# Appendix

## Contents



## A. Proofs of the Generic Algorithms

**Proposition** 2.2 "Let $p$ be a strictly improving mapping. Then any optimal solution $x^*$ of (1) must satisfy $p_v(x^*_v) = x^*_v$, $v \in \mathcal{V}$".

*Proof.* Let $p$ be a strictly improving mapping and $x^* \in \arg\min_{x\in\mathcal{X}} \langle f, \delta(x)\rangle$ such that $p(x) \neq x$. Then from Definition 2.1 follows $\langle f, \delta(p(x^*))\rangle < \langle f, \delta(x^*)\rangle$, which contradicts optimality of $x^*$. □

**Proposition** 2.5 "If mapping $p$ is strictly $\Lambda$-improving then it is strictly improving."

*Proof.* From Definition 2.4 follows that for all $\mu \in \Lambda$ such that $[p]\mu \neq \mu$ there holds $\langle f, [p]\mu\rangle < \langle f, \mu\rangle$. Since $\Lambda \supseteq \mathcal{M}$ it holds also for $\mu = \delta(x)$ for all $x \in \mathcal{X}$, which proves the proposition. □

**Proposition** 2.6: "$p \in \mathbb{S}_f$ iff $(\forall v \in \mathcal{V} \;\; \forall i \in \mathcal{O}^*_v) \;\; p_v(i) = i$".

*Proof.* Let $p \in \mathbb{S}_f$. Assume for contradiction that $(\exists v \in \mathcal{V} \;\; \exists i \in \mathcal{O}^*_v) \;\; p_v(i) \neq i$. Since $i \in \mathcal{O}^*_v$ there exists $\mu \in \mathcal{O}^*$ such that $\mu_v(i) > 0$. It's image $\mu' = [p]\mu$ has $\mu_v(i) = 0$ due to $p_v(i) \neq i$ by evaluating the extension (3). This contradicts to $[p]\mu = \mu$.

Now let $(\forall v \in \mathcal{V} \;\; \forall i \in \mathcal{O}^*_v) \;\; p_v(i) = i$. Clearly, $[p]\mu = \mu$ holds for all $\mu$ on the support set given by $(\mathcal{O}^*_v \mid v \in \mathcal{V})$, hence for $\mathcal{O}^*$. It remains to show that the value of minimum in (6) is zero. For $\mu \in \mathcal{O}^*$ we have $[p]\mu = \mu$ and the objective in (6), $\langle (I - [p])^\mathsf{T} f, \mu\rangle = \langle f, \mu - [p]\mu\rangle$ vanishes. □

**Proposition** 2.7: "Mapping $p \in \mathcal{P}^{2,y}$ is strictly $\Lambda$-improving for the cost vector $f \in \mathbb{R}^\mathcal{I}$ iff there holds $(\forall v \in \mathcal{V}) \;\; \mathcal{O}^*_v \cap \mathcal{Y}_v = \varnothing$".

*Proof.* Follows by Definition 2.4 and Proposition 2.6. □

**Proposition** 3.1: "Algorithm 1 is polynomial and returns a strictly $\Lambda$-improving mapping $p \in \mathbb{S}_f \cap \mathcal{P}^{2,y}$".

*Proof.* Solving the verification LP in every iteration as well as finding the support sets of all optimal solutions $\mathcal{O}^*_v$ is polynomial. These sets equal to the support set of any strict relative interior optimal solution, *i.e.*, a solution found by an interior point method, see, *e.g.*, [32].

At every iteration, if the algorithm has not terminated yet, at least one of the sets $\mathcal{Y}_v$ strictly shrinks in line 9. Therefore the algorithm terminates in at most $\sum_v (|\mathcal{X}_v| - 1)$ iterations. On termination $p \in \mathbb{S}_f$ by Proposition (2.7). □

**Theorem** 3.2: "Mapping $p$ returned by Algorithm 1 is the maximum of $\mathbb{S}_f \cap \mathcal{P}^{2,y}$".

*Proof.* Two following lemmas form a basis for the proof.

**Lemma A.1** (special case of [26], Theorem 3(b))**.** Let $q \in \mathbb{S}_f$; $p \geq q$ and $\mathcal{O}^*$ is the set of optimal relaxed labelings, $\operatorname{argmin}_{\mu\in\Lambda} \langle (I - [p])^\top f, \mu\rangle$, *i.e.*, as in line 5 of Algorithm 1. Then $(\forall \mu \in \mathcal{O}^*) \;\; q(\mu) = \mu$.

Additionally, if $q \in \mathcal{P}^{2,y}$ and $\mathcal{O}^*_v = \{i \in \mathcal{X}_v \mid (\exists \mu \in \mathcal{O}^*) \; \mu_v(i) > 0\}$ then

$$(\forall v \in \mathcal{V}, \forall i \in \mathcal{O}^*_v) \;\; q_v(i) = i. \tag{16}$$

*Proof.* We prove the additional part. It follows similarly to Proposition 2.7. Assume for contradiction that $(\exists v \in \mathcal{V} \;\; \exists i \in \mathcal{O}^*_v) \;\; q_v(i) \neq i$. Since $i \in \mathcal{O}^*_v$ there exists $\mu \in \mathcal{O}^*$ such that $\mu_v(i) > 0$. It's image $\mu' = [q]\mu$ has $\mu_v(i) = 0$ due to $q_v(i) \neq i$ by evaluating the extension (3). This contradicts to $[q]\mu = \mu$, the direct statement of [26, Theorem 3(b)]. □

**Lemma A.2.** Algorithm 1 maintains invariant that $(\forall q \in \mathbb{S}_f \cap \mathcal{P}^{2,y}) \;\; p \geq q$.

*Proof.* We prove by induction. The statement holds trivially for the first iteration. Assume it is true for the current iteration $t$. Let $p^t$ denote the mapping $p$ computed in line 3 on iteration $t$. Then for any $q \in \mathbb{S}_f \cap \mathcal{P}^{2,y}$ holds $p^t \geq q$ and therefore Lemma A.1 applies. We show that line 9 only prunes maps that are not in $\mathcal{P}^{2,y} \cap \mathbb{S}_f$ as follows.

Let $p^{t+1}$ be the mapping on the next iteration, *i.e.* computed by line 3 after pruning line 9.

Assume for contradiction that $\exists q \in \mathcal{P}^{2,y} \cap \mathbb{S}_f$ such that $p^{t+1} \ngeq q$. By negating the definition and expanding,

$$(\exists v \in \mathcal{V}) \;\; p^{t+1}_v(\mathcal{X}_v) \neq q_v(\mathcal{X}_v), \tag{17a}$$

$$(\exists v \in \mathcal{V} \; \exists i \in \mathcal{X}_v) \;\; i \in p^{t+1}_v(\mathcal{X}_v) \;\; \wedge \;\; i \notin q_v(\mathcal{X}_v), \tag{17b}$$

$$(\exists v \in \mathcal{V} \; \exists i \in \mathcal{X}_v) \;\; p^{t+1}_v(i) = i \;\; \wedge \;\; q_v(i) \neq i. \tag{17c}$$

If $i$ was pruned in line 9, $i \in \mathcal{O}^*_v$, then it must be that

$q_v(i) = i$, which contradicts to (17c). Therefore

$$(\exists v \in \mathcal{V}\ \exists i \in \mathcal{X}_v \backslash \mathcal{O}_v^*)\ \ p_v^{t+1}(i) = i \ \ \wedge \ \ q_v(i) \neq i. \quad (18)$$

However, in this case $p_v^{t+1}(i) = p_v^t(i) = i$ and $p^t \geq q$ fails to hold, which contradicts to the assumption of induction. Therefore $p^{t+1} \geq q$ holds by induction on every iteration. □

By Proposition 3.1 the algorithm terminates and returns a map in $\mathbb{S}_f \cap \mathcal{P}^{2,y}$. By Lemma A.2 the returned map $p$ satisfies $p \geq q$ for all $q \in \mathbb{S}_f \cap \mathcal{P}^{2,y}$. Therefore $q$ is the element of $\mathbb{S}_f \cap \mathcal{P}^{2,y}$ which is larger or equal to any other element of this set. It is the maximum. □

**Proposition** 4.2: "Let $\mathcal{O}_v(\varphi) = \mathcal{O}_v^*$ for all $v \in \mathcal{V}$. Then $f^\varphi$ is arc consistent".

*Proof.* Condition $\mathcal{O}_v(\varphi) = \mathcal{O}_v^*$ implies that $\varphi$ satisfies strict complementarity with some primal optimal solution $\mu$. The strict complementarity implies that $(\forall i \in \mathcal{X}_v)$ $(f_u^\varphi(i) = 0 \Rightarrow \mu_u(i) > 0)$. By feasibility of $\mu$, there must hold $(\forall v \in \mathrm{nb}(u))$ $(\exists j \in \mathcal{X}_v)$ $\mu_{uv}(i,j) > 0$. And by using complementary slackness again, it must be that $f_{uv}^\varphi(i,j) = 0$. Similarly, the second condition of arc consistency is verified. It follows that $f^\varphi$ is arc consistent. □

**Proposition** 4.3: "Algorithm 2 terminates in a finite number of iterations and delivers a strictly $\Lambda$-improving mapping $p \in \mathbb{S}_f \cap \mathcal{P}^{2,y}$".

*Proof.* Since $\mathcal{O}_v^* \subseteq \mathcal{O}_v(\varphi)$, Algorithm 2 prunes a superset of maps pruned by Algorithm 1. Algorithm 2 is finite, because in case the termination condition in line 5 is not satisfied at least one of the sets $\mathcal{Y}_v$ shrinks in line (9). Since $\mathcal{O}_v^* \subseteq \mathcal{O}_v(\varphi)$ from the stopping condition $(\forall v \in \mathcal{V})$ $\mathcal{O}_v(\varphi) \cap \mathcal{Y}_v = \varnothing$ follows $(\forall v \in \mathcal{V})$ $\mathcal{O}_v^* \cap \mathcal{Y}_v = \varnothing$, which is sufficient for $p$ to be strictly $\Lambda$-improving according to Proposition 2.7. □

## B. Proofs of the Reduction

In this section we prove the reduction Theorem 5.1 and the correctness of heuristics (pruning of negative labelings and individual nodes) that are based on Lemma 5.2. These theorems revisit the key elements on which Algorithm 1 builds: the verification LP (2.7), Proposition 2.6 and Lemma A.1, which either certifies $p \in \mathbb{S}_f$ or allows to make pruning. A correct pruning can be done when we have a guarantee to preserve all strictly improving maps $q$, assuming $q \leq p$. Therefore theorems in this section are formulated for such pairs.

The proof chain considers adjustments to the cost vector that preserve the set of mappings that are strictly improving for the cost vector. These adjustments do not in general preserve optimal solutions to the associated LP relaxation. In that the reduction is different from an equivalent transformation.

**Theorem B.1.** Let $q \leq p$. Then $q \in \mathbb{S}_f$ iff $q \in \mathbb{S}_g$ for $g = (I - [p])^\mathsf{T} f$.

*Proof.* Let $Q = [q]$, $P = [p]$. Because $q \leq p$ there holds $PQ = P$. It implies $(I-P)(I-Q) = (I-Q)$. Therefore,

$$\langle g, (I-Q)\mu \rangle = \langle (I-P)^\mathsf{T} f, (I-Q)\mu \rangle \quad (19)$$
$$= \langle f, (I-P)(I-Q)\mu \rangle = \langle f, (I-Q)\mu \rangle. \quad (20)$$

Assume $\mu \in \Lambda$ is such that $Q\mu \neq \mu$. Then equality (19) ensures that $\langle g, (I-Q)\mu \rangle > 0$ iff $\langle f, (I-Q)\mu \rangle > 0$. The theorem follows from definition of $\mathbb{S}_f$, $\mathbb{S}_g$. □

The reduction that we further build is based on the following result.

**Theorem B.2** (Characterization [26])**.** Let $P = [p]$, $p$ idempotent component-wise. Then

$$(\forall \mu \in \Lambda)\ \langle f, P\mu \rangle \leq \langle f, \mu \rangle \quad (21)$$

iff exists reparametrization $\varphi$ such that

$$P^\mathsf{T} f^\varphi \leq f^\varphi. \quad (22)$$

The reduction in Theorem 5.1 is proven in several steps. The following theorem assumes arbitrary mapping $q$, not necessarily in $\mathcal{P}^{2,y}$ and take as input sets $U_u$ that are subsets of *immovable* labels. In the context of Theorem 5.1, we will use $U_u = \mathcal{X}_u \backslash \mathcal{Y}_u$.

**Theorem B.3** (Reduction 1)**.** Let $q\colon \mathcal{X} \to \mathcal{X}$ component-wise idempotent. Let $U_u \subseteq \{i \in \mathcal{X}_u \mid q_u(i) = i\}$ for $u \in \mathcal{V}$. Let $g_{uv}(i,j) = 0$ for $(i,j) \in U_u \times U_v$. Let $\bar{g}$ be defined by

$$\bar{g}_v = g_v,\ v \in \mathcal{V}; \quad (23a)$$

$$\bar{g}_{uv}(i,j) = \begin{cases} \min\limits_{i' \in U_u} g_{uv}(i',j), & i \in U_u, j \notin U_v\,, \\ \min\limits_{j' \in U_v} g_{uv}(i,j'), & i \notin U_u, j \in U_v\,, \\ g_{uv}(i,j), & \text{otherwise.} \end{cases} \quad (23b)$$

Then $q \in \mathbb{S}_g$ iff $q \in \mathbb{S}_{\bar{g}}$.

*Proof.* Note, unary components of $g$ and $\bar{g}$ are equal. We will prove by separate pairwise components.

Direction $\Leftarrow$. Let us prove the following inequality:

$$(\forall ij \in \mathcal{X}_{uv})\ g_{uv}(q_u(i), q_v(j)) - \bar{g}_{uv}(q_u(i), q_v(j)) \leq g_{uv}(i,j) - \bar{g}_{uv}(i,j). \quad (24)$$

We need to consider only cases where $\bar{g}_{uv}(i,j) \neq g_{uv}(i,j)$. Let $i \in U_u$ and $j \notin U_v$ (the remaining case is symmetric). In this case $q_u(i) = i$. Substituting $\bar{g}$ we have to prove

$$g_{uv}(i, q_v(j)) - \min_{i' \in U_u} g_{uv}(i', q_v(j)) \quad (25)$$
$$\leq g_{uv}(i,j) - \min_{i' \in U_u} g_{uv}(i',j).$$

Clearly, LHS is zero by assumption about $g$. At the same time RHS is non-negative since $i \in U_s$. The inequality holds. Inequality (24) implies (by multiplication of pairwise inequalities and unary inequalities with respective components of $\mu$ and summing) that

$$(\forall \mu \in \Lambda)\ \langle g, Q\mu \rangle - \langle \bar{g}, Q\mu \rangle \leq \langle g, \mu \rangle - \langle \bar{g}, \mu \rangle. \quad (26)$$

Inequality (26) is equivalent to

$$\langle \bar{g}, (I - Q)\mu \rangle \leq \langle g, (I - Q)\mu \rangle. \quad (27)$$

Whenever the LHS of (27) is strictly positive then so is the RHS and therefore from $q \in \mathbb{S}_{\bar{g}}$ follows $q \in \mathbb{S}_g$.

Direction $\Rightarrow$. Assume $q \in \mathbb{S}_g$. By Theorem B.2, there exist dual multipliers $\varphi$ such that for $g' := g - A^{\mathsf{T}}\varphi$ the following component-wise inequalities hold:

$$\begin{aligned} &(\forall u \in \mathcal{V},\ \forall i \in \mathcal{X}_u)\ g'_u(q_u(i)) \leq g'_u(i); \quad (28) \\ &(\forall uv \in \mathcal{E},\ \forall ij \in \mathcal{X}_{uv})\ g'_{uv}(q_u(i), q_v(j)) \leq g'_{uv}(i,j). \end{aligned}$$

Let us expand the pairwise inequality in the case $i \in U_u$, $j \notin U_v$. Let $q_v(j) = j^*$. Using $q_u(i) = i$ we obtain

$$\begin{aligned} &g_{uv}(i, j^*) - \varphi_{uv}(i) - \varphi_{vu}(j^*) \quad (29) \\ \leq &g_{uv}(i,j) - \varphi_{uv}(i) - \varphi_{vu}(j). \end{aligned}$$

Terms $\varphi_{uv}(i)$ cancel:

$$g_{uv}(i, j^*) - \varphi_{vu}(j^*) \leq g_{uv}(i,j) - \varphi_{vu}(j). \quad (30)$$

We take $\min$ over $i \in U_u$ of both sides:

$$\min_{i \in U_u} g_{uv}(i, j^*) - \varphi_{vu}(j^*) \leq \min_{i \in U_u} g_{uv}(i,j) - \varphi_{vu}(j). \quad (31)$$

Finnaly we can subtract $\varphi_{uv}(i)$ from both sides to obtain

$$\bar{g}^{\varphi}_{uv}(i, j^*) \leq \bar{g}^{\varphi}_{uv}(i,j). \quad (32)$$

The case when $i \notin U_u$, $j \in U_v$ is symmetric. In the remaining cases, $\bar{g}^{\varphi}_{uv}(i,j) = \bar{g}(i,j) - \varphi_{uv}(i) - \varphi_{vu}(j) = g(i,j) - \varphi_{uv}(i) - \varphi_{vu}(j) = g^{\varphi}(i,j)$. In total, $\bar{g}^{\varphi}$ satisfies all component-wise inequalities as $g'$ in (28). By Theorem B.2,

$$(\forall \mu \in \Lambda)\ \langle \bar{g}, Q\mu \rangle \leq \langle \bar{g}, \mu \rangle. \quad (33)$$

We have shown that $\langle \bar{g}, (I - Q)\mu \rangle \geq 0$. It remains to prove that the inequality holds strictly when $Q\mu \neq \mu$. It can be shownthat for $q \in \mathbb{S}_g$ at least one of unary inequalities (28) from the support of $\mu$ holds strictly and therefore inequality (33) is also strict. □

**Theorem B.4** (Reduction 2)**.** Let $q\colon \mathcal{X} \to \mathcal{X}$ component-wise idempotent. Let $\bar{g} = g - \Delta^+$ and $\Delta^+ \in \mathbb{R}^{\mathcal{I}}$ has zero unary components and pairwise components as follows:

$$\begin{aligned} \Delta^+_{uv}(i,j) = \max\big\{0, g_{uv}(i,j) + g_{uv}(q_u(i), q_v(j)) \quad (34) \\ - g_{uv}(i, q_v(j)) - g_{uv}(q_u(i), j)\big\}. \end{aligned}$$

Then $q \in \mathbb{S}_g$ iff $q \in \mathbb{S}_{\bar{g}}$.

*Proof.* Note, unary components of $g$ and $\bar{g}$ are equal. We will prove by separate pairwise components.

Direction $\Leftarrow$. Let us prove the inequality (24). It is equivalent to

$$\Delta^+_{uv}(q_u(i), q_v(j)) \leq \Delta^+_{uv}(i,j). \quad (35)$$

Consider the following cases:
- $q_u(i) = i$ or $q_v(j) = j$: In this case $\Delta^+_{uv}(i,j) = 0 = \Delta^+_{uv}(q_u(i), q_v(j))$ by substitution.
- $q_u(i) \neq i$, $q_v(j) \neq j$: By idempotency, it must be that $q_u(q_u(i)) = q_u(i)$ and $q_v(q_v(j)) = q_v(j)$. It follows that $\Delta^+_{uv}(q_u(i), q_v(j)) = 0$ and $\Delta^+_{uv}(i,j) \geq 0$ by definition.

Inequality (24) implies (by multiplication of pairwise inequalities and unary inequalities with respective components of $\mu$ and summing) that

$$(\forall \mu \in \Lambda)\ \langle g, Q\mu \rangle - \langle \bar{g}, Q\mu \rangle \leq \langle g, \mu \rangle - \langle \bar{g}, \mu \rangle. \quad (36)$$

Note, cost vector $\bar{g}$ satisfying (36) is called *auxiliary* for $g$ in [15, 23]. Inequality (36) is equivalent to

$$\langle \bar{g}, (I - Q)\mu \rangle \leq \langle g, (I - Q)\mu \rangle. \quad (37)$$

Whenever the LHS of (37) is strictly positive then so is the RHS and therefore from $q \in \mathbb{S}_{\bar{g}}$ follows $q \in \mathbb{S}_g$.

Direction $\Rightarrow$. Assume $q \in \mathbb{S}_g$. By Theorem B.2, there exist dual multipliers $\varphi$ such that for $g' := g - A^{\mathsf{T}}\varphi$ the following component-wise inequalities hold:

$$\begin{aligned} &(\forall u \in \mathcal{V},\ \forall i \in \mathcal{X}_u)\ g'_u(q_u(i)) \leq g'_u(i); \quad (38) \\ &(\forall uv \in \mathcal{E},\ \forall ij \in \mathcal{X}_{uv})\ g'_{uv}(q_u(i), q_v(j)) \leq g'_{uv}(i,j). \end{aligned}$$

Let $\bar{g}' := \bar{g} - A^{\mathsf{T}}\varphi = g - \Delta^+ - A^{\mathsf{T}}\varphi = g' - \Delta^+$. Let us show that component-wise inequalities (38) hold for $\bar{g}'$. Clearly they hold for unary components and for pairwise components where $\Delta^+_{uv}(i,j) = 0$. Let $uv \in \mathcal{E}$ and $\Delta^+_{uv}(i,j) > 0$. Let $i' = q_u(i)$ and $j' = q_u(j)$. It must be that $i' \neq i$ and $j' \neq j$. Let us denote $a = g'_{uv}(i', j')$, $b = g'_{uv}(i', j)$, $c = g'_{uv}(i, j')$ and $d = g'_{uv}(i', j)$. By idempotency of $q$, there holds $0 = \Delta^+_{uv}(i', j) = \Delta^+_{uv}(i, j') = \Delta^+_{uv}(i', j')$. Let $\bar{d} := g'_{uv}(i,j) - \Delta^+_{uv}(i,j) = d - (a + d - b - c) = b + c - a$. From (38) we have that $a \leq b, c, d$. It follows that $2a \leq b + c$ or $a \leq b + c - a = \bar{d}$. We proved that $\bar{g}'_{uv}(q_u(i), q_v(j)) \leq \bar{g}'_{uv}(i,j)$. In total, $\bar{g}'$ satisfies all component-wise inequalities, same as $g'$ in (28). By Theorem B.2,

$$(\forall \mu \in \Lambda)\ \langle \bar{g}, Q\mu \rangle \leq \langle \bar{g}, \mu \rangle. \quad (39)$$

We have shown that $\langle \bar{g}, (I - Q)\mu \rangle \geq 0$. The inequality holds strictly when $Q\mu \neq \mu$. In this case for $q \in \mathbb{S}_g$ at least one of unary inequalities (38) from the support of $\mu$ holds strictly and therefore inequality (39) is also strict. □

**Remark.** Note, taking into account that $g_{vu}(i', j') = 0$ for $i' \in \mathcal{X}_v \backslash \mathcal{Y}_v$, $j' \in \mathcal{X}_u \backslash \mathcal{Y}_u$, the reduction that we obtained in all cases can be interpreted as forcing the mixed discrete derivative

$$g_{vu}(i, j') + g_{vu}(i', j) - g_{vu}(i, j) - g_{vu}(i', j') \quad (40)$$

to be non-negative. The cost vector $\bar{g}$ is therefore a (partial) submodular *truncation* of $g$. One can notice certain similarity with construction of auxiliary problems by Kovtun [15], where full submodularity is enforced. We essentially proved that a part of Kovtun's construction of constructing an auxiliary problem is optimal.

**Theorem B.5.** Let $q, p \in \mathcal{P}^{2,y}$, $q \leq p$. Let $\bar{g}$ be defined by (13). Then $q \in \mathbb{S}_f$ iff $q \in \mathbb{S}_{\bar{g}}$.

*Proof.* Let $g = (I - P)^\mathsf{T} f$. By Theorem B.1, $q \in \mathbb{S}_f$ iff $q \in \mathbb{S}_g$. We need to consider only pairwise terms. Let $uv \in \mathcal{E}$. Since $q \leq p$, if $p_u(i) = i$ then necessarily $q_u(i) = i$. Let $p$ be defined using sets $\mathcal{Y}_u$ as in (8). The reduction $\bar{g}$ in (13) will be composed of reductions by Theorem B.3 and Theorem B.4.

From $g = (I - P)^\mathsf{T} f$ we have that for $i \in \mathcal{X}_u \backslash \mathcal{Y}_u$ and $j \in \mathcal{X}_v \backslash \mathcal{Y}_v$ $g_{uv}(i, j) = 0$. Conditions of Theorem B.3 are satisfied with $U_s = \mathcal{X}_u \backslash \mathcal{Y}_u$. We obtain part of the reduction (13) for cases when $i \notin \mathcal{Y}_u$ or $j \notin \mathcal{Y}_v$. Let us denote the reduced vector $\bar{g}'$. Applying Theorem B.4 to it, we obtain $\bar{g}$ as defined in (13). $\square$

We are ready to show the statements claimed in §5.

**Theorem 5.1 (Reduction).** The theorem formulated in §5 states that mapping $q \leq p$ is in $\mathbb{S}_f$ iff $q(\mathcal{O}_v^*) = \mathcal{O}_v^*$.

*Proof.* By Theorem B.5, $q \in \mathbb{S}_f$ iff $q \in \mathbb{S}_{\bar{g}}$, where $\bar{g}$ is defined by (13). There holds characterization $q \in \mathbb{S}_{\bar{g}}$ iff $q_u(O_u^*) = O_u^*$, where $O_u^*$ is defined by (12). $\square$

**Lemma 5.2:** "Let $q \in \mathbb{S}_f \cap \mathcal{P}^{2,y}$, $q \leq p$, $Q = [q]$. Let $\Lambda' \subseteq \Lambda$ and $Q(\Lambda') \subseteq \Lambda'$. Let $\bar{g}$ be defined by (13) (depends on $p$) and let $\mathcal{O}^* = \operatorname{argmin}_{\mu \in \Lambda'} \langle \bar{g}, \mu \rangle$. Then $(\forall v \in \mathcal{V})$ $q_v(\mathcal{O}_v^*) = \mathcal{O}_v^*$."

*Proof.* Note, unlike Theorem 5.1, this lemma states a necessary condition only. Let $\mu \in \mathcal{O}^*$. Assume for contradiction that $Q\mu \neq \mu$. In this case, by Theorem 5.1, we have that $\langle \bar{g}, Q\mu \rangle < \langle \bar{g}, \mu \rangle$. Since $\mu \in \Lambda'$ and $Q(\Lambda') \subseteq \Lambda$ there holds $Q\mu \in \Lambda'$. It follows that $Q\mu$ is a feasible solution of a better cost than $\mu$ which contradicts optimality of $\mu$. It must be therefore that $Q\mu = \mu$. The claim $q(\mathcal{O}_v^*) = O_v^*$ follows. $\square$

We have achieved the following. Suppose that we solve the verification LP with the reduced cost vector $\bar{g}$. If there holds $p_v(O_v^*) = O_v^*$ for all $v$, then $p \in \mathbb{S}_f$ by Theorem 5.1. In the opposite case, Theorem 5.1 asserts that for all strictly relaxed-improving mappings $q \leq p$ there must hold $q_v(O_v^*) = O_v^*$. Therefore the reduced verification LP is valid to be used in the step 5 of Algorithm 1.

A practical aspect of the reduction is that, *e.g.*, TRW-S is able to find a labeling with a negative cost much faster since we have decreased many edge costs. Additionally, minimization over polytope $\Lambda_x$ with reduced $g$ is a purely submodular problem (recall that $y_v \in \mathcal{X}_v \backslash \mathcal{Y}_v$). We thus can solve it with a regular minimum cut and not QPBO minimum cut.

# C. Termination with AC Solvers

For completeness, we give two theorems when the solver is applied to the verification problem with and without reduction. Both results apply when the solver has found an arc consistent solution. The theorem with reduction allows to obtain somewhat stronger guarantees. The guarantees are necessary in order to show that our Algorithm 2 terminates. In the next section we prove guarantees for TRW-S also in the case when it has not converged yet.

**Theorem C.1.** Consider the verification LP defined by $g = (I - P^\mathsf{T})f$. Let $g^\varphi$ be an arc-consistent reparametrization. Then at least one of the two conditions is satisfied:

(a) $LB(\varphi) \overset{\text{def}}{=} g_0 + \sum_u \varphi_u = 0$ and $\varphi$ is dual optimal;
(b) $(\exists u \in \mathcal{V})$ $\mathcal{O}_u(\varphi) \cap \mathcal{Y}_u \neq \varnothing$.

*Proof.* Assume (b) does not hold: $(\forall u \in \mathcal{V})$ $\mathcal{O}_u(\varphi) \cap \mathcal{Y}_u = \varnothing$. Then for each node $u$ there is a label $z_u \in \mathcal{O}_u(\varphi) \backslash \mathcal{Y}_u$. By arc consistency, for each edge $uv$ there is a label $j \in \mathcal{O}_u(\varphi) \subseteq \mathcal{X}_u \backslash \mathcal{Y}_u$ such that $g_{uv}^\varphi(z_u, j) = 0$ and similarly, there exists $i \in \mathcal{O}_u(\varphi) \subseteq \mathcal{X}_u \backslash \mathcal{Y}_u$ such that $g_{uv}^\varphi(i, z_v) = 0$.

By construction, $g_{uv}(i, j) = 0$ for all $ij \in \mathcal{X}_{uv} \backslash \mathcal{Y}_{uv}$ and therefore the following modularity equality holds:

$$g_{uv}^\varphi(z_u, z_v) + g_{uv}^\varphi(i, j) = g_{uv}^\varphi(z_u, j) + g_{uv}^\varphi(i, z_v). \quad (41)$$

From local minimality of $(z_u, j)$ we have

$$g_{uv}^\varphi(z_u, j) \leq g_{uv}^\varphi(i, j). \quad (42)$$

By adding (41) and (42) we obtain

$$g_{uv}^\varphi(z_u, z_v) \leq g_{uv}^\varphi(i, z_v) \quad (43)$$

and hence $(z_u, z_v)$ is locally minimial too: $g_{uv}^\varphi(z_u, z_v) = 0$. Therefore $\delta(z)$ and dual point $\varphi$ satisfy complementarity slackness and hence they are primal-dual optimal and $LB(\varphi) = E_g(z) = 0$. $\square$

We can now show the result needed to establish termination of Algorithm 2 based on a solver delivering arc consistency, but not necessarily solving the LP relaxation in all cases.

**Lemma 4.4**: "Let $(\forall v \in \mathcal{V})$ $\mathcal{O}_v(\varphi) \cap \mathcal{Y}_v = \varnothing$ hold for an arc consistent dual vector $\varphi$. Then $\varphi$ is dual optimal".

*Proof.* Corollary from Theorem C.1. $\square$

For the reduced problem $\bar{g}$, a stronger condition is satisfied.

**Theorem C.2.** Consider the reduced problem $\bar{g}$ defined in Theorem 5.1. Let $\bar{g}^{\varphi}$ be an arc-consistent reparametrization. Then at least one of the two conditions is satisfied:

(a) $LB(\varphi) = 0$ and $\varphi$ is optimal;

(b) $(\exists u \in \mathcal{V})\ \ \mathcal{O}_u(\bar{g}^{\varphi}) \subseteq \mathcal{Y}_u$.

*Proof.* Since for problem $\bar{g}$ all labels in $\mathcal{X}_u \backslash \mathcal{Y}_u$ are equivalent w.r.t. unary and pairwise costs, we can contract these sets and w.l.o.g. assume that $\mathcal{X}_u \backslash \mathcal{Y}_u = \{y_u\}$.

Assume (b) does not hold. Then for each node $u$ label $y_u$ is locally minimal: $\bar{g}^{\varphi}(y_u) = \min_i \bar{g}^{\varphi}_u(i)$. By arc consistency, for each edge $uv$ there is a label $j \in \mathcal{X}_v$ such that pair $(y_u, j)$ is locally minimal and similarly there exists $i \in \mathcal{X}_v$ such that pair $(i, y_v)$ is locally minimal. From partial submodularity, inequality (40), follows that the pair $(y_u, y_v)$ is locally minimal as well. Therefore, integer labeling $y$ and dual point $\varphi$ satisfy complementarity slackness and hence they are primal-dual optimal and $LB(\varphi) = E_{\bar{g}}(y) = 0$. ☐

Both proofs above are analogues to how fixed points of TRW-S or max-sum diffusion are shown to correspond to exact solutions for submodular problems [22], *i.e.*, we use the same argument to construct an integer feasible solution that satisfy complementary slackness.

# D. Implementation with TRW-S

When we consider specifically TRW-S there are several questions regarding correctness and efficiency:

- TRW-S may not achieve arc consistency in finite time. Can we stop it earlier? Will there be some progress possible in terms of sets $\mathcal{Y}_u$?
- Can we exploit efficient distance transforms on the reduced problem?

In this section we answer all the above questions positively.

## D.1. Algorithm Details

The algorithm will make use of the reduction, pruning cuts and the warm start. The warm start is motivated by that every next outer iteration of Algorithm 1 may result in just a small adjustment to the problem $\bar{g}$ and therefore it is desirable to reuse the messages (reparametrization) in TRW-S from the previous outer iteration. We first give a full specification in Algorithm 3 and state formal properties of the algorithm. Then we prove the claims by deriving some new properties of TRW-S.

**Theorem D.1** (Correctness of Algorithm 3)**.** For any stopping condition in line 4, the algorithm terminates in at most $\sum_v (|\mathcal{X}_v| - 1)$ outer iterations and outputs $p \in \mathbb{S}_f$.

**Theorem D.2.** Message passing in TRW-S for reduced edge term $\bar{g}_{uv}$ can be computed in $O(|\mathcal{Y}_u| + |\mathcal{Y}_v|)$ extra time compared to the message passing for the original term $f_{uv}$.

**Algorithm 3:** Efficient Iterative Pruning with TRW-S

**Input**: Problem $f \in \mathbb{R}^{\mathcal{I}}$, test labeling $y \in \mathcal{X}$;
**Output**: Improving map $q \in \mathcal{P}^{2,y} \cap \mathbb{S}_f$;
1 $(\forall s \in \mathcal{V})\ \mathcal{Y}_u := \mathcal{X}_s \backslash \{y_s\}$;
2 Set $\varphi$ to the initial reparametrization / messages if available;
3 **repeat**
4   Construct reduced characterization problem $\bar{g}$ for problem $f$ and sets $(\mathcal{Y}_u \mid u \in \mathcal{V})$ according to Table 6;
5   **repeat**
6     Run 10 TRW-S iterations with problem $\bar{g}$ and reparametrization $\varphi$. Provides $LB_{\mathrm{T}}$, $\varphi$, $O_u$ and best $x$;
7     **if** $(\forall s \in \mathcal{V})\ O_u \cap \mathcal{Y}_u = \varnothing$ **then return** $q$;
8     **if** $E_{\bar{g}}(x) < 0$ **then**
9       Apply pruning cut with $x$;
10       Apply single node pruning;
11       **goto** step 4 to rebuild $\bar{g}$;
12   **until** any stopping condition (*e.g.*, iteration limit);
13   Mark immovable: $(\forall u \in \mathcal{V})\ \mathcal{Y}_u := \mathcal{Y}_u \backslash O_u$;
14   Apply single node pruning;

*Proof* (of Theorem D.1). Using Margin Theorem D.3 below, either current mapping $q$ is returned or whenever we decide to stop TRW-S iterations, the set of minimal labels $O_u$ for some $s$ contains a new label to be marked as immovable. It follows that Algorithm 3 terminates. The convergence of TRW-S does not affect correctness of our algorithm, only a possibly non-maximal map $q$ is found due to non-optimal dual point used in step 13. The termination condition in step 7 indicates that $y$ is an optimal primal solution and in this case it must be that $LB = 0$, therefore current $\varphi$ is an optimal dual point. In this case $O^*_u \subseteq O_u$, where set $O^*_u$ is the support set of all primal optimal solutions (see Algorithm 1). We therefore have $O^*_u \subseteq O_u \subseteq \mathcal{X}_u \backslash \mathcal{Y}_u$ for all $u$. It follows that stopping condition of generic Algorithm 1 is satisfied as well and hence $q \in \mathbb{S}_f$. ☐

Note, whenever the algorithm found a pruning cut or used local pruning conditions, no loss of maximality occurs. In our experiments for some instances, the algorithm finished before ever reaching the step 13. In such cases the reduction $q \in \mathbb{S}_f$ is the maximum.

## D.2. Properties of TRW-S

In this section we give several theoretical guarantees for TRW-S when it is applied to solve the verification problem $(I - [p]^{\mathsf{T}})f$ or to its reduced version.

TRW-S algorithm, though observed to always converge in practice to a fixed point, has very weak theoretical convergence guarantees. It is know that there is a convergent subsequence who's limit satisifies Week Tree Agreement

(WTA) [12]. A finite iteration epsilon variant of this statement exists [21]. In either case we have no guarantee to obtain an arc consistent solution required to prove Proposition 4.3. The major obstacles are: (i) WTA is achieved only in the limit and (ii) there is no guarantee that the set of labels which are in WTA (let alone its AC subset) converges as well.

Fortunately, TRW-S enjoys very useful for us properties especially when it is applied to the reduced verification problem. We first give a short specification of TRW-S. In our notation, the order of TRW-S updates is fully specified by orientations of edges in $\mathcal{E}$. For a serial implementation, this order can be completed to a total order on $\mathcal{V}$ by defining $u < v$ iff $uv \in \mathcal{E}$. We specify only the forward pass of TRW-S as Algorithm 4 and note that the backward pass is obtained by reversing all edges. In Algorithm 4, $n_u$ is the

**Algorithm 4:** TRW-S Forward Iteration, *c.f.* [12, Fig. 3]

**Input**: Problem $f \in \mathbb{R}^{\mathcal{I}}$, reparametrization $\varphi$;
**Output**: Updated reparametrization $\varphi$, $LB_T$, best labeling $x$, locally optimal labels $\mathcal{O}_u(\varphi)$;

**for** $u \in \mathcal{V}$ **do**
  **for** $v \in \mathrm{nb}(u) \mid uv \in \mathcal{E}$ **do**
    $\varphi_{vu}(j) := \min\limits_i \Big[\frac{1}{n_u} f^\varphi_u(i) - \varphi_{uv}(i) + f_{uv}(i,j)\Big]$;
$\mathcal{O}_u(\varphi) := \mathrm{argmin}_i\, f^\varphi_u(i)$; $x_u \in \mathcal{O}_u(\varphi)$;
$LB_{\mathrm{T}}(\varphi) := \sum_{u\in\mathcal{V}} \frac{n_{\mathrm{term}}(u)}{n_u} \min\limits_i f^\varphi_u(i)$;

number of chains that meet in node $u$ and $n_{\mathrm{term}}(u)$ is the number of chains that terminate in $u$.

Viewed as computing the division of costs between the chains, TRW-S has the following properties. To a chain $\tau$ there is associated its oriented graph $(\mathcal{V}^\tau, \mathcal{E}^\tau)$ and its share of the unary terms in the decomposition, $f^\tau$. We assume thet each pairwise term $f_{uv}$ is associated to exactly one chain passing through edge $uv$.

After the first backward pass, the algorithm maintains the following invariants. Consider the step of processing vertex $u^* \in \mathcal{V}$. For a chain $\tau$ which passes through edge $uv \in \mathcal{E}$ there holds:

- If $u^* < u$, message $\varphi_{uv}(i)$ is equal to the *right* min-marginals of the chain:

$$\varphi_{uv}(i) = \min_{\substack{(x_w \mid w>v)\\ x_u=i}} \Big(\sum_{\substack{u\in\mathcal{V}^\tau\\ w>u}} f^\tau_w(x_w) + \sum_{\substack{u'v'\in\mathcal{E}^\tau\\ u'\leq u}} f_{u'v'}(x_{u'}, x_{v'})\Big). \tag{44}$$

- If $u^* > u$, message $\varphi_{vu}(j)$ is equal to the *left* min-marginals of the chain:

$$\varphi_{vu}(j) = \min_{\substack{(x_w \mid w<v)\\ x_v=j}} \Big(\sum_{\substack{w\in\mathcal{V}^\tau\\ w<v}} f^\tau_w(x_w) + \sum_{\substack{u'v'\in\mathcal{E}^\tau\\ u'\leq u}} f_{u'v'}(x_{u'}, x_{v'})\Big). \tag{45}$$

Weak tree agreement is the condition when the lower bound cannot be improved further by TRW-S [12]. It requires that among all optimal assignments of the chains there is a consistent subset.

We now introduce a measure of how far the TRW-S algorithm is from a fixed integer solution $y$. For a node $u$ we define *node margin* of $u$ as the value

$$m_u(\varphi) = \min_i (f^\varphi_u(i) - f^\varphi_u(y_u)). \tag{46}$$

The margin is related to the set $\mathcal{O}_u(\varphi)$ of local minimizers. If it is negative, $\mathcal{O}_u(\varphi)$ does not contain $y_u$ and the negative value measures how well it is separated. Let us define the *problem margin* as the value

$$m(\varphi) = \min_{s\in\mathcal{V}} m_u(\varphi). \tag{47}$$

For the convenience of analysis, we continue to denote the cost vector to which Algorithm 4 is applied as $f$. In the context of Algorithm 3, we temporarily let $f := \bar{g}$.

**Theorem D.3** (Margin for TRW-S)**.** For the reduced verification problem $f$ and reparametrization $\varphi$ build by TRW-S there holds:

$$LB_{\mathrm{T}}(\varphi) = 0 \ \text{ iff } \ m_u(\varphi) \geq 0. \tag{48}$$

*Proof.* $(\Rightarrow)$ Let $LB_T(\varphi) = 0$. By construction, $E_f(y) = 0$ so the duality gap is zero and therefore $\varphi$ is dual optimal and $\delta(y)$ is primal optimal. By complementary slackness it must be that $(\forall u \in \mathcal{V})$ $f^\varphi_u(y_u)$ is locally minimal. Therefore problem margin is exactly zero. This part of the proof holds also for non-reduced verification problem and any reparametrization such that $LB_T(\varphi) = 0$.

$(\Leftarrow)$ Let now the problem margin be non-negative. Let us consider edge $uv \in \mathcal{E}$ and let $f^\tau$ be the cost vector of the monotonic chain $\tau$ passing through edge $uv$ at the beginning of iteration processing node $v$. Since $u$ was already processed, the value $f^\varphi_u$ is proportional to min-marginals of chain $\tau$ in $u$ and we have $m_u(\varphi) \geq 0$. It follows that there is a minimizer $x \in \min_x \langle f^\tau, \delta(x)\rangle$ such that $x_u = y_u$. For this minimizer there holds

$$x \in \operatorname*{argmin}_{\tilde{x} \mid \tilde{x}_v = x_v} \langle f^\tau, \delta(\tilde{x})\rangle. \tag{49}$$

This is because we constrained $\tilde{x}$ to coincide with the global chain minimizer in $v$. This conditional minimizer however clearly does not depend on $f^\tau_v$, the unary term associated to chain $\tau$.

Consider the operation of averaging min-marginals over node $v$. By assumption, after averaging operation, $m_v \geq 0$ and therefore the local minimum over each chain is attained in $y_v$. The averaging operation only changes the value $f^\tau_v$. The conditional minimizer

$$x' \in \underset{\tilde{x} | \tilde{x}_v = y_v}{\operatorname{argmin}} \langle f^\tau, \delta(\tilde{x}) \rangle \tag{50}$$

does not depend on $f^\tau_v$.

Let $r$ precede $u$ in $\tau$. Assume for contradiction that $x'_u \neq y_u$. Because $u$ has been processed, the value $\varphi_{ur}(i)$ is the left-min-marginal for chain $\tau$ in node $u$ and label $i \in \mathcal{X}_u$. Denote $\overrightarrow{f^\tau_u} = \varphi_{ur}(i) + f^\tau_u(i)$. The relations (49) and (50) can be expressed as

$$y_u \in \underset{i}{\operatorname{argmin}} \big(\overrightarrow{f^\tau_u} + f_{uv}(i, x_v)\big); \tag{51a}$$

$$x'_u \in \underset{i}{\operatorname{argmin}} \big(\overrightarrow{f^\tau_u} + f_{uv}(i, y_v)\big). \tag{51b}$$

From (51a) we have that $\overrightarrow{f^\tau_u}(y_u) + f_{uv}(y_u, x_v) \leq \overrightarrow{f^\tau_u}(x'_u) + f_{uv}(x'_u, x_v)$. We use now the partial submodularity fulfilled for $f$,

$$f_{uv}(y_u, x_v) + f_{uv}(x'_u, y_v) \geq f_{uv}(x'_u, x_v) \tag{52}$$

and obtain that

$$\overrightarrow{f^\tau_u}(y_u) \leq \overrightarrow{f^\tau_u}(x'_u) + f_{uv}(x'_u, y_v). \tag{53}$$

Adding on the LHS the term $f_{uv}(y_u, y_v) = 0$, we get

$$\overrightarrow{f^\tau_u}(y_u) + f_{uv}(y_u, y_v) \leq \overrightarrow{f^\tau_u}(x'_u) + f_{uv}(x'_u, y_v). \tag{54}$$

This inequality allows to conclude that $y_u$ is a also a minimizer to (51b). Therefore either $x'_u = y_u$ or $y_u$ is an equally good substitute.

We have shown that as soon as there is zero margin, for each chain the part of the minimizer over already processed nodes can be selected equal to $y$. By induction, at the end of the sweep, $y$ is a minimizer for each chain. Hence WTA is achieved. In this case, there is zero integrality gap and $LB_T = E_f(y) = 0$. □

It follows that on *every* iteration (but the first initializing one) of TRW-S either the problem margin is negative and therefore $(\exists u)\ \mathcal{O}_u(\varphi) \cap \mathcal{Y}_u \neq \varnothing$ (there is always something to prune) or WTA is achieved and the algorithm terminates with $y$ being optimal integer solution. This proves Theorem D.1. Furthermore, if the stopping condition in line 12 is the iteration limit, Algorithm 3 runs in polynomial time.

### D.3. Efficient Message Passing

Suppose that some pairwise potentials are specially structured, so that messages in message passing algorithms can be computed in linear time in the number of labels instead of quadratic time, see [5]. This is the case for many potentials with a linear ordering, e.g. absolute differences, squared distance and truncated versions thereof, including the Potts model.

| | |
|---|---|
| $i \notin \mathcal{Y}_u$ | $\bar{g}_u(i) = 0;$ |
| $i \in \mathcal{Y}_u$ | $\bar{g}_u(i) = f_u(i) - f_u(y_u);$ |
| $i \notin \mathcal{Y}_u\ j \notin \mathcal{Y}_v$ | $\bar{g}_{uv}(i,j) = 0;$ |
| $i \notin \mathcal{Y}_u\ j \in \mathcal{Y}_v$ | $\bar{g}_{uv}(i,j) = \Delta_{vu}(j)$, $\Delta_{vu}(j) = \min_{i' \notin \mathcal{Y}_u} \big[f_{uv}(i',j) - f_{uv}(i',y_v)\big];$ |
| $i \in \mathcal{Y}_u\ j \notin \mathcal{Y}_v$ | $\bar{g}_{uv}(i,j) = \Delta_{uv}(i)$, $\Delta_{uv}(i) = \min_{j' \notin \mathcal{Y}_v} \big[f_{uv}(i,j') - f_{uv}(y_u,j')\big];$ |
| $i \in \mathcal{Y}_u\ j \in \mathcal{Y}_v$ | $\bar{g}_{uv}(i,j) = \min\Big\{f_{uv}(i,j) - f_{uv}(y_u,y_v),\ \Delta_{vu}(j) + \Delta_{uv}(i)\Big\}.$ |

Table 6. Components of the Reduced Verification Problem

It would be very helpful to be able to compute messages fast in these cases also for our reduced verification problem $\bar{g}$. The components of $\bar{g}$ specified by Theorem 5.1 can be expressed directly in components of $f$ as proposed in Table 6. Passing a message on edge $uv$ amounts to calculate an expression of the form

$$\varphi_{vu}(j) := \min_{i \in \mathcal{X}_u} \Big[a(i) + \bar{g}_{uv}(i,j)\Big] \tag{55}$$

for some vector $a \in \mathbb{R}^{\mathcal{X}_u}$. For $j \notin \mathcal{Y}_v$, substituting pairwise terms of $\bar{g}$ it is expanded as

$$\begin{aligned} \varphi_{vu}(j) &= \min_{i \in \mathcal{X}_u} \Big[a(i) + \Delta_{uv}(i)\Big] \\ &= \min_{i \in \mathcal{X}_u} \Big[a(i) + \Delta_{uv}(i)\Big] + \Delta_{vu}(j), \end{aligned} \tag{56}$$

where we defined that $\Delta_{uv}(U_u) = \Delta_{vu}(U_v) = 0$.

For $j \in \mathcal{Y}_v$, substituting pairwise terms of $\bar{g}$ and denoting $c = f_{uv}(y_u, y_v)$ the message is expanded as

$$\begin{aligned} &\min\Big\{\min_{i \in U_u} a(i) + \Delta_{vu}(j), \\ &\quad \min_{i \in \mathcal{Y}_u} \Big[a(i) + \min\big\{f_{uv}(i,j) - c, \Delta_{uv}(i) + \Delta_{vu}(j)\big\}\Big]\Big\} \\ &= \min\Big\{\min_{i \in \mathcal{Y}_u} \big[a(i) + f_{uv}(i,j)\big] - c, \\ &\qquad \min_{i \in \mathcal{X}_u \setminus \mathcal{Y}_u} \big[a(i) + \Delta_{uv}(i)\big] + \Delta_{vu}(j)\Big\}. \end{aligned} \tag{57}$$

We therefore need to calculate (56) for all $j \in \mathcal{X}_u$ because the expression reoccurs in (57) for $j \in \mathcal{Y}_v$. Then for $j \in \mathcal{Y}_v$ it remains to take the minimum of a regular message and $\varphi_{vu}(j)$ by (56).

**Algorithm 5:** Message Passing for Pruning

**Input**: Unary term $a \colon \mathcal{Y}_u \cup \{y_u\} \to \mathbb{R}$;
**Output**: Message $\varphi_{vu} \colon \mathcal{Y}_v \cup \{y_v\} \to \mathbb{R}$;

`/* Offset constant in (56) */`

$m_1 := \min\limits_{i \in \mathcal{Y}_u \cup \{y_u\}} \big[a(i) + \Delta_{uv}(i)\big]$;

`/* Message passing for f */`

$(\forall j \in \mathcal{Y}_v)\ \varphi_{vu}(j) := \min\limits_{i \in \mathcal{Y}_u \cup \{y_u\}} \big[a(i) + f_{uv}(i,j)\big]$;

`/* Message correction for ḡ */`

$(\forall j \in \mathcal{Y}_v)\ \varphi_{vu}(j) := \min\{\varphi_{vu}(j) - c, \Delta_{vu}(j) + m_1\}$;

$\varphi_{vu}(y_v) := m_1$;

We thus have reduced the message passing for $\bar{g}$ to the message passing for $f$, $O(|\mathcal{X}_u| + |\mathcal{X}_v|)$ operations to calculate expression (56) and $O(|\mathcal{X}_v|)$ operations for the outer minimum. Therefore we can compute the message for the modified energy in time $O(|\mathcal{X}_u| + |\mathcal{X}_v|)$. We can take this complexity down to the theoretical perfection as follows. Recall that labels in $\mathcal{X}_u \backslash \mathcal{Y}_u$ can be contracted to a single one. In this case calculating (56) for all $j \in \mathcal{Y}_u$ takes only $O(|\mathcal{Y}_u| + |\mathcal{Y}_v|)$ time. Using the non-uniform min-convolution algorithm of [37] the message passing for the labels in the set $\mathcal{Y}_u$ to labels in $\mathcal{Y}_v$ can be implemented in $O(|\mathcal{Y}_u| + |\mathcal{Y}_v|)$ time. We have obtained the complexity matching to the total number of active labels in the problem. The more labels are marked as immovable in the course of the algorithm (sets $\mathcal{Y}_u$ reduce), the less work is required. The final message update is specified in Algorithm 5, where the contracted labels $\mathcal{X}_u \backslash \mathcal{Y}_u$ are represented by $y_u$. In the implementation for each edge $uv$ we need to store the offsets $\Delta_{uv} \colon \mathcal{Y}_u \to \mathbb{R}$ and $\Delta_{vu} \colon \mathcal{Y}_v \to \mathbb{R}$ and $c \in \mathbb{R}$.

## E. Detailed Experimental Evaluation

**Datasets and Evaluation.** We give a brief characterization of all 38 test problem instances and report the obtained total percentage of persistent variables of our and competing methods in Table 2. The datasets `mrf-stereo` and `mrf-photomontage` originate from the Middlebury MRF benchmark [31]. The `color-seg` and `color-seg-n4` datasets were taken from the OpenGM MRF benchmark [10, 9], `ProteinFolding` originates from [16, 36]. All datasets are made available in the OpenGM-format [10, 9].

Detailed quantitative experimental evaluation can be found in Table 7. In addition to the per-label measure of partial optimality (15), to allow for future comparisons we report also the *logarithmic measure*. It is motivated by the fact that eliminating one label in a variable with say 2 states brings more information than eliminating one label in a variable with 100 states. We propose to measure the total decrease of the number of configurations of the search

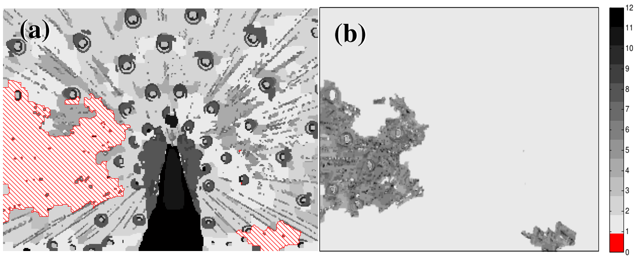


Figure 2. Instance `pfau`. (a) Proved part of optimal solution (red solution not determined / non-unique). (b) Reminder of the optimization problem: number of remaining labels in every pixel.

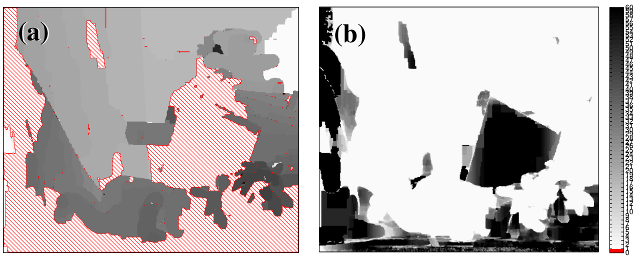


Figure 3. Instance `ted`. (a) Proved part of optimal solution (red = solution not determined / non-unique). (b) Reminder of the optimization problem: number of remaining labels in every pixel.

space, *e.g.*, from $|\mathcal{X}|$ to $p(|\mathcal{X}|)$, in the logarithmic domain:

$$1 - \frac{\log \prod_{v \in \mathcal{V}} |p_v(\mathcal{X}_v)|}{\log \prod_{v \in \mathcal{V}} |\mathcal{X}_v|} = 1 - \frac{\sum_{v \in \mathcal{V}} \log |p_v(\mathcal{X}_v)|}{\sum_{v \in \mathcal{V}} \log |\mathcal{X}_v|}. \quad (58)$$

In Figures 2-4 we give examples where the method was performing well. Figures 5 and 6, on the contrary reveal some cases of very poor performance. For example for `photomontage/pano` instance, we report 80% solution completeness, but these 80% only correspond to trivial hard constraints in the problem. Other methods perform worse mainly because they consider determining complete optimal labels only (`[29]-TRWS`) or intervals of labels (`MQPBO`).

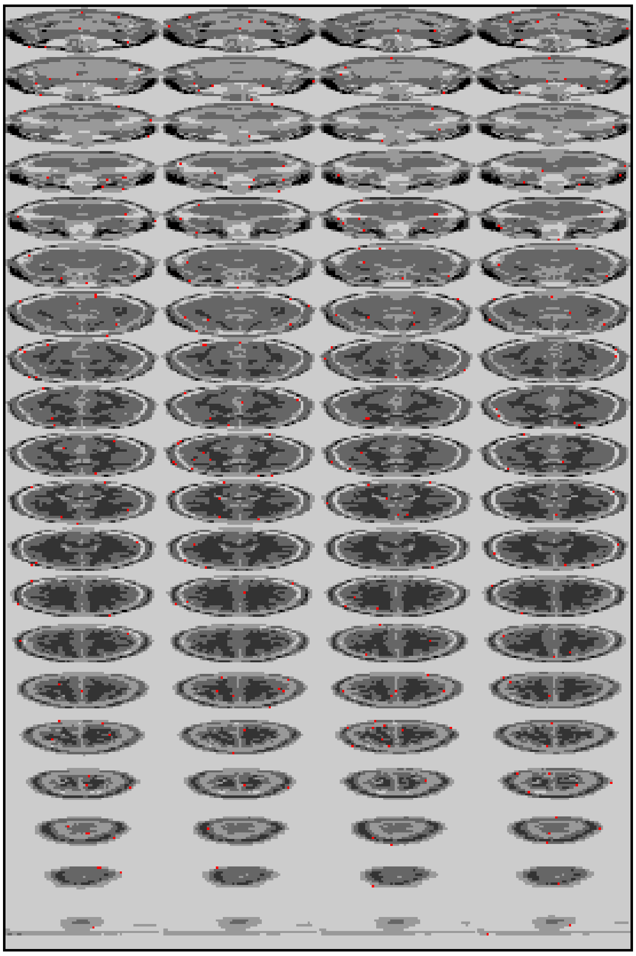

Figure 4. Instance `brain-9mm/brain-0`. Slices of the 3D volumetric problem. Proved part of optimal solution (red = solution not determined / non-unique).

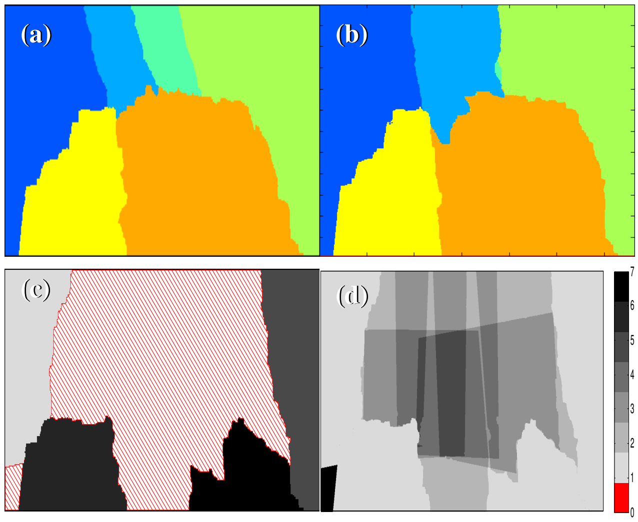


Figure 5. Instance `pano`: label encodes the image index for photomontage. (a) (b) Two labelings by TRW-S with slightly different initializations. It is clear that there is high ambiguity. (c) Part of the solution that was proved optimal and unique. (d) Reminder of the optimization problem (number of non-eliminated labels). It is clear that the method essentially removed hard constraints implied by different fields of view of images composing the panorama.

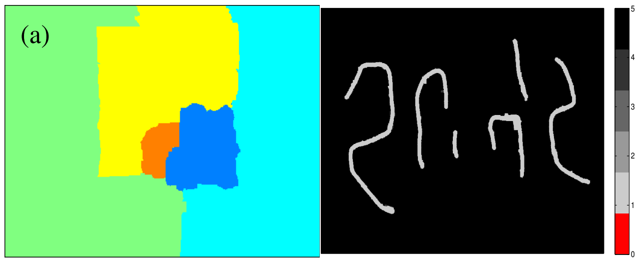


Figure 6. Instance `family`. (a) Labeling by TRW-S. (b) Reminder of the optimization problem (number of labels per pixel). It is clear that the method essentially only followed the hard constraints corresponding to the scribbles provided by the user and constrained very few pixels ontop of that.

| Instance | Algorithm | Time needed overall (s) | Time for initial solution (s) | #iterations Algorithm 1,2 | #iterations TRWS | Logarithmic percentage partial optimality | Percentage excluded labels |
|---|---|---|---|---|---|---|---|
| | | | ProteinFolding | | | | |
| 1CKK | Our-CPLEX | 2757.62 | 1177.62 | 5 | † | 14.24% | 27.04% |
| | Our-TRWS | 5.76 | 5.00 | 3 | 1000+15 | 13.83% | 26.53% |
| | MQPBO-10 | 5670.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 825.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 2502.69 | 2493.65 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 47.57 | 30.19 | 2 | 288+185 | 0.00% | 0.00% |
| 1CM1 | Our-CPLEX | 4070.00 | 992.15 | 7 | † | 8.38% | 34.28% |
| | Our-TRWS | 6.03 | 4.70 | 4 | 1000+65 | 8.07% | 29.98% |
| | MQPBO-10 | 5520.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 723.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 2388.46 | 2198.04 | 3 | † | 0.00% | 0.00% |
| | [29]-TRWS | 51.33 | 21.60 | 3 | 242+358 | 0.00% | 0.00% |
| 1SY9 | Our-CPLEX | 2628.72 | 416.74 | 5 | † | 25.34% | 51.30% |
| | Our-TRWS | 6.88 | 5.50 | 4 | 1000+15 | 28.06% | 57.98% |
| | MQPBO-10 | 7494.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 2112.00 | 0.00 | 0 | † | 0.00% | 0.11% |
| | [29]-CPLEX | 1067.46 | 910.90 | 4 | † | 0.00% | 0.00% |
| | [29]-TRWS | 66.73 | 46.77 | 5 | 400+174 | 0.00% | 0.00% |
| 2BBN | Our-CPLEX | 9677.42 | 5476.81 | 5 | † | 2.12% | 8.58% |
| | Our-TRWS | 10.00 | 8.60 | 3 | 1000+10 | 2.64% | 14.17% |
| | MQPBO-10 | 1736.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 2429.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 9776.60 | 9771.18 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 54.21 | 42.90 | 2 | 242+146 | 0.00% | 0.00% |
| 2BCX | Our-CPLEX | 36222.90 | 6998.66 | 5 | † | 4.81% | 15.66% |
| | Our-TRWS | 9.14 | 7.90 | 3 | 1000+55 | 4.39% | 14.21% |
| | MQPBO-10 | 1008.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 1288.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 11419.60 | 11409.90 | 2 | † | 0.00% | 0.00% |
| | [29]-TRWS | 55.26 | 39.60 | 2 | 252+194 | 0.00% | 0.00% |
| 2BE6 | Our-CPLEX | 1381.60 | 765.84 | 4 | † | 9.14% | 17.68% |
| | Our-TRWS | 4.67 | 3.91 | 4 | 1000+60 | 8.96% | 15.12% |
| | MQPBO-10 | 3728.00 | 0.00 | 0 | † | 0.00% | 0.05% |
| | MQPBO | 540.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 1552.95 | 1552.88 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 40.12 | 28.12 | 2 | 363+230 | 0.00% | 0.00% |
| 2F3Y | Our-CPLEX | 3628.90 | 2546.68 | 5 | † | 6.22% | 10.66% |
| | Our-TRWS | 5.83 | 5.20 | 3 | 1000+10 | 8.39% | 13.74% |
| | MQPBO-10 | 5138.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 928.00 | 0.00 | 0 | † | 0.00% | 0.05% |
| | [29]-CPLEX | 4618.78 | 4618.76 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 41.87 | 33.03 | 3 | 321+164 | 0.00% | 0.00% |
| 2FOT | Our-CPLEX | 7458.75 | 1996.55 | 5 | † | 4.10% | 11.64% |
| | Our-TRWS | 6.25 | 5.30 | 4 | 1000+25 | 4.01% | 11.01% |
| | MQPBO-10 | 4961.00 | 0.00 | 0 | † | 0.00% | 0.09% |
| | MQPBO | 1054.00 | 0.00 | 0 | † | 0.00% | 0.07% |
| | [29]-CPLEX | 4473.58 | 4440.51 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 61.92 | 44.42 | 2 | 398+222 | 0.00% | 0.00% |
| 2HQW | Our-CPLEX | 5721.95 | 1946.20 | 6 | † | 10.30% | 17.30% |
| | Our-TRWS | 6.49 | 4.80 | 6 | 1000+160 | 8.33% | 18.08% |
| | MQPBO-10 | 7228.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 1193.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 2163.98 | 2161.07 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 44.07 | 35.46 | 2 | 382+121 | 0.00% | 0.00% |
| 2O60 | Our-CPLEX | 12085.40 | 3007.95 | 6 | † | 4.22% | 12.81% |
| | Our-TRWS | 7.74 | 6.50 | 3 | 1000+55 | 4.94% | 15.55% |
| | MQPBO-10 | 7516.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 1997.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 6137.07 | 6128.14 | 1 | † | 0.00% | 0.00% |
| | [29]-TRWS | 93.87 | 46.42 | 2 | 352+369 | 0.00% | 0.00% |
| 3BXL | Our-CPLEX | 3247.11 | 915.86 | 7 | † | 4.97% | 17.18% |
| | Our-TRWS | 7.44 | 5.90 | 4 | 1000+60 | 4.66% | 12.35% |

| Instance | Algorithm | Time needed overall (s) | Time for initial solution (s) | #iterations Algorithm 1,2 | #iterations TRWS | Logarithmic percentage partial optimality | Percentage excluded labels |
|---|---|---|---|---|---|---|---|
| | MQPBO-10 | 6709.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 1291.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-CPLEX | 1776.23 | 1598.07 | 2 | † | 0.00% | 0.00% |
| | [29]-TRWS | 44.71 | 25.52 | 2 | 227+216 | 0.00% | 0.00% |
| pdb1b25 | Our-CPLEX | 1599.67 | 55.01 | 28 | † | 76.76% | 84.05% |
| | Our-TRWS | 5.18 | 2.92 | 18 | 530+150 | 83.00% | 87.84% |
| | MQPBO-10 | 27.00 | 0.00 | 0 | † | 0.00% | 2.53% |
| | MQPBO | 2.00 | 0.00 | 0 | † | 0.00% | 1.99% |
| | [29]-CPLEX | 324.64 | 72.11 | 14 | † | 18.84% | 22.32% |
| | [29]-TRWS | 119.71 | 27.62 | 14 | 443+1238 | 18.92% | 22.34% |
| pdb1d2e | Our-CPLEX | 154.76 | 25.44 | 5 | † | 97.30% | 97.98% |
| | Our-TRWS | 1.67 | 1.13 | 7 | 420+75 | 96.97% | 98.25% |
| | MQPBO-10 | 12.00 | 0.00 | 0 | † | 0.00% | 4.61% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 2.74% |
| | [29]-CPLEX | 483.55 | 34.59 | 25 | † | 55.53% | 58.94% |
| | [29]-TRWS | 83.82 | 6.16 | 47 | 190+2775 | 55.69% | 58.98% |
| pdb1fmj | Our-CPLEX | 99.33 | 12.35 | 7 | † | 92.58% | 94.90% |
| | Our-TRWS | 1.05 | 0.60 | 14 | 540+135 | 83.18% | 87.09% |
| | MQPBO-10 | 6.00 | 0.00 | 0 | † | 0.00% | 2.92% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 2.04% |
| | [29]-CPLEX | 77.30 | 16.97 | 11 | † | 15.94% | 18.83% |
| | [29]-TRWS | 16.67 | 3.10 | 11 | 186+677 | 16.18% | 18.91% |
| pdb1i24 | Our-TRWS | 0.06 | 0.02 | 2 | 60+5 | 99.73% | 99.94% |
| | MQPBO-10 | 3.00 | 0.00 | 0 | † | 0.00% | 2.85% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 3.43% |
| | [29]-CPLEX | 5.66 | 5.66 | 0 | † | 100.00% | 100.00% |
| | [29]-TRWS | 0.82 | 0.82 | 0 | 115+0 | 100.00% | 100.00% |
| pdb1iqc | Our-CPLEX | 111.58 | 18.51 | 5 | † | 99.10% | 99.63% |
| | Our-TRWS | 0.74 | 0.40 | 8 | 200+35 | 96.39% | 97.10% |
| | MQPBO-10 | 8.00 | 0.00 | 0 | † | 0.00% | 6.06% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 4.90% |
| | [29]-CPLEX | 229.09 | 24.36 | 20 | † | 35.32% | 41.15% |
| | [29]-TRWS | 36.03 | 4.06 | 28 | 169+2058 | 40.50% | 45.56% |
| pdb1jmx | Our-CPLEX | 142.20 | 15.52 | 9 | † | 97.24% | 98.69% |
| | Our-TRWS | 0.67 | 0.29 | 10 | 200+75 | 93.46% | 95.83% |
| | MQPBO-10 | 8.00 | 0.00 | 0 | † | 0.00% | 3.76% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 3.73% |
| | [29]-CPLEX | 121.71 | 16.21 | 19 | † | 35.86% | 39.98% |
| | [29]-TRWS | 20.02 | 3.59 | 24 | 188+1098 | 35.26% | 39.12% |
| pdb1kgn | Our-CPLEX | 196.03 | 17.97 | 10 | † | 89.22% | 93.23% |
| | Our-TRWS | 1.37 | 0.76 | 9 | 400+170 | 88.92% | 93.16% |
| | MQPBO-10 | 9.00 | 0.00 | 0 | † | 0.00% | 3.24% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 2.27% |
| | [29]-CPLEX | 161.57 | 24.37 | 12 | † | 39.42% | 39.67% |
| | [29]-TRWS | 53.49 | 6.45 | 17 | 268+1824 | 13.20% | 13.36% |
| pdb1kwh | Our-CPLEX | 105.77 | 9.63 | 10 | † | 79.09% | 85.64% |
| | Our-TRWS | 0.46 | 0.27 | 8 | 440+50 | 76.53% | 83.26% |
| | MQPBO-10 | 5.00 | 0.00 | 0 | † | 0.00% | 2.99% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 3.43% |
| | [29]-CPLEX | 51.33 | 12.89 | 9 | † | 25.54% | 31.15% |
| | [29]-TRWS | 9.15 | 2.43 | 8 | 208+401 | 25.43% | 31.13% |
| pdb1m3y | Our-CPLEX | 73.60 | 18.58 | 3 | † | 98.54% | 99.47% |
| | Our-TRWS | 0.79 | 0.65 | 3 | 340+10 | 97.53% | 99.08% |
| | MQPBO-10 | 8.00 | 0.00 | 0 | † | 0.00% | 6.38% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 5.72% |
| | [29]-CPLEX | 120.60 | 25.18 | 14 | † | 31.05% | 27.97% |
| | [29]-TRWS | 28.19 | 4.82 | 12 | 200+1135 | 31.08% | 27.98% |
| pdb1qks | Our-CPLEX | 138.12 | 15.19 | 8 | † | 98.30% | 98.93% |
| | Our-TRWS | 0.30 | 0.12 | 4 | 80+20 | 98.57% | 99.38% |
| | MQPBO-10 | 9.00 | 0.00 | 0 | † | 0.00% | 5.09% |
| | MQPBO | 0.00 | 0.00 | 0 | † | 0.00% | 3.68% |
| | [29]-CPLEX | 96.77 | 15.82 | 12 | † | 28.18% | 26.37% |

| Instance | Algorithm | Time needed overall (s) | Time for initial solution (s) | #iterations Algorithm 1,2 | #iterations TRWS | Logarithmic percentage partial optimality | Percentage excluded labels |
|---|---|---|---|---|---|---|---|
| | [29]-TRWS | 27.99 | 3.24 | 15 | 161+1154 | 30.63% | 28.46% |
| | | | | color-seg | | | |
| colseg-cow3 | Our-TRWS | 66.30 | 48.10 | 6 | 1000+140 | 99.96% | 99.97% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 99.89% |
| | MQPBO-10 | 206.00 | 0.00 | 0 | † | 0.00% | 43.55% |
| | MQPBO | 24.00 | 0.00 | 0 | † | 0.00% | 32.06% |
| | [29]-TRWS | 7530.72 | 690.55 | 14 | 826+6056 | 99.95% | 99.95% |
| colseg-cow4 | Our-TRWS | 91.26 | 48.31 | 13 | 1000+310 | 99.92% | 99.93% |
| | Kovtun | 2.00 | 0.00 | 0 | † | 0.00% | 99.90% |
| | MQPBO-10 | 46.00 | 0.00 | 0 | † | 0.00% | 0.56% |
| | MQPBO | 40.00 | 0.00 | 0 | † | 0.00% | 0.37% |
| | [29]-TRWS | 7395.03 | 742.58 | 10 | 848+6349 | 99.80% | 99.80% |
| colseg-garden4 | Our-TRWS | 0.49 | 0.15 | 5 | 70+20 | 99.91% | 99.94% |
| | Kovtun | 0.00 | 0.00 | 0 | † | 0.00% | 94.96% |
| | MQPBO-10 | 14.00 | 0.00 | 0 | † | 0.00% | 4.27% |
| | MQPBO | 1.00 | 0.00 | 0 | † | 0.00% | 0.21% |
| | [29]-TRWS | 33.68 | 6.75 | 5 | 167+488 | 99.89% | 99.89% |
| | | | | color-seg-n4 | | | |
| clownfish-small | Our-TRWS | 1.72 | 0.68 | 3 | 80+10 | >99.99% | >99.99% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 74.11% |
| | MQPBO-10 | 536.00 | 0.00 | 0 | † | 0.00% | 15.83% |
| | MQPBO | 41.00 | 0.00 | 0 | † | 0.00% | 4.67% |
| | [29]-TRWS | 151.98 | 30.01 | 6 | 223+610 | 99.97% | 99.97% |
| crops-small | Our-TRWS | 1.87 | 1.02 | 2 | 120+5 | 100.00% | 100.00% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 64.70% |
| | MQPBO-10 | 577.00 | 0.00 | 0 | † | 0.00% | 14.32% |
| | MQPBO | 33.00 | 0.00 | 0 | † | 0.00% | 0.71% |
| | [29]-TRWS | 677.08 | 34.88 | 40 | 260+3578 | 99.00% | 99.00% |
| fourcolors | Our-TRWS | 0.57 | 0.08 | 2 | 20+5 | 99.96% | 99.97% |
| | Kovtun | 0.00 | 0.00 | 0 | † | 0.00% | 69.52% |
| | MQPBO-10 | 37.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | MQPBO | 3.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-TRWS | 31.28 | 2.60 | 8 | 34+238 | 99.92% | 99.92% |
| lake-small | Our-TRWS | 1.28 | 0.43 | 2 | 50+5 | 100.00% | 100.00% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 74.87% |
| | MQPBO-10 | 607.00 | 0.00 | 0 | † | 0.00% | 15.31% |
| | MQPBO | 31.00 | 0.00 | 0 | † | 0.00% | 6.65% |
| | [29]-TRWS | 13.75 | 13.75 | 0 | 95+-95 | 100.00% | 100.00% |
| palm-small | Our-TRWS | 2.48 | 1.37 | 3 | 160+10 | >99.99% | >99.99% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 68.65% |
| | MQPBO-10 | 510.00 | 0.00 | 0 | † | 0.00% | 0.48% |
| | MQPBO | 19.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-TRWS | 846.27 | 39.97 | 19 | 291+4582 | 98.20% | 98.20% |
| penguin-small | Our-TRWS | 1.21 | 0.54 | 2 | 90+5 | 100.00% | 100.00% |
| | Kovtun | 0.00 | 0.00 | 0 | † | 0.00% | 91.99% |
| | MQPBO-10 | 193.00 | 0.00 | 0 | † | 0.00% | 1.42% |
| | MQPBO | 13.00 | 0.00 | 0 | † | 0.00% | 1.03% |
| | [29]-TRWS | 15.67 | 15.67 | 0 | 152+-152 | 100.00% | 100.00% |
| pfau-small | Our-TRWS | 18.77 | 7.22 | 48 | 950+470 | 89.43% | 93.41% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 5.59% |
| | MQPBO-10 | 591.00 | 0.00 | 0 | † | 0.00% | 0.70% |
| | MQPBO | 16.00 | 0.00 | 0 | † | 0.00% | 0.00% |
| | [29]-TRWS | 799.08 | 79.34 | 44 | 654+10857 | 10.43% | 10.43% |
| snail | Our-TRWS | 0.79 | 0.23 | 2 | 50+5 | 99.99% | 99.99% |
| | Kovtun | 0.00 | 0.00 | 0 | † | 0.00% | 97.77% |
| | MQPBO-10 | 7.00 | 0.00 | 0 | † | 0.00% | 77.91% |
| | MQPBO | 1.00 | 0.00 | 0 | † | 0.00% | 58.35% |
| | [29]-TRWS | 46.20 | 6.47 | 5 | 83+332 | 99.98% | 99.98% |
| strawberry-glass-2-small | Our-TRWS | 1.35 | 0.60 | 2 | 80+5 | 100.00% | 100.00% |
| | Kovtun | 1.00 | 0.00 | 0 | † | 0.00% | 54.99% |
| | MQPBO-10 | 528.00 | 0.00 | 0 | † | 0.00% | 2.78% |
| | MQPBO | 39.00 | 0.00 | 0 | † | 0.00% | 0.00% |

| Instance | Algorithm | Time needed overall (s) | Time for initial solution (s) | #iterations Algorithm 1,2 | #iterations TRWS | Logarithmic percentage partial optimality | Percentage excluded labels |
|---|---|---|---|---|---|---|---|
| | [29]-TRWS | 311.54 | 31.00 | 11 | 259+1721 | 99.31% | 99.31% |
| | | | mrf-photomontage | | | | |
| family-gm | Our-TRWS | 286.40 | 93.08 | 77 | 1000+1265 | 4.75% | 4.80% |
| | MQPBO-10 | 1087.00 | 0.00 | 0 | † | 0.00% | 4.41% |
| | MQPBO | 90.00 | 0.00 | 0 | † | 0.00% | 4.34% |
| | [29]-TRWS | 12726.45 | 1291.11 | 50 | 1015+22483 | 4.41% | 4.41% |
| pano-gm | Our-TRWS | 320.00 | 112.17 | 59 | 1000+1105 | 67.73% | 79.17% |
| | MQPBO-10 | 646.00 | 0.00 | 0 | † | 0.00% | 28.06% |
| | MQPBO | 97.00 | 0.00 | 0 | † | 0.00% | 40.37% |
| | [29]-TRWS | 14360.45 | 1871.14 | 33 | 911+11193 | 27.55% | 27.55% |
| | | | mrf-stereo | | | | |
| ted-gm | Our-TRWS | 231.97 | 72.67 | 119 | 1000+715 | 67.27% | 72.05% |
| | [29]-TRWS | 3837.51 | 436.30 | 28 | 689+10383 | 38.13% | 38.13% |
| tsu-gm | Our-TRWS | 19.75 | 14.67 | 10 | 670+75 | 99.91% | 99.94% |
| | [29]-TRWS | 9277.99 | 267.55 | 54 | 377+17421 | 0.39% | 0.39% |
| ven-gm | Our-TRWS | 108.73 | 94.44 | 9 | 1000+40 | 0.01% | 0.02% |
| | [29]-TRWS | 14737.47 | 1451.83 | 55 | 993+16592 | 0.00% | 0.00% |

Table 7: Detailed experimental evaluation for Algorithm 1 utilising CPLEX [8] as a subsolver, denoted as `Our-CPLEX`, Algorithm 2 utilising TRW-S [12] as a subsolver, denoted as `Our-TRWS`, their counterparts from [29] denoted by `[29]-CPLEX` and `[29]-TRWS` and MQPBO [11] run for one iteration with predefined label order, denoted by `MQPBO`, and run 10 iterations in 10 random label orders, denoted by `MQPBO-10`.