UDC 519.157

A. Shekhovtsov, V. Hlaváč

On Partial Opimality by Auxiliary Submodular Problems

Доказаны определенные соотношения между тремя различными методами минимизации энергии. Предложено новое достаточное условие частичной оптимальности, основанное на *LP*-релаксации и названное *LP*-автаркией.

Some relations between three different energy minimization techniques are proved. A new sufficient condition of the optimal partial assignment which is based on the LP-relaxation and called LP-autarky is suggested.

Доведено певні співвідношення між трьома різними методами оптимізації енергії. Запропоновано нову достатню умову часткової оптимальності, яка базується на *LP*-релаксації і названа *LP*-автаркією.

Abstract

In this work, we prove several relations between three different energy minimization techniques. A recently proposed methods for determining a provably optimal partial assignment of variables by Ivan Kovtun (IK), the linear programming relaxation approach (LP) and the popular expansion move algorithm by Yuri Boykov. We propose a novel sufficient condition of optimal partial assignment, which is based on LP relaxation and called LP-autarky. We show that methods of Kovtun, which build auxiliary submodular problems, fulfill this sufficient condition. The following link is thus established: the LP relaxation cannot be tightened by IK. For non-submodular problems this is a non-trivial result. In the case of two labels, LP relaxation provides optimal partial assignment, known as persistency, which, as we show, dominates IK. Relating IK with an expansion move, we show that the set of fixed points of expansion move with any «truncation» rule for the initial problem and the problem restricted by one-vs-all method of IK would coincide – i.e. the expansion move cannot be improved by this method. In the case of two labels, expansion move with a particular truncation rule coincide with one-vs-all method.

1. Introduction

1.1. Energy Minimization

In this work[1] we consider the minimization problem of the following form:

$$\min_{x\in\mathcal{L}}\left[f_0+\sum_{s\in\mathcal{V}}f_s(x_s)+\sum_{st\in\mathcal{E}}f_{st}(x_{st})\right]=\min_{x\in\mathcal{L}}\boldsymbol{f}(x). \qquad (1)$$

Keywords: energy minimization, partial optimality, persistency, max-sum, WCSP, MRF, autarky, LP-relaxation, expansion move.

Here, $\mathcal{V}$ is a finite set and $\mathcal{E}\subset\mathcal{V}\times\mathcal{V}$. A concatenated vector of all variables $x=(x_s\mid s\in\nu)$ is called a *labeling*. Variable x_s takes its values in a discrete *domain* $\mathcal{L}_s$, called *labels*. Labeling x takes values in $\mathcal{L}$, the Cartesian product of all domains $\mathcal{L}_s$. In this paper all $\mathcal{L}_s$ will have the same number of labels, but may have different associated orderings, etc. Notation *st* denotes the ordered pair (s,t) and x_{st} denotes the pair of corresponding variables, (x_s,x_t). The objective is composed of term $f_0\in\mathbb{R}$ and functions $f_s:\mathcal{L}_s\to\mathbb{R}$ and $f_{st}:\mathcal{L}_s\times\mathcal{L}_t\to\mathbb{R}$.

The problem (1) is considered in several fields. It is also known as the labeling problem, the Weighted Constraint Satisfaction (WCSP) and for the case of two labels ($|L_s|=2,\forall s$) as the pseudo-Boolean[2] optimization [1]. Our terminology comes from considering probabilistic models in the form of Gibbs distribution. There is certain difference between problems with two labels and more than two labels, the later will be referred to as *multi-label* problems.

1.2. A Partial Optimality

An energy minimization (1) is an NP-hard problem in general. Techniques which allow us to find a «part of the optimal» labeling are of our central interest here. The idea is that it may be

[1]The work was supported bu EU projects FP7-ICT-247870 NIFTi and FP7-ICT-247525 HUMAVIPS and the Czech project 1M0567 CAK.

[2]Variables $x_s\in\{0,1\}$ are regarded as Boolean in this case and a «pseudo» emphasize, that a real-valued rather than Boolean function of these variables is considered.

possible to fix a part of variables to take certain labels such that any optimal labeling will provably have the same partial assignment.

More precisely, we consider a subset of variables $\mathcal{A} \subset \mathcal{V}$ and the assignment of labels over this subset $y = (y_s | s \in \mathcal{A})$. The pair $(\mathcal{A}, y)$ is called a *strong optimal partial assignment* (strong persistency [3]), if for any minimizer x^* it holds $x^*_{\mathcal{A}} = y$, where notation $x^*_{\mathcal{A}}$ is the restriction of x^* to $\mathcal{A}$, $(x^*_s | s \in \mathcal{A})$. Likewise, if there exist at least one minimizer x^*, for which $x^*_{\mathcal{A}} = y$ holds we say that $(\mathcal{A}, y)$ is a *weak optimal partial assignment*.

Two or more strong optimal partial assignments can be combined together, because each of them preserves all optimal solutions. This is not true for weak assignments, even if they assign different variables, – they may not share any globally optimal solutions in common. However, if we want to find a minimizer of (1) (or at least «localize» it as much as possible), a weak optimal partial assignment could be more helpful – the best one assigns all variables.

1.3. Domain Constraints

The idea of the optimal partial assignment naturally extends to constraining a variable to a subset of labels $K_s \subset \mathcal{L}_s$. Let $\mathcal{A} \subset \mathcal{V}$, let $K_s \subset \mathcal{L}_s$, $\forall s \in \mathcal{A}$. Let $\mathbf{K}$ be the Cartesian product of K_s, $s \in \mathcal{A}$. We say that a pair $(\mathcal{A}, \mathbf{K})$ is a strong (resp. weak) optimal constraint if $x^*_{\mathcal{A}} \in \mathbf{K}$ for all (resp. at least one) minimizer x^*. This type of constraints is called *domain constraints*. Obviously, it includes partial assignment as a special case.

1.4. Autarkies

Some domain constraints follow from more specific properties called «autarkies». This term occurs in [3] for two-label problems and we consider its extension [15] to multi-label problems.

Let $\mathcal{L}_s = \{0,1\dots L\}$ $\forall s \in \mathcal{V}$, $L \in \mathbb{N}$. Let $x, y \in \mathcal{L}$. Define component-wise minimum and maximum of two labellings:

$$(x \wedge y)_s = \min(x_s, y_s), \quad (2a)$$

$$(x \vee y)_s = \max(x_s, y_s). \quad (2b)$$

A pair ($x^{\min} \in \mathcal{L}$, $x^{\max} \in \mathcal{L}$) such that $x^{\min} \le x^{\max}$ (component-wise) is called a *weak autarky* for problem (1), if

$$\forall x \in \mathcal{L} \ f((x \vee x^{\min}) \wedge x^{\max}) \le \boldsymbol{f}(x). \quad (3)$$

If additionally for any $x \ne (x \vee x^{\min}) \wedge x^{\max}$ strict inequality

$$\boldsymbol{f}((x \vee x^{\min}) \wedge x^{\max}) < \boldsymbol{f}(x) \quad (4)$$

holds, then the autarky is called *strong*.

The autarky provides domain constraints with $K_s = [x^{\min}_s, \dots, x^{\max}_s]$. For any minimizer x^*, we have that $\hat{x} = (x^* \vee x^{\min}) \wedge x^{\max}$ is a minimizer as well, and $\hat{x}_s \in K_s$. A strong autarky guarantees additionally that x^* must itself satisfy $x^*_s \in K_s$. Indeed, if it was not true then $\hat{x} \ne x^*$ and $\boldsymbol{f}(\hat{x}) < \boldsymbol{f}(x^*)$, which is a contradiction. Therefore a weak (resp. strong) autarky provides a weak (resp. strong) domain constraint.

A determining whether a given pair $(x^{\min}, x^{\max})$ is a strong autarky is an NP-hard decision problem [3].

The autarkies can be combined together. A *join* of two autarkies (x^1, x^2), (y^1, y^2) is the pair $(x^1 \vee y^1, x^2 \wedge y^2)$. For strong autarkies, the result is a strong autarky and this operation is commutative, associative and idempotent, so that it defines a semilattice.

Proof. From definition of autarkies, we have

$$\begin{aligned} &\boldsymbol{f}((((x \vee x^1) \wedge x^2) \vee y^1) \wedge y^2) \le \\ &\le \boldsymbol{f}((x \vee x^1) \wedge x^2) \le \boldsymbol{f}(x). \end{aligned} \quad (5)$$

Note, that for $x^1 \le x^2$ we have $(x \vee x^1) \wedge x^2 = (x \wedge x^2) \vee x^1$. We can rewrite the labeling in the left hand side (LHS) as follows

$$\begin{aligned} &((x \vee x^1) \wedge x^2) \vee y^1) \wedge y^2 = \\ &((x \wedge x^2) \vee (x^1 \vee y^1)) \wedge y^2 \doteq \\ &(x \wedge (x^2 \wedge y^2)) \vee (x^1 \vee y^1), \end{aligned} \quad (6)$$

where doted equality holds if $y^2 \ge x^1$. This is satisfied for strong autarkies, because it would be a contradiction that all optimal labellings are below y^2 and above x^1. □

Thus there exists an autarky, which provides the maximal amount of domain constraints among strong autarkies. It is the join of all strong autarkies.

It is also possible to join «non-contradictive» weak autarkies together, but let us leave it aside for now.

We will consider a special cases of autarkies with «one-side constraints», of the form $(x^{\min}, L)$ or $(0, x^{\max})$, where L and 0 represent the labeling with all components equal to L (resp. 0). For such autarkies inequality (4) simplifies, because $x \vee 0 = x$ and $x \wedge L = x$. Methods [10, 11] compute strong autarkies of this form. By taking the join of two strong autarkies $(x^{\min}, L)$ and $(0, x^{\max})$ we can obtain a strong autarky $(x^{\min}, x^{\max})$. However, the reverse is not true: if $(x^{\min}, x^{\max})$ is a strong autarky, it does not imply that $(x^{\min}, L)$ or $(0, x^{\max})$ is an autarky. And it is the case that other methods (roof-dual [1] in the case of two-label problem and its multi-label extension [15]) can find an autarky of the form $(x^{\min}, x^{\max})$, which is not a join of two one-side autarkies.

1.5. Submodular Problems

A function $\boldsymbol{f}$ is called *submodular* if

$$\forall x, y \in \mathcal{L}\ \boldsymbol{f}(x \vee y) + \boldsymbol{f}(x \wedge y) \le \boldsymbol{f}(x) + \boldsymbol{f}(y). \quad (7)$$

In the case $\boldsymbol{f}$ is defined by (1), it is submodular iff (see *e.g.* [18]) $\forall st \in \mathcal{E}, \forall x_{st}, y_{st} \in \mathcal{L}_{st} = \mathcal{L}_s \times \mathcal{L}_t$

$$f_{st}(x_{st}) + f_{st}(y_{st}) \ge f_{st}(x_{st} \wedge y_{st}) + f_{st}(x_{st} \vee y_{st}). \quad (8)$$

Minimizing a pairwise submodular function reduces to mincut problem [7], [13]. Let $\boldsymbol{f}$ be submodular and x^* be its minimizer. Then we have the following properties:

$$\boldsymbol{f}(x \vee x^*) \le \boldsymbol{f}(x), \quad (9a)$$

$$\boldsymbol{f}(x \wedge x^*) \le \boldsymbol{f}(x). \quad (9b)$$

They easily follow from submodularity, noting that $\boldsymbol{f}(x \vee x^*) \ge \boldsymbol{f}(x^*)$ and $\boldsymbol{f}(x \vee x^*) \ge \boldsymbol{f}(x^*)$. So, in fact, any pair of optimal solutions (x^{1*}, x^{2*}) is a weak autarky for this problem. Moreover, if we let

$$x^{\min} = \bigwedge \arg\min_x \boldsymbol{f}(x), \quad (10a)$$

$$x^{\max} = \bigvee \arg\min_x \boldsymbol{f}(x), \quad (10b)$$

where argmin is the set of minimizers, we see that both $x^{\min}$ and $x^{\max}$ are minimizers of $\boldsymbol{f}$ and that $(x^{\min}, x^{\max})$ is a strong autarky for $\boldsymbol{f}$. In fact, it is the join of all strong autarkies for $\boldsymbol{f}$. This strong autarky can be determined from a solution of the corresponding maxflow problem.

2. An Approach by Kovtun

In this section, we review the techniques [10, 11] for building autarkies (and hence domain constrains) by constructing auxiliary problems. We take a somewhat different perspective on these results, however, our statements and proofs here are in a sense equivalent to ones given in [10, 11].

Theorem 1. *Let* $\boldsymbol{f} = \boldsymbol{g} + \boldsymbol{h}$, *let* $(x^{\min}, x^{\max})$ *be a strong autarky for* $\boldsymbol{g}$ *and a weak autarky for h. Then* $(x^{\min}, x^{\max})$ *is a strong autarky for* $\boldsymbol{f}$.

Proof. We have

$$\boldsymbol{f}((x \vee x^{\min}) \wedge x^{\max}) = \boldsymbol{g}((x \vee x^{\min}) \wedge x^{\max}) + \\ + \boldsymbol{h}((x \vee x^{\min}) \wedge x^{\max}) \le \boldsymbol{g}(x) + \boldsymbol{h}(x), \quad (11)$$

and the inequality is strict if $(x^{\min}, x^{\max})$ is strong for either $\boldsymbol{h}$ or $\boldsymbol{g}$. □

The idea of auxiliary problems is to construct a submodular $\boldsymbol{g}$, for which, as we know, a strong autarky $(x^{\min}, L)$ can be found by choosing $x^{\min}$ as the lowest minimizer of $\boldsymbol{g}$, given by (10a). The trick is to find such g that $(x^{\min}, L)$ is at the same time an autarky for $h = \boldsymbol{f} - \boldsymbol{g}$. The following sufficient conditions were proposed [11]:

Statement 1. *Let* $\boldsymbol{h}$ *satisfy*

$$\forall s \in \mathcal{V}, x_s \in \mathcal{L}, \hat{x}_s \in K_s$$
$$h_s(x_s \vee \hat{x}_s) \le h_s(\hat{x}_s) \quad (12a)$$

$$\forall st \in \mathcal{E}, x_{st} \in \mathcal{L}, \hat{x}_{st} \in K_{st}$$
$$h_{st}(x_{st} \vee \hat{x}_{st}) \le h_{st}(x_{st}). \quad (12b)$$

Then for any $x^{\min}$ such that $x_s^{\min} \in K_s$, the pair $(x^{\min}, L)$ is a weak autarky for $\boldsymbol{h}$. If additionally

$$\forall s \in \mathcal{V}, \hat{x}_s \in K_s, x_s < \hat{x}_s$$
$$h_s(x_s \vee \hat{x}_s) < h_s(\hat{x}_s), \quad (12c)$$

then $(x^{\min}, L)$ is a strong autarky.

Proof. For any $x \in \mathcal{L}$, summing corresponding inequalities from (12a) and (12b), we obtain

$$\sum_s h_s(x_s \vee x_s^{\min}) + \sum_{st} h_{st}(x_{st} \vee x_{st}^{\min}) \le \sum_s h_s(x_s) + \sum_{st} h_{st}(x_{st}). \quad (13)$$

If $x \vee x^{\min} \ne x$, then $\exists s \in \mathcal{V}\ x_s < x_s^{\min}$ and (11) implies strict inequality. □

Two practical methods were proposed [11] to construct $\boldsymbol{g}$ and $(K_s | s \in \mathcal{V})$. We first describe a more general approach.

Algorithm 1: Sequential construction of $\boldsymbol{g}$, $(K_s \mid s \in \mathcal{V})$, [11]

---

1. Start with $K_s = \varnothing$, $s \in \mathcal{V}$;
2. Find g such that $\boldsymbol{h} = \boldsymbol{f} - \boldsymbol{g}$ satisfies (12) and $\boldsymbol{g}$ satisfies submodularity constraints (8);
3. Find $x^{\min} = \wedge \arg\min_x \boldsymbol{g}(x)$;
4. If $x_s^{\min} \in K_s$ for all $s \in \mathcal{V}$ then stop;
5. Set $K_s \leftarrow K_s \cup \{x_s^{\min}\}$ $\forall s \in \mathcal{V}$ and go to step 2.

In step 2 for each edge $st \in \mathcal{E}$ a system of linear inequalities in g_{st} has to be solved. While [11] provides an explicit solution, for our consideration it will not be necessary. When the algorithm stops, g is submodular and $(x^{\min}, L)$ is a strong autarky for $\boldsymbol{g}$ and a weak autarky for $\boldsymbol{f} - \boldsymbol{g}$. By Theorem 1, it is a strong autarky for $\boldsymbol{f}$. It may stop, however, with $x_s^{\min} = 0$ for all s, so that efficiently no constraints are derived. Being a polynomial algorithm it cannot have a guarantee to simplify the problem (1).

A simpler non-iterative method proposed in [10] is shown in Algorithm 2. It attempts to identify nodes s where the label L is better than any other label. The constructed auxiliary problem $\boldsymbol{g}$ has a property that its lowest minimizer $x^{\min} = \wedge \arg\min_x \boldsymbol{g}(x)$ is guaranteed to satisfy $x^{\min} \in K_s$ $\forall s \in \mathcal{V}$. (because all costs $(g_{st}(i,j) | i < L, j < L)$ are equal and $g_s(0) \le g_s(i)\ \forall s \in \mathcal{V}, \forall i < L$, see proof in [10]). Therefore $(x^{\min}, L)$ is a weak autarky for $\boldsymbol{f} - \boldsymbol{g}$ and Theorem (1) applies.

Both methods allow us to choose various orderings of sets $\mathcal{L}_s$. Strong domain constraints derived from various orderings can be then combined.

Algorithm 2: One vs all method, [10, 11]

---

1. For each s chose such ordering of $\mathcal{L}_s$ that $0 \in \arg\min_{i \ne L} f_s(i)$;
2. Set $g_s = f_s$, $s \in \mathcal{V}$;
3. Set $K_s = \{0, L\}$
4. Set $g_{st}(i,j) = \begin{cases} a_{st}, & i = L,\ j = L, \\ b_{st}, & i = L,\ j \ne L, \\ c_{st}, & i \ne L,\ j = L, \\ d_{st}, & i \ne L,\ j \ne L, \end{cases}$

where $a_{st}, b_{st}, c_{st}, d_{st}$ are such that $f_{st} - g_{st}$ satisfy (12b) and submodularity constraints. One of the solutions is as follows:

$$\begin{aligned} a_{st} &= f_{st}(L,L), \\ b_{st} &= \min_{j \ne L} f_{st}(L,j), \\ c_{st} &= \min_{i \ne L} f_{st}(i,L), \\ d_{st} &= \min(b_{st} + c_{st} - a_{st}, \min_{i \ne L, j \ne L} [f_{st}(i,j) \\ &\quad + \min\{b_{st} - f_{st}(L,j), c_{st} - f_{st}(i,L)\}]). \end{aligned} \quad (14)$$

---

3. LP-autarkies

In this section we introduce a special subclass of autarkies, which preserve optimal solutions of the LP-relaxation. Unlike with general autarkies, the membership to this subclass is polynomially verifiable. We show that autarkies constructed by algorithms 1, 2 belong to this subclass. This has useful implications for LP relaxation.

3.1. LP Relaxation

Let $\phi(x)$ be a vector with components $\phi(x)_0 = 1$, $\phi(x)_{s,i} = [[x_s = i]]$ and $\phi(x)_{st,ij} = [[x_{st} = ij]]$, where $[[\cdot]]$ is 1 if the expression inside is true and 0 otherwise. Let f denote a vector with components f_0, $f_{s,i} = f_s(i)$ and $f_{st,ij} = f_{st}(ij)$. With respect to components of energy functions we will be using this index and parenthesis notations completely interchangeably. Let $\langle \cdot, \cdot \rangle$ denote a scalar product. Then we can write energy minimization as

$$\min_{x \in \mathcal{L}} \langle f, \phi(x) \rangle. \quad (15)$$

Its relaxation to a linear program is written as

$$\min_{\mu\in\Lambda}\langle f,\mu\rangle, \tag{16}$$

where Λ is the *local polytope*. It approximates $conv\{\phi(x)|x\in\mathcal{L}\}$ from the outside, see [18] for more detail. It is given by the linear constraints

$$\begin{aligned}
&\mu_0=1,\\
&\mu_{s,i}\ge 0,\quad \mu_{st,ij}\ge 0,\\
&\sum_{ij\in\mathcal{L}_{st}}\mu_{st,ij}=1 \quad \forall st\in\mathcal{E},\\
&\sum_{j\in\mathcal{L}_t}\mu_{st,ij}=\mu_{s,i} \quad \forall i\in\mathcal{L}_s, st\in\mathcal{E},\\
&\sum_{i\in\mathcal{L}_s}\mu_{st,ij}=\mu_{t,j} \quad \forall j\in\mathcal{L}_t, st\in\mathcal{E}.
\end{aligned} \tag{17}$$

Vector $\mu\in\Lambda$ is called a *relaxed* labeling.

3.2. LP-autarky

We now extend the notion of autarky to relaxed labellings.

Definition 1. *A binary operation* $\bar{\wedge}:\Lambda\times\mathcal{L}\to\Lambda$, *is defined as follows. Let* $y\in\mathcal{L}$ *and* $\mu\in\Lambda$. *Then* $\nu=\mu\bar{\wedge}y\in\Lambda$ *is constructed as*:

$$\nu_{s,i}=\begin{cases}\mu_{s,i}, & i<y_s,\\ \sum_{i'\ge y_s}\mu_{s,i'}, & i=y_s,\\ 0, & i>y_s;\end{cases} \tag{18a}$$

$$\nu_{st,ij}=\begin{cases}\mu_{st,ij}, & i<y_s, j<y_t,\\ \sum_{i'\ge y_s}\mu_{st,i'j}, & i=y_s, j<y_t,\\ \sum_{j'\ge y_t}\mu_{st,ij'}, & i<y_s, j=y_t,\\ \sum_{\substack{i'\ge y_s\\ j'\ge y_t}}\mu_{st,i'j'}, & i=y_s, j=y_t,\\ 0, & i>y_s \ or\ j>y_t.\end{cases} \tag{18b}$$

By construction, the relaxed labeling ν has non-zero weights only for labels «below» *y*: $\nu_{s,i}=0$ for $i>y_s$ and the same for pairs *st*, *ij*. Let us check that $\nu\in\Lambda$.

Proof. Normalization constraint:

$$\sum_i\nu_{s,i}=\sum_{i<y_s}\mu_{s,i}+\sum_{i'\ge y_s}\mu_{s,i'}=\sum_i\mu_{s,i}=1. \tag{19}$$

Marginalization constraint:

$$\sum_i\nu_{st,ij}=\begin{cases}\sum_{i<y_s}\mu_{st,ij}+\sum_{i'\ge y_s}\mu_{st,i'j}, & j<y_t,\\ \sum_{\substack{i<y_s\\ j'\ge y_t}}\mu_{st,ij}+\sum_{\substack{i'\ge y_s\\ j'\ge y_t}}\mu_{st,i'j'}, & j=y_t,\\ 0, & j>y_t,\end{cases} \tag{20}$$

$$=\nu_{t,j}.$$ □

The operation $\nu=\mu\underline{\vee}y$ is defined completely similarly, having singleton components

$$(\mu\underline{\vee}y)_{s,i}=\begin{cases}\mu_{s,i}, & i>y_s,\\ \sum_{i'\le y_s}\mu_{s,i'}, & i=y_s,\\ 0, & i<y_s.\end{cases} \tag{21}$$

Definition 2. *We say that a pair* $(x^{\min},x^{\max})$ *is a weak LP-autarky for f, if*

$$\forall\mu\in\Lambda\quad \langle f,(\mu\bar{\wedge}x^{min})\underline{\vee}x^{max}\rangle\le\langle f,\mu\rangle. \tag{22}$$

If additionally for all μ such that $(\mu\bar{\wedge}x^{\min})\underline{\vee}$ $\underline{\vee}x^{\max}\ne\mu$ the strict inequality holds then we say that it is a strong LP-autarky.

3.3. The Properties of LP-autarkies

Statement 2. *Any weak* (*resp. strong*) *LP-autarky is a weak (resp. strong) autarky.*

Proof. By substituting $\mu=\phi(x)$. □

Statement 3. *Checking whether* $(x^{\min},x^{\max})$ *is an LP-autarky for f can be solved in a polynomial time.*

Proof. By construction, $(\mu\bar{\wedge}x^{\min})\underline{\vee}\,x^{\max}$ is a linear map in μ, let us denote it $A\mu$. Inequality (22) holds iff

$$\min_{\mu\in\Lambda}\langle f,\mu-A\mu\rangle\ge 0, \tag{23}$$

which is a linear program. To verify whether *A* is a strong LP-autarky we need to solve

$$\min\langle f,\mu-A\mu\rangle>0$$
$$s.t.\begin{cases}\mu\in\Lambda,\\ \sum_s\sum_{x^{\min}\le i\le x^{\max}}\mu_{s,i}<|\mathcal{V}|.\end{cases} \tag{24}$$

□

Statement 4. *If* ***f*** *is submodular*, *then*

$$\forall\mu\in\Lambda,\forall y\in\mathcal{L}$$
$$\langle\mu,f\rangle+\langle\phi(y),f\rangle\ge\langle\mu\bar{\wedge}y,f\rangle+\langle\mu\underline{\vee}y,f\rangle. \tag{25}$$

Proof. Scalar products in (25) are composed of sums of singleton terms and pairwise terms. We first show that sums of singleton terms are equal, expanding singleton terms in the right hand side (RHS):

$$\begin{aligned}&\sum_s\sum_i[(\mu\bar{\wedge} y)_{s,i} \quad+(\mu\underline{\vee} y)_{s,i}]f_s(i)=\\&\sum_s\sum_{i<y_s}\mu_{s,i}f_s(i) \quad+\sum_s\sum_{i'\geq y_s}\mu_{s,i'}f_s(y_s)+\\&\sum_s\sum_{i>y_s}\mu_{s,i}f_s(i) \quad+\sum_s\sum_{i'\leq y_s}\mu_{s,i'}f_s(y_s)=\\&\sum_s\sum_i\mu_{s,i}f_s(i) \quad+\sum_s(\sum_{i'}\mu_{s,i'})f_s(y_s)=\\&\sum_s\sum_i\mu_{s,i}f_s(i) \quad+\sum_s\sum_i[[i=y_s]]f_s(y_s).\end{aligned} \tag{26}$$

Now consider submodularity constraints:

$$\begin{aligned}&\forall st\in\mathcal{E},\forall ij\in\mathcal{L}_{st},\forall y_{st}\in\mathcal{L}_{st}\\&f_{st}(ij)+f_{st}(y_{st})\geq f_{st}(ij\wedge y_{st})+f_{st}(ij\vee y_{st}).\end{aligned} \tag{27}$$

Multiplying this inequality by $\mu_{st,ij}$ and summing over ij, we obtain on the LHS:

$$\begin{aligned}&\sum_{ij}\mu_{st,ij}f_{st}(ij)+f_{st}(y_{st})=\\&=\sum_{ij}\mu_{st,ij}f_{st}(ij)+\sum_{ij}[[ij=y_{st}]]f_{st}(y_{st})\end{aligned} \tag{28}$$

and on the RHS:

$$\begin{aligned}&\sum_{ij}\mu_{st,ij}[f_{st}(ij\wedge y_{st})+f_{st}(ij\vee y_{st})]=\\&=\sum_{ij}[(\mu\bar{\wedge} y)_{st,ij}+(\mu\bar{\vee} y)_{st,ij}]f_{st}(ij),\end{aligned} \tag{29}$$

where the equality is verified as follows:

$$\begin{aligned}&\sum_{ij}\mu_{st,ij}f_{st}(ij\wedge y_{st})=\\&=\sum_{\substack{i<y_s\\ j<y_t}}\mu_{st,ij}f_{st}(ij)+\sum_{\substack{i\geq y_s\\ j<y_t}}\mu_{st,ij}f_{st}(y_s,j)+\\&+\sum_{\substack{i<y_s\\ j\geq y_t}}\mu_{st,ij}f_{st}(i,y_t)+\sum_{\substack{i\geq y_s\\ j\geq y_t}}\mu_{st,ij}f_{st}(y_{st})=\\&=\sum_{ij}(\mu\bar{\wedge} y)_{st,ij}f_{st,ij}.\end{aligned} \tag{30}$$

The term with $\underline{\vee}$ is rewritten similarly. By summing inequalities (28) $\geq$ (29) over $st\in\mathcal{E}$ and adding equalities (26) of the singleton terms, we get the result. □

Statement 5. *Let* ***f*** *be submodular and* $x^*\in\arg\min_x \boldsymbol{f}(x)$. *Then* $\forall\mu\in\Lambda$

$$\left\langle\mu\bar{\wedge} x^*,f\right\rangle\leq\left\langle\mu,f\right\rangle, \tag{31a}$$

$$\left\langle\mu\underline{\vee} x^*,f\right\rangle\leq\left\langle\mu,f\right\rangle. \tag{31b}$$

Proof. Let us show (37). For submodular problems LP-relaxation (21) is tight. Thus for any $\nu\in\Lambda$ there holds $\left\langle\nu,f\right\rangle\geq\boldsymbol{f}(x^*)=\left\langle\varphi(x^*),f\right\rangle$. In particular, for $\nu=\mu\bar{\wedge} y$ we have $\left\langle\mu\bar{\wedge} y,f\right\rangle\geq$ $\geq\left\langle\phi(x^*),f\right\rangle$, which when combined with (31) implies the statement. □

Statement 6. *Let* $(x^{\min},L)$ *be a strong LP-autarky for* ***f****, then*:

$$\begin{aligned}&\forall s\in\mathcal{V},\ \forall i<x_s^{\min},\ \forall\mu^*\in\\&\in\arg\min_{\mu\in\Lambda}\left\langle\mu,f\right\rangle\mu^*_{s,i}=0.\end{aligned} \tag{32}$$

Proof. Let $\mu^*\in\arg\min_{\mu\in\Lambda}\left\langle\mu,f\right\rangle$ and $\mu^*_{s,i}>0$. Then $\mu^*\underline{\vee} x^{\min}\neq\mu^*$ and $\boldsymbol{f}(\mu^*\underline{\vee} x^{\min})<\boldsymbol{f}(\mu^*)$, which contradicts optimality of μ^*. □

3.4. Implications for the Algorithms 1, 2

We have already seen in the statement 5 that for a submodular problem ***g***, taking y as a minimizer (resp. the lowest minimizer) of ***g*** gives a weak (resp. strong) LP-autarky (y,L). Let us show that statement 1 extends to LP-autarkies too. This would imply that autarkies derived by algorithms 1, 2 are in fact LP-autarkies for $\boldsymbol{f}=\boldsymbol{g}+\boldsymbol{h}$.

Statement 7. *Let* ***h*** *satisfy inequalities* (12). *Then for any* $y\in\mathcal{L}$ *such that* $y_s\in K_s$*, the pair* (y,L) *is a weak LP-autarky for* ***h****.*

Proof. Let $\mu\in\Lambda$. From inequality (12a) we have

$$\sum_s\sum_i((\mu\underline{\vee} y)_{s,i}-\mu_{s,i})h_{s,i}\leq 0. \tag{33}$$

Multiplying (12b) by $\mu_{st,ij}$ and summing over $ij\in\mathcal{L}_{st}$ and over $st\in\mathcal{E}$ we obtain

$$\sum_{st}\sum_{ij}[(\mu\underline{\vee} y)_{st,ij}-\mu_{st,ij}]h_{st,ij}\leq 0. \tag{34}$$

Adding (33) and (34), we get:

$$\left\langle\mu\underline{\vee} y-\mu,h\right\rangle\leq 0, \tag{35}$$

which is equivalent to (22). □

We have shown that algorithms 1, 2 derive domain constraints in the form of strong LP-autarkies. We know too that optimal solutions of

LP-relaxation will obey domain constraints derived via strong LP-autarkies. Note, while algorithms 1, 2 depend on the ordering of the labels, solutions of the LP-relaxation does not. Hence.

Corollary 1. *Let* $(K_s \subset \mathcal{L}_s | s \in \mathcal{V})$ *be a strong domain constraint derived by Algorithms* 1, 2 *any ordering of sets* $\mathcal{L}_s$*. Then the set of optimal solutions of LP relaxation with and without these domain constraints would coincide.*

We proved that LP relaxation cannot be tightened by algorithms 1, 2. It may only be simplified by eliminating all variables which are guaranteed to be 0 in every optimal solution. This may be useful in practical methods solving LP relaxation.

For problems with two labels, the following relation also holds. Let $\Lambda^* = \arg\min_{\mu\in\Lambda}\langle f, \mu\rangle$. Let

$$\begin{aligned} x_s^{\min} &= \min\{i \mid \exists\mu^* \in \Lambda^* \ \ \mu_{s,i} > 0\}, \\ x_s^{\max} &= \max\{i \mid \exists\mu^* \in \Lambda^* \ \ \mu_{s,i} > 0\}, \end{aligned} \quad (36)$$

then $(x^{\min}, x^{\max})$ is a strong autarky for $\boldsymbol{f}$. This is the *roof-dual* autarky [1]. Because for any other autarky derived via the algorithms 1 and 2 the statement 6 holds, we conclude that the roof-dual autarky dominates algorithms 1 and 2.

4. Expansion Move

An expansion move algorithm [4] seeks to improve the current solution x by considering a *move*, which for every $s \in \mathcal{V}$ either keeps the current label x_s or changes it to the label k.

Algorithm 3: Expansion-Move [4]

---

1. Let $x \in \mathcal{L}$, let $k \in \mathcal{L}$. The *move energy* function $\boldsymbol{g}(z)$ of binary configuration $z \in \{0,1\}^{\mathcal{V}}$ is defined by:

$$\begin{aligned} &g_0 = f_0, \quad g_s(0) = f_s(x_s), \quad g_s(1) = f_s(k), \\ &g_{st}(1,1) = f_{st}(k,k), \quad g_{st}(1,0) = f_{st}(k,x_t), \\ &g_{st}(0,1) = f_{st}(x_s,k), \quad g_{st}(0,0) = f_{st}(x_s,x_t). \end{aligned} \quad (37)$$

2. Let $z^* \in \arg\min_z \boldsymbol{g}(z)$;

3. If $\boldsymbol{g}(z^*) < \boldsymbol{g}(0)$, assign $x_s \leftarrow \begin{cases} x_s, & \text{if } z_s = 0, \\ k, & \text{if } z_s = 1. \end{cases}$

If the above procedure is repeated for all labels $k \in \mathcal{L}$ and no improvement to x is found then x is said to be a fixed point of this method.

In the case $\boldsymbol{f}$ is a metric energy [4], the move energy g is submodular for arbitrary x and step 2 is easy.

Statement 8. *Let* $\boldsymbol{f}$ *be metric* [4]. *Let* $(x^{\min}, L)$ *be a strong autarky for* $\boldsymbol{f}$ *such that* $x_s^{\min} \in \{0, L\}$, $\forall s \in \mathcal{V}$. *Then for any fixed point* x *of the expansion-move algorithm there holds*

$$x \geq x^{\min}. \quad (38)$$

Proof. Assume $\exists s \in \mathcal{V}$ such that $x_s < x^{\min}$. Then $\boldsymbol{f}(x \vee x^{\min}) < \boldsymbol{f}(x)$ and since $x_s \in \{1, L\}$, it is

$$x_s \vee x_s^{\min} = \begin{cases} x_s, & x_s^{\min} = 1, \\ L, & x_s^{\min} = L, \end{cases} \quad (39)$$

which is a valid expansion move from x to label $k = L$, strictly improving the energy. □

In the case when a move energy is not submodular, it can be «truncated» to make it submodular while still preserving the property that the move does not increase $f(x)$ [12]. Let $\Delta_{st} = g_{st}(1,1) + g_{st}(0,0) - g_{st}(0,1) - g_{st}(1,0)$. Pair st is submodular iff $\Delta_{st} < 0$.

Definition 3. *The truncation* $\boldsymbol{g}'$ *of* $\boldsymbol{g}$ *is different from* $\boldsymbol{g}$ *only in non-submodular pairwise components of* g, *which are set as*:

$$\begin{aligned} g'_{st,00} &= g_{st,00} - \beta_{st}\Delta_{st}, \\ g'_{st,01} &= g_{st,01} + \alpha_{st}\Delta_{st}, \\ g'_{st,10} &= g_{st,10} + (1 - \alpha_{st} - \beta_{st})\Delta_{st}, \\ g'_{st,11} &= g_{st,11}, \end{aligned} \quad (40)$$

where α_{st} and β_{st} are free parameters, satisfying $\alpha_{st} \geq 0$, $\beta_{st} \geq 0$, $\alpha_{st} + \beta_{st} \leq 1$.

It is easy to verify that $\boldsymbol{g}'$ is submodular, and

$$\boldsymbol{g}(z) - \boldsymbol{g}(0) \leq \boldsymbol{g}'(z) - \boldsymbol{g}'(0), \quad (41)$$

saying that increase in g is no more than increase in $\boldsymbol{g}'$ when changing from 0 to z.

Proof. *of* (41). By construction of $\boldsymbol{g}'$, for all $st \in \mathcal{E}$ such that $\Delta_{st} > 0$, enumerating all z_{st},

$$\begin{aligned} &g'_{st,00} - g'_{st,00} = 0, \\ &g'_{st,01} - g'_{st,00} = g_{st,01} - g_{st,00} + (\alpha_{st} + \beta_{st})\Delta_{st}, \\ &g'_{st,10} - g'_{st,00} = g_{st,10} - g_{st,00} + (1 - \alpha_{st})\Delta_{st}, \\ &g'_{st,11} - g'_{st,00} = g_{st,11} - g_{st,00} + \beta_{st}\Delta_{st}, \end{aligned} \quad (42)$$

we see that only positive values are added on RHS. It is also seen that the added positive values do only increase with β_{st}. This means that the truncation with $\beta_{st} > 0$ (let's denote it $\boldsymbol{g}^{\alpha,\beta}$) is *never better* than the truncation with $\beta = 0$ (let's denote it g^{α}): $\forall z$

$$\boldsymbol{g}(z) - \boldsymbol{g}(0) \le \boldsymbol{g}^{\alpha}(z) - \boldsymbol{g}^{\alpha}(0) \le \boldsymbol{g}^{\alpha,\beta}(z) - \boldsymbol{g}^{\alpha,\beta}(0). \quad (43)$$

Similarly, the truncation with $\alpha = 0$, $\beta = 1$ ($g^{0,1}$) is not better than the truncation $\boldsymbol{g}^{\alpha,\beta}$:

$$\boldsymbol{g}^{\alpha,\beta}(z) - \boldsymbol{g}^{\alpha,\beta}(0) \le \boldsymbol{g}^{0,1}(z) - \boldsymbol{g}^{0,1}(0). \quad (44)$$

This is verified by examining components:

$$g^{0,1}_{st}(z_{st}) - g^{0,1}_{st}(0) - g^{\alpha,\beta}_{st}(z_{st}) + g^{\alpha,\beta}_{st}(0) = \\ = \begin{cases} 0, & z_{st} = 00, \\ \Delta_{st}(1-(\alpha+\beta)), & z_{st} = 01, \\ \Delta_{st}(1-(1-\alpha)), & z_{st} = 10, \\ \Delta_{st}(1-\beta), & z_{st} = 11, \end{cases} \quad (45) \\ \ge \ 0.$$

If z is an improving move for $\boldsymbol{g}^{0,1}$ then it is also an improving move for any truncation. □

We have the following result about the Algorithm 2.

Statement 9. *Let* $(x^{\min}, L)$ *be a strong autarky for* $\boldsymbol{f}$ *obtained by Algorithm* 2*. Let* x *be a fixed point of the expansion-move algorithm with any truncation rule. Then*

$$x \ge x^{\min}. \quad (46)$$

Proof. We will prove that the statement holds for truncation $(\alpha = 0, \beta = 1)$. We need to show that for a move from x to $x \vee x^{\min}$ the truncated energy decreases at least as much as does auxiliary problem built by Algorithm 2. This can be verified by inspecting pairwise components for the 4 cases $z_{st} = 00, 01, 10, 11$. □

5. Conclusion

We propose a novel representation of methods [10, 11] as deriving domain constraints via LP-autarkies. This allows for comparison with other methods deriving domain constraints in the same form [3, 15] and establishing the relations with the common methods of (approximate) optimization. We also believe that «label domination» condition proposed by [5] can be interpreted in the same framework, allowing for the theoretical comparison and or for the design of combined methods.

Our results open several directions for improvements. A direct improvement to the Algorithm 2 can be obtained as follows. The Algorithm 2 constructs a multi-label auxiliary problem, which is equivalent to a two-label problem (since we know that there is a minimizer with $x^*_s \in \{0, L\}$, $\forall s \in \mathcal{V}$). For two label problems, we also know that the autarky constructed by roof-dual dominates the autarky by truncation, so it will be better to set

$$d_{st} = \min_{i \ne L, j \ne L} [f_{st}(i,j) \\ + \min\{b_{st} - f_{st}(L,j), c_{st} - f_{st}(i,L)\}] \quad (47)$$

and solve for roof-dual using reduction to max-flow [2]. This would be a non-submodular auxiliary problem.

We can also attempt to construct auxiliary problem with mixed submodular and supermodular terms as in [15] or design an algorithm which will propose an autarky in some greedy way and then verify it via solving linear program (24).

1. *Boros E.*, *Hammer P.L.* Pseudo-boolean optimization. // Discrete Applied Mathematics, 2002. – 123(1-3). – P. 155–225.
2. *Boros E.*, *Hammer P.L.*, *Sun X.* Network flows and minimization of quadratic pseudo-Boolean functions // Technical Report RRR 17-1991, RUTCOR, May 1991.
3. *Boros E.*, *Hammer P.L.*, *Tavares G.* Preprocessing of unconstrained quadratic binary optimization // Ibid. 10-2006, RUTCOR, Apr. 2006.
4. *Boykov Y.*, *Veksler O.*, *Zabih R.* Fast approximate energy minimization via graph cuts // IEEE Transactions on Pattern Analysis and Machine Intelligence. – Nov. 2001. – **23**(11). – P. 1222–1239.
5. *Desmet J.*, *Maeyer M.D.*, *Hazes B.* et al. The dead-end elimination theorem and its use in protein side-chain positioning. Nature, 356:539–542, 1992.
6. *Hammer P.L.*, *Hansen P.*, *Simeone B.* Roof duality, complementation and persistency in quadratic 0-1 optimization. Math. Programming. – 1984. – P. 121–155.
7. *Hiroshi Ishikawa*. Exact optimization for Markov random fields with convex priors // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2003. – N 25(10). – P. 1333–1336.
8. *Kolmogorov V.* Convergent tree-reweighted message passing for energy minimization / Cowell R.G., Ghahramani Z., eds. // AI and Statistics, Society for Artificial Intelligence and Statistics, 2005. – P. 182–189.

End on the page 93

9. *Kolmogorov V.*, *Wainwright M.* On the optimality of a tree-reweighted max-product message passing // To appear in 21st Conf. on Uncertainty in Artificial Intelligence (UAI), July 2005.
10. *Kovtun I.* Partial optimal labeling search for a NP-hard subclass of (max, +) problems // DAGM-Symposium. – 2003. – P. 402–409.
11. *Kovtun I.* Image segmentation based on sufficient conditions of optimality in NP-complete classes of structural labelling problem. PhD thesis // IRTC ITS National Academy of Sciences, Ukraine, 2004. In Ukrainian.
12. *Digital* tapestry. In CVPR'05: Proceedings of the 2005 / C. Rother, S. Kumar, V. Kolmogorov et al. // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). – USA: Washington, DC, 2005. – **1**. – P. 589–596.
13. *Schlesinger D.*, *Flach B.* Transforming an arbitrary minsum problem into a binary one. Research Report TUD-FI06-01, Dresden University of Technology, April 2006.
14. *Schlesinger M.I.* A syntactic analysis of two-dimensional visual signals in the presence of noise // Cybernetics and Systems Analysis. – 1976. – N 12. – P. 612–628.
15. *LP-relaxation* of binarized energy minimization / A. Shekhovtsov, V. Kolmogorov, P. Kohli et al. // Research Report CTU–CMP–2007–27, Czech Technical University, 2008.
16. *Wainwright M.*, *Jaakkola T.*, *Willsky A.* Exact MAP estimates by (hyper)tree agreement / S. Thrun, S. Becker, K. Obermayer, ed. // Advances in Neural Information Processing Systems 15, MIT Press, 2003. – P. 809–816.
17. *Werner T.* A linear programming approach to max-sum problem: A review. Research Report CTU–CMP–2005–25, Center for Machine Perception, Czech Technical University, Dec. 2005.
18. *Werner T.* A linear programming approach to max-sum problem: A review // IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(7):1165–1179, July 2007.

●