Gain a problem, and an LP relaxation. If the relaxation is weak we can add more constraints in an ad hoc way (as inspired by the bad examples and the structure of the problem). But there are also ways to add constraints automatically, thus these two ways of looking at it.

Gain a combinatorial optimization problem (such as the cube $\{0,1\}_n$), suppose we have an LP relaxation. And suppose this relaxation has the property that the $K_4 \cap \{0,1\}_n = \emptyset$. Then we can generate feasible inequalities mechanically. Many ways. Here are some.

1. Gomory-Cut (cutty planes): if current constraints are $\exists x \geq \beta$ with $x \in \mathbb{Z}^n$, then can infer $\exists x \geq \lfloor \beta \rfloor$. So do this for base LP to get a "lift". How many times do we need to lift until get only convex hull of integer solns?

   - Gomory: Finite bound on the "route".
   - Fenebrond & Schuyl : $O(n^2 \log n)$, for 0-1 polyhedra.
   - Problem: don't know how to take lift in polyhedral

   (BTW, Matching polytope = lift of base polytope $\leq$ times).

   for non-bip

2. Lovasz Schrijver / Sherali Ahmed / Lassere

   - All known to have rank $n$. Can be implemented with $O(n^4 \log n)$ for $k$ lift.
   - We'll talk about Sherali Ahmed (LP) & Lassere / Lassere (SDP).

   Sherali Ahmed
What if we those in constrants that are valid, generated by the following steps:

Add the constrants obtained by the products of variables + linearify + circunlize in non-linear

e.g., we introduce variables $X_{ij} = x_i \cdot x_j$

$x_{ijk} = x_i \cdot x_j \cdot x_k$ etc.

What kind of constrants?

Obtained by taking \( \forall x \geq 0 \) and multiplying by \( TT(x) \cdot TT(1-x) \)

and setting \( x^2 = x \) (valid for 0-1 solutions) and linearifying.

**Example:** Stable set \( \max \sum x_i \)

\[ s.t. \ x_i + x_j \leq 1 \quad \forall i,j \in E \\
0 \leq x_i \leq 1 \quad \forall i \in V \]

Now multiply each constrant by \( x_i \), and by \( (1-x_i) \) etc.

So get \( (x_i + x_j)x_i \leq x_i \)

\[ \Leftrightarrow x_i^2 + x_i x_j \leq x_i \Leftrightarrow x_i x_j = 0 \\
\text{(since } x_i = x_i) \]

Define vars \( X_{ij} \) and this says \( X_{ij} = 0 \quad \forall i,j \in E \)

but also \( x_{ik} + x_{jk} \leq x_k \)

and so have a variable for each set \( S \) of size \( \leq 2 \).

\[ x_S \]

(Don't forget to use \( 0 \leq x_i \leq 1 \) to get \( 0 \leq x_i \cdot TT(x) \cdot TT(1-x) \leq \)

\[ \forall x \in S \quad \text{best} \]

Can do this for k-level SA

by taking \( 1S + 1T \leq k \) and along the operators
Does this help?

On the 5-th cycle:

\[ x_1 + x_2 \leq 1 \]
\[ x_2 + x_3 \leq 1 \]
\[ x_3 + x_4 \leq 1 \]
\[ x_4 + x_5 \leq 1 \]
\[ x_5 + x_1 \leq 1 \]

\[ \Rightarrow (\sum_{i=1}^{5} x_i)^2 \leq 1 \]
\[ x_1 + x_2 + x_3 + x_4 + x_5 \leq 2 \]

So the multiplying by \( \prod_{i} x_i \prod(1-x_i) \) is not too useful, but it's where you use \( x_i^2 = x_i \) (only for \( 0 \leq x_i \leq 1 \)).

And then linearize (which is messy but necessary for phrasing).

Intuition: get "local distributions" over sets of size \( S \leq k \) (for take \( k \)-levels) i.e.

For distribution \( \rho \) on set \( S \) of size \( |S| \leq k \).

subject to \( \text{Pr}_{\rho(S)}[\bigwedge_{i \in S} (x_i = 1)] = x_3 + x_4 + x_5 \leq 5 \)

And consistency (i.e. if \( S \subseteq S' \) then \( \text{Pr}_{\rho(S)}[\bigwedge_{i \in S'} (x_i = 1)] = \text{Pr}_{\rho(S')}[\bigwedge_{i \in S'} (x_i = 1)] \).

Imp: These distributions satisfy all constraints on \( \leq k \) vars. (and more)

So locally "looks like" solutions.

Of course, these may not be a global one that agrees with the local solutions.

But maybe the integrality gap has improved.

Bad news: for Ind-Set, does not capture clique constraints (which the standard SDP, i.e. convex relaxation, captures).

\[ \text{for Max Cut, integrality gap remains } 2 - \varepsilon \text{ even after } O(n) \text{ rounds}. \]

Good news: Helps for bounded TW problems, (Ind-Set, Sparsest Cut etc.).

And for other problems where "remembering small state" works.
Looserre idea: Similar idea, but also SDP constraints.

e.g. for independent set \[ \max \sum_{i} ||x_i||^2 \]

\[ \text{ s.t. } \langle x_i, x_j \rangle = 0 \quad \forall \quad i \neq j \]

\[ \langle x_s, x_t \rangle = \langle x_{s'}, x_{t'} \rangle \quad \text{if} \quad \text{SUB} = \text{S'} \times \text{T'} \]

\[ ||x_\phi||^2 = 1. \]

Says things like \[ \langle x_i, x_\phi \rangle = \langle x_i, x_i \rangle \]

\[ \langle x_i, x_{\phi \cup j} \rangle = \langle x_i, x_j \rangle \]

and

all use that intended solutions satisfy \[ x_i^2 = x_i \]

Gives a lot more power. Don't know bad examples any more, for lots of problems.

In fact, all "standard" bad examples fail — for many problems.

This is the optimal LP for constraint satisfaction problems (unless UC fails).

Bad examples: \( k \)-XOR — linear rounds still cannot detect satisfiable instances from unsat.

Dynamic programming problems like knapsack not exactly solved.

But still — don't know how to use it, don't know bad examples, for many problems.
Here's a different perspective coming from polynomial optimization.

Want to solve  \( \max \{ P(x) \} \)

\[ \text{st } f(x) \geq 0 \quad \forall f \in \mathcal{F} \]

semialgebraic set

\[ g(x) = 0 \quad \forall g \in \mathcal{G}. \]

This set denoted by \( \mathcal{K} \) is not convex in general, may not even be connected.

\[ \text{eg. } \sum_{i=1}^{n} x_i^2 = 1 \text{ (surface of unit ball) or } x_i = x_i + v_i \text{ (cube).} \]

Let's make it a convex problem by considering a measure \( \mu \) on \( \mathcal{K} \), now we want to optimize

\[ \max_{x \in \mathcal{K}, \mu} \mathbb{E}_{\mu}[P(x)] \]

\[ \text{st } \mu \text{ is supported on } \mathcal{K}. \]

Too many degrees of freedom! Suppose \( P(x) \) has degree \( d \). Then suppose we say that

\[ \mathbb{E}_{\mu}[x^k] = L(z^k) \quad \text{and } L \text{ is a linear map (no exponent to all degree-d polynomials by)} \]

\[ L(\varphi(x)) = \sum_{k \in \mathbb{Z}^n} \varphi_k \cdot L(x^k) \]

Then we want to maximize

\[ \max L(\varphi(x)) \]

\[ \text{st } L(1) = 1. \]

\( L \) is a linear map

\[ \mu \text{ is a measure on } \mathcal{K} \text{ that has moments } L(z^k). \]

This is a hard problem. So let's relax the problem slightly.

For all \( f \in \mathcal{F} \), add constraint \( L(f^2) \geq 0 \quad \forall q \text{ of degree } d - \deg(f) \)

For all \( g \in \mathcal{G} \), add constraint \( L(g^2) = 0 \quad \forall q \text{ of degree } d - \deg(g) \).

So finally we have:

\[ \max L(\varphi(x)) \]

\[ \text{st } L(1) = 1, \]

\[ L(\varphi(z^k) \geq 0. \quad \forall q \text{ of degree } d - \deg(f) \quad \forall f \in \mathcal{F} \]

\[ L(g^2) = 0 \quad \forall q \text{ of degree } d - \deg(g) \quad \forall g \in \mathcal{G}. \]
Claim: This is an SDP with \( n \times d \) size.

If sketch:

For the \( g_s \), sufficient to check
\[
L(g_s^T \Phi^{1/2}) = 0 \implies \sum_{x} g_s \cdot L(x^T \Phi g_s) = 0.
\]

Variables in an LP!

For each \( f \),
\[
L(f(x) q(x^T) = \sum_{b,i} f(x) q_b \cdot q_i \cdot L(x^T \Phi g_s) \geq 0 \quad \forall g_s \in \mathbb{R} \}
\]
\[
\iff \left( \sum_{x} f(x) L(x^T \Phi g_s) \right) \geq 0. \quad \text{psd ness.}
\]

Dual SOS

\[
\begin{align*}
\inf \lambda & \\
\text{st.} & \quad \mu = \lambda - \mathbf{1}^T \Phi \\
& \quad \lambda \in \mathbb{R} \\
& \quad \nu \in \mathcal{Q}_d(F \cup \mathbb{R}_b) \\
& \quad \omega \in \mathcal{I}_d(G)
\end{align*}
\]

Notation:
\[
\begin{align*}
\mathcal{Q}_d(F) & = \text{"quadratic module" of degree } \leq d. \text{ wrt } F \\
& = \text{cone } \left( f q^2 : q \in \mathbb{R}^{d}\text{-deg}(x) \right) \\
\text{Id}(G) & = \text{"ideal" of degree } \leq d. \text{ wrt } G \\
& = \text{span } \left( f q^2 : q \in \mathbb{R}^{d}\text{-deg}(x) \right)
\end{align*}
\]

Weak duality: given \( L \), and \( (\lambda, \nu, \omega) \)
\[
L(\mu) = L(\lambda - \mathbf{1}^T \Phi) = \lambda - L(\nu) - L(\omega) \leq \lambda.
\]

Note: In general, we could throw in more conditions, e.g.
\[
L((T f) g^2) > 0 \quad \forall T \in F.
\]

But don't need this for "well conditioned" ("Archimedean") problems.
Before we give the convergence result of Lasserre (which says that as $d \to \infty$, the degree of relaxation and the degree of $O(\delta)$ both converge to $O(\delta)$).

Let's give some perspective.

LP duality (Farkas' Lemma) says that

dsp. $A x \geq b$ implies $x \geq 0$, i.e., $\alpha \geq 0$ for all points in $K = \{x \in \mathbb{R}^n \mid A x \geq b\}$

then we can write the inequality $A x \geq b$ by convex combination of $\sum \alpha_i x \geq b_i 3_i$.

Equivalently, if $A x \geq b$ is inconsistent, i.e., if no $x$ s.t. $A x \geq b$, then there exists a linear combo of constraints that proves $-1 \geq 0$.

What about poly systems in general?

Positivstellensatz [Schmüdgen]

- If $K_{FG}$ is compact, then if $f(x)$ is positive on $K_{FG}$, then

  
  $$f(x) = \sum_{i=0}^{\infty} h_i(x) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{i!j!k!} \prod_{i 

  \text{I(G) ideal members}

  \text{or in the contrapositive form. if } K_{FG} \text{ is empty, then}

  -1 = \text{SOS} + \alpha(x) \text{ or this form.}

  \text{I(G)}

  \text{Positivstellensatz [Putinar]}

  - If $K_{FG}$ is Archimedean, then don't need the products of the $f$s. i.e., with the

  Archimedean assumption, if $p(x) > 0$ on $K_{FG}$

  $$p(x) = \text{SOS} \text{ I(G)} + Q(FU \{1\})$$

  Note: No bound on the degrees of the polynomials on the right.!! Every $p(x)$, $f(x), g(x)$ all have low degrees.
What in Archimedean?

It means that the statement $\|x\|^2 \leq R$ can be proven in the same system.

That is $\exists R : \|x\|^2 \in [0, R]$.

This implies that $K_{R,\infty}$ is compact.

OK. So we have the theorem saying the system is sound and complete. But the degrees of the proofs may not be controlled. And the convergence theorem of Lasserre says pretty much that as the $d \to \infty$, the SOS/Lasserre systems converge to OPT.

So direction in TCS: What statements have low degree proofs? (and use small numbers — see Ryan's paper) that set came.

E.g., the GW max-cut result implies that we prove

\[ \text{MAX Cut} - 0.878 \text{ SDP} \leq \text{SOS}. \text{ of small degree} \]

\[ \implies \text{SDP} \geq \frac{\text{MAX Cut}}{0.878} \text{ and so integrality gap is not too large}. \]

- Can we close that gap yet better with higher degree SOS?
- What limits to this general technique?