SUMMARY

- We consider linear programming (LP) relaxations over the local (LOC) and triplet (TRI) polytopes.
- 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. We use a primal perturbation argument to demonstrate from first principles:
  
  
  \[ \text{LP+TRI is tight (attains an optimum at an integral vertex) for almost balanced models.} \]
  
  

OPTIMIZING OVER POLYTOPES

- We optimize a score determined by singleton \( \{ \theta_i \} \) and edge \( \{ W_{ij} \} \) potentials concatenated in a vector \( \theta \).
- We maximize \( \theta \cdot q = \sum_{i \in V} \theta_i q_i + \sum_{(i,j) \in E} W_{ij} q_{ij} \) over singleton \( \{ q_i \} \) and edge \( \{ q_{ij} \} \) marginals.
- Singleton potentials \( \{ \theta_i \} \) may take any value, often determined by data.
- Edge potentials: if \( W_{ij} > 0 \) then the edge is attractive; this is equivalent to submodular cost for the edge.
- LOC enforces pairwise consistency by requiring for every edge: \( \max(0, q_i + q_j - 1) \leq q_{ij} \leq \min(q_i, q_j) \).
- TRI adds four inequalities for every triplet of variables. Both LOC and TRI introduce fractional vertices.

Proof idea:

If a model is almost balanced (a property of the vector \( \theta \) of potentials) then after any non-integral optimum vertex \( \hat{\theta} \) is proposed, we demonstrate an explicit small perturbation \( p \) s.t. \( \hat{\theta} + p \) and \( \hat{\theta} - p \) remain in TRI, while \( \hat{\theta} = \frac{1}{2}(\hat{\theta} - p) + \frac{1}{2}(\hat{\theta} + p) \) and hence \( \hat{\theta} \) cannot be a vertex.

KEY STEPS IN THE PROOF

- We may assume an almost attractive model: all edges are attractive except for some which are incident to the special variable \( s \).
- If \( s \) is held to a fixed marginal \( x \in \{0, 1\} \), while all other marginals are optimized, some edge marginals ‘behave as attractive edges’ by taking their highest possible value in LOC, i.e. \( q_{ij} = \min(q_i, q_j) \).
- We prove a structural result: any edge which is not ‘behaving attractive’ must be in a binding triplet constraint together with the special variable \( s \).
- This allows us to construct an explicit perturbation \( p \) by \( p \) while remaining within TRI, unless all marginals take a simple form in \( \{0, x, 1-x, 1\} \).
- Using this we show: let \( F_{\text{TRI}}(x) \) be the constrained optimum in TRI holding the marginal of \( s \) to value \( x \), then \( F_{\text{TRI}}(x) \) is linear for an almost balanced model.

PREVIOUS WORK ON ALMOST BALANCED MODELS

- Remarkably, a different method involving perfect graphs (Weller, 2015) also works for almost balanced models.
- Weller (2015) has a composition result: if the method applies for two submodels, then it will apply if the submodels are pasted on any one variable.
- For LP+TRI, we show a stronger composition result: in addition, one may paste on an edge, if the edge includes special variable \( s \) in each submodel.
- Further, LP+TRI is known to be exact for any model with treewidth 2. Thus, LP+TRI dominates as a polynomial-time method for exact MAP inference.