Given a problem, and an LP relaxation, if the relaxation is weak we can add more constraints in an additive way (or inspired by the bad examples, and the structure of the problem). But there are also ways to add constraints automatically; thus, there are two ways of looking at it.

Given a combinatorial optimization problem (over the cube $\{0,1\}^n$), suppose we have an LP relaxation. And suppose this relaxation has the property that the $K_{\text{LP}} \cap \{0,1\}^n = \text{Polytope of convex hull of solution set of feasible solutions}$. Then we can generate feasible inequalities mechanically. Many ways; here are some.

1. Gomory-Chvátal (cutting planes): if current constraints for $Ax \geq b$ with $x \in \mathbb{Z}^n$, then we can infer $Ax \geq [b]$. So do this for base LP to get a "lift". How many times do we need to lift until we get only convex hull of integer solns?
   - Chvátal: finite bound on the "radius"
   - Frenkel and Schrijver: $O(n^{20}n^{19})$, for 0-1 polytope.
   - Problem: don't know how to take lift in polyhedral

   [BTW, Matching polytope = lift of base polytope + slacks]

2. Lovász, Schrijver, Shekali Androu, Lassere.
   - All known to have rank $n$. Can be implemented with $O(n^2 \log n)$ for $k$ lift.
   - We'll talk about Shekali, Androu (LP) & Lassere (SDP).

---

Shekali Androu

Gomory, Cottle, Tardos, Ellipsoid, Lassere, Polyhedral, Shekali, Androu
What if we choose constraints that are valid, generated by the following steps:

Add all constraints obtained by multiplying variables + linearizing + circumscribing variables.

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

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

Robot constraints?

Obtained by taking $a_i x_i \geq 0$ and multiplying by $T(x_i) T(x_j)$

and setting $x_i^2 = x_i$ (valid for 0-1 solutions) and linearizing.

Example: Stable set

$$\text{max } \sum x_i$$

$$\text{st } x_i + x_j \leq 1 \forall i \neq j \in V$$

$$0 \leq x_i \leq 1 \forall i \in V$$

Now multiply each constraint by $x_i$, and by $(1-x_i)$ etc.

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

$\Leftrightarrow x_i^2 + x_i x_j \leq x_i \Leftrightarrow x_i x_j = 0.$

(see $x_i^2 = x_i$).

Define vars $X_{ij}$ and this says $X_{ij} = 0 \forall i \neq j \in V$

but also $x_i x_k + x_j x_k \leq x_k$

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 $0 \leq x_i T(x_i) T(x_j) \leq x_i \text{ for } S 	ext{ of size } \leq 2$)

Can do this for t-level SA

by taking $1S + 1T \leq t$ and doing the operation

---
Does this help?

On the 5th cycle:

\[ x_1 + x_2 \leq 1 \quad x_1 \]
\[ x_2 + x_3 \leq 1 \quad (1-x_1) \]
\[ x_3 + x_4 \leq 1 \quad x_1 \]
\[ x_4 + x_5 \leq 1 \quad (1-x_1) \]
\[ x_5 + x_1 \leq 1 \quad x_1 \]

\[ \Rightarrow (x_1^2 + x_3 x_2) \leq 1 \]
\[ x_2 + x_3 - x_2 x_3 - x_3 x_2 \leq 1 - x_1 \]
\[ x_3 x_5 + x_3 x_4 \leq x_1 \]
\[ x_4 + x_5 \leq x_1 \]
\[ x_1 x_5 + (1-x_1)^2 \leq 1 \]

\[ \Rightarrow x_1 + x_3 + x_4 + x_5 \leq 2 \]

So the multiplying by $\prod_{i \leq s} x_i^{r_i}$ is not too useful, but its where you can $x_i^2 = x_i$ (only for $0 \leq s \leq 5$).

And then linearize (which is easy, but necessary for polynomial).

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

For a distribution as on set $B$ of size $1 \leq k$.

\[ \text{st. } \mathbb{P}_{x \in B} \left[ \bigwedge_{i \leq s} (x_i = 1) \right] = x_S', \quad s' \leq s. \]

And continuity also (i.e. if $s' \leq s$ then $\mathbb{P}_{x \in B} \left[ \bigwedge_{i \leq s'} (x_i = 1) \right] = \mathbb{P}_{x \in B} \left[ \bigwedge_{i \leq s} (x_i = 1) \right]$).

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

So locally "look alike" solutions.

Of course, there may not be a global $\mathbb{P}$ that agrees with the local solutions.

But may be the integrality gap has improved.

Bad news: for Ind-Set, does not capture clique constraints (which essentially basic SDP, i.e. degree-$2$ function, captures).  

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

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

and for other problems where "remembering small state" works.
Laurent Scho: Similar idea, but also SDP constraints.

e.g. for independent set $\max \sum \|x_i\|^2$

$\mathbf{s.t.} \quad \langle x_i, x_j \rangle = 0 \quad \forall i \neq j$

$\langle x_s, x_t \rangle = \langle x_{S'}, x_{T'} \rangle \text{ if } SUT = S'UT'$

$\|x_p\| = 1.$

Says things like $\langle x_i, x_p \rangle = \langle x_i, x_i \rangle$ to scale down to 1.

and $\langle x_i, x_{p+j} \rangle = \langle x_i, x_j \rangle$

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 UGC 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 \( \max P(x) \)
\[
\text{st } f(0) \geq 0 \quad \forall f \in F \quad \text{semi-algebraic set} \\
g(x) = 0 \quad \forall g \in G.
\]

This set denoted by \( K_F G \) is not convex in general, may not even be connected.

eg. \( \sum_x x^2 = 1 \) (surface of unit ball) or \( x^2 = x + vi \) (cube).

Let's make it a convex problem by considering a measure \( \mu \) on \( K_F G \), now we want to optimize

\[
\max_{x \in X \mu} \mathbb{E}[P(x)]
\]
\[
\text{st } \mu \text{ is supported on } K_F G
\]

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

\[
\mathbb{E}[x^k] = L(z^k) \quad \text{and} \quad L \text{ is a linear map (not extending to all degree-}d \text{ polynomials by)}
\]

\[
L(g(x)) = \sum_b \beta_b L(\phi^b(x))
\]

Then we want to maximize

\[
\max \quad L(p(x)) \\
\text{st } L(1) = 1.
\]

\( L \) is a linear map that has moments \( L(z^k) \).

Let's suppose \( \mu \) is a measure on \( K_F G \) that has moments \( L(z^k) \).

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

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

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

so finally we have

\[
\left\{ \begin{array}{l}
\max P(x) \\
\text{st } L(1) = 1, \\
L(\phi^2) \geq 0. \quad \text{// think of this as 120.} \\
L(\phi^3) \geq 0 \quad \forall q \text{ of degree } \leq d - \deg(f) \quad \forall f \in F \\
L(g^2) = 0 \quad \forall q \text{ of degree } \leq d - \deg(g) \quad \forall g \in G.
\end{array} \right.
\]
Claim: this is an SDP with $n^{(d)}$ size.

If sketch:

1. For the $g$s, sufficient to check $L(g^x \cdot p) = 0 \Rightarrow \sum_x g_x \cdot L(x^+ \cdot p) = 0$. Variables in an LP!

2. For rays $f$, $L(f(x) \cdot q(x)^+ \cdot p) = \sum_{b \in \mathbb{R}^d} f_b \cdot q_b \cdot q^+ \cdot L(x^+ \cdot p+b) \geq 0 \forall q \in \mathbb{R}^d$,

   $\Rightarrow \left( \sum_x f(x) \cdot L(x^+ \cdot p+b) \right)_{b \in \mathbb{R}^d} \geq 0$. Psd cone.

Dual SOS degree-$d$

Inf $\lambda$

St. $p = \lambda - \varphi + w$

$\lambda \in \mathbb{R}$

$\varphi \in Q_d(F \cup G)$

$w \in I_d(G)$

Notation: $Q_d(F)$ = "quadratic module" of degree $\leq d$. wrt $F$

$$= \text{cone} \left( f q^2 : q \in \mathbb{R}^{d-\deg} [x] \right)$$

Total degree $\frac{d}{2}$ smaller hand.

$I_d(G)$ = "ideal" of degree $\leq d$. wrt $G$

$$= \text{span} \left( f q : q \in \mathbb{R}^{d-\deg} [x] \right)$$

Weak duality: given $L$, and $(\lambda, \varphi, w)$

$$L(p(w)) = L(\lambda - \varphi + w) = \lambda - L(\varphi) + L(w) \leq \lambda \Rightarrow \lambda \geq 0.$$

Note: in general, we could throw in more conditions, eg

$$L((\mathcal{T} f) q^2) \geq 0 \quad \forall \mathcal{T} \in \mathcal{F}.$$

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

Let's give some perspective

LP duality (Farkas' Lemma) says that $Ax \geq b$ implies $x \geq 0$, i.e. $x \geq 0 \iff \text{ primal}$

then we can write the inequality $x \geq 0$ by a combo of linear and also use $\geq 0$.

Equivalently, if $Ax \geq b$ is inconsistent, i.e. if no sol'n to $Ax \geq b$,

then $A$ a linear combo of constraints that proves $-1 \geq 0$.

What about poly systems in general?

Positivstellensatz [Schmüdgen]

$K_{F,G}$ be compact. then if $f(x)$ is positive on $K_{F,G}$ then

$$f(x) = \sum_{i=0}^{1} h_i(x) + \sum_{i=0}^{1} f_i(x) q_i(x)$$

$I(G)$ ideal members

or in the contrapositive form, if $K_{F,G}$ is empty then

$$-1 = \sum_{i=0}^{1} h_i(x) + \sum_{i=0}^{1} f_i(x) q_i(x)$$

$I(G)$

Positiv Stellensatz [Putinar]

If $K_{F,G}$ is Archimedean, then don't need the products of the $f$'s. i.e. with the add'l Archimedean assumption, if $p(x) \geq 0$ on $K_{F,G}$

$$p(x) \geq \sum_{i=0}^{1} h_i(x) + \sum_{i=0}^{1} f_i(x) q_i(x)$$

$I(G)$

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

It means that the statement $\|x\|^2 \leq R$ can be proven in the same system. That is $\exists R, \|x\|^2 \in \mathbb{R} + \mathcal{O}(FU213)$. This implies that $K_{\Phi\Theta}$ 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 WLS/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 can be.

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

$$\text{MAXCut} - 0.878 \text{ SDP} \leq \text{SOS}, \text{ of small degree}$$

$$\implies \text{SDP} \geq \frac{\text{MAXCut}}{0.878} \text{ and so integrality gap is not too large.}$$

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