Conditions Beyond Treewidth for Tightness of Higher-Order LP Relaxations

Mark Rowland, Aldo Pacchiano and Adrian Weller

SUMMARY

- We consider linear programming (LP) relaxations for the problem of MAP inference for binary pairwise graphical models.
- We show that a low-treewidth condition obtained by Wainwright and Jordan (2004) guaranteeing tightness across all models with a given topology is not necessary.
- We also strengthen an earlier result regarding tightness of the triplet-consistent polytope $\mathbb{L}_3(G)$ for almost-balanced models.

MAP INFERENCE AND LP RELAXATIONS

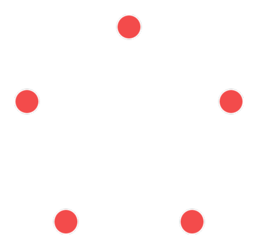
A binary pairwise graphical model is specified by a graph $G = (V, E)$, and a set of potentials $(\theta_i)_{i\in V}$, $(W_{ij})_{ij\in E}$. This gives a probabiltiy distribution for a set of binary random variables $(X_v)_{v\in V}$ - for any $x_V \in \{0,1\}^V$:

$$p(x_V) = \frac{1}{Z}\exp\left(\sum_{i\in V}\theta_i x_i + \sum_{ij\in E} W_{ij}x_i x_j\right)$$

MAP inference is the problem of finding a most likely configuration.

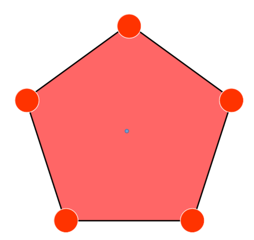
Combinatorial problem

$$\max_{x\in\{0,1\}^V}\left[\sum_{i\in V}\theta_i \mathbb{1}_{x_i=1} + \sum_{ij\in E} W_{ij}\mathbb{1}_{x_i=1,x_j=1}\right]$$



Equivalent linear program

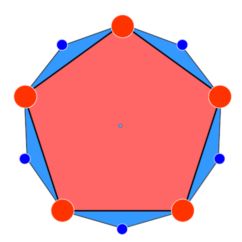
$$\max_{q\in\mathbb{M}(G)}\left[\sum_{i\in V}\theta_i q_i + \sum_{ij\in E} W_{ij}q_{ij}\right]$$



Marginal polytope $\mathbb{M}(G)$: enforce global consistency on marginals $(q_i)_{i\in V}$ and $(q_{ij})_{ij\in E}$

Relaxed linear program

$$\max_{q\in\mathbb{L}_r(G)}\left[\sum_{i\in V}\theta_i q_i + \sum_{ij\in E} W_{ij}q_{ij}\right]$$



Sherali-Adams polytope $\mathbb{L}_r(G)$: enforce consistency over each cluster of r variables on (pseudo)marginals $(q_i)_{i\in V}$ and $(q_{ij})_{ij\in E}$

This yields a polytope relaxation $\mathbb{M}(G) \subseteq \mathbb{L}_r(G)$, and we therefore have

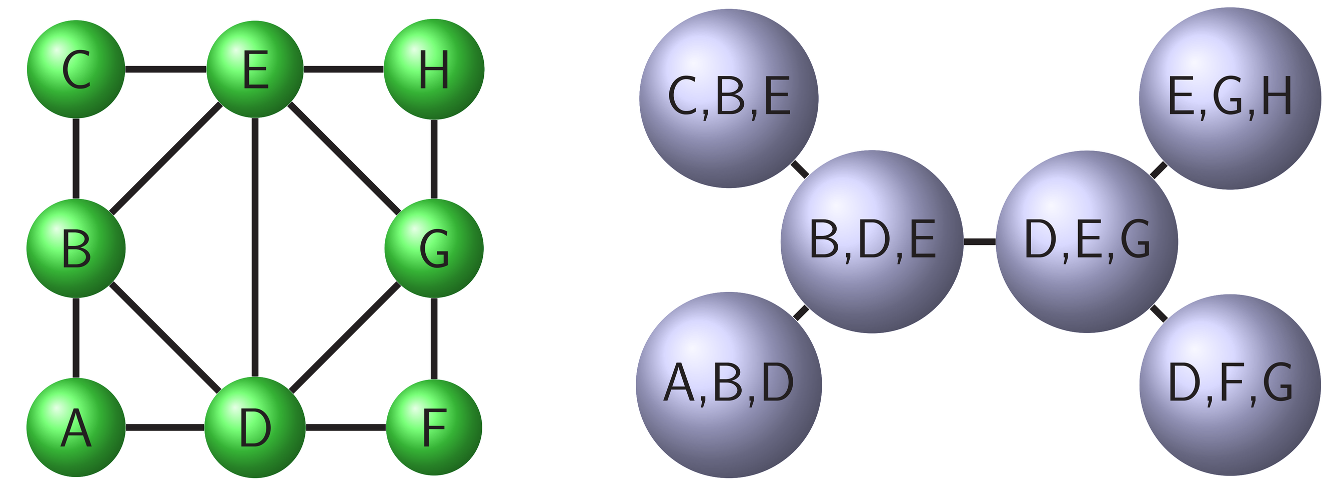
$$\max_{q\in\mathbb{M}(G)}\left[\sum_{i\in V}\theta_i q_i + \sum_{ij\in E} W_{ij}q_{ij}\right] \leq \max_{q\in\mathbb{L}_r(G)}\left[\sum_{i\in V}\theta_i q_i + \sum_{ij\in E} W_{ij}q_{ij}\right]$$

When we have equality above, we say that $\mathbb{L}_r(G)$ is **tight** for the graphical model concerned, and MAP inference can be performed efficiently by optimising over $\mathbb{L}_r(G)$.

Wainwright and Jordan (2004) showed that if a graph G has treewidth $\leq r$ then $\mathbb{L}_{r+1}(G)$ is tight **for all binary pairwise models on** G. Weller (2016) showed that this condition is necessary for $r = 1, 2$. **We show that it is not necessary for $r = 3$, by investigating graphical models on minimal forbidden minors.**

TREEWIDTH

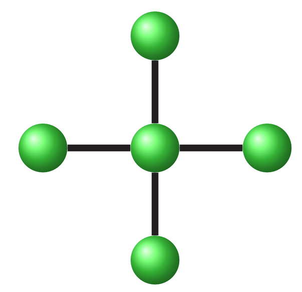
The *treewidth* of a graph G is the minimum width over all tree decompositions of G - intimately connected to the junction tree algorithm.



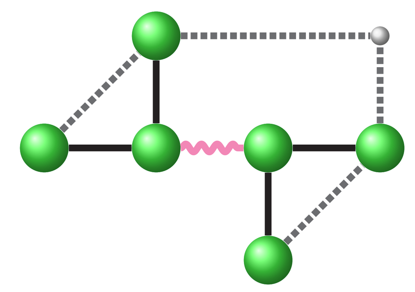
A graph of treewidth 2, and an optimal tree decomposition of the graph.

FORBIDDEN MINOR CONDITIONS

A graph G' is a *minor* of another graph G if G' is obtainable from G by deleting vertices and edges, and identifying vertices connected by an edge:



is a minor of



Each of the properties
- G has treewidth $\leq r$
- $\mathbb{L}_r(G)$ is tight for all possible binary pairwise graphical models on G

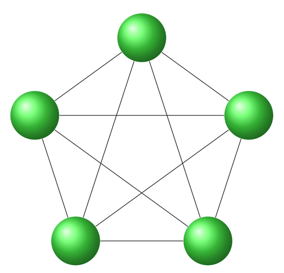
is **minor-closed**: true for $G \implies$ true for all minors of G.

A powerful result of Robertson and Seymour (2004) states that any graph property that is minor-closed is characterised by a *finite set of minimal forbidden minors* - "minimal obstructions" to the property.

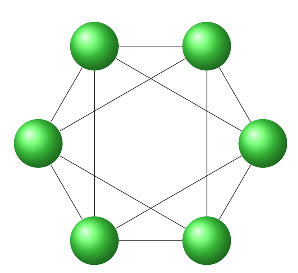
MINIMAL FORBIDDEN MINORS FOR LP TIGHTNESS AND LOW TREEWIDTH

Treewidth	Min. forbidden minors
1	K_3
2	K_4
3	K_5, O_6, M_8, Y_5 (see below)

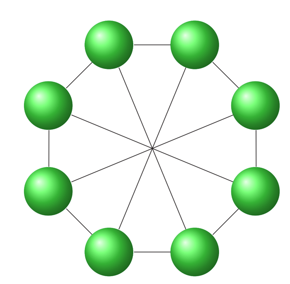
LP Relaxation	Min. forbidden minors
$\mathbb{L}_2(G)$	K_3
$\mathbb{L}_3(G)$	K_4
$\mathbb{L}_4(G)$	K_5 + ? (**NOT** O_6, M_8, Y_5)



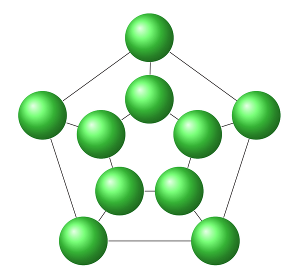
K_5



Octahedral graph, O_6



Wagner graph, M_8

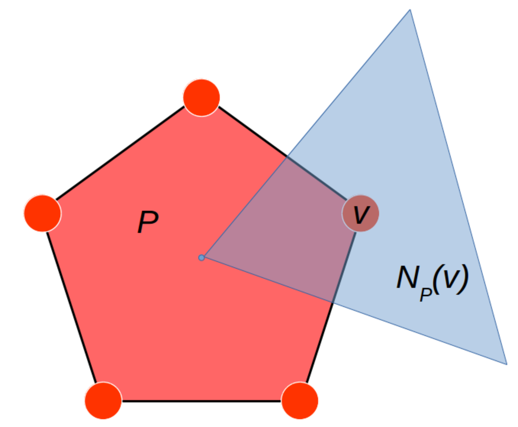


Pentagonal prism graph, Y_5

GEOMETRY OF SHERALI-ADAMS POLYTOPES

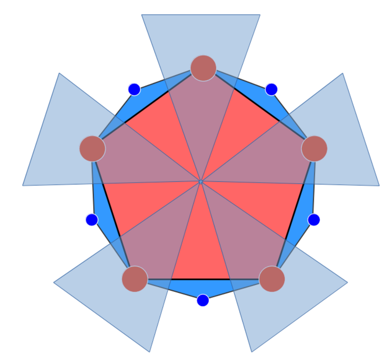
For a polytope $P \subset \mathbb{R}^d$ and an extremal point $v \in P$, the **normal cone** to P at v is

$$N_P(v) = \left\{c \in \mathbb{R}^d \,\middle|\, v \in \arg\max_{x\in P}\langle c, x\rangle\right\}$$



This gives a succinct characterisation of graphical models for which a Sherali-Adams relaxation is tight, namely:

$$\cup_{v\in \text{Vertices}(\mathbb{M}(G))} N_{\mathbb{L}_r(G)}(v)$$



We use this perspective to provide new proofs of several well-known results for tightness of $\mathbb{L}_2(G)$.

We also have the following decomposition of the entire space of potentials, which proves useful in our analysis of models over the minimal treewidth 3 forbidden minors:

$$\mathbb{R}^{V\cup E} = \cup_{v\in\text{Vertices}(\mathbb{M}(G))} N_{\mathbb{M}(G)}(v)$$

DECOMPOSITION AND LP CHECKS

For each of the treewidth 3 minimal forbidden minors $G = (V, E)$, we are interested in whether or not there exists a binary pairwise graphical model which is not tight for $\mathbb{L}_4(G)$. Writing $c = ((\theta_i)_{i\in V}, (W_{ij})_{ij\in E})$, this can be determined by solving the following optimisation problem:

$$\max_{c\in\mathbb{R}^{V\cup E}}\left[\max_{q\in\mathbb{L}_4(G)}\langle c, q\rangle - \max_{q\in\mathbb{M}(G)}\langle c, q\rangle\right]$$

Using the geometric considerations described above, we decompose this into a manageable number of LPs to check, yielding our main result:

The only treewidth 3 minimal forbidden minor that has models for which $\mathbb{L}_4(G)$ is not tight is K_5.

REFERENCES

N. Robertson and P. Seymour. Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325–357, 2004.

M. Wainwright and M. Jordan. Treewidth-based conditions for exactness of the Sherali-Adams and Lasserre relaxations. *Univ. California, Berkeley, Technical Report*, 671:4, 2004.

A. Weller. Characterizing tightness of LP relaxations by forbidding signed minors. In *Uncertainty in Artifical Intelligence (UAI)*, 2016.

ACKNOWLEDGEMENTS

MR acknowledges support by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L016516/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis. AW acknowledges support by the Alan Turing Institute under the EPSRC grant EP/N510129/1.