cho $u_i v_k \in E(G)$ và k $\notin \Big(\sigma(\tau_1), ..., \sigma(\tau_{k_i - 1})\Big)$ (nói cách khác, $N_i(\sigma, \tau)$ là số khả năng còn lại cho $\sigma(i)$ khi đã biết $\sigma(\tau_1), ..., \sigma(\tau_{k_i-1}))$. Vì $deg(u_i) = d_i$ đồng thời u_i phải được ghép cặp trong σ nên $1 \leq N_i(\sigma, \tau) \leq d_i$ với mọi $\sigma \in \Sigma$. Tiến hành chi tập $suppX$ thành các tập con $E_{i,j}^{(\tau)}$ sao cho :

$$\Big(\sigma(\tau_1), ..., \sigma(\tau_{k_i - 1})\Big) \in E_{i,j}^{(\tau)} <=> N_i(\sigma, \tau) = j, \text{ với } 1 \leq j \leq d_i$$

Coi $N_i(\sigma,\tau)$ là một biến ngẫu nhiên trên Σ, ta có:

$$Prob(X \in E_{i,j}^{(\tau)}) = Prob(N_i(\sigma,\tau) = j)$$

Từ mệnh đề **(C)**, với τ cố định:

$$\begin{aligned}
H(\sigma(i)|\sigma(\tau_1),...,\sigma(\tau_{k_i-1})) &\le \sum_{j=1}^{d_i} Prob(|N_i(\sigma,\tau)| = j)\log_2 j \\
&= \sum_{j=1}^{d_i} \log_2 j \sum_{\sigma\in\Sigma} P(\sigma).P(N_i(\sigma,\tau) = j|\sigma) \text{ (Do } \sigma \in \Sigma \text{ là 1 nhóm đầy đủ)} \\
&= \sum_{j=1}^{d_i} \log_2 j \sum_{\sigma\in\Sigma} \frac{P(N_i(\sigma,\tau) = j|\sigma)}{|\Sigma|}
\end{aligned}$$

Kết hợp với ...

$$H(\sigma(1),...,\sigma(n)) \le \sum_{j=1}^{d_i} \log_2 j \left(\frac{1}{n!|\Sigma|} \sum_{\sigma\in\Sigma} \sum_{\tau\in S_n} P(N_i(\sigma,\tau) = j|\sigma) \right) \tag{4}$$

Xét $\sum_{\tau\in S_n} P(N_i(\sigma,\tau) = j|\sigma)$

Với mỗi σ cố định $\in \Sigma$ và trên từng hoán vị $\tau \in S_n$, $N_i(\sigma,\tau)$ nhận các giá trị từ 1 tới d_i với xác suất như nhau là $\frac{1}{d_i}$, vì $N_i(\sigma,\tau)$ chỉ phụ thuộc vào vị trí của $\sigma(i)$ trong hoán vị τ (do σ đã cố định), tương ứng với các đỉnh k thỏa mãn $u_i v_k \in E(G)$ (Nếu $\sigma(i)$ là lân cận gần nhất của i theo thứ tự của hoán vị (giả sử là τ_1). Khi đó $N_i(\sigma,\tau_1) = d_i$ và $P(N_i(\sigma,\tau_1) = d_i|\sigma) = \frac{1}{d_i}$; Nếu $\sigma(i)$ là lân cận gần thứ hai của i theo thứ tự của hoán vị (giả sử τ_2). Khi đó $N_i(\sigma,\tau_2) = d_i - 1$ và $P(N_i(\sigma,\tau_2) = d_i - 1|\sigma) = \frac{1}{d_i}$,...tương tự với n! hoán vị của S_n .Do đó, giá trị của biểu thức trên bằng:

$$\sum_{\tau\in S_n} P(N_i(\sigma,\tau) = j|\sigma) = n!.\frac{1}{d_i} = \frac{n!}{d_i}$$

Và tổng ở (2) trở thành:

$$\sum_{j=1}^{d_i} \log_2 j \left(\frac{1}{n!|\Sigma|} \sum_{\sigma\in\Sigma} \frac{n!}{d_i} \right) = \sum_{j=1}^{d_i} \log_2 j \left(\frac{1}{n!|\Sigma|}.|\Sigma|.\frac{n!}{d_i} \right) = \sum_{j=1}^{d_i} \log_2 j.\frac{1}{d_i} = \frac{\log_2 d_i!}{d_i}$$

$\sigma(\tau_1), \sigma(\tau_2), ..., \sigma(\tau_n)$

$H(\sigma(\tau_1),..,\sigma(\tau_n)) = \sum_{i=1}^{n} H(\sigma(\tau_i)|\sigma(\tau_1),...,\sigma(\tau_{k_i-1}))$

$H(\sigma(\tau_1),..,\sigma(\tau_n)) = \frac{1}{n!} \sum_{\tau} \left(\sum_{i=1}^{n} H(\sigma(\tau_i)|\sigma(\tau_1),...,\sigma(\tau_{k_i-1})) \right)$

$\Theta(n^{\log_2 2})$