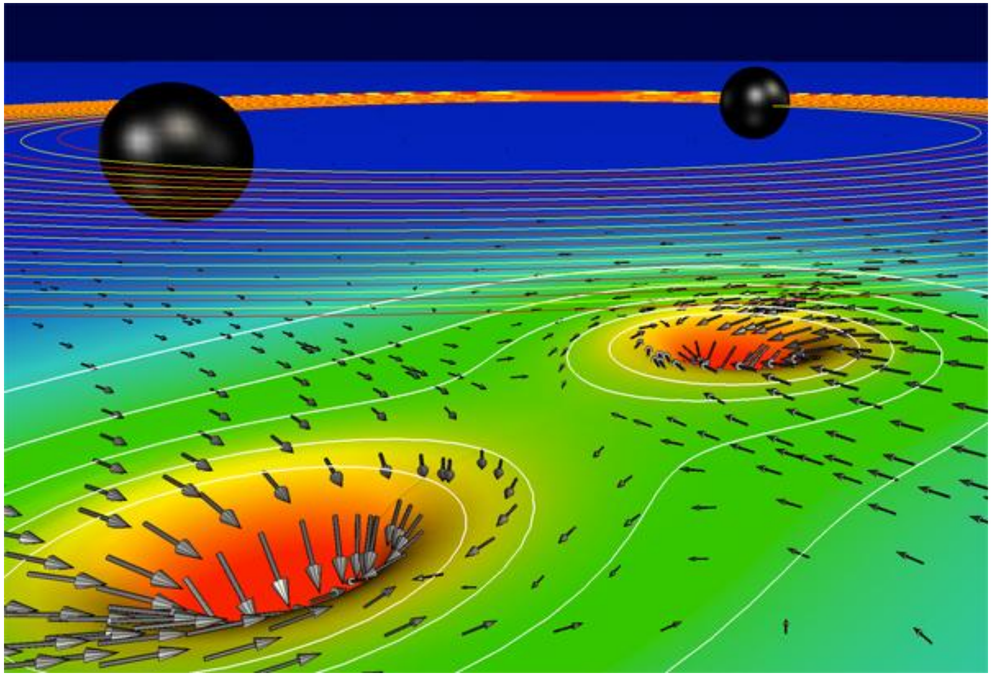
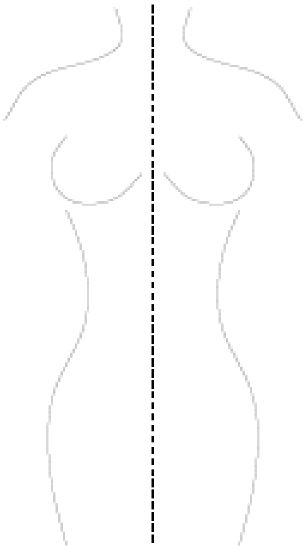
Compressive Symmetry

Matteo Fischetti, University of Padova
Leo Liberti, LIX, Ecole Polytechnique, France

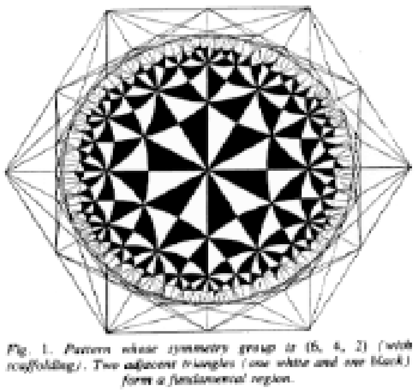


Motivation

Nature,



people,



scientists

all like symmetry

Why not us?

Symmetry for dummies

- Consider a generic optimization problem of the form

$$v(P) := \min\{f(x) : x \in F(P)\}$$

$$\text{where } F(P) \subseteq \mathcal{R}^n \text{ and } f : \mathcal{R}^n \to \mathcal{R}$$

- A symmetry permutation is an index permutation

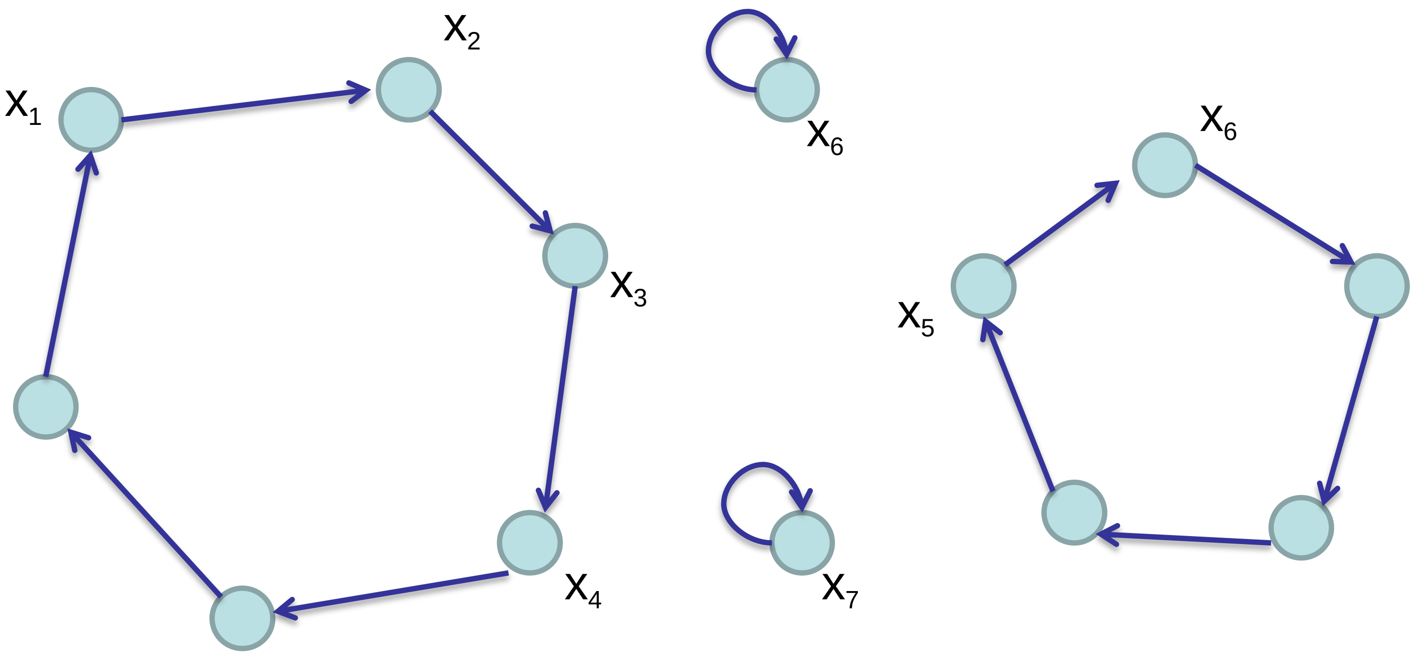
$$\pi : \{1, \cdots, n\} \to \{1, \cdots, n\}$$

such that

$$x \in F(P) \Rightarrow x' \in F(P) \quad \text{and} \quad f(x') = f(x)$$

$$\text{where } x'_{\pi(j)} := x_j \quad \text{for all } j \in \{1, \cdots, n\}$$

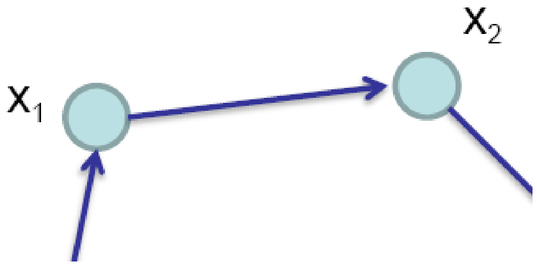
Symmetry permutation (illustration)



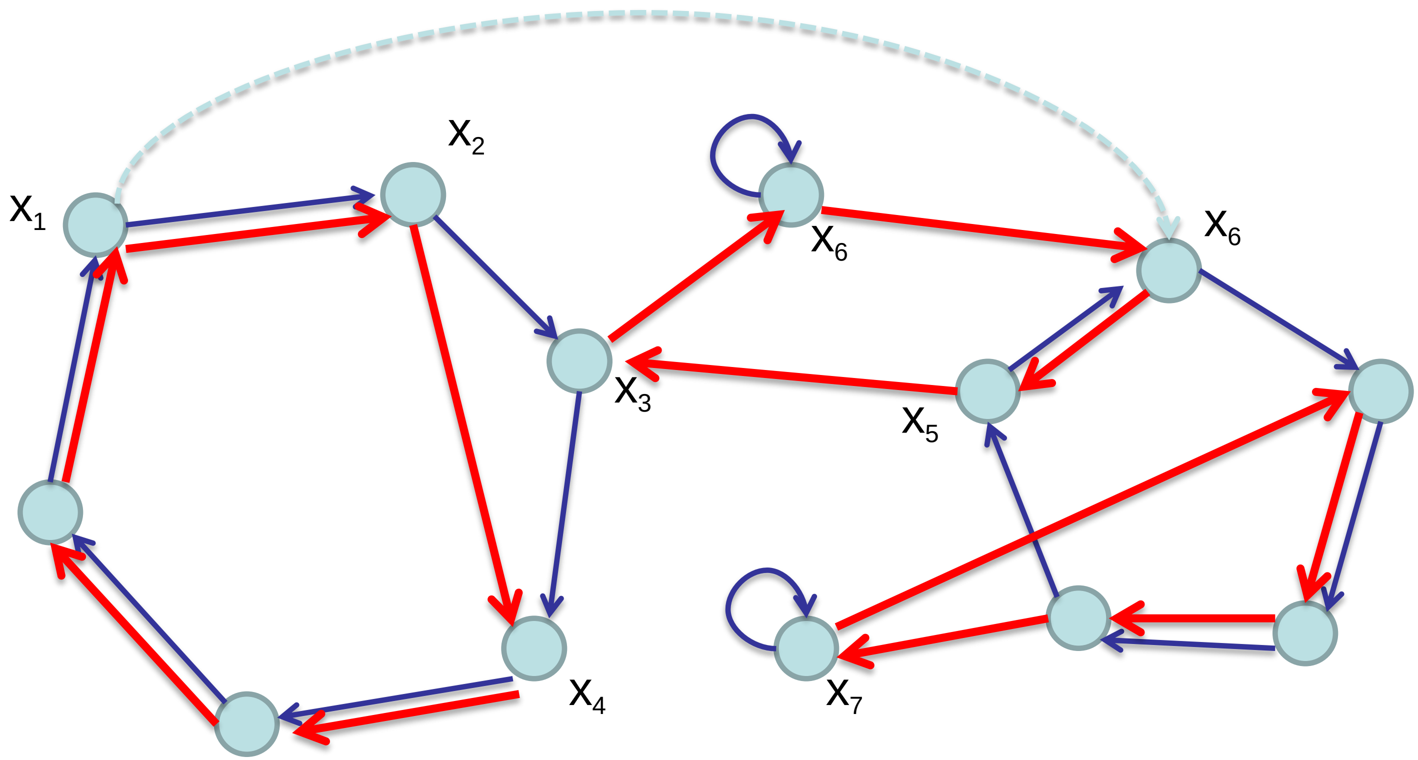
permutation = node covering by disjoint directed cycles

Symmetry group

- **Symmetry group G**: finite collection of symmetry permutations closed under composition and inversion
- **Generators** of G: set of symmetry permutations whose composition (and inversion) yields G
- **Orbits of G**: indices *i* and *j* belong to a same orbit iff there exist *π* in G such that *π(i) = j*



Generators and orbits (example)



orbits = strongly-connected components of generator graph

Nice, but… how to compute them?

- In some cases, a suitable (sub) group is known a priori
- Otherwise, it can be computed starting from the available mathematical formulation of the problem …

  … hoping the user was not so clever to use sophisticated tricks (e.g., lifted cuts) that hide symmetry
- Reasonably fast in practice through sw such as saucy, nauty, etc.
- So, let’s assume a suitable symmetry group G has been computed with “reasonable” overhead

Symmetry and convex optimization

- Assume both *F(P)* and *f* are convex
- Fact (Parrilo, 2003): an optimal solution $\overline{x}$ exists such that

$$\overline{x}_j = \text{ const. within each orbit } O_h$$

Argument:

- if *f* is <u>strictly</u> convex, a unique optimal solution *x** exists, so each π in G must leave *x** unchanged → equal value within each orbit
- if not, just consider a second-level strictly convex function ρ to break ties…

Claim: there exists an optimal solution $\overline{x}$ such that

$$\overline{x}_j = \text{ const. within each orbit } O_h$$

Indeed, let x^* be an optimal solution with minimum

$$\rho(x^*) := \sum_{j=1}^{n} (x_j^*)^2$$

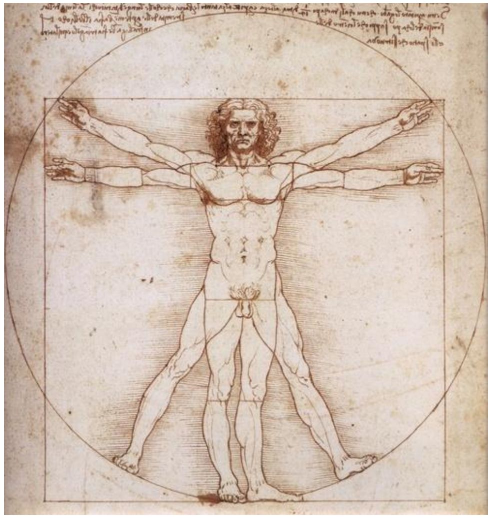
Assume, by contradiction, that there exist $i, j \in O_h$ such that $x_i^* \neq x_j^*$

- As i, j belong a same orbit, there exists π in G such that $\pi(i) = j$.
- Let y^* be an optimal solution obtained from x^* by applying π, where $x^* \neq y^*$ and $\rho(y^*) = \rho(x^*)$.
- By convexity, the point $z^* := (x^* + y^*)/2$ is an optimal solution as well, while $\rho(z^*) < (\rho(x^*) + \rho(y^*))/2 = \rho(x^*)$ because ρ is strictly convex.

$$\begin{array}{rl} & \;\; i \;\; j \\ x^* = & (2, 1, 1, 1)(3, 4) \quad (5, 5)(1, 1, 1, 1) \\ y^* = & (1, 2, 1, 1)(4, 3) \quad (5, 5)(1, 1, 1, 1) \\ & |\leftarrow \;\; O_1 \;\; \rightarrow| \quad |\leftarrow \;\; O_2 \;\; \rightarrow| \end{array}$$

Symmetry helps in the convex case

- Because of the above, the only unknowns in the convex case are the *k* values of $\overline{x}$ inside $O_1, \ldots, O_k$



- Exact reformulation
  1. introduce additional variables

$$z_h = \sum_{j \in O_h} x_j, \quad h = 1, \cdots, k$$

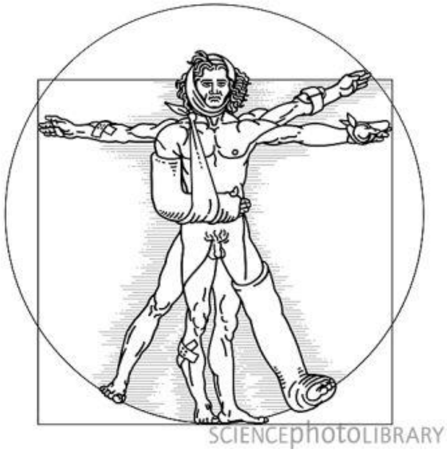
  2. project the formulation on the z-space, by just replacing

$$x_j \to z_h / |O_h| \qquad \text{for all } j \in O_h, \; h = 1, \cdots, k$$

  3. solve the projected model on the z-space

Symmetry hurts in the nonconvex case

- Unfortunately, the average point $\overline{x}$ can be infeasible/nonoptimal in the nonconvex case → **reformulation does not work**!
- Even worse: in the discrete case, **enumeration** is tricked by symmetry (symmetric subtrees can be visited again and again…)
- Possible remedies:
- a**. Break symmetry** somehow
  → a potentially useful property
  is not fully exploited!



- b. Modify branching rules (**isomorphism pruning, orbital branching**) to take advantage of symmetry → a powerful idea!

Symmetry in convex MI(N)LP

- Consider the convex Integer (N)LP

$$(P) \quad v(P) := \min\{f(x) : x \in F(P),\ x \text{ integer}\}$$

where $F(P) \subseteq \mathcal{R}^n$ and $f : \mathcal{R}^n \to \mathcal{R}$ are both **convex.**

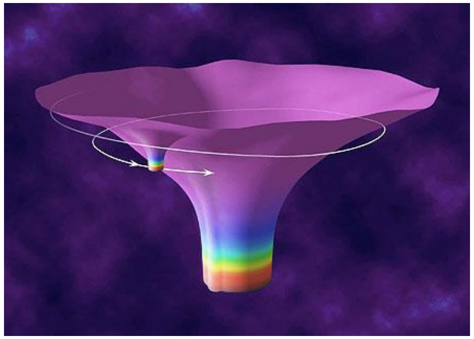
(mixed-integer case very similar, with integer/continuous orbits)

- Feasible set is **nonconvex** → reformulation on the z-space is **not** exact

- Can we exploit the symmetry group anyway?

- E.g., within an enumerative method, at each node compute the symmetry group after branching, and solve the convex continuous <u>relaxation</u> on the z. (instead of x-) space

Orbital Shrinking



- **Idea**: relax "individual integrality" into "surrogate integrality"

- **EXTEND:** Introduce additional z-variables $z_h = \sum_{j \in O_h} x_j, \quad h = 1, \cdots, k$
  along with the integrality requirement z_h integer, $h = 1, \cdots, k$

- **RELAX**: Remove the integrality requirement on *x* (**but not on *z***) to obtain a "**blurred relaxation**" → still a convex MI(N)LP with the same symmetry group

- **SHRINK**: Reformulate exactly the blurred relaxation on the z-space by just replacing

  $$x_j \to z_h/|O_h| \qquad \text{for all } j \in O_h, \ h = 1, \cdots, k$$

  → still a convex MI(N)LP but of smaller size and with **no symmetry left**

- **SOLVE**: Solve the shrunken MI(N)LP relaxation to get a lower bound → hopefully much easier than solving the original problem [smaller/no symmetry]

A familiar example

- Consider the Asymmetric TSP on a complete digraph…
  … and take an instance with **symmetric** arc costs $c_{ij} = c_{ji}$
- Very inefficient because of symmetry → orbital shrinking will automatically detected orbits

$$(x_{12}, x_{21})\,(x_{13}, x_{31}) \cdots (x_{ij},\; x_{ji}) \cdots$$

- … and introduce orbital integer variables $z_{\{ij\}} := x_{ij} + x_{ji}$
- 2-node SECs → $z_{\{ij\}} = x_{ij} + x_{ji} \leq 1$ → $z_{\{ij\}} \in \{0, 1\}$
- In this case, orbital shrinking yields an **exact reformulation**: optimize on the *z*-space (STSP), get an optimal (integer) *z**, and then optimize on the *x*-space with restriction $x_{ij} + x_{ji} = z^*_{\{ij\}}$ to get an optimal *x**

Discussion

- Can we expect to get a tight relaxation in all cases?
- Certainly **NOT** when
  - a single orbit (or just few) exists → 100% symmetrical formulations
  - Removing detailed integrality on the single *x*'s oversimplifies the problem → trivial relaxation on the *z*-space
    - e.g. bin packing problem with 2-index (item,bin) x-variables
- Hopefully **YES** when the blurred relaxation still has a structure that requires nontrivial branching/cuts to be solved
  - rich structure induced by integrality of the *z* var.s only

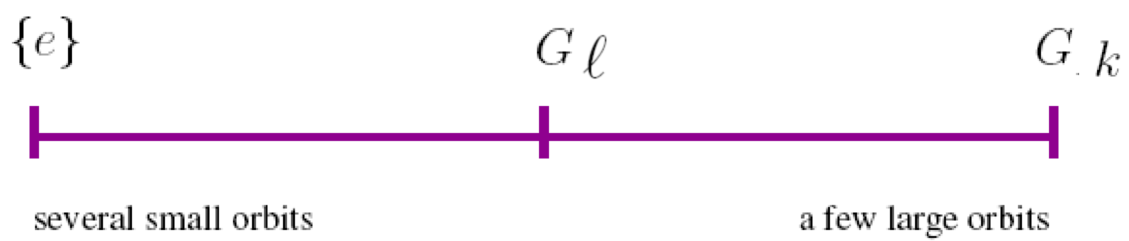
Illustrative experiments

- Testbed: Margot's website (100% symmetrical instances)

- Working with a subgroup of *G* generated by a subset of generators
  - Tradeoff between size and expected tightness of the shrunken relaxation

- Idea:
  1. sort the k (say) generators somehow → generators with small cycles first…
  2. consider the subgroup of G induced by the first ℓ (say) generators

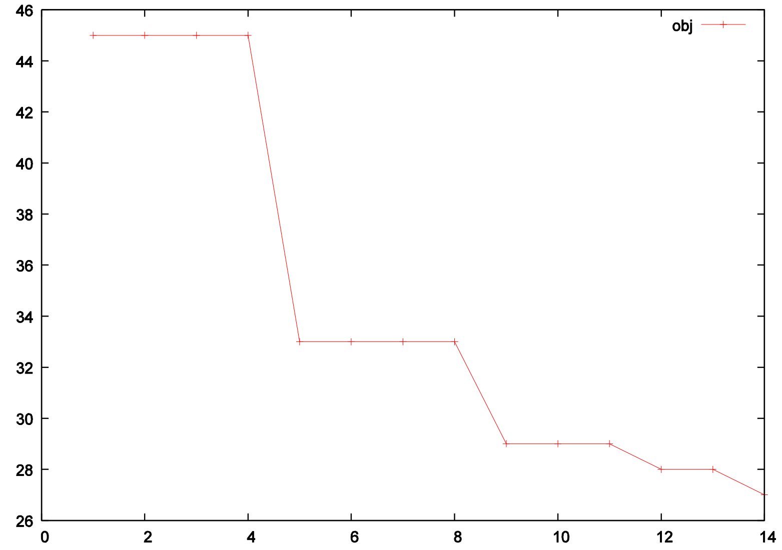
Consider a “dial” to test intermediate situations:

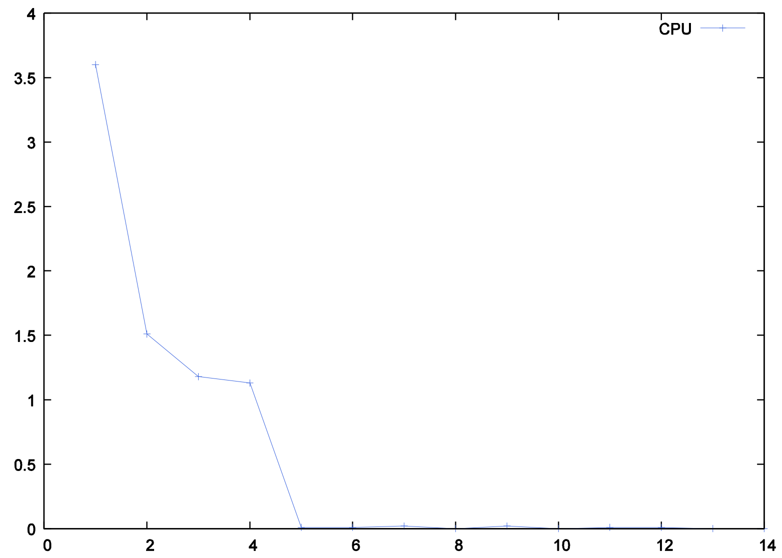


Typical behaviour

sts81

ℓ/k	obj	CPU
1/14	45	3.60
2/14	45	1.51
3/14	45	1.18
4/14	45	1.13
5/14	33	0.01
6/14	33	0.01
7/14	33	0.02
8/14	33	0.00
9/14	29	0.02
10/14	29	0.00
11/14	29	0.01
12/14	28	0.01
13/14	28	0.00
14/14	27	0.00





Hand-picked thresholds

Instance	G_ℓ	LP	CPU
ca36243	49	48	0.07
clique9	∞	36	0.06
cod105	-16	-18	4.91
cod105r	-13	-15	0.25
cod83	-26	-28	0.12
cod83r	-22	-25	4.44
cod93	-48	-51	3.07
cod93r	-46	-47	2.74
cov1075	19	18	3.03
cov1076	44	43	185.83
cov954	28	26	0.45
mered	∞	140	0.12
O4_35	∞	70	0.07
oa36243	∞	48	0.75
oa77247	∞	112	0.00
of5_14_7	∞	35	0.13
of7_18_9	∞	63	0.04
pa36243	-44	-48	1.26
sts135	60	45	0.05
sts27	12	9	0.01
sts45	24	15	0.39
sts63	27	21	0.00
sts81	33	27	0.00

Automatically-chosen thresholds

Instance	ℓ/k	best	G_1	G_ℓ	CPU	cpx_t	LP
ca36243	3/6	49*	49	48	0.02		48
clique9	5/15	∞*	∞	∞	**0.06**	0.17	36
cod105	3/11	-12*	*limit*	**-14.09**†	*limit*		-18
cod105r	3/10	-11*	-11	**-11**	**24.12**	28.36	-15
cod83	3/9	-20*	-21	**-24**	16.78	**9.54**	-28
cod83r	3/7	-19*	-21	**-22**	**4.44**	7.85	-25
cod93	3/10	-40		**-46.11**†	*limit*		-51
cod93r	3/8	-38	-39	**-44**	**271.74**	446.48	-47
cov1075	3/9	20*	20	**19**	**3.03**	79.79	18
cov1076	3/9	45	44	43	2.78		43
cov954	3/8	30*	28	26	0.11		26
mered	21/31	∞*	∞	∞	**0.15**	3.37	140
O4_35	3/9	∞*	∞	70	0.00		70
oa36243	3/6	∞*	∞	48	0.01		48
oa77247	3/7	∞*	∞	∞	**0.10**	265.92	112
of5_14_7	7/9	∞*	∞	35	0.00		35
of7_18_9	7/16	∞*	∞	∞	**0.09**	0.15	63
pa36243	3/6	-44*	-44	-48	0.01		-48
sts135	3/8	106	75	**60**	**0.11**	109.81	45
sts27	4/8	18*	14	**12**	0.01	1.05	9
sts45	2/5	30*	24	15	0.00		15
sts63	4/9	45*	36	**27**	**0.02**	1.99	21
sts81	5/14	61	45	**33**	**0.01**	3.92	27

Research questions

- Identify relevant **classes of problems** suitable to orbital shrinking (i.e., with a rich structure on the z-space left after shrinking)
- Exploit **cuts** taken from the shrunken formulation on the z var.s
- Conditions under which the shrunken relaxation is in fact **exact**
- Full **integration** of orbital shrinking within an exact solution scheme
- Use as a **heuristic**: find an optimal z, fix it, and optimize on x …