Characterizing Tightness of LP Relaxations by Forbidding Signed Minors

Adrian Weller, University of Cambridge

Computational and Biological Learning
University of Cambridge

SUMMARY

- **We examine MAP inference for binary pairwise graphical models. We show that many results on tightness of linear programming (LP) relaxations may be elegantly and compactly characterized (necessary and sufficient conditions for all valid potentials, efficiently testable) by forbidding certain minors in the graph of the model. This provides a connection to an important body of work in structural graph theory.**
- **We obtain stronger results by considering (i) the signed graph, where edges are signed as either attractive or repulsive; and (ii) the suspension graph, where singleton potentials are converted to edge potentials to a new added variable.**
- **Our strongest result is that the LP relaxation on the triplet-consistent polytope (LP+TRI) is tight for all potentials respecting edge signs iff the signed suspension graph does not contain an odd-K_5 (complete graph with all edges repulsive) as a signed minor.**

BACKGROUND AND MOTIVATION

- In a binary pairwise graphical model, each edge is signed as attractive (pulls variables toward the same value) or repulsive (pushes variables to different values). In an attractive model, all edges are attractive. A balanced model can be 'flipped to attractive'. An almost balanced model can be rendered balanced by deleting one variable.
- Many applications in computer vision use models which are 'close to attractive', such as image denoising or foreground-background segmentation.

Example: foreground-background segmentation

(Weizmann horse database)

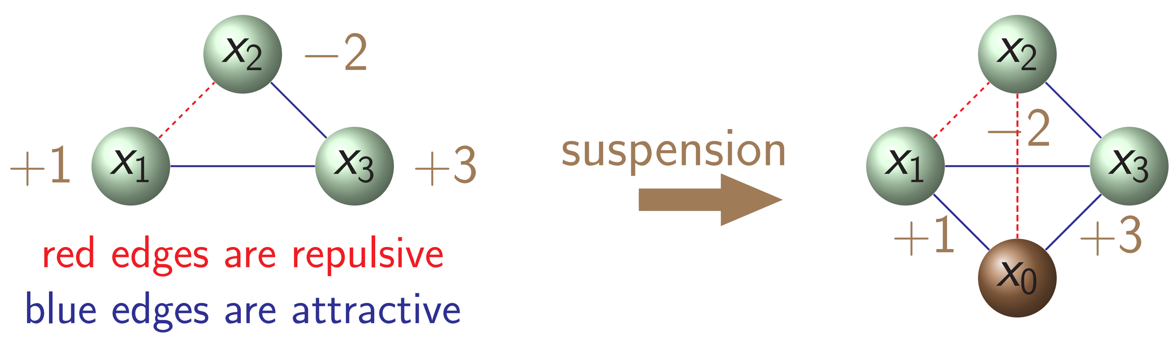
LP Relaxations for MAP Inference

- Binary variables $X_1, \dots, X_n \in \{0,1\}$. Write $x = (x_1, \dots, x_n) \in \{0,1\}^n$ for one configuration. Probability $p(x) \propto \exp[\text{score}(x)]$.
- $\text{Score}(x) = -\sum_{i \in V} \theta_i x_i - \sum_{(i,j) \in E} W_{ij} \mathbb{1}[x_i \neq x_j]$, $\quad W_{ij} > 0$ is attractive, $= w \cdot z$, where $w = (-\theta_1, \dots, -\theta_n, \dots, -W_{ij}, \dots)$.
- The convex hull of the 2^n binary solutions is the *marginal polytope* $\mathbb{M}$.
- MAP inference may be framed as the LP to find $\mu^* \in \arg\max_{\mu \in \mathbb{M}} w \cdot \mu$.
- Optimizing over $\mathbb{M}$ is intractable since it has an exponential number of facets. Hence, we relax $\mathbb{M}$ to a simpler, larger space that yields an upper bound on the true maximum. If an integral vertex is returned, then this must be an optimum of the original problem and the LP is *tight*.

- We consider linear programming (LP) relaxations over the local (LOC) and triplet (TRI) polytopes. These are the first two levels in the Sherali-Adams hierarchy.
- LP+LOC enforces pairwise consistency and is widely used. It is known that for balanced models, LP+LOC is *tight* (attains an optimum at an integral vertex).
- However, in practice, LP+LOC often yields a fractional solution, while LP relaxations using higher order cluster consistency are often tight (Sontag et al., 2008).
- Here we improve our theoretical understanding of this phenomenon by characterizing theoretically when LP relaxations are guaranteed to be tight.

SUSPENSION GRAPH ∇G, COMPLETE GRAPHS K_n AND ODD-K_n

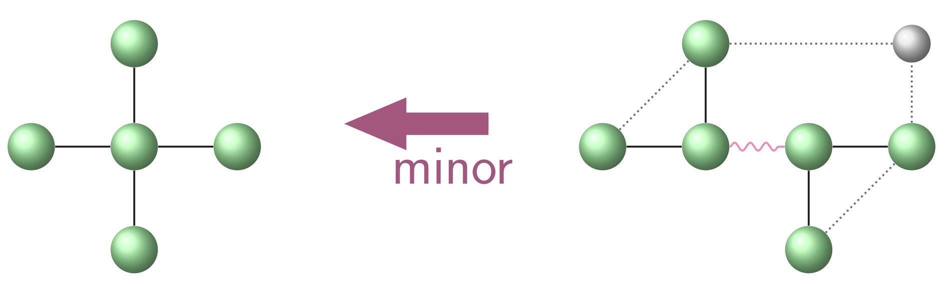
To obtain the *suspension graph* ∇G of a model, a vertex X_0 is added to the graph G. Edges to X_0 encode the singleton potentials of the original model.



Complete graphs K_n and odd-K_n

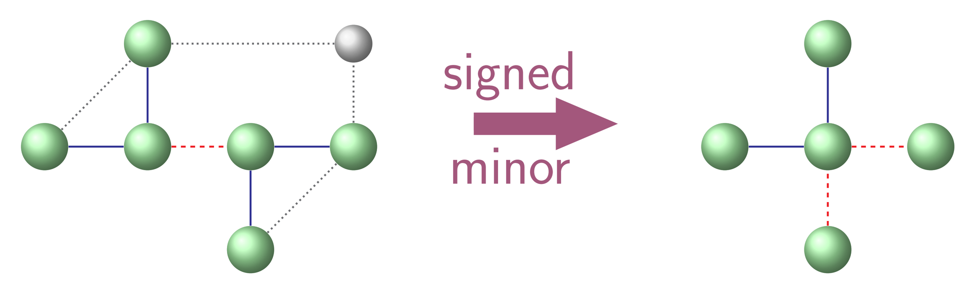
The unsigned graph K_n is the complete graph on n vertices, i.e. all vertices are adjacent.
The signed graph *odd-K_n* is the signed complete graph on n vertices where all edges are repulsive.

MINORS AND SIGNED MINORS

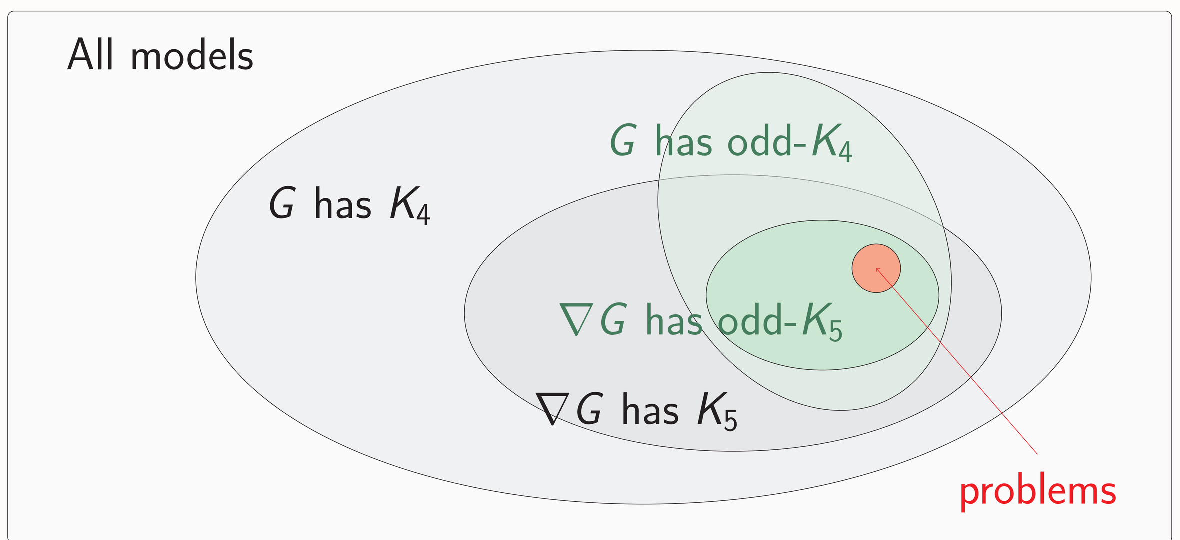


In an unsigned graph, a minor is obtained as a subgraph but in addition an edge may be *contracted*.

In a signed graph, a signed minor is obtained in the same way except: (i) only attractive edges may be contracted; (ii) any *resigning* is allowed (flipping a node changes incident edges attractive ↔ repulsive).



RESULTS



← Problems are models for which LP+TRI is not tight. Our strongest result shows that all problems lie within the set of models where the suspension graph ∇G contains an odd-K_5 as a signed minor.

The table to the right → summarizes results characterizing tightness of various LP relaxations by forbidden minors.

Forbidden minors

Sherali-Adams	cluster size	Graph G Unsigned	Graph G Signed	Suspension graph ∇G Unsigned	Suspension graph ∇G Signed
LOC	$\mathcal{L}_2$	K_3	odd-K_3		
TRI	$\mathcal{L}_3$	K_4	**odd-K_4**	K_5	**odd-K_5**
	$\mathcal{L}_4$	K_5+?	odd-K_5+?	K_6+?	odd-K_6+?

Our main new results build on work by Guenin (2001).
For $\mathcal{L}_4$, results are not yet known.

COMPARISON TO EARLIER WORK

- Weller et al. (2016) showed that LP+TRI is tight for any almost balanced model, and for certain 'pasted' combinations of such models.
- Weller (2016) showed that TRI is universally rooted, and used this to extend the above result to show that LP+TRI is tight for any model whose suspension graph is 2-almost balanced (may be rendered balanced by deleting 2 vertices; again the same pasting operation preserves tightness).
- The conditions here are significantly stronger: they enlarge the set of models for which LP+TRI is guaranteed to be tight, they are necessary as well as sufficient, and they are efficient to test.

REFERENCES

B. Guenin. A characterization of weakly bipartite graphs. *Journal of Combinatorial Theory, Series B*, 83(1):112–168, 2001.
D. Sontag, T. Meltzer, A. Globerson, T. Jaakkola, and Y. Weiss. Tightening LP relaxations for MAP using message passing. In *UAI*, pages 503–510, 2008.
A. Weller. Uprooting and rerooting graphical models. In *International Conference on Machine Learning (ICML)*, 2016.
A. Weller, M. Rowland, and D. Sontag. Tightness of LP relaxations for almost balanced models. In *Artificial Intelligence and Statistics (AISTATS)*, 2016.