Maximum Persistency via Iterative Relaxed Inference with Graphical Models

Alexander Shekhovtsov, Paul Swoboda, Bogdan Savchynskyy
TU Graz IST Austria TU Dresden

"Maximum Persistency via Iterative Relaxed Inference with Graphical Models" CVPR'15
A.S.: "Higher Order Maximum Persistency and Comparison Theorems", CVIU'16

May 23, 2016

Persistency (Partial Optimality)

ILP	LP
$\min c^{\mathsf{T}} x$	$\min c^{\mathsf{T}} x$
$Ax \leq b$	$Ax \leq b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$

1. When the solution to LP is integer?

Persistency (Partial Optimality)

ILP	LP
$\min c^\mathsf{T} x$	$\min c^\mathsf{T} x$
$Ax \le b$	$Ax \le b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$

1. When the solution to LP is integer?

$$x = (0, 1, 1, 0, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{3}, 1, 0, \dots)$$

2. Is the integer part of an optimal solution to LP optimal for ILP?

Persistency (Partial Optimality)

ILP	LP
$\min c^{\mathsf{T}} x$	$\min c^{\mathsf{T}} x$
$Ax \leq b$	$Ax \leq b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$

1. When the solution to LP is integer?

$$x = (0, 1, 1, 0, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{3}, 1, 0, \dots)$$

2. Is the integer part of an optimal solution to LP optimal for ILP?
3. Is a part of the integer part is optimal for ILP?

Persistency (Partial Optimality)

ILP	LP
$\min c^{\mathsf{T}} x$	$\min c^{\mathsf{T}} x$
$Ax \leq b$	$Ax \leq b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$

1. When the solution to LP is integer?

$$x = (0, 1, 1, 0, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{3}, 1, 0, \dots)$$

2. Is the integer part of an optimal solution to LP optimal for ILP?
3. Is a part of the integer part is optimal for ILP?
4. Sufficient conditions for a part of an optimal solution to LP to be optimal for ILP?

Persistency (Partial Optimality)

ILP	LP
$\min c^\mathsf{T} x$	$\min c^\mathsf{T} x$
$Ax \leq b$	$Ax \leq b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$

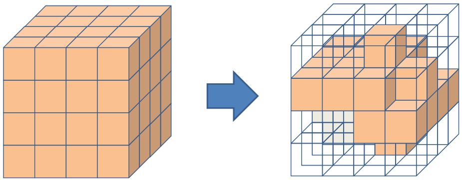
1. When the solution to LP is integer?

$$x = (0, 1, 1, 0, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{3}, 1, 0, \dots)$$

2. Is the integer part of an optimal solution to LP optimal for ILP?
3. Is a part of the integer part is optimal for ILP?
4. Sufficient conditions for a part of an optimal solution to LP to be optimal for ILP?
5. Find the largest part satisfying sufficient conditions.

Persistency (Partial Optimality)

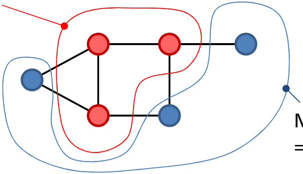
ILP	LP
$\min c^{\mathsf{T}} x$	$\min c^{\mathsf{T}} x$
$Ax \leq b$	$Ax \leq b$
$x \in \{0,1\}^n$	$x \in [0,1]^n$



Outline

- Introduction
  - ✓ Persistency (Partial Optimality)
  - Vertex Packing, QPBO, Energy Minimization
  - Optimization-Based Methods for Persistency
- Generalized Sufficient Conditions
  - Improving Substitution
  - Relaxed-Improving Substitution
  - Generality
- Maximizing Persistency
  - Optimization-Based Formulation
  - Discrete Cutting Plane
  - OpenGM Benchmark

Vertex Packing / Maximum Independent Set



Maximum Weighted Vertex Packing

- $(\mathcal{V}, \mathcal{E})$ – an undirected graph;
- *Vertex Packing* is a subset $P \subset \mathcal{V}$ for which $u, v \in P \Rightarrow (u, v) \notin \mathcal{E}$;
- Weights $c\colon \mathcal{V} \to \mathbb{R}$;
- Problem:

$$\max_x \sum_{v \in \mathcal{V}} c_v x_v \tag{VP}$$
$$(\forall uv \in \mathcal{E})\ x_u + x_v \leq 1,$$
$$(\forall v \in \mathcal{V})\ x_v \in \{0, 1\}.$$

Vertex Packing / Maximum Independent Set

Relaxing the integrality constraints:

$$\max_{\mu} \sum_{v \in \mathcal{V}} c_v \mu_v \qquad \text{(VPL)}$$
$$(\forall uv \in \mathcal{E}) \ \mu_u + \mu_v \leq 1,$$
$$(\forall v \in \mathcal{V}) \ \mu_v \geq 0.$$

Theorems

- (Balinski, 1965; Lorentzen, 1966): Any basic feasible solution to (VLP) is $\{0, \frac{1}{2}, 1\}$-valued.
- (Edmonds and Pulleyblank) (VLP) reduces to a maxflow problem on a related symmetric bipartite graph;
- (Nemhauser and Trotter, 1975): Variables which assume binary values in an optimum (VLP) solution retain the same values in an optimum (VP) solution.
- (Picard and Queyranne, 1977): There exists a unique maximum set of variables that are integer valued in an optimal solution to (VLP).

QPBO

Quadratic pseudo-Boolean Optimization (QPBO)

- $(\mathcal{V}, \mathcal{E})$ – an undirected graph;
- Weights $a\colon \mathcal{V} \cup \mathcal{E} \to \mathbb{R}$;
- Problem:

$$\min_x \sum_{v\in\mathcal{V}} a_v x_v + \sum_{uv\in\mathcal{E}} a_{uv} x_u x_v$$

$$(\forall v \in \mathcal{V})\ x_v \in \{0, 1\}.$$

- Generalizes Vertex Packing (let $a_{uv} = B$, a big number; $a = -c_v$).

QPBO

Natural linear relaxation: $x_s \to \mu_s \in [0, 1]$, $x_s x_t \to \mu_{st} \in [0, 1]$ (lifting)

$$\min_{\mu\colon \mathcal{V}\cup\mathcal{E}\to[0,1]} \sum_{v\in\mathcal{V}} a_v \mu_v + \sum_{uv\in\mathcal{E}} a_{uv}\mu_{uv} \qquad \text{(LP)}$$

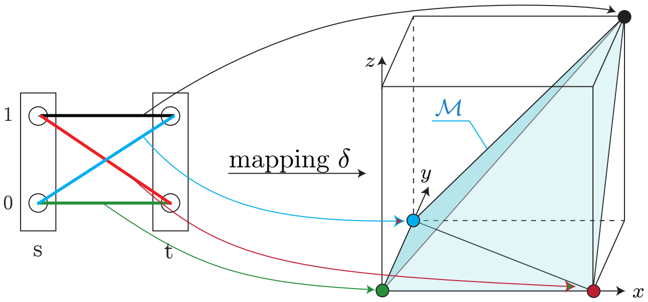
s.t. $(\forall uv \in \mathcal{E})\ \mu_u + \mu_v - 1 \leq \mu_{uv} \leq \min(\mu_u, \mu_v)$ (local convex hulls).

QPBO

Natural linear relaxation: $x_s \to \mu_s \in [0, 1]$, $x_s x_t \to \mu_{st} \in [0, 1]$ (lifting)

$$\min_{\mu\colon \mathcal{V}\cup\mathcal{E}\to[0,1]} \sum_{v\in\mathcal{V}} a_v \mu_v + \sum_{uv\in\mathcal{E}} a_{uv}\mu_{uv} \quad \text{(LP)}$$

s.t. $(\forall uv \in \mathcal{E})\ \mu_u + \mu_v - 1 \leq \mu_{uv} \leq \min(\mu_u, \mu_v)$ (local convex hulls).



QPBO

Natural linear relaxation: $x_s \to \mu_s \in [0, 1]$, $x_s x_t \to \mu_{st} \in [0, 1]$ (lifting)

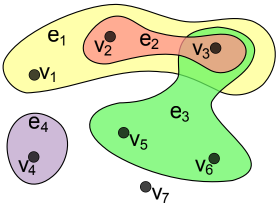
$$\min_{\mu\colon \mathcal{V}\cup\mathcal{E}\to[0,1]} \sum_{v\in\mathcal{V}} a_v \mu_v + \sum_{uv\in\mathcal{E}} a_{uv}\mu_{uv} \tag{LP}$$

$$\text{s.t. } (\forall uv \in \mathcal{E})\ \mu_u + \mu_v - 1 \le \mu_{uv} \le \min(\mu_u, \mu_v)$$ (local convex hulls).

Theorems

- Each extreme point of the feasible set is $\{0, \frac{1}{2}, 1\}$-valued.
- (Hammer et al., 1984; Boros et al., 1991): LP reduces to a maxflow problem;
- **Weak Persistency** (Hammer et al., 1984): Variables μ_v which assume binary values in an optimum (LP) solution retain the same values in an ILP solution.
- **Strong Persistency** (Hammer et al., 1984): Variables μ_v which assume binary values in **all** optimal (LP) solutions retain the same values in **all** optimal ILP solutions.

0-1 Polynomial Programming



A hypergraph (courtesy of wikipedia).

0-1 Polynomial Programming / pseudo-Boolean Optimization

- $(\mathcal{V}, \mathcal{E})$ – a hypergraph, $\mathcal{E} \subset 2^{\mathcal{V}}$;
- Weights $f \colon \mathcal{E} \to \mathbb{R}$;
- Problem:

$$\min_{x \in \{0,1\}^{\mathcal{V}}} \sum_{\mathrm{C} \in \mathcal{E}} f_{\mathrm{C}} \prod_{v \in \mathrm{C}} x_v. \tag{PP}$$

- Any pseudo-Boolean function can be represented as a multilinear polynomial.

0-1 Polynomial Programming

- Quadratization techniques
  - + 0-1 PP can be reduced to QPBO with auxiliary variables Boros and Hammer (2001), Ishikawa (2011), Fix et al. (2011)
  - + Can apply roof dual relaxation (combinatorial, persistency)
  - – Relaxation of reduced problem is looser, multiple reductions

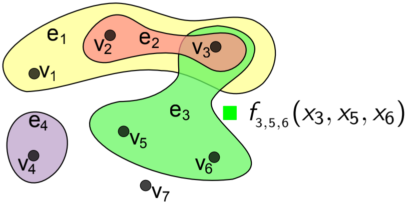
0-1 Polynomial Programming

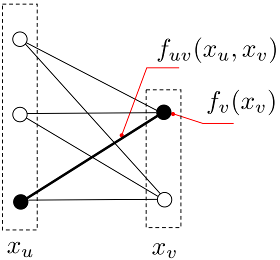
- Quadratization techniques
  - \+ 0-1 PP can be reduced to QPBO with auxiliary variables Boros and Hammer (2001), Ishikawa (2011), Fix et al. (2011)
  - \+ Can apply roof dual relaxation (combinatorial, persistency)
  - – Relaxation of reduced problem is looser, multiple reductions
- Special Relaxations: (bi)submodular relaxations (Kolmogorov, 2012)
  - \+ extreme feasible solutions are half-integral;
  - \+ reduces to sum of (bi)submodular functions minimization (combinatorial);
  - \+ all integer variables are persistent;
  - – Relatively loose, multiple choices

0-1 Polynomial Programming

- Quadratization techniques
  - + 0-1 PP can be reduced to QPBO with auxiliary variables Boros and Hammer (2001), Ishikawa (2011), Fix et al. (2011)
  - + Can apply roof dual relaxation (combinatorial, persistency)
  - – Relaxation of reduced problem is looser, multiple reductions
- Special Relaxations: (bi)submodular relaxations (Kolmogorov, 2012)
  - + extreme feasible solutions are half-integral;
  - + reduces to sum of (bi)submodular functions minimization (combinatorial);
  - + all integer variables are persistent;
  - – Relatively loose, multiple choices
- Tighter relaxations, *e.g.* relaxation of Sherali and Adams (1990)
  - – optimal solutions are not half-integral in general;
  - – no combinatorial method to solve;
  - – not persistent in general;

Energy Minimization / Graphical Model





Energy Minimization / Weighted Constraint Satisfaction

- $(\mathcal{V}, \mathcal{E})$ - a hypergraph;
- $\mathcal{X}_v$ - a finite set of *labels*, $v \in \mathcal{V}$;
- Costs $f_{\mathrm{C}} \colon \prod_{v \in \mathrm{C}} \mathcal{X}_v \to \mathbb{R}$, $\mathrm{C} \in \mathcal{E}$;
- Energy: $E_f(x) = \sum_{\mathrm{C} \in \mathcal{E}} f_{\mathrm{C}}(x_{\mathrm{C}})$;
- Probloem: $\min_{x \in \mathcal{X}} E_f(x)$;

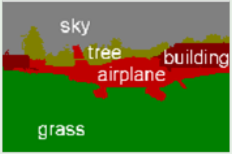
Energy Minimization

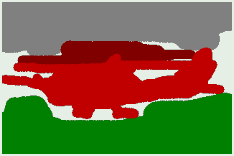
Example: Potts Model for Object Class Segmentation

- $\mathcal{V}$ - set of pixels; $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ neighboring pixels;
- $\mathcal{X}_s = \{1, \dots K\}$ – class label;
- $E_f(x) = \sum_{s \in \mathcal{V}} f_s(x_s) + \sum_{st \in \mathcal{E}} \lambda_{st} [\![x_s \neq x_t]\!]$.

Image

Ground Truth





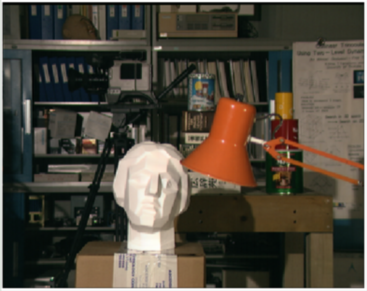
(MSRC object class segmentation)

Energy Minimization

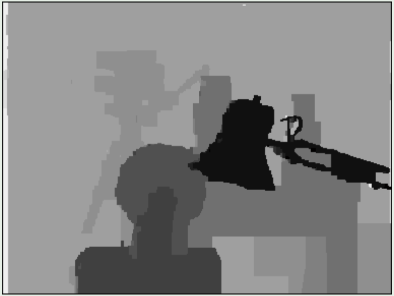
Example: Potts Model for Stereo

- $\mathcal{V}$ - set of pixels; $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ neighboring pixels;
- $\mathcal{X}_s = \{1, \dots K\}$ – disparity value;
- $E_f(x) = \sum_{s \in \mathcal{V}} f_s(x_s) + \sum_{st \in \mathcal{E}} \lambda_{st} [\![x_s \neq x_t]\!]$.

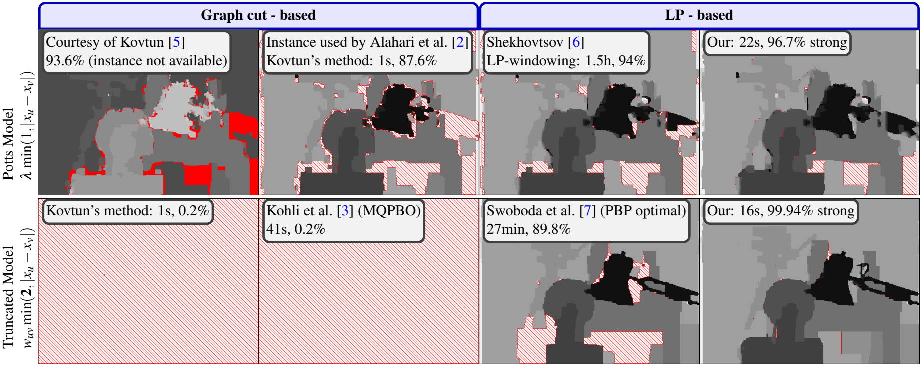
Reference (Left) Image



Depth Reconstruction



Persistency - Sufficient Conditions



- Model 1 (Kovtun'03, Alahari et al.'10): Potts, strong unaries with window aggregation
- Model 2 (Szeliski et al., 2008): Nearly Potts, per-pixel unaries

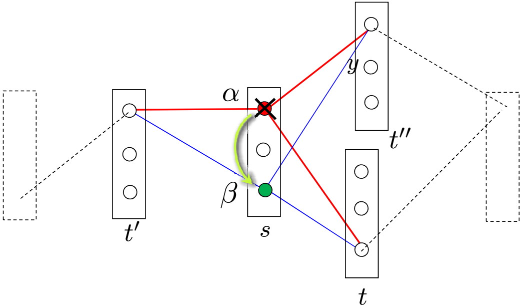
Generalized Sufficient Conditions for Persistency

Improving Substitution

- $E_f : \mathcal{X} \to \mathbb{R}$, $x \in \mathcal{X}$
- Substitution: $(x_1, x_2, \ldots, x_v, \ldots x_n) \to (x_1, x_2, \ldots, \alpha, \ldots x_n)$
- Denote as $x[v \leftarrow \alpha]$
- If $E_f(x[v \leftarrow \alpha]) \leq E_f(x)$ for all x then $x_v = \alpha$ is optimal!

Improving Substitution

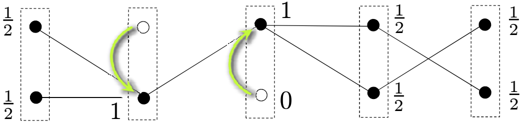
- $E_f : \mathcal{X} \to \mathbb{R}$, $x \in \mathcal{X}$
- Substitution: $(x_1, x_2, \ldots, x_v, \ldots x_n) \to (x_1, x_2, \ldots, \alpha, \ldots x_n)$
- Denote as $x[v \leftarrow \alpha]$
- If $E_f(x[v \leftarrow \alpha]) \leq E_f(x)$ for all x then $x_v = \alpha$ is optimal!
- If $E_f(x[v \leftarrow \beta]) \leq E_f(x[v \leftarrow \alpha])$ for all x then $x_v = \alpha$ can be thrown away!



- Substitutability (constraint programming) dominance (valued constraint satisfaction) dead end elimination (bioinformatics)

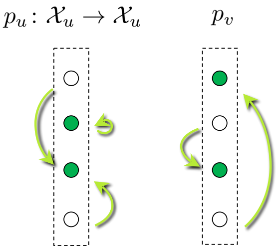
Improving Substitution

- Substitute simultaneously:
- Let $x[\mathcal{A} \leftarrow y_{\mathcal{A}}]_v = y_v$ for $v \in \mathcal{A}$ and x_v for $v \in \mathcal{V} \backslash \mathcal{A}$.
- If $E_f(x[\mathcal{A} \leftarrow y_{\mathcal{A}}]) \leq E_f(x)$ for all x then $y_{\mathcal{A}}$ is a part of an optimal assignment.



- Autarky in QPBO
- Verifying whether $y_{\mathcal{A}}$ satisfies condition is NP-hard.

Simultaneous Improving Substitution

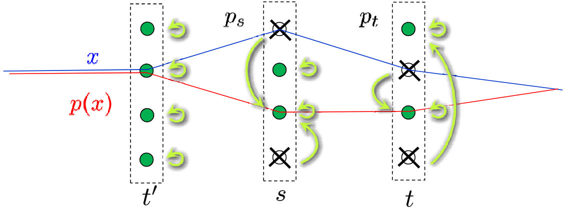


$p\colon \mathcal{X} \to \mathcal{X}$ node-wise

Definition

Substitution p is improving if
$(\forall x) \quad E_f(p(x)) \leq E_f(x)$.

- If x is optimal then $p(x)$ is optimal. Search space can be reduced.



Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify

Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify
- Equivalent to: $\min_{x \in \mathcal{X}} (E_f(x) - E_f(p(x))) \geq 0$

Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify
- Equivalent to: $\min_{x \in \mathcal{X}} (E_f(x) - E_f(p(x))) \geq 0$
- Lift: $\min_{\mu \in \delta(\mathcal{X})} (\langle f, \mu \rangle - \langle f, P\mu \rangle) \geq 0$

Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify
- Equivalent to: $\min_{x \in \mathcal{X}} (E_f(x) - E_f(p(x))) \geq 0$
- Lift: $\min_{\mu \in \delta(\mathcal{X})} (\langle f, \mu \rangle - \langle f, P\mu \rangle) \geq 0$
- Relax: $\min_{\mu \in \Lambda} \langle f - P^{\mathsf{T}} f, \mu \rangle \geq 0$, Λ - any tractable polytope containing $\delta(\mathcal{X})$

Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify
- Equivalent to: $\min_{x \in \mathcal{X}} (E_f(x) - E_f(p(x))) \geq 0$
- Lift: $\min_{\mu \in \delta(\mathcal{X})} (\langle f, \mu \rangle - \langle f, P\mu \rangle) \geq 0$
- Relax: $\min_{\mu \in \Lambda} \langle f - P^{\mathsf{T}} f, \mu \rangle \geq 0$, Λ - any tractable polytope containing $\delta(\mathcal{X})$

Definition:

Substitution p is *relaxed-improving* if $\min_{\mu \in \Lambda} \langle (I - P^{\mathsf{T}}) f, \mu \rangle \geq 0$

- Polynomial to verify.

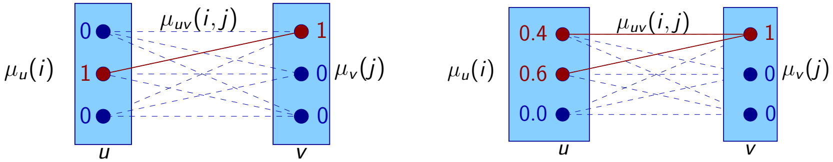
Sufficient Conditions for Peristency

- Improving mapping: $(\forall x \in \mathcal{X})\ E_f(p(x)) \leq E_f(x)$ – NP hard to verify
- Equivalent to: $\min_{x \in \mathcal{X}} (E_f(x) - E_f(p(x))) \geq 0$
- Lift: $\min_{\mu \in \delta(\mathcal{X})} (\langle f, \mu \rangle - \langle f, P\mu \rangle) \geq 0$
- Relax: $\min_{\mu \in \Lambda} \langle f - P^{\mathsf{T}} f, \mu \rangle \geq 0$, Λ - any tractable polytope containing $\delta(\mathcal{X})$

Definition:

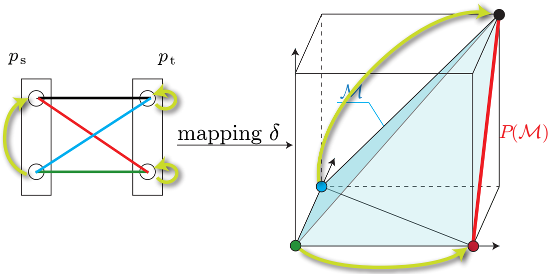
Substitution p is *relaxed-improving* if $\min_{\mu \in \Lambda} \langle (I - P^{\mathsf{T}}) f, \mu \rangle \geq 0$

- Polynomial to verify.



Relaxed Improving Substitution

- Substitution $p\colon \mathcal{X} \to \mathcal{X}$ can be represented in the lifted space:



- Linear mapping P is the *extension* of $p\colon \mathcal{X} \to \mathcal{X}$,
- An oblique projection onto a facet.

Generality of Sufficient Conditions

- Sufficient condition for persistency
- Can be verified by solving LP over $\Lambda \supset \delta(\mathcal{X})$
- Tightens with relaxation:
  For $\Lambda' \subset \Lambda$, if p is improving on Λ then it is improving on Λ'.

Generality of Sufficient Conditions

Theorems (Shekhovtsov (2014, 2015))

Relaxed-improving condition with natural (local) relaxations are satisfied for a.o.f.:

pairwise multilabel	Simple DEE (Goldstein, 1994)	✓
	MQPBO (Kohli et al., 2008)	✓
	Kovtun (2003) one-agains-all	✓
	Kovtun (2011) iterative	✓
	Swoboda et al. (2014)*	✓
	Roof dual / QPBO Hammer et al. (1984)	✓
higher order pseudo-Boolean	Reductions: HOCR (Ishikawa, 2011), (Fix et al., 2011)	FLP
	Bisubmodular relaxations (Kolmogorov, 2010)**	BLP
	Generalized Roof Dualilty (Kahl and Strandmark, 2011)	FLP
	Persistency by Adams et al. (1998)	FLP

BLP = Basic LP Relaxation Werner (2007); Thapper and Živný (2013);
FLP = Full Local LP Relaxation, equivalent to Sherali and Adams (1990);

*Swoboda et al. (2014) is higher order but the comparison proof is for pairwise case. **Result holds for sum of bisubmodular functions over the same hypergraph as the BLP relaxation.

Maximizing Persistency

Maximum Persistency

- Given that verification problem is polynomially solvable,
- which method is better?

Maximum Persistency

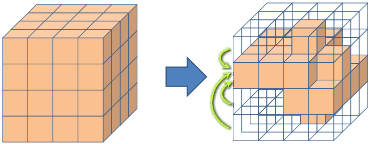
- Given that verification problem is polynomially solvable,
- which method is better?

Maximum Persistency Problem

Find the substitution $p\colon \mathcal{X} \to \mathcal{X}$ that delivers the maximum problem reduction:

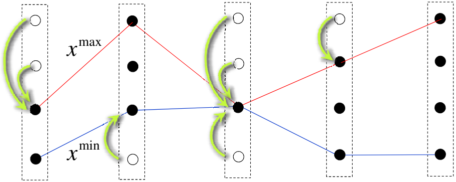
$$\min_{p \in \mathcal{P}} \sum_{u \in \mathcal{V}} |p(\mathcal{X}_u)| \quad \text{s.t. } p \text{ is relaxed-improving,}$$

$\mathcal{P}$ - class of mappings.

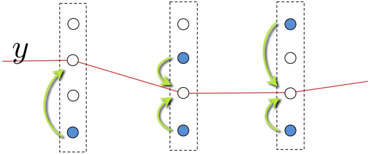


Restricted Class of Mappings

- Clamping to an interval. Order-dependent



- Fix a test labeling y and substitute any subset $\mathcal{Y}_v \subset \mathcal{X}_v$ with y_v. Order independent



- Lattice (nesting) of substitutions in both cases

Restricted Class of Mappings

- Can find the maximum (eliminating most of variables) (strictly) Λ- improving substitution in these cases for any Λ!

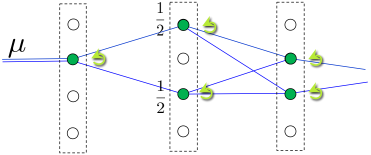
Subsets substituting class covers

pairwise multilabel	Simple DEE (Goldstein, 1994)	
	MQPBO (Kohli et al., 2008)	
	Kovtun (2003) one-agains-all	
	Kovtun (2011) iterative	
	Swoboda et al. (2014)*	
higher order pseudo-Boolean	Roof dual / QPBO Hammer et al. (1984)	
	Reductions: HOCR (Ishikawa, 2011), (Fix et al., 2011)	FLP
	Bisubmodular relaxations (Kolmogorov, 2010)**	BLP
	Generalized Roof Dualilty (Kahl and Strandmark, 2011)	FLP
	Persistency by Adams et al. (1998)	FLP

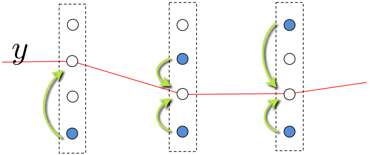
Discrete Cutting Plane

Theorem

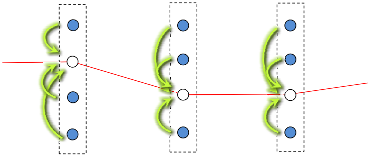
Let μ be a solution to LP-relaxation: $\mu \in \operatorname{argmin}_{\mu\in\Lambda}\langle f, \mu\rangle$ and $p\colon \mathcal{X} \to \mathcal{X}$ be (strictly) relaxed-improving. Then $P\mu = \mu$.



- Initialize test labeling y from μ



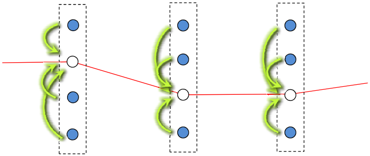
Discrete Cutting Plane



Algorithm

- Start with a mapping p that substitutes *everything* with y

Discrete Cutting Plane

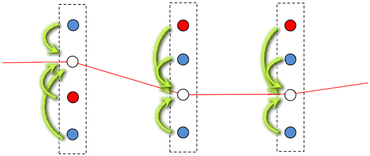


Algorithm

- Start with a mapping p that substitutes *everything* with y
- Auxiliary problem $g = (I - P^{\mathsf{T}})f$
- Check relaxed-improving conditions by solving LP:

$$\min_{\mu \in \Lambda} \langle g, \mu \rangle \overset{?}{\geq} 0$$

Discrete Cutting Plane



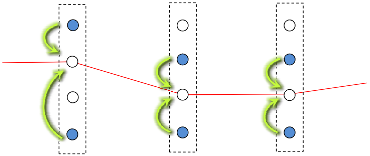
Algorithm

- Start with a mapping p that substitutes *everything* with y
- Auxiliary problem $g = (I - P^{\mathsf{T}})f$
- Check relaxed-improving conditions by solving LP:

$$\min_{\mu \in \Lambda} \langle g, \mu \rangle \quad \overset{?}{\geq} \quad 0$$

- If not satisfied, determine the most violating solution μ

Discrete Cutting Plane



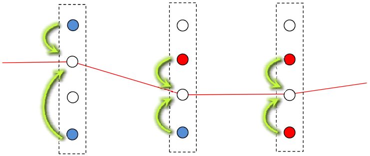
Algorithm

- Start with a mapping p that substitutes *everything* with y
- Auxiliary problem $g = (I - P^{\mathsf{T}})f$
- Check relaxed-improving conditions by solving LP:

$$\min_{\mu \in \Lambda} \langle g, \mu \rangle \quad \overset{?}{\geq} \quad 0$$

- If not satisfied, determine the most violating solution μ
- Make the solution μ immovable by p

Discrete Cutting Plane



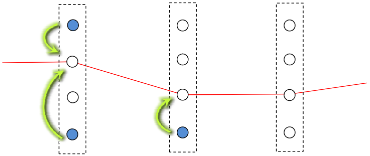
Algorithm

- Start with a mapping p that substitutes *everything* with y
- Auxiliary problem $g = (I - P^{\mathsf{T}})f$
- Check relaxed-improving conditions by solving LP:

$$\min_{\mu \in \Lambda} \langle g, \mu \rangle \quad \overset{?}{\geq} \quad 0$$

- If not satisfied, determine the most violating solution μ
- Make the solution μ immovable by p

Discrete Cutting Plane



Algorithm

- Start with a mapping p that substitutes *everything* with y
- Auxiliary problem $g = (I - P^{\mathsf{T}})f$
- Check relaxed-improving conditions by solving LP:

$$\min_{\mu \in \Lambda} \langle g, \mu \rangle \overset{?}{\geq} 0$$

- If not satisfied, determine the most violating solution μ
- Make the solution μ immovable by p

Correctness and Optimality

Main Properties:

- Runs in polynomial time;
- Finds the maximum relaxed improving substitution when the LP solver is e.g. the interior point method (uses strict complementarity).
- Correct even with sub-optimal (no convergence guarantees) LP solvers
- Correct with dual suboptimal solvers (we use TRW-S by Kolmogorov (2006))
- Can be implemented as incremental

Efficiency

- solving relaxed inference approximately even once is slow
- Fast block-coordinate algorithms TRW-S not finitely converging

How can we iterate such relaxed inference?

Fast implementation with TRW-S

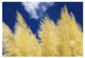
- Incremental: reuse reparametrizations φ
- Guaranteed to prune something even after 1 iteration of TRW-S (there is a blocking constraint not yet pruned)
- An optimal pruning is often possible before the dual is solved (cuts)
- Problem reductions preserving the sufficient condition
- Fast message passing for $(I - P^{\mathsf{T}})f$ with reductions

Combined Effect of Speedups

Instance	Initialization (1000 it.)	Extra time for persistency				
		no speedups	+reduction	+node pruning	+labeling pruning	+fast msgs
Protein folding `1CKK`	8.5s	268s (26.53%)	168s (26.53%)	2.0s (26.53%)	2.0s (26.53%)	2.0s (26.53%)
colorseg-n4 `pfau-small`	9.3s	439s (88.59%)	230s (93.41%)	85s (93.41%)	76s (93.41%)	19s (93.41%)

Experiments

OpenGM Benchmark



ColorSegmentation (N8)
J. Lellmann et.al.
converted by J. Lellmann and J.H. Kappes

Object Segmentation
K. Alahari et.al.
converted by J.H. Kappes

MRF Stereo
R. Szeliski et.al.
converted by J.H. Kappes

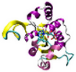
ColorSegmentation
K. Alahari et.al.
converted by J.H. Kappes

MRF Photomontage
R. Szeliski et.al.
converted by J.H. Kappes

MRF Inpainting
R. Szeliski et.al.
converted by J.H. Kappes

Chinese Characters
S. Nowozin et.al.
converted by S. Nowozin and J. H. Kappes

Scene Decomposition
Gould et.al.
converted by S. Nowozin and J. H. Kappes



Protein Folding
Yanover et. al.
converted by Joerg Kappes

Brain 3mm
J. H. Kappes et.al.
converted by J. H. Kappes

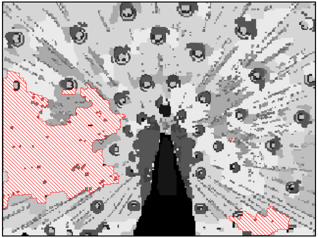
Geometric Surface Labeling (3)
Gallagher et.al.
converted by D. Batra and J. H. Kappes

Problem family	#I	#L	#V	MQPBO		MQPBO-10		Kovtun		[29]-TRWS		Our-TRWS	
mrf-stereo	3	16-60	> 100000	†		†		†		2.5h	13%	117s	**73.56%**
mrf-photomontage	2	5-7	≤ 514080	93s	22%	866s	16%	†		3.7h	16%	483s	**41.98%**
color-seg	3	3-4	≤ 424720	22s	11%	87s	16%	**0.3s**	98%	1.3h	**>99%**	61.8s	**99.95%**
color-seg-n4	9	3-12	≤ 86400	22s	8%	398s	14%	**0.2s**	67%	321s	90%	4.9s	**99.26%**
ProteinFolding	21	≤ 483	≤ 1972	685s	2%	2705s	2%	†		48s	18%	9.2s	**55.70%**
object-seg	5	4-8	68160	3.2s	0.01%	†		**0.1s**	93.86%	138s	98.19%	2.2s	**100%**

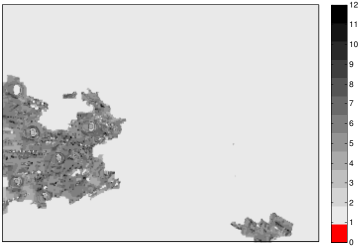
OpenGM Benchmark

Input image (Potts model Color Segmentation)

Proved optimal part



Reminder (number of labels)

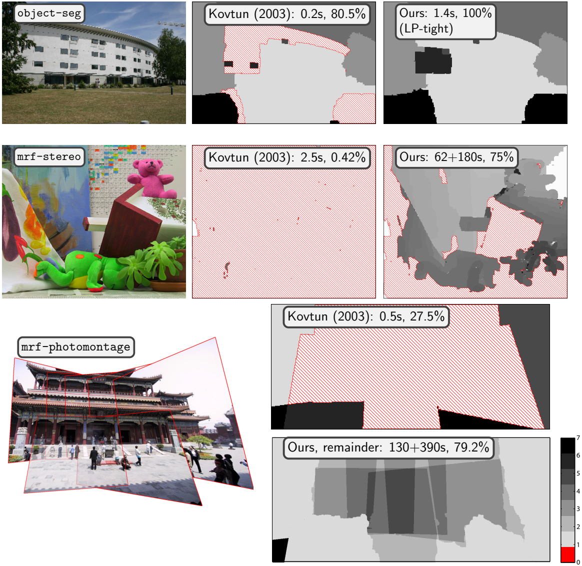


OpenGM Benchmark

Easy

Hard

Very hard



Conclusion

- We find a part of an optimal solution in polynomial time
- New general sufficient condition
- Covers many methods in the literature
- Developed an efficient algorithm (implementation available, matlab interface)
- In a sense, we converted a method without guarantees (TRW-S) into a method with guarantees at a reasonable overhead

M. L. Balinski, "Integer programming: Methods, uses, computations," *Management Science*, vol. 12, no. 3, pp. 253–313, 1965.

L.-C. Lorentzen, "Notes on covering of arcs by nodes in an undirected graph," 1966.

G. L. Nemhauser and L. E. Trotter, Jr., "Vertex packings: Structural properties and algorithms," *Mathematical Programming*, vol. 8, pp. 232–248, 1975.

J.-C. Picard and M. Queyranne, "On the integer-valued variables in the linear vertex packing problem," *Mathematical Programming*, vol. 12, no. 1, pp. 97–101, 1977. [Online]. Available: http://dx.doi.org/10.1007/BF01593772

P. Hammer, P. Hansen, and B. Simeone, "Roof duality, complementation and persistency in quadratic 0-1 optimization," *Mathematical Programming*, pp. 121–155, 1984.

E. Boros, P. L. Hammer, and X. Sun, "Network flows and minimization of quadratic pseudo-Boolean functions," RUTCOR, Tech. Rep. RRR 17-1991, May 1991.

E. Boros and P. Hammer, "Pseudo-Boolean optimization," RUTCOR, Tech. Rep., October 2001.

H. Ishikawa, "Transformation of general binary MRF minimization to the first-order case," *PAMI*, vol. 33, no. 6, pp. 1234–1249, June 2011.

A. Fix, A. Gruber, E. Boros, and R. Zabih, "A graph cut algorithm for higher-order Markov random fields," in *ICCV*, 2011, pp. 1020–1027.

V. Kolmogorov, "Generalized roof duality and bisubmodular functions," *Discrete Applied Mathematics*, vol. 160, no. 4-5, pp. 416–426, 2012.

H. D. Sherali and W. P. Adams, "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems," *SIAM J. Discrete Math.*, vol. 3, no. 3, pp. 411–430, 1990.

R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *PAMI*, vol. 30, no. 6, pp. 1068–1080, 2008.

A. Shekhovtsov, "Maximum persistency in energy minimization," in *CVPR*, June 2014, pp. 1162–1169.

——, "Higher order maximum persistency and comparison theorems," *CVIU*, 2015.

R. F. Goldstein, "Efficient rotamer elimination applied to protein side-chains and related spin glasses," *Biophysical Journal*, vol. 66, no. 5, pp. 1335–1340, May 1994.

P. Kohli, A. Shekhovtsov, C. Rother, V. Kolmogorov, and P. Torr, "On partial optimality in multi-label MRFs," in *ICML*, A. McCallum and S. Roweis, Eds. Omnipress, 2008, pp. 480–487.

I. Kovtun, "Partial optimal labeling search for a NP-hard subclass of (max, +) problems," in *DAGM-Symposium*, 2003, pp. 402–409.

——, "Sufficient condition for partial optimality for (max, +) labeling problems and its usage," *Control Systems and Computers*, no. 2, 2011, special issue.

P. Swoboda, B. Savchynskyy, J. Kappes, and C. Schnrr, "Partial optimality by pruning for MAP-inference with general graphical models," in *CVPR*, June 2014, pp. 1170–1177.

V. Kolmogorov, "Generalized roof duality and bisubmodular functions," in *NIPS*, 2010, pp. 1144–1152.

F. Kahl and P. Strandmark, "Generalized roof duality for pseudo-Boolean optimization," in *ICCV*, 2011, pp. 255–262.

W. P. Adams, J. B. Lassiter, and H. D. Sherali, "Persistency in 0-1 polynomial programming," *Mathematics of Operations Research*, vol. 23, no. 2, pp. 359–389, Feb. 1998.

T. Werner, "A linear programming approach to max-sum problem: A review," *PAMI*, vol. 29, no. 7, pp. 1165–1179, July 2007.

J. Thapper and S. Živný, "The complexity of finite-valued CSPs," in *Symposium on the Theory of Computing (STOC)*. ACM, 2013, pp. 695–704. [Online]. Available: http://www.cs.ox.ac.uk/Stanislav.Zivny/homepage/publications/tz13stoc-preprint.pdf

V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *PAMI*, vol. 28(10), pp. 1568–1583, october 2006.

A. Shekhovtsov and V. Hlaváč, "On partial opimality by auxiliary submodular problems," *Control Systems and Computers*, no. 2, 2011, special issue.

A. Shekhovtsov, V. Kolmogorov, P. Kohli, V. Hlavac, C. Rother, and P. Torr, "LP-relaxation of binarized energy minimization," Czech Technical University, Tech. Rep. CTU–CMP–2007–27, 2008.

P. Swoboda, B. Savchynskyy, J. Kappes, and C. Schnörr, "Partial optimality via iterative pruning for the Potts model," in *SSVM*, 2013.

K. Alahari, P. Kohli, and P. H. S. Torr, "Reduce, reuse & recycle: Efficiently solving multi-label MRFs," in *CVPR*, 2008.

M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends in Machine Learning*, vol. 1, no. 1-2, pp. 1–305, 2008.

A. M. Koster, S. Hoesel, and A. Kolen, "The partial constraint satisfaction problem: Facets and lifting theorems," *Operations Research Letters*, vol. 23, pp. 89–97(9), 1998.

C. Chekuri, S. Khanna, J. Naor, and L. Zosin, "Approximation algorithms for the metric labeling problem via a new linear programming formulation," in *In Symposium on Discrete Algorithms*, 2001, pp. 109–118.

C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer, "Optimizing binary MRFs via extended roof duality," in *CVPR*, 2007.

I. Kovtun, "Image segmentation based on sufficient conditions of optimality in NP-complete classes of structural labelling problem," Ph.D. dissertation, IRTC ITS National Academy of Sciences, Ukraine, 2004, in Ukrainian.

E. Boros and P. Hammer, "Pseudo-Boolean optimization," *Discrete Applied Mathematics*, vol. 1-3, no. 123, pp. 155–225, 2002.

J. Desmet, M. D. Maeyer, B. Hazes, and I. Lasters, "The dead-end elimination theorem and its use in protein side-chain positioning," *Nature*, vol. 356, pp. 539–542, 1992.

P. Orponen and H. Mannila, "On approximation preserving reductions: Complete problems and robust measures (revised version)," Tech. Rep., 1990.