Uprooting and Rerooting Graphical Models

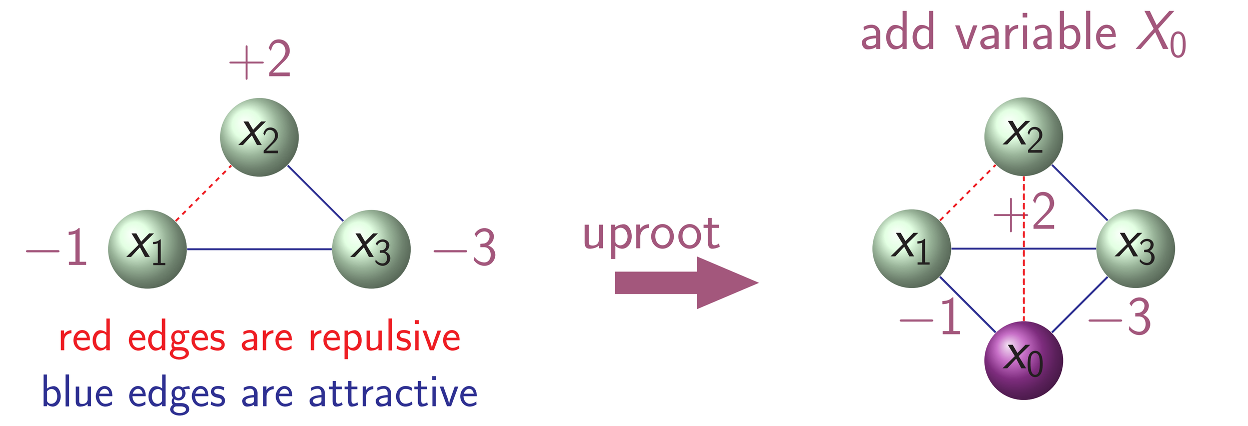
Adrian Weller, University of Cambridge

SUMMARY

- **We focus on binary pairwise graphical models (i.e. binary variables $X_1, \dots, X_n \in \{0,1\}^n$ with singleton and edge potentials).**
- **We show that each model is in an equivalence class, allowing a user easily to select whichever model is most beneficial for analysis or inference.**
- **This generalizes earlier theoretical results and is useful in practice, particularly for dense models with weak singleton potentials.**
- **The approach is related to clamping but demonstrates new insights and results, and obtains a clamping 'for free'.**

UPROOTING

Original model M — *Uprooted* model M^+, add variable X_0



Transform singleton potentials ⟶ edge potentials to X_0

Score for an edge iff its end variables are different

M^+ is fully symmetric with no singleton potentials

M is M^+ with X_0 *clamped* to 0, write $M = M_0$

If we don't clamp, each config of $M \to$ pair of configs of M^+ with the same score, e.g.

$$\begin{array}{ccc} x_1 & x_2 & x_3 \\ 1 & 0 & 1 \end{array} \to \left\{ \begin{array}{cccc} x_0 & x_1 & x_2 & x_3 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array} \right.$$

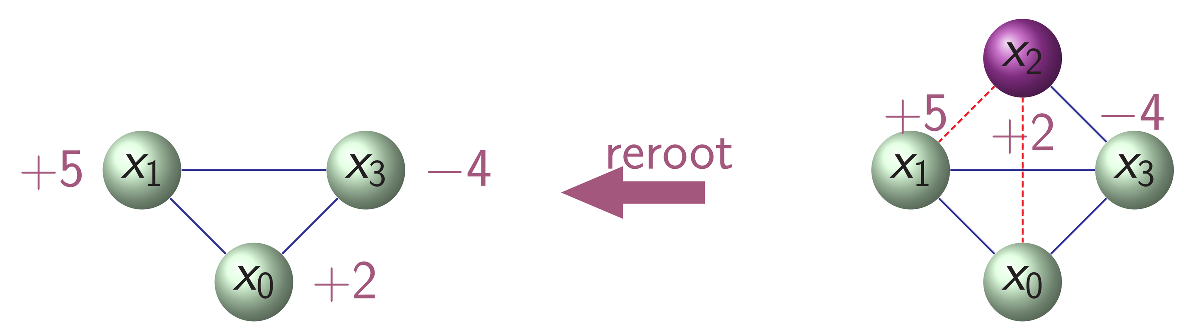
M^+ config				edges: ✓ if ends different					
x_0	x_1	x_2	x_3	e_{01}	e_{02}	e_{03}	e_{12}	e_{13}	e_{23}
0	0	0	0						
0	0	0	1			✓		✓	✓
0	0	1	0		✓		✓		✓
0	0	1	1		✓	✓	✓	✓	
0	1	0	0	✓			✓	✓	
0	1	0	1	✓		✓	✓		✓
0	1	1	0	✓	✓			✓	✓
0	1	1	1	✓	✓	✓			
1	0	0	0	✓	✓	✓			
1	0	0	1	✓	✓			✓	✓
1	0	1	0	✓		✓	✓		✓
1	0	1	1	✓			✓	✓	
1	1	0	0		✓	✓	✓	✓	
1	1	0	1		✓		✓		✓
1	1	1	0			✓		✓	✓
1	1	1	1						

⟵ Each color pair of configs of M^+ has the same score.

⟶ The original model $M = M_0$ is M^+ 'rooted at' $x_0 = 0$. This is one way to select one config from each color pair.

M^+ config				edges: ✓ if ends different					
x_0	x_1	x_2	x_3	e_{01}	e_{02}	e_{03}	e_{12}	e_{13}	e_{23}
0	0	0	0						
0	0	0	1			✓		✓	✓
0	0	1	0		✓		✓		✓
0	0	1	1		✓	✓	✓	✓	
0	1	0	0	✓			✓	✓	
0	1	0	1	✓		✓	✓		✓
0	1	1	0	✓	✓			✓	✓
0	1	1	1	✓	✓	✓			
1	0	0	0	✓	✓	✓			
1	0	0	1	✓	✓			✓	✓
1	0	1	0	✓		✓	✓		✓
1	0	1	1	✓			✓	✓	
1	1	0	0		✓	✓	✓	✓	
1	1	0	1		✓		✓		✓
1	1	1	0			✓		✓	✓
1	1	1	1						

REROOTING



Rerooted model M_2
attractive

How should we pick a good root variable?

- Same as choosing a variable to clamp in M^+.
- Rerooting substitutes an implicit initial clamp choice for a well chosen one 'for free'.
- Several existing good methods, including *maxW*.
  Idea: break heavy cycles and also 'spread influence'.
- We introduce *maxtW*: strength of an edge weight saturates, works well in our context

Here we *reroot* M^+ at $x_2 = 0$ to obtain the equivalent model M_2. Observe that again, this selects one config from each color pair.

M^+ config				edges: ✓ if ends diff					
x_0	x_1	x_2	x_3	e_{01}	e_{02}	e_{03}	e_{12}	e_{13}	e_{23}
0	0	**0**	0						
0	0	**0**	1			✓		✓	✓
0	0	1	0		✓		✓		✓
0	0	1	1		✓	✓	✓	✓	
0	1	**0**	0	✓			✓	✓	
0	1	**0**	1	✓		✓	✓		✓
0	1	1	0	✓	✓			✓	✓
0	1	1	1	✓	✓	✓			
1	0	**0**	0	✓	✓	✓			
1	0	**0**	1	✓	✓			✓	✓
1	0	1	0	✓		✓	✓		✓
1	0	1	1	✓			✓	✓	
1	1	**0**	0		✓	✓	✓	✓	
1	1	**0**	1		✓		✓		✓
1	1	1	0			✓		✓	✓
1	1	1	1						

Rerooting Observations

- Rerooted models $\{M_i\}$ form an equivalence class, each has the same 'parent' uprooted model M^+.
- Score-preserving 1-1 correspondence $\forall i$, configs of M_i $\leftrightarrow$ configs of M_0.
- Inference (exact or approx) on *any* M_i allows simple recovery of info for M_0.
  - ⋆ Each M_i has the same partition function.
  - ⋆ MAP config of $M_i \to$ recover MAP config of M_0.
  - ⋆ Marginals of $M_i \to$ recover marginals of M_0.
- Inference may be much faster / more accurate on some M_i.
- Singleton and edge potentials are essentially the same, only appear different due to choice of rooting.

EXPERIMENTS ON RANDOM MODELS (USING BETHE APPROX) AND DISCUSSION

Empirical results for complete graphs (see right):

- Rerooting is very effective, better for low singleton potentials θ_i.
- *maxW* and *maxtW* both perform well.

For grids (see lower panels to right):

- Rerooting is less effective, still helps for low θ_i.
- *maxtW* performs much better than *maxW*.

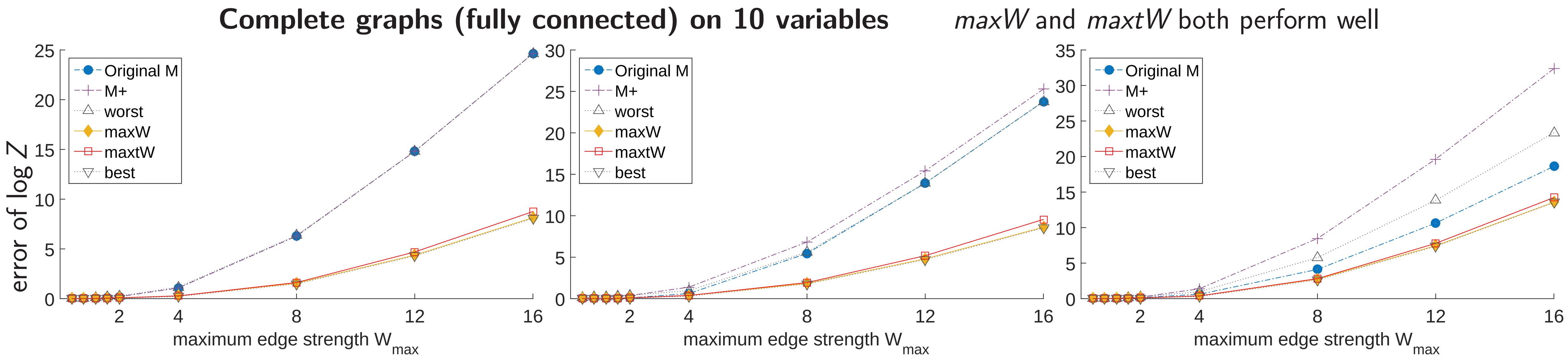
Implications of Rerooting

We can generalize or improve many results:

- e.g. max flow / min cut can now be used for models where M^+ is almost attractive (almost balanced) since then $\exists i$ s.t. M_i is attractive (balanced).
- e.g. bounds (on marginals, partition function) can be improved by considering different rerootings.

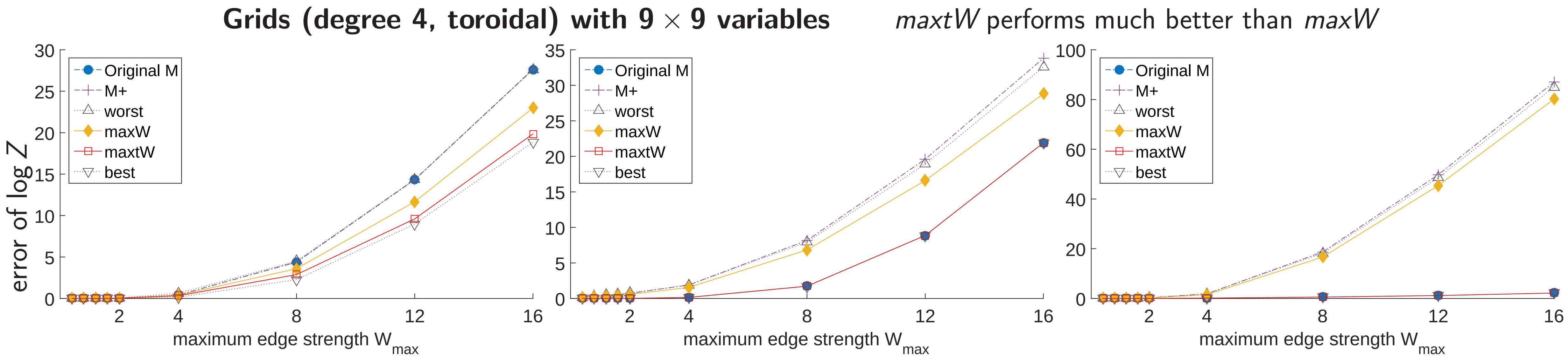
Reveals an intriguing perspective on TRI (triplet-consistent polytope): TRI is universally rooted.

Complete graphs (fully connected) on 10 variables — *maxW* and *maxtW* both perform well



low singleton $\theta_i \sim [-0.1, 0.1]$ — medium singleton $\theta_i \sim [-2, 2]$ — singleton + edge potentials scale together

Grids (degree 4, toroidal) with 9×9 variables — *maxtW* performs much better than *maxW*



low singleton $\theta_i \sim [-0.1, 0.1]$ — medium singleton $\theta_i \sim [-2, 2]$ — singleton + edge potentials scale together