MAP Estimation in Binary MRFs using Bipartite Multi-Cuts

Sashank J. Reddi Sunita Sarawagi
Sundar Vishwanathan

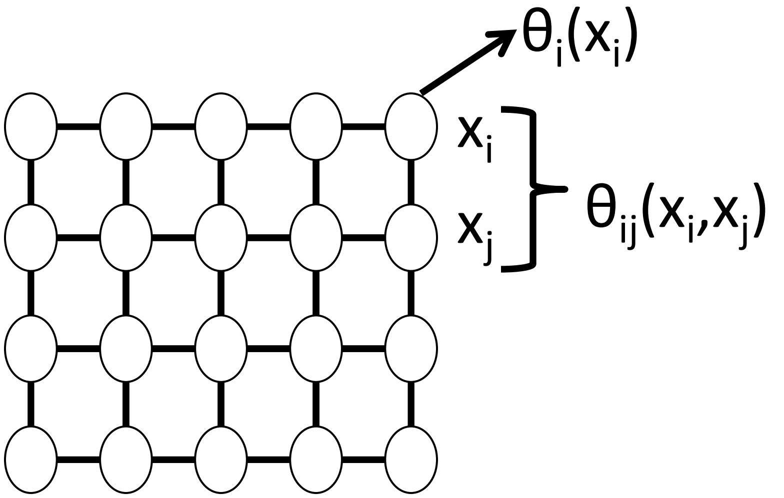
Indian Institute of Technology, Bombay

MAP Estimation

- Energy Function

$$E(x \mid \theta) = \sum_{i \in \mathcal{V}} \theta_i(x_i) + \sum_{(i,j) \in \mathcal{E}} \theta_{ij}(x_i, x_j)$$

$$G = (\mathcal{V}, \mathcal{E}),\ x \in \{0,1\}^n$$



- MAP Estimation: Find the labeling which minimizes the energy function
  - NP Hard in general

Popular Approximation

- Based on this LP relaxation (Pairwise LP Relaxation)

$$\min_{\mu} \sum_{i,x_i} \theta_i(x_i)\mu_i(x_i) + \sum_{(i,j),x_i,x_j} \theta_{ij}(x_i,x_j)\mu_{ij}(x_i,x_j)$$

$$\sum_{x_j} \mu_{ij}(x_i,x_j) = \mu_i(x_i) \quad \forall (i,j) \in \mathcal{E}, \forall x_i \in \{0,1\}$$

$$\sum_{x_i} \mu_i(x_i) = 1 \quad \forall i \in \mathcal{V}$$

$$\mu_{ij}(x_i,x_j) \geq 0 \quad \forall (i,j) \in \mathcal{E}, \forall x_i, x_j \in \{0,1\}$$

- Two approaches for efficiently solving the LP
  - Message passing algorithms (e.g. TRW-S)
  - Graph Cut based algorithms (e.g. QPBO)

Tightening Pairwise LP Relaxation

- MPLP (David Sontag et.al 2008), Cycle Repairing Algorithm (Nikos Komodakis et.al 2008)

Tightening Pairwise LP Relaxation

- MPLP (David Sontag et.al 2008), Cycle Repairing Algorithm (Nikos Komodakis et.al 2008)
- MPLP uses higher order relaxation

Tightening Pairwise LP Relaxation

- MPLP (David Sontag et.al 2008), Cycle Repairing Algorithm (Nikos Komodakis et.al 2008)
- MPLP uses higher order relaxation

$$\min_{\mu} \sum_{i,x_i} \theta_i(x_i)\mu_i(x_i) + \sum_{(i,j),x_i,x_j} \theta_{ij}(x_i,x_j)\mu_{ij}(x_i,x_j)$$

$$\sum_{x_j} \mu_{ij}(x_i,x_j) = \mu_i(x_i) \quad \forall (i,j) \in \mathcal{E}, \forall x_i \in \{0,1\}$$

$$\sum_{x_i} \mu_i(x_i) = 1 \quad \forall i \in \mathcal{V}$$

$$\mu_{ij}(x_i,x_j) \geq 0 \quad \forall (i,j) \in \mathcal{E}, \forall x_i, x_j \in \{0,1\}$$

Tightening Pairwise LP Relaxation

- MPLP (David Sontag et.al 2008), Cycle Repairing Algorithm (Nikos Komodakis et.al 2008)
- MPLP uses higher order relaxation

$$\min_{\mu} \sum_{i,x_i} \theta_i(x_i)\mu_i(x_i) + \sum_{(i,j),x_i,x_j} \theta_{ij}(x_i,x_j)\mu_{ij}(x_i,x_j)$$

$$\sum_{x_j} \mu_{ij}(x_i,x_j) = \mu_i(x_i) \quad \forall (i,j) \in \mathcal{E}, \forall x_i \in \{0,1\}$$

$$\sum_{x_i} \mu_i(x_i) = 1 \quad \forall i \in \mathcal{V}$$

$$\mu_{ij}(x_i,x_j) \geq 0 \quad \forall (i,j) \in \mathcal{E}, \forall x_i, x_j \in \{0,1\}$$

$$\boxed{\sum_{x_k} \mu_{ijk}(x_i,x_j,x_k) = \mu_{ij}(x_i,x_j)} \quad \forall i,j,k \in \mathcal{V}$$

MAP Estimation via Graph Cuts

- Construct a specialized graph for the particular energy function

MAP Estimation via Graph Cuts

- Construct a specialized graph for the particular energy function
- Minimum cut on this graph minimizes the energy function

MAP Estimation via Graph Cuts

- Construct a specialized graph for the particular energy function
- Minimum cut on this graph minimizes the energy function
- Exact MAP in polynomial time when the energy function is sub-modular
  - $\theta_{ij}(0,1) + \theta_{ij}(1,0) - \theta_{ij}(0,0) - \theta_{ij}(1,1) \geq 0$

Assumption on Parameters

- Symmetric, that is $\theta_{ij}(0,0) = \theta_{ij}(1,1)$ and $\theta_{ij}(0,1) = \theta_{ij}(1,0)$

Assumption on Parameters

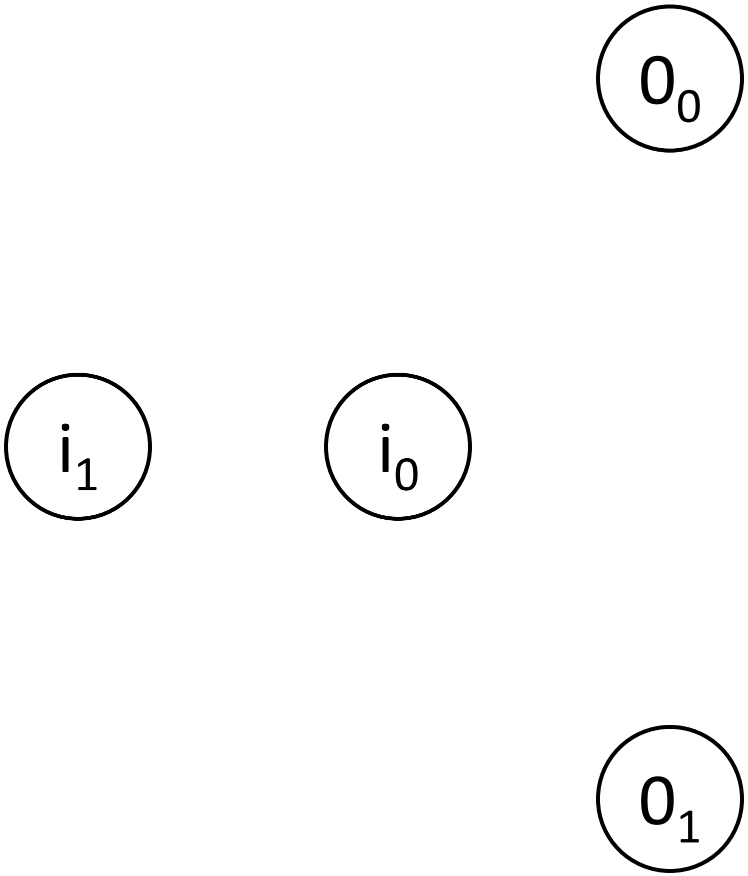
- Symmetric, that is $\theta_{ij}(0,0) = \theta_{ij}(1,1)$ and $\theta_{ij}(0,1) = \theta_{ij}(1,0)$
- Zero-normalized, that is $\min_{x_i} \theta_i(x_i) = 0$ and $\min_{x_i, x_j} \theta_{ij}(x_i, x_j) = 0$

Assumption on Parameters

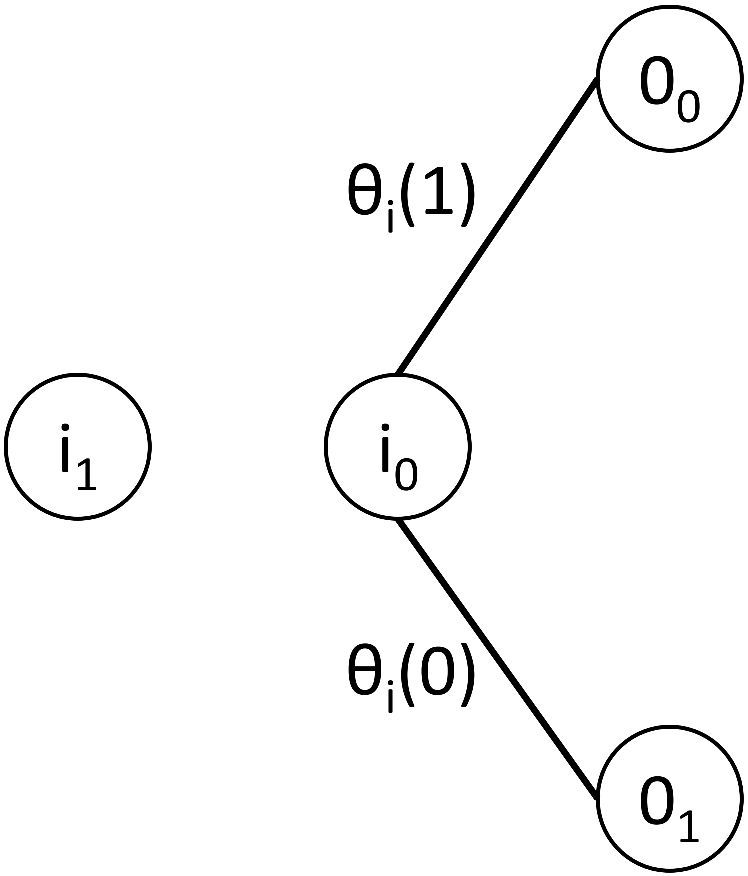
- Symmetric, that is $\theta_{ij}(0,0) = \theta_{ij}(1,1)$ and $\theta_{ij}(0,1) = \theta_{ij}(1,0)$
- Zero-normalized, that is $\min_{x_i} \theta_i(x_i) = 0$ and $\min_{x_i, x_j} \theta_{ij}(x_i, x_j) = 0$
- Any Energy function can be transformed in to an equivalent energy function of this form

Graph Construction

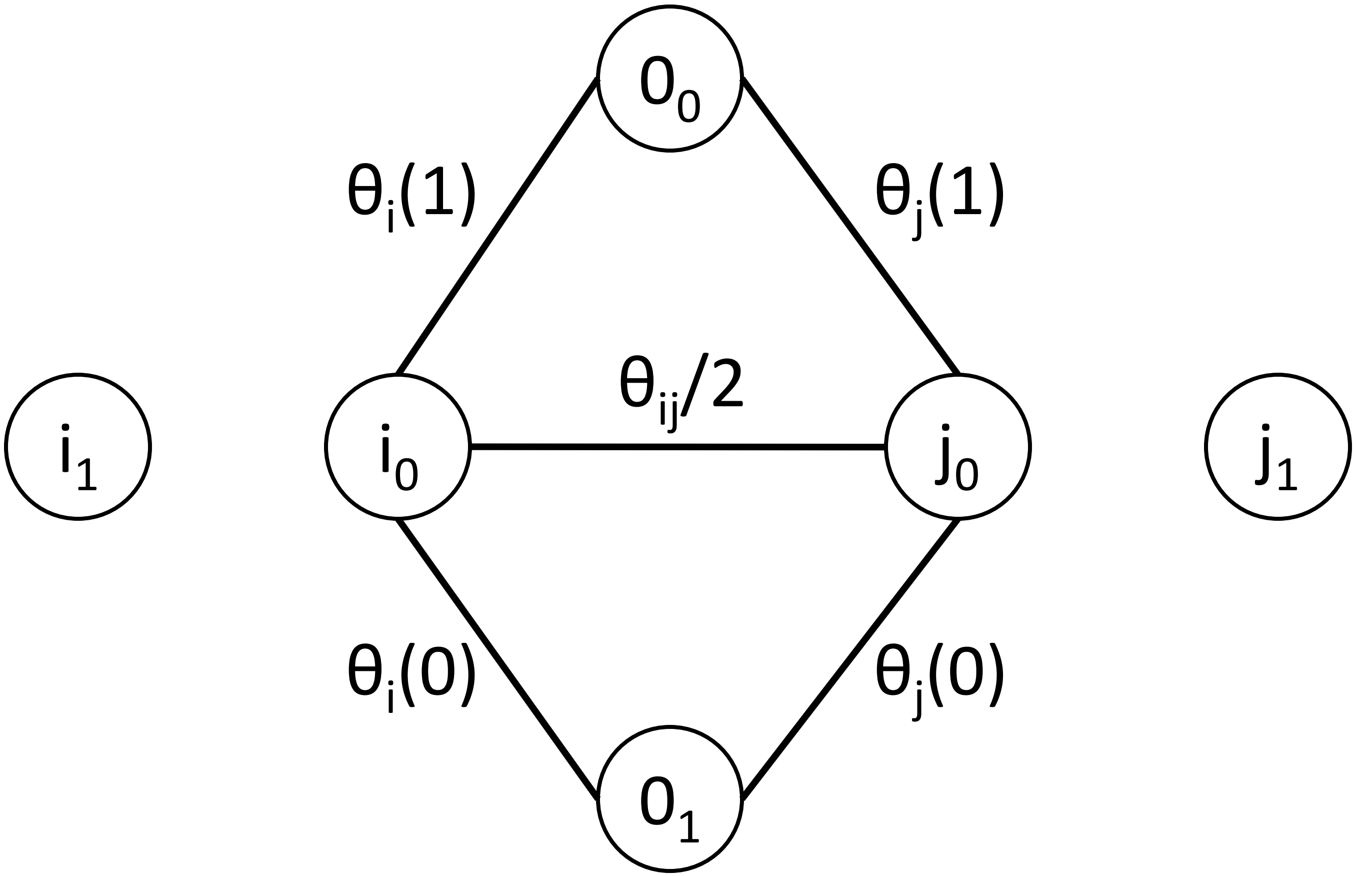
Graph Construction



Graph Construction

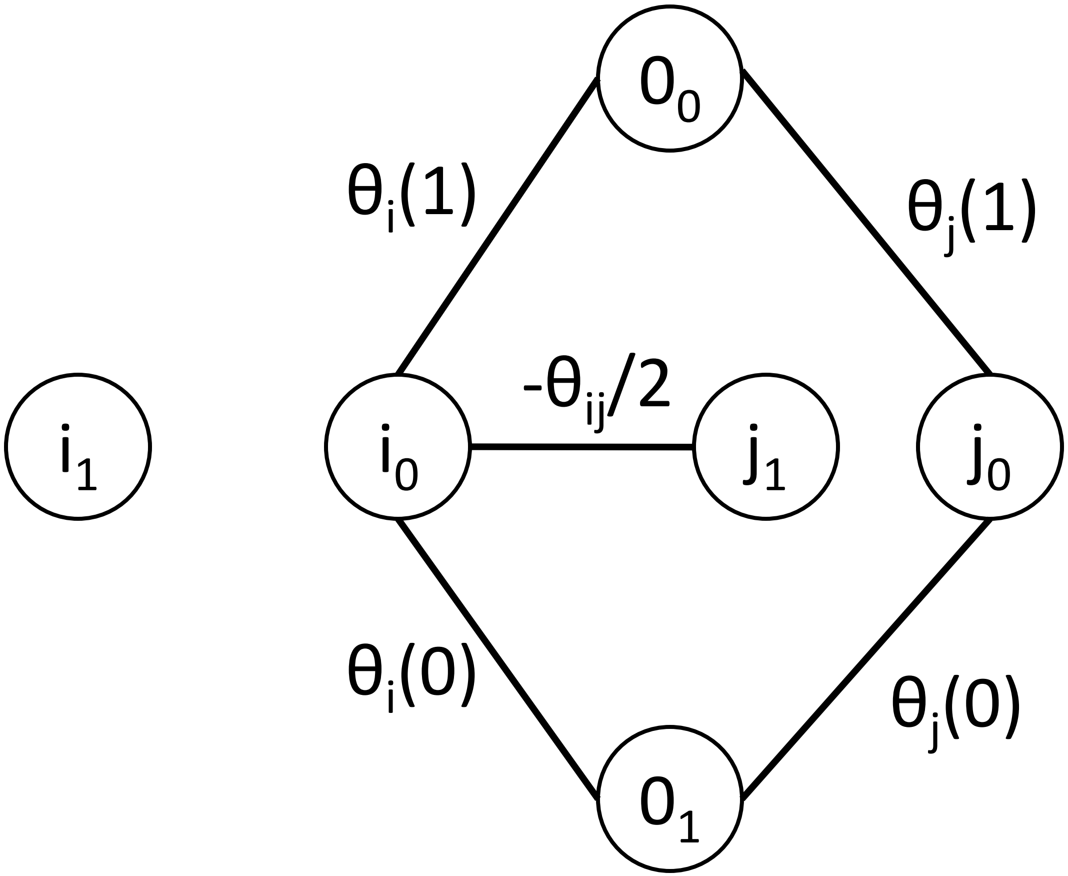


Graph Construction



$$\theta_{ij} = \theta_{ij}(0,1) + \theta_{ij}(1,0) - \theta_{ij}(0,0) - \theta_{ij}(1,1) > 0$$

Graph Construction



$$\theta_{ij} = \theta_{ij}(0,1) + \theta_{ij}(1,0) - \theta_{ij}(0,0) - \theta_{ij}(1,1) < 0$$

Bipartite Multi-Cut problem

- Given an undirected graph $J = (N, A)$ with non-negative edge weights and *k* ST pairs

Bipartite Multi-Cut problem

- Given an undirected graph $J = (N, A)$ with non-negative edge weights and *k* ST pairs
  - Objective: Find the minimum cut with divides the graph into 2 regions and separates the ST pairs

Bipartite Multi-Cut problem

- Given an undirected graph $J = (N, A)$ with non-negative edge weights and *k* ST pairs
  - Objective: Find the minimum cut with divides the graph into 2 regions and separates the ST pairs
- LP and SDP Relaxations (Sreyash Kenkre *et.al 2006)*

Bipartite Multi-Cut problem

- Given an undirected graph $J = (N, A)$ with non-negative edge weights and *k* ST pairs
  - Objective: Find the minimum cut with divides the graph into 2 regions and separates the ST pairs

- LP and SDP Relaxations (Sreyash Kenkre *et.al 2006)*
  - LP Relaxation gives $O(\log k)$ approximation
  - SDP Relaxation gives $O(\sqrt{\log(k)}\log(\log(k)))$ approximation

Bipartite Multi-Cut problem

- Given an undirected graph $J = (N, A)$ with non-negative edge weights and *k* ST pairs
  - Objective: Find the minimum cut with divides the graph into 2 regions and separates the ST pairs

- LP and SDP Relaxations (Sreyash Kenkre *et.al 2006)*
  - LP Relaxation gives $O(\log k)$ approximation
  - SDP Relaxation gives $O(\sqrt{\log(k)}\log(\log(k)))$ approximation

- Bipartite Multi-Cut vs Multi-Cut
  - Additional constraint on the number of regions the graph is cut

BMC LP

- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$
\begin{aligned}
\min \sum_{e \in \mathcal{E}_H} w_e d_e \\
D_{i_0}(i_1) &= 1 \\
d_e + D_u(i_s) - D_u(j_t) &\geq 0 \\
d_e + D_u(j_t) - D_u(i_s) &\geq 0 \\
D_{i_0}(j_0) &= D_{i_1}(j_1) \\
D_{i_0}(j_1) &= D_{i_1}(j_0) \\
D_u(i_s) &\geq 0 \\
d_e &\geq 0
\end{aligned}
$$

BMC LP

- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$\min \sum_{e \in \mathcal{E}_H} w_e d_e$$

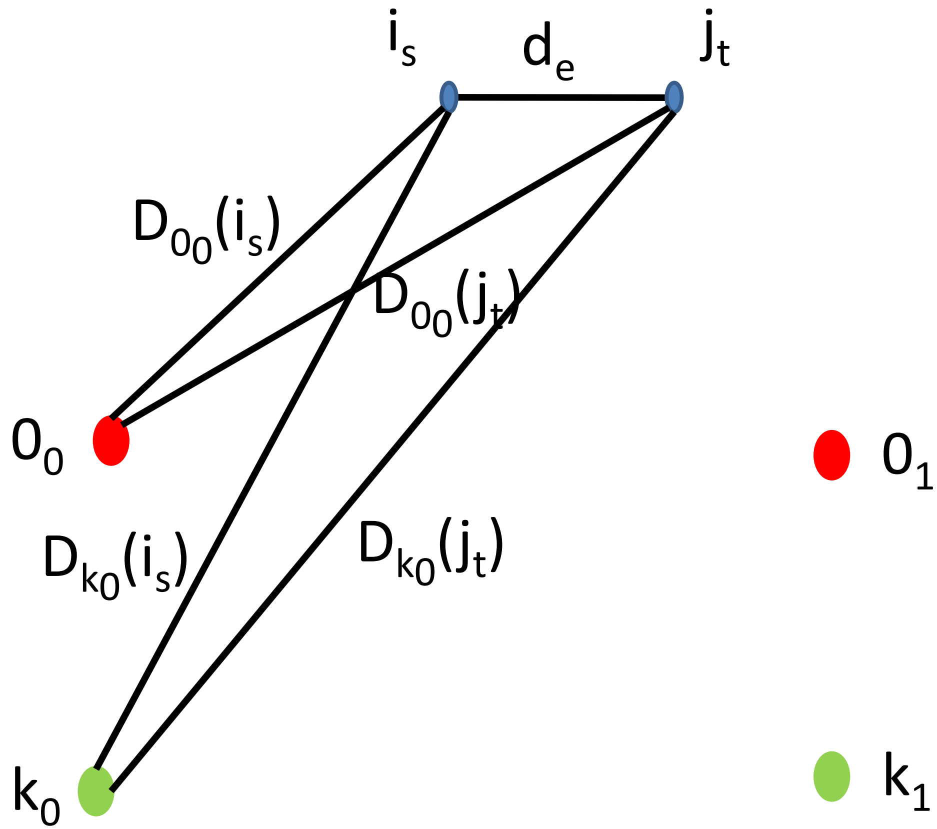
$$\begin{aligned}
D_{i_0}(i_1) &= 1 \\
d_e + D_u(i_s) - D_u(j_t) &\geq 0 \\
d_e + D_u(j_t) - D_u(i_s) &\geq 0 \\
D_{i_0}(j_0) &= D_{i_1}(j_1) \\
D_{i_0}(j_1) &= D_{i_1}(j_0) \\
D_u(i_s) &\geq 0 \\
d_e &\geq 0
\end{aligned}$$

BMC LP

- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$\min \sum_{e \in \mathcal{E}_H} w_e d_e$$

$$\begin{array}{rcl}
D_{i_0}(i_1) & = & 1 \\
d_e + D_u(i_s) - D_u(j_t) & \geq & 0 \\
d_e + D_u(j_t) - D_u(i_s) & \geq & 0 \\
D_{i_0}(j_0) & = & D_{i_1}(j_1) \\
D_{i_0}(j_1) & = & D_{i_1}(j_0) \\
D_u(i_s) & \geq & 0 \\
d_e & \geq & 0
\end{array}$$



BMC LP

- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$
\begin{aligned}
\min \sum_{e \in \mathcal{E}_H} w_e d_e & & \\
D_{i_0}(i_1) &= 1 \\
d_e + D_u(i_s) - D_u(j_t) &\geq 0 \\
d_e + D_u(j_t) - D_u(i_s) &\geq 0 \\
D_{i_0}(j_0) &= D_{i_1}(j_1) \\
D_{i_0}(j_1) &= D_{i_1}(j_0) \\
D_u(i_s) &\geq 0 \\
d_e &\geq 0
\end{aligned}
$$

BMC LP

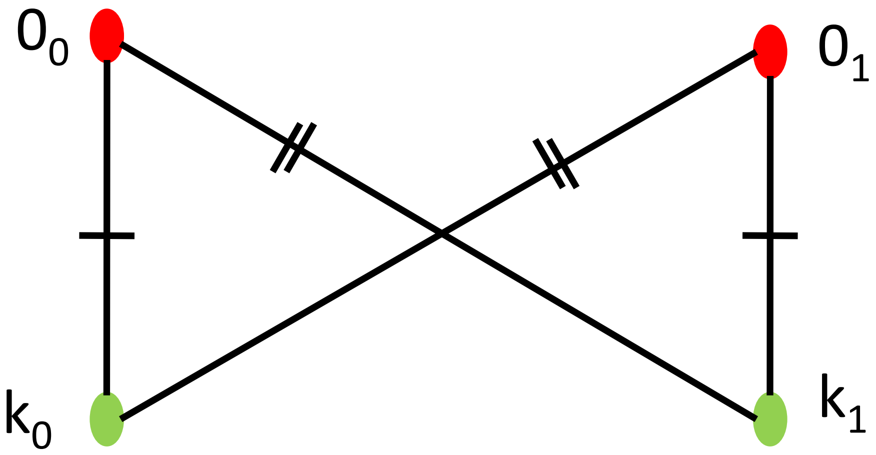
- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$\min \sum_{e \in \mathcal{E}_H} w_e d_e$$

$$\begin{aligned} D_{i_0}(i_1) &= 1 \\ d_e + D_u(i_s) - D_u(j_t) &\geq 0 \\ d_e + D_u(j_t) - D_u(i_s) &\geq 0 \end{aligned}$$

$$\begin{aligned} D_{i_0}(j_0) &= D_{i_1}(j_1) \\ D_{i_0}(j_1) &= D_{i_1}(j_0) \end{aligned}$$

$$\begin{aligned} D_u(i_s) &\geq 0 \\ d_e &\geq 0 \end{aligned}$$

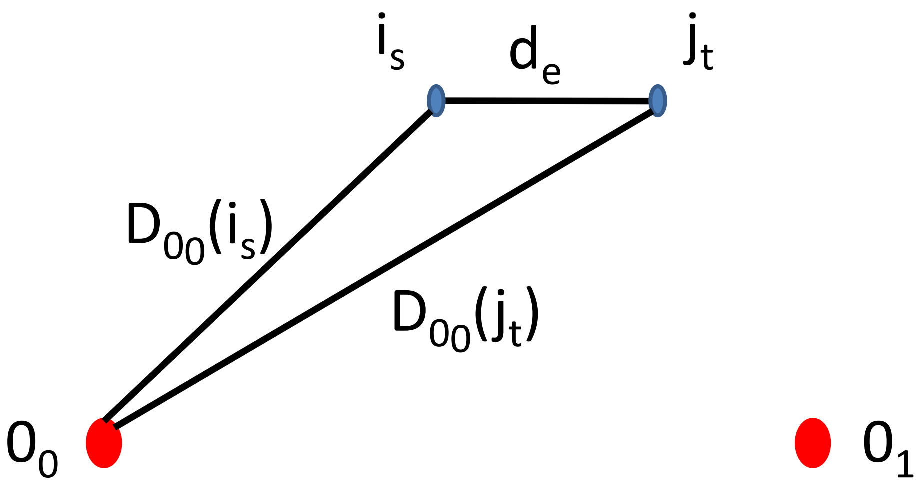


BMC LP

- (i_0, i_1) are the ST pairs in the Bipartite Multi-Cut problem
- *Terminals* = $\{0_0, 0_1, ..., k_0, k_1\}$

$$\min \sum_{e \in \mathcal{E}_H} w_e d_e$$

$$\begin{array}{rcl}
D_{i_0}(i_1) & = & 1 \\
d_e + D_{0_0}(i_s) - D_{0_0}(j_t) & \geq & 0 \\
d_e + D_{0_0}(j_t) - D_{0_0}(i_s) & \geq & 0 \\
D_u(i_s) & \geq & 0 \\
d_e & \geq & 0
\end{array}$$



BMC LP

- Infeasible to solve using LP solvers
  - Large number of constraints

BMC LP

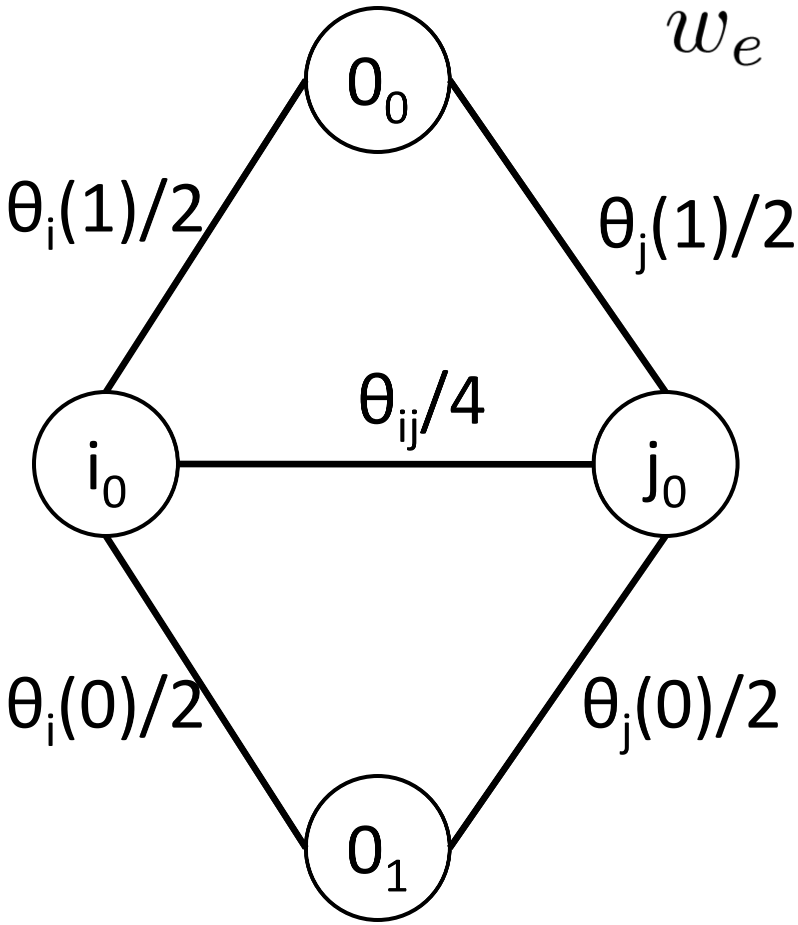
- Infeasible to solve using LP solvers
  - Large number of constraints
- Combinatorial Algorithm
  - Solving LP for general graphs may not be easy

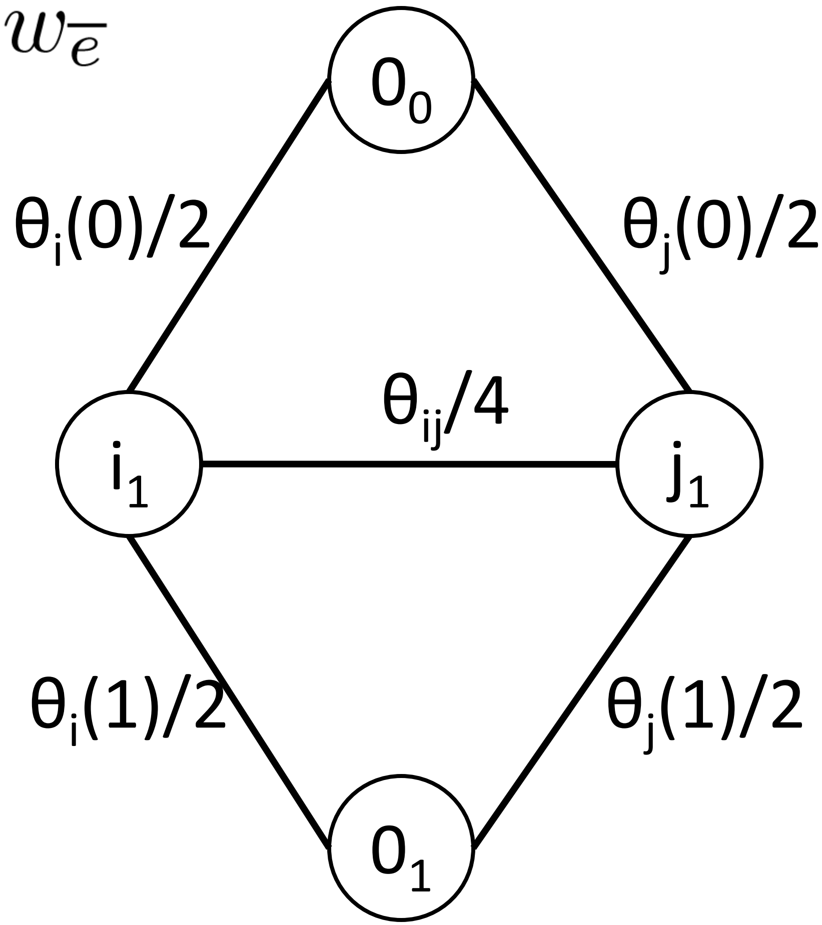
BMC LP

- Infeasible to solve using LP solvers
  - Large number of constraints
- Combinatorial Algorithm
  - Solving LP for general graphs may not be easy
  - Our graph for MAP estimation is special ➔ we exploit it to design an efficient algorithm

Symmetric Graph Construction

$$w_e = w_{\overline{e}}$$





$$\theta_{ij} = \theta_{ij}(0,1) + \theta_{ij}(1,0) - \theta_{ij}(0,0) - \theta_{ij}(1,1)$$

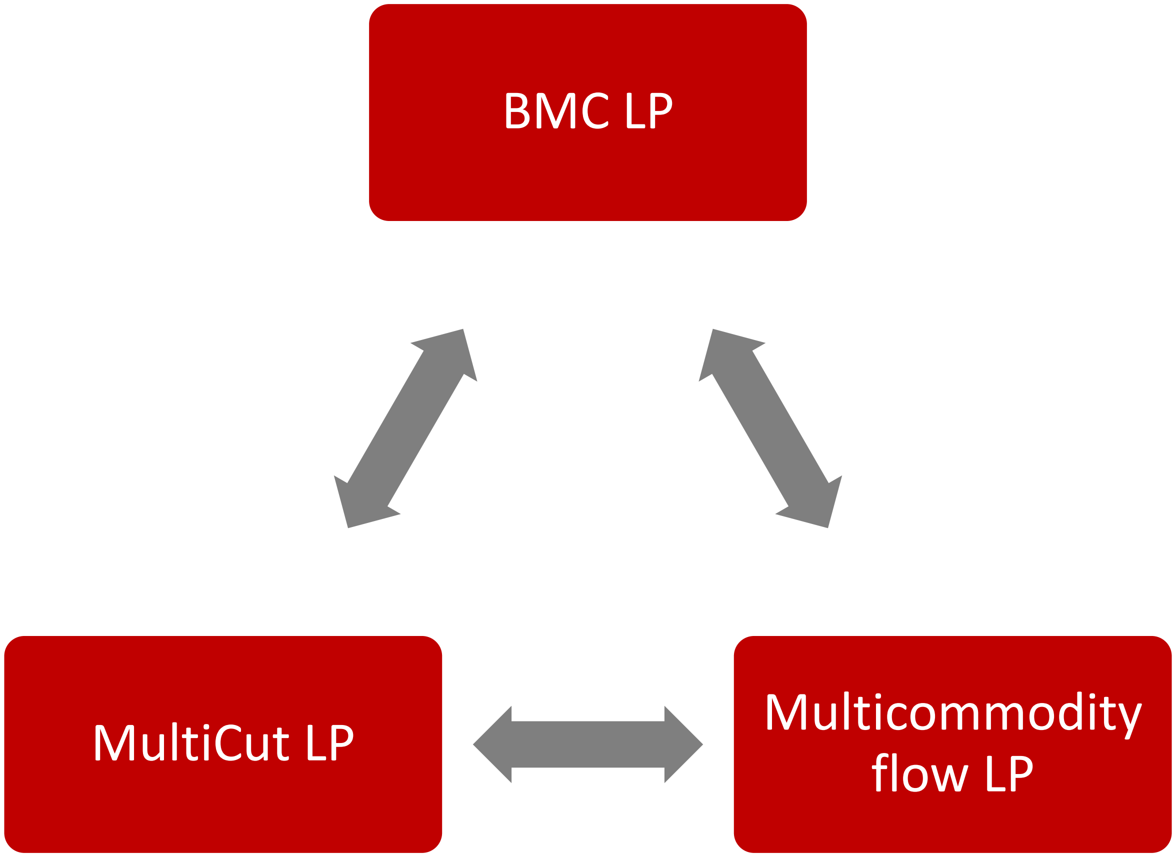
Symmetric Graph Construction

- More ST pairs are added

Symmetric Graph Construction

- More ST pairs are added
- Trade off is combinatorial algorithm

Approach of the Algorithm



Multi-Cut LP

- $\mathcal{P}$ denotes all paths between pair of vertices in ST

Multi-Cut LP

- $\mathcal{P}$ denotes all paths between pair of vertices in ST
- $\mathcal{P}_e$ denotes the set of paths in $\mathcal{P}$ which contain edge e

Multi-Cut LP

- $\mathcal{P}$ denotes all paths between pair of vertices in ST
- $\mathcal{P}_e$ denotes the set of paths in $\mathcal{P}$ which contain edge e

$$\min \sum_{e \in \mathcal{E}_{\mathcal{H}}} w_e d_e$$

$$\sum_{e \in P} d_e \geq 1 \quad \forall P \in \mathcal{P}$$

$$d_e \geq 0 \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

Multi – Cut LP

$$\max_f \sum_{P \in \mathcal{P}} f_P$$

$$\sum_{P \in \mathcal{P}_e} f_P \leq w_e \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

$$f_P \geq 0 \quad \forall P \in \mathcal{P}$$

Multi -Commodity Flow LP

Multi-Cut LP

- $\mathcal{P}$ denotes all paths between pair of vertices in ST
- $\mathcal{P}_e$ denotes the set of paths in $\mathcal{P}$ which contain edge e

$$\min \sum_{e \in \mathcal{E}_{\mathcal{H}}} w_e d_e$$

$$\sum_{e \in P} d_e \geq 1 \quad \forall P \in \mathcal{P}$$

$$d_e \geq 0 \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

Multi – Cut LP

$$\max_f \sum_{P \in \mathcal{P}} f_P$$

$$\sum_{P \in \mathcal{P}_e} f_P \leq w_e \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

$$f_P \geq 0 \quad \forall P \in \mathcal{P}$$

Multi -Commodity Flow LP

- Can be solved approximately i.e ε-approximation for any error parameter ε

Symmetric BMC LP (BMC-Sym LP)

$$
\begin{aligned}
\min \sum_{e \in \mathcal{E}_{\mathcal{H}}} w_e d_e & \\
\sum_{e \in P} d_e \geq 1 & \quad \forall P \in \mathcal{P} \\
d_e \geq 0 & \quad \forall e \in \mathcal{E}_{\mathcal{H}}
\end{aligned}
$$

$$
d_e = d_{\overline{e}} \quad \forall e \in \mathcal{E}_{\mathcal{H}}
$$

Symmetric BMC LP (BMC-Sym LP)

$$\min \sum_{e \in \mathcal{E}_{\mathcal{H}}} w_e d_e$$

$$\sum_{e \in P} d_e \geq 1 \quad \forall P \in \mathcal{P}$$

$$d_e \geq 0 \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

$$d_e = d_{\overline{e}} \quad \forall e \in \mathcal{E}_{\mathcal{H}}$$

- Very similar to Multi-Cut LP except for the symmetric constraints

Equivalence of LPs

Theorem – When the constructed graph is symmetric, BMC LP, BMC-Sym LP and Multi-Cut LP are equivalent

Equivalence of LPs

Theorem – When the constructed graph is symmetric, BMC LP, BMC-Sym LP and Multi-Cut LP are equivalent

Proof Outline

- Any feasible solution of each of the LP can be transformed into a feasible solution of other LPs without changing the objective value

Combinatorial Algorithm

Primal Step

Dual Step

Combinatorial Algorithm

Primal Step

Dual Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$

Dual Step

- Let $f_P = \min w_e$

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$

Dual Step

- Let $f_P = \min w_e$
- Update the flow in the path by f_P

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$

Dual Step

- Let $f_P = \min w_e$
- Update the flow in the path by f_P
- Update flow in complementary path

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$
- Update $d_e = d_e(1 + \frac{\epsilon f_P}{w_e})\ \forall e \in P$

Dual Step

- Let $f_P = \min w_e$
- Update the flow in the path by f_P
- Update flow in complementary path

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1) \ \forall (i_0, i_1) \in ST$
- Update $d_e = d_e(1 + \frac{\epsilon f_P}{w_e}) \ \forall e \in P$
- Update Complementary path

Dual Step

- Let $f_P = \min w_e$
- Update the flow in the path by f_P
- Update flow in complementary path

Combinatorial Algorithm

Primal Step

- Find shortest path P between $(i_0, i_1)\ \forall (i_0, i_1) \in ST$
- Update $d_e = d_e(1 + \frac{\epsilon f_P}{w_e})\ \forall e \in P$
- Update Complementary path

Dual Step

- Let $f_P = \min w_e$
- Update the flow in the path by f_P
- Update flow in complementary path

- Converges in $O(\epsilon^{-2} k m^2)$
- Can be improved to $O(\epsilon^{-2} m^2)$ (Fleischer 1999)

Relationship with Cycle Inequalities

- BMC LP is closely related to cycle inequalities
  - BMC LP with slight modification is equivalent to cycle inequalities

Relationship with Cycle Inequalities

- BMC LP is closely related to cycle inequalities
  - BMC LP with slight modification is equivalent to cycle inequalities
- Relates our work to many recent works solving cycle inequalities

Empirical Results

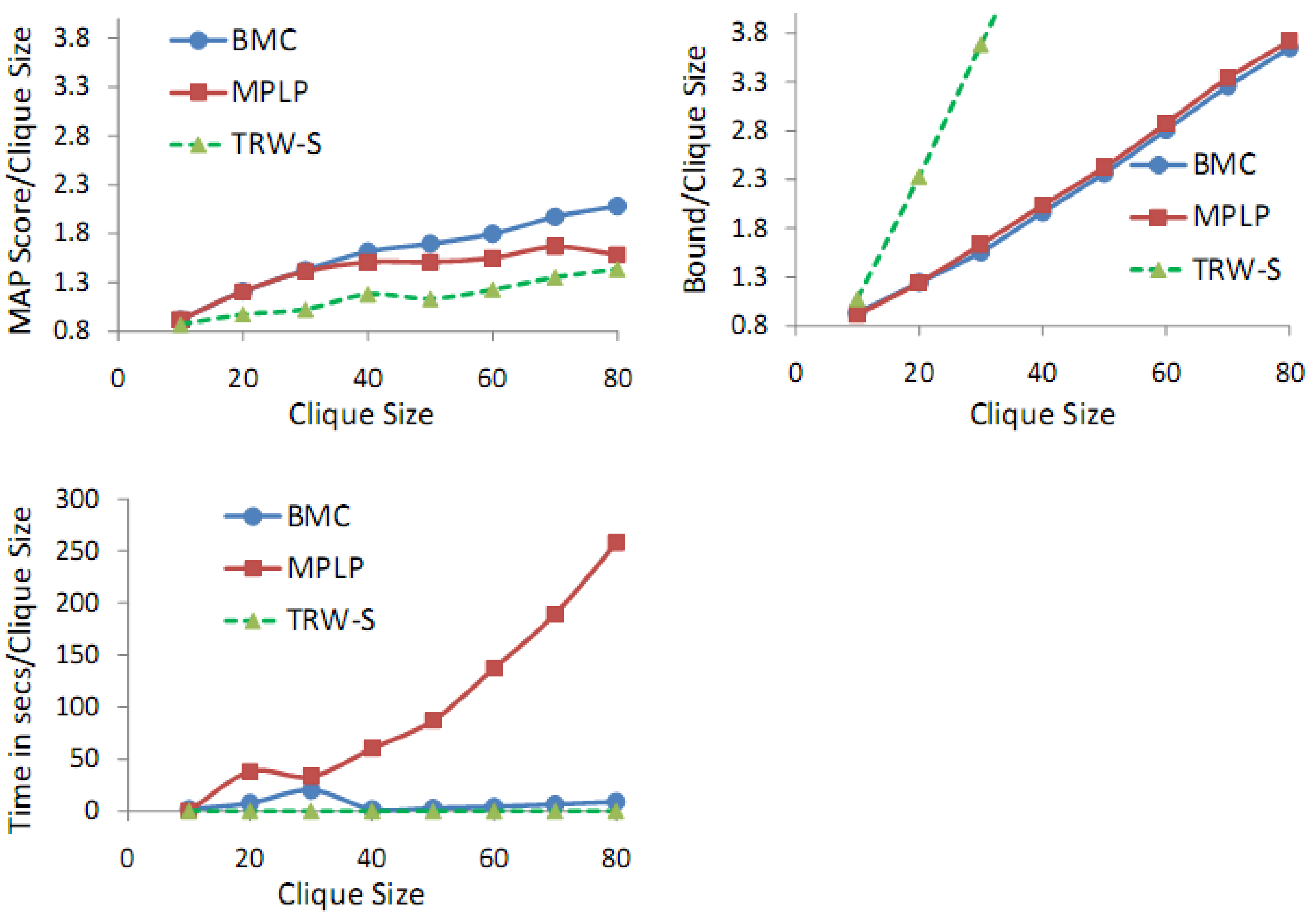
- Data Sets
  - Synthetic Problems
    - Clique Graphical models of various sizes
  - Benchmark Data Set
    - Max Cut Graphs of various sizes and density

Empirical Results

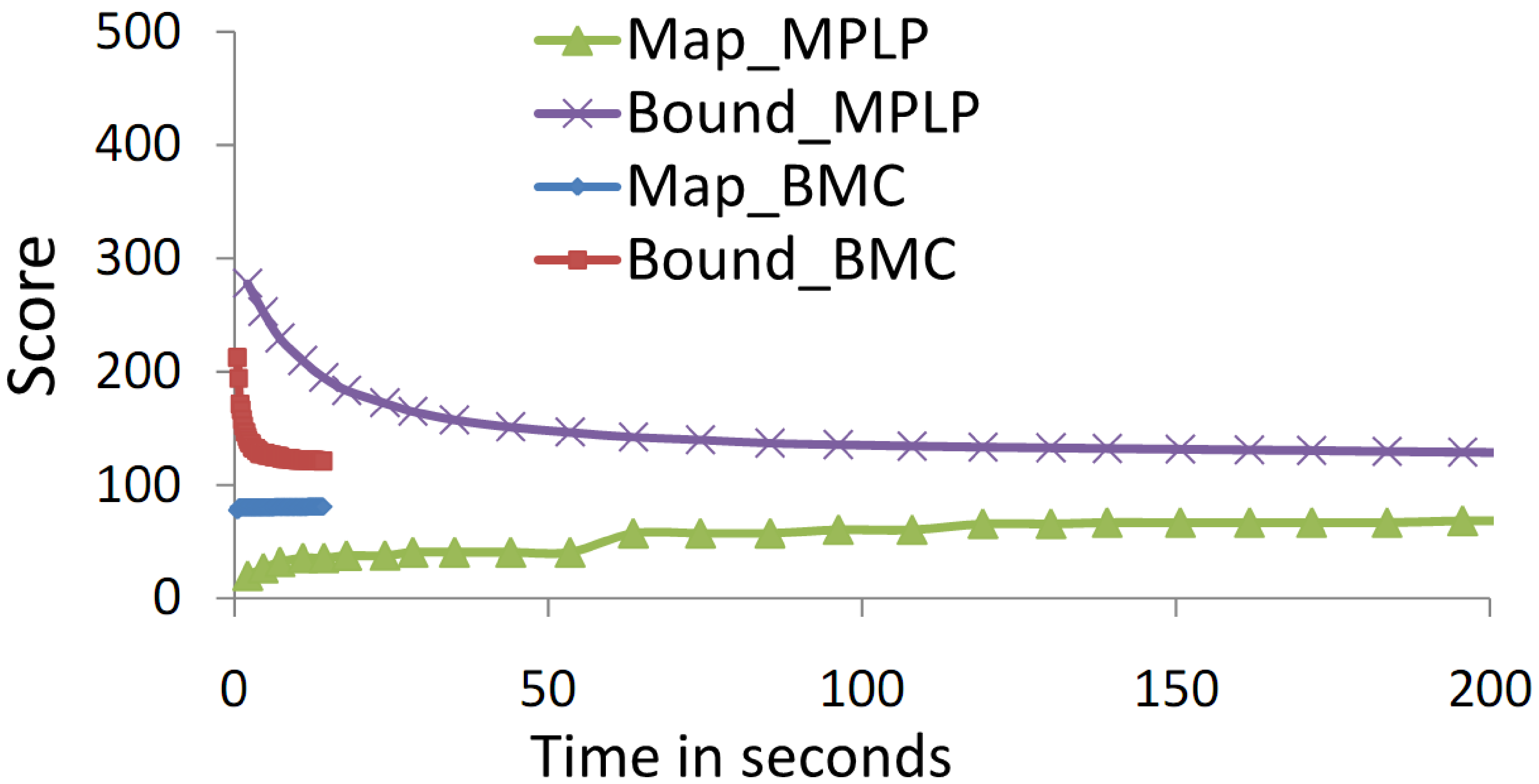
- Data Sets
  - Synthetic Problems
    - Clique Graphical models of various sizes
  - Benchmark Data Set
    - Max Cut Graphs of various sizes and density

- Algorithms
  - TRW-S
  - BMC
    - $\varepsilon = 0.02$
  - MPLP
    - *1000* iterations or up to convergence

Empirical Results



Empirical Results



Convergence Comparison of BMC and MPLP

Empirical Results

		Bound			Time in seconds		
Graph	density	BMC	MPLP	TRW-S	BMC	MPLP	TRW-S
pm1s	0.1	131	200	257	45	43	0.005
pw01	0.1	2079	2397	2745	48	46	0.006
w01	0.1	720	1115	1320	46	41	0.004
g05	0.5	1650	1720	2475	761	317	0.021
pw05	0.5	9131	9195	13696	699	1139	0.021
w05	0.5	2245	2488	6588	737	1261	0.021
pw09	0.9	16493	16404	24563	106	2524	0.041
w09	0.9	4073	4095	11763	123	2671	0.053
pm1d	0.99	842	924	2463	12	1307	0.047

Empirical Results

		Bound			Time in seconds		
Graph	density	BMC	MPLP	TRW-S	BMC	MPLP	TRW-S
pm1s	0.1	131	200	257	45	43	0.005
pw01	0.1	2079	2397	2745	48	46	0.006
w01	0.1	720	1115	1320	46	41	0.004
g05	0.5	1650	1720	2475	761	317	0.021
pw05	0.5	9131	9195	13696	699	1139	0.021
w05	0.5	2245	2488	6588	737	1261	0.021
pw09	0.9	16493	16404	24563	106	2524	0.041
w09	0.9	4073	4095	11763	123	2671	0.053
pm1d	0.99	842	924	2463	12	1307	0.047

Conclusion & Future Work

- Conclusion

Conclusion & Future Work

- Conclusion
  - MAP estimation can be reduced to Bipartite Multi-Cut problem
  - Algorithms for multi commodity flow can be used for MAP estimation

Conclusion & Future Work

- Conclusion
  - MAP estimation can be reduced to Bipartite Multi-Cut problem
  - Algorithms for multi commodity flow can be used for MAP estimation
  - BMC LP is closely related to cycle inequalities

Conclusion & Future Work

- Conclusion
  - MAP estimation can be reduced to Bipartite Multi-Cut problem
  - Algorithms for multi commodity flow can be used for MAP estimation
  - BMC LP is closely related to cycle inequalities
- Future Work
  - Extensions to multi-label graphical models

Conclusion & Future Work

- Conclusion
  - MAP estimation can be reduced to Bipartite Multi-Cut problem
  - Algorithms for multi commodity flow can be used for MAP estimation
  - BMC LP is closely related to cycle inequalities
- Future Work
  - Extensions to multi-label graphical models
  - Combinatorial algorithm for solving SDP relaxation

Conclusion & Future Work

- Conclusion
  - MAP estimation can be reduced to Bipartite Multi-Cut problem
  - Algorithms for multi commodity flow can be used for MAP estimation
  - BMC LP is closely related to cycle inequalities
- Future Work
  - Extensions to multi-label graphical models
  - Combinatorial algorithm for solving SDP relaxation
- Acknowledgement

  (Travel Support)

  Google™