Discrete Inference and Learning Lecture 3

MVA

2017 – 2018

http://thoth.inrialpes.fr/~alahari/disinflearn

Slides based on material from Pushmeet Kohli, Nikos Komodakis, M. Pawan Kumar

Recap

The st-mincut problem

Connection between st-mincut and energy minimization?

What problems can we solve using st-mincut?

st-mincut based Move algorithms

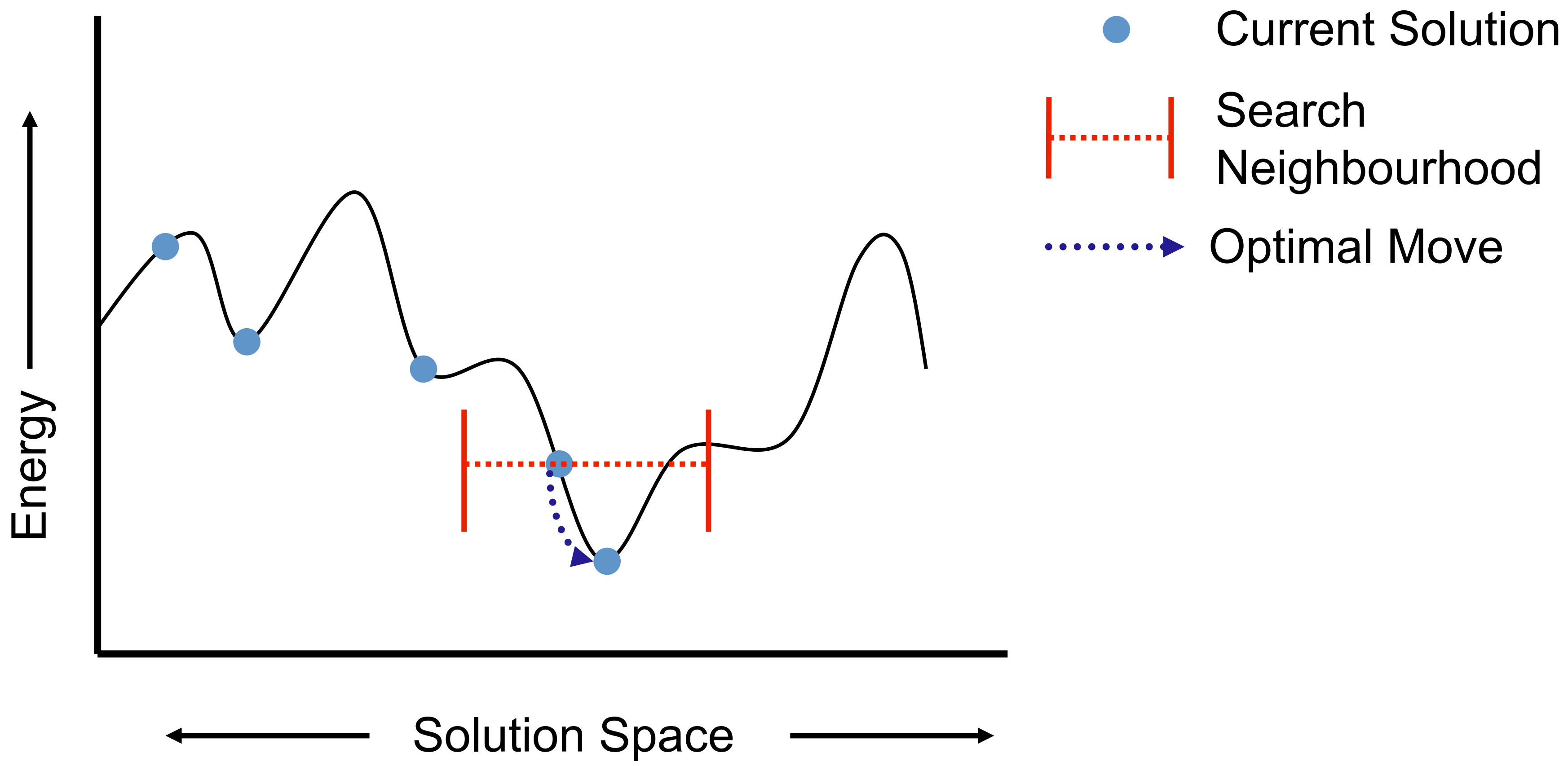
St-mincut based Move algorithms

$$E(\mathbf{y}) = \sum_{i} \theta_i (y_i) + \sum_{i,j} \theta_{ij} (y_i, y_j)$$

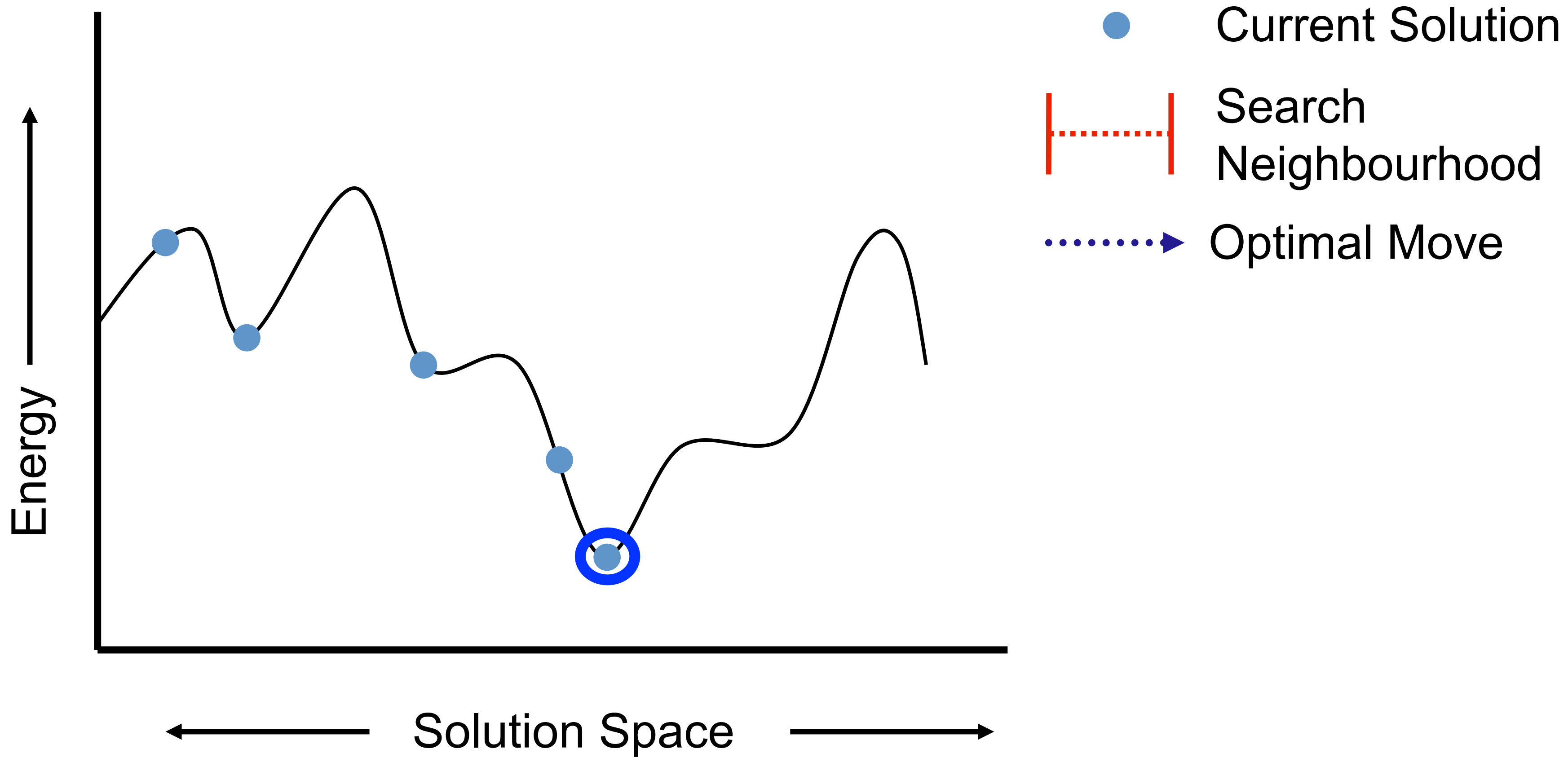
y ϵ Labels L = {l_1, l_2, … , l_k}

- Commonly used for solving **non-submodular** multi-label problems
- Extremely efficient and produce good solutions
- Not Exact: Produce local optima

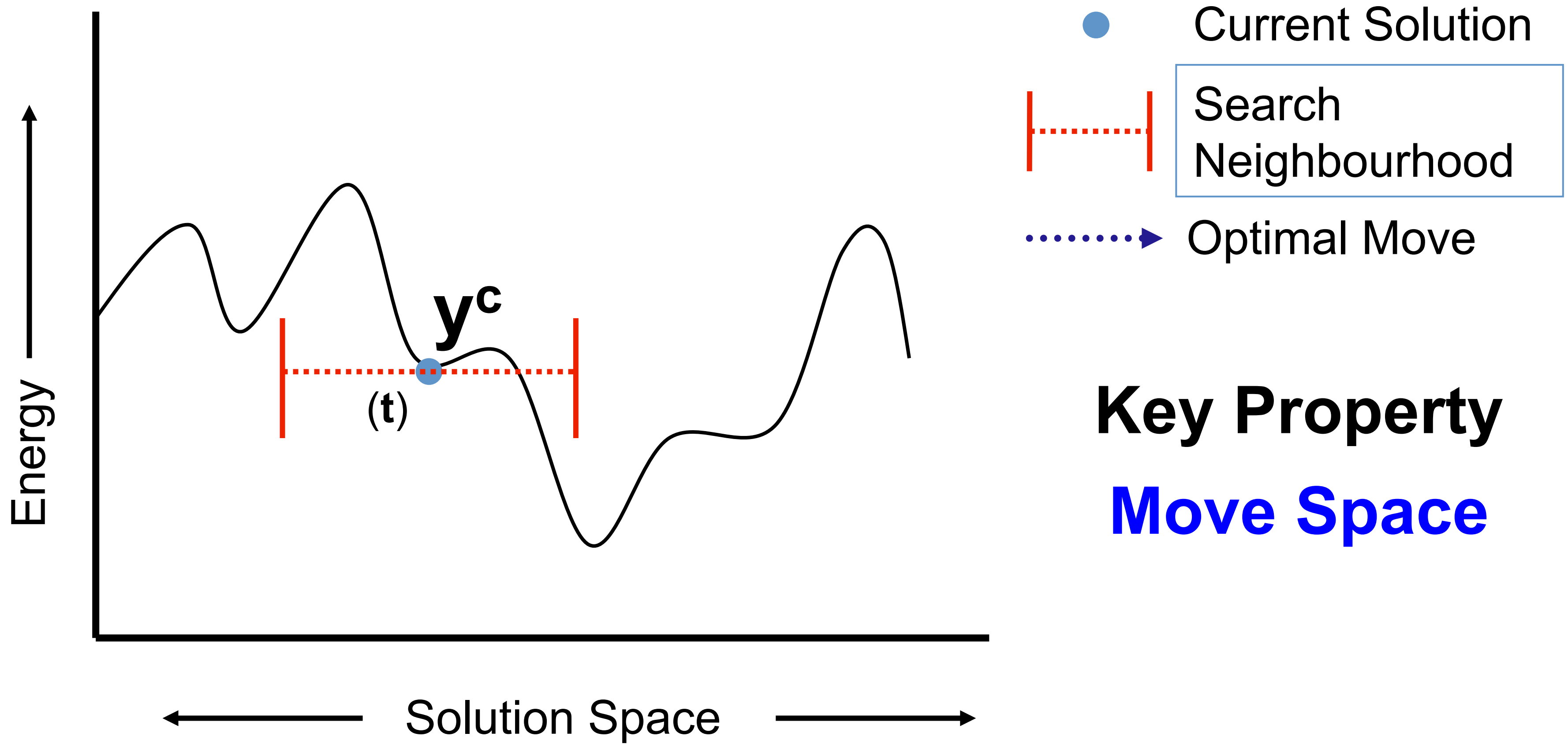
Move Making Algorithms



Move Making Algorithms



Computing the Optimal Move



Key Property

Move Space

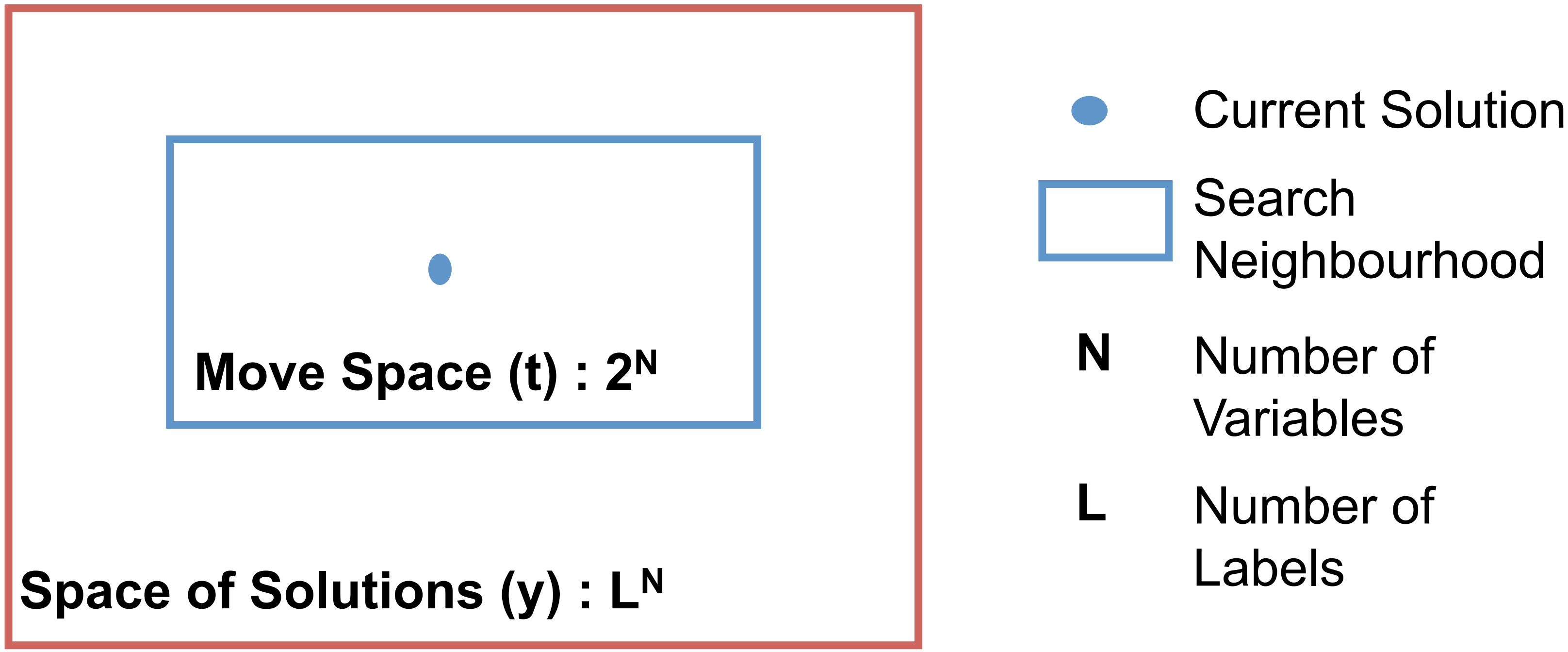
Bigger move space →
- **Better solutions**
- **Finding the optimal move hard**

Moves using Graph Cuts

Expansion and Swap move algorithms

[Boykov Veksler and Zabih, PAMI 2001]

- **Makes a series of changes to the solution (moves)**
- **Each move results in a solution with smaller energy**

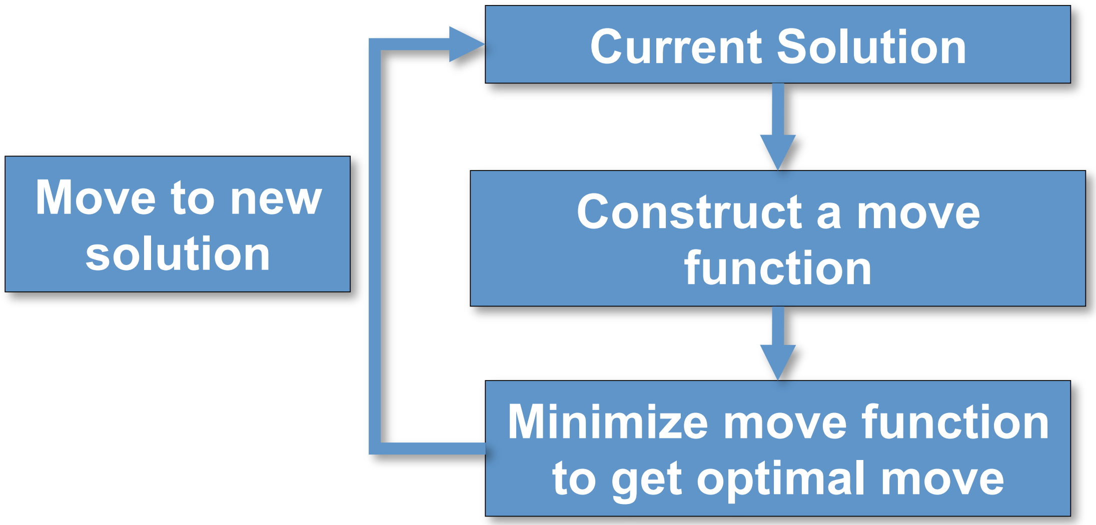


Moves using Graph Cuts

Expansion and Swap move algorithms

[Boykov Veksler and Zabih, PAMI 2001]

- **Makes a series of changes to the solution (moves)**
- **Each move results in a solution with smaller energy**



How to minimize move functions?

General Binary Moves

$$y = t\, y^1 + (1- t)\, y^2$$

New solution — Current Solution — Second solution

$$E_m(t) = E(t\, y^1 + (1- t)\, y^2)$$

Minimize over move variables t to get the optimal move

Move energy is a submodular QPBF (Exact Minimization Possible)

Boykov, Veksler and Zabih, PAMI 2001

Expansion Move

- Variables take label α or retain current label

→ Tree
Ground
House
Sky

Status: Initialize with Tree

[Boykov, Veksler, Zabih]

Expansion Move

- **Variables take label *α* or retain current label**

Tree
→ Ground
House
Sky

Status: Expand Ground

[Boykov, Veksler, Zabih]

Expansion Move

- Variables take label *α* or retain current label

Tree
Ground
→ House
Sky

Status: Expand House

[Boykov, Veksler, Zabih]

Expansion Move

- Variables take label α or retain current label

Tree
Ground
House
Sky

Status: Expand Sky

[Boykov, Veksler, Zabih]

Expansion Move

- Variables take label *α* or retain current label

- Move energy is submodular if:
  - Unary Potentials: Arbitrary
  - Pairwise potentials: Metric

$$\theta_{ij}(l_a, l_b) \geq 0$$

$$\theta_{ij}(l_a, l_b) = 0 \quad \text{iff} \quad a = b$$

Semi metric

Examples: Potts model, Truncated linear

Cannot solve truncated quadratic

[Boykov, Veksler, Zabih]

Expansion Move

- Variables take label *α* or retain current label

- Move energy is submodular if:
  - Unary Potentials: Arbitrary
  - Pairwise potentials: Metric

$$\theta_{ij}(l_a,l_b) + \theta_{ij}(l_b,l_c) \geq \theta_{ij}(l_a,l_c)$$

Triangle Inequality

Examples: Potts model, Truncated linear

Cannot solve truncated quadratic

[Boykov, Veksler, Zabih]

Swap Move

- Variables labeled *α*, *β* can swap their labels

Swap Sky, House

[Boykov, Veksler, Zabih]

Swap Move

- **Variables labeled *α*, *β* can swap their labels**

Swap Sky, House

Tree
Ground
→ House
→ Sky

[Boykov, Veksler, Zabih]

Swap Move

- Variables labeled α, β can swap their labels

- Move energy is submodular if:
  - Unary Potentials: Arbitrary
  - Pairwise potentials: Semimetric

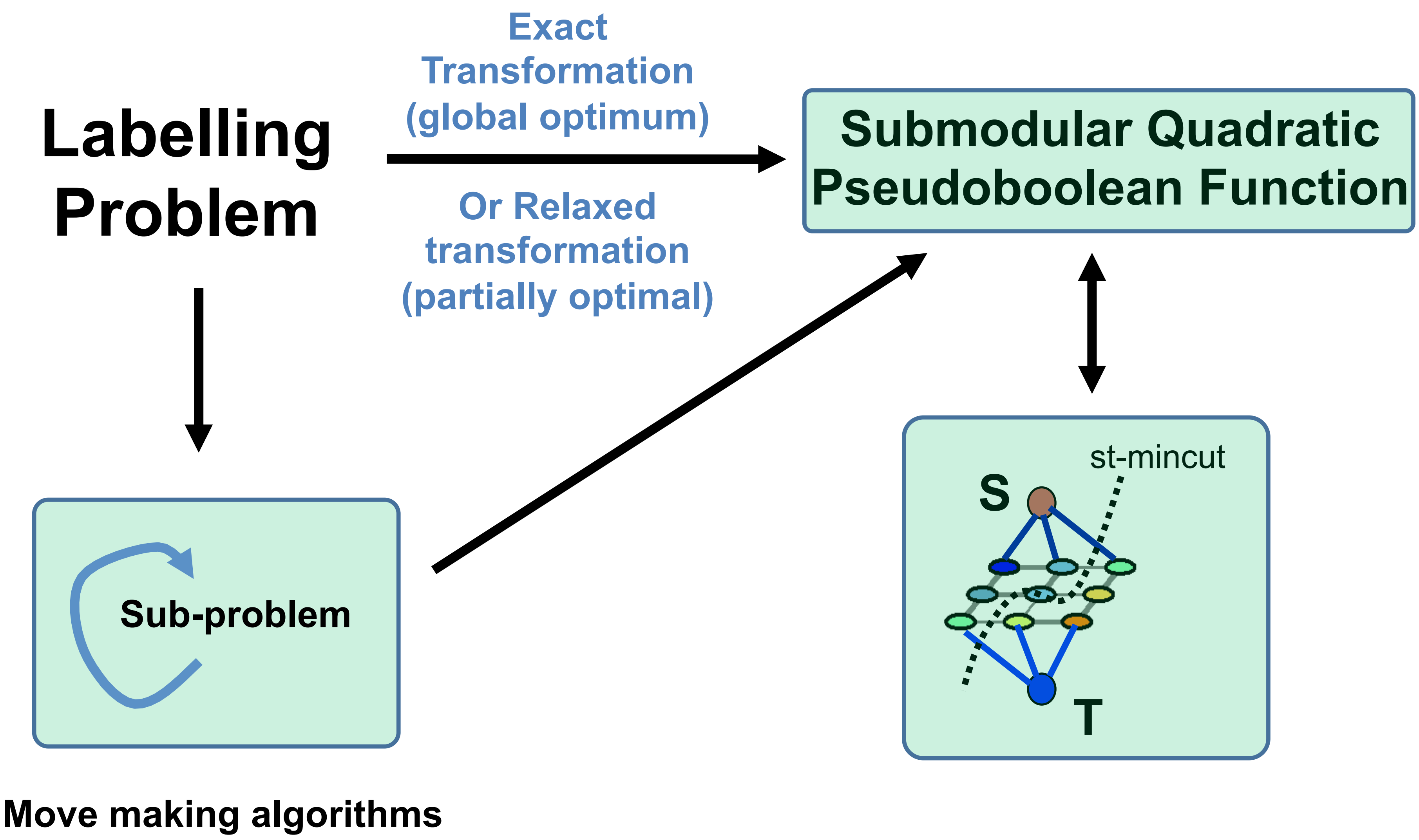
$$\theta_{ij}(l_a, l_b) \geq 0$$

$$\theta_{ij}(l_a, l_b) = 0 \longleftrightarrow a = b$$

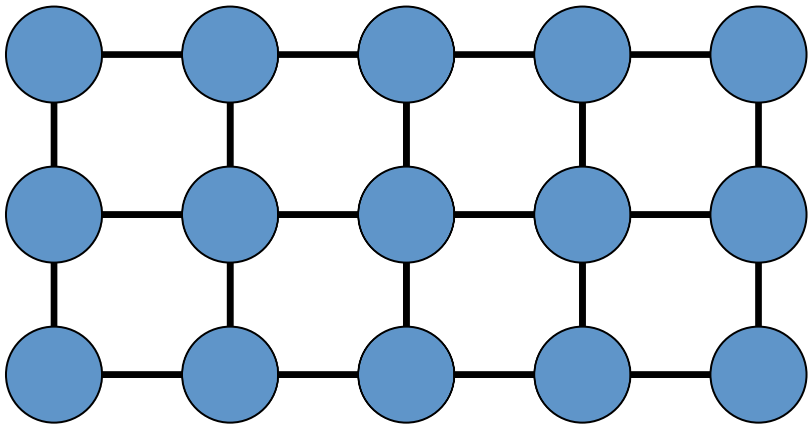
Examples: Potts model, Truncated Convex

[Boykov, Veksler, Zabih]

Summary

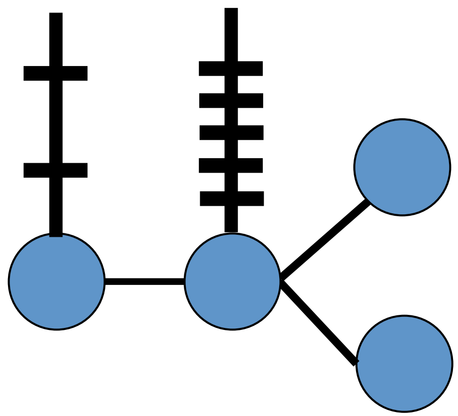

Labelling Problem
Exact Transformation (global optimum)
Or Relaxed transformation (partially optimal)
Submodular Quadratic Pseudoboolean Function
Sub-problem
st-mincut
S
T
Move making algorithms

Where do we stand ?



Grid graph -
submodular, 2-label: Use graphcuts
“metric”: Use expansion

otherwise: Use TRW,
dual decomposition,
relaxation

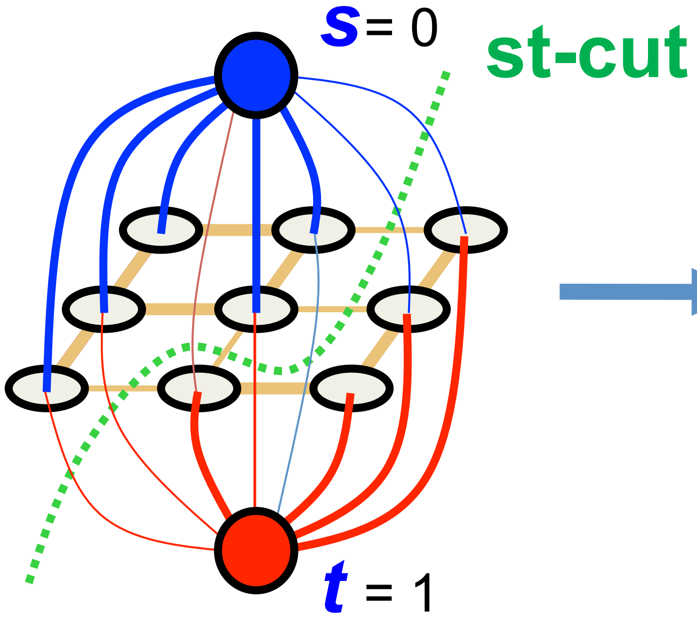


Chain/Tree, 2/multi-label: Use BP

Dynamic Energy Minimization

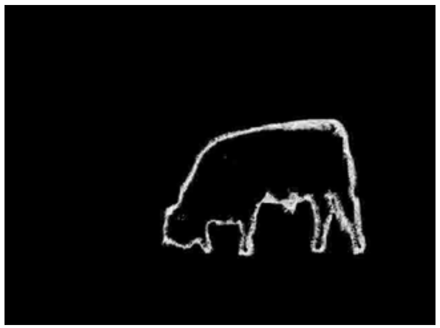
Image Segmentation in Video

$E(x)$



x^*

Video frame



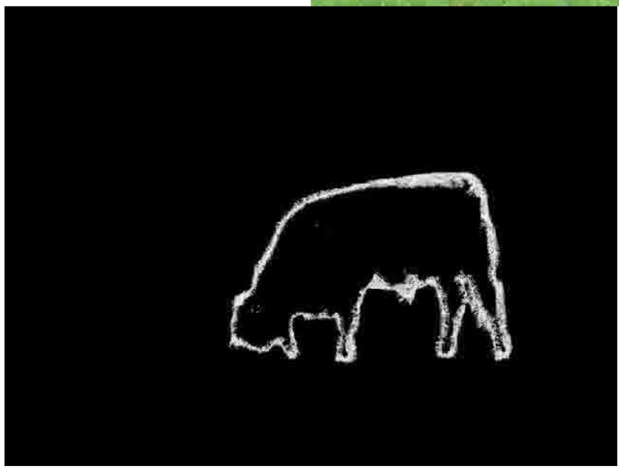
Flow

Global Optimum

Image Segmentation in Video

Video frame

Flow



Global Optimum

Dynamic Energy Minimization

E_A —minimize→ S_A

Can we do better?

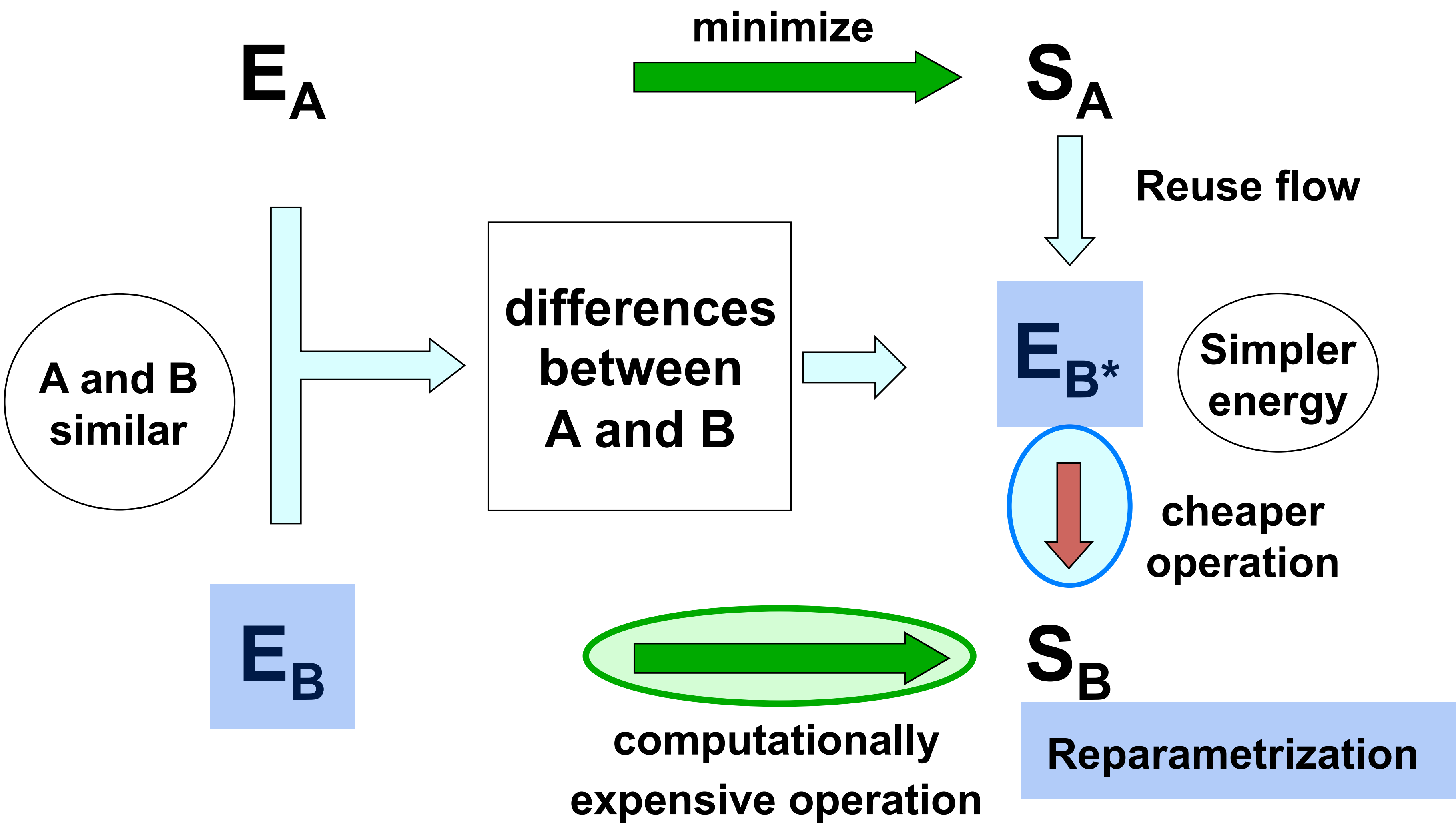
Recycling Solutions

E_B → S_B

computationally expensive operation

Boykov & Jolly ICCV'01, Kohli & Torr (ICCV05, PAMI07)

Dynamic Energy Minimization



Boykov & Jolly ICCV'01, Kohli & Torr (ICCV05, PAMI07)

Dynamic Energy Minimization

Original Energy

$$E(a_1,a_2) = 2a_1 + 5\bar{a}_1 + 9a_2 + 4\bar{a}_2 + 2a_1\bar{a}_2 + \bar{a}_1 a_2$$

Reparameterized Energy

$$E(a_1,a_2) = 8 + \bar{a}_1 + 3a_2 + 3\bar{a}_1 a_2$$

New Energy

$$E(a_1,a_2) = 2a_1 + 5\bar{a}_1 + 9a_2 + 4\bar{a}_2 + 7a_1\bar{a}_2 + \bar{a}_1 a_2$$

New Reparameterized Energy

$$E(a_1,a_2) = 8 + \bar{a}_1 + 3a_2 + 3\bar{a}_1 a_2 + 5a_1\bar{a}_2$$

Boykov & Jolly ICCV’01, Kohli & Torr (ICCV05, PAMI07)

Outline

- Reparameterization (lecture 1)
- Belief Propagation (lecture 1)
- Tree-reweighted Message Passing
  - Integer Programming Formulation
  - Linear Programming Relaxation and its Dual
  - Convergent Solution for Dual
  - Computational Issues and Theoretical Properties

First…

Recap of Integer Linear Program

Integer Linear Program

$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$

s.t. $A\,\mathbf{x} \leq \mathbf{b}$

$\mathbf{x}$ is an integer vector

Every element of **x** is an integer

Integer Linear Program

$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$

s.t. $A\,\mathbf{x} \leq \mathbf{b}$

$\mathbf{x} \in Z^n$

Every element of **x** is an integer

Example

$\max_{\mathbf{x}} x_1 + x_2$

s.t. $x_1 \geq 0$

$x_2 \geq 0$

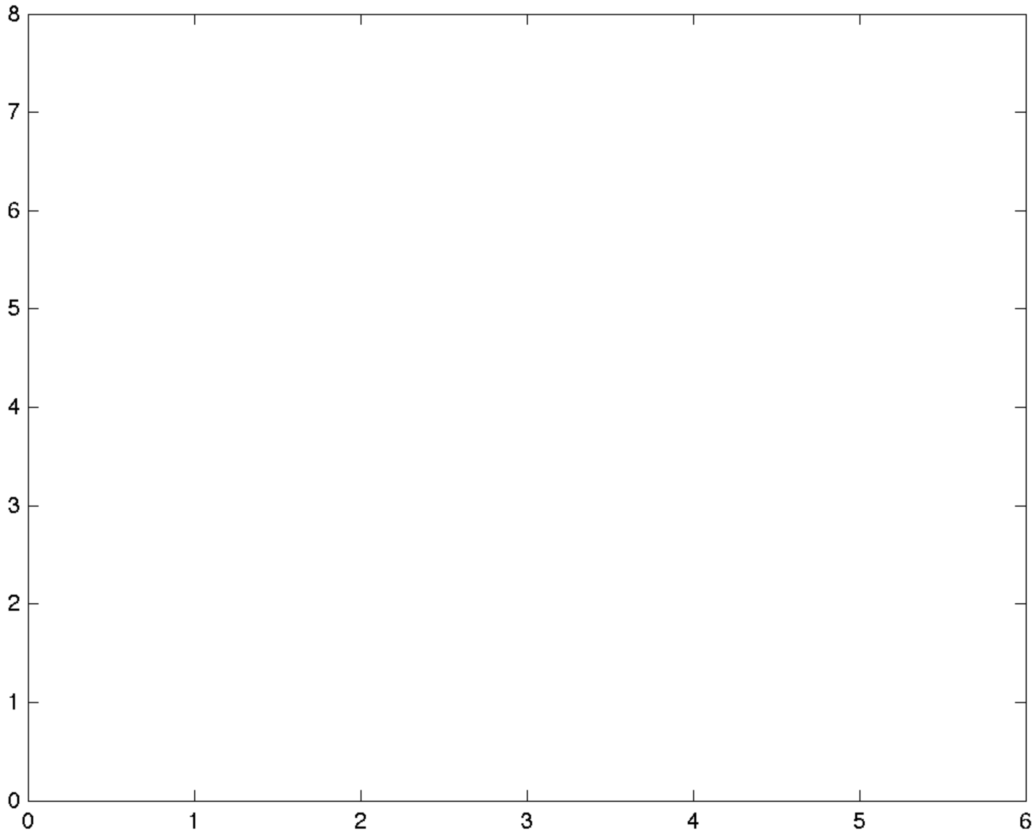
$4x_1 - x_2 \leq 8$

$2x_1 + x_2 \leq 10$

$5x_1 - 2x_2 \geq -2$

$\mathbf{x} \in Z^n$

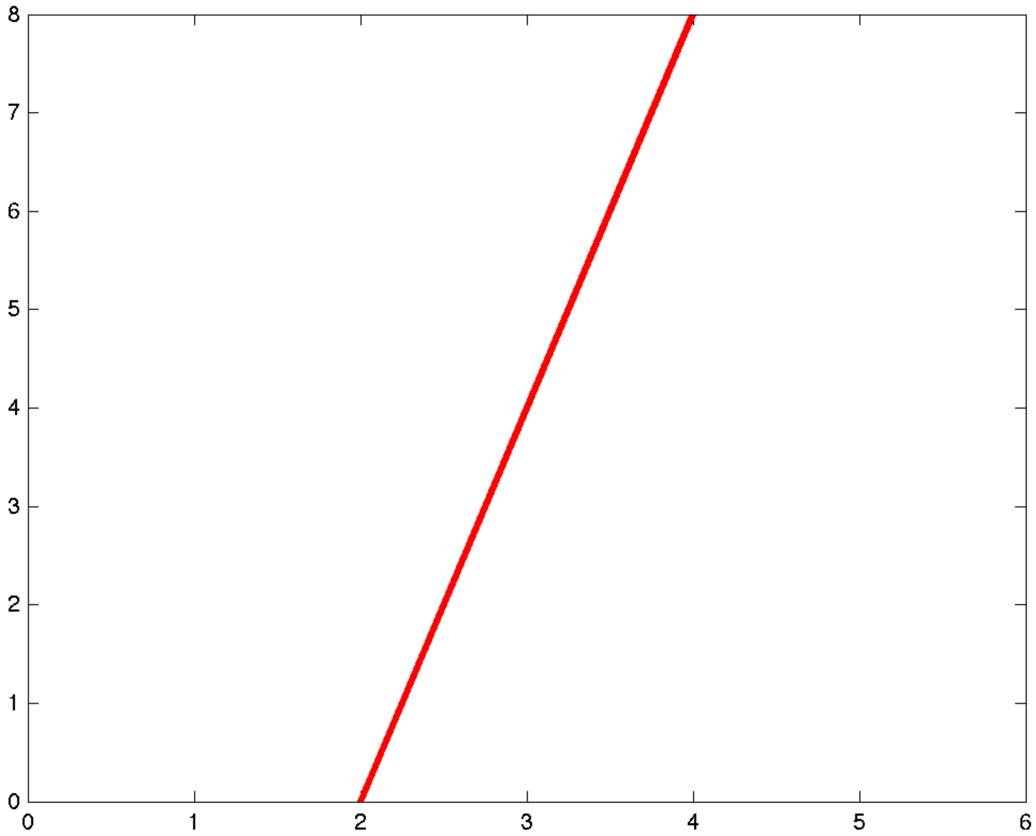
Example



$x_1 \geq 0$

$x_2 \geq 0$

Example

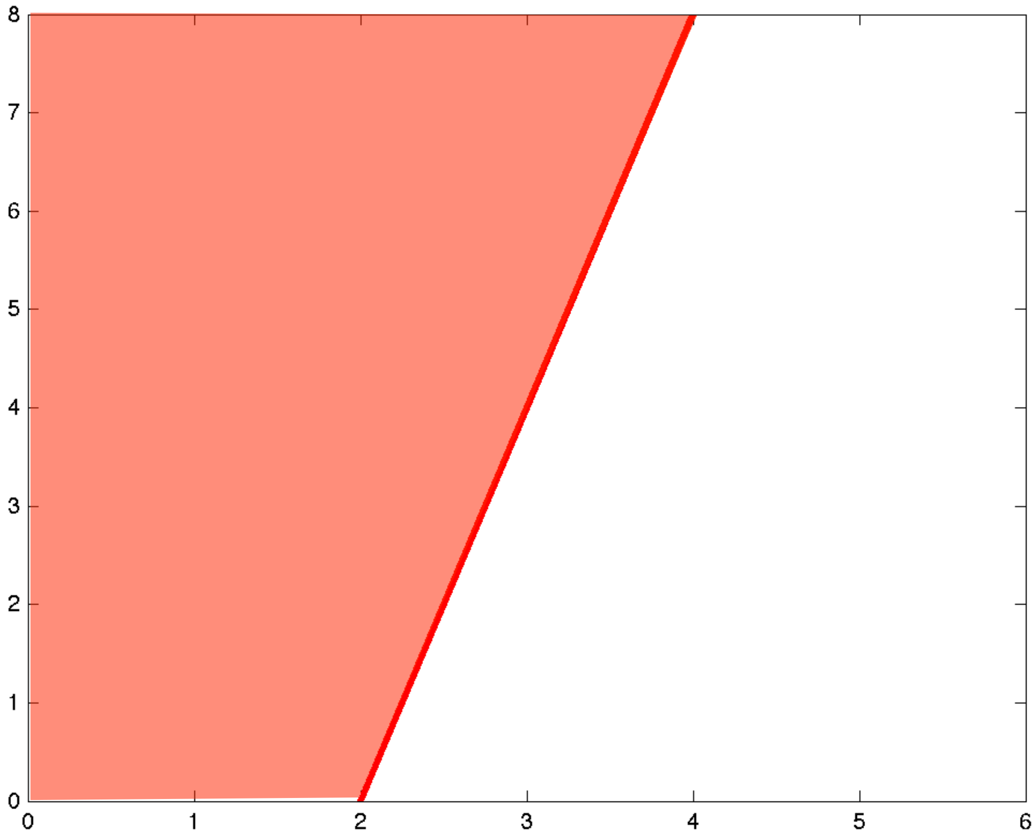


$x_1 \geq 0$

$x_2 \geq 0$

$4x_1 - x_2 = 8$

Example

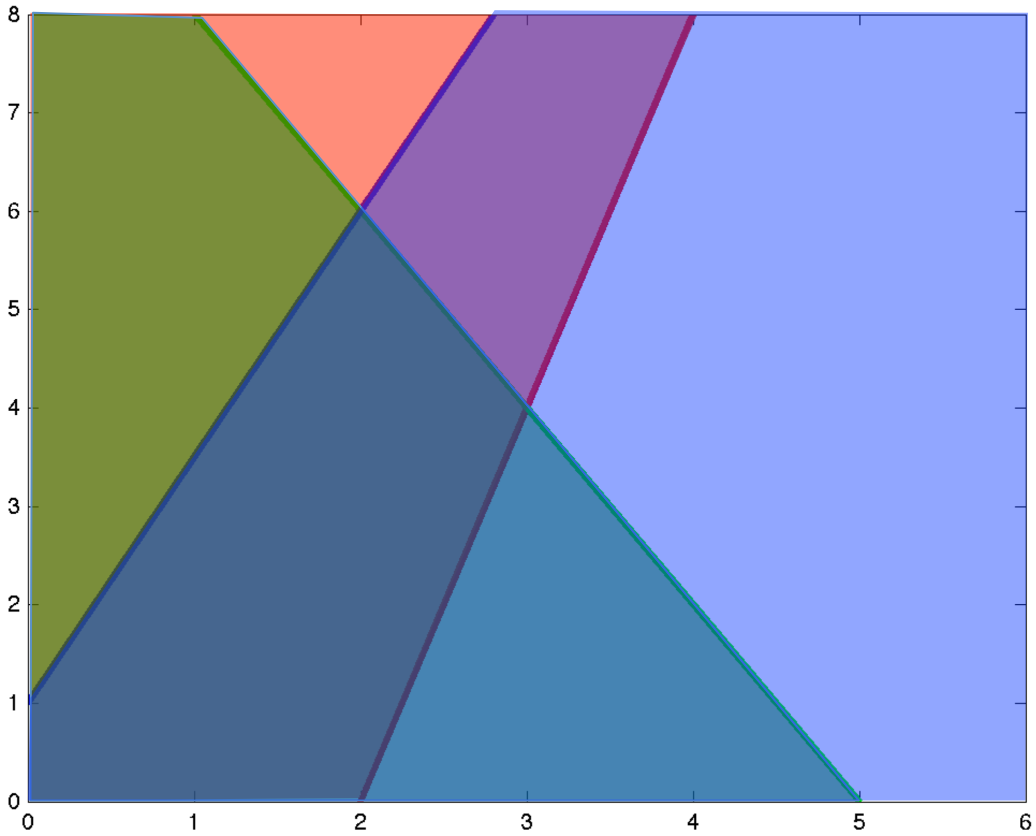


$x_1 \geq 0$

$x_2 \geq 0$

$4x_1 - x_2 \leq 8$

Example



$x_1 \geq 0$

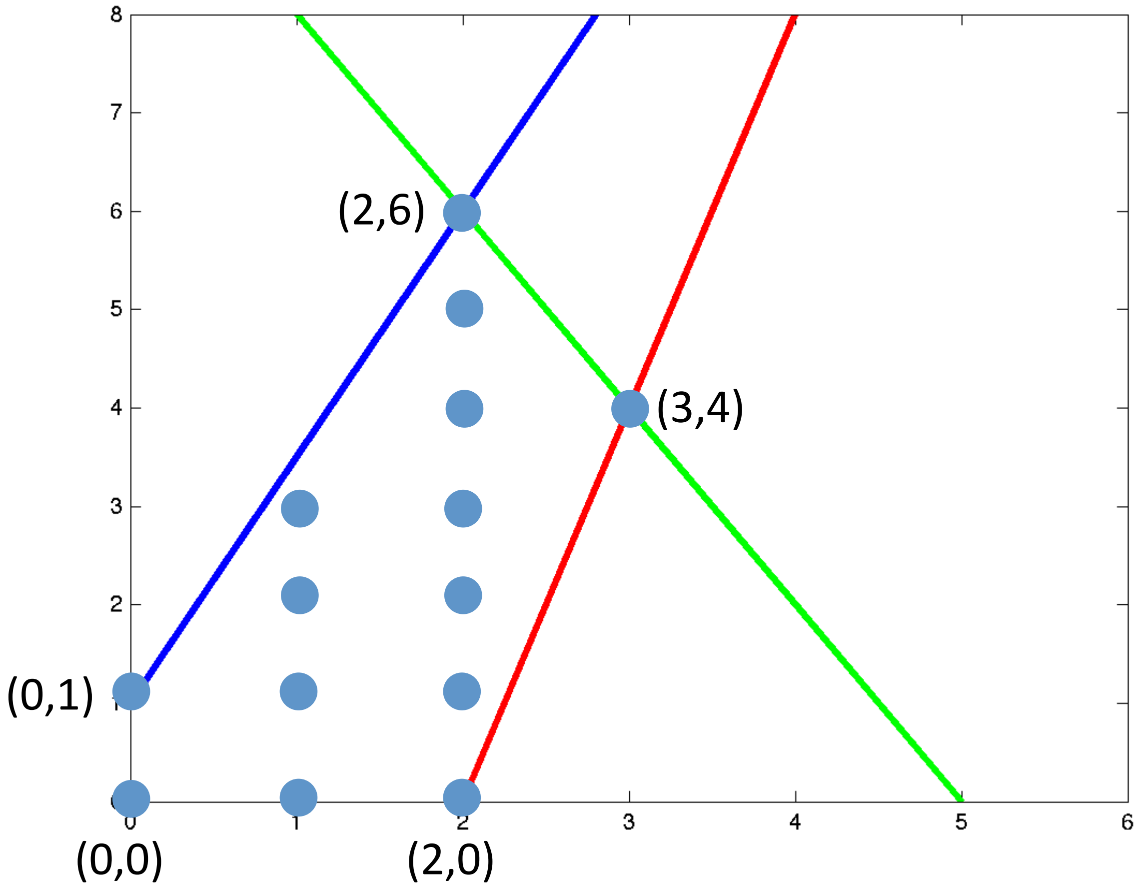
$x_2 \geq 0$

$4x_1 - x_2 \leq 8$

$2x_1 + x_2 \leq 10$

$5x_1 - 2x_2 \geq -2$

Example



$x_1 \geq 0$

$x_2 \geq 0$

$4x_1 - x_2 \leq 8$

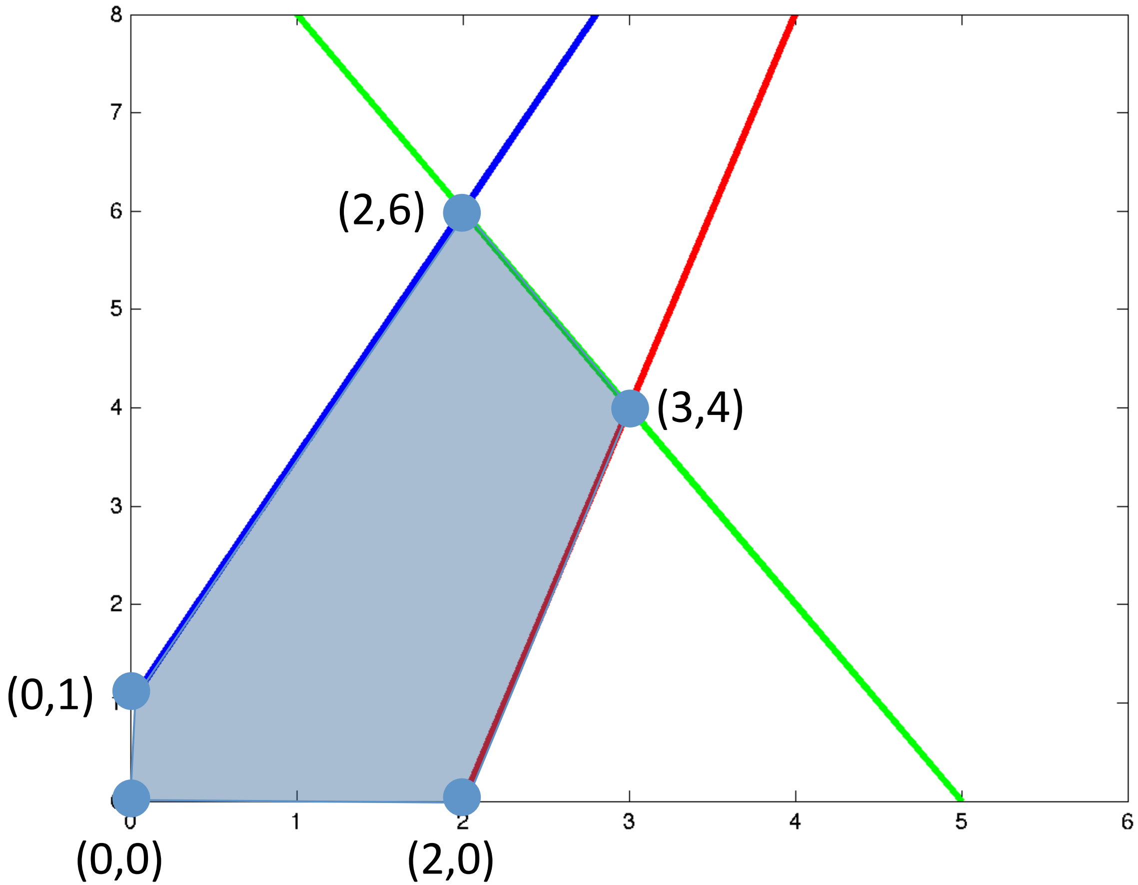
$2x_1 + x_2 \leq 10$

$5x_1 - 2x_2 \geq -2$

$\mathbf{x} \in Z^n$

$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$

Example



$x_1 \geq 0$

$x_2 \geq 0$

$4x_1 - x_2 \leq 8$

$2x_1 + x_2 \leq 10$

$5x_1 - 2x_2 \geq -2$

~~$\mathbf{x} \in Z^n$~~

$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$

Why? **True in general?**

Why is the solution at a vertex?

Let x^* be the minimum. Since each point in a polytope is a convex combination of the vertices $v_1, \ldots, v_t$, we have

$$x^* = \lambda_1 v_1 + \ldots + \lambda_t v_t$$

and the objective value at optimality can be expressed as

$$cx^* = \lambda_1 * (cv_1) + \ldots + \lambda_t(cv_t).$$

Assume that the minimum is not at a vertex, i.e.,

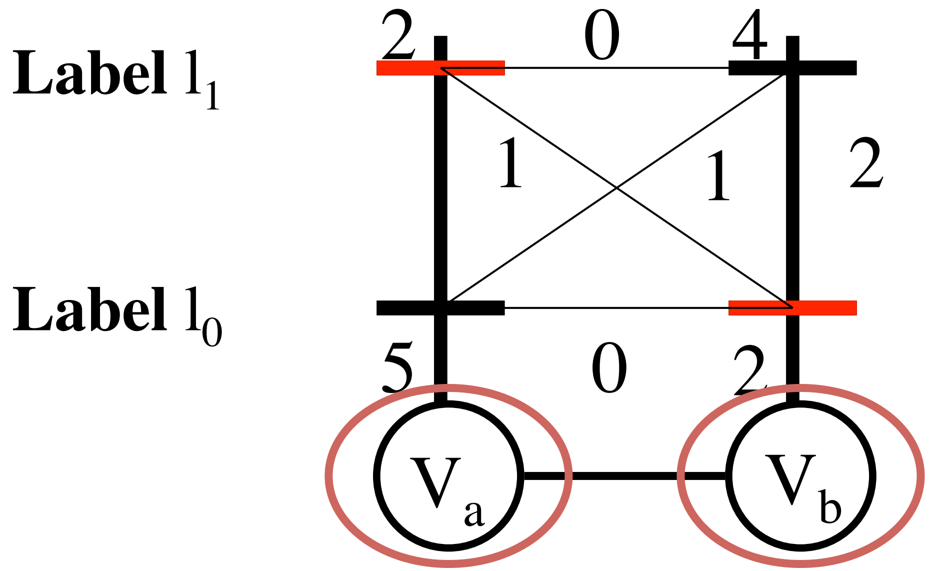
$$cx^* < cv_i \quad \forall i : 1 \leq i \leq t.$$

It follows that

$$\begin{aligned} cx^* &= \lambda_1 * (cv_1) + \ldots + \lambda_t(cv_t) \\ &> \lambda_1 * (cx^*) + \ldots + \lambda_t(cx^*) \\ &> (\lambda_1 + \ldots + \lambda_t)(cx^*) \\ &> cx^*. \end{aligned}$$

Hence, it must be the case that $x^* = v_i$ for some $1 \leq i \leq t$.

Integer Programming Formulation



Unary Potentials

$\theta_{a;0} = 5$ $\quad \theta_{b;0} = 2$

$\theta_{a;1} = 2$ $\quad \theta_{b;1} = 4$

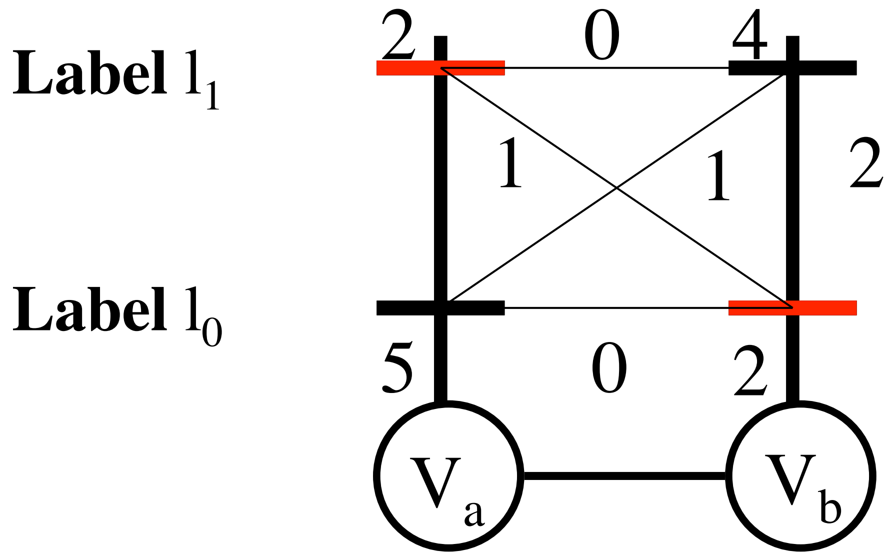
Labeling

f(a) = 1 $\quad y_{a;0} = 0 \quad y_{a;1} = 1$

f(b) = 0 $\quad y_{b;0} = 1 \quad y_{b;1} = 0$

Any f(.) has equivalent boolean variables $y_{a;i}$

Integer Programming Formulation



Unary Potentials

$\theta_{a;0} = 5$ $\theta_{b;0} = 2$

$\theta_{a;1} = 2$ $\theta_{b;1} = 4$

Labeling

$f(a) = 1$ $y_{a;0} = 0$ $y_{a;1} = 1$

$f(b) = 0$ $y_{b;0} = 1$ $y_{b;1} = 0$

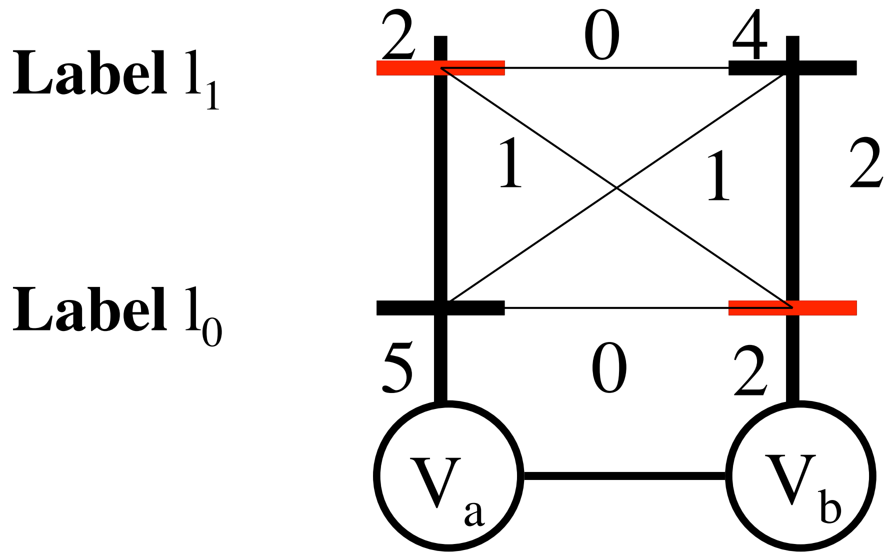
Find the optimal variables $y_{a;i}$

Integer Programming Formulation

Unary Potentials

$\theta_{a;0} = 5$ $\theta_{b;0} = 2$

$\theta_{a;1} = 2$ $\theta_{b;1} = 4$



Sum of Unary Potentials

$$\sum_a \sum_i \theta_{a;i} \, y_{a;i}$$

$$y_{a;i} \in \{0,1\}, \text{ for all } V_a, l_i$$

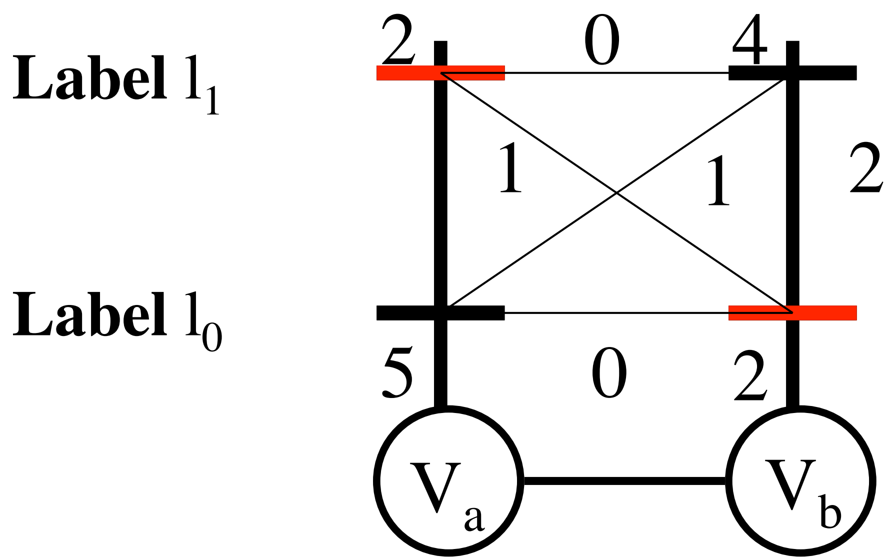
$$\sum_i y_{a;i} = 1, \text{ for all } V_a$$

Integer Programming Formulation

Pairwise Potentials

$\theta_{ab;00} = 0 \quad \theta_{ab;01} = 1$

$\theta_{ab;10} = 1 \quad \theta_{ab;11} = 0$



Sum of Pairwise Potentials

$$\sum_{(a,b)} \sum_{ik} \theta_{ab;ik} \, y_{a;i} y_{b;k}$$

$$y_{a;i} \in \{0,1\}$$

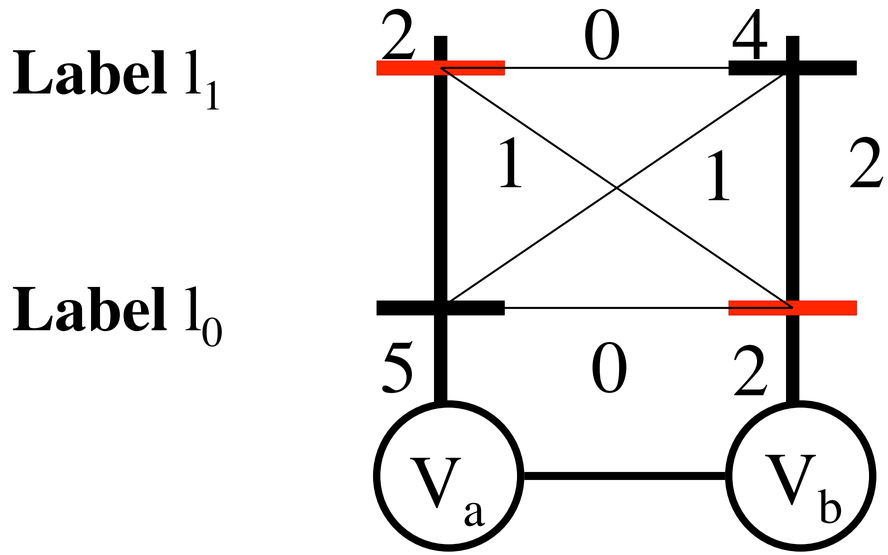
$$\sum_i y_{a;i} = 1$$

Integer Programming Formulation

Pairwise Potentials

$\theta_{ab;00} = 0 \quad \theta_{ab;01} = 1$

$\theta_{ab;10} = 1 \quad \theta_{ab;11} = 0$



Sum of Pairwise Potentials

$$\sum_{(a,b)} \sum_{ik} \theta_{ab;ik} \, y_{ab;ik}$$

$$y_{a;i} \in \{0,1\} \quad y_{ab;ik} = y_{a;i} \, y_{b;k}$$

$$\sum_i y_{a;i} = 1$$

Integer Programming Formulation

$$\min \sum_a \sum_i \theta_{a;i} \, y_{a;i} + \sum_{(a,b)} \sum_{ik} \theta_{ab;ik} \, y_{ab;ik}$$

$$y_{a;i} \in \{0,1\}$$

$$\sum_i y_{a;i} = 1$$

$$y_{ab;ik} = y_{a;i} \, y_{b;k}$$

Integer Programming Formulation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in \{0,1\}$$

$$\sum_i y_{a;i} = 1$$

$$y_{ab;ik} = y_{a;i}\, y_{b;k}$$

$$\boldsymbol{\theta} = [\ \ldots\ \theta_{a;i}\ \ldots.\ ;\ \ldots\ \theta_{ab;ik}\ \ldots.]$$

$$\mathbf{y} = [\ \ldots\ y_{a;i}\ \ldots.\ ;\ \ldots\ y_{ab;ik}\ \ldots.]$$

Integer Programming Formulation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in \{0,1\}$$

$$\sum_i y_{a;i} = 1$$

$$y_{ab;ik} = y_{a;i}\, y_{b;k}$$

Solve to obtain MAP labeling y*

Integer Programming Formulation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in \{0,1\}$$

$$\sum_i y_{a;i} = 1$$

$$y_{ab;ik} = y_{a;i}\, y_{b;k}$$

But we can't solve it in general

Outline

- Reparameterization (lecture 1)
- Belief Propagation (lecture 1)
- Tree-reweighted Message Passing
  - Integer Programming Formulation
  - Linear Programming Relaxation and its Dual
  - Convergent Solution for Dual
  - Computational Issues and Theoretical Properties

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$\boxed{y_{a;i} \in \{0,1\}}$$

$$\sum_i y_{a;i} = 1$$

$$\boxed{y_{ab;ik} = y_{a;i}\, y_{b;k}}$$

Two reasons why we can't solve this

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

$$\boxed{y_{ab;ik} = y_{a;i}\, y_{b;k}}$$

One reason why we can't solve this

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

$$\sum_k y_{ab;ik} = \sum_k y_{a;i}\, y_{b;k}$$

One reason why we can't solve this

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

$$\sum_k y_{ab;ik} = y_{a;i} \sum_k y_{b;k} = 1$$

One reason why we can't solve this

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

$$\boxed{\sum_k y_{ab;ik} = y_{a;i}}$$

One reason why we can't solve this

Linear Programming Relaxation

$$\min \theta^T \mathbf{y}$$

$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

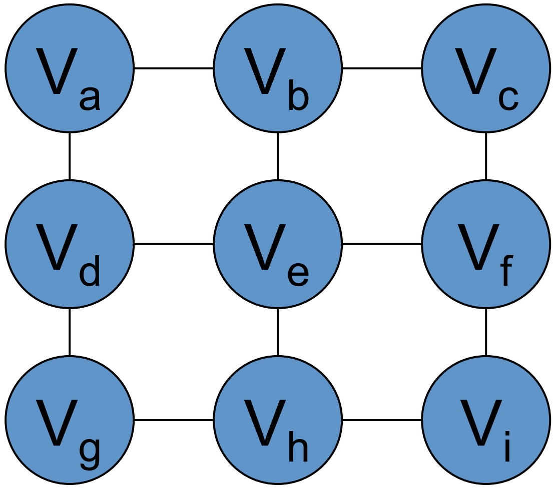
$$\sum_k y_{ab;ik} = y_{a;i}$$

No reason why we can't solve this*

* memory requirements, time complexity

Dual of the LP Relaxation

Wainwright et al., 2001



θ

$$\min \theta^T \mathbf{y}$$

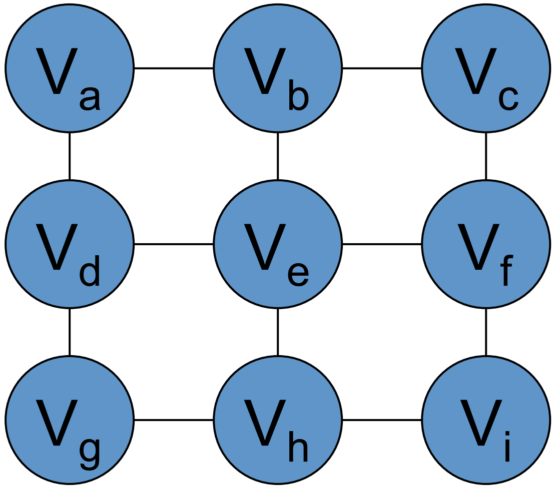
$$y_{a;i} \in [0,1]$$

$$\sum_i y_{a;i} = 1$$

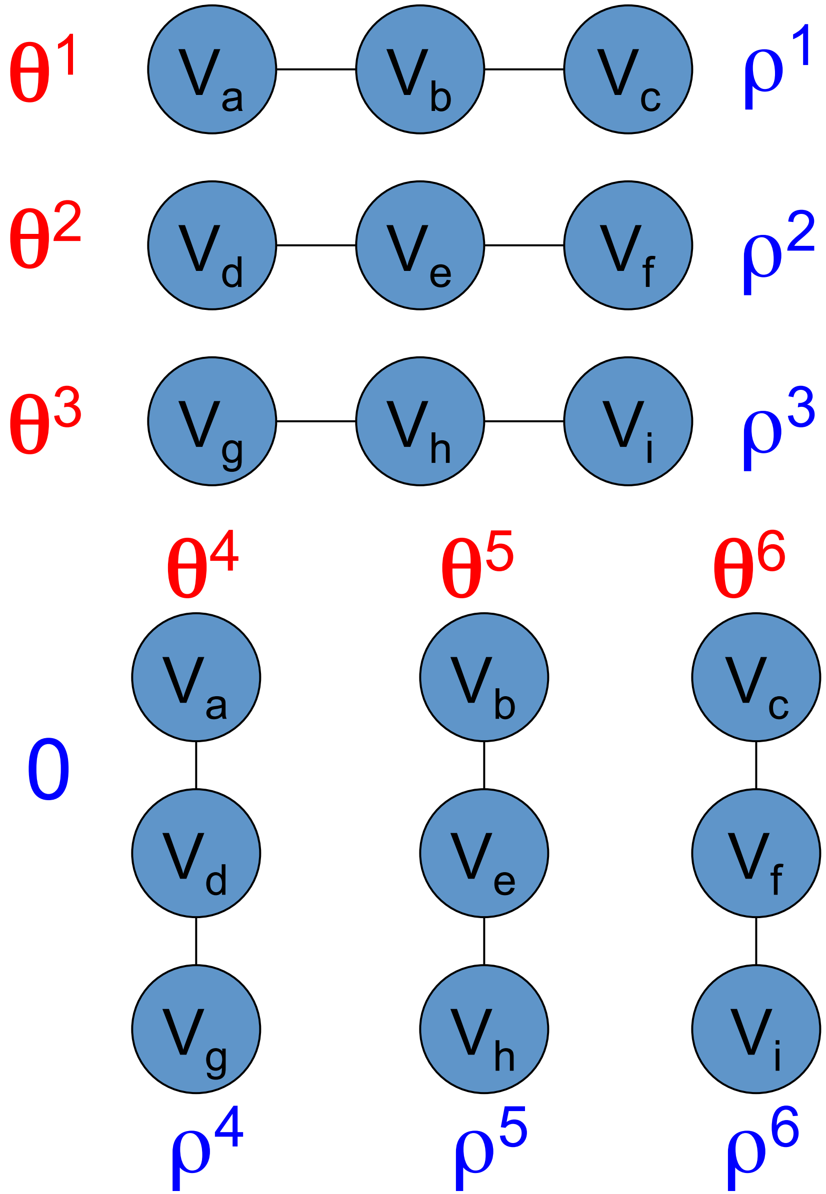
$$\sum_k y_{ab;ik} = y_{a;i}$$

Dual of the LP Relaxation

Wainwright et al., 2001



θ

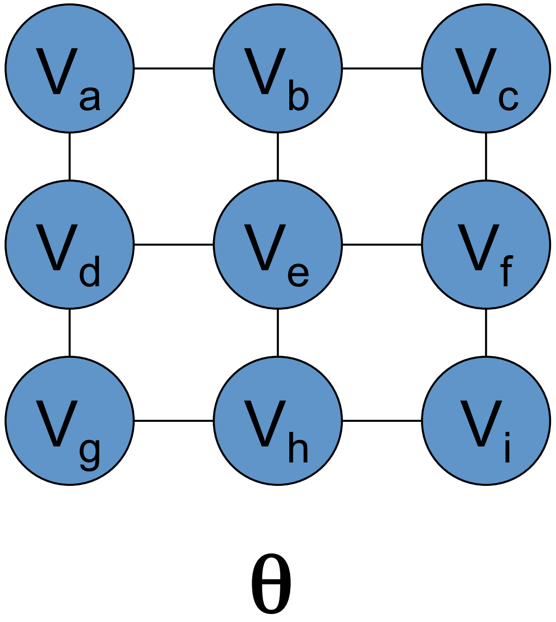


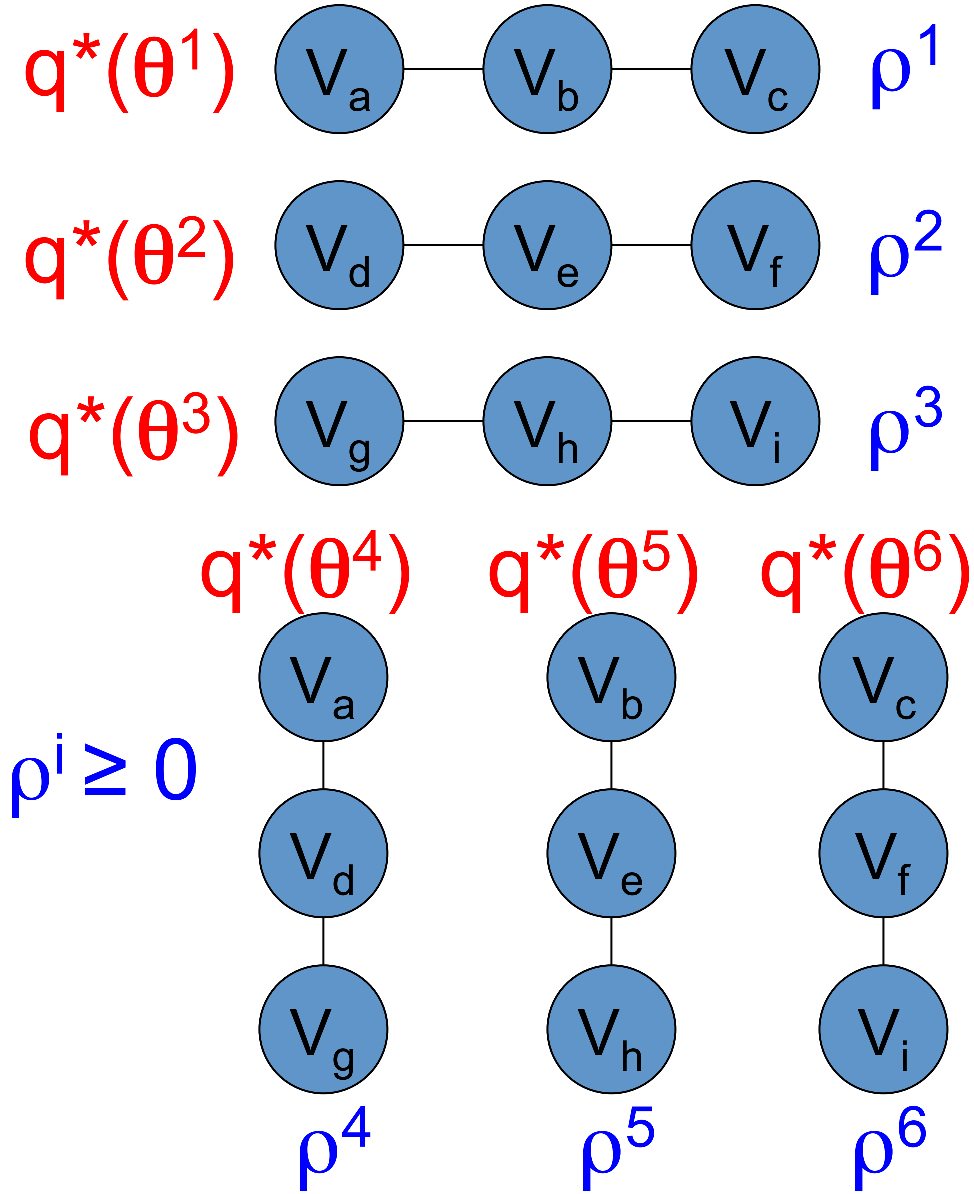
$\rho^i \geq 0$

$\sum \rho^i \theta^i = \theta$

Dual of the LP Relaxation

Wainwright et al., 2001





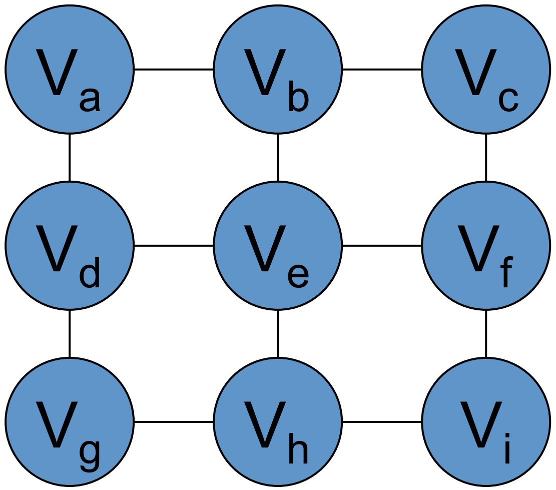
Dual of LP

$$\max \sum \rho^i q^*(\theta^i)$$

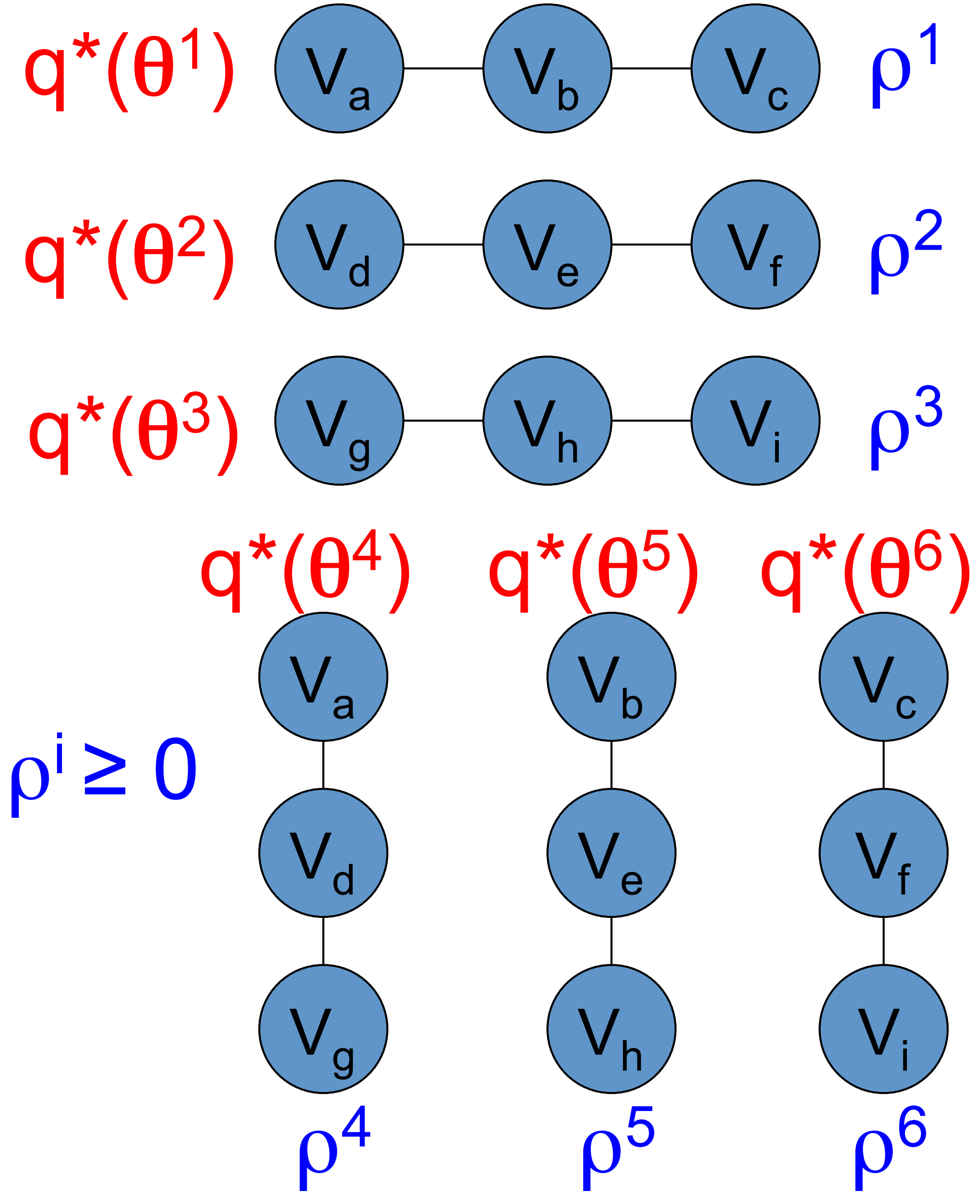
$$\sum \rho^i \theta^i = \theta$$

Dual of the LP Relaxation

Wainwright et al., 2001



θ



Dual of LP

$$\max \sum \rho^i q^*(\theta^i)$$

$$\sum \rho^i \theta^i \equiv \theta$$

Dual of the LP Relaxation

Wainwright et al., 2001

$$\max \sum \rho^i q^*(\theta^i)$$

$$\sum \rho^i \theta^i \equiv \theta$$

I can easily compute $q^*(\theta_i)$

I can easily maintain reparam constraint

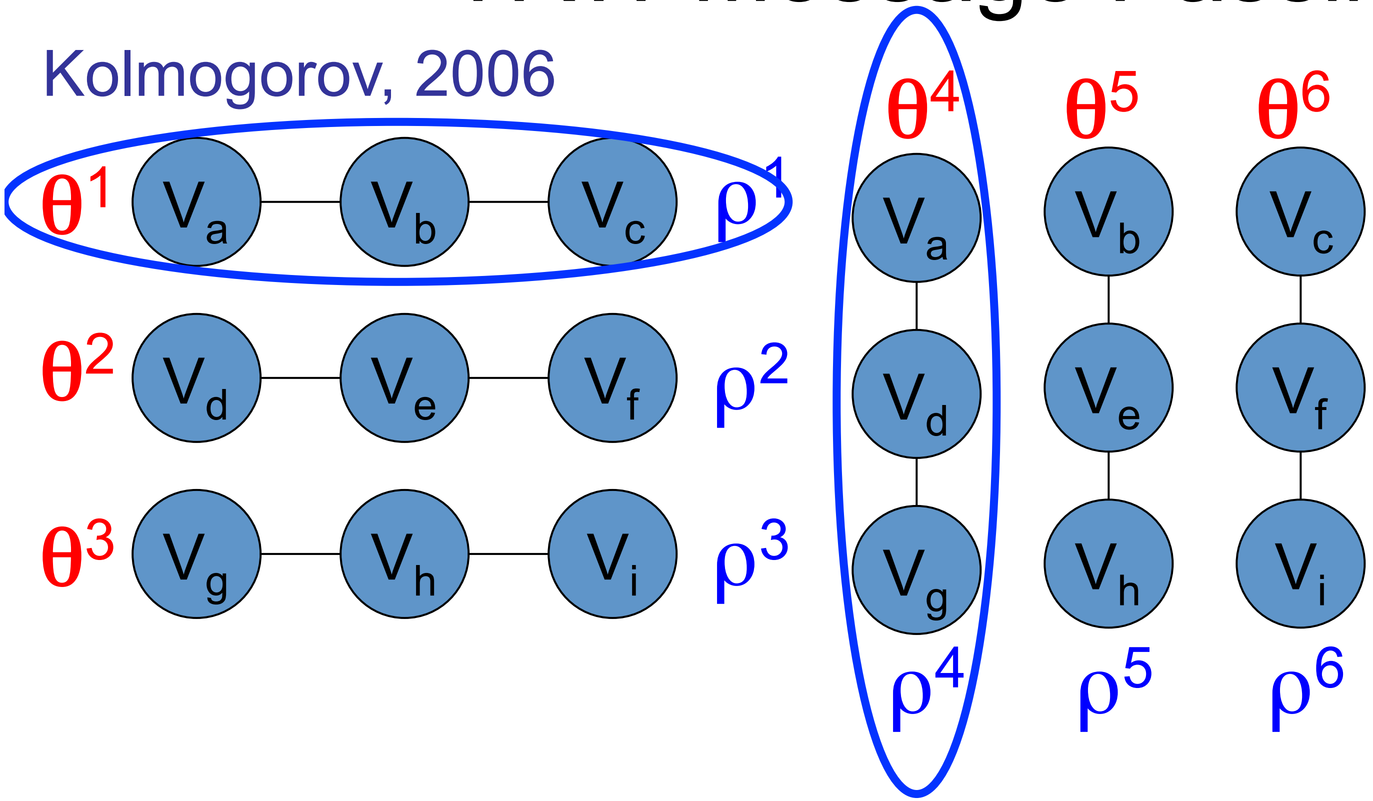
So can I easily solve the dual?

Outline

- Reparameterization (lecture 1)
- Belief Propagation (lecture 1)
- Tree-reweighted Message Passing
  - Integer Programming Formulation
  - Linear Programming Relaxation and its Dual
  - Convergent Solution for Dual
  - Computational Issues and Theoretical Properties

TRW Message Passing

Kolmogorov, 2006



Pick a variable

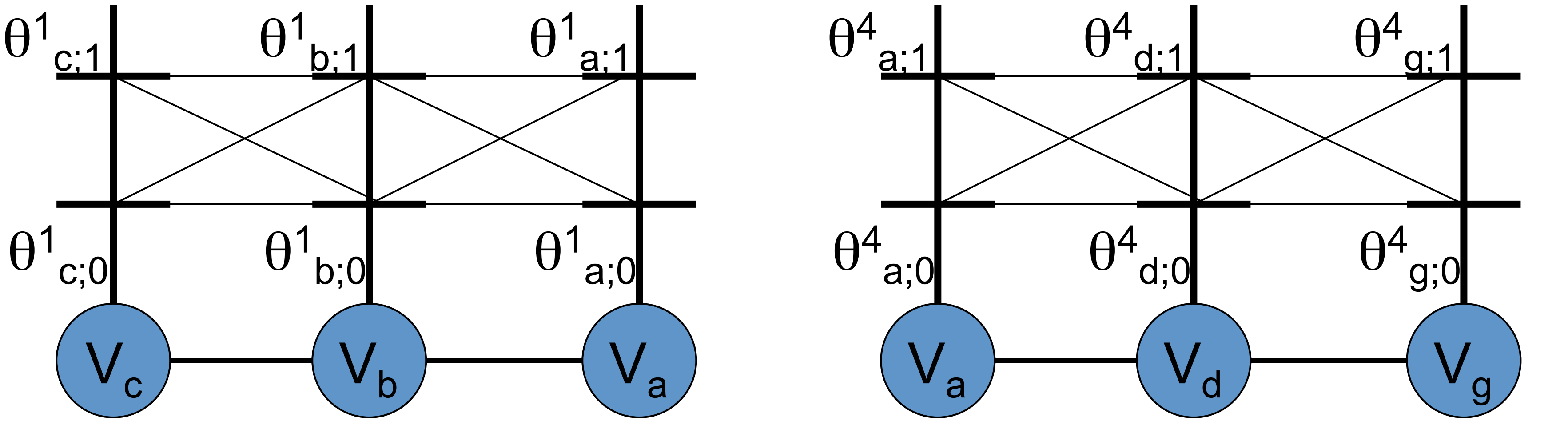
V_a

$$\sum \rho^i q^*(\theta^i)$$

$$\sum \rho^i \theta^i \equiv \theta$$

TRW Message Passing

Kolmogorov, 2006

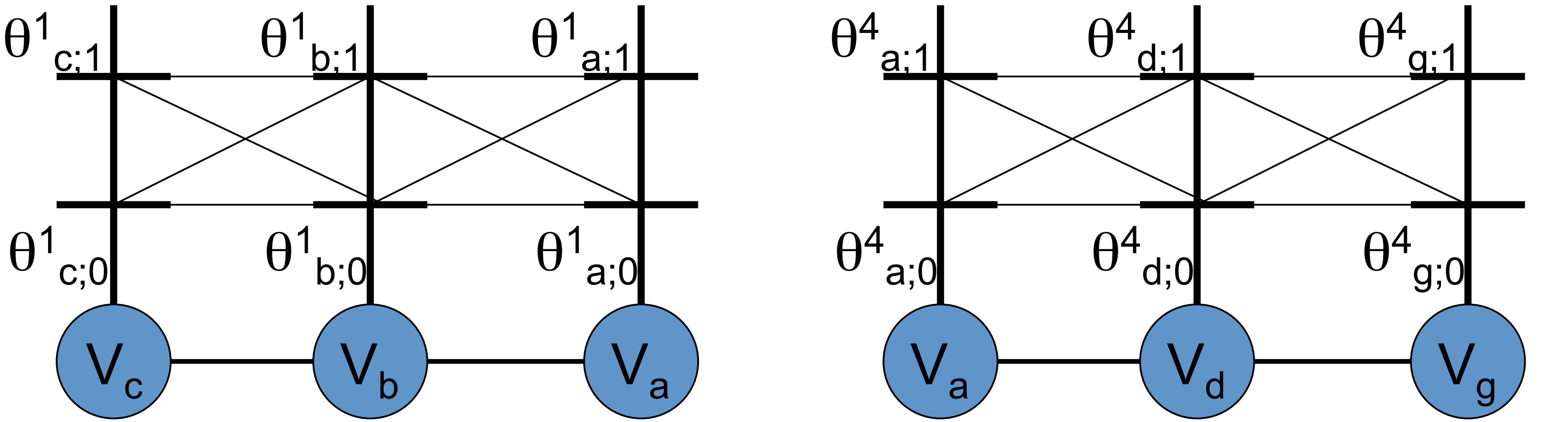


$$\sum \rho^i q^*(\boldsymbol{\theta}^i)$$

$$\sum \rho^i \boldsymbol{\theta}^i \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



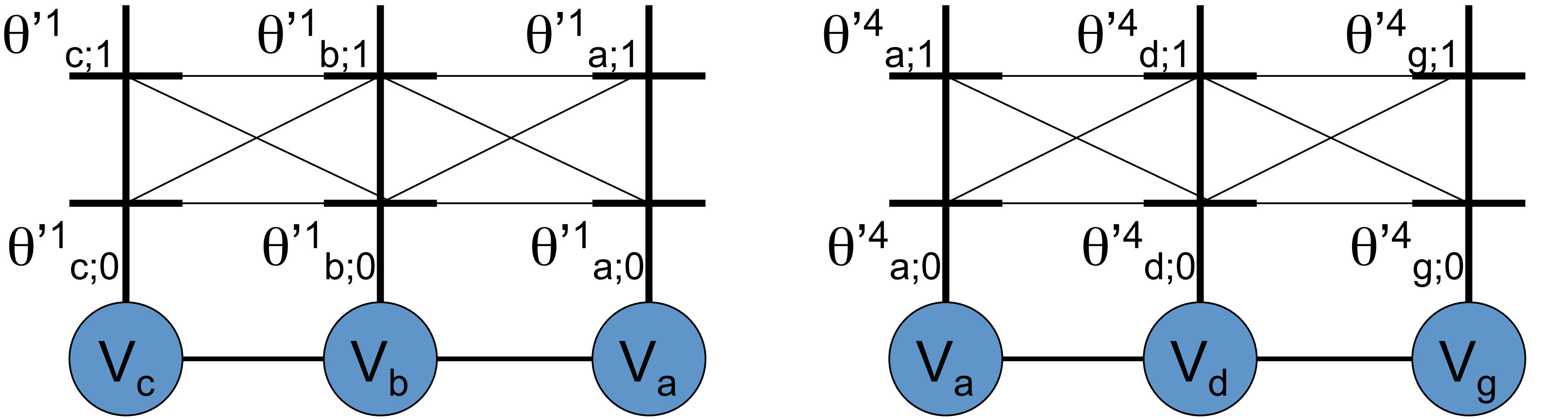
Reparameterize to obtain min-marginals of V_a

$$\rho^1 q^*(\boldsymbol{\theta}^1) + \rho^4 q^*(\boldsymbol{\theta}^4) + K$$

$$\rho^1\boldsymbol{\theta}^1 + \rho^4\boldsymbol{\theta}^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



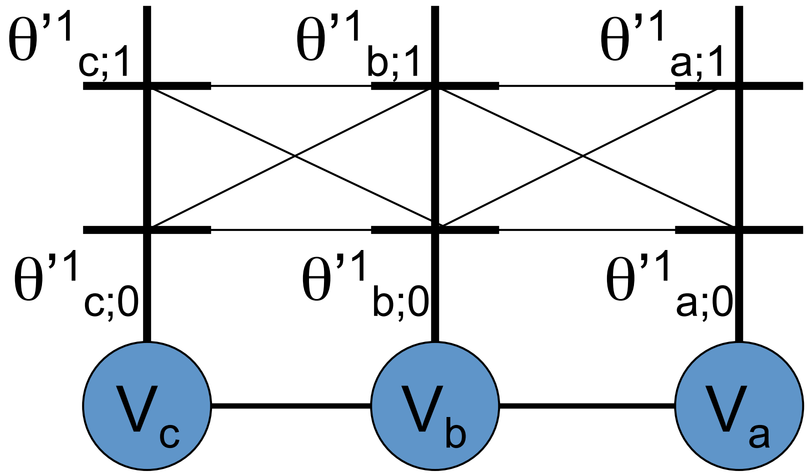
One pass of Belief Propagation

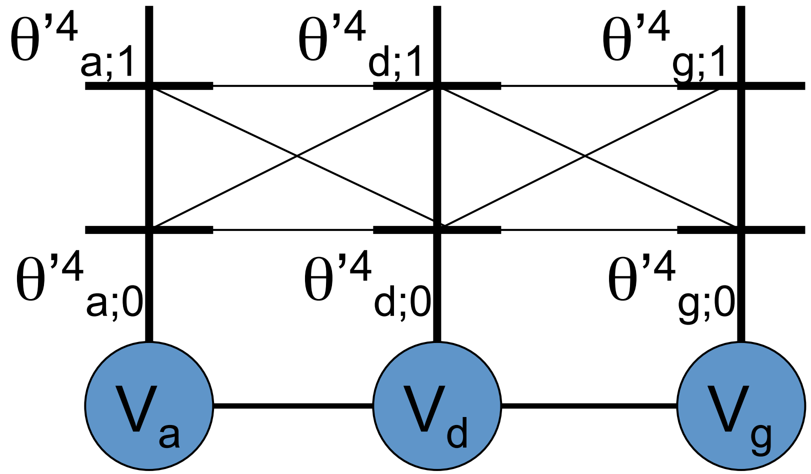
$$\rho^1 q^*(\theta'^1) + \rho^4 q^*(\theta'^4) + K$$

$$\rho^1\theta'^1 + \rho^4\theta'^4 + \theta^{rest}$$

TRW Message Passing

Kolmogorov, 2006





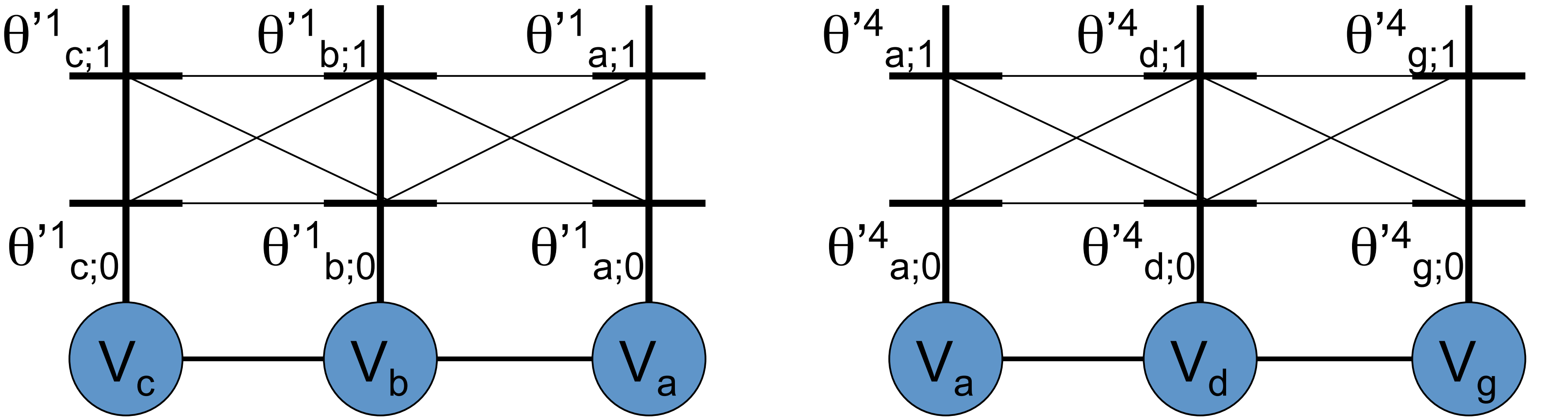
Remain the same

$$\rho^1 q^*(\theta'^1) + \rho^4 q^*(\theta'^4) + K$$

$$\rho^1\theta'^1 + \rho^4\theta'^4 + \theta^{rest} \equiv \theta$$

TRW Message Passing

Kolmogorov, 2006

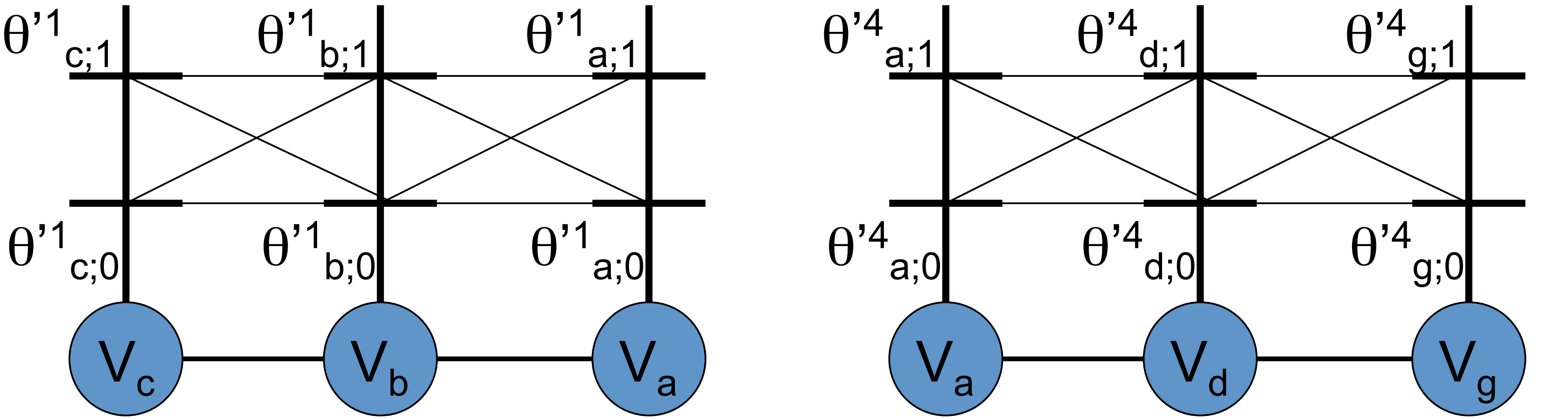


$$\rho^1 \min\{\theta'^1_{a;0}, \theta'^1_{a;1}\} + \rho^4 \min\{\theta'^4_{a;0}, \theta'^4_{a;1}\} + K$$

$$\rho^1 \boldsymbol{\theta}'^1 + \rho^4 \boldsymbol{\theta}'^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



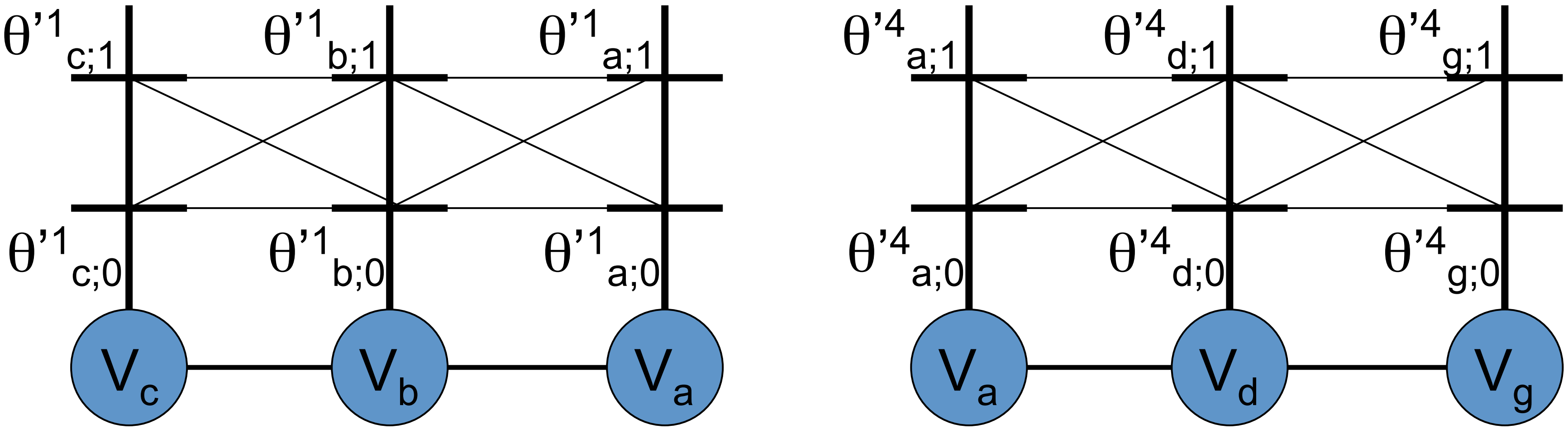
Compute weighted average of min-marginals of V_a

$$\rho^1 \min\{\theta'^1_{a;0}, \theta'^1_{a;1}\} + \rho^4 \min\{\theta'^4_{a;0}, \theta'^4_{a;1}\} + K$$

$$\rho^1 \boldsymbol{\theta}'^1 + \rho^4 \boldsymbol{\theta}'^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



$$\theta''_{a;0} = \frac{\rho^1\theta'^1_{a;0} + \rho^4\theta'^4_{a;0}}{\rho^1 + \rho^4}$$

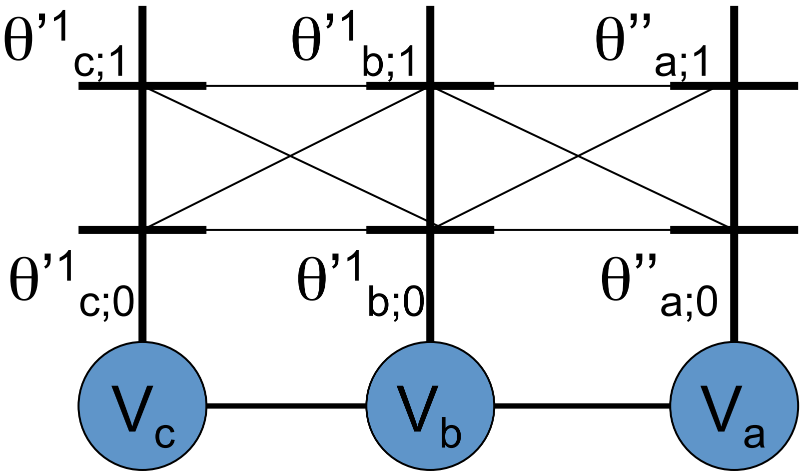
$$\theta''_{a;1} = \frac{\rho^1\theta'^1_{a;1} + \rho^4\theta'^4_{a;1}}{\rho^1 + \rho^4}$$

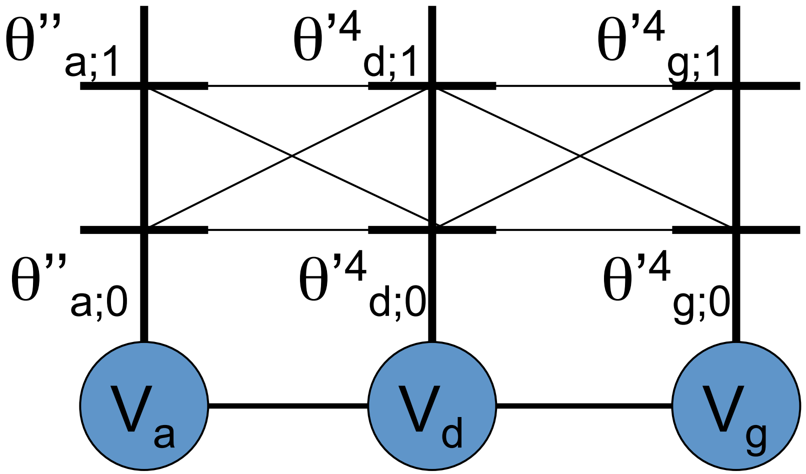
$$\rho^1 \min\{\theta'^1_{a;0},\theta'^1_{a;1}\} + \rho^4 \min\{\theta'^4_{a;0},\theta'^4_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}'^1 + \rho^4\boldsymbol{\theta}'^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006





$$\theta''_{a;0} = \frac{\rho^1\theta'^1_{a;0} + \rho^4\theta'^4_{a;0}}{\rho^1 + \rho^4}$$

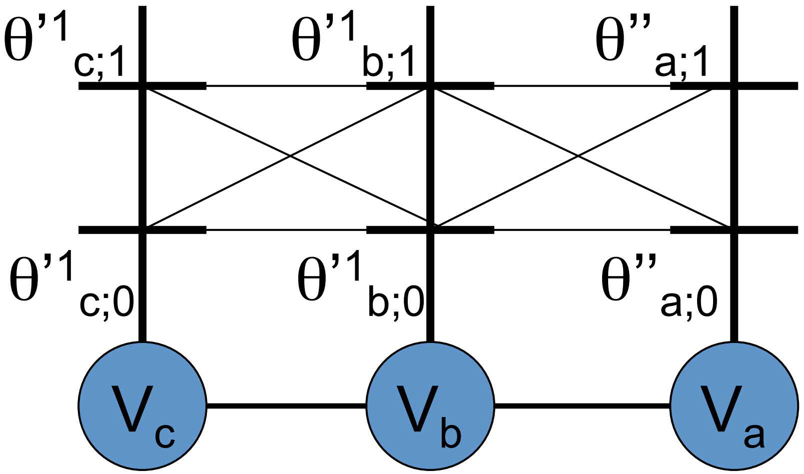
$$\theta''_{a;1} = \frac{\rho^1\theta'^1_{a;1} + \rho^4\theta'^4_{a;1}}{\rho^1 + \rho^4}$$

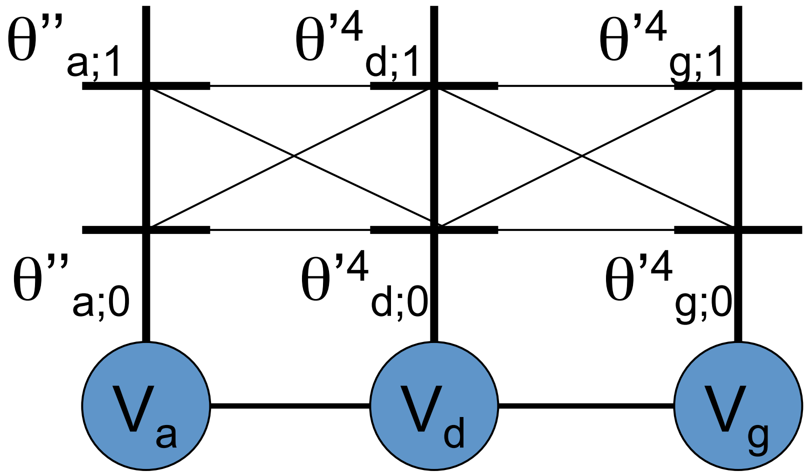
$$\rho^1 \min\{\theta'^1_{a;0}, \theta'^1_{a;1}\} + \rho^4 \min\{\theta'^4_{a;0}, \theta'^4_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest}$$

TRW Message Passing

Kolmogorov, 2006





$$\theta''_{a;0} = \frac{\rho^1\theta'^1_{a;0} + \rho^4\theta'^4_{a;0}}{\rho^1 + \rho^4}$$

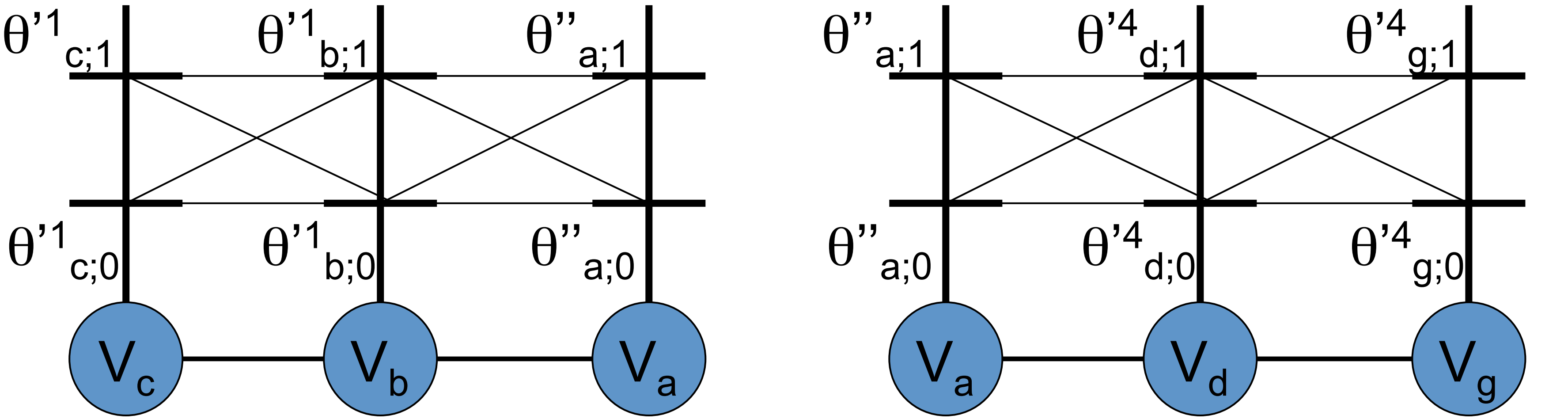
$$\theta''_{a;1} = \frac{\rho^1\theta'^1_{a;1} + \rho^4\theta'^4_{a;1}}{\rho^1 + \rho^4}$$

$$\rho^1 \min\{\theta'^1_{a;0}, \theta'^1_{a;1}\} + \rho^4 \min\{\theta'^4_{a;0}, \theta'^4_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



$$\theta''_{a;0} = \frac{\rho^1\theta'^1_{a;0} + \rho^4\theta'^4_{a;0}}{\rho^1 + \rho^4}$$

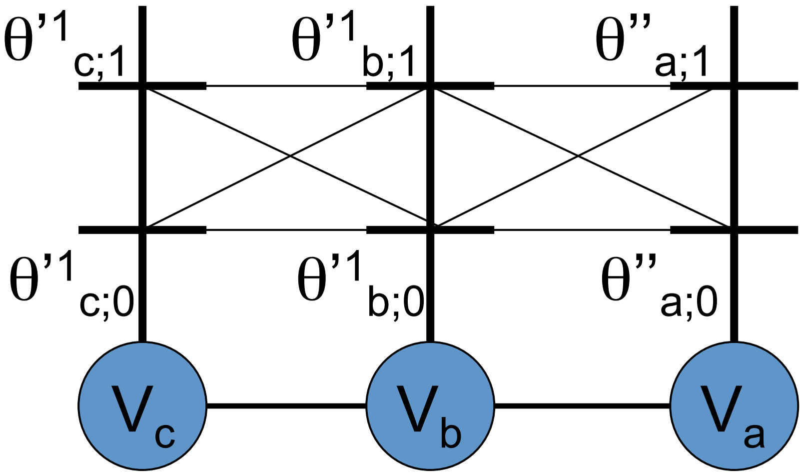
$$\theta''_{a;1} = \frac{\rho^1\theta'^1_{a;1} + \rho^4\theta'^4_{a;1}}{\rho^1 + \rho^4}$$

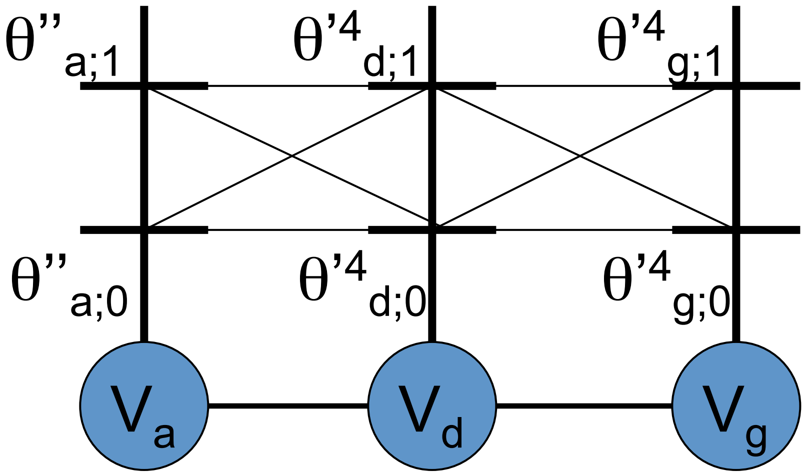
$$\rho^1 \min\{\theta''_{a;0}, \theta''_{a;1}\} + \rho^4 \min\{\theta''_{a;0}, \theta''_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006





$$\theta''_{a;0} = \frac{\rho^1\theta'^1_{a;0} + \rho^4\theta'^4_{a;0}}{\rho^1 + \rho^4}$$

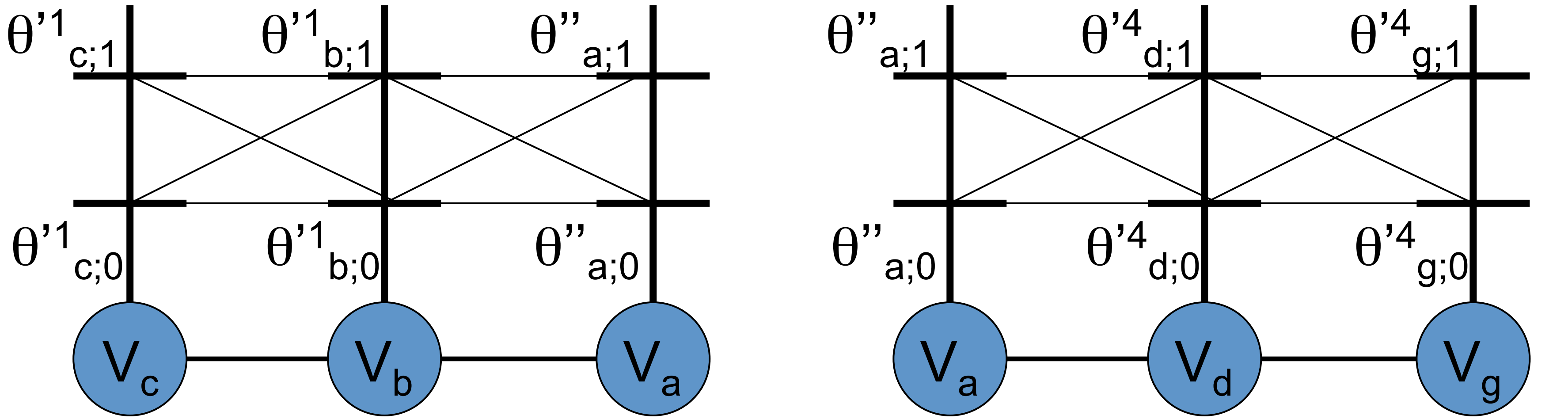
$$\theta''_{a;1} = \frac{\rho^1\theta'^1_{a;1} + \rho^4\theta'^4_{a;1}}{\rho^1 + \rho^4}$$

$$(\rho^1 + \rho^4)\min\{\theta''_{a;0}, \theta''_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



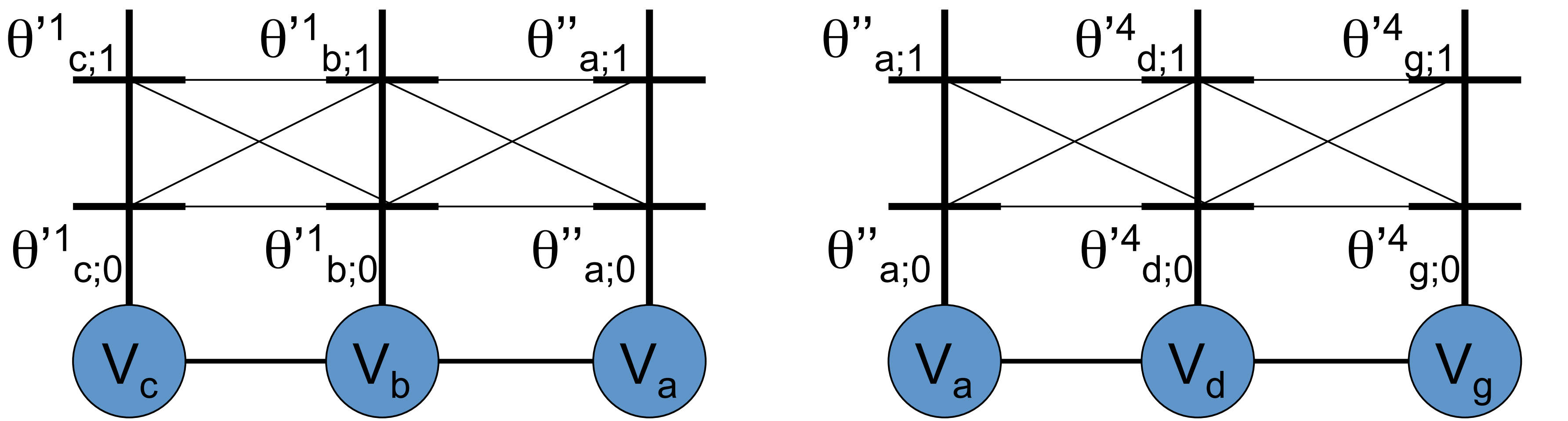
$$\min\{p_1+p_2,\ q_1+q_2\} \geq \min\{p_1,\ q_1\} + \min\{p_2,\ q_2\}$$

$$(\rho^1 + \rho^4) \min\{\theta''_{a;0},\ \theta''_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Kolmogorov, 2006



Objective function increases or remains constant

$$(\rho^1 + \rho^4) \min\{\theta''_{a;0}, \theta''_{a;1}\} + K$$

$$\rho^1\boldsymbol{\theta}''^1 + \rho^4\boldsymbol{\theta}''^4 + \boldsymbol{\theta}^{rest} \equiv \boldsymbol{\theta}$$

TRW Message Passing

Initialize θ^i. Take care of reparam constraint

Choose random variable V_a

Compute min-marginals of V_a for all trees

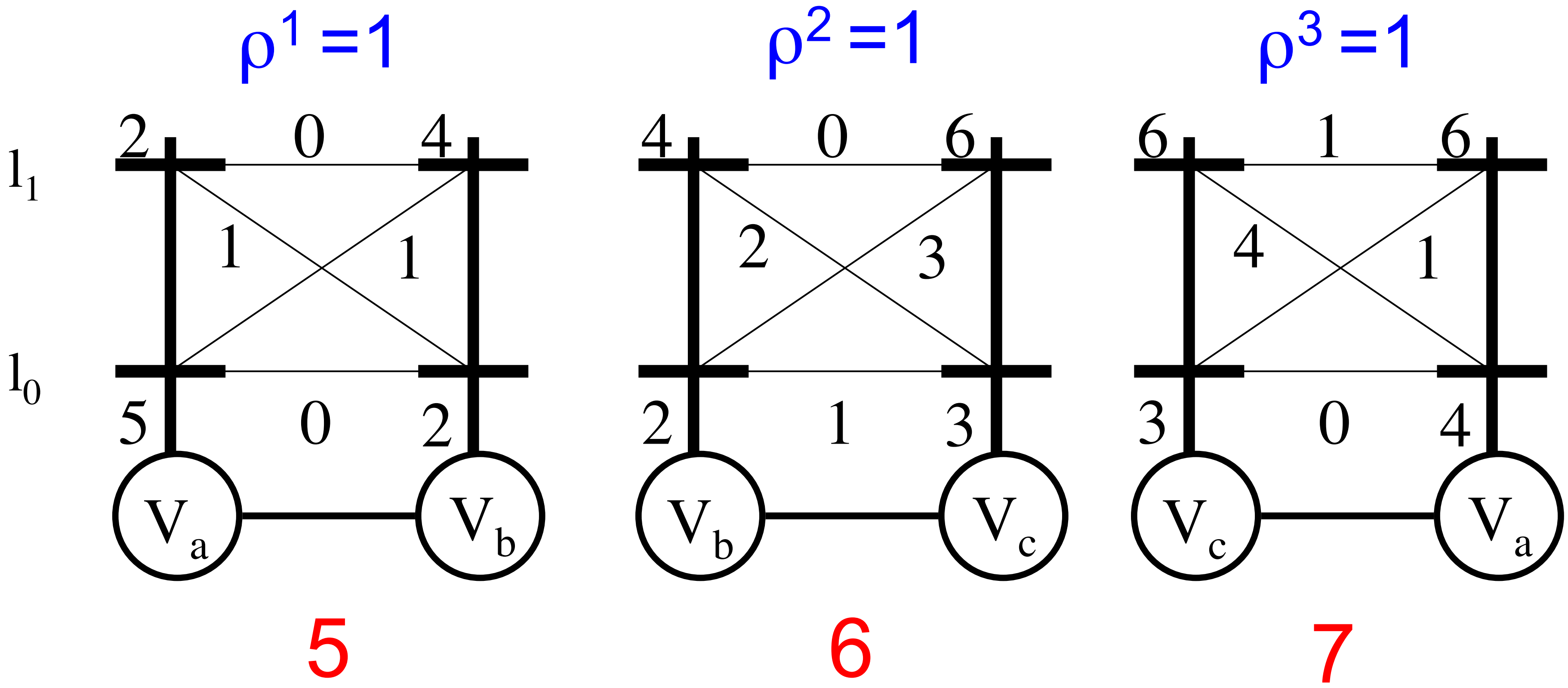
Node-average the min-marginals

REPEAT

Can also do edge-averaging

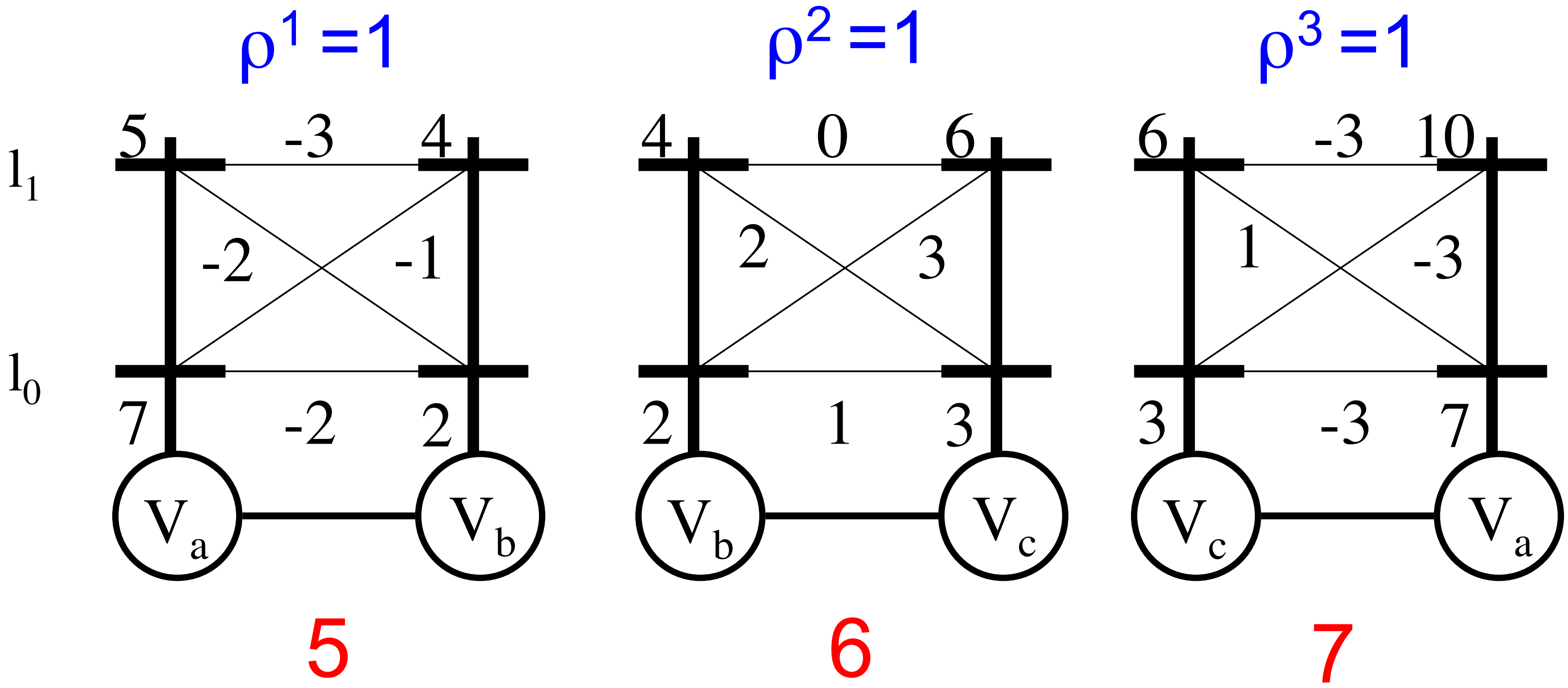
Kolmogorov, 2006

Example 1



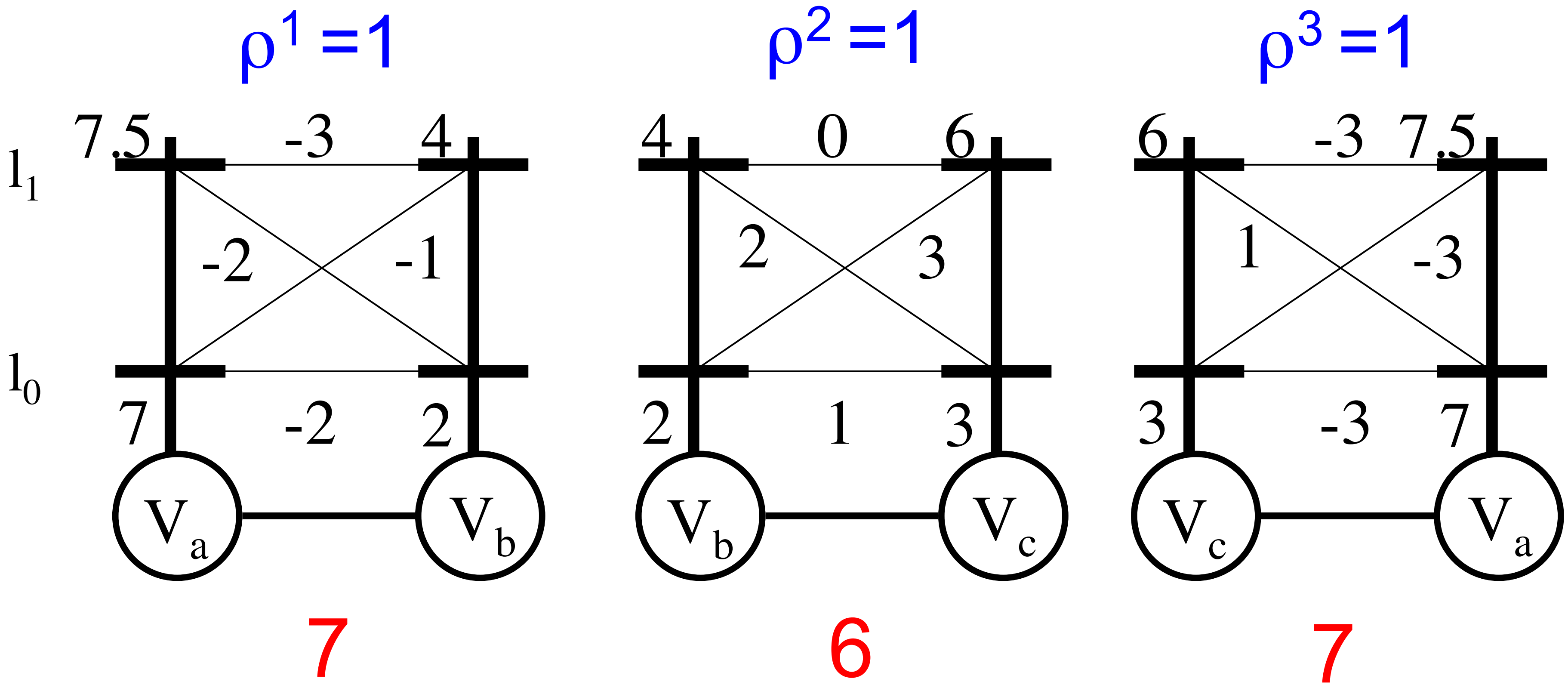
Pick variable V_a. Reparameterize.

Example 1



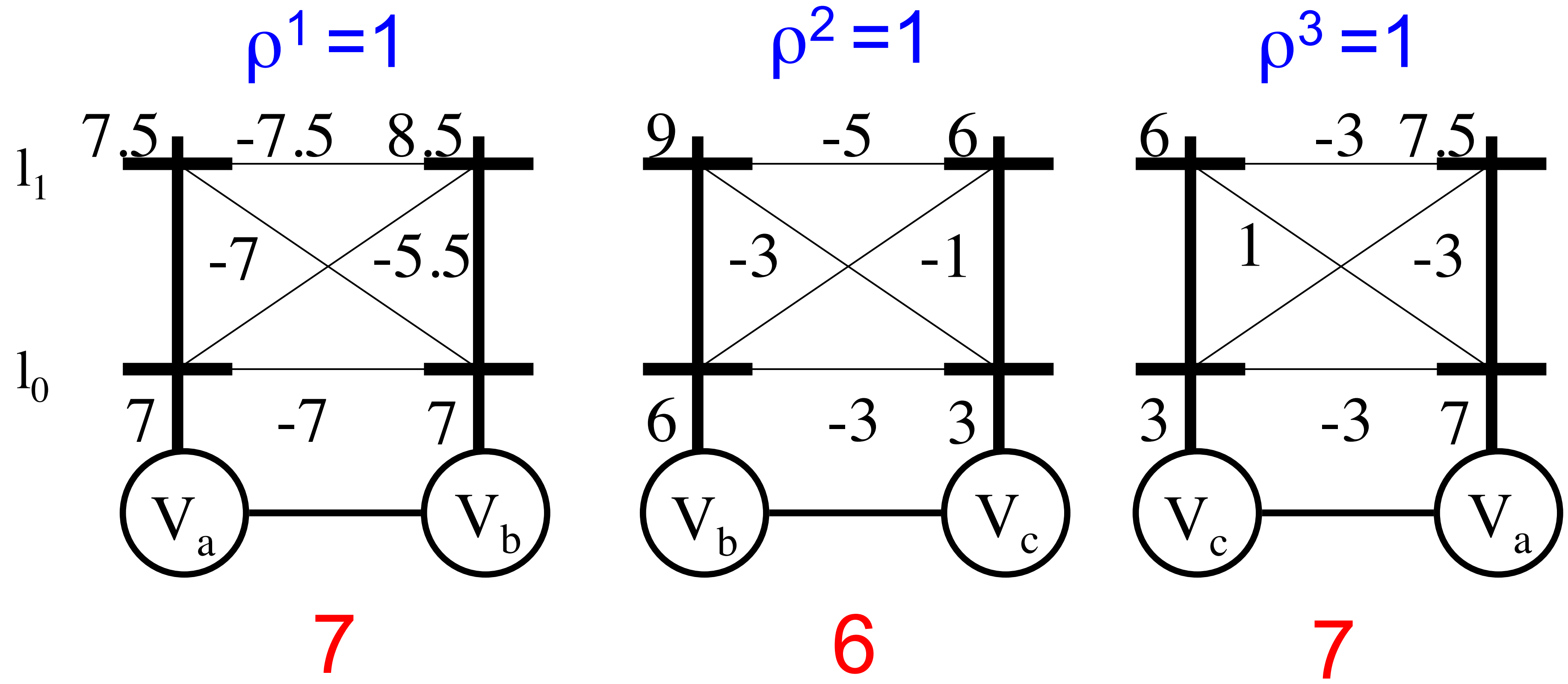
Average the min-marginals of V_a

Example 1



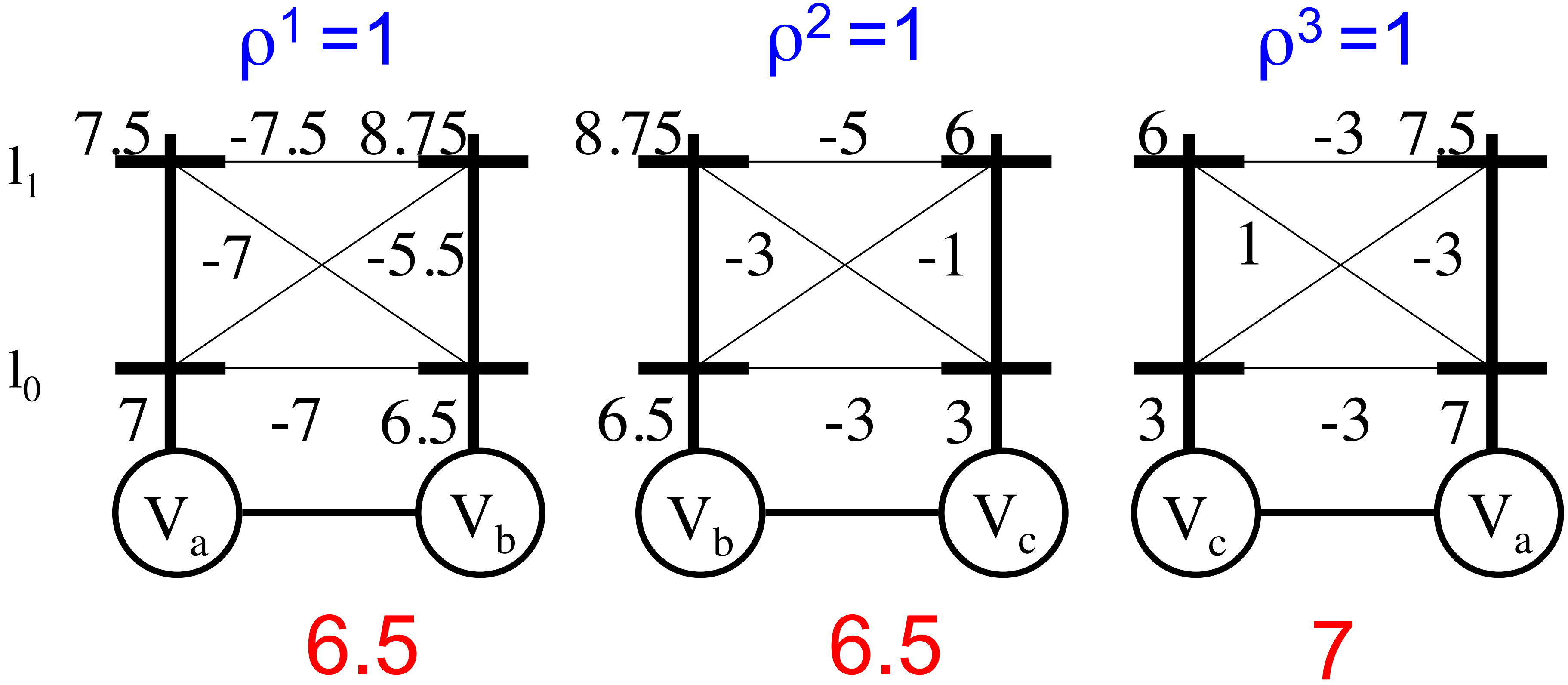
Pick variable V_b. Reparameterize.

Example 1



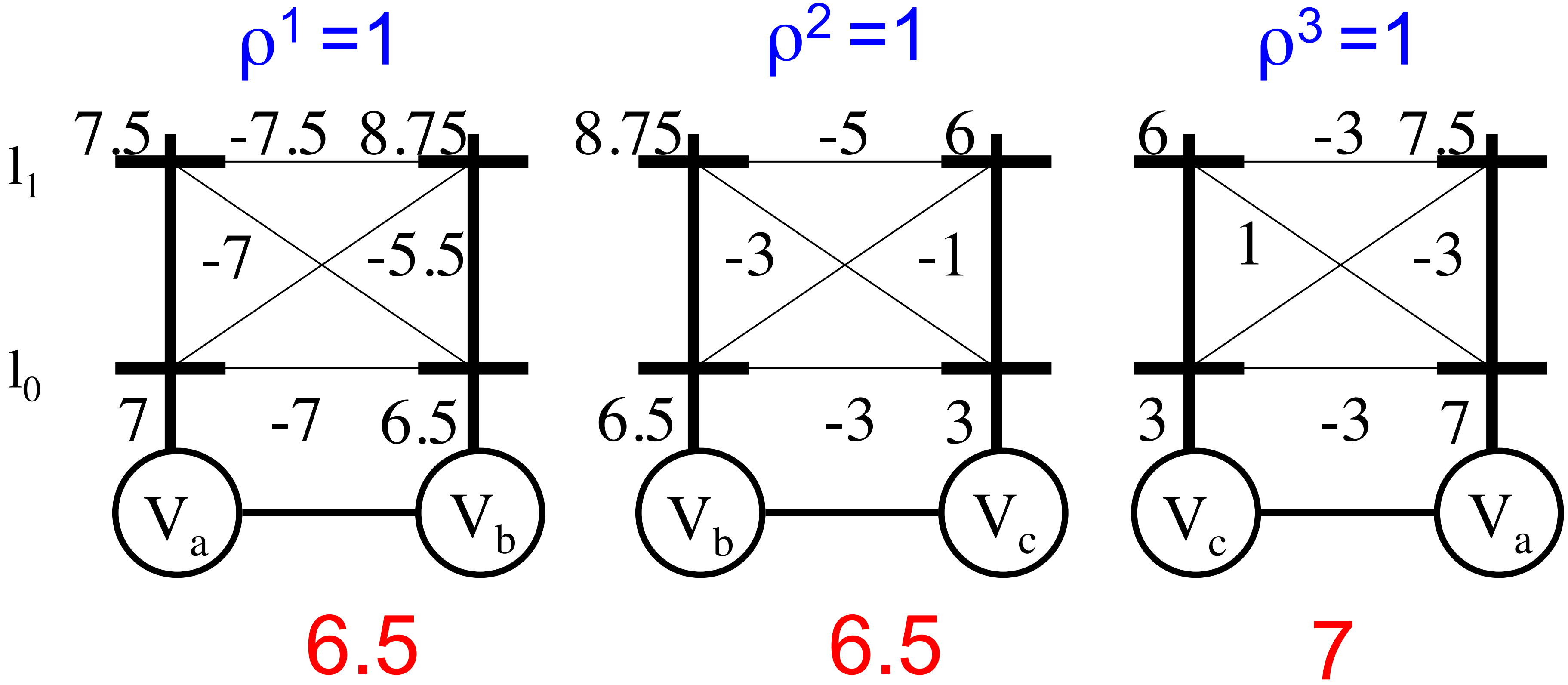
Average the min-marginals of V_b

Example 1



Value of dual does not increase

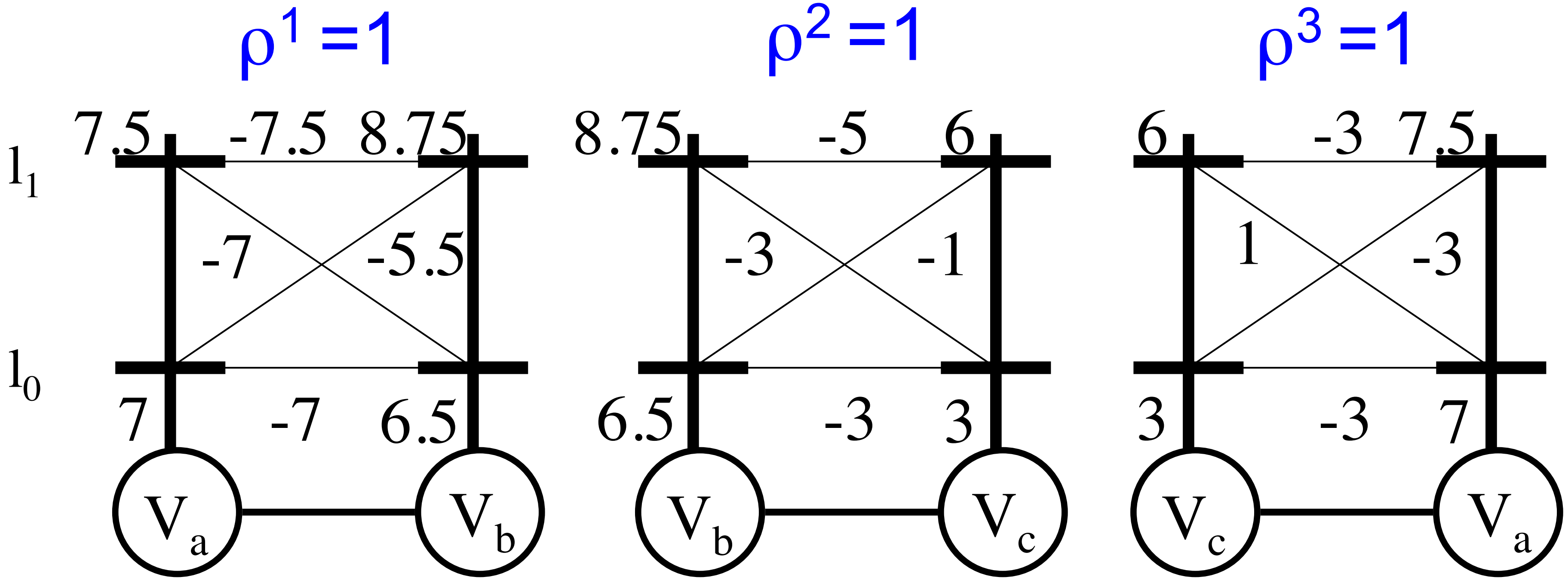
Example 1



Maybe it will increase for V_c

NO

Example 1

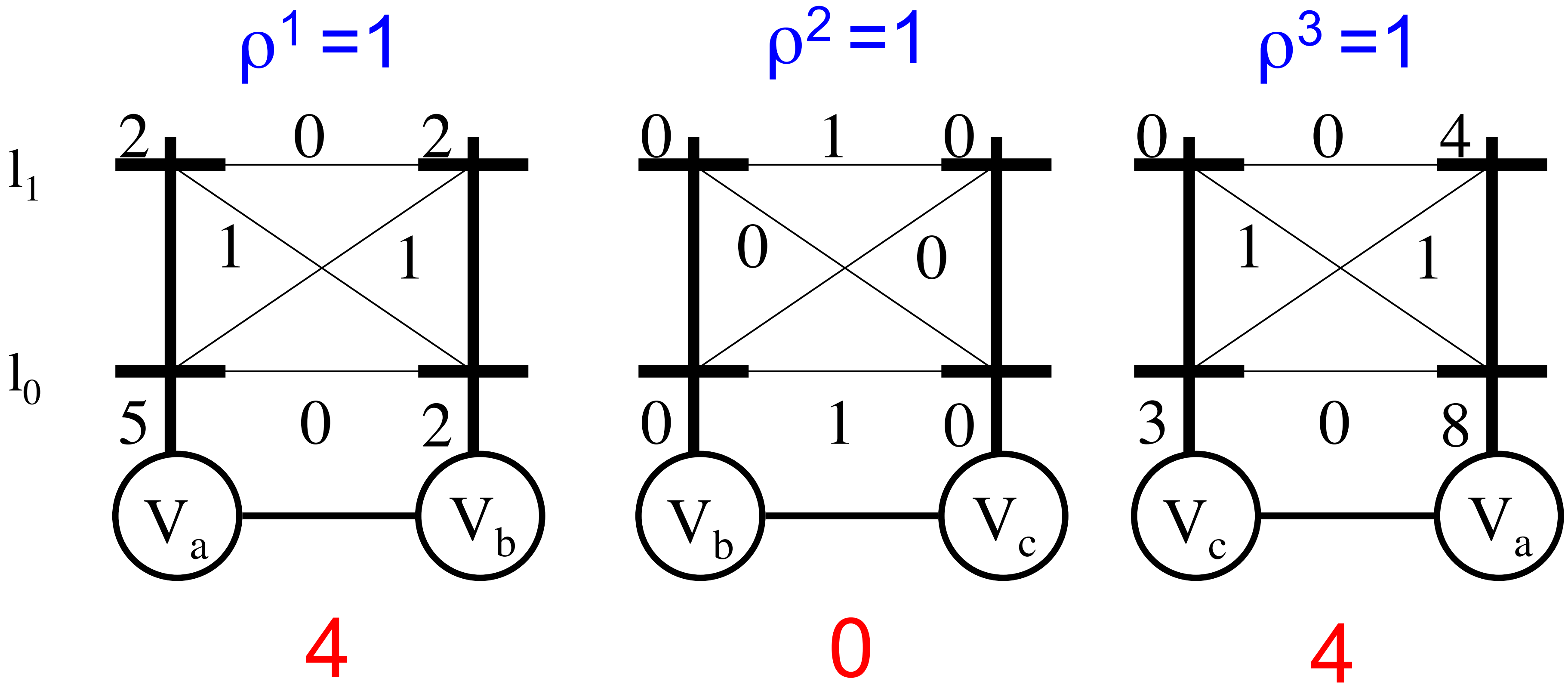


$f^1(a) = 0$ $f^1(b) = 0$ $f^2(b) = 0$ $f^2(c) = 0$ $f^3(c) = 0$ $f^3(a) = 0$

Strong Tree Agreement

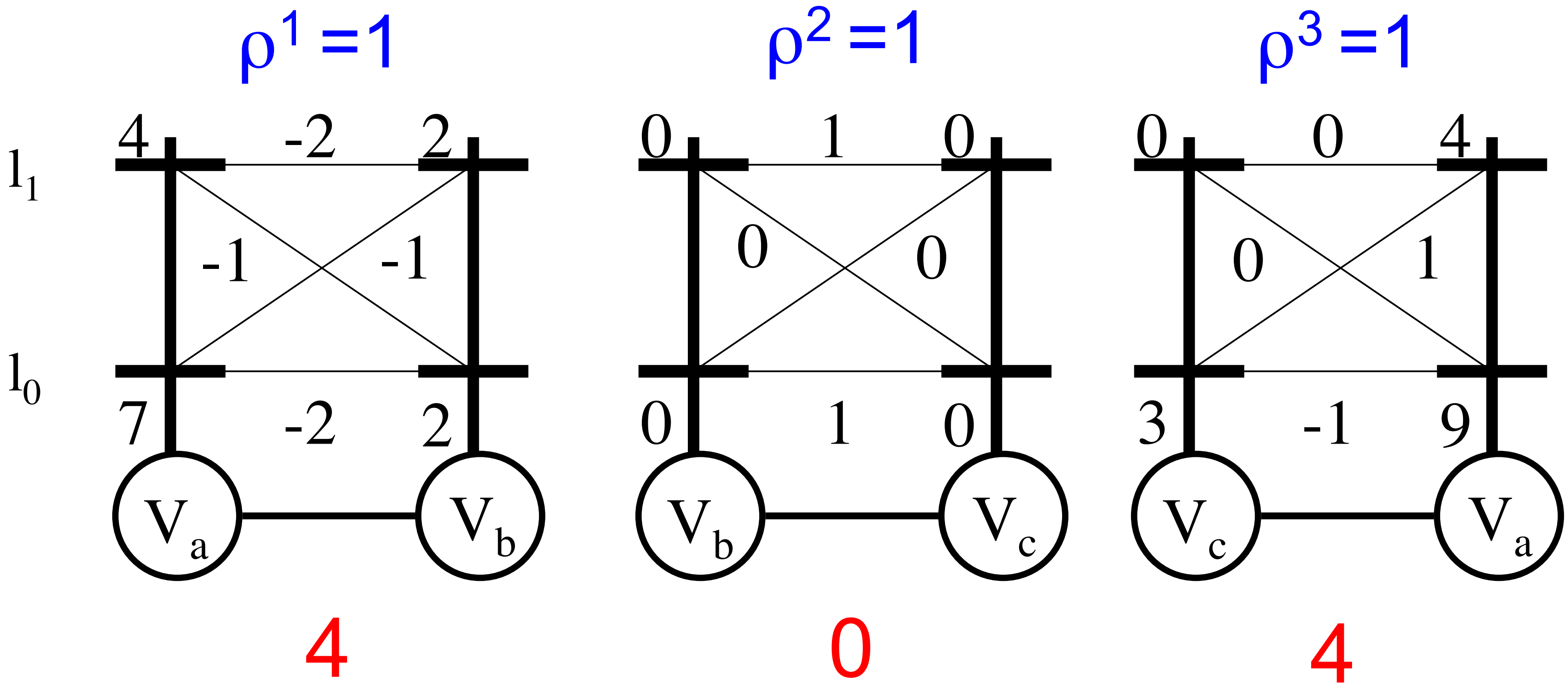
Exact MAP Estimate

Example 2



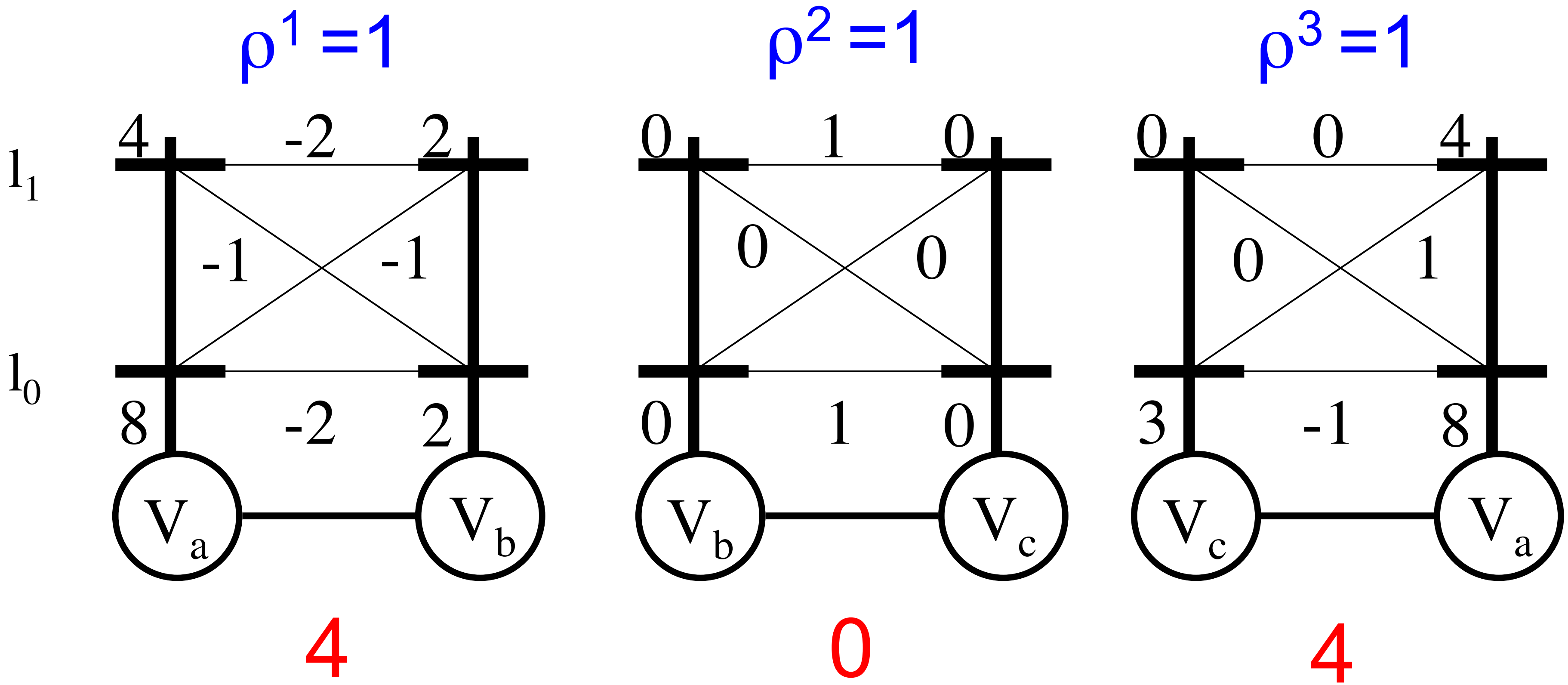
Pick variable V_a. Reparameterize.

Example 2



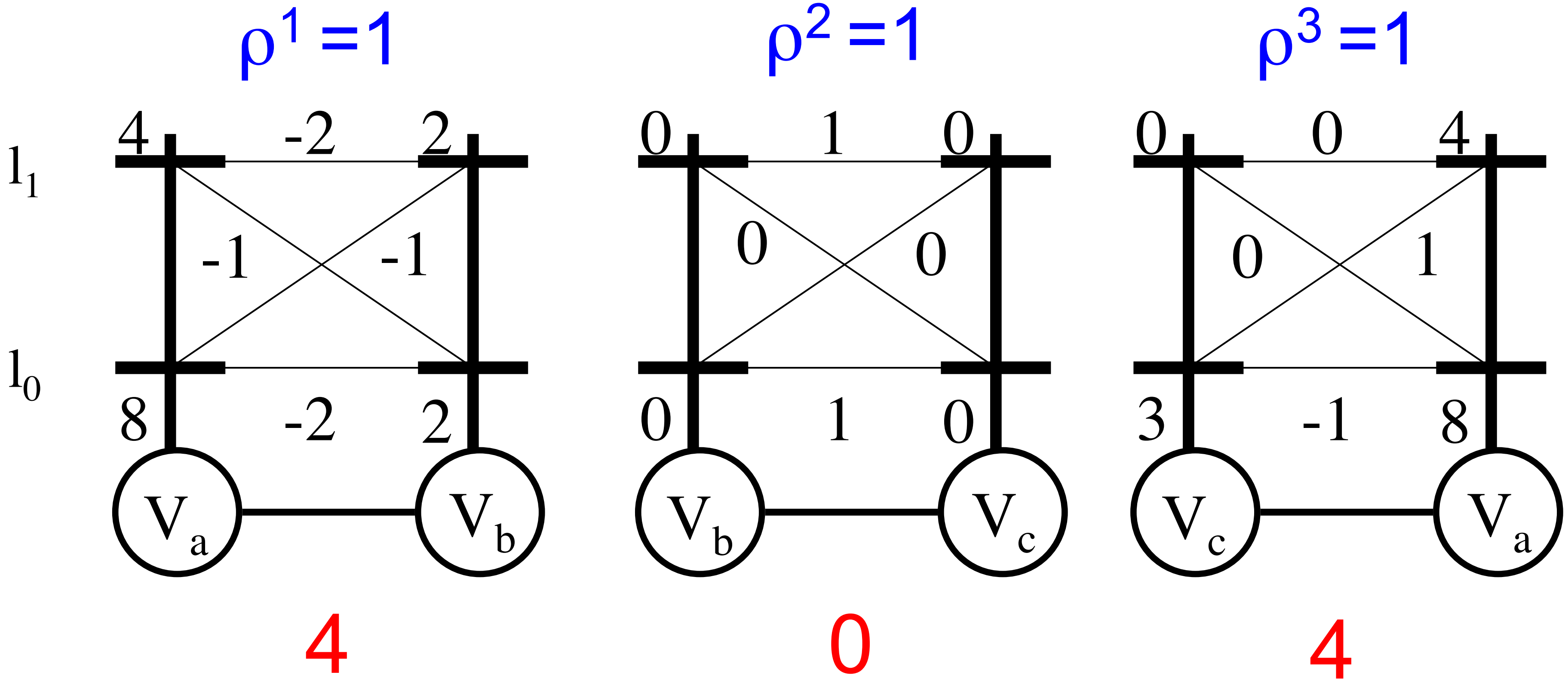
Average the min-marginals of V_a

Example 2



Value of dual does not increase

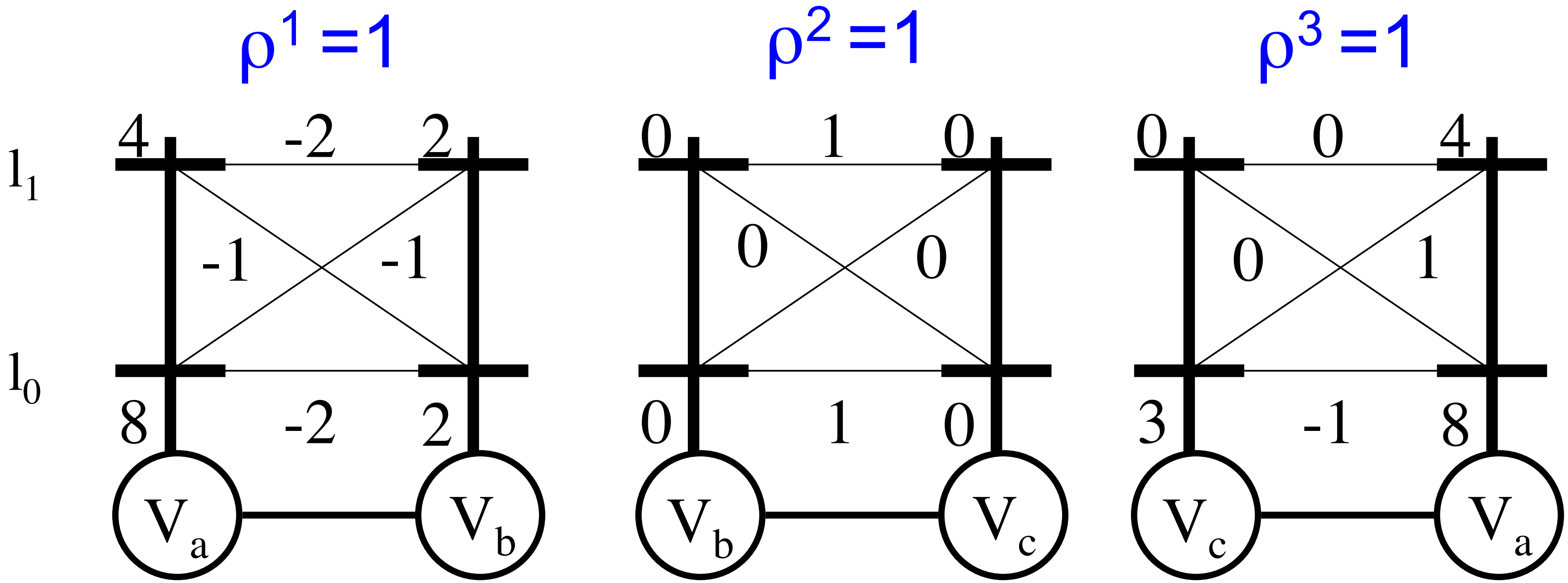
Example 2



Maybe it will decrease for V_b or V_c

NO

Example 2



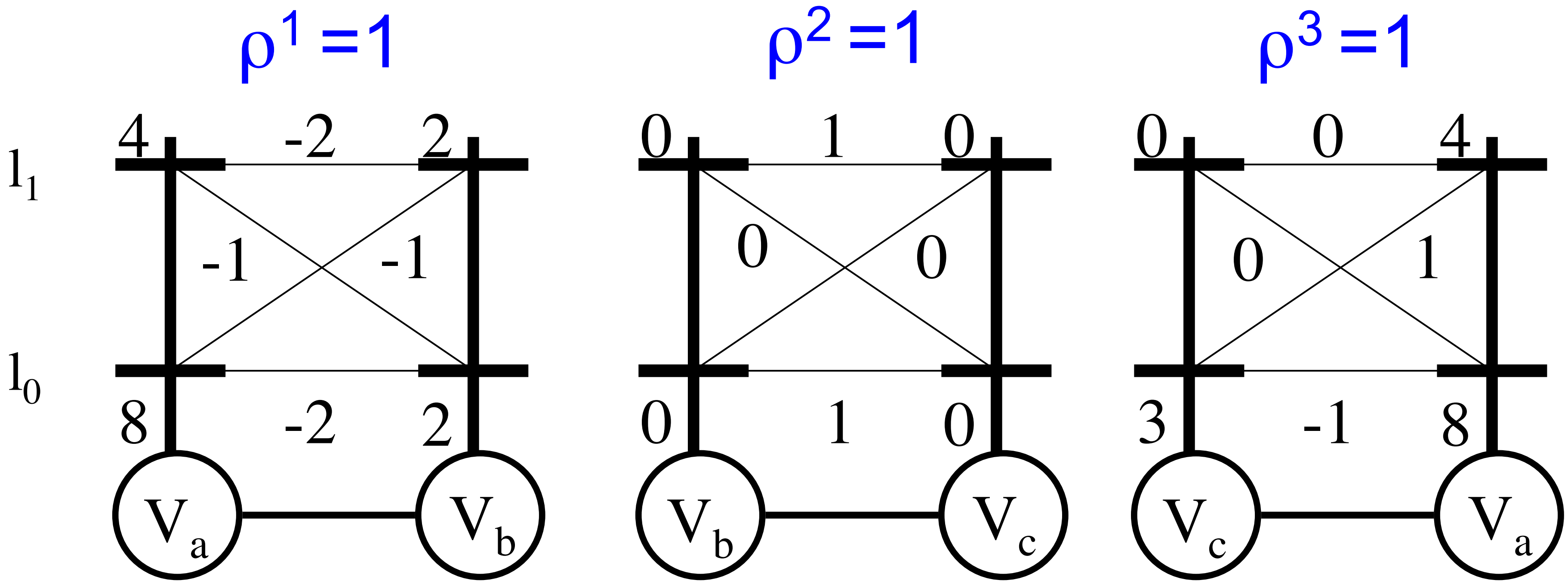
$f^1(a) = 1$ $f^1(b) = 1$ $f^2(b) = 1$ $f^2(c) = 0$ $f^3(c) = 1$ $f^3(a) = 1$

$f^2(b) = 0$ $f^2(c) = 1$

Weak Tree Agreement

Not Exact MAP Estimate

Example 2



$f^1(a) = 1$ $f^1(b) = 1$ $f^2(b) = 1$ $f^2(c) = 0$ $f^3(c) = 1$ $f^3(a) = 1$

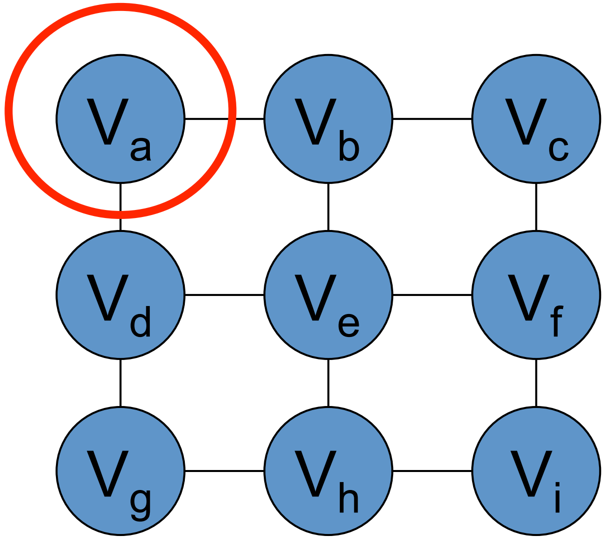
$f^2(b) = 0$ $f^2(c) = 1$

Weak Tree Agreement

Convergence point of TRW

Obtaining the Labeling

Only solves the dual. Primal solutions?

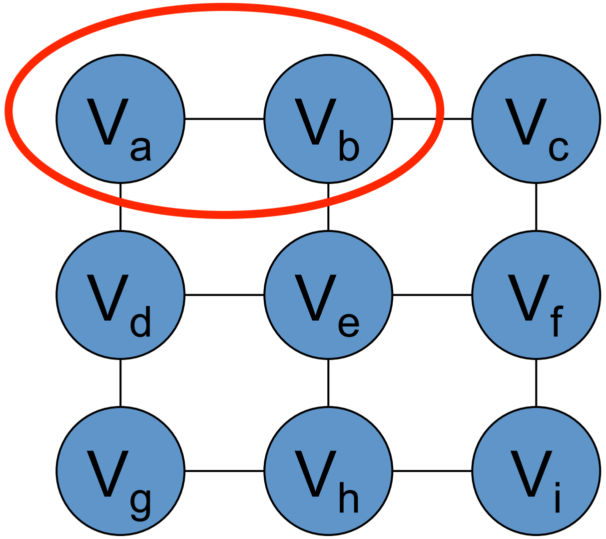


Fix the label of V_a

$$\theta' = \sum \rho^i \theta^i \equiv \theta$$

Obtaining the Labeling

Only solves the dual. Primal solutions?



Fix the label of V_b

$$\theta' = \sum \rho^i \theta^i \equiv \theta$$

Continue in some fixed order

Meltzer et al., 2006

Outline

- Reparameterization (lecture 1)
- Belief Propagation (lecture 1)
- Tree-reweighted Message Passing
  - Integer Programming Formulation
  - Linear Programming Relaxation and its Dual
  - Convergent Solution for Dual
  - Computational Issues and Theoretical Properties

Computational Issues of TRW

Basic Component is Belief Propagation

- Speed-ups for some pairwise potentials

  Felzenszwalb & Huttenlocher, 2004

- Memory requirements cut down by half

  Kolmogorov, 2006

- Further speed-ups using monotonic chains

  Kolmogorov, 2006

Theoretical Properties of TRW

- Always converges, unlike BP

  Kolmogorov, 2006

- Strong tree agreement implies exact MAP

  Wainwright et al., 2001

- Optimal MAP for two-label submodular problems

$$\theta_{ab;00} + \theta_{ab;11} \leq \theta_{ab;01} + \theta_{ab;10}$$

Kolmogorov and Wainwright, 2005

Summary

- Trees can be solved exactly - BP
- No guarantee of convergence otherwise - BP
- Strong Tree Agreement - TRW-S
- Submodular energies solved exactly - TRW-S
- TRW-S solves an LP relaxation of MAP estimation