Product rules in Semidefinite Programming

Rajat Mittal, Mario Szegedy and Troy Lee

Rutgers University

Abstract. In recent years we have witnessed the proliferation of semidefinite programming bounds in combinatorial optimization [1,5,8], quantum computing [9,2,3,6,4,16] and even in complexity theory [7,17,18]. Examples to such bounds include the semidefinite relaxation for the maximal cut problem [5], and the quantum value of multi-prover interactive games [3,4,16]. The first semidefinite programming bound, which gained fame, arose in the late seventies and was due to László Lovász [11], who used his theta number to compute the Shannon capacity of the five cycle graph. As in Lovász's upper bound proof for the Shannon capacity and in other situations the key observation is often the fact that the new parameter in question is multiplicative with respect to the product of the problem instances. In some recent results authors show the parallel repetition theorem for quantum interactive games using product properties of semidefinite relaxation [4,16]. Our goal is to classify those semidefinite programming instances for which the optimum is multiplicative under a naturally defined product operation. The product operation we define generalizes the ones used in [11] and [4,16]. We find conditions under which the product rule always holds and give examples for cases when the product rule does not hold. These conditions are lated extended to cover the remaining cases too [17,18].

1 Introduction

The Shannon capacity of a graph G is defined by $\lim_{n\to\infty} \mathrm{stbl}(G^n)^{1/n}$, where $\mathrm{stbl}(G)$ denotes the maximal independence set size of G. In his seminal paper of 1979, L. Lovász solved the open question that asked if the Shannon capacity of the five cycle, C_5 is $\sqrt{5}$ [11]. The proof was based on that $\mathrm{stbl}(C_5^2) = 5$ and that the independence number of any graph G is upper bounded by a certain semidefinite programming bound, that he called $\vartheta(G)$. Lovász showed that $\vartheta(C_5) = \sqrt{5}$, and that ϑ is multiplicative: $\vartheta(G \times G') = \vartheta(G) \times \vartheta(G')$, for any two graphs, G and G'. These facts together with the super-multiplicativity of $\mathrm{stbl}(G)$ are clearly sufficient to imply the conjecture.

Another prominent application of semidefinite programming is interactive games in classical and quantum complexity. Many articles have used semidefinite relaxations to bound the value of these games. Since it can be shown that these relaxations approximate the value of the game quite well, parallel repetition for these semidefinite programs can be converted into parallel repetition of interactive multi-prover games. Initially Feige and Lovasz used this method to bound classical value of these games[17]. Few years back R. Cleve, W. Slofstra,

F. Unger and S. Upadhyay showed that these programs can exactly model quantum XOR games and so came up with perfect parallel repetition theorem for these games [4]. Recently Kempe, Regev and Toner [16] gave good semidefinite relaxations for unique games and also used a similar approach to prove parallel repetition theorem for these games.

These successful applications of semidefinite programming bounds together with other ones, such as bounding acceptance probabilities achievable with various computational devices for independent copies of a given computational problem (generally known as "direct sum theorems"), point to the great use of product theorems for semidefinite programming.

In spite of these successes we do not know of any work which systematically investigates the conditions under which such product theorems hold. This is what we attempt to do in this article. While we do not manage to classify all cases, we hope that our study will serve as a starting point for such investigations. We define a brand of semidefinite programming instances with significantly large subclasses that obey the product rule. In Section 4 we describe some cases when product theorems hold, while in Proposition 2 we give an example when it does not. Later an attempt is made to give a necessary condition for these product theorems. We also raise several questions that intuit that product theorems always hold for "positive" instances, although that what should be the notion of positivity is not yet clear. Our goal is to provoke ideas, and set the scene for what one day might hopefully becomes a complete classification.

2 Affine semidefinite program instances

We will investigate a brand of semidefinite programming instances, which is described by a triplet $\pi = (J, A, b)$, where

- J is a matrix of dimension $n \times n$;
- $\mathbf{A} = (A^{(1)}, \ldots, A^{(m)})$ is a list of m matrices, each of dimension $n \times n$. We may view $\mathbf{A}$ as a three-dimensional matrix A_{kij} of dimensions $n \times n \times m$, where the last index corresponds to the upper index in the list;
- b is a vector of length m.

With π we associate a semidefinite programming instance with optimal value $\alpha(\pi)$:

$$\alpha(\pi) = \{\max J * X \mid \mathbf{A}X = b \quad \text{and} \quad X \succeq 0\} \tag{1}$$

We define dimension of the instance as the dimension of $\mathbf{A}$. Here variable matrix X has the same dimension ($n \times n$) as J and also the elements of the list $\mathbf{A}$. To avoid complications we assume that all matrices involved are symmetric. The operator that we denote by $*$ is the dot product ($tr(J^T X) = \sum_{ij} J_{ij} X_{ij}$) of matrices, so it results in a scalar. The set of m linear constraints are often of some simple form, e.g. in the case of Lovász's theta number all constraints are either of the form $X_{ij} = 0$ or $Tr(X) = 1$. In our framework the constraints can generally be of the form $\sum_{i,j} A_{kij} X_{ij} = b_k$, and the only restriction they have compared to the most general form of semidefinite programming instances is that all relations are strictly equations as opposed to inequalities *and* equations. These types of instances we call *affine*. In our notation the "scalar product" $\mathbf{A}X$ simply means the vector $(A^{(1)} * X, \ldots, A^{(m)} * X)$.

We will need the dual of π, which we denote by π^* (for the method to express the dual see for example [13]):

$$\{\min y.b \mid \; y\mathbf{A} - J \succeq 0\} \tag{2}$$

where y is a row vector of length m. Here $y\mathbf{A}$ is the matrix $\sum_{k=1}^{m} y_k A^{(k)}$.

3 Product instance

Definition 1. *Let $\pi_1 = (J_1, \mathbf{A}_1, b_1)$ and $\pi_2 = (J_2, \mathbf{A}_2, b_2)$ be two semidefinite instances with dimensions (n_1, n_1, m_1) and (n_2, n_2, m_2), respectively. We define the product instance as $\pi_1 \times \pi_2 = (J_1 \otimes J_2, \mathbf{A}_1 \otimes \mathbf{A}_2, b_1 \otimes b_2)$, where $\mathbf{A}_1 \otimes \mathbf{A}_2$ is by definition the list $(A_1^{(k)} \otimes A_2^{(l)})_{k,l}$ of length $m_1 m_2$ of $n_1 n_2 \times n_1 n_2$ matrices. The product instance has dimensions $(n_1 n_2, n_1 n_2, m_1 m_2)$.*

Although the above is a fairly natural definition, as it was pointed out in [10] in the special case of the Lovász's theta number, a slightly different definition gives the same optimal value, which is useful in some cases. The idea is that in lucky cases, when b_1 and/or b_2 have zeros, we may add new equations (extra to ones in Definition 1) to the primal system representing the product instance without changing its optimum value. The new instances that arise this way we call *weak product* and denote by "$\times_w$," even though there is a little ambiguity in the definition (it will only be clear from the context to an individual instance which equations we wish to add). Since if we add extra constraints to a maximization problem, the objective value does not increase, we have that

Proposition 1. $\alpha(\pi_1 \times_w \pi_2) \leq \alpha(\pi_1 \times \pi_2)$.

In Section 6 we give precise definitions for weak products and investigate their properties further. For the forthcoming sections we restrict ourselves to the product as defined in Definition 1. The natural question to ask here is whether two instances of semidefinite programs multiply well. i.e.

Definition 2. *Two semidefinite programming instances π_1 and π_2 are said to follow product theorem iff*
$\alpha(\pi_1 \times_w \pi_2) = \alpha(\pi_1 \times \pi_2)$

The rest of the article focuses on finding out conditions (sufficient, necessary) and examples for instances which follow product theorem.

3.1 Counterexample to product theorem

In this section we give an example when the product theorem does not hold. The example is the maximal eigenvalue function of a matrix, which, in contrast to the similar notion of spectral norm, is not multiplicative. Indeed, let M be a matrix with maximal eigenvalue 1 and minimal eigenvalue -2. Then, using the fact that under tensor product the spectra of matrices multiply, we get that $M \otimes M$ has maximal eigenvalue $4 \neq 1^2$ (the corresponding spectral norms would be 2 for M and 4 for $M \otimes M$).

Proposition 2. *The maximal eigenvalue of a matrix can be formulated as the optimal value of an affine semidefinite programming instance. This instance is not multiplicative.*

Proof. First notice that

$$\text{max eigenvalue}(M) = \{\min \lambda \mid \lambda I - M \succeq 0\}. \tag{3}$$

This is a dual (minimization) instance. Observe that $m = 1$, $n' = n$, $\mathbf{A} = (I)$, $J = M$ and $b = 1$. For the sake of completeness we describe the primal problem:

$$\text{max eigenvalue}(M) = \{\max \sum_{1 \leq i,j \leq n} M_{ij} X_{ij} \mid \text{Tr} X = 1; \;\; X \succeq 0\}. \tag{4}$$

The product instance associated with two matrices, M_1 and M_2, has parameters $I = I_1 \otimes I_2$, $M = M_1 \otimes M_2$ and $b = 1$. Since I is an identity matrix of appropriate dimensions, the optimum value of this instance is exactly the maximal eigenvalue of $M_1 \otimes M_2$. On the other hand, as was stated in the beginning of the section, the maximal eigenvalue problem is not multiplicative.

3.2 Product solution

Definition 3. *A subclass $\mathcal{C}$ of affine instances is said to obey the* product rule *if $\alpha(\pi_1 \times \pi_2) = \alpha(\pi_1)\alpha(\pi_2)$ for every $\pi_1, \pi_2 \in \mathcal{C}$.*

In section 3.1 we have given an example to an affine instance whose square does not obey the product rule. Therefore, for the product rule to hold we need to look for proper subclasses of all affine instances.

Let π_1 and π_2 be two affine instances with optimal solutions X_1 and X_2 for the primal and optimal solutions y_1 and y_2 for the dual. The well known duality theorem for semidefinite programming states that the value of the dual agrees with the value of the primal (given existence of slater points [15]). The first instinct for proving the product theorem would be to show that $X_1 \otimes X_2$ is a solution of the product instance with objective value $\alpha(\pi_1)\alpha(\pi_2)$, and $y_1 \otimes y_2$ is a solution of the dual of the product instance with the same value. The above two potential solutions for the product instance and its dual we call *product-solution* and *dual product-solution*. In other words, in order to show that the product rule holds for π_1 and π_2 it is sufficient to prove:

1. Feasibility of the product-solution: $(\mathbf{A}_1 \otimes \mathbf{A}_2)(X_1 \otimes X_2) = b_1 \otimes b_2$;
2. Feasibility of the dual product-solution: $y_1 \otimes y_2(\mathbf{A}_1 \otimes \mathbf{A}_2) - J_1 \otimes J_2 \succeq 0$;

3. Objective value of the primal product-solution: $(J_1 \otimes J_2) * (X_1 \otimes X_2) = (J_1 * X_1)(J_2 * X_2)$;
4. Objective value of the dual product-solution: $(y_1 \otimes y_2).(b_1 \otimes b_2) = (y_1.b_1)(y_2.b_2)$.

We also need the positivity of $X_1 \otimes X_2$, but this is automatic from the positivity of X_1 and X_2. Which of 1–4 fail to hold in general? Basic linear algebra gives that conditions 1, 3 and 4 hold without any further assumption. Note that we require the existence of slater points for duality theorem ($\alpha(\pi) = \alpha(\pi^*)$) in case of semidefinite programming. It says either there exist primal feasible positive definite X or dual feasible y s.t. $y\mathbf{A} - J \succ 0$ ([15]). But slater points exist in most of the cases and their existence can be shown depending upon specific instance. So we will assume that they exist, thus we already have that:

Proposition 3. *Let π_1 and π_2 be two affine instances. Then $\alpha(\pi_1 \times \pi_2) \geq \alpha(\pi_1)\alpha(\pi_2)$.*

In what follows, we will examine cases when Condition 2 also holds.

4 Some Sufficient Conditions

From previous section, any property (of $\mathbf{A}, J, b$) which imply Condition 2 (dual feasibility) will be a sufficient condition for product theorem to hold. So in this section we present some sufficient conditions for dual feasibility. Later we also derive a necessary condition for dual feasibility (which is also sufficient if we restrict our attention to an instance and its square) in the next section, but the latter expression uses y_1 and y_2, like dual feasibility itself. It remains a task for the future to develop a necessary and sufficient condition whose criterion is formulated solely in terms of the problem instances π_1 and π_2.

These sufficient conditions are connected to each other. The proofs are presented independently for better intuition. At this section's end, there will be a table explicitly stating all the sufficient conditions in this section. The connection between them will also be explained. The proof of these connections will be clear from the proofs given below.

4.1 Positivity of matrix J

Our first simple condition is the positivity of J.

Theorem 1. *Assume that both J_1 and J_2 are positive semidefinite. Then $\alpha(\pi_1 \times \pi_2) = \alpha(\pi_1)\alpha(\pi_2)$.*

Proof. As we noted in Section 3.2 it is sufficient to show that Condition 2 of that section holds. By our assumptions on y_1 and y_2 we have that $y_1\mathbf{A}_1 - J_1$ and $y_2\mathbf{A}_2 - J_2$ are positive semi-definite. So $y_1\mathbf{A}_1 + J_1$ and $y_2\mathbf{A}_2 + J_2$ are also positive semi-definite, since they arise as sums of two positive matrices. For instance, $y_1\mathbf{A}_1 + J_1 = (y_1\mathbf{A}_1 - J_1) + 2J_1$. The above implies that

$$(y_1\mathbf{A}_1 - J_1) \otimes (y_2\mathbf{A}_2 + J_2) = y_1\mathbf{A}_1 \otimes y_2\mathbf{A}_2 - J_1 \otimes y_2\mathbf{A}_2 + y_1\mathbf{A}_1 \otimes J_2 - J_1 \otimes J_2 \succeq 0. \quad (5)$$

Also

$$(y_1\mathbf{A}_1+J_1)\otimes(y_2\mathbf{A}_2-J_2) = y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-y_1\mathbf{A}_1\otimes J_2+J_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2 \succeq 0 \quad (6)$$

Taking the average of the right hand sides of Equations (5) and (6) we obtain that

$$y_1\mathbf{A}_1 \otimes y_2\mathbf{A}_2 - J_1 \otimes J_2 \succeq 0, \quad (7)$$

which is the desired Condition 2. (Note: It is easy to see that $y_1\mathbf{A}_1 \otimes y_2\mathbf{A}_2 = y_1 \otimes y_2(\mathbf{A}_1 \otimes \mathbf{A}_2)$.)

Lovász theta number ([11]) is an example that falls into this category. Consider the definition of Lovász theta number in [13]. Then J is the all $1's$ matrix, which is positive semidefinite. The matrix remains positive definite even if we consider the weighted version of the theta number [10], in which case J is of the form ww^T for some column vector w.

We remark that slater points definitely exist for the case $J \succeq 0$. If y is feasible then cy where $c >> 0$ is a slater point.

4.2 All $A^{(k)}$ are block diagonal, and J is block anti-diagonal

The argument in the previous section is applicable whenever $y_c\mathbf{A}_c + J_c$ $(c \in \{1,2\})$ are known to be positive semidefinite matrices. Let us state this explicitly:

Lemma 1. *Whenever $y_c\mathbf{A}_c + J_c$ $(c \in \{1,2\})$ are positive definite, where y_1 and y_2 are the optimal solutions of π_1^* and π_2^*, respectively, then the product theorem holds for π_1 and π_2.*

This is the avenue Cleve et. al. take in [4]. Following their lead, but slightly generalizing their argument we show:

Lemma 2. *For a semidefinite programming instance $\pi = (\mathbf{A}, J, b)$ if the matrix J is block anti-diagonal and if y is a feasible solution of the dual such that $y\mathbf{A}$ is block diagonal then $y\mathbf{A} + J \succeq 0$.*

Block diagonal and anti-diagonal matrices have the following structure:

Block anti-diagonality	Block diagonality
$\begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix}$	$\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$

In our definition block diagonal and anti-diagonal matrices have two by two blocks. We require that if J is block anti-diagonal and $y\mathbf{A}$ is block-diagonal, then their rows and columns be divided to blocks in exactly the same way.

We will prove our claim by contradiction. Suppose $y\mathbf{A}$ and J are of the required form but $y\mathbf{A} + J$ is not positive semidefinite. Then there exists a vector

w, in block form $w = (w', w'')$ for which $w^T(y\mathbf{A} + J)w$ is negative (we treat all vectors as column vectors). Define $v = (w', -w'')$. Now

$$\begin{aligned} v^T(y\mathbf{A} - J)v = \\ (w', -w'')^T y\mathbf{A}(w', -w'') - (w', -w'')^T J(w', -w'') = \\ (w', w'')^T y\mathbf{A}(w', w'') + (w', w'')^T J(w', w'') = \\ w^T(y\mathbf{A} + J)w < 0. \end{aligned}$$

This implies that $y\mathbf{A} - J$ is not positive semidefinite, which is a contradiction since by our assumption y is a solution of π^*. We can generalize the proof for case when J is of the form $J_1 + J_2$, where J_1 is of the form as before and J_2 is positive semidefinite ($y\mathbf{A}$ should still be block diagonal). Notice that the block diagonality of $y\mathbf{A}$ automatically holds if $\mathbf{A} = (A^{(1)}, \ldots, A^{(m)})$, where each $A^{(k)}$ is block diagonal. We summarize the findings of this section in the following theorem:

Theorem 2. *Let $\pi_1 = (\mathbf{A}_1, J_1, b_1)$ and $\pi_2 = (\mathbf{A}_2, J_2, b_2)$ be affine instances such that for $c \in \{1, 2\}$:*

1. *$\mathbf{A}_c = (A_c^{(1)}, \ldots, A_c^{(m)})$, where each $A_c^{(k)}$ is block diagonal;*
2. *$J_c = J'_c + J''_c$ $(c \in \{1, 2\})$, where J'_c is block anti-diagonal and J''_c is positive.*

(All blocked matrices have the same block divisions.) Then for π_1 and π_2 the product theorem holds.

Again we note that existence of slater points is required for this theorem to hold. Matrix I serves as a Slater point for Cleve et. al. [4]. Also observe that the product defined here ([4]) is slightly different from our case. This format difference can be handled using the approach from [16].

4.3 Unified Condition

We notice that if $y\mathbf{A} + J \succeq 0$ then product theorem holds. Lets try to give sufficient conditions for this to hold. If $y\mathbf{A} - J \succeq 0$, then for any orthogonal matrix S, we can say

$$\begin{aligned} S(y\mathbf{A} - J)S^{-1} \succeq 0 \\ \Rightarrow yS\mathbf{A}S^{-1} - SJS^{-1} \succeq 0 \\ \Rightarrow yS\mathbf{A}S^{-1} + J - SJS^{-1} - J \succeq 0 \\ \Rightarrow yS\mathbf{A}S^{-1} + J \succeq SJS^{-1} + J \end{aligned}$$

Clearly if $\exists S$ s.t. $S\mathbf{A} = \mathbf{A}S, SJS^{-1} + J \succeq 0$, then $y\mathbf{A} + J \succeq 0$. Both the cases mentioned above follow from this single condition. If we take $S = I$ then we have the condition that $J \succeq 0$. Also if

$$S = \begin{matrix} I & 0 \\ 0 & -I \end{matrix}$$

We get the condition of block-diagonality and block anti-diagonality.

Now we summarize the contents of this section through a table. The first row in table (Lemma 1) is the most basic and strongest among the four. Since all the proofs go through proving this condition, all other conditions follow from condition 1. From section 4.3 we know condition 2 ($\exists S$ s.t. $S\mathbf{A}S^{-1} = \mathbf{A} \wedge SJS^{-1} + J \succeq 0$) implies condition 3 and condition 4. It is interesting to find out what other criteria can emerge from condition 2.

Table 1. Sufficient conditions for product theorem

<table>
<tr><td colspan="2">1. $y\mathbf{A} + J \succeq 0$</td></tr>
<tr><td colspan="2">2. $\exists S$ s.t. $S\mathbf{A}S^{-1} = \mathbf{A} \wedge SJS^{-1} + J \succeq 0$</td></tr>
<tr><td>3. $J \succeq 0$</td><td>4. J block-antidiagonal, $\mathbf{A}$ block diagonal.</td></tr>
</table>

5 A necessary condition for the feasibility of $y_1 \otimes y_2$

In this section we show that the condition in Lemma 1 is not only sufficient, but also necessary (or at least "half of it"), if we insist on the "first instinct" proof method. The only exception is the trivial case when $y_c\mathbf{A}_c - J_c$ is identically zero. But in that case product theorem will hold and so can be dealt with separately.

Lemma 3. *Let us assume $y_c\mathbf{A}_c - J_c$ is not identically zero for $c \in \{1, 2\}$. For two instances π_1 and π_2, let y_1 and y_2 be optimal solutions of π_1^* and π_2^*, respectively. Then $y_1 \otimes y_2$ is a feasible solution of the dual of the product instance (i.e. Condition 2 of section 3.2 holds) only if at least one of $y_c\mathbf{A}_c + J_c$ ($c \in \{1, 2\}$) are positive semidefinite.*

Proof. Let us assume the contrary. Then we have vectors v_c ($c \in \{1, 2\}$) such that $v_c^T(y_c\mathbf{A}_c + J_c)v_c < 0$ ($c \in \{1, 2\}$). Our assumptions imply that $v_c^T(y_c\mathbf{A}_c - J_c)v_c \geq 0$ ($c \in \{1, 2\}$). Let us call this a weaker condition.

Now we will show that, we have vectors w_c ($c \in \{1, 2\}$) such that $w_c^T(y_c\mathbf{A}_c + J_c)w_c < 0$ and $w_c^T(y_c\mathbf{A}_c - J_c)w_c > 0$ ($c \in \{1, 2\}$). Notice that now in the second inequality we have strictly greater than zero condition. Let us call this a strong condition.

Say v satisfies our weak condition and not strong condition (if it does then we are done). So pick any direction u and define $w = v + \delta u$ where δ is really small. So

$$
\begin{aligned}
w^T(y\mathbf{A}+J)w &< 0 \quad \text{(Since } \delta \text{ can be made arbitrarily small)}\\
w^T(y\mathbf{A}-J)w &\geq 0 \text{ (because } yA-J \text{ is positive semidefinite)}\\
\Rightarrow w^T(y\mathbf{A}-J)w &= 0 \quad \text{(else } w \text{ satisfies the strong condition)}\\
\Rightarrow (v+\delta u)^T(y\mathbf{A}-J)(v+\delta u) &= 0\\
\Rightarrow (\delta u)^T(y\mathbf{A}-J)(\delta u) &= 0 \qquad\qquad (v^T(y\mathbf{A}-J)v=0)
\end{aligned}
$$

This follows because $y\mathbf{A}-J$ is positive semidefinite and $v^T(y\mathbf{A}-J)v=0$, then $(y\mathbf{A}-J)v=0$ (any positive semidefinite matrix can be written in the form M^TM). But now $u^T(y\mathbf{A}-J)u$ for any direction u so we conclude that $y\mathbf{A}-J=0$. This proves that we have w which satisfy the strong condition.

From strong condition it follows that

$$
\begin{aligned}
&(w_1\otimes w_2)^T((y_1\mathbf{A}_1-J_1)\otimes(y_2\mathbf{A}_2+J_2))(w_1\otimes w_2)<0\\
\Rightarrow\ &(w_1\otimes w_2)^T(y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2-J_1\otimes y_2\mathbf{A}_2+y_1\mathbf{A}_1\otimes J_2)(w_1\otimes w_2)<0\\
\Rightarrow\ &(w_1\otimes w_2)^T(y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2)(w_1\otimes w_2)+\\
&(w_1\otimes w_2)^T(y_1\mathbf{A}_1\otimes J_2-J_1\otimes y_2\mathbf{A}_2)(w_1\otimes w_2)<0
\end{aligned}
$$

By similar argument, considering now the inequality

$$(w_1\otimes w_2)^T((y_1\mathbf{A}_1+J_1)\otimes(y_2\mathbf{A}_2-J_2))(w_1\otimes w_2)<0,$$

we can show that

$$
\begin{aligned}
&(w_1\otimes w_2)^T(y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2)(w_1\otimes w_2)+\\
&(w_1\otimes w_2)^T(-y_1\mathbf{A}_1\otimes J_2+J_1\otimes y_2\mathbf{A}_2)(w_1\otimes w_2)<0
\end{aligned}
$$

By averaging the two inequalities we get that

$$(w_1\otimes w_2)^T(y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2)(w_1\otimes w_2)<0$$

This contradicts to the assumption of the lemma that $y_1\otimes y_2$ is a feasible solution of $\pi_1\times\pi_2$ (which in turn implies that $y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2$ is positive definite).

Now consider the case when $y_c\mathbf{A}_c-J_c=0$. If exactly one of $y_c\mathbf{A}_c-J_c$ is zero (say $y_1\mathbf{A}_1-J_1$), then $(y_1\mathbf{A}_1-J_1)\otimes(y_2\mathbf{A}_2+J_2)$ is zero. Hence

$$y_1\mathbf{A}_1\otimes y_2\mathbf{A}_2-J_1\otimes J_2=(y_1\mathbf{A}_1+J_1)\otimes(y_2\mathbf{A}_2-J_2) \tag{8}$$

Then dual is feasible if and only if $y_1\mathbf{A}_1+J_1$ is positive semidefinite. In case both $y_c\mathbf{A}_c-J_c$ are zero. Then J can be written as linear combination of A^i matrices. Since every A^i*X is fixed, the objective value of the primal is same for every feasible X, i.e. y^Tb (here y is the vector such that $y\mathbf{A}-J=0$). Same holds true for the product problem and can be easily shown that these constant objective values multiply.

One might suspect that the full converse of Lemma 1 holds, i.e. in the case of the feasibility of $y_1\otimes y_2$ both $y_1\mathbf{A}_1+J_1$ and $y_2\mathbf{A}_2+J_2$ should be positive semi-definite, but in the next section we give a counter-example to this.

5.1 Counterexample and necessary condition

It is educational to see where the condition of Proposition 2 fails with respect to the example of section 3.1. Recall that $J = M$, $\mathbf{A} = (I)$ and $y = \lambda$ (the maximal eigenvalue of M). The point is that even when $\lambda I - M$ is positive, $\lambda I + M$ is not necessarily. We can extend the above example to show that in Lemma 3 we cannot exchange the "one of" to "both." Let M_1 be the matrix with eigenvalues -2 and 1 and let M_2 be the matrix with eigenvalues 0 and 1. Then $y_1 = 1$ and $y_2 = 1$, so $y_1 \otimes y_2 = 1$, which is a solution of

$$\{\min \lambda \mid \lambda I - M_1 \otimes M_2 \succeq 0\}, \tag{9}$$

even though $I + M_1$ is not positive semidefinite.

Notice on the other hand, if M is positive then $\lambda I - M \succeq 0 \Rightarrow \lambda I + M \succeq 0$, and indeed the maximum eigenvalue of positive matrices multiply under tensor product. As a perhaps far-fetched conjecture we ask:

Conjecture 1. For an affine instance $\pi = (\mathbf{A}, J, b)$ define

$$\alpha^+(\pi) = \{\max |J * X| \mid \mathbf{A}X = b \quad \text{and} \quad X \succeq 0\}.$$

Is it true that α^+ is always multiplicative? Here α^+ represents a generalized "spectral norm."

But a counterexample can be given for this conjecture also. For that counterexample define matrix norm as the sum of two largest eigenvalues of that matrix. Then it can be expressed as a semidefinite program ([15]). It can be seen that Conjecture 1 does not hold in this case.

6 The weak product

A surprising observation about the theta number of Lovász, well described in [10], is that it is multiplicative with two different notions of products:

Definition 4 (Strong product "$\times$" of graphs). *(u', u'') −− (v', v'') or (u', u'') = (v', v'') in $G' \times G''$ if and only if (u' −− v' or $u' = v'$ in G') and (u'' −− v'' or $u'' = v''$ in G'').*

and

Definition 5 (Weak product "$\times_w$" of graphs). *$G' \times_w G'' = \overline{\overline{G'} \times \overline{G''}}$.*

Recall that $\vartheta(G)$ is defined by [13] (by J we denote the matrix with all 1 elements):

$$\vartheta(G) = \{\max J * X \mid I * X = 1;\ \forall (i,j) \in E(G):\ X_{i,j} = 0;\ X \succeq 0\}. \tag{10}$$

That is, every edge gives a new linear constraint, increasing m by one. In general, $E(G' \times_w G'') \supseteq E(G' \times G'')$, because (u', u'') −− (v', v'') is an edge of $G' \times G''$ if and only if both of its projections are edges or identical coordinates, but

$(u', u'') \neq (v', v'')$. On the other hand, $(u', u'') -- (v', v'')$ is an edge of $G' \times_w G''$ if and only if there exists at least one projection which is an edge.

It is easy to see that the constraint in Expression (10) for $\vartheta(G' \times G'')$ has a constraint for every constraint pair in the corresponding expression for G' and G', so the strong product is the one that corresponds to our usual product notion that appears in previous sections. In contrast, when we write down Expression (10) for $\vartheta(G' \times_w G'')$, we see a lot of extra constrains.

How do they arise? In general, assume that we know that the product solution $X_1 \otimes X_2$ is the optimal solution for $\pi_1 \times \pi_2$ (which is indeed the case under the conditions we considered in earlier sections). Assume furthermore that some coordinate i of b_1 is zero. Then $A_1^{(i)} * X_1 = 0$. Now we may take any $n_2 \times n_2$ matrix B, and it will hold that

$$(A_1^{(i)} \otimes B) * (X_1 \otimes X_2) = (A_1^{(i)} * X_1)(B * X_2) = 0.$$

Therefore adding matrices of the form $A_1^{(i)} \otimes B$ to $\mathbf{A}_1 \otimes \mathbf{A}_2$ and setting the the corresponding entry of the longer b vector of the product instance to zero will not influence the objective value. The same can be said about about exchanging the roles of π_1 and π_2.

We can easily see that the weak product in the case of the theta number arises this way. That what equations to the product system we wish to add this way is a matter of taste, and we believe it depends on the specific class of semidefinite programming instances under study. We summarize the finding of this section in the following proposition

Proposition 4. *Assume that for affine instances π_1 and π_2 the multiplicative rule holds. Then if define a system $\pi_1 \times_w \pi_2$ that we call "weak product" by conveniently adding arbitrary number of new constrains to the system that follow the construction rules described above (in particular, every added constraint should be associated with a zero entry of b_1 or b_2), the multiplicative rule will also hold for the weak product.*

The above lemma explains why the theta number of Lovász is multiplicative with respect to the weak product of graphs.

7 Nonnegativity constraint

There are still some cases [17,18], which are not explained by the sufficient conditions given above. The main difference is the format of semidefinite programming they consider and the product defined. So now we consider programs of the following form:

$$\begin{aligned} \alpha(\pi) = \max\nolimits_X J \bullet X & \text{ such that} \\ \mathbf{A} \bullet X &= b \\ \mathbf{B} \bullet X &\geq \mathbf{0} \\ X &\succeq 0 \end{aligned}$$

Here both $\mathbf{A}$ and $\mathbf{B}$ are vectors of matrices, and $\mathbf{0}$ denotes the all 0 vector.

We should point out a subtlety here. A program of this form can be equivalently written as an affine program by suitably extending X and modifying $\mathbf{A}$ accordingly to enforce the $\mathbf{B} \bullet X \geq \mathbf{0}$ constraints through the $X \succeq 0$ condition. The catch is that two equivalent programs do not necessarily lead to equivalent product instances. We explicitly separate out the non-negativity constraints here so that we can define the product as follows: for two programs, $\pi_1 = (J_1, \mathbf{A}_1, b_1, \mathbf{B}_1)$ and $\pi_2 = (J_2, \mathbf{A}_2, b_2, \mathbf{B}_2)$ we say

$$\pi_1 \times \pi_2 = (J_1 \otimes J_2, \mathbf{A}_1 \otimes \mathbf{A}_2, b_1 \otimes b_2, \mathbf{B}_1 \otimes \mathbf{B}_2).$$

Notice that the equality constraints and non-negativity constraints do not interact in the product, which is usually the intended meaning of the product of instances.

It is again straightforward to see that $\alpha(\pi_1 \times \pi_2) \geq \alpha(\pi_1)\alpha(\pi_2)$, thus we focus on the reverse inequality. The following theorem captures the product theorems of Feige-Lovász [17] and discrepancy [18].

Theorem 3. *Let $\pi_1 = (J_1, \mathbf{A}_1, b_1, \mathbf{B}_1)$ and $\pi_2 = (J_2, \mathbf{A}_2, b_2, \mathbf{B}_2)$ be two semidefinite programs for which strong duality holds. Suppose the following two conditions hold:*

1. *(Bipartiteness) There is a partition of rows and columns into two sets such that, with respect to this partition, J_i and all the matrices of $\mathbf{B}_i$ are block anti-diagonal, and all the matrices of $\mathbf{A}_i$ are block diagonal, for $i \in \{1, 2\}$.*
2. *There are non-negative vectors u_1, u_2 such that $J_1 = u_1^T \mathbf{B}_1$ and $J_2 = u_2^T \mathbf{B}_2$.*

Then $\alpha(\pi_1 \times \pi_2) \leq \alpha(\pi_1)\alpha(\pi_2)$.

Proof. To prove the theorem it will be useful to consider the dual formulations of π_1 and π_2. Dualizing in the standard fashion, we find

$$\begin{aligned} \alpha(\pi_1) = \min_{y_1} \; & y_1^T b_1 \text{ such that} \\ & y_1^T \mathbf{A}_1 - (z_1^T \mathbf{B}_1 + J_1) \succeq 0 \\ & z_1 \geq 0 \end{aligned}$$

and similarly for π_2. Fix y_1, z_1 to be vectors which realizes this optimum for π_1 and similarly y_2, z_2 for π_2. The key observation of the proof is that if we can also show that

$$y_1^T \mathbf{A}_1 + (z_1^T \mathbf{B}_1 + J_1) \succeq 0 \text{ and } y_2^T \mathbf{A}_2 + (z_2^T \mathbf{B}_2 + J_2) \succeq 0 \tag{11}$$

then we will be done. Let us for the moment assume 11 and see why this is the case.

If equation 11 holds, then we also have

$$\begin{aligned} \left(y_1^T \mathbf{A}_1 - (z_1^T \mathbf{B}_1 + J_1)\right) \otimes \left(y_2^T \mathbf{A}_2 + (z_2^T \mathbf{B}_2 + J_2)\right) \succeq 0 \\ \left(y_1^T \mathbf{A}_1 + (z_1^T \mathbf{B}_1 + J_1)\right) \otimes \left(y_2^T \mathbf{A}_2 - (z_2^T \mathbf{B}_2 + J_2)\right) \succeq 0 \end{aligned}$$

Averaging these equations, we find

$$(y_1 \otimes y_2)^T (\mathbf{A}_1 \otimes \mathbf{A}_2) - \big((z_1^T \mathbf{B}_1 + J_1) \otimes (z_2^T \mathbf{B}_2 + J_2)\big) \succeq 0.$$

Let us work on the second term. We have

$$\begin{aligned}(z_1^T \mathbf{B}_1 + J_1) \otimes (z_2^T \mathbf{B}_2 + J_2) &= (z_1 \otimes z_2)^T (\mathbf{B}_1 \otimes \mathbf{B}_2) + z_1^T \mathbf{B}_1 \otimes J_2 + J_1 \otimes z_2^T \mathbf{B}_2 \\ &\quad + J_1 \otimes J_2 \\ &= (z_1 \otimes z_2)^T (\mathbf{B}_1 \otimes \mathbf{B}_2) + (z_1 \otimes u_2)^T \mathbf{B}_1 \otimes \mathbf{B}_2 \\ &\quad + (u_1 \otimes z_2)^T \mathbf{B}_1 \otimes \mathbf{B}_2 + J_1 \otimes J_2.\end{aligned}$$

Thus if we let $v = z_1 \otimes z_2 + z_1 \otimes u_2 + u_1 \otimes z_2$ we see that $v \geq 0$ as all of z_1, z_2, u_1, u_2 are, and also

$$(y_1 \otimes y_2)^T \otimes (\mathbf{A}_1 \otimes \mathbf{A}_2) - (v^T (\mathbf{B}_1 \otimes \mathbf{B}_2) + J_1 \otimes J_2) \succeq 0.$$

Hence $(y_1 \otimes y_2, v)$ form a feasible solution to the dual formulation of $\pi_1 \times \pi_2$ with value $(y_1 \otimes y_2)(b_1 \otimes b_2) = \alpha(\pi_1)\alpha(\pi_2)$.

It now remains to show that 11 follows from the condition of the theorem. Given $y\mathbf{A} - (z^T \mathbf{B} + J) \succeq 0$ and the bipartiteness condition of the theorem, we will show that $y\mathbf{A} + (z^T \mathbf{B} + J) \succeq 0$. This argument can then be applied to both π_1 and π_2.

We have that $y^T \mathbf{A}$ is block diagonal and $z^T \mathbf{B} + J$ is block anti-diagonal with respect to the same partition. Hence for any vector $x^T = \begin{bmatrix} x_1 \; x_2 \end{bmatrix}$, we have

$$\begin{bmatrix} x_1 \; x_2 \end{bmatrix} \left(y^T \mathbf{A} - (z^T \mathbf{B} + J)\right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \; -x_2 \end{bmatrix} \left(y^T \mathbf{A} + (z^T \mathbf{B} + J)\right) \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$$

Thus the positive semidefiniteness of $y\mathbf{A} + (z^T \mathbf{B} + J)$ follows from that of $y\mathbf{A} - (z^T \mathbf{B} + J)$.

One may find the condition that J lies in the positive span of $\mathbf{B}$ in the statement of Theorem 3 somewhat unnatural. If we remove this condition, however, a simple counterexample shows that the theorem no longer holds. Consider the program

$$\begin{aligned}\alpha(\pi) = &\max_X \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \bullet X \\ &\text{such that } I \bullet X = 1, \begin{bmatrix} 0 \; 1 \\ 0 \; 0 \end{bmatrix} \bullet X \geq 0, \begin{bmatrix} 0 \; 0 \\ 1 \; 0 \end{bmatrix} \bullet X \geq 0, X \succeq 0.\end{aligned}$$

Here I stands for the 2-by-2 identity matrix. This program satisfies the bipartiteness condition of Theorem 3, but J does not lie in the positive span of the matrices of $\mathbf{B}$. It is easy to see that the value of this program is zero. The program $\pi \times \pi$, however, has positive value as $J \otimes J$ does not have any negative entries but is the matrix with ones on the main anti-diagonal. For details about how this theorem applies to cases of [17,18], please refer to [19].

8 Some open problems

We formulate some open problems all coming from the intuition that there must be a notion of "positive" affine instances for which the product theorem always holds. The next question relates to monotonicity:

Conjecture 2. Let $\pi_1 = (\mathbf{A}_1, J_1, b_1)$ and $\pi_2 = (\mathbf{A}_2, J_2, b_2)$ be the affine instances for which the product theorem holds. Then it also holds for the instance pair $\pi_1' = (\mathbf{A}_1, J_1 + J, b_1)$ and $\pi_2' = (\mathbf{A}_2, J_2 + J', b_2)$, where J and J' are positive matrices.

The following question suggests that the more negative J is, the more special $\mathbf{A}$ has to be. In particular, if J is not positive then at least some $\mathbf{A}$ is excluded.

Conjecture 3. For every strictly non-positive J (i.e. J has a negative eigenvalue) there are $\mathbf{A}$ and b such that for the instance $\pi = (\mathbf{A}, J, b)$ it holds that $\alpha(\pi^2) \neq \alpha(\pi)^2$.

On the other hand, we may conjecture that whether the product theorem holds or not is entirely independent of b:

Conjecture 4. Let $\pi_1 = (\mathbf{A}_1, J_1, b_1)$ and $\pi_2 = (\mathbf{A}_2, J_2, b_2)$ be the affine instances for which the product theorem holds. Then it also holds for the instance pair $\pi_1' = (\mathbf{A}_1, J_1, b_1 + b)$ and $\pi_2' = (\mathbf{A}_2, J_2, b_2 + b')$ for any b and b'.

Other area of research is to try for more general *composition* theorems: in this setting, if one has a lower bound on the complexity of $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^k \to \{0,1\}$, one would like to obtain a lower bound on $(f \circ g)(\boldsymbol{x}) = f(g(x_1), \ldots, g(x_n))$. What we have studied so far in looking at tensor products corresponds to the special cases where f is the PARITY or AND function, depending on if the objective matrix is a sign matrix or a 0/1 valued matrix. One example of such a general composition theorem is known for the adversary method, a semidefinite programming quantity which lower bounds quantum query complexity. There it holds that $\mathrm{ADV}(f \circ g) \geq \mathrm{ADV}(f)\mathrm{ADV}(g)$ [**?**,**?**]. It would be interesting to develop a theory to capture these cases as well.

9 Conclusions

We have started to systematically investigate product theorems for affine instances of semidefinite programming. Our theorems imply the important results of Cleve. et al. [4], theta number of Lovász [11], Feige et.al. [17] and others. Although their proof came both logically and chronologically first, the mere fact that the proposed theory has such immediate consequences, in our opinion serves as a worthwhile motivation for its development. Added to this that various direct sum results for different computational models would also be among the immediate consequences of the theory, we conclude that we have hit upon a basic research topic with immediate and multiple applications in computer science.

The issue, therefore, at this point is not the number of potential applications, which seems abundant, but rather the relative scarcity of positive results. In the paper we have formulated conjectures that we hope will raise interest in researchers who intend to study this topic further.

We would like to thank anonymous referees for their invaluable suggestions.

References

1. S. Arora, S. Rao, U. Vazirani: Expander Flows, Geometric Embeddings and Graph Partitioning. Proceedings of Symposium on the Theory of Computing, 2004.
2. Howard Barnum, Michael E. Saks, Mario Szegedy: Quantum query complexity and semi-definite programming. IEEE Conference on Computational Complexity 2003: 179-193
3. Richard Cleve, Peter Hoyer, Benjamin Toner, John Watrous: Consequences and Limits of Nonlocal Strategies. IEEE Conference on Computational Complexity 2004: 236-249
4. R. Cleve and W. Slofstra and F. Unger and S. Upadhyay *Strong parallel repetition Theorem for Quantum XOR Proof Systems* Aug 2006: quant-ph.
5. M.X. Goemans and D.P. Williamson, Approximation Algorithms for MAX-3-CUT and Other Problems Via Complex Semidefinite Programming, Journal of Computer and System Sciences (Special Issue for STOC 2001), 68, 442–470, 2004. Preliminary version in Proceedings of 33rd STOC, Crete, 443–452, 2001.
6. Peter Hoyer, Troy Lee and Robert Spalek: Negative weights makes adversaries stronger, quant-ph/0611054
7. Sophie Laplante, Troy Lee, Mario Szegedy: The Quantum Adversary Method and Classical Formula Size Lower Bounds. Computational Complexity 15(2): 163-196 (2006)
8. David R. Karger, Rajeev Motwani, Madhu Sudan: Approximate Graph Coloring by Semidefinite Programming. J. ACM 45(2): 246-265 (1998)
9. Kitaev(unpublished proof). Quantum Coin Tossing. Slides at the archive of MSRI Berkeley.
10. Donald E. Knuth *The Sandwich Theorem* 1994 : Electron. J. Combin..
11. László Lovász: *On the Shannon capacity of a graph* Jan 1979 : IEEE Transactions on Information Theory.
12. Ran Raz: *A Parallel Repetition Theorem*, SIAM Journal of Computing 27(3) (1998) pp. 763-803
13. Mario Szegedy: *A note on the theta number of Lovász and the generalized Delsarte bound* 1994: FOCS.
14. B. S. Tsirelson: *Quantum analogues of the Bell inequalities: The case of two spatially separated domains* Journal of Soviet Mathematics, 36:557-570, 1987.
15. L. Vandenberghe and S. Boyd: *Semidefinite Programming* SIAM review, 38(1):49-95, March 1996.
16. J. Kempe, O. Regev, B. Toner: The Unique Game Conjecture with Entangled Provers is False. quant-ph/arXiv:0710.0655.
17. U. Feige and L. Lovász: Two-prover one-round proof systems: their power and their problems. In *Proceedings of the 24th ACM Symposium on the Theory of Computing*, pages 733–744. ACM, 1992.
18. T. Lee, A. Shraibman, and R. Špalek: A direct product theorem for discrepancy. In *Proceedings of the 23rd IEEE Conference on Computational Complexity*. IEEE, 2008.

19. T. Lee and R. Mittal: *Product theorems via semidefinite programming* Icalp 2008.