ADVANCED STUDY ON SDP APPLICATIONS

ZHIRONG LI ZHIRONGLI@GMAIL.COM

1. Geometric Representation of Graphs

1.1. **Unit Distance Representation of Graphs.** We may want to map each node of a graph $G = (V, E)$ to a point in d-dimensional real Euclidean space such that for every edge in G, the distance in this d-dimensional space is always equal to a unit distance. I.e., $\|v^{(i)} - v^{(j)}\|_2 = 1, \forall \{i, j\} \in E$. Petersen graph can be drawn in a plane and have the unit distance representation. We may interest in finding the answer for questions like for any given d, what is the radius of the smallest Euclidean ball containing such unit distance representation of graph G. In engineering, we may also be given the prescribed distances $w_{ij} \in \mathbb{R}^{\mathcal{S}}_{+}$ for $\mathcal{S} \subseteq E$ and the task is to find the smalled d such that those prescribed distance requirements can be satisfied. We have the following nice theorem stating the property that a special SDP problem formulation can give us the radius of the smallest Euclidean ball containing unit distance representation of G in $\mathbb{R}^{|V|}$.

Theorem 1.1. *Suppose graph $G = (V, E)$ is given. The optimal solution for the following SDP exists, and it is attained with the optimal value equal to the square of the radius of the smallest Euclidean ball containing the unit distance representation of G.*

$$
\begin{aligned}
t(G) := \quad & \min t \\
st. \quad & diag(X) - t\bar{e} \leq 0 \\
& X_{ii} - 2X_{ij} + X_{jj} = 1 \quad \forall \{i, j\} \in E \\
& X \in \Sigma^{|V|}_{+}
\end{aligned}
$$

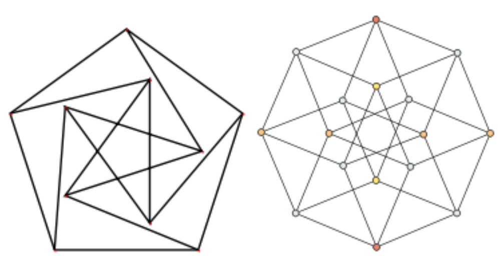

Figure 1.1. Petersen Graph and Hypercube $Q4$ in its unit distance representation form

Proof. We can check the dual problem of the above. Let us define

$$A^{\{i,j\}} := e_i e_i^T + e_j e_j^T - \left(e_i e_j^T + e_j e_i^T\right) = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & 1 & 0 & -1 & \vdots \\ \vdots & 0 & \cdots & 0 & \vdots \\ \vdots & -1 & 0 & 1 & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}_{|V|\times|V|} , \forall \{i,j\} \in E.$$

The dual problem can be obtained by observing the Lagrangian multipliers of the original (P) problem, and apply the first order necessary optimality conditions if we are unsure about what the exact form it should look like. We will associate KKT multiplier z_{ij} for each unit distance constraint, $y \in \mathbb{R}_-^{|V|}$ for the inequality constraint for the radius of the Euclidean ball. The Lagrangian of (P) is

$$\mathcal{L}(X, t, z_{i,j}, -y) = t - \sum_{\{i,j\}\in E} z_{ij}(X_{ii} - 2X_{ij} + X_{jj} - 1) - (-y)^T(t\bar{e} - diag(X)).$$

Therefore $\nabla\mathcal{L}_X(X, t, z_{ij}, -y) = 0$ implies $0_{|V|\times|V|} = \sum_{\{i,j\}\in E} z_{ij} A^{\{i,j\}} + Diag(y)$. $\nabla\mathcal{L}_t(X, t, z_{ij}, -y) = 0$ indicates $1 = -\bar{e}^T y$ (free variable t indicates equality constraint). $\nabla\mathcal{L}_{z_{ij}} = 0$ and $\nabla\mathcal{L}_{-y} = 0$ can show us $X_{ii} - 2X_{ij} + X_{jj} = 1$ and $t\bar{e} - diag(X) - S = 0$ (if add slack variable S) as expected. By the KKT conditions ,we know KKT multipliers$-y$ for inequality constraints must be nonnegative thus we have $y \leq 0$ (or $\leq$ constraint corresponds to non-positive variable y). By the KKT complementarity condition, we can derive $(-y)^T(t\bar{e} - diag(X) = S) = 0$ since it will guarantee that the Lagrangian of (P) will have the same optimal objective value as the original problem (P). Therefore for the Lagrangian dual, $\mathcal{L}(z_{ij}, -y) = \sum_{\{i,j\}\in E} z_{ij}$(similar to LP) , and its full dual can be formulated as

$$\begin{aligned} &\max \textstyle\sum_{\{i,j\}\in E} z_{ij} \\ st. \quad &Diag(y) + \textstyle\sum_{\{i,j\}\in E} z_{ij} A^{\{i,j\}} \preceq 0 \quad \forall \{i,j\} \in E \\ &-\bar{e}^T y = 1 \\ &y \leq 0 \end{aligned}$$

The first inequality follows from the rule "nonnegative variables X corresponds to $\preceq$ constraint" and the coefficient for matrix variable X in the objective function is $0_{|V|\times|V|}$. If we add slack matrix variable $S \in \Sigma_+^{|V|}$,the inequality constraint can be changed to its equality form. In general, the dual constraints can be derived from the Lagrangian of the primal by taking first derivative on it to get the exact form. Pay attention to that dual SDP has Slater point $\bar{y} = -\frac{1}{|V|}, \bar{z}_{ij} = 0$, and primal SDP has Slater point $\bar{X} = \frac{1}{2}I, t = \frac{1}{2} + \epsilon$ for any $\epsilon > 0$. By the corollary of the strong duality theorem, both the primal and dual SDP problem have optimal solution and the optimal objective values coincide. Assume we already have a unit distance representation of G with the smallest radius $r := \max\left\{\|v^{(i)}\|_2\right\}$where $\left\{v^{(i)} \in \mathbb{R}^{|V|} : i \in V\right\}$is the representation itself. We can define $B^T := \left[v^{(1)}, \cdots v^{(|V|)}\right]$ then set $\bar{X} := BB^T$. It is obvious $\bar{X} \succeq 0$ and $\bar{X}_{ij} = \left\langle v^{(i)}, v^{(j)}\right\rangle \forall i, j$. Therefore, $diag(\bar{X}) \leq r^2\bar{e}$ and $\bar{X}_{ii} - 2\bar{X}_{ij} + \bar{X}_{jj} = 1$. So $(\bar{X}, r^2)$ is feasible in the primal SDP problem with the objective value r^2. Similarly, let $(\bar{X}, \bar{t})$ be the optimal solution of the primal SDP, we can always Cholesky decompose $\bar{X} = BB^T$. Let $v^{(i)}$ denote the i-th column of B^T, we can claim $\left\{v^{(i)} : i \in V\right\}$form a unit distance representation of G, and $\|v^{(i)}\|_2 \leq \sqrt{\bar{t}}$. Thus completes the proof of above theorem. That is, the optimal value of the SDP is precisely the square-root of the smallest Euclidean ball in $\mathbb{R}^{|V|}$ containing a unit distance representation of G. □

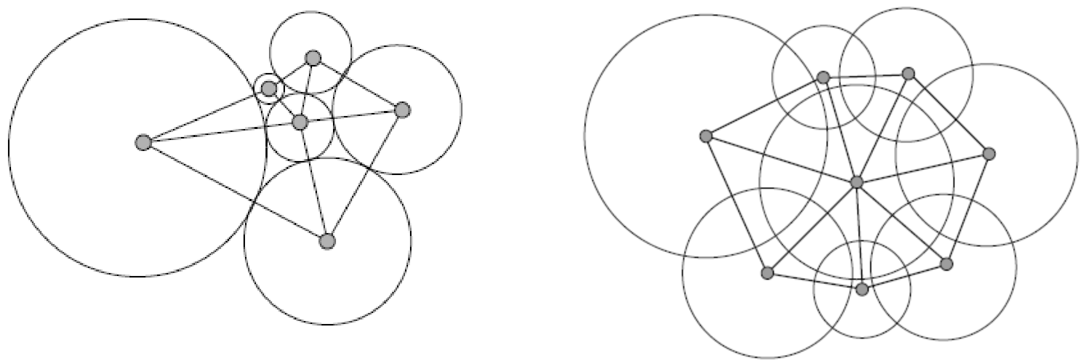

FIGURE 1.2. Koebe-Andreev-Thurston Circle Packing Theorem

1.2. **Hypersphere Representation of Graphs.** It is very similar to unit distance representation of graph except that we require every node lies exactly on the hypersphere. The problem is given by

$$\begin{aligned} t(G) := \quad & \min t \\ st. \quad & diag(X) = t\bar{e} \\ & X_{ii} - 2X_{ij} + X_{jj} = 1 \quad \forall \{i,j\} \in E \\ & X \in \Sigma_+^{|V|} \end{aligned}$$

and we also are provided with the following theorem.

Theorem 1.2. *Let $G = (V, E)$ be given. The optimal solution for above SDP problem exists, and the optimal value is attained with the optimal value equal to the square of the radius of the smallest Euclidean ball containing a smallest hypersphere representation of G.*

Koebe-Andreev-Thurston's Circle Packing Theorem states that every planar graph can be represented in such a way that its nodes correspond to disjoint disks which touch if and only if the corresponding nodes are adjacent. Actually there exists a corresponding representation of the dual graph G^* by disks such that intersecting edges of G and G^* are represented by disks whose boundaries intersect orthogonally as shown in the following figure.

1.3. **Orthonormal Representations of Graphs.** $\left\{u^{(i)} \in \mathbb{R}^d : i \in V\right\}$ is called an orthonormal representation of graph if $\|u^{(i)}\|_2 = 1$ for $\forall i \in V$ and $\left\langle u^{(i)}, u^{(j)} \right\rangle = 0$ for $\forall \{i,j\} \in \bar{E}$. We will find there is some correspondence between orthonormal representation of G and hypersphere representation of $\bar{G}$. I.e., Suppose $\left\{v^{(i)} \in \mathbb{R}^d : i \in V\right\}$ is the smallest hypersphere representation of $\bar{G}$ with radius t. Then we have $\|v^{(i)}\|_2^2 = t$, $\forall i \in V$ and $\left\langle v^{(i)}, v^{(j)} \right\rangle = \frac{\|v^{(i)}\|_2^2 + \|v^{(j)}\|_2^2 - \|v^{(i)} - v^{(j)}\|_2^2}{2} = \frac{2t-1}{2}, \forall \{i,j\} \in \bar{E}$. Since $\bar{X} := \frac{1}{2}I, \bar{t} := \frac{1}{2}$ is a feasible point in the SDP formulation of hypersphere representation of G, we can claim $t \leq \frac{1}{2}$. We can give an orthonormal representation of G in lifted space $\mathbb{R}^{d+1}$ based on $\left\{v^{(i)} \in \mathbb{R}^d : i \in V\right\}$. Assume the zeroth coordinate is set to $\sqrt{\frac{1}{2} - t}$ and we embed the original graph G in the hyperplane $\left\{x \in \mathbb{R}^{d+1} : x_0 = \sqrt{\frac{1}{2} - t}\right\}$.

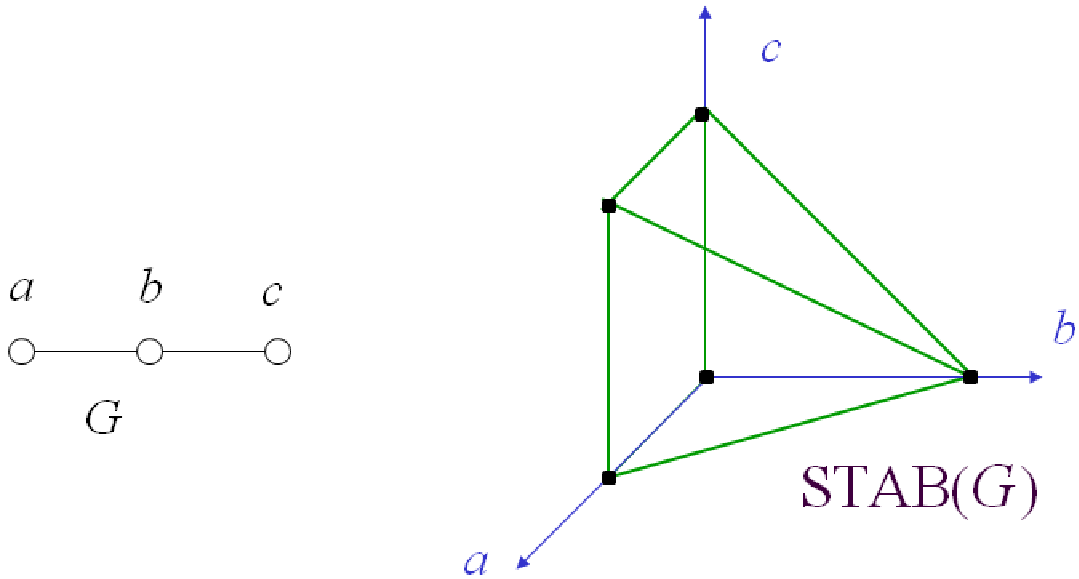

FIGURE 1.3. Stable Set Polytope

Let $u^{(i)} := \sqrt{2}\begin{bmatrix} \sqrt{\frac{1}{2}-t} \\ v^{(i)} \end{bmatrix}, \forall i \in V$, then we have $\|u^{(i)}\|_2^2 = 2(\frac{1}{2}-t+t) = 1$ and $\left\langle u^{(i)}, u^{(j)} \right\rangle = 2(\frac{1}{2}-t+\frac{2t-1}{2}) = 0, \forall \{i,j\} \in \bar{E}$ as desired. Similarly, every orthonormal representation of $\bar{G}$ gives a hypersphere representation of G as well by letting $\left\{u^{(i)} \in \mathbb{R}^{d+1} : i \in V\right\}$and defining $v^{(i)} := \frac{1}{\sqrt{2}}u^{(i)}, \forall i \in V$. This is because $\|v^{(i)} - v^{(j)}\|_2^2 = \frac{1}{2}\|u^{(i)}\|_2^2 + \frac{1}{2}\|u^{(j)}\|_2^2 = 1$ and $\left\langle v^{(i)}, v^{(j)} \right\rangle = 0, \forall \{i,j\} \in \bar{\bar{E}} = E$. Therefore $\left\{v^{(i)} : i \in V\right\}$ is a hypersphere representation of G contained in $B(0, \sqrt{\frac{1}{2}})$.

Orthonormal representation of graphs have a deep connection to stable set problem. A subset $\mathcal{S} \subseteq V$ is a stable set of G if for every $\{i,j\} \in E$ at most one of i, j is in $\mathcal{S}$. The stability number of G is the cardinality of the largest cardinality stable set in G and is denoted by $\alpha(G) := \max\{|\mathcal{S}| : \mathcal{S}\, is\, a\, stable\, set\, in\, G\}$. However, computing $\alpha(G)$ is an $\mathcal{NP}$-hard problem. Moreover, it is even hard to approximate $\alpha(G)$ in polynomial time, that is no constant ratio, polynomial time approximation algorithm exists unless $\mathcal{P} = \mathcal{NP}$. We can define a stable set polytope by taking convex hull on all incidence vectors $x \in \{0,1\}^{|V|}$ of each stable set.

$$STAB(G) := conv\left\{x \in \{0,1\}^{|V|} : x\, is\, an\, incidence\, vector\, of\, a\, stable\, set\, in\, G\right\}.$$

We can define a fractional stable set polytope as

$$FRAC(G) := \left\{x \in [0,1]^{|V|} : x_i + x_j \leq 1, \forall \{i,j\} \in E\right\}$$

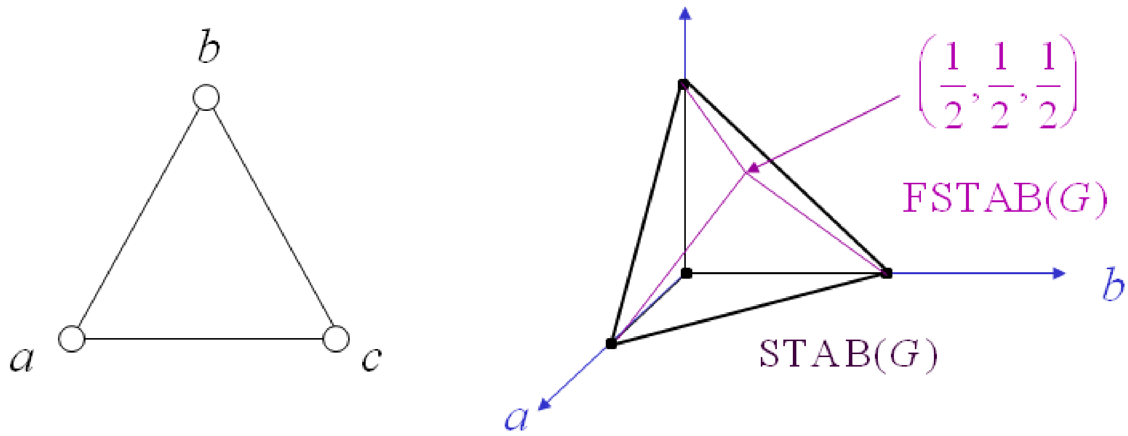

FIGURE 1.4. Fractional Stable Set Polytope ($STAB(G) \subseteq FRAC(G)$)

and then we can represent stable set polytope as $STAB(G) = conv\left(FRAC(G) \cap \{0,1\}^{|V|}\right)$. For any clique $\mathcal{C}$ in G, the clique inequality says that

$$\sum_{i\in\mathcal{C}} x_i \leq 1.$$

We can define a clique matrix $A_{clq}(G)$ whose rows are the incidence vectors of cliques in G. Thus A_{clq} is a $0,1$ matrix with $|V|$ columns and as many rows as the number of cliques in G including singletons and edges. Then we can have the clique polytope of G as given by

$$CLQ(G) := \left\{x \in \mathbb{R}_+^{|V|} : A_{clq}(G)x \leq \bar{e}\right\}.$$

We now turn our focus to an orthonormal representation of G by $\left\{u^{(i)} \in \mathbb{R}^d : i \in V\right\}$. Then for any $c \in \mathbb{R}^d$ such that $\|C\|_2 = 1$, the inequality

$$\sum_{i\in V}\left(c^T u^{(i)}\right)^2 x_i \leq 1$$

is called an orthonormal representation constraint. Based on above constraint, we can define a compact, convex set with nonempty interior theta body of G as

$$TH(G) := \left\{x \in \mathbb{R}_+^{|V|} : x\, satisfies\, all\, orthonormal\, representation\, constraints\, for\, G\right\}.$$

Clearly $0 \in TH(G)$ and $e_i \in TH(G)$ (since $\sum_{i\in V}\left(c^T u^{(i)}\right)^2 x_i = \left(c^T u^{(i)}\right)^2 \leq \|c\|_2^2 \cdot \|u^{(i)}\|_2^2 = 1$) thus it is not empty. We have the following theorem indicating the relationships between stable set polytope, theta body polytope, clique polytope and fractional stable set polytope.

Theorem 1.3. *For every graph $G = (V, E)$, we have $STAB(G) \subseteq TH(G) \subseteq CLQ(G) \subseteq FRAC(G)$.*

Proof. We prove $STAB(G) \subseteq TH(G)$ first. Suppose $\mathcal{S} \subseteq V$ is a stable set in G. Let $\left\{u^{(i)} : i \in V\right\}$ be an arbitrary orthonormal representation of G and let c be a vector such that $\|c\|_2 = 1$.Let $\mathcal{U}_{\mathcal{S}}^T$ be an $|\mathcal{S}| \times |V|$ matrix with columns $u^{(i)}$. Then $\sum_{i\in\mathcal{S}}\left(c^T u^{(i)}\right)^2 = \|\mathcal{U}_{\mathcal{S}} c\|_2^2 \leq \|c\|_2^2 = 1$ thus for every stable set in G its incidence vector x satisfies $\sum_{i\in V}\left(c^T u^{(i)}\right)^2 x_i \leq 1$ for every orthonormal representation of G

and every unit vector c. Therefore theta body of G is a relaxation of stable set polytope. Secondly, let's prove $TH(G) \subseteq CLQ(G)$. Let $\mathcal{C} \subseteq V$ be a clique in G. We can always have a set of orthonormal vectors $\left\{u^{(i)} : i \in V \backslash \mathcal{C}\right\} \cup \{c\}$ and set $u^{(i)} := c, \forall i \in \mathcal{C}$. Notice that $\left\{u^{(i)} : i \in V\right\}$forms an orthonormal representation of G and the corresponding orthonormal representation constraint is given by

$$\sum_{i \in V} \left(c^T u^{(i)}\right)^2 x_i \le 1 \Leftrightarrow \sum_{i \in \mathcal{C}} \left(c^T c\right)^2 x_i \le 1 \Leftrightarrow \sum_{i \in \mathcal{C}} x_i \le 1$$

Thus every clique inequality indeed is a direct result of an orthonormal representation constraint. So we have $TH(G) \subseteq CLQ(G)$. $CLQ(G) \subseteq FRAC(G)$ is obvious since $FRAC(G)$ is defined by the non-negativity constraints and clique constraints with only one or two nodes. $\square$

We can approximate any linear function over $FRAC(G)$ in polynomial time while the relaxation is usually very week. It is $\mathcal{NP}$-hard to find the largest clique in a graph and linear optimization over $CLQ(G)$ is also $\mathcal{NP}$-hard. The next theorem shows optimizing over $TH(G)$ is tractable.

Suppose we are given a graph $G = (V, E)$ and a weight vector $w \in \mathbb{R}_+^{|V|}$. We have the following definition

$$\theta(G, w) := \max \left\{w^T x : x \in TH(G)\right\}$$

and $W \in \Sigma^{|V|}$ is defined component-wise as $W_{ij} := \sqrt{w_i w_j}$.

Theorem 1.4. *Let $G = (V, E)$ be a given graph with a weight weight vector $w \in \mathbb{R}_+^{|V|}$. Then TFAE:*
(i) $\theta(G, w)$
(ii) $\min_{all\, orth.\, representations} \max_{i \in V} \left\{\frac{w_i}{\left(c^T u^{(i)}\right)^2}\right\}$
(iii) $\min \left\{\eta : diag(S) = 0;\ S_{ij} = 0, \forall \{i, j\} \in \bar{E};\ \eta I - S \succeq W\right\}$
(iv) $\max \left\{\langle W, X\rangle : X_{ij} = 0, \forall \{i, j\} \in E;\ \langle I, X\rangle = 1; X \succeq 0\right\}$

Proof. First let's assume $w \neq 0$ otherwise the proof is trivial since all optimal values are equal to 0. First, let's prove (i)⇒(ii). We know that $TH(G)$ is nonempty and compact, and $w^T x$ is continuous over $TH(G)$ thus maximizer $\exists \bar{x} \in TH(G)$ such that $w^T \bar{x} = \theta(G, w)$. Let $\left\{\bar{u}^{(i)} : i \in V\right\}$be an orthonormal representation of G and $\bar{c}$ be the unit vector to attain the min-max in (ii), for every i, $\bar{c}^T \bar{u}^{(i)} \neq 0$. We will establish the following fact

$$\begin{aligned} (i) &= \theta(G, w) = \sum_{i \in V} w_i \bar{x}_i = \sum \frac{w_i}{\left(\bar{c}^T \bar{u}^{(i)}\right)^2} \left(\bar{c}^T \bar{u}^{(i)}\right)^2 \bar{x}_i \\ &\le \left(\max_{i \in V} \left\{\frac{w_i}{\left(\bar{c}^T \bar{u}^{(i)}\right)^2}\right\}\right) \underbrace{\sum_{i \in V} \left(\bar{c}^T \bar{u}^{(i)}\right)^2 \bar{x}_i}_{\le 1} \le \max_{i \in V} \left\{\frac{w_i}{\left(\bar{c}^T \bar{u}^{(i)}\right)^2}\right\} = (ii) \end{aligned}$$

Secondly, we should notice that (iii) and (iv) are primal (iv)-dual (iii) SDP's. Notice that $\bar{X} := \frac{1}{|V|} I$ and $\bar{S} := 0, \bar{\eta} := \eta_1(W) + 1$for sufficient large $\eta_1$yield Slater point for primal and dual SDP, respectively. By the corollary of strong duality theorem, (iii)=(iv).

Thirdly, we prove (ii)≤(iii). Suppose $\bar{S}$ and $\bar{\eta}$ yield an optimal solution for (D) SDP problem. That is, $\bar{\eta} = (iii)$. Then we have

$$\bar{\eta} \underbrace{\ge}_{weak\, duality} \left\langle W, \bar{X}\right\rangle = \frac{1}{|V|} \bar{e}^T \underbrace{w}_{>0} > 0.$$

Let $Y \in \mathbb{R}^{|V| \times |V|}$ be the Cholesky decomposition matrix of $YY^T = \bar{\eta} I - \bar{S} - W \succeq 0$. Let $v^{(i)}$ denote the ith column of Y^T as before. Then we get

$$\left\langle v^{(i)}, v^{(i)} \right\rangle = \left(\bar{\eta} I - \bar{S} - W\right)_{ii} = \bar{\eta} - w_i, \forall i \in V$$

and

$$\left\langle v^{(i)}, v^{(j)} \right\rangle = -\sqrt{w_i w_j}, \forall \{i,j\} \in \bar{E}.$$

We can claim Y is not full rank. Otherwise, we could end up with $\bar{\eta} I - \bar{S} - W \succ 0$ and there exists a small enough $\epsilon > 0$ such that the pair $\bar{S}, (\bar{\eta} - \epsilon)$ are still feasible in D which contradicts $\bar{\eta}$ is the optimal value of (D). Since $rank(Y) < |V|$, there must exist $c \in \mathbb{R}^V$such that $\|c\|_2 = 1$ and $Yc = 0$(i.e., $\left\langle c, v^{(i)} \right\rangle = 0, \forall i \in V$). We can further define $u^{(i)} := \frac{1}{\sqrt{\bar{\eta}}} \left(\sqrt{w_i} c + v^{(i)}\right), \forall i \in V$. It is clear $\|u^{(i)}\|_2^2 = \frac{1}{\bar{\eta}} \left(w_i + \bar{\eta} - w_i\right) = 1$ for every $i \in V$ and $\left\langle u^{(i)}, u^{(j)} \right\rangle = \frac{1}{\bar{\eta}} \left(\sqrt{w_i w_j} - \sqrt{w_i w_j}\right) = 0$ for $\forall \{i,j\} \in \bar{E}$. Therefore $\left\{u^{(i)} : i \in V\right\}$ gives an orthonormal representation of G. Thus

$$(ii) \leq \max_{i \in V, w_i > 0} \left\{ \frac{w_i}{\left(c^T u^{(i)}\right)^2} \right\} = \frac{w_i}{\frac{w_i}{\bar{\eta}}} = \bar{\eta} = (iii).$$

Finally let's prove $(iv) < (i)$. Let X^* be an optimal solution of (P). Then there exists $Y \in \mathbb{R}^{|V| \times |V|}$ such that $YY^T = X^*$. Let $y^{(i)}$ denote the ith column of Y^Tas before and let $u^{(i)} := \frac{y^{(i)}}{\|y^{(i)}\|_2}, \forall i \in V$ such that $y^{(i)} \neq 0$. For the remaining $i \in V$ such that $y^{(i)} = 0$, define $\{u^{(i)}\}$as an orthonormal basis for $span\left\{u^{(i)} : y^{(i)} \neq 0\right\}^{\perp}$. Now $\left\{u^{(i)} : i \in V\right\}$is an orthonormal representation of $\bar{G}$ ($\left\langle u^{(i)}, u^{(j)} \right\rangle = 0, \forall \{i,j\} \in E$ since $\left\langle y^{(i)}, y^{(j)} \right\rangle = 0, \forall (i,j) \in E$). We define $c := \frac{1}{\sqrt{\langle W, X^* \rangle}} Y^T \begin{bmatrix} \sqrt{w_1} \\ \sqrt{w_2} \\ \vdots \\ \sqrt{w_{|V|}} \end{bmatrix}$ then

$$\|c\|_2^2 = \frac{1}{\langle W, X^* \rangle} \begin{bmatrix} \sqrt{w_1} & \sqrt{w_2} & \cdots & \sqrt{w_{|V|}} \end{bmatrix} \underbrace{YY^T}_{=X^*} \begin{bmatrix} \sqrt{w_1} \\ \sqrt{w_2} \\ \vdots \\ \sqrt{w_{|V|}} \end{bmatrix} = \frac{\langle W, X^* \rangle}{\langle W, X^* \rangle} = 1.$$

Define $U^T := \begin{bmatrix} u^{(1)} & u^{(2)} & \cdots & u^{(|V|)} \end{bmatrix}$and we have the following two lemmas to complete the proof.

Lemma 1: $Uc \odot Uc \in TH(G)$.

To prove this lemma, we simply show the vector $Uc \odot Uc$ satisfies all orthonormal representation constraints for graph G. Let $\left\{v^{(i)} : i \in V\right\}$be an arbitrary orthonormal representation of G and d be an arbitrary unit vector. We have the following two facts $\left\langle cd^T, cd^T \right\rangle = \left(c^T c\right)\left(d^T d\right) = 1$ and

$$\left\langle u^{(i)} \left(v^{(i)}\right)^T, u^{(j)} \left(v^{(j)}\right)^T \right\rangle = \left\langle u^{(i)}, u^{(j)} \right\rangle \left\langle v^{(i)}, v^{(j)} \right\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Thus $\sum_{i \in V} \left(d^T v^{(i)}\right)^2 \left(c^T u^{(i)}\right)^2 = \sum_{i \in V} \left(\left\langle cd^T, u^{(i)} \left(v^{(i)}\right)^T \right\rangle\right) \leq \left\langle cd^T, cd^T \right\rangle = 1$.

Lemma 2: With the above definitions, we have

$$\langle W, X^* \rangle = \left[\sum_{i \in V} \sqrt{w_i} \|y^{(i)}\|_2 \left(c^T u^{(i)}\right)\right]^2$$

To prove lemma 2, we observe that

$$\langle W, X^* \rangle \leq \underbrace{\left(\sum_{i \in V} \|y^{(i)}\|_2^2 \right)}_{= tr(X^*) = \langle I, X^* \rangle = 1} \left[\sum_{i \in V} w_i \left(c^T u^{(i)} \right)^2 \right] = \sum_{i \in V} w_i \left(c^T u^{(i)} \right)^2 \underbrace{\leq}_{Uc \odot Uc \in TH(G)} \theta(G, w)$$

□

A graph is called perfect if for every node induced subgraph H of G, the clique number of H and the chromatic number of H coincide: $\omega(H) = \chi(H)$. An odd-hole is a chord-less odd cycle of length at least 5. An odd-anti-hole is the complement of an odd-hole.

Theorem 1.5. *For every graph $G = (V, E)$, $[TH(G)]^* \cap \mathbb{R}_+^{|V|} = TH(\bar{G})$. That is , the polar of $TH(G)$ when restricted to nonnegative orthant, coincides with the $TH(\cdot)$ set of the complement of G.*

We also have the equivalent representation of $TH(G)$ as the projection of the feasible region of a tractable SDP problem:

Theorem 1.6. *Let $G = (V, E)$ be an undirected graph. Then*

$$TH(G) = \left\{ x \in \mathbb{R}^{|V|} : \begin{pmatrix} 1 \\ x \end{pmatrix} = Ye_0;\, Y_{ij} = 0, \forall \{i, j\} \in E;\, Ye_0 = diag(Y);\, Y \in \Sigma_+^{\{0\} \cup V} \right\}$$

Proof. First let's prove $\bar{x} \in TH(G)$ implies there exist $c \in \mathbb{R}^{|V|}$ with $\|c\|_2 = 1$ and an orthonormal representation of $\bar{G}$, $\left\{ u^{(i)} : i \in V \right\}$ such that $\bar{x}_i = \left(c^T u^{(i)} \right)^2$for every $i \in V$. Then for $i, j \in V$ consider $Y_{ij} := \sqrt{\bar{x}_i \bar{x}_j} \left\langle u^{(i)}, u^{(j)} \right\rangle$. □

We also have the nice theorem about node-symmetric graph.

Theorem 1.7. *Let $G = (V, E)$ be an undirected graph. We say that G is node-symmetric(vertex-transitive) if the automorphism group of G acts transitively on V. If G is such a node-symmetric graph, then $\theta(G)\theta(\bar{G}) = |V|$ where $\theta(G) := \theta(G, \bar{e})$.*

The well known sandwich theorem states that:

Theorem 1.8. *For every graph G*

$$\begin{array}{rcl} \theta(G) & = \max \left\langle \bar{e}\bar{e}^T, X \right\rangle & = \min t \\ & X_{ij} = 0, \forall \{i, j\} \in E & diag(Z) = (t-1)\bar{e} \\ & \langle I, X \rangle = 1 & Z_{ij} = -1, \forall \{i, j\} \in \bar{E} \\ & X \succeq 0 & Z \succeq 0 \end{array}$$

Moreover,

$$\alpha(G) \leq \Theta(G) \leq \theta(G) \leq \chi(\bar{G})$$

for all graphs G where $\Theta(G) := \lim_{k \to \infty} \left[\alpha(G^k) \right]^{1/k}$.The products of undirected graphs $G = (V, E), H = (W, F)$ is defined as

$$G \otimes H := (V(G \otimes H), E(G \otimes H))$$

where

$$V(G \otimes H) := V \times W$$

and

$$E(G \otimes H) := \left\{ \left\{ \{i, u\}, \{j, v\} \right\} : (\{i, j\} \in E \cap \{u, v\} \in F) \cup (\{i, j\} \in E \cap u = v) \cup (i = j \cap \{u, v\} \in F) \right\}.$$

2. Lift and Project Method

For combinatorial optimization, we probably face the problem like maximizing or minimizing $c^T x$ over $F \subseteq \{0,1\}^d$ where $c \in \mathbb{R}^d$. The integer-valued feasible region F can also be described in terms of a polytope $P \subseteq [0,1]^d$ such that $F = P \cap \{0,1\}^d$. We are interested in doing optimization over the convex hull of F with the hope to get a good result after this type of relaxation. Let P_I denote the relaxed feasible region $P_I := conv(F) = conv\left(P \cap \{0,1\}^d\right)$ as described above. The original integer programming problem is converted to solving a linear programming problem. In some cases, we may have a compact representation of exponentially many inequality constraints in a lifted space with polynomially number of inequalities. The benefit we can get from lies in that we could have more options if we take certain liftings to Σ^d to get a tight SDP relaxations for some very hard problems. The lift and project method that will be introduced in this section starts from P and recursively generates tighter and tighter relaxation of P_I. Moreover, this relaxation process will converge to P_I. The name for this method comes from the procedure to embed the original feasible region to a higher dimension hyperplane then project back onto $\mathbb{R}^d$. Let $\mathbb{Q} := cone\left\{x \in \mathbb{R}^{d+1} : x \in \{0,1\}^{d+1}, x_0 = 1\right\}$ and $\mathcal{K} \subset \mathbb{Q}$ denote a polyhedral cone obtained from a polytope $P \subseteq [0,1]^d$ via homogenization using a new variable x_0. That is, suppose we have P defined as $P := \left\{x \in \mathbb{R}^d : Ax \le b, 0 \le x \le \bar{e}\right\}$, then we get $\mathcal{K} := \left\{\begin{bmatrix} x_0 \\ x \end{bmatrix} \in \mathbb{R}^{d+1} : Ax \le x_0 b, 0 \le x \le x_0 \bar{e}\right\}$. Notice that we embed P x_0-times larger in every hyperplane in $\mathbb{R}^{d+1}$ with the zeroth coordinate equal to x_0 since $\mathcal{K}$ is a conic representation of P. Similarly, let $\mathcal{K} \in \mathbb{R}^{d+1}$ be the convex cone as above and the element of $\mathcal{K}$ be denoted as $y = \begin{bmatrix} y_0 & y_1 & \cdots & y_d \end{bmatrix}^T =: \begin{bmatrix} x_0 \\ x \end{bmatrix}$ where $x \in \mathbb{R}^d$ such that $\left\{x \in \mathbb{R}^d : \begin{bmatrix} 1 \\ x \end{bmatrix} \in \mathcal{K}\right\} \subseteq [0,1]^d$. Thus $\mathcal{K}_I$ is also a homogenization of P_I.

2.1. Lovasz-Schrijver Procedure.

Lovasz-Schrijver Procedure. We have a SDP representation of theta body of G: Let $G = (V, E)$ be an undirected graph, then

$$TH(G) = \left\{x \in \mathbb{R}^{|V|} : \begin{pmatrix} 1 \\ x \end{pmatrix} = Ye_0;\ Y_{ij} = 0, \forall \{i,j\} \in E;\ Ye_0 = diag(Y);\ Y \in \Sigma_+^{\{0\}\cup V}\right\}$$

The constraint $Y_{ij} = 0, \forall \{i,j\} \in E$ is not quite general thus we can replace this constraint with a more general geometric object such as $FRAC(G)$'s conic form in $[0,1]^{\{0\}\cup V}$. I.e.,

$$Ye_i \in cone(FRAC(G)) := \left\{\begin{bmatrix} x_0 \\ x \end{bmatrix} \in \mathbb{R}^{\{0\}\cup V} : x_i + x_j \le x_0, \forall \{i,j\} \in E,\ 0 \le x \le x_0 \bar{e}\right\}$$

We may also strengthen our relaxation by adding $Y(e_0 - e_i) \in cone(FRAC(G))$ since when $rank(Y) = 1$, i.e.,

$$x \in P \cap \{0,1\}^{|V|}, Y = \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 & x^T \end{bmatrix},$$

$Ye_0 = \begin{bmatrix} 1 \\ x \end{bmatrix}$, $Y(e_0 - e_i) = \begin{bmatrix} 1 - x_i \\ x - x_i x \end{bmatrix} = (1 - x_i)\begin{bmatrix} 1 \\ x \end{bmatrix}$ both belong to the cone generated by the fractional stable set polytope. Let $P, \mathcal{K}$ be as above, we will define three operators $M_+(\mathcal{K})$, $M(\mathcal{K})$ and $M_0(\mathcal{K})$ (they are all convex cones) and corresponding $N_+(\mathcal{K})$, $N(\mathcal{K})$ and $N_0(\mathcal{K})$ respectively.

$$\begin{aligned}\underbrace{M_+(\mathcal{K})}_{\cap} &:= \left\{Y \in \Sigma_+^{1+d} : diag(Y) = Ye_0;\ Ye_i, Y(e_0 - e_i) \in \mathcal{K}, \forall i \in \{1, 2, \cdots d\}\right\} \\ \underbrace{M(\mathcal{K})}_{\cap} &:= \left\{Y \in \Sigma^{1+d} : diag(Y) = Ye_0;\ Ye_i, Y(e_0 - e_i) \in \mathcal{K}, \forall i \in \{1, 2, \cdots d\}\right\} \\ M_0(\mathcal{K}) &:= \left\{Y \in \mathbb{R}^{(1+d)\times(1+d)} : diag(Y) = Ye_0 = Y^T e_0;\ Ye_i, Y(e_0 - e_i) \in \mathcal{K}, \forall i \in \{1, 2, \cdots d\}\right\}\end{aligned}$$

and

$$\begin{aligned}N_+(\mathcal{K}) &:= \{Ye_0 : Y \in M_+(\mathcal{K})\} \\ N(\mathcal{K}) &:= \{Ye_0 : Y \in M(\mathcal{K})\} \\ N_0(\mathcal{K}) &:= \{Ye_0 : Y \in M_0(\mathcal{K})\} \qquad .\end{aligned}$$

We have the following lemma saying the relationships between $\mathcal{K}, \mathcal{K}_I, N_+(\mathcal{K}), N(\mathcal{K})$ and $N_0(\mathcal{K})$.

Lemma 2.1. *Let $\mathcal{K}, \mathcal{K}_I$ be as above, then $\mathcal{K}_I \subseteq N_+(\mathcal{K}) \subseteq N(\mathcal{K}) \subseteq N_0(\mathcal{K}) \subseteq \mathcal{K}$*

Proof. Firstly, let's prove $N_0(\mathcal{K}) \subseteq \mathcal{K}$. Let $\bar{y} \in N_0(\mathcal{K})$, by the definition of $N_0(\mathcal{K})$, there exists $\bar{Y} \in M_0(\mathcal{K})$ such that $\bar{Y}e_0 = \bar{y}$. Notice that $\bar{Y}e_0 = \underbrace{\bar{Y}(e_0 - e_i)}_{\in\mathcal{K}} + \underbrace{\bar{Y}e_i}_{\in\mathcal{K}} \in \mathcal{K}$. Secondly, let's prove $\mathcal{K}_I \subseteq N_+(\mathcal{K})$. Let $\bar{y} \in \mathcal{K}_I$ such that $\bar{y} \in \{0,1\}^{1+d}$. If $\bar{y}_0 = 0$, then $\bar{y} = 0$ and $\bar{Y} := 0$ shows $\bar{y} \in N_+(\mathcal{K})$. Otherwise we may assume $\bar{y}_0 = 1$ without loss of generality. Let $\bar{Y} := \bar{y}\bar{y}^T$ where $\bar{y} := \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_d \end{bmatrix}$, then $\bar{Y} = \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix}, x \in F$.

Whence $diag(\bar{Y}) = \begin{bmatrix} 1 \\ x_1^2 \\ x_2^2 \\ \vdots \\ x_d^2 \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = \bar{Y}e_0$ and $\bar{Y}e_i = \underbrace{x_i}_{\geq 0} \underbrace{\begin{bmatrix} 1 \\ x \end{bmatrix}}_{\in\mathcal{K}_I} \in \mathcal{K}$, $\bar{Y}(e_0 - e_i) = \underbrace{(1 - x_i)}_{\geq 0} \underbrace{\begin{bmatrix} 1 \\ x \end{bmatrix}}_{\in\mathcal{K}_I} \in \mathcal{K}$ for $\forall i \in \{1, \cdots d\}$. $\bar{Y} = \bar{y}\bar{y}^T \succeq 0$ thus $\bar{Y} \in M_+(\mathcal{K})$ and $\bar{Y}e_0 = \bar{y} \in N_+(\mathcal{K})$. □

Lemma 2.2. *Let $P, P_I, \mathcal{K}, \mathcal{K}_I$ be as above, then*

$$N_0(\mathcal{K}) \subseteq \left(\mathcal{K} \cap \left\{y \in \mathbb{R}^{d+1} : y_i = 0\right\}\right) + \left(\mathcal{K} \cap \left\{y \in \mathbb{R}^{d+1} : y_i = y_0\right\}\right), \forall i \in \{1, 2 \cdots d\}.$$

Also

$$N_0(P) \subseteq conv\left\{\left(P \cap \left\{x \in \mathbb{R}^d : x_i = 0\right\}\right) \cup \left(P \cap \left\{x \in \mathbb{R}^d : x_i = 1\right\}\right)\right\}, \forall i \in \{1, 2 \cdots d\}.$$

Proof. Let $\bar{y} \in N_0(\mathcal{K})$, then $\exists \bar{Y} \in M_0(\mathcal{K})$ such that $\bar{Y}e_0 = \bar{y}$ and

$$\bar{y} = \bar{Y}e_0 = \underbrace{\bar{Y}(e_0 - e_i)}_{\in\mathcal{K}\cap\{y\in\mathbb{R}^{d+1}:y_i=0\}} + \underbrace{\bar{Y}e_i}_{\in\mathcal{K}\cap\{y\in\mathbb{R}^{d+1}:y_i=y_0\}}$$

as required. With above lemma, we immediately get

$$N_+(P) \subseteq N(P) \subseteq N_0(P) \subseteq \bigcap_{i=1}^{d} conv\left[\left(P \cap \left\{x \in \mathbb{R}^d : x_i = 0\right\}\right) \cup \left(P \cap \left\{x \in \mathbb{R}^d : x_i = 1\right\}\right)\right]$$

. □

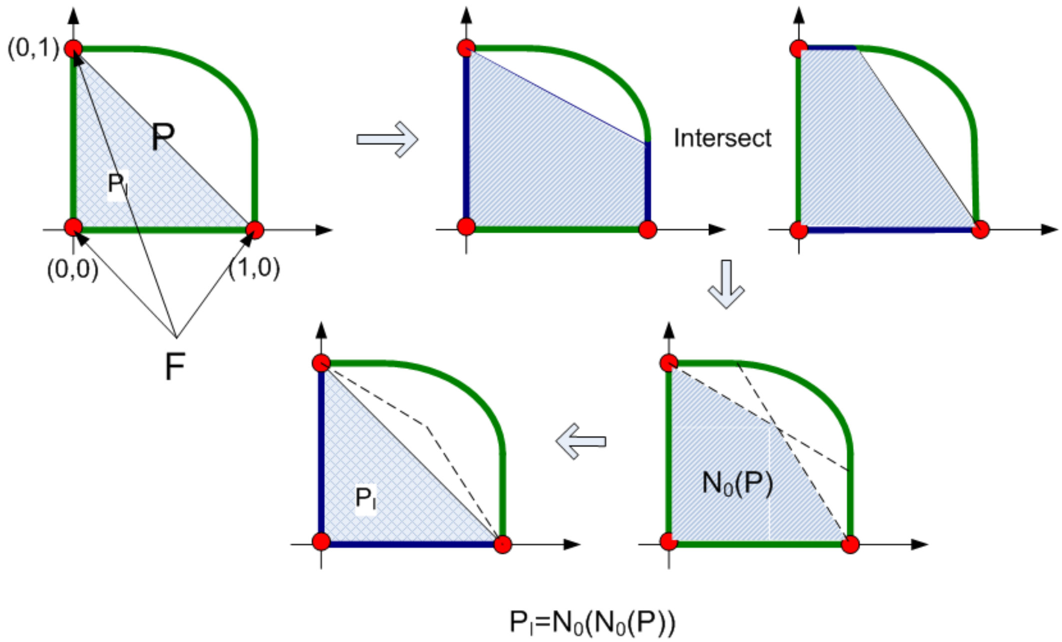

FIGURE 2.1. N_0-Operator to get P_I

Theorem 2.3. *Let P be as above. Then*

$$P \supseteq N_0(P) \supseteq N_0^2(P) \supseteq \cdots\cdots \supseteq N_0^d(P) = P_I.$$

Similarly for N as well as N_+.

We now do an example to see the lift-and-project procedures will always terminate with the convex hull P_I in at most d steps.

2.2. **Balas-Ceria-Cornuejols Procedure.** We have the main observation $N_0(P) \subseteq conv\left\{\left(P \cap \left\{x \in \mathbb{R}^d : x_i = 0\right\}\right) \cup \left(P \cap \left\{\right.\right.\right.$ $\{1, 2 \cdots d\}$ thus we can define a weaker operator

$$N_{(j)}(P) := conv\left\{\left(P \cap \left\{x \in \mathbb{R}^d : x_j = 0\right\}\right) \cup \left(P \cap \left\{x \in \mathbb{R}^d : x_j = 1\right\}\right)\right\}, \forall j \in \{1, 2 \cdots d\}.$$

We can also define $J := \{j_1, j_2, \cdots j_k\} \subseteq \{1, 2, \cdots d\}$and denote $N_{(J)} := N_{(j_k)}\left(N_{j_{k-1}}\left(\cdots\cdots N_{(j_1)}(P)\cdots\cdots\right)\right).$
It is easy to show

$$N_{(J)}(P) = conv\left(P \cap \{x_j \in \{0, 1\}, \forall j \in J\}\right).$$

Let P be as above then $N_{(\{1,2,\cdots\cdots d\})}(P) = P_I$. This procedure is called Balas-Ceria-Cornuejols Procedure.

2.3. **Sherali-Adams (Reformulation-Linearization) Procedure.**