CS 880: Complexity of Counting Problems 03/20/2012

Lecture 17: Reduction to Discrete Unitary Matrix (Step 2.1)

Instructor: Jin-Yi Cai Scribe: Chetan Rao

Let matrix $\mathbf{A}$ be the bipartization of an $m \times n$ matrix $\mathbf{B}$ i.e. $\mathbf{A}$ is the $(m+n)\times(m+n)$ matrix -

$$\mathbf{A} = \begin{pmatrix} 0 & \mathbf{B} \\ \mathbf{B}^T & 0 \end{pmatrix}$$

Let $\boldsymbol{\mu} = \{\mu_1, \mu_2, \ldots, \mu_s\}$ and $\boldsymbol{\nu} = \{\nu_1, \nu_2, \ldots, \nu_t\}$ be two decreasing sequences of positive rational numbers of lengths $s \geq 1$ and $t \geq 1$, respectively i.e. $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ satisfy $\mu_1 > \mu_2 > \ldots > \mu_s$ and $\nu_1 > \nu_2 > \ldots > \nu_t$. Let $\boldsymbol{m} = \{m_1, m_2, \ldots, m_s\}$ and $\boldsymbol{n} = \{n_1, n_2, \ldots, n_t\}$ be two sequences of positive integers such that $m = \sum_{i=1}^{s} m_i$ and $n = \sum_{i=1}^{t} n_i$.

The rows of $\mathbf{B}$ are indexed by $\mathbf{x} = (x_1, x_2)$ where $x_1 \in [s]$ and $x_2 \in [m_{x_1}]$ and the columns of $\mathbf{B}$ are indexed by $\mathbf{y} = (y_1, y_2)$ where $y_1 \in [t]$ and $y_2 \in [n_{y_1}]$. Then, for all $\mathbf{x}, \mathbf{y}$, we have

$$B_{\mathbf{x},\mathbf{y}} = B_{(x_1,x_2),(y_1,y_2)} = \mu_{x_1}\nu_{y_1}S_{\mathbf{x},\mathbf{y}}$$

where $\mathbf{S} = \{S_{\mathbf{x},\mathbf{y}}\}$ is an $m \times n$ matrix in which every entry $(S_{\mathbf{x},\mathbf{y}})$ is a root of unity (power of ω_N).

$$\mathbf{B} = \begin{pmatrix} \mu_1\mathbf{I}_{m_1} & & & \\ & \mu_2\mathbf{I}_{m_2} & & \\ & & \ddots & \\ & & & \mu_s\mathbf{I}_{m_s} \end{pmatrix} \begin{pmatrix} \mathbf{S}_{(1,*),(1,*)} & \mathbf{S}_{(1,*),(2,*)} & \cdots & \mathbf{S}_{(1,*),(t,*)} \\ \mathbf{S}_{(2,*),(1,*)} & \mathbf{S}_{(2,*),(2,*)} & \cdots & \mathbf{S}_{(2,*),(t,*)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{(s,*),(1,*)} & \mathbf{S}_{(s,*),(2,*)} & \cdots & \mathbf{S}_{(s,*),(t,*)} \end{pmatrix} \begin{pmatrix} \nu_1\mathbf{I}_{n_1} & & & \\ & \nu_2\mathbf{I}_{n_2} & & \\ & & \ddots & \\ & & & \nu_t\mathbf{I}_{n_t} \end{pmatrix}$$

where $\mathbf{I}_k$ denotes the $k \times k$ identity matrix.

Also let

$$I \equiv \bigcup_{i\in[s]} \{(i,j) | j \in [m_i]\} \quad \text{and} \quad J \equiv \bigcup_{i\in[t]} \{(i,j) | j \in [n_i]\}$$

Given a vector $\mathbf{x} \in I$ and $j \in [t]$, we let $\mathbf{S}_{\mathbf{x},(j,*)}$ denote the j^{th} block of the $\mathbf{x}^{th}$ row vector of $\mathbf{S}$:

$$\mathbf{S}_{\mathbf{x},(j,*)} = (S_{\mathbf{x},(j,1)}, \ldots, S_{\mathbf{x},(j,n_j)}) \in \mathbb{C}^{n_j}$$

Similarly, given $\mathbf{y} \in J$ and $i \in [s]$, we let $\mathbf{S}_{(i,*),\mathbf{y}}$ denote the i^{th} block of the $\mathbf{y}^{th}$ column vector of $\mathbf{S}$:

$$\mathbf{S}_{(i,*),\mathbf{y}} = (S_{(i,1),\mathbf{y}}, \ldots, S_{(i,m_i),\mathbf{y}}) \in \mathbb{C}^{m_i}$$

Suppose $(\mathbf{A}, (N, \boldsymbol{\mu}, \boldsymbol{\nu}, \mathbf{m}, \mathbf{n})$ are as defined above. Then we have the following lemma -

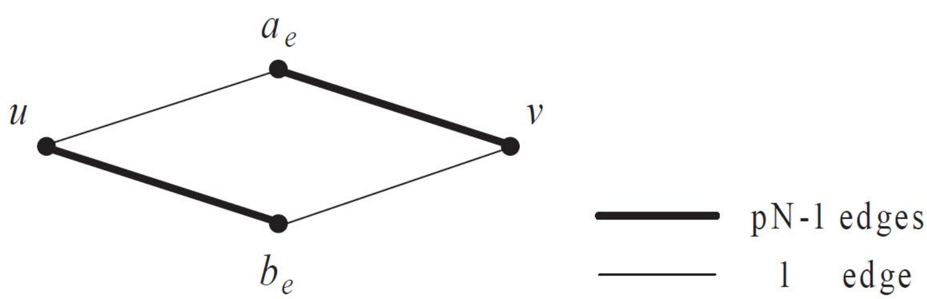

Figure 1: [2] Gadget for constructing graph $G^{[p]}, p \geq 1$.

Lemma 1. $\mathsf{EVAL}(\mathbf{A})$ *is* #P−*hard or the following conditions are satisfied by* $(\mathbf{A}, (N, \boldsymbol{\mu}, \boldsymbol{\nu}, \mathbf{m}, \mathbf{n})$*:*

- *For all two rows* $\mathbf{x}, \mathbf{x}' \in I$*, either* $\mathbf{S}_{\mathbf{x},*} = \omega_N^k \cdot \mathbf{S}_{\mathbf{x}',*}$ *for some integer* k *or for every* $j \in [t]$*,*
$$\langle \mathbf{S}_{\mathbf{x},(j,*)}, \mathbf{S}_{\mathbf{x}',(j,*)} \rangle = 0$$
- *For all two columns* $\mathbf{y}, \mathbf{y}' \in J$*, either* $\mathbf{S}_{*,\mathbf{y}} = \omega_N^k \cdot \mathbf{S}_{*,\mathbf{y}'}$ *for some integer* k *or for every* $i \in [s]$*,*
$$\langle \mathbf{S}_{(i,*),\mathbf{y}}, \mathbf{S}_{(i,*),\mathbf{y}'} \rangle = 0$$

Proof. Assume that $\mathsf{EVAL}(\mathbf{A})$ is not #P-hard. We prove that any two given rows are linearly dependent by ω_N^k for some integer k. The proofs for the columns is similar.

Let $G = (V, E)$ be an undirected graph. For each $p \geq 1$, we construct a new graph $G^{[p]}$ by replacing every edge $e = (u, v) \in E$ with a gadget as shown in Figure 1. More precisely, we add two vertices a_e, b_e for every edge $e \in E$. $G^{[p]} = (V^{[p]}, E^{[p]})$ is defined as follows -

$$V^{[p]} = V \cup \{a_e, b_e | e \in E\}$$

and $E^{[p]}$ contains the following edges for every edge $e = (u, v) \in E$:

- single edges (u, a_e) and (b_e, v).
- $(pN - 1)$ multiple edges between (u, b_e) and (a_e, v).

The construction of $G^{[p]}$ for each $p \geq 1$, gives us an $(m + n) \times (m + n)$ matrix $\mathbf{A}^{[p]}$ such that for all undirected graphs G, we have -

$$Z_{\mathbf{A}^{[p]}}(G) = Z_{\mathbf{A}}(G^{[p]})$$

Hence, we have $\mathsf{EVAL}(\mathbf{A}^{[p]}) \leq \mathsf{EVAL}(\mathbf{A})$ and $\mathsf{EVAL}(\mathbf{A}^{[p]})$ is also not #P-hard. The entries of $\mathsf{EVAL}(\mathbf{A}^{[p]})$ are as follows -

$$A^{[p]}_{(0,\mathbf{u}),(1,\mathbf{v})} = A^{[p]}_{(1,\mathbf{v}),(0,\mathbf{u})} = 0, \qquad \forall \mathbf{u} \in I, \mathbf{v} \in J$$

Thus, $\mathbf{A}^{[p]}$ is a block diagonal matrix with 2 blocks of $m \times m$ and $n \times n$ i.e.

$$\mathbf{A}^{[p]} = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$$

with the upper-left $m \times m$ block having the following entries:

$$\begin{aligned} A^{[p]}_{(0,\mathbf{u}),(0,\mathbf{v})} &= \left(\sum_{\mathbf{a}\in J} A^{[p]}_{(0,\mathbf{u}),(1,\mathbf{a})} (A^{[p]}_{(0,\mathbf{v}),(1,\mathbf{a})})^{pN-1}\right) \left(\sum_{\mathbf{b}\in J} (A^{[p]}_{(0,\mathbf{u}),(1,\mathbf{b})})^{pN-1} A^{[p]}_{(0,\mathbf{v}),(1,\mathbf{b})}\right) \\ &= \left(\sum_{\mathbf{a}\in J} B_{\mathbf{u},\mathbf{a}} (B_{\mathbf{v},\mathbf{a}})^{pN-1}\right) \left(\sum_{\mathbf{b}\in J} (B_{\mathbf{u},\mathbf{b}})^{pN-1} B_{\mathbf{v},\mathbf{b}}\right) \end{aligned}$$

for all $\mathbf{u}, \mathbf{v} \in I$. The factor $B_{\mathbf{u},\mathbf{a}}$ is -

$$B_{\mathbf{u},\mathbf{a}} = \mu_{u_1} \nu_{a_1} S_{\mathbf{u},\mathbf{a}}$$

which leads to -

$$\begin{aligned} \sum_{\mathbf{a}\in J} B_{\mathbf{u},\mathbf{a}} (B_{\mathbf{v},\mathbf{a}})^{pN-1} &= \sum_{\mathbf{a}\in J} \mu_{u_1} \nu_{a_1} S_{\mathbf{u},\mathbf{a}} (\mu_{v_1} \nu_{a_1})^{pN-1} \overline{S_{\mathbf{v},\mathbf{a}}} \\ &= \mu_{u_1} \mu_{v_1}^{pN-1} \sum_{\mathbf{a}\in J} \nu_{a_1}^{pN} S_{\mathbf{u},\mathbf{a}} \overline{S_{\mathbf{v},\mathbf{a}}} \\ &= \mu_{u_1} \mu_{v_1}^{pN-1} \sum_{\mathbf{i}\in[t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle \end{aligned}$$

and

$$\sum_{\mathbf{b}\in J} (B_{\mathbf{u},\mathbf{b}})^{pN-1} B_{\mathbf{v},\mathbf{b}} = \mu_{u_1}^{pN-1} \mu_{v_1} \sum_{\mathbf{i}\in[t]} \nu_i^{pN} \overline{\langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle}$$

As a result, we have

$$A^{[p]}_{(0,\mathbf{u}),(0,\mathbf{v})} = (\mu_{u_1} \mu_{v_1})^{pN} \left| \sum_{\mathbf{i}\in[t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle \right|^2 \tag{1}$$

We can prove a similar result for the lower-right $n \times n$ block. Thus, $\mathbf{A}^{[p]}$ is a non-negative real matrix. Also, if $\mathbf{u} = \mathbf{v}$, then the inner product in equation 1 is equal to n_i.

Since $\mathsf{EVAL}(\mathbf{A}^{[p]})$ is not #P-hard, by the dichotomy theorem of Bulatov and Grohe [1],

$$\left| \sum_{\mathbf{i}\in[t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle \right| = \left\{ \begin{array}{c} 0 \\ \sum_{i\in[t]} n_i \cdot \nu_i^{pN} \end{array} \right.$$

If the vectors $\mathbf{S}_{\mathbf{u},*}$ and $\mathbf{S}_{\mathbf{v},*}$ are linearly dependent, then there must exist an integer $\theta_{\mathbf{u},\mathbf{v}} \in [0, N-1]$ such that $\mathbf{S}_{\mathbf{u},*} = \omega_N^{\theta_{\mathbf{u},\mathbf{v}}} \cdot \mathbf{S}_{\mathbf{v},*}$ (as the entries of $\mathbf{S}$ are all powers of unity - ω_N). Moreover, we need all these $\theta_{\mathbf{u},\mathbf{v}} = \theta$ for all vectors $\mathbf{u}, \mathbf{v}$ to get the equality:

$$\left| \sum_{\mathbf{i} \in [t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle \right| = \left| \sum_{\mathbf{i} \in [t]} \nu_i^{pN} n_i \cdot \omega_N^{\theta_{\mathbf{u},\mathbf{v}}} \right| = \sum_{\mathbf{i} \in [t]} n_i \cdot \nu_i^{pN}$$

and we are done.

On the other hand, assuming that the vectors $\mathbf{S}_{\mathbf{u},*}$ and $\mathbf{S}_{\mathbf{v},*}$ are linearly independent, we have

$$\left| \sum_{\mathbf{i} \in [t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle \right| < \sum_{\mathbf{i} \in [t]} n_i \cdot \nu_i^{pN}, \quad \text{for any } p \geq 1.$$

otherwise it contradicts the assumption that the vectors are linearly independent. The only other possible value of this term is 0 and hence:

$$\sum_{\mathbf{i} \in [t]} \nu_i^{pN} \langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle = 0, \quad \text{for all } p \geq 1.$$

Since $\nu_1 > \nu_2 > \ldots > \nu_t$ is strictly distinct and decreasing, by using the Vandermonde matrix, we have

$$\langle \mathbf{S}_{\mathbf{u},(i,*)} \mathbf{S}_{\mathbf{v},(i,*)} \rangle = 0, \quad \text{for all } i \geq [t].$$

□

This leads to the following corollary:

Corollary 1. *For all $i \in [s]$ and $j \in [t]$, the rank of the $(i,j)^{th}$ block matrix $\mathbf{S}_{(i,*),(j,*)}$ of $\mathbf{S}$ has exactly the same rank as $\mathbf{S}$.*

Proof. We make use of Lemma 1 to establish that $\text{rank}(\mathbf{S}_{(1,*),(1,*)}) = \text{rank}(\mathbf{S})$. Without loss of generality, this is sufficient to prove the corollary.

First, we use Lemma 1 to show that

$$\text{rank} \begin{pmatrix} \mathbf{S}_{(1,*),(1,*)} \\ \mathbf{S}_{(2,*),(1,*)} \\ \vdots \\ \mathbf{S}_{(s,*),(1,*)} \end{pmatrix} = \text{rank}(\mathbf{S})$$

Consider any h $(= \text{rank}(\mathbf{S}))$ rows of $\mathbf{S}$ which are linearly independent. Among them, since any two, $\mathbf{S}_{\mathbf{x},(*,*)}$ and $\mathbf{S}_{\mathbf{y},(*,*)}$, are linearly independent, the two subvectors $\mathbf{S}_{\mathbf{x},(1,*)}$ and $\mathbf{S}_{\mathbf{y},(1,*)}$ are orthogonal. Therefore, the corresponding h rows of the matrix on the left-hand side are pairwise orthogonal and the rank is at least h. Since it cannot be greater than the rank of the matrix $\mathbf{S}$, it must be exactly the same.

Following a similar argument, we can show that

$$\text{rank}(\mathbf{S}_{(1,*),(1,*)}) = \text{rank}\begin{pmatrix} \mathbf{S}_{(1,*),(1,*)} \\ \mathbf{S}_{(2,*),(1,*)} \\ \vdots \\ \mathbf{S}_{(s,*),(1,*)} \end{pmatrix}$$

which completes the proof that $\text{rank}(\mathbf{S}_{(1,*),(1,*)}) = \text{rank}(\mathbf{S})$. □

If $h = \text{rank}(\mathbf{S})$, then by Corollary 1, there must exist h indices $1 \leq i_1 < \ldots < i_h \leq m_1$ and $1 \leq j_1 < \ldots < j_h \leq n_1$ such that the sub-matrix of $\mathbf{S}$ - $\{(1,i_1),\ldots,(1,i_h)\} \times \{(1,j_1),\ldots,(1,j_h)\}$ has full rank h. Without loss of generality we can assume that these indices are the first h indices i.e. $i_k = j_k = k$ for all $k \in [h]$. The matrix $\mathbf{H}$ is used to represent the $h \times h$ matrix: $H_{i,j} = S_{(1,i),(1,j)}$.

By Lemma 1 and Corollary 1, for every index $\mathbf{x} \in I$, there exists two unique integers $j \in [h]$ and $k \in [0 : N-1]$ such that

$$\mathbf{S}_{\mathbf{x},*} = \omega_N^k \cdot \mathbf{S}_{(1,j),*} \tag{2}$$

Similarly, for every index $\mathbf{y} \in J$, there exists two unique integers $j \in [h]$ and $k \in [0 : N-1]$ such that

$$\mathbf{S}_{*,\mathbf{y}} = \omega_N^k \cdot \mathbf{S}_{*,(1,j)} \tag{3}$$

This gives us a partition set of $\{0\} \times I$ and $\{1\} \times J$ respectively:

$$\mathcal{R}_0 = \{R_{(0,i,j),k} | i \in [s], j \in [h], k \in [0 : N-1]\}$$
$$\mathcal{R}_1 = \{R_{(1,i,j),k} | i \in [t], j \in [h], k \in [0 : N-1]\}$$

as follows: For every $\mathbf{x} \in I$, $(0,\mathbf{x}) \in R_{(0,i,j),k}$ if $i = x_1$ and $\mathbf{x}, j, k$ satisfy (2) and for every $\mathbf{y} \in J$, $(1,\mathbf{y}) \in R_{(1,i,j),k}$ if $i = y_1$ and $\mathbf{y}, j, k$ satisfy (3) respectively.

By Corollary 1, we have

$$\bigcup_{k \in [0:N-1]} R_{(0,i,j),k} \neq \phi, \quad \text{for all } i \in [s], j \in [h]$$
$$\bigcup_{k \in [0:N-1]} R_{(1,i,j),k} \neq \phi, \quad \text{for all } i \in [t], j \in [h]$$

Further, we define $(\mathbf{C}, \mathfrak{D})$ and use the Cyclotomic Reduction Lemma (refer previous lectures) to show that

$$\mathsf{EVAL}(\mathbf{C}, \mathfrak{D}) \equiv \mathsf{EVAL}(\mathbf{A})$$

Firstly, we define a matrix $\mathbf{F}$ (of size $sh \times th$) and represent $\mathbf{C}$ as a bipartisation of this matrix.

$$\mathbf{F} = \begin{pmatrix} \mu_1 \mathbf{I} & & & \\ & \mu_2 \mathbf{I} & & \\ & & \ddots & \\ & & & \mu_s \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{H} & \mathbf{H} & \cdots & \mathbf{H} \\ \mathbf{H} & \mathbf{H} & \cdots & \mathbf{H} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H} & \mathbf{H} & \cdots & \mathbf{H} \end{pmatrix} \begin{pmatrix} \nu_1 \mathbf{I} & & & \\ & \nu_2 \mathbf{I} & & \\ & & \ddots & \\ & & & \nu_t \mathbf{I} \end{pmatrix}$$

where $\mathbf{I}$ is the $h \times h$ identity matrix. Alternately, $\mathbf{F}$ is defined as

$$F_{\mathbf{x},\mathbf{y}} = \mu_{x_1}\nu_{y_1}H_{x_2,y_2} = \mu_{x_1}\nu_{y_1}S_{(1,x_2),(1,y_2)}, \text{ for all } \mathbf{x} = (x_1 \in [s], x_2 \in [h]), \mathbf{y} = (y_1 \in [t], y_2 \in [h])$$

The matrix $\mathbf{C}$ is defined as

$$\mathbf{C} = \begin{pmatrix} 0 & \mathbf{F} \\ \mathbf{F}^T & 0 \end{pmatrix}$$

The term $\mathfrak{D}$ is defined as $\mathfrak{D} = \{\mathbf{D}^{[0]}, \ldots, \mathbf{D}^{[N-1]}\}$ is a sequence of N diagonal matrices with the same size of $\mathbf{C}$ and defined by:

$$D^{[r]}_{(0,\mathbf{x}} = \sum_{k=0}^{N-1} |R_{(0,x_1,x_2),k}| \cdot \omega_N^{kr} \text{ and } D^{[r]}_{(1,\mathbf{y}} = \sum_{k=0}^{N-1} |R_{(1,y_1,y_2),k}| \cdot \omega_N^{kr}$$

for all $r \in [0 : N-1], \mathbf{x} = (x_1, x_2) \in [s] \times [h]$ and $\mathbf{y} = (y_1, y_2) \in [t] \times [h]$.

Applying the Cyclotomic Reduction Lemma, we then have

Lemma 2. $\mathsf{EVAL}(\mathbf{A}) \equiv \mathsf{EVAL}(\mathbf{C}, \mathfrak{D})$

Proof. We show that the matrix $\mathbf{A}$ can be generated by the partition (of $[m]$) $\mathcal{R} = \mathcal{R}_0 \cup \mathcal{R}_1$. This is sufficient to prove the lemma (with the aid of Cyclotomic Reduction Lemma).

Let $\mathbf{x}, \mathbf{x}' \in I$, $(0, \mathbf{x}) \in R_{(0,x_1,j),k}$ and $(0, \mathbf{x}') \in R_{(0,x'_1,j'),k'}$. Since $\mathbf{A}$ and $\mathbf{C}$ are bipartisations of $\mathbf{B}$ and $\mathbf{F}$, respectively, we have

$$A_{(0,\mathbf{x}),(0,\mathbf{x}')} = C_{(0,x_1,j),(0,x'_1,j')} = 0$$

As a result, we have

$$A_{(0,\mathbf{x}),(0,\mathbf{x}')} = C_{(0,x_1,j),(0,x'_1,j')} \cdot \omega_N^{k+k'}$$

Let $\mathbf{x} \in I, (0, \mathbf{x}) \in R_{(0,x_1,j),k}, \mathbf{y} \in J, (1, \mathbf{y}) \in R_{(0,y_1,j'),k'}$ for some j, k, j', k'. Then by (2) and (3),

$$A_{(0,\mathbf{x}),(1,\mathbf{y})} = \mu_{x_1}\nu_{y_1}S_{\mathbf{x},\mathbf{y}} = \mu_{x_1}\nu_{y_1}S_{(1,j),\mathbf{y}} \cdot \omega_N^k = \mu_{x_1}\nu_{y_1}S_{(1,j),(1,j')} \cdot \omega_N^{k+k'} = C_{(0,x_1,j),(0,y_1,j')} \cdot \omega_N^{k+k'}$$

Similarly, we can generate the lower-left block of $\mathbf{A}$ from $\mathbf{C}$ using $\mathcal{R}$. Also, the construction of $\mathfrak{D}$ resulted from $\mathcal{R} = \mathcal{R}_0 \cup \mathcal{R}_1$ and hence the lemma follows from the Cyclotomic Reduction Lemma. □

References

[1] Andrei A. Bulatov and Martin Grohe, *The complexity of partition functions*, Theoretical Computer Science, Volume 348(2-3): 148–186, 2005.

[2] Jin-Yi Cai, Xi Chen and Pinyan Lu, *Graph Homomorphisms with Complex Values: A Dichotomy Theorem*, In Proceedings of the 37th International Colloquium Conference on Automata, Languages and Programming (ICALP'10), Lecture Notes in Computer Science, Volume 6198: 275–286. Springer, 2010.