# A dual ascent framework for Lagrangean decomposition of combinatorial problems


Paul Swoboda[1] Jan Kuske[2] Bogdan Savchynskyy[3]

[1] IST Austria [2] Heidelberg University, Germany [3] TU Dresden, Germany

e-mail: pswoboda@ist.ac.at


## Goal and contribution

**Goal:** Efficient dual block coordinate ascent solvers for a wide range of large scale combinatorial problems from computer vision and beyond.

**Contribution:**

- We propose a new class of general LP-relaxations of ILP-problems, called **I**nteger **R**elaxed **P**airwise **S**eparable **L**inear **P**rograms (IRPS-LP).
  - (IRPS-LP) generalize the local polytope relaxation for CRFs.
  - (IRPS-LP) allows explicit modelling of allowed configurations, instead of forbidding them with $\infty$-costs. This leads to more compact formulations and lower computational cost.
  - Relaxations of the multicut [4] (poster #75) and graph matching [5] (poster #69 tommorow afternoon) problems can be written as (IRPS-LP).
- We analyze dual block coordinate ascent for (IRPS-LP).
  - We prove monotonical improvement of a dual lower bound.
  - Popular algorithms TRWS [2], SRMP [3] and MPLP [1] for inference in CRFs are special cases.

**Practical impact:**

- Efficient algorithms for multicut and graph matching have been written in our framework (for extensive details see corresponding posters).
- C++ implementation available at `https://github.com/pawelswoboda/LP_MP`.

## Problem formulation (IRPS-LP)

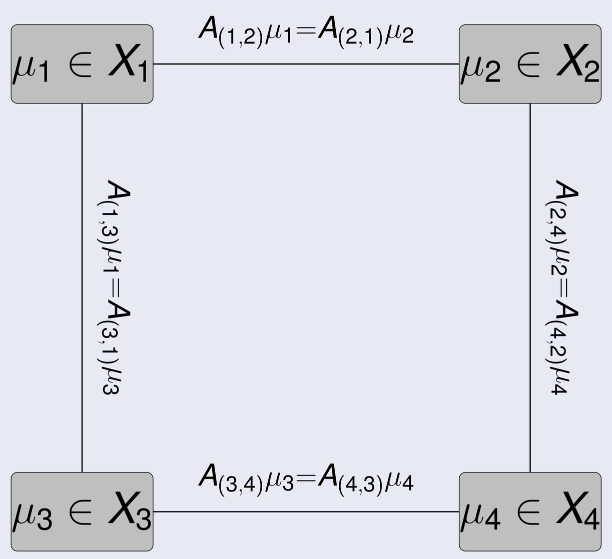
**Factor graph** $\mathbb{G} = (\mathbb{F}, \mathbb{E})$:

- *factors* $i \in \mathbb{F}$ with variables $\mu_i \in X_i \subseteq \{0,1\}^{d_i}$
- *couplings* $ij \in \mathbb{E} \subseteq \binom{\mathbb{F}}{2}$ with constraints $A_{(i,j)}\mu_i = A_{(j,i)}\mu_j$.

**(IRPS-LP)**:

$$\min_{\mu \in \Lambda_{\mathbb{G}}} \sum_{i=1}^{k} \langle \theta_i, \mu_i \rangle$$

$$\Lambda_{\mathbb{G}} := \left\{ (\mu_1 \dots \mu_k) \,\middle|\, \begin{array}{ll} \mu_i \in \mathrm{conv}(X_i) & \forall i \in \mathbb{F} \\ A_{(i,j)}\mu_i = A_{(j,i)}\mu_j & \forall ij \in \mathbb{E} \end{array} \right\}.$$




## Example: MAP-inference in CRFs as (IRPS-LP)

**CRF**:

- Graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$,
- label space $X = \prod_{u \in \mathrm{V}} X_u$,
- $\forall u \in \mathrm{V}$: unary costs $\theta_u : X_u \to \mathbb{R}$,
- $\forall uv \in \mathrm{E}$: pairwise costs $\theta_{uv} : X_u \times X_v \to \mathbb{R}$.

**MAP-inference**:

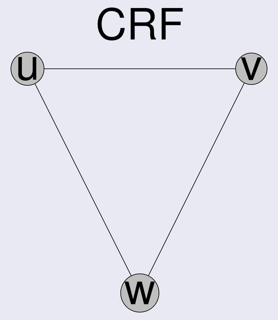
$$\min_{x \in X} \sum_{u \in \mathrm{V}} \theta_u(x_u) + \sum_{uv \in \mathrm{E}} \theta_{uv}(x_{uv}).$$

**(IRPS-LP)**: $\mathbb{F} = \mathrm{V} \cup \mathrm{E}$, $\mathbb{E} = \{\{u, uv\}, \{v, uv\} : uv \in \mathrm{E}\}$. Simplex: $\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i) = 1\}$.

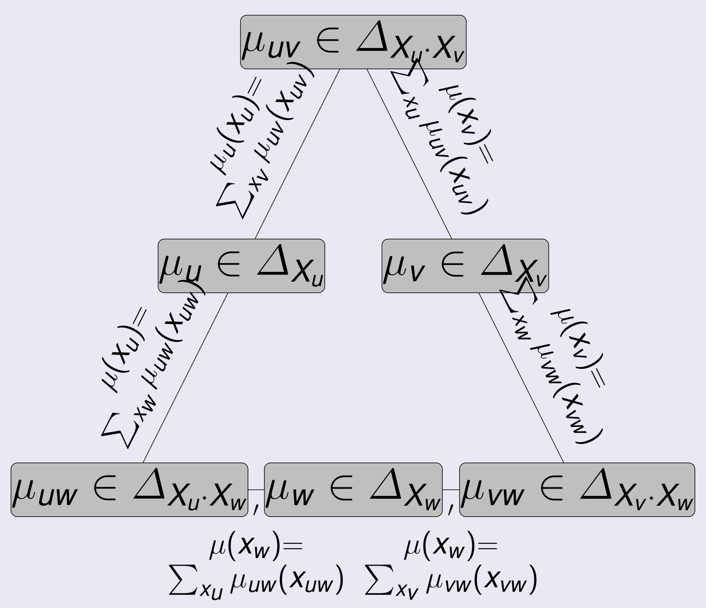
$$\min_{\mu \in \mathrm{L}_\mathrm{G}} \langle \theta, \mu \rangle := \sum_{u \in \mathrm{V}} \langle \theta_u, \mu_u \rangle + \sum_{uv \in \mathrm{E}} \langle \theta_{uv}, \mu_{uv} \rangle$$

$$\mathrm{L}_\mathrm{G} = \left\{ \begin{array}{ll} \underline{\mu_u \in \mathrm{conv}(X_u)}: & \mu_u \in \Delta_{|X_u|}, u \in \mathrm{V} \\ \underline{\mu_{uv} \in \mathrm{conv}(X_{uv})}: & \mu_{uv} \in \Delta_{|X_{uv}|}, uv \in \mathrm{E} \\ \underline{A_{(uv,u)}\mu_{uv} = A_{(u,uv)}\mu_u}: & \sum_{x_v \in X_v} \mu_{uv}(x_u, x_v) = \mu_u(x_u). \end{array} \right.$$

CRF




CRF as (IRPS-LP)

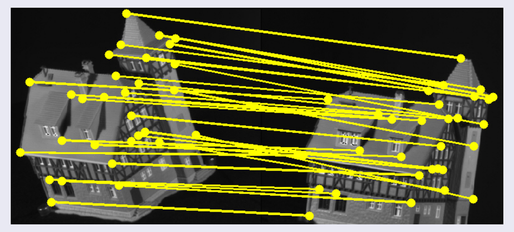



## Example: graph matching

**Given**:

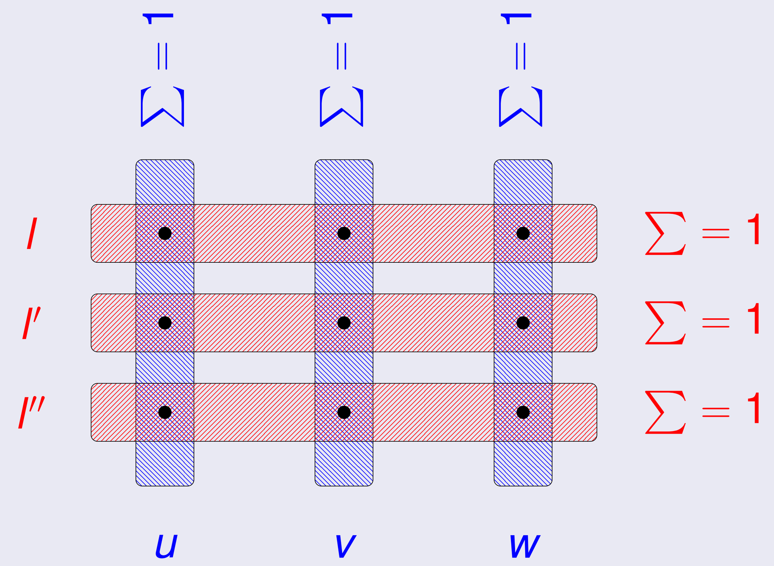
- CRF.
- Label spaces $X_u \subset \mathcal{L}$ are part of an *universe*.

$$\underbrace{\min_{x \in X} \sum_{u \in \mathrm{V}} \theta_u(x_u) + \sum_{uv \in \mathrm{E}} \theta_{uv}(x_{uv})}_{\text{CRF}} \quad \text{s.t.} \quad \underbrace{x_u \neq x_v \quad \forall u \neq v}_{\text{uniqueness constraints}}$$



**Example**: $V = \{u, v, w\}$, $\mathcal{L} = \{l, l', l''\}$.

- **Unary factors**: Each node takes a label.
- **Label factors**: Each label can be taken once.




**(IRPS-LP)**:

$$\min_{\mu \in \mathrm{L}_\mathrm{G}, \tilde{\mu}} \langle \theta, \mu \rangle$$

$$\text{s.t. } \tilde{\mu} \in \left\{ \forall s \in \mathcal{L} : \sum_{u \in \mathrm{V}} \tilde{\mu}_u(s) \leq 1 \right\}$$

additional (IRPS-LP) factors for uniqueness constraint

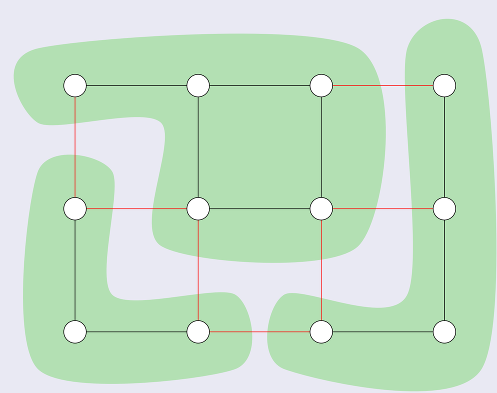
$\mu_u(s) = \tilde{\mu}_s(u)$, $s \in X_u$ couple CRF and uniqueness constraints.

## Example: multicut

**Optimization problem**:

- Weighted graph $G = (V, E)$.
- Find partition of $G$.

$$\min_{x \in \{0,1\}^{|E|}} \sum_{e \in E} \theta_e x_e, \quad \text{s.t. } \forall \text{ cycles C} : \underbrace{\forall e' \in C : \sum_{e \in C \setminus \{e'\}} x_e \geq x_{e'}}_{\text{cycle inequalities}}.$$

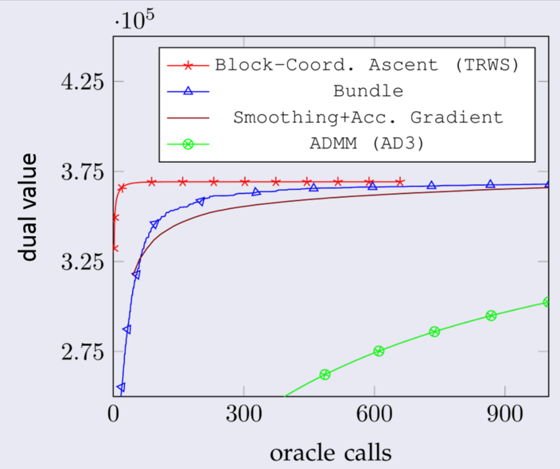


**(IRPS-LP)**:

- Factors: $\mathbb{F} = E \cup$ 3-cycle$(E)$.
- Couplings: $\mathrm{E} = \{\{e, C\} : C \in \text{3-cycle}(E), e \in C\}$

## Motivation: Dual block coordinate ascent for CRFs

- **Observation**: TRWS [2] is state of the art among LP-based methods for MAP-inference for CRFs:
- **Contribution**: Develop dual block coordinate ascent methods inspired by TRWS [2] for (IRPS-LP).




## Dual lower bound

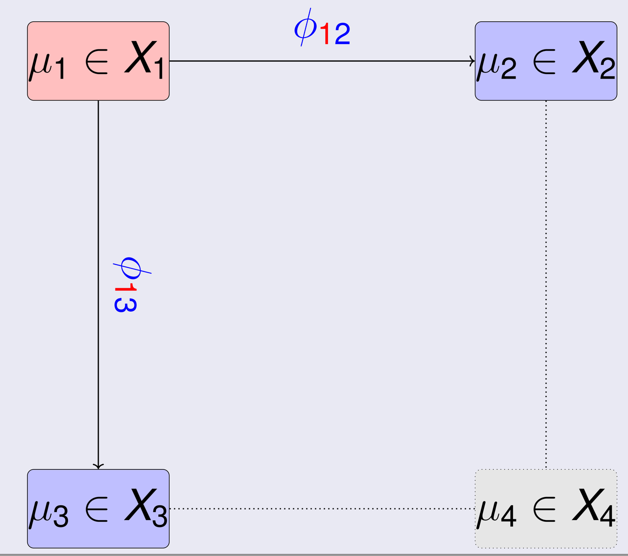
Dualize primal problem w.r.t. coupling constraints:

- Primal variables $P := \left\{ \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} : \mu \in \begin{pmatrix} \mathrm{conv}(X_1) \\ \vdots \\ \mathrm{conv}(X_k) \end{pmatrix} \right\}$.
- Coupling constraints: $A\mu = (A_{(i,j)}\mu_i - A_{(j,i)}\mu_j)_{ij \in \mathbb{E}} = 0$.
- Maximize dual lower bound: $\max_\phi [D(\phi) := \min_{\mu \in P} \langle \theta, \mu \rangle + \langle \phi, A\mu \rangle]$.
- Reparametrization: $\theta^\phi := \theta + A^\top \phi$.

## Elementary updates – Admissible messages

- For factor $i \in BF$ take subset of neighbors $N \subseteq \{j : ij \in \mathbb{E}\}$.
- Consider dual variables $\phi_{ij}$ related to couplings $j \in N$.
- Change $(\phi_{ij})_{j \in N}$ by $\Delta$ such that $D(\phi - \Delta) \geq D(\phi)$.

Illustration in the figure:
factor 1, neighbors $N = \{2, 3\}$.



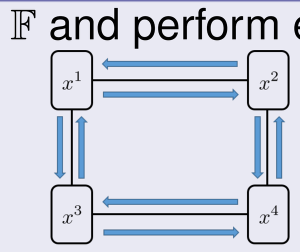

## Elementary updates – Admissible messages

**Computation of** $\Delta$:

- $\hat{x}_i \in \arg\min_{x_i \in X_i} \langle \theta^\phi_i, x_i \rangle$ : opt. configuration before update.
- $\hat{x}_i \in \arg\min_{x_i \in X_i} \langle \theta^{\phi - \Delta}_i, x_i \rangle$ : opt. configuration after update.
- $\langle \theta^{\phi-\Delta}_i, \hat{x}_i \rangle \geq \langle \theta^\phi_i, \hat{x}_i \rangle$ : costs do not decrease.
- $\Delta = \arg\max_{\Delta\text{–admissible}} \langle \delta, \theta^{\phi-\Delta}_i \rangle$, $\delta(s) \begin{cases} > 0, & \hat{x}_i(s) = 1 \\ < 0, & \hat{x}_i(s) = 0 \end{cases}$

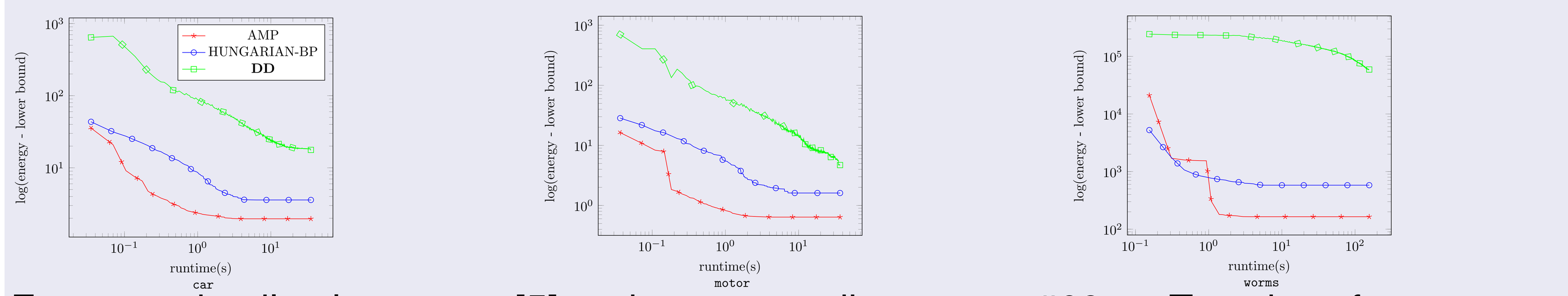
## Algorithm

Visit all factors $i \in \mathbb{F}$ and perform elementary updates:



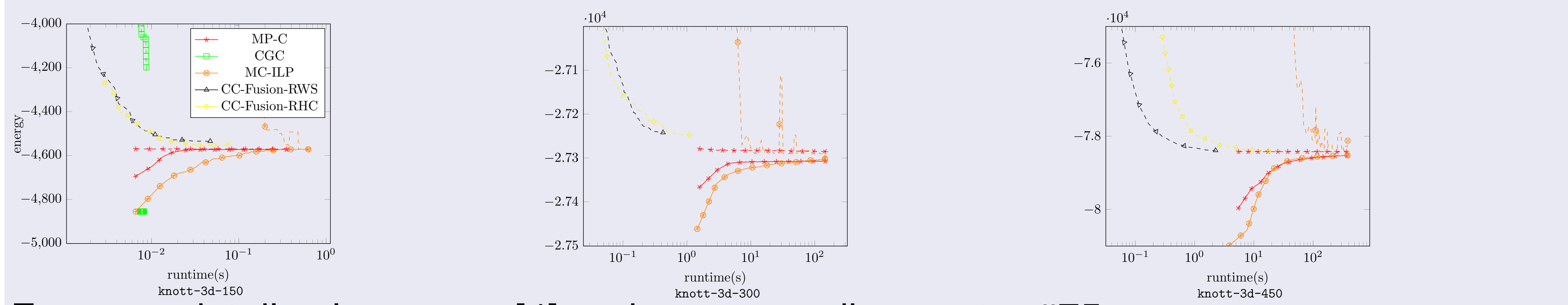
Special cases: TRWS [2], SRMP [3] and MPLP [1] for MAP-inference in CRFs.

## Results: graph matching




For more details please see [5] and corresponding poster #69 on Tuesday afternoon.

## Results: multicut




For more details please see [4] and corresponding poster #75.

## References


[1] A. Globerson and T. S. Jaakkola. Fixing max-product: Convergent message passing algorithms for MAP LP-relaxations. In *NIPS*, pages 553–560, 2007.

[2] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(10):1568–1583, 2006.

[3] V. Kolmogorov. A new look at reweighted message passing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(5):919–930, 2015.

[4] P. Swoboda and B. Andres. A message passing algorithm for the minimum cost multicut problem. In *CVPR*, 2017.

[5] P. Swoboda, C. Rother, H. Abu Alhaija, D. Kainmueller, and B. Savchynskyy. Study of Lagrangean decomposition and dual ascent solvers for graph matching. In *CVPR*, 2017.



**Acknowledgements.** The first author was supported by the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no 616160, the second author by DFG Grant GRK 1653 and the last author by the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No 647769) and the DFG Grant "ERBI" SA 2640/1-1.