Universality of the Local Marginal Polytope

Daniel Průša

(joint work with Tomáš Werner)

Center for Machine Perception
Czech Technical University
Prague, Czech Republic

The 35th Pattern Recognition and Computer Vision Colloquium
October 16, 2014

Prague

Overview

1. Introduction to min-sum problem, its usage in computer vision.

Overview

1. Introduction to min-sum problem, its usage in computer vision.
2. Linear programming (LP) relaxation of the problem.

Overview

1. Introduction to min-sum problem, its usage in computer vision.
2. Linear programming (LP) relaxation of the problem.
3. How hard is to solve the LP relaxation, what are fundamental limitations.

Min-sum problem
(a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V} \Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v) \Big].$$

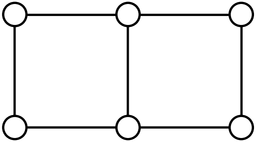
All weights $f_u(k), f_{uv}(k,\ell) \in \mathbb{R} \cup \{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.

Min-sum problem (a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V} \Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v) \Big].$$

All weights $f_u(k), f_{uv}(k,\ell) \in \mathbb{R} \cup \{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.

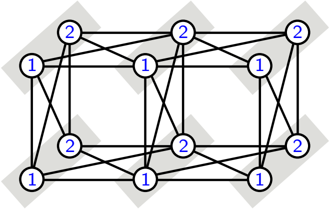


Min-sum problem
(a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V} \Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v) \Big].$$

All weights $f_u(k), f_{uv}(k,\ell) \in \mathbb{R} \cup \{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.

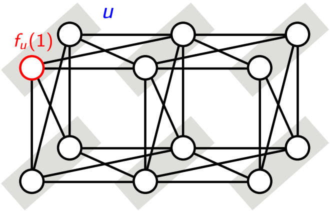


Min-sum problem
(a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V} \Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v) \Big].$$

All weights $f_u(k), f_{uv}(k, \ell) \in \mathbb{R} \cup \{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.

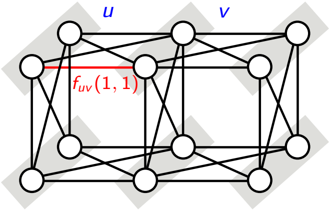


Min-sum problem
(a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V}\Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v)\Big].$$

All weights $f_u(k), f_{uv}(k,\ell) \in \mathbb{R}\cup\{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.

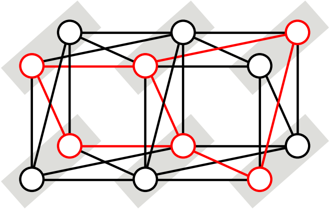


Min-sum problem
(a.k.a. MAP inference in graphical models or discrete energy minimization problem)

Pairwise min-sum problem with graph (V, E) and label set K:

$$\min_{\mathbf{k}\in K^V} \Big[\sum_{u\in V} f_u(k_u) + \sum_{\{u,v\}\in E} f_{uv}(k_u, k_v) \Big].$$

All weights $f_u(k), f_{uv}(k,\ell) \in \mathbb{R} \cup \{\infty\}$ form a vector $\mathbf{f}$.
Problem instance is defined by $(V, E, K, \mathbf{f})$.



Min-sum problem in computer vision

Segmentation

Stereo (correspondences)

Multiview reconstruction, surface fitting, shape matching, deconvolution, texture restoration, super resolution, . . .

Complexity of min-sum problem

In general, *NP*-hard.

Certain classes of instances are tractable.

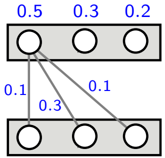
- min-sum problems on trees (restricting structure of graph)
- submodular min-sum problems (restricting weight functions f)
- ...

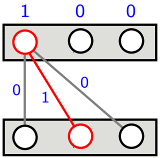
Linear programming relaxation of min-sum problem

LP relaxation = linear optimization over local marginal polytope:

$$\begin{aligned} \langle \mathbf{f}, \boldsymbol{\mu} \rangle &\to \min \\ \sum_{k \in K} \mu_u(k) &= 1, && u \in V \\ \sum_{\ell \in K} \mu_{uv}(k, \ell) &= \mu_u(k), && \{u, v\} \in E,\ k \in K \\ \boldsymbol{\mu} &\geq \mathbf{0} \end{aligned}$$

where in scalar product $\langle \mathbf{f}, \boldsymbol{\mu} \rangle$ we define $\infty \cdot 0 = 0$.
Components $\mu_u(k)$ and $\mu_{uv}(k, \ell)$ of $\boldsymbol{\mu}$ are pseudomarginals.





Solving the LP relaxation

2 labels

- the optimal solution is half-integral (pseudomarginals in $\{0, \frac{1}{2}, 1\}$)
- efficiently solvable by max-flow/min-cut algorithms [Boros & Hammer 1991]

Solving the LP relaxation

2 labels

- the optimal solution is half-integral (pseudomarginals in $\{0, \frac{1}{2}, 1\}$)
- efficiently solvable by max-flow/min-cut algorithms [Boros & Hammer 1991]

3+ labels

- general solvers (simplex, interior point): quadratic space complexity $\Rightarrow$ not applicable for large-scale instances

Solving the LP relaxation

2 labels

- the optimal solution is half-integral (pseudomarginals in $\{0, \frac{1}{2}, 1\}$)
- efficiently solvable by max-flow/min-cut algorithms [Boros & Hammer 1991]

3+ labels

- general solvers (simplex, interior point): quadratic space complexity $\Rightarrow$ not applicable for large-scale instances
- convergent message-passing algorithms: linear space complexity, but only local optimum, convergence rate not known

Solving the LP relaxation

2 labels

- the optimal solution is half-integral (pseudomarginals in $\{0, \frac{1}{2}, 1\}$)
- efficiently solvable by max-flow/min-cut algorithms [Boros & Hammer 1991]

3+ labels

- general solvers (simplex, interior point): quadratic space complexity $\Rightarrow$ not applicable for large-scale instances
- convergent message-passing algorithms: linear space complexity, but only local optimum, convergence rate not known
- recently, other algorithms with linear space (using subgradients [Komodakis et al. 2010], bundle methods [Kappes et al. 2012], steepest descent methods [Schwing et al. 2012, 2014], etc). But these are considerably slower than message-passing.

Solving the LP relaxation

2 labels

- the optimal solution is half-integral (pseudomarginals in $\{0, \frac{1}{2}, 1\}$)
- efficiently solvable by max-flow/min-cut algorithms [Boros & Hammer 1991]

3+ labels

- general solvers (simplex, interior point): quadratic space complexity $\Rightarrow$ not applicable for large-scale instances
- convergent message-passing algorithms: linear space complexity, but only local optimum, convergence rate not known
- recently, other algorithms with linear space (using subgradients [Komodakis et al. 2010], bundle methods [Kappes et al. 2012], steepest descent methods [Schwing et al. 2012, 2014], etc). But these are considerably slower than message-passing.

Is there a chance of inventing something better?

Reductions inside the class P

$X \leq_P Y$ (problem X is polynomial time reducible to problem Y)

Assuming, X is a well known problem, what does it say about Y?
Why it can be difficult to design a special efficient algorithm for Y?

Reductions inside the class P

$X \leq_P Y$ (problem X is polynomial time reducible to problem Y)

Assuming, X is a well known problem, what does it say about Y?
Why it can be difficult to design a special efficient algorithm for Y?

1. [stronger argument] Proposing a very fast algorithm for Y might result in a new, faster algorithm for X.

Reductions inside the class P

$X \leq_P Y$ (problem X is polynomial time reducible to problem Y)

Assuming, X is a well known problem, what does it say about Y?
Why it can be difficult to design a special efficient algorithm for Y?

1. [stronger argument] Proposing a very fast algorithm for Y might result in a new, faster algorithm for X.
2. [weaker argument] Proposing an algorithm for Y might bring a new principle for solving X.

Reductions inside the class P

$X \leq_P Y$ (problem X is polynomial time reducible to problem Y)

Assuming, X is a well known problem, what does it say about Y? Why it can be difficult to design a special efficient algorithm for Y?

1. [stronger argument] Proposing a very fast algorithm for Y might result in a new, faster algorithm for X.
2. [weaker argument] Proposing an algorithm for Y might bring a new principle for solving X.

In our case, X is general LP, Y is the LP relaxation of min-sum problem.

Linear programming - history

Simplex algorithm [Dantzig 1947]

- quadratic space, polynomial time not guaranteed

Linear programming - history

Simplex algorithm [Dantzig 1947]

- quadratic space, polynomial time not guaranteed

Ellipsoid algorithm [Khachiyan 1979]

- first polynomial time algorithm for LP

Linear programming - history

Simplex algorithm [Dantzig 1947]

- quadratic space, polynomial time not guaranteed

Ellipsoid algorithm [Khachiyan 1979]

- first polynomial time algorithm for LP

Karmarkar's algorithm [Karmarkar 1984]

- interior point method
- fastest known algorithm for LP
  $\mathcal{O}(n^{3.5} L^2 \log L \log \log L)$

Main result

Theorem (Průša-Werner-CVPR2013)

Any linear program can be reduced in linear time to the LP relaxation of a pairwise min-sum problem with 3 labels.

Main result

Theorem (Průša-Werner-CVPR2013)

Any linear program can be reduced in linear time to the LP relaxation of a pairwise min-sum problem with 3 labels.

Consequences:

- Finding an efficient algorithm to solve LP relaxation of min-sum problem might be as hard as improving the complexity of the best known algorithm for LP.

Main result

Theorem (Průša-Werner-CVPR2013)

Any linear program can be reduced in linear time to the LP relaxation of a pairwise min-sum problem with 3 labels.

Consequences:

- Finding an efficient algorithm to solve LP relaxation of min-sum problem might be as hard as improving the complexity of the best known algorithm for LP.
- LP relaxation of min-sum problem with 3+ labels is inherently more complex than for 2 labels.

Elementary min-sum problems

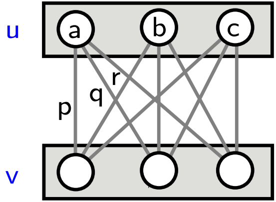
The reduction is done by combining elementary min-sum problems.

- They perform simple operations on unary pseudomarginals.

The reduction is done by combining elementary min-sum problems.

- They perform simple operations on unary pseudomarginals.

- Depicting a pair $\{u, v\} \in E$ with $|K| = 3$ labels:

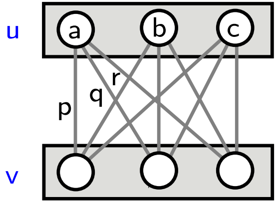


$$p + q + r = a$$
$$a + b + c = 1$$

Elementary min-sum problems

The reduction is done by combining elementary min-sum problems.

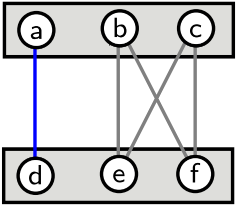
- They perform simple operations on unary pseudomarginals.
- Depicting a pair $\{u, v\} \in E$ with $|K| = 3$ labels:



$$p + q + r = a$$
$$a + b + c = 1$$

- Visible edges have weights $f_{uv}(k, \ell) = 0$.
  Invisible edge have weights $f_{uv}(k, \ell) = \infty$, implying $\mu_{uv}(k, \ell) = 0$.

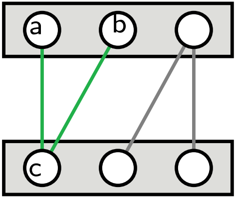
Elementary min-sum problem COPY



Enforces $a = d$.

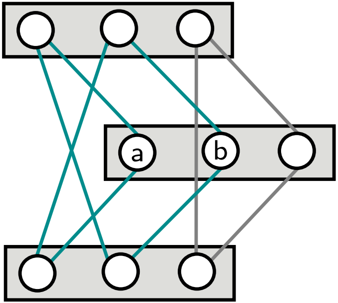
Precisely:
Given any feasible unary pseudomarginals a, b, c, d, e, f,
feasible pairwise pseudomarginals exist if and only if $a = d$.

Elementary min-sum problem ADDITION



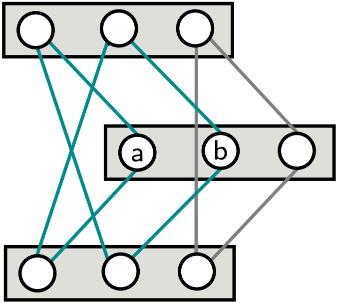
Enforces $c = a + b$.

Elementary min-sum problem EQUALITY

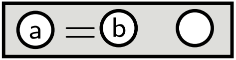


Enforces $a = b$.

Elementary min-sum problem EQUALITY

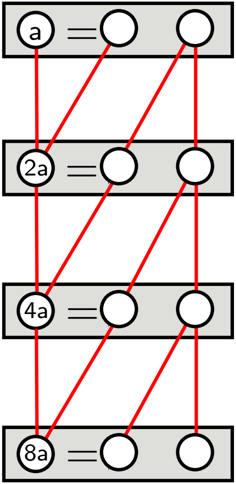


Enforces $a = b$.



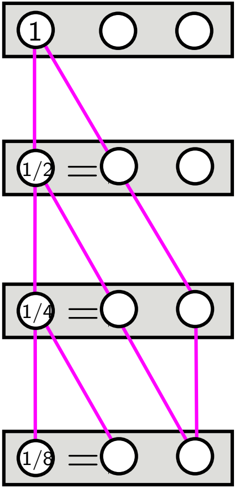
shorthand

Elementary min-sum problem POWERS



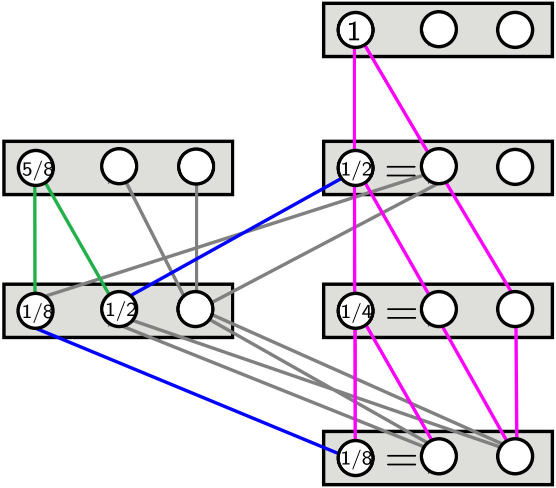
Constructs unary pseudomarginals with values $2^i a$ for $i = 0, \ldots, d$, where d is the depth of the problem.

Elementary min-sum problem NegPowers



Constructs unary pseudomarginals with values 2^{-i} for $i = 0, \ldots, d$.

Example of combining elementary min-sum problems



Constructs a unary pseudomarginal with value $5/8 = 5 \cdot 2^{-d}$.
Similarly, we can construct any multiple of 2^{-d} (not greater than 1).

The input LP

The input of the reduction is the LP

$$\min\{\, \langle \mathbf{c}, \mathbf{x} \rangle \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$$

where $\mathbf{A} \in \mathbb{Z}^{m \times n}$, $\mathbf{b} \in \mathbb{Z}^m$, $\mathbf{c} \in \mathbb{Z}^n$, $m \leq n$.

The input LP

The input of the reduction is the LP

$$\min\{ \langle \mathbf{c}, \mathbf{x} \rangle \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \}$$

where $\mathbf{A} \in \mathbb{Z}^{m \times n}$, $\mathbf{b} \in \mathbb{Z}^m$, $\mathbf{c} \in \mathbb{Z}^n$, $m \leq n$.

Before reduction, the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is rewritten as

$$\mathbf{A}^+\mathbf{x} = \mathbf{A}^-\mathbf{x} + \mathbf{b}$$

where all entries of $\mathbf{A}^+, \mathbf{A}^-, \mathbf{b}$ are non-negative and $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$.

Bounding the variable ranges

Lemma

Let $\mathbf{x}$ *be a vertex of the polyhedron* $\{\, \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$*. Then each component* x_j *of* $\mathbf{x}$ *satisfies either* $x_j = 0$ *or* $M^{-1} \leq x_j \leq M$*, where*

$$\begin{aligned} M &= m^{m/2}(B_1 \times \cdots \times B_{n+1}) \\ B_j &= \max\{1, |a_{1j}|, \ldots, |a_{mj}|\}, \quad j = 1, \ldots, n \\ B_{n+1} &= \max\{1, |b_1|, \ldots |b_m|\}. \end{aligned}$$

Bounding the variable ranges

Lemma

Let $\mathbf{x}$ *be a vertex of the polyhedron* $\{\, \mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$. *Then each component* x_j *of* $\mathbf{x}$ *satisfies either* $x_j = 0$ *or* $M^{-1} \leq x_j \leq M$, *where*

$$\begin{aligned} M &= m^{m/2}(B_1 \times \cdots \times B_{n+1}) \\ B_j &= \max\{1, |a_{1j}|, \ldots, |a_{mj}|\}, \quad j = 1, \ldots, n \\ B_{n+1} &= \max\{1, |b_1|, \ldots |b_m|\}. \end{aligned}$$

Lemma

Let the polyhedron $\{\, \mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ *be bounded. Then for any* $\mathbf{x}$ *from the polyhedron, each component of* $\mathbf{A}^{+}\mathbf{x}$ *and* $\mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ *is not greater than* $N = M(B_1 + \cdots + B_{n+1})$.

Initializing the reduction

The reduction algorithm:

- Its input is $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, assuming w.l.o.g. that the polyhedron $\{\, \mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ is bounded.
- Its output will be a min-sum problem $(V, E, K, \mathbf{f})$ with $V = \{1, \ldots, |V|\}$ and $K = \{1, 2, 3\}$.

Initializing the reduction

The reduction algorithm:

- Its input is $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, assuming w.l.o.g. that the polyhedron $\{\, \mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ is bounded.
- Its output will be a min-sum problem $(V, E, K, \mathbf{f})$ with $V = \{1, \dots, |V|\}$ and $K = \{1, 2, 3\}$.

The algorithm is initialized as follows:

1. For each variable x_j in the input LP, introduce a new object j into V and set $f_j(1) = c_j$.
   (Pseudomarginal $\mu_j(1)$ will represent variable x_j.)

Initializing the reduction

The reduction algorithm:

- Its input is $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, assuming w.l.o.g. that the polyhedron $\{\, \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ is bounded.
- Its output will be a min-sum problem $(V, E, K, \mathbf{f})$ with $V = \{1, \dots, |V|\}$ and $K = \{1, 2, 3\}$.

The algorithm is initialized as follows:

1. For each variable x_j in the input LP, introduce a new object j into V and set $f_j(1) = c_j$.
   (Pseudomarginal $\mu_j(1)$ will represent variable x_j.)
2. For each such object $j \in V$, build $\textsc{Powers}$ with the depth $d_j = \lfloor \log_2 B_j \rfloor$ based on label 1.

Initializing the reduction

The reduction algorithm:

- Its input is $(\mathbf{A}, \mathbf{b}, \mathbf{c})$, assuming w.l.o.g. that the polyhedron $\{\, \mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ is bounded.
- Its output will be a min-sum problem $(V, E, K, \mathbf{f})$ with $V = \{1, \ldots, |V|\}$ and $K = \{1, 2, 3\}$.

The algorithm is initialized as follows:

1. For each variable x_j in the input LP, introduce a new object j into V and set $f_j(1) = c_j$.
   (Pseudomarginal $\mu_j(1)$ will represent variable x_j.)
2. For each such object $j \in V$, build $\textsc{Powers}$ with the depth $d_j = \lfloor \log_2 B_j \rfloor$ based on label 1.
3. Build $\textsc{NegPowers}$ with the depth $d = \lceil \log_2 N \rceil$.

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

1. Construct pseudomarginals with values $a_{ij}^{+}x_j$ and $a_{ij}^{-}x_j$ by summing selected values from the POWERS.

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

1. Construct pseudomarginals with values $a_{ij}^{+}x_j$ and $a_{ij}^{-}x_j$ by summing selected values from the POWERS.
2. Construct a pseudomarginal with value $2^{-d}b_i$ by summing selected values from the NEGPOWERS. (The number 2^{-d} plays the rôle of the unit.)

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

1. Construct pseudomarginals with values $a_{ij}^{+}x_j$ and $a_{ij}^{-}x_j$ by summing selected values from the POWERS.
2. Construct a pseudomarginal with value $2^{-d}b_i$ by summing selected values from the NEGPOWERS. (The number 2^{-d} plays the rôle of the unit.)
3. Sum the terms on each side of the equation by repetitively applying ADDITION and COPY.

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

1. Construct pseudomarginals with values $a_{ij}^{+}x_j$ and $a_{ij}^{-}x_j$ by summing selected values from the POWERS.
2. Construct a pseudomarginal with value $2^{-d}b_i$ by summing selected values from the NEGPOWERS. (The number 2^{-d} plays the rôle of the unit.)
3. Sum the terms on each side of the equation by repetitively applying ADDITION and COPY.
4. Enforce equality of the two sides of the equation by COPY.

Encoding the equality constraints

Each equation

$$a_{i1}^{+}x_1 + \cdots + a_{in}^{+}x_n = a_{i1}^{-}x_1 + \cdots + a_{in}^{-}x_n + b_i$$

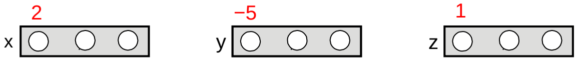
of the system $\mathbf{A}^{+}\mathbf{x} = \mathbf{A}^{-}\mathbf{x} + \mathbf{b}$ is encoded as follows:

1. Construct pseudomarginals with values $a_{ij}^{+}x_j$ and $a_{ij}^{-}x_j$ by summing selected values from the POWERS.
2. Construct a pseudomarginal with value $2^{-d}b_i$ by summing selected values from the NEGPOWERS. (The number 2^{-d} plays the rôle of the unit.)
3. Sum the terms on each side of the equation by repetitively applying ADDITION and COPY.
4. Enforce equality of the two sides of the equation by COPY.

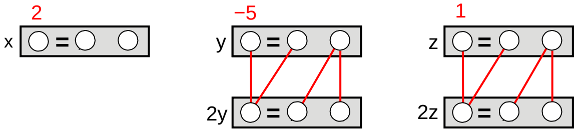
Finally, set $f_i(k) = 0$ for all $i > n$ or $k > 1$.

$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$

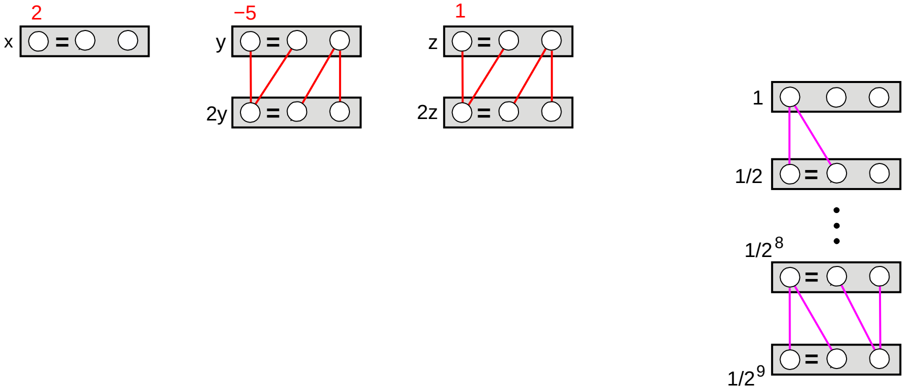
$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$



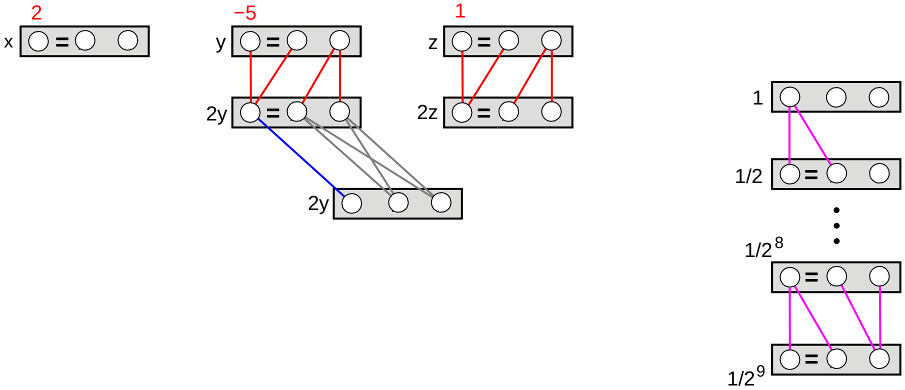
$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$



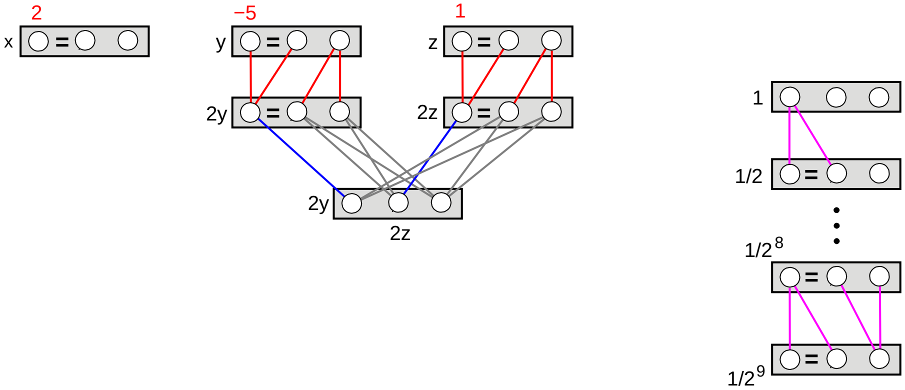
$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$



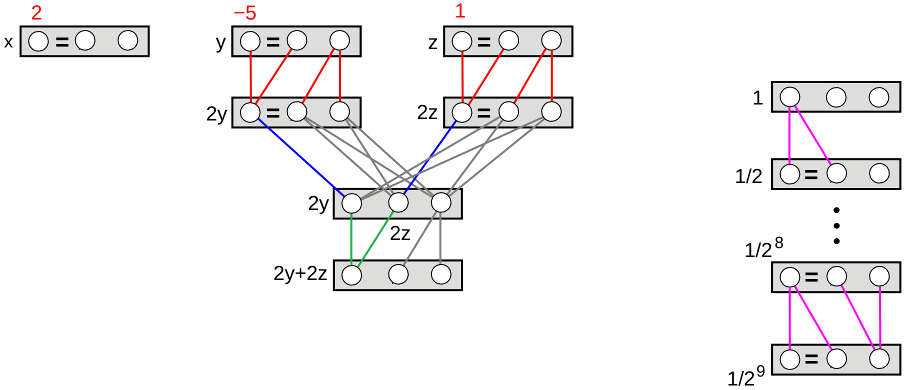
$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$

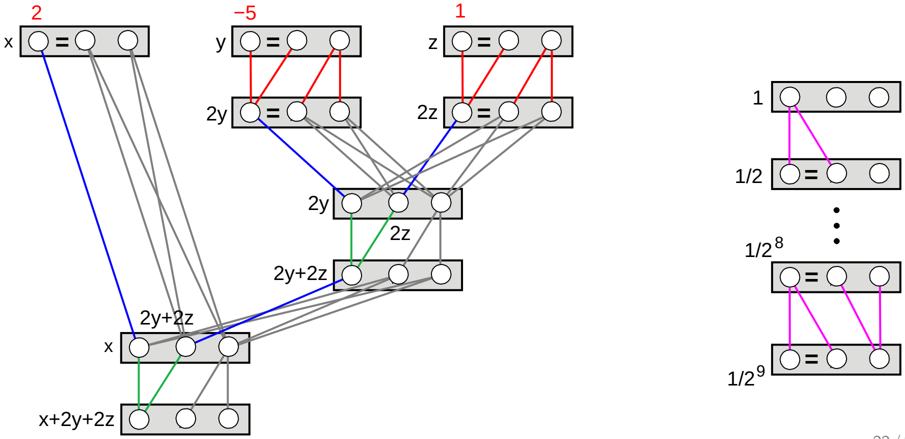


$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$

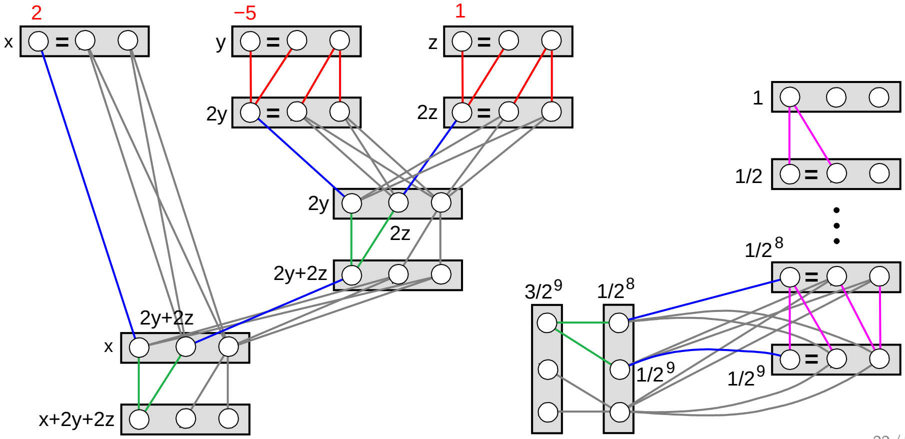


$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$

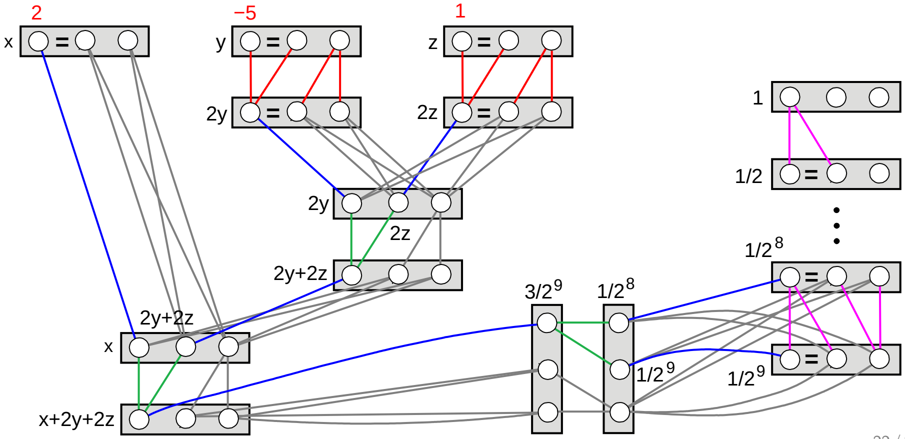



2
−5
1
x
y
z
2y
2z
2y
2z
2y+2z
2y+2z
x
x+2y+2z
1
1/2
1/2 8
1/2 9

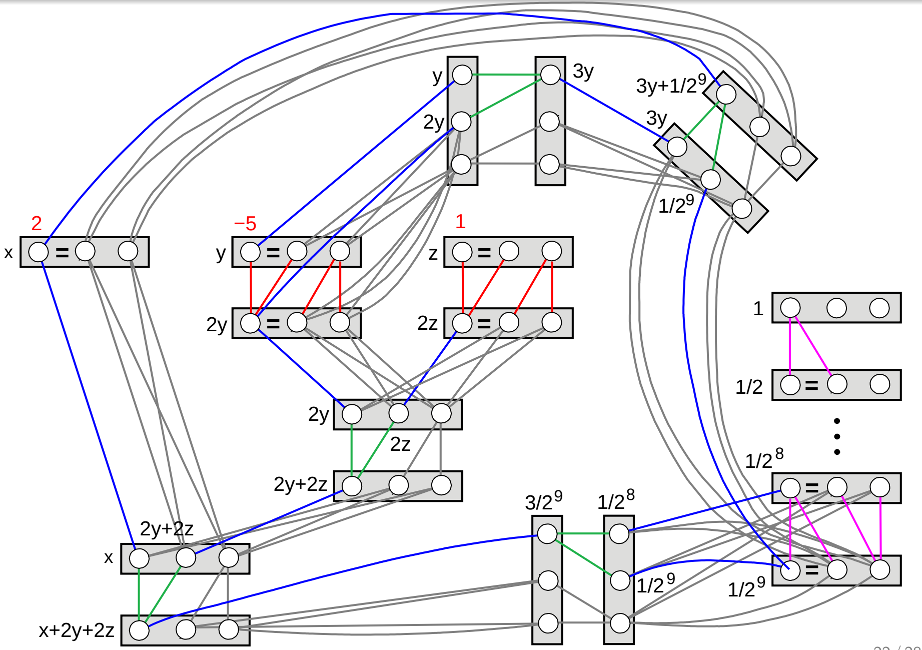
$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$



$\min\{\,2x-5y+z \mid x+2y+2z=3;\; x=3y+1;\; x,y,z\geq 0\,\}$



$\min\{\, 2x - 5y + z \mid x + 2y + 2z = 3;\ x = 3y + 1;\ x, y, z \geq 0 \,\}$



Complexity of the reduction

Let L be the number of bits of the binary representation of $(\mathbf{A}, \mathbf{b}, \mathbf{c})$.
Want to prove that the reduction time is $\mathcal{O}(L)$.

Complexity of the reduction

Let L be the number of bits of the binary representation of $(\mathbf{A}, \mathbf{b}, \mathbf{c})$.
Want to prove that the reduction time is $\mathcal{O}(L)$.

This is easy:

- Let the output of the reduction be $(V, E, K, \mathbf{f})$.
- Clearly, the reduction time is $\mathcal{O}(|E|)$.
- Clearly, $|E| = \mathcal{O}(|V|)$.
- Thus we need to prove $|V| = \mathcal{O}(L)$.
- For that, it suffices to prove that the numbers $d_j = \lceil \log_2 B_j \rceil$ and $d = \lceil \log_2 N \rceil$ are $\mathcal{O}(L)$.

Other results

Corollary

Every polytope is (up to scale) a coordinate-erasing projection of a face of a local marginal polytope with 3 labels, whose description can be computed from the description of the original polytope in linear time.

Other results

Corollary

Every polytope is (up to scale) a coordinate-erasing projection of a face of a local marginal polytope with 3 labels, whose description can be computed from the description of the original polytope in linear time.

If only finite weights are allowed ($f_u(k), f_{uv}(k, \ell) \in \mathbb{R}$) then:

Theorem

Any linear program reduces in time and space $\mathcal{O}(L^2)$ to a linear optimization over a local marginal polytope with 3 labels.

Planar graphs

Vision applications usually induce sparse, planar graphs (like grids).

Is it possible to reduce every LP to a min-sum problem with the underlying planar graph?

Planar graphs

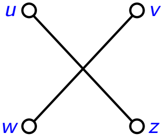
Vision applications usually induce sparse, planar graphs (like grids).

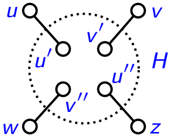
Is it possible to reduce every LP to a min-sum problem with the underlying planar graph?

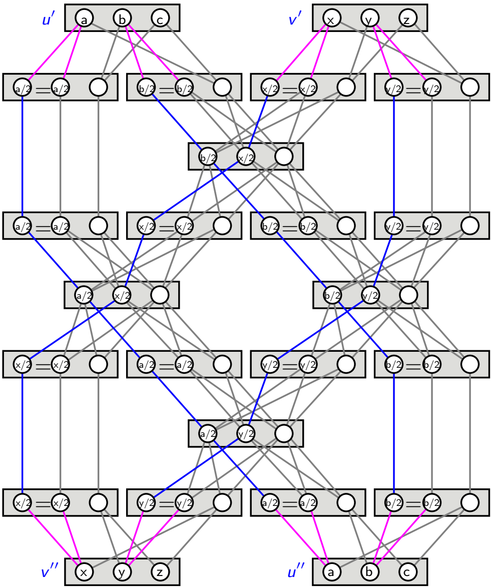
Theorem (Průša-Werner-PAMI2014)

Every LP reduces to a linear optimization (with infinite costs) over a local marginal polytope with 3 labels over a planar graph. The size of the output and the reduction time are $\mathcal{O}(mL)$.

Planar graphs – eliminating one edge crossing





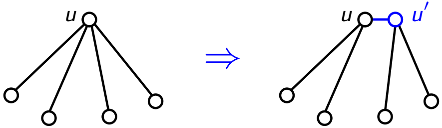


Reduction to a grid

Theorem (Tamassia 1989)

Any planar graph $G = (V, E)$ with maximal node degree 4 can be embedded in linear time into a grid with the area $\mathcal{O}(|V|^2)$.

All degrees of nodes in a planar min-sum problem can be reduced to 3 (for a chosen node, create its copies and distribute incident edges among them).



Publications

- D. Průša, T. Werner: Universality of the Local Marginal Polytope. *CVPR*, 2013.
- S. Živný, T. Werner, D. Průša: The Power of LP Relaxation for MAP Inference. A chapter in: *Advanced Structured Prediction*, MIT Press, 2014 (To appear).
- D. Průša, T. Werner: Universality of the Local Marginal Polytope. *IEEE Transactions on PAMI*, 2014 (Early access).
- D. Průša, T. Werner: How Hard is the LP Relaxation of the Potts Min-Sum Labeling Problem? *EMMCVPR*, 2015 (To appear).