Non-commutative polynomial optimization

S. Pironio[1], M. Navascuès[2], A. Acín[3]

[1] Laboratoire d'Information Quantique (Brussels)

[2] Department of Mathematics (Bristol)

[3] ICFO-The Institute of Photonic Sciences (Barcelona)

MEGA 2011, Stockholm University, June 2, 2011

Our motivation: solve optimization problems arising in quantum physics.
They can be viewed as polynomial optimization problems in non-commuting variables (i.e. operators or matrices).

↓

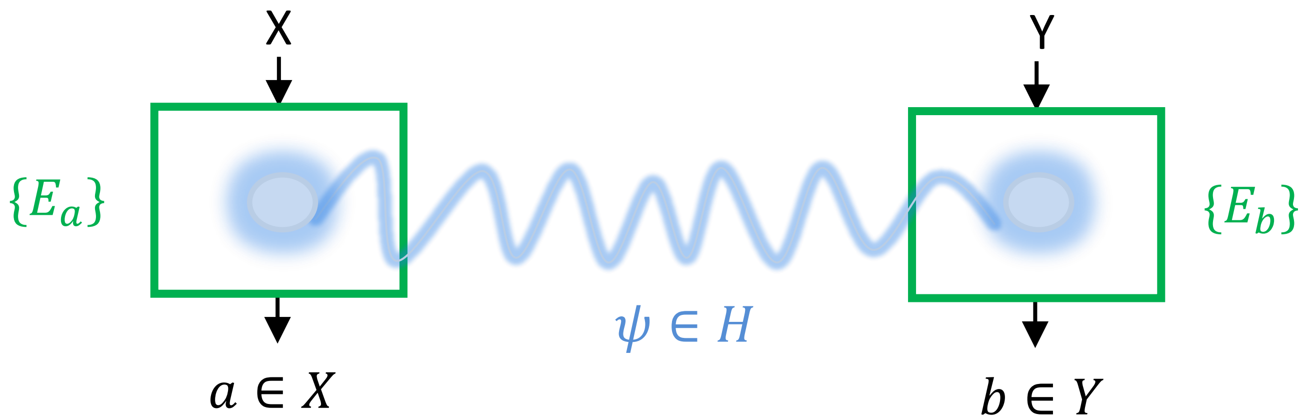
We introduced a hierarchy of semidefinite programming (SDP) relaxations which generate a sequence of lower bounds that converges to the optimum.
P., Navascuès, Acín, SIAM J. Opt. 20, 2157 (2010)

The method can be seen as a non-commutative generalization of the SDP relaxations introduced by Lasserre and Parrilo for polynomial optimization.

The method is very effective on some instances. For some class of problems, it is the only approach at our disposal.

The non-commutative setting is very natural for polynomial optimization ; it contains the commutative case, but gives more intuitive proofs of many results (*asymptotic convergence, optimality certificates, optimizer extraction*).

Quantum correlations between separated systems



Experiment characterized by joint probabilities $P(ab|XY) = \langle\psi, E_a E_b \psi\rangle$

In many problems, quantities of interest are linear functions $\sum_{ab} c_{ab} P(ab|XY)$

$$\longrightarrow \quad \max_{\psi, E, H} \quad \langle \psi, \sum_{ab} c_{ab} E_a E_b \psi \rangle$$

subject to

$$\langle \psi, \psi \rangle = 1$$

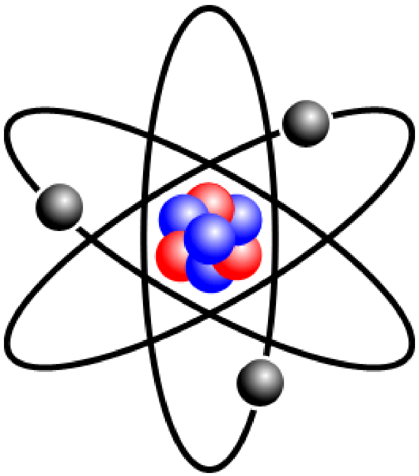
$$E_a E_{a'} = \delta_{aa'} E_a \qquad \forall\, a, a' \in X, \forall\, X \quad \text{(idem for B)}$$

$$\sum_{a \in X} E_a = I \qquad \forall X \qquad \text{(idem for B)}$$

$$[E_a, E_b] = 0 \qquad \forall\, a, b$$

Note: $\dim(H)$ is not fixed

Quantum chemistry



Fundamental problem:

Compute ground-state energy of atom or molecule comprised of N electrons that can occupy M orbitals.

$$\min_{\psi, a, a^*, H} \quad \langle \psi, \sum_{ijkl} h_{ijkl} a_i^* a_j^* a_k a_l \, \psi \rangle$$

subject to

$$\langle \psi, \psi \rangle = 1$$

$$\{a_i, a_j\} = 0 \qquad i, j = 1, \dots, M$$

$$\{a_i^*, a_j^*\} = 0 \qquad i, j = 1, \dots, M$$

$$\{a_i^*, a_j\} = \delta_{ij} \qquad i, j = 1, \dots, M$$

$$\left(\sum_i a_i^* a_i - N \right) \psi = 0$$

Note: H completely fixed by anti-commutation relations

Non-commutative polynomial optimization problem

$$
\begin{array}{lll}
\min\limits_{\psi,X,H} & \langle\psi, p(X)\psi\rangle & \\
\text{subject to} & \langle\psi,\psi\rangle = 1 & \\
& q_i(X) \succcurlyeq 0 & i = 1,\dots,m_q \\
& r_j(X)\psi = 0 & j = 1,\dots,m_r \\
& \langle\psi, s_k(X)\psi\rangle \geq 0 & k = 1,\dots,m_s
\end{array}
$$

$X = (X_1, X_2, \dots, X_n)$ is a set of bounded operators in a separable Hilbert space H .

$p(X), q_i(X), r_j(X), s_k(X)$ are (hermitian) polynomials of bounded degree d in the variables X

Example: $p(X) = 1 + 3X_1^2 + 4X_2X_1^*X_2 - X_1X_3 - 4\,X_3X_1$ (+ c.t.)

The aim is to find a vector ψ and a set of operators (i.e., matrices) X satisfying the constraints and minimizing the objective function. The size of these objects, i.e., the dimension of the underlying Hilbert space, is not fixed.

Why “non-commutative polynomial optimization” ?

$$\begin{aligned} \min_{\psi,X,H} \quad & \langle \psi, p(X), \psi \rangle \\ \text{subject to} \quad & \langle \psi, \psi \rangle = 1 \\ & q_i(X) \succcurlyeq 0 \qquad i = 1, \dots, m_q \\ & [X_j, X_k] = 0 \quad j,k = 1, \dots, n \end{aligned}$$

Since the operators X commute, they generate an abelian algebra that is unitarily equivalent to an algebra of diagonal operators.

It is then not difficult to show that the above problem reduces to the standard polynomial optimization problem

$$\begin{aligned} \min_{x} \quad & p(x) \\ \text{subject to} \quad & q_i(x) \geq 0 \quad i = 1, \dots, m_q \end{aligned}$$

in the scalar variables $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Moments and moment matrices

Let $w(X)$ be a monomial,
i.e., a world built out of the $2n$ letters $X_1, \dots, X_n, X_1^*, \dots, X_n^*$.

Example: $w(X) = 1,\ X_1,\ X_2^* X_3 X_1, \dots$

Any polynomial $p(X)$ of degree d can then be expanded as

$$p(X) = \sum_{|w| \leq d} p_w\, w(X), \qquad \text{where } p_w \in \mathbb{R}, \mathbb{C}.$$

Given a vector ψ and a set of operators X in some Hilbert space,
define for each world w the moment

$$y_w = \langle \psi, w(X)\psi \rangle \in \mathbb{R}, \mathbb{C}\,.$$

Let $y = (y_w)_{|w| \leq 2t}$ be a finite sequence of moments corresponding to words of length $\leq 2t$.
Define $M_t(y)$ as the moment matrix with rows and columns labeled by words of length $\leq t$, and with entries

$$[M_t(y)]_{v,w} = y_{v^* w}$$

Moment matrices are positive semidefinite

Proof: $M_t(y) \succcurlyeq 0 \quad \Leftrightarrow \quad z^* M_t(y) z \geq 0$ for all vectors $z \in \mathbb{R}^t, \mathbb{C}^t$.

$$z^* M_t(y) z = \sum_{v,w} z_v^* \, [M_t(y)]_{v,w} \, z_w = \sum_{v,w} z_v^* \, y_{v^* w} \, z_w$$

$$= \sum_{v,w} z_v^* \, \langle \psi, v^*(X) w(X) \, \psi \rangle \, z_w = \langle \psi, z^*(X) z(X) \, \psi \rangle \geq 0,$$

where $\mathrm{z}(\mathrm{X}) = \sum_w z_w \, w(X)$.

Further properties of moments

Let $y_w = \langle \psi, w(X)\psi \rangle$ be the moments associated to a feasible solution of the problem

$$\begin{aligned}
\min_{\psi, X, H} \quad & \langle \psi, p(X), \psi \rangle \\
\text{subject to} \quad & \langle \psi, \psi \rangle = 1 \\
& q_i(X) \succcurlyeq 0 \qquad i = 1, \dots, m_q \\
& r_j(X)\psi = 0 \qquad j = 1, \dots, m_r \\
& \langle \psi, s_k(X)\psi \rangle \geq 0 \qquad k = 1, \dots, m_s.
\end{aligned}$$

Then

$$\langle \psi, p(X), \psi \rangle = \sum_w p_w \, \langle \psi, w(X), \psi \rangle = \sum_w p_w \, y_w$$

$$\langle \psi, \psi \rangle = y_1 = 1$$

$$\langle \psi, v^*(X) r_j(X) \psi \rangle = \sum_w r_{j,w} \, y_{vw} = 0$$

$$\langle \psi, s_k(X)\psi \rangle = \sum_w s_{k,w} \, y_w \geq 0$$

Localizing matrices

The condition $q_i(X) \succcurlyeq 0$ implies that the localizing matrix $M_t(q_i y)$ with rows and columns labeled by words of length $\leq k$, and with entries

$$[M_t(q_i y)]_{v,w} = \sum_u q_{i,u}\, y_{v^* u w}$$

is positive semidefinite.

Proof: $M_t(q_i y) \succcurlyeq 0 \quad \Leftrightarrow \quad z^* M_t(q_i y) z \geq 0$ for all vectors $z \in \mathbb{R}^t, \mathbb{C}^t$.

$$z^* M_t(q_i y) z = \sum_{v,w} z_v^* \, [M_t(q_i y)]_{v,w} \, z_w = \sum_{v,w} \sum_u z_v^* q_{i,u} \, y_{v^* u w} \, z_w$$

$$= \sum_{v,w} z_v^* \, \langle \psi, v^*(X) q_i(X) w(X) \, \psi \rangle \, z_w = \langle \psi, z^*(X) q_i(X) z(X) \, \psi \rangle \geq 0,$$

where $\mathrm{z}(\mathrm{X}) = \sum_w z_w \, w(X)$.

Putting everything together

To any feasible solution ψ, X, H of our optimization problem

we can associate finite sequences $y = (y_w)_{|w| \leq 2t}$ of moments (one for each value of $2t \geq d$) such that.

$$\min_{\psi, X, H} \quad \langle \psi, p(X), \psi \rangle \quad \rightarrow \quad \sum_w p_w \, y_w$$

$$M_t(y) \succcurlyeq 0$$

$$\text{subject to} \quad \langle \psi, \psi \rangle = 1 \quad \rightarrow \quad y_1 = 1$$

$$q_i(X) \succcurlyeq 0 \quad \rightarrow \quad M_{t-d}(q_i y) \succcurlyeq 0$$

$$r_j(X)\psi = 0 \quad \rightarrow \quad \sum_w r_{j,w} \, y_{vw} = 0$$

$$\langle \psi, s_k(X)\psi \rangle \geq 0 \quad \rightarrow \quad \sum_w s_{k,w} \, y_w \geq 0$$

Hierachy of SDP relaxations

$$
\begin{aligned}
p_* = \min_{\psi, X, H} \quad & \langle \psi, p(X), \psi \rangle \\
\text{subject to} \quad & \langle \psi, \psi \rangle = 1 \\
& q_i(X) \succcurlyeq 0 \\
& r_j(X)\psi = 0 \\
& \langle \psi, s_k(X)\psi \rangle \geq 0
\end{aligned}
$$

$\Rightarrow$

$$
\begin{aligned}
p_t = \min_{y} \quad & \sum_w p_w \, y_w \\
\text{subject to} \quad & M_t(y) \succcurlyeq 0 \\
& y_1 = 1 \\
& M_{t-d}(q_i y) \succcurlyeq 0 \\
& \sum_w r_{j,w} \, y_{vw} = 0 \\
& \sum_w s_{k,w} \, y_w \geq 0
\end{aligned}
$$

$$p_* \geq \cdots \geq p_{t+1} \geq p_t$$

Asymptotic convergence

Archimedean assumption

Suppose that the polynomials $q_i(X)$ are such that

$$C - \sum_k X_k X_k^* = \sum_i f_i(X) f_i^*(X) + \sum_{ij} g_{ij}(X) q_i(X) g_{ij}^*(X)$$

for some $C \geq 0$ and some polynomials $f_i(X)$, $g_{ij}(X)$.

This implies that any set of operators X satisfying $q_i(X) \succcurlyeq 0$ (i.e. any feasible solution of the optimization problem) must be bounded: $\sum_k X_k X_k^* \preccurlyeq C$.

Then

$$\lim_{t \to \infty} p_t = p_*$$

The proof is constructive. From the sequence of optimal solutions of the SDP relaxations, we show how to construct optimal ψ and X achieving p^*. The optimal ψ and X are infinite-dimensional.

Optimality at a finite relaxation step

Rank-loop condition

If the moment matrix at relaxation step t satisfies

$$\text{rank}\, M_t(y) = \text{rank}\, M_{k-d}(y)$$

where d is the maximal degree of the polynomials $q_i(X) \succcurlyeq 0$,

then

$$\boldsymbol{p_t = p^*}$$

Furthermore, there is an explicit procedure to build the optimal ψ and X out of the moment matrix $M_t(y)$. They are defined in a vector space H of dimension

$$\mathbf{dim}\, \boldsymbol{H} = \mathbf{rank}\, \boldsymbol{M_t(y)}$$

Dealing with equality constraints

Suppose that the problems contains polynomial equality constraints $e_i(X) = 0$.

A clever way to deal with such constraints is to express every polynomial modulo the ideal $I = \sum_i f_i e_i g_i$, that is, to work using a monomial basis B for the quotient ring P/I.

All results still hold if the relaxations are build from such a monomial basis B.

Link with commutative polynomial optimization

$$\begin{aligned} \min_{\psi, X, H} \quad & \langle \psi, p(X), \psi \rangle \\ \text{subject to} \quad & \langle \psi, \psi \rangle = 1 \\ & q_i(X) \succcurlyeq 0 \\ & [X_j, X_k] = 0 \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \min_{x} \quad & p(x) \\ \text{subject to} \quad & q_i(x) \geq 0 \end{aligned}$$

↓

Equality constraints:
use commutative monomial basis:
$1, X_1, X_2, X_1^2, X_1 X_2, X_2^2, \ldots$

↓

SDP relaxations	=	Lasserre- Parrilo SDP relaxations
Rank condition for optimality	=	Curto – Fialkow flat extension
Optimizer extraction	=	Henrion - Lasserre

Dual approach and sums of squares

$$\begin{array}{ll} \min\limits_{\psi,X,H} & \langle \psi, p(X), \psi \rangle \\ \text{subject to} & \langle \psi, \psi \rangle = 1 \\ & q_i(X) \succcurlyeq 0 \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \max\limits_{\lambda,X,H} & \lambda \\ \text{subject to} & p(X) - \lambda \succcurlyeq 0 \\ & q_i(X) \succcurlyeq 0 \end{array}$$

Helton and McCullough positivstenllensatz for non−commutative polynomials :

$$p(X) - \lambda \succcurlyeq 0 \text{ on } q_i(X) \succcurlyeq 0$$

$$\Updownarrow$$

$$p(X) - \lambda = \textstyle\sum_i f_i(X) f_i^*(X) + \sum_{ij} g_{ij}(X) q_i(X) g_{ij}^*(X) = \text{S.O.S}$$

Dual of the moment-based SDP relaxations

➔ SDP relaxations:

$$\begin{array}{ll} \lambda_t = \max\limits_{\lambda.f_i,g_{ij}} & \lambda \\ \text{subject to} & p(X) - \lambda = \text{S.O.S} \\ & \deg(\text{S.O.S}) \leq 2t\,, \end{array}$$

$$\lambda_t \leq \lambda_{t+1} \leq .. \leq p_* \text{ and } \lim_{t\to\infty} \lambda_t = p_*$$

Applications

Quantum correlations between separate subsystems: method already introduced in Navascuès, P., Acín 07, Navascuès, P. Acín 08, Doherty et al 08.

- Effective: in practice, convergences observed at low-order relaxations Vertesi-Pal: tested 241 Bell inequalities, 3rd relaxation yields optimum for 225
- The fact that it provides *lower-bounds* is very important (quantum crypto).
- Basically, only algorithm that we have.

Quantum chemistry: low-order relaxations coincide with « reduced-density-matrix methods »

- Good: robustness, high-accuracy.
- Bad: computational time, memory consumption.

Other problems (many-body physics, Weyl algebra, …) : to explore !

References

- P., Navascuès, Acín, SIAM J. Opt. 20, 2157 (2010).

- Navascuès, P., Acín, in Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications, edited by M. Anjos and J. Lasserre, forthcoming.