Supplementary Material
A Primal-Dual Solver for Large-Scale Tracking-by-Assignment

A.1 Project website

Our project website at https://vislearn.github.io/libct contains additional information. At the time writing there we distribute: (*i*) The source code of our cell-tracking solver, (*ii*) information about how to obtain the datasets, and (*iii*) the model parameters that we have used.

A.2 Tracking-by-Assignment formulation and cost computation

A description of the mathematical model of the tracking-by-assignment formulation was already given in Section 2. Even though the reasoning in the paper has no restrictions on the costs, the cost assignment is a crucial step when using the method in practice. In the following we describe the cost computation that we have used for preparation of this paper, especially for obtaining the results in our experimental evaluation in Section 6.

The cost θ_u associated with each segmentation variable $u \in \hat{\mathcal{V}}_{\mathsf{det}}$ is based on image and object features of the underlying segmentation hypothesis. All segmentation hypotheses are assigned negative costs in order to promote selection as part of a tracking solution, i. e., a segmentation hypothesis with higher negative cost is more likely to be picked as part of a solution. Similar to Jug et al. (2014a,b) and Kaiser et al. (2018), the cost θ_u of any segment hypothesis is chosen according to its area and convexity according to the following rule

$$\theta_u = -\alpha_{\mathsf{det}} \cdot a(u) + \beta_{\mathsf{det}} \cdot \big(|a_{\mathcal{C}}(u) - a(u)|\big) + \gamma_{\mathsf{det}} \cdot \max\big(0, \big(a(u) - A\big)\big)^2 , \tag{14}$$

where α_{det}, β_{det} and γ_{det} are free coefficients, $a(u)$ is the area of the hypothesis u, $a_{\mathcal{C}}(u)$ is the area of convex hull of that hypothesis u, and A is a free parameter that denotes the upper limit of the range of reasonable object (segment) sizes.

The costs for all transitions between time steps (moves and divisions) are set up to reflect the knowledge of biological experts. For any $u \xrightarrow{w} v \in \hat{\mathcal{E}}_{\mathsf{move}}$ the associated cost $\theta_{u \xrightarrow{w} v}$ is given by a function which takes segment size and displacement (of segment centre of mass) between consecutive time points into account. The cost for a move variable can be written as

$$\theta_{u \xrightarrow{w} v} = \alpha_{\mathsf{move}} \cdot \Delta a(u, v) + \beta_{\mathsf{move}} \cdot \Delta p(u, v) , \tag{15}$$

where α_{move} and β_{move} are free coefficients, Δa and Δp represent the change in area and in squared position between two consecutive time points, respectively.

Let $u \overset{w}{\rightrightarrows} v/v' \in \hat{\mathcal{E}}_{\mathsf{div}}$. The cost $\theta_{u \overset{w}{\rightrightarrows} v/v'}$ for division variable additionally accounts for the fact that a dividing cell typically splits into two equally sized daughter cells, and that the cumulative volume of the daughter cells roughly equals the volume of the mother cell. The division variable cost is given by

$$\begin{aligned} \theta_{u \overset{w}{\rightrightarrows} v/v'} = \alpha_{\mathsf{div}} + \beta_{\mathsf{div}} \cdot \Delta a_{mds}(u, v, v') + \gamma_{\mathsf{div}} \cdot \Delta a(v, v') + \kappa_{\mathsf{div}} \cdot \Delta a(v, v')^2 + \\ + 0.5 \cdot \rho_{\mathsf{div}} \cdot (\Delta p(u, v)^2 + \Delta p(u, v')^2) + \sigma_{\mathsf{div}} \cdot \Delta p(v, v')^2 + \tau_{\mathsf{div}} \cdot \Delta r(u, v, v') , \end{aligned} \tag{16}$$

where α_{div}, β_{div}, γ_{div}, κ_{div}, ρ_{div}, σ_{div} and τ_{div} are free coefficients, $\Delta a_{mds}(u, v, v') := |a(u) - a(v) - a(v')|$ is the change of area between mother and daughter cells, and $\Delta r(u, v, v')$ is the difference in angular orientation between mother cell and daughter cells. Overall, the transition costs discourage the deviation from the above mentioned biological rules for any decision variable. Transition costs are positive and in order to collect the reward (negative costs) for a segmentation hypothesis, a solution needs to pay the price for explaining the past and future of this segment.

Additionally, it is possible for a cell to appear/disappear along the image border (cells moving in/out of the field of view) but costs of appearance and disappearance are set to be higher for cells further away from the image boundary. For sake of simplicity our description in Section 2 does not include decision variables for appearance or disappearance events. However, our formulation allows to deactivate all incoming (outgoing) transition variables for a segment to model cell appearance (disappearance), see (3). The costs for appearance and disappearance described below can be incorporated by simply shifting the costs of incoming and outgoing transition variables and the affected segmentation variable by a constant factor. The cost $\theta_{\mathsf{app}}(u)$ and $\theta_{\mathsf{dis}}(u)$ for an appearance or disappearance of segmentation hypothesis $u \in \hat{\mathcal{V}}_{\mathsf{det}}$ are given by

$$\theta_{\mathsf{app}}(u) = \alpha_{\mathsf{app}} \cdot a(u) + \beta_{\mathsf{app}} \cdot \sqrt{d_b(u)} + \gamma_{\mathsf{app}} \cdot d_b(u) \ , \tag{17}$$

$$\theta_{\mathsf{dis}}(u) = \alpha_{\mathsf{dis}} \cdot a(u) + \beta_{\mathsf{dis}} \cdot \sqrt{d_b(u)} + \gamma_{\mathsf{dis}} \cdot d_b(u) \ , \tag{18}$$

where α_{app}, α_{dis}, β_{app}, β_{dis}, γ_{app} and γ_{dis} are free coefficients and $d_b(u)$ represents the distance of the centre of mass of hypothesis u to the closest image boundary.

All free coefficients and the parameter A are set to sensible values by the engineer of the proposed system. The values we have used for all reported results are available online at our project website.

A.3 Source code of our cell-tracking solver

We implemented the suggested solving scheme in a modern C++ library. The source code of this implementation is publicly available and we plan to incorporate further improvements in the future. To make the results presented in this paper reproducible, the repository holding the source code also contains a fixed version which we used during the preparation of this paper.

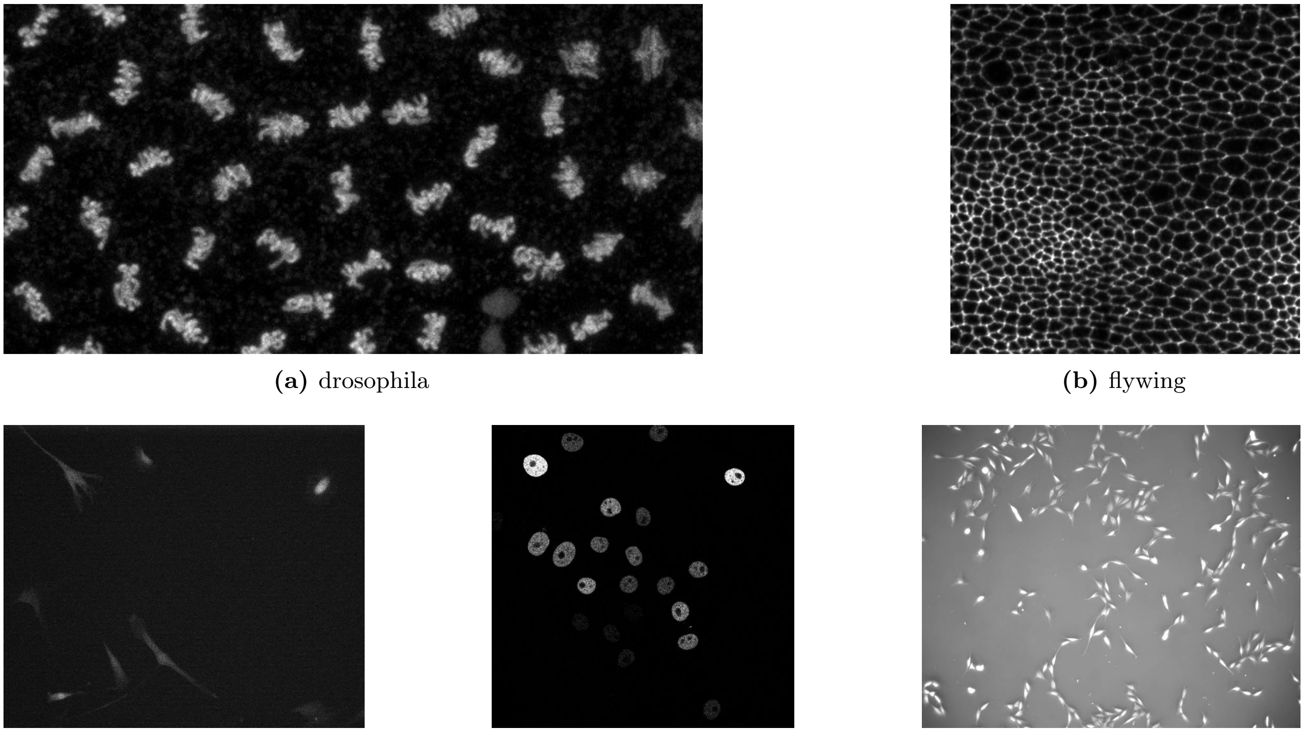
Along with the library we provide Python 3 bindings which allow to feed a text file representation of cell-tracking problems into the native library to run the solver.

For further information about the implementation and the text formats, please refer to the `README` file that is bundled with the source code.

Source code repository: https://github.com/vislearn/libct

A.4 Detailed information about the datasets

A description of all used datasets can be found in Section 6. Instructions how to obtain the datasets can be found on our project website. There we also distribute the resulting optimization problems for each cell-tracking instance in a text format and provide all model parameters.



(a) drosophila **(b)** flywing

(c) Fluo-C2DL-MSC **(d)** Fluo-N2DH-GOWT1 **(e)** PhC-C2DL-PSC

Figure 3: Example images of datasets that have been used for the evaluation.

instance	#timesteps	$\#\frac{\text{detections}}{\text{time step}}$	$\#\frac{\text{conflicts}}{\text{time step}}$	transitive conflict clique
drosophila	252	323.3 ± 62.9	161.9 ± 62.9	2.0 ± 0.3
flywing-100-1	100	2041.1 ± 358.0	2138.8 ± 358.0	753.2 ± 997.2
flywing-100-2	100	2223.4 ± 258.2	1831.6 ± 258.2	131.4 ± 511.7
flywing-245	245	3317.2 ± 326.6	2733.5 ± 326.6	54.7 ± 373.9
Fluo-C2DL-MSC-1	48	115.1 ± 6.9	41.4 ± 6.9	12.3 ± 8.3
Fluo-C2DL-MSC-2	48	52.2 ± 4.9	18.8 ± 4.9	9.8 ± 8.8
Fluo-N2DH-GOWT1-1	92	168.4 ± 1.6	24.9 ± 1.6	7.3 ± 1.6
Fluo-N2DH-GOWT1-2	92	207.1 ± 4.9	36.8 ± 4.9	7.5 ± 2.0
PhC-C2DL-PSC-1	426	1551.4 ± 482.7	576.7 ± 482.7	3.4 ± 1.4
PhC-C2DL-PSC-2	426	1249.8 ± 372.6	455.8 ± 372.6	3.5 ± 1.4

Table 2: Characteristics of all used tracking problem instances.

A.5 Detailed convergence plots

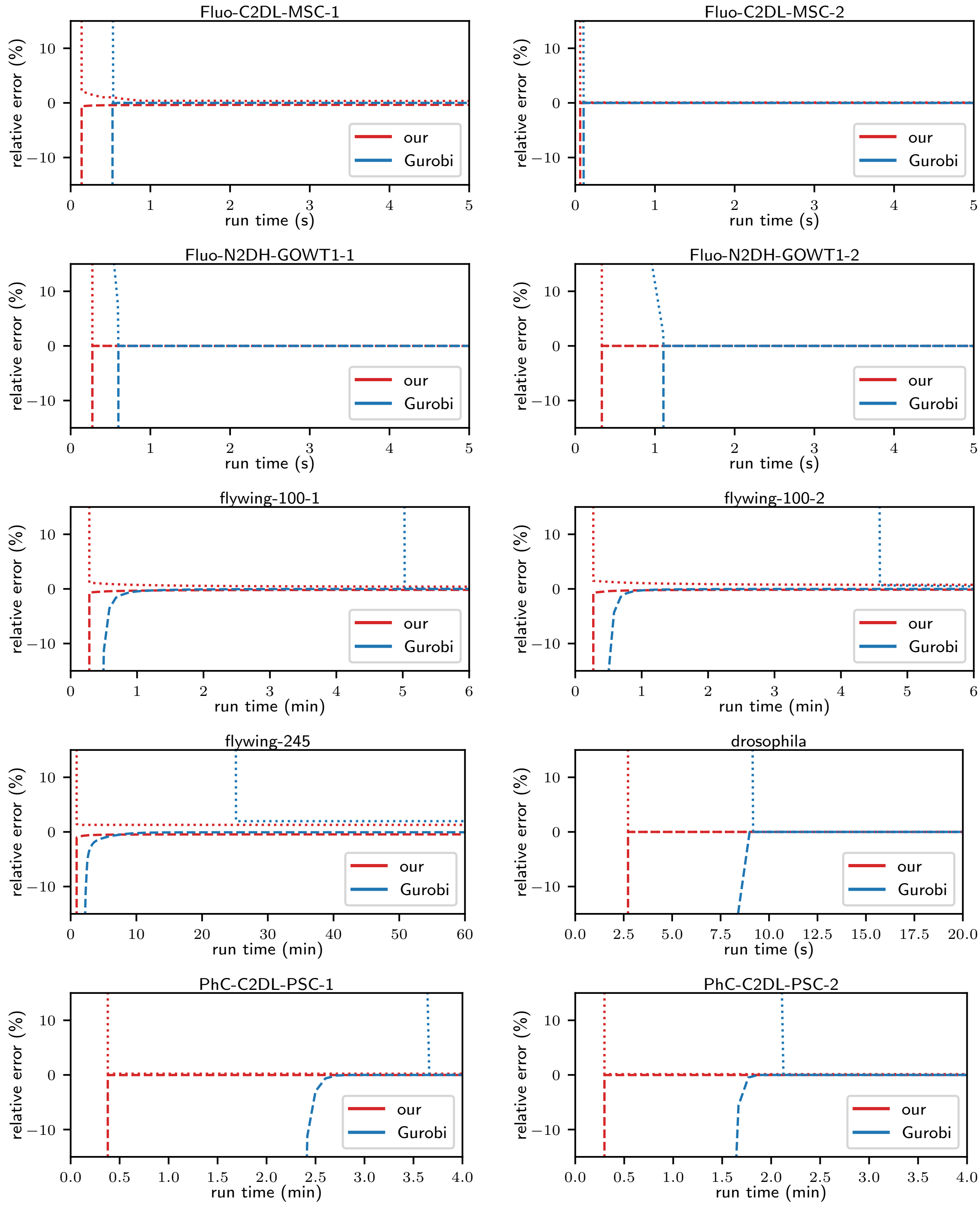


Figure 4: Comparison of lower-bound (dashed ==) and upper-bound (dotted ::::) convergence for our solver and Gurobi. We obtain high-quality solutions after only a few iterations. For more information see section 6.

A.6 Proofs of mathematical statements

Lemma 1. *The optimization objective* $E(\theta,x)=\sum_{v\in\mathcal{V}_{\mathsf{det}}}\langle\theta_v,x_v\rangle+\sum_{c\in\mathcal{V}_{\mathsf{conf}}}\langle\theta_c,x_c\rangle$*, cf.* (10)*, is equivalent to*

$$\begin{aligned}E(\theta,x)=&\sum_{v\in\mathcal{V}_{\mathsf{det}}}\theta_{v,\mathsf{det}}\,x_{v,\mathsf{det}}+\sum_{\substack{e=u\rightarrow v\\ \in\mathcal{E}_{\mathsf{move}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)+\sum_{\substack{e=u\rightarrow v\\ \in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\\ &+\sum_{\substack{e=u\rightrightarrows v/w\\ \in\mathcal{E}_{\mathsf{div}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)+\sum_{\substack{e=u\rightrightarrows v/w\\ \in\mathcal{E}_{\mathsf{div}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{\substack{e=u\rightrightarrows v/w\\ \in\mathcal{E}_{\mathsf{div}}}}\theta_{w,\mathsf{in}}(e)\,x_{w,\mathsf{in}}(e)+\sum_{\substack{e=v\lightning c\\ \in\mathcal{E}_{\mathsf{conf}}}}\theta_c(v)\,x_c(v)\,.\end{aligned}\tag{19}$$

Proof. First, we apply the definition of $\mathcal{X}_v$ for all $v\in\mathcal{V}_{\mathsf{det}}$ as well as the definition of $\mathcal{X}_c$ for all $c\in\mathcal{V}_{\mathsf{conf}}$. Next, we write the inner products in an explicit form.

$$\begin{aligned}\sum_{v\in\mathcal{V}_{\mathsf{det}}}\langle\theta_v,x_v\rangle&=\sum_{v\in\mathcal{V}_{\mathsf{det}}}\Big(\langle\theta_{v,\mathsf{det}},x_{v,\mathsf{det}}\rangle+\langle\theta_{v,\mathsf{in}},x_{v,\mathsf{in}}\rangle+\langle\theta_{v,\mathsf{out}},x_{v,\mathsf{out}}\rangle\Big)\\ &=\sum_{v\in\mathcal{V}_{\mathsf{det}}}\theta_{v,\mathsf{det}}\,x_{v,\mathsf{det}}+\sum_{v\in\mathcal{V}_{\mathsf{det}}}\sum_{e\in\mathsf{in}(v)}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{u\in\mathcal{V}_{\mathsf{det}}}\sum_{e\in\mathsf{out}(u)}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)\end{aligned}\tag{20}$$

$$\sum_{c\in\mathcal{V}_{\mathsf{conf}}}\langle\theta_c,x_c\rangle=\sum_{c\in\mathcal{V}_{\mathsf{conf}}}\sum_{\substack{v\in\mathcal{V}_{\mathsf{det}}:\\ v\lightning c\in\mathcal{E}_{\mathsf{conf}}}}\theta_c(v)\,x_c(v)=\sum_{v\lightning c\in\mathcal{E}_{\mathsf{conf}}}\theta_c(v)\,x_c(v)\tag{21}$$

We can now use the definition of $\mathsf{in}(\cdot)$ and $\mathsf{out}(\cdot)$ to expand the corresponding sums in (20).

$$\begin{aligned}\sum_{v\in\mathcal{V}_{\mathsf{det}}}\sum_{e\in\mathsf{in}(v)}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)&=\sum_{v\in\mathcal{V}_{\mathsf{det}}}\sum_{\substack{u\in\mathcal{V}_{\mathsf{det}}:\\ e=\\ u\rightarrow v\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{v\in\mathcal{V}_{\mathsf{det}}}\sum_{\substack{u,w\in\mathcal{V}_{\mathsf{det}}:\\ e=\\ u\rightrightarrows v/w\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{v\in\mathcal{V}_{\mathsf{det}}}\sum_{\substack{u,w\in\mathcal{V}_{\mathsf{det}}:\\ e=\\ u\rightrightarrows w/v\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)\\ &=\sum_{\substack{e=\\ u\rightarrow v\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{\substack{e=\\ u\rightrightarrows v/w\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)+\sum_{\substack{e=\\ u\rightrightarrows w/v\in\mathcal{E}_{\mathsf{move}}}}\theta_{v,\mathsf{in}}(e)\,x_{v,\mathsf{in}}(e)\end{aligned}\tag{22}$$

$$\begin{aligned}\sum_{u\in\mathcal{V}_{\mathsf{det}}}\sum_{e\in\mathsf{out}(u)}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)&=\sum_{u\in\mathcal{V}_{\mathsf{det}}}\sum_{\substack{v\in\mathcal{V}_{\mathsf{det}}:\\ e=\\ u\rightarrow v\in\mathcal{E}_{\mathsf{move}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)+\sum_{u\in\mathcal{V}_{\mathsf{det}}}\sum_{\substack{v,w\in\mathcal{V}_{\mathsf{det}}:\\ e=\\ u\rightrightarrows v/w\in\mathcal{E}_{\mathsf{move}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)\\ &=\sum_{\substack{e=\\ u\rightarrow v\in\mathcal{E}_{\mathsf{move}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)+\sum_{\substack{e=\\ u\rightrightarrows v/w\in\mathcal{E}_{\mathsf{move}}}}\theta_{u,\mathsf{out}}(e)\,x_{u,\mathsf{out}}(e)\end{aligned}\tag{23}$$

Substituting the terms in (10) by (20), (21), (22) and (23) results in equation (19). □

Corollary 1. *For any* $x\in\mathcal{X}$ *the optimization objective* $E(\theta,x)$ *is equivalent to*

$$\begin{aligned}E(\theta,x)=&\sum_{v\in\mathcal{V}_{\mathsf{det}}}\Big(\theta_{v,\mathsf{det}}(e)+\sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ v\lightning c\in\mathcal{E}_{\mathsf{conf}}}}\theta_c(v)\Big)\,x_{v,\mathsf{det}}(e)+\sum_{\substack{e=u\rightarrow v\\ \in\mathcal{E}_{\mathsf{move}}}}\Big(\theta_{u,\mathsf{out}}(e)+\theta_{v,\mathsf{in}}(e)\Big)\,x_{u,\mathsf{out}}(e)+\\ &+\sum_{\substack{e=u\rightrightarrows v/w\\ \in\mathcal{E}_{\mathsf{div}}}}\Big(\theta_{u,\mathsf{out}}(e)+\theta_{v,\mathsf{in}}(e)+\theta_{w,\mathsf{in}}(e)\Big)\,x_{u,\mathsf{out}}(e)\end{aligned}\tag{24}$$

Proof. Due to $x\in\mathcal{X}$ we know that the coupling constraints (9) hold. This means that for a given edge $e=u\rightarrow v\in\mathcal{E}_{\mathsf{move}}$ it holds that $x_{u,\mathsf{out}}(e)=x_{v,\mathsf{in}}(e)$ and similarly for divisions and conflict edges. We can now regroup the expression (19) of Lemma 1 and sort all terms by elements of vector x to directly obtain (24). □

Proposition 1. $\forall x \in \mathcal{X},\ \lambda \in \Lambda$: $E(\theta, x) = E(\theta^\lambda, x)$.

Proof. Due to $x \in \mathcal{X}$ we can apply Coralarry 1 and hence know that $E(\theta, x)$ is equivalent to (24). Corolarry 1 also holds for $E(\theta^\lambda, x)$ and we obtain

$$E(\theta^\lambda, x) = \sum_{v \in \mathcal{V}_{\mathsf{det}}} \Big(\theta^\lambda_{v,\mathsf{det}}(e) + \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}:\\ v \nleftrightarrow c \in \mathcal{E}_{\mathsf{conf}}}} \theta^\lambda_c(v)\Big)\, x_{v,\mathsf{det}}(e) + \sum_{\substack{e=u \to v\\ \in \mathcal{E}_{\mathsf{move}}}} \Big(\theta^\lambda_{u,\mathsf{out}}(e) + \theta^\lambda_{v,\mathsf{in}}(e)\Big)\, x_{u,\mathsf{out}}(e) + \\ + \sum_{\substack{e=u \rightrightarrows v/w\\ \in \mathcal{E}_{\mathsf{div}}}} \Big(\theta^\lambda_{u,\mathsf{out}}(e) + \theta^\lambda_{v,\mathsf{in}}(e) + \theta^\lambda_{w,\mathsf{in}}(e)\Big)\, x_{u,\mathsf{out}}(e)\,. \tag{25}$$

By definition of the reparametrized costs θ^λ we can simplify each of the following terms into

$$\forall\, v \in \mathcal{V}_{\mathsf{det}}: \quad \theta^\lambda_{v,\mathsf{det}}(e) + \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}:\\ v \nleftrightarrow c \in \mathcal{E}_{\mathsf{conf}}}} \theta^\lambda_c(v) = \theta_{v,\mathsf{det}}(e) - \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}:\\ v \nleftrightarrow c \in \mathcal{E}_{\mathsf{conf}}}} \lambda(v \nleftrightarrow c) + \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}:\\ v \nleftrightarrow c \in \mathcal{E}_{\mathsf{conf}}}} \Big(\theta^\lambda_c(v) + \lambda(v \nleftrightarrow c)\Big) \\ = \theta_{v,\mathsf{det}}(e) + \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}:\\ v \nleftrightarrow c \in \mathcal{E}_{\mathsf{conf}}}} \theta_c(v) \tag{26}$$

$$\forall\, e = u \to v \in \mathcal{E}_{\mathsf{move}}: \quad \theta^\lambda_{u,\mathsf{out}}(e) + \theta^\lambda_{v,\mathsf{in}}(e) = \theta_{u,\mathsf{out}}(e) - \lambda(e) + \theta_{v,\mathsf{in}}(e) + \lambda(e) \\ = \theta_{u,\mathsf{out}}(e) + \theta_{v,\mathsf{in}}(e) \tag{27}$$

$$\forall\, e = u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}: \quad \theta^\lambda_{u,\mathsf{out}}(e) + \theta^\lambda_{v,\mathsf{in}}(e) + \theta^\lambda_{w,\mathsf{in}}(e) = \theta_{u,\mathsf{out}}(e) - \lambda_v(e) - \lambda_w(e) + \theta_{v,\mathsf{in}}(e) + \lambda_v(e) + \theta_{w,\mathsf{in}}(e) + \lambda_w(e) \\ = \theta_{u,\mathsf{out}}(e) + \theta_{v,\mathsf{in}}(e) + \theta_{w,\mathsf{in}}(e) \tag{28}$$

Note that all tuples/triples of λ have been cancelling out each other. We can now insert (26), (27) and (28) into (25) and obtain the same expression as the right-hand side of (24). Due to Coralarry 1 we now that the very same expression is equivalent to $E(\theta, x)$, hence $E(\theta^\lambda, x) = E(\theta, x)$. $\square$

Proposition 2. *Dualizing all coupling constraints* (9) *in the objective* (10) *yields the* Lagrange dual problem $\max_{\lambda \in \Lambda} D(\lambda)$*, where*

$$D(\lambda) := \sum_{u \in \mathcal{V}_{\mathsf{det}}} \min_{x_u \in \mathcal{X}_u} \langle \theta_u^\lambda, x_u \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \min_{x_c \in \mathcal{X}_c} \langle \theta_c^\lambda, x_c \rangle \,. \tag{11}$$

Proof. To recap, the primal optimization problem is defined as the following, cf. (9) and (10):

$$\min_{x \in \{0,1\}^n} \Big[E(\theta, x) = \sum_{u \in \mathcal{V}_{\mathsf{det}}} \langle \theta_u, x_u \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \langle \theta_c, x_c \rangle \Big] \quad \text{s.t.} \quad \begin{cases} x_{u,\mathsf{det}} = x_c(u) & \forall\, u \lightning c \in \mathcal{E}_{\mathsf{conf}} \\ x_{u,\mathsf{out}}(u \to v) = x_{v,\mathsf{in}}(u \to v) & \forall\, u \to v \in \mathcal{E}_{\mathsf{move}} \\ x_{u,\mathsf{out}}(u \rightrightarrows v/w) = x_{v,\mathsf{in}}(u \rightrightarrows v/w) & \forall\, u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}} \\ x_{u,\mathsf{out}}(u \rightrightarrows v/w) = x_{w,\mathsf{in}}(u \rightrightarrows v/w) & \forall\, u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}} \end{cases} \tag{29}$$

We are now dualizing all the constraints of (29) by introducing a Lagrangean multipler for each equality constraint in (29). In total we have $|\mathcal{E}_{\mathsf{conf}}| + |\mathcal{E}_{\mathsf{move}}| + 2\,|\mathcal{E}_{\mathsf{div}}|$ constraints, so to assign a Lagrangean multiplier to each constraint we will write $\lambda \in \Lambda = \mathbb{R}^{|\mathcal{E}_{\mathsf{conf}}| + |\mathcal{E}_{\mathsf{move}}| + 2\,|\mathcal{E}_{\mathsf{div}}|}$, see the definition in the main paper. The Lagrange dual function augmented by the Lagrange multipliers now reads

$$\begin{aligned} D(\lambda) = \min_{x \in \mathcal{X}} \Big[& E(\theta, x) + \sum_{u \lightning c \in \mathcal{E}_{\mathsf{conf}}} \big(x_c(u) - x_{u,\mathsf{det}}\big)\, \lambda(u \lightning c) + \sum_{\substack{e= \\ u \to v \in \mathcal{E}_{\mathsf{move}}}} \big(x_{v,\mathsf{in}}(e) - x_{u,\mathsf{out}}(e)\big)\, \lambda(e) + \\ & + \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} \big(x_{v,\mathsf{in}}(e) - x_{u,\mathsf{out}}(e)\big)\, \lambda_v(e) + \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} \big(x_{w,\mathsf{in}}(e) - x_{u,\mathsf{out}}(e)\big)\, \lambda_w(e) \Big] \,, \end{aligned}$$

$$\begin{aligned} D(\lambda) = \min_{x \in \mathcal{X}} \Big[& E(\theta, x) + \sum_{u \lightning c \in \mathcal{E}_{\mathsf{conf}}} x_c(u)\, \lambda(u \lightning c) + \sum_{\substack{e= \\ u \to v \in \mathcal{E}_{\mathsf{move}}}} x_{v,\mathsf{in}}(e)\, \lambda(e) + \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} x_{v,\mathsf{in}}(e)\, \lambda_v(e) + \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} x_{w,\mathsf{in}}(e)\, \lambda_w(e) - \\ & - \sum_{\substack{e= \\ u \lightning c \in \mathcal{E}_{\mathsf{conf}}}} x_{u,\mathsf{det}}(e)\, \lambda(e) - \sum_{\substack{e= \\ u \to v \in \mathcal{E}_{\mathsf{move}}}} x_{u,\mathsf{out}}(e)\, \lambda(e) - \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} x_{u,\mathsf{out}}(e)\, \lambda_v(e) - \sum_{\substack{e= \\ u \rightrightarrows v/w \in \mathcal{E}_{\mathsf{div}}}} x_{u,\mathsf{out}}(e)\, \lambda_w(e) \Big] \,. \end{aligned} \tag{30}$$

We can now apply Lemma 1 to replace the term $E(\theta, x)$ by (19) in (30). After regrouping the terms and sorting them by elements of x we get

$$\begin{aligned} D(\lambda) = \min_{x \in \mathcal{X}} \Big[& \sum_{v \in \mathcal{V}_{\mathsf{det}}} \overbrace{\Big(\theta_{v,\mathsf{det}} - \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}: \\ c \lightning v \in \mathcal{E}_{\mathsf{conf}}}} \lambda(c \lightning v)\Big)}^{\theta^\lambda_{v,\mathsf{det}}} x_{v,\mathsf{det}} + \sum_{\substack{e = u \to v \\ \in \mathcal{E}_{\mathsf{move}}}} \overbrace{\Big(\theta_{u,\mathsf{out}}(e) - \lambda(e)\Big)}^{\theta^\lambda_{u,\mathsf{out}}(e)} x_{u,\mathsf{out}}(e) + \sum_{\substack{e = u \to v \\ \in \mathcal{E}_{\mathsf{move}}}} \overbrace{\Big(\theta_{v,\mathsf{in}}(e) + \lambda(e)\Big)}^{\theta^\lambda_{v,\mathsf{in}}(e)} x_{v,\mathsf{in}}(e) + \\ & + \sum_{\substack{e = u \rightrightarrows v/w \\ \in \mathcal{E}_{\mathsf{div}}}} \underbrace{\Big(\theta_{u,\mathsf{out}}(e) - \lambda_v(e) - \lambda_w(e)\Big)}_{\theta^\lambda_{u,\mathsf{out}}(e)} x_{u,\mathsf{out}}(e) + \sum_{\substack{e = u \rightrightarrows v/w \\ \in \mathcal{E}_{\mathsf{div}}}} \underbrace{\Big(\theta_{v,\mathsf{in}}(e) + \lambda_v(e)\Big)}_{\theta^\lambda_{v,\mathsf{in}}(e)} x_{v,\mathsf{in}}(e) + \\ & + \sum_{\substack{e = u \rightrightarrows v/w \\ \in \mathcal{E}_{\mathsf{div}}}} \underbrace{\Big(\theta_{w,\mathsf{in}}(e) + \lambda_w(e)\Big)}_{\theta^\lambda_{w,\mathsf{in}}(e)} x_{w,\mathsf{in}}(e) + \sum_{\substack{e = v \lightning c \\ \in \mathcal{E}_{\mathsf{conf}}}} \underbrace{\Big(\theta_c(v) + \lambda(e)\Big)}_{\theta^\lambda_c(e)} x_c(v) \Big] \,. \end{aligned} \tag{31}$$

Due to Lemma 1 we know that (31) is equivalent to $D(\lambda) = \min_{x \in \mathcal{X}} E(\theta^\lambda, x) = \min_{x \in \mathcal{X}} \big[\sum_{u \in \mathcal{V}_{\mathsf{det}}} \langle \theta_u^\lambda, x_u \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \langle \theta_c^\lambda, x_c \rangle\big]$ which is our unconstrained objective function for the Lagrange dual of (29).

As we want to maximize the dual function $D(\lambda)$ with respect to $\lambda \in \Lambda$ the Lagrange dual problem reads

$$\max_{\lambda \in \Lambda} \min_{x \in \mathcal{X}} \Big[\sum_{u \in \mathcal{V}_{\mathsf{det}}} \langle \theta_u^\lambda, x_u \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \langle \theta_c^\lambda, x_c \rangle \Big] = \max_{\lambda \in \Lambda} \Big[\sum_{u \in \mathcal{V}_{\mathsf{det}}} \min_{x_v \in \mathcal{X}_v} \langle \theta_u^\lambda, x_u \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \min_{x_c \in \mathcal{X}_c} \langle \theta_c^\lambda, x_c \rangle \Big] \,. \tag{32}$$

□

Proposition 3. *Dual updates* $\Delta \in \{\Delta_u^{\leftarrow}, \Delta_u^{\rightarrow}, \Delta_u^{\uparrow} \mid u \in \mathcal{V}_{\mathsf{det}}\} \cup \{\Delta_c^{\uparrow} \mid c \in \mathcal{V}_{\mathsf{conf}}\}$ *monotonically increase the dual function, i. e.* $\forall \lambda \in \Lambda \colon D(\lambda) \leq D(\lambda + \Delta)$.

Proof. For all possible choices of Δ we want to show

$$D(\lambda) \leq D(\lambda + \Delta),$$

for any fixed $\lambda \in \Lambda$, which is equivalent to $0 \leq D(\lambda + \Delta) - D(\lambda)$. Without loss of generality we can assume $\lambda = 0$, since any reparametrization is linear, and, therefore, $\theta^{\lambda+\Delta} = \left(\theta^{\lambda}\right)^{\Delta}$. So we can just redefine θ to match θ^{λ}. Thus, it suffices to prove

$$0 \leq D(\Delta) - D(0), \tag{33}$$

for all possible choices of Δ.

Case 1: Let $\Delta = \Delta_c^{\uparrow}$, $c \in \mathcal{V}_{\mathsf{conf}}$ arbitrary but fixed. Recall that for all $u \in c$, $e = u \not\leftrightarrow c$:

$$\Delta_c^{\uparrow}(e) := -\theta_c(u) + \frac{1}{2}\left[\langle \theta_c, z_c^{\star} \rangle + \langle \theta_c, z_c^{\star\star} \rangle\right], \text{ with } z_c^{\star} = \arg\min_{x \in \mathcal{X}_c} \langle \theta_c, x \rangle, \; z_c^{\star\star} = \arg\min_{x \in \mathcal{X}_c \setminus \{z_c^{\star}\}} \langle \theta_c, x \rangle .$$

For convenience, let $B_c := \frac{1}{2}\left[\langle \theta_c, z_c^{\star} \rangle + \langle \theta_c, z_c^{\star\star} \rangle\right]$. Note that $\langle \theta_c, z_c^{\star} \rangle \leq B_c \leq \langle \theta_c, x \rangle$ for all $x \in \mathcal{X}_c \setminus \{z_c^{\star}\}$ by definition of $z_c^{\star}$. We now rewrite the difference $D(\Delta_c^{\uparrow}) - D(0)$:

$$\begin{aligned}
D(\Delta_c^{\uparrow}) - D(0) &= \sum_{d \in \mathcal{V}_{\mathsf{conf}}} \min_{x \in \mathcal{X}_d} \langle \theta_d^{\Delta_c^{\uparrow}}, x \rangle + \sum_{v \in \mathcal{V}_{\mathsf{det}}} \min_{x \in \mathcal{X}_v} \langle \theta_v^{\Delta_c^{\uparrow}}, x \rangle - \left[\sum_{d \in \mathcal{V}_{\mathsf{conf}}} \min_{x \in \mathcal{X}_d} \langle \theta_d, x \rangle + \sum_{v \in \mathcal{V}_{\mathsf{det}}} \min_{x \in \mathcal{X}_v} \langle \theta_v, x \rangle \right] \\
&= \min_{x \in \mathcal{X}_c} \langle \theta_c^{\Delta_c^{\uparrow}}, x \rangle - \min_{x \in \mathcal{X}_c} \langle \theta_c, x \rangle + \sum_{u \in c} \min_{x \in \mathcal{X}_u} \langle \theta_u^{\Delta_c^{\uparrow}}, x \rangle - \sum_{u \in c} \min_{x \in \mathcal{X}_u} \langle \theta_u, x \rangle \\
&= \min_{x \in \mathcal{X}_c} \sum_{u \in c} \left[\theta_c(u) - \theta_c(u) + B_c\right] \cdot x(u) - \langle \theta_c, z_c^{\star} \rangle + \sum_{u \in c} \left[\min_{x \in \mathcal{X}_u} \langle \theta_u^{\Delta_c^{\uparrow}}, x \rangle - \min_{x \in \mathcal{X}_u} \langle \theta_u, x \rangle \right] \\
&= \min\{0, B_c\} - \langle \theta_c, z_c^{\star} \rangle + \sum_{u \in c} \left[\min_{x \in \mathcal{X}_u} \left(\langle \theta_u, x \rangle + [\theta_c(u) - B_c] \cdot x_{\mathsf{det}} \right) - \min_{x \in \mathcal{X}_u} \langle \theta_u, x \rangle \right]
\end{aligned}$$

If $z_c^{\star}(u) = 0$ for all $u \in c$, Equation (33) holds, as in this case $\theta_c(u) \geq B_c \geq 0$ for all $u \in c$. So we are left with the case that there exists $u^{\star} \in c$ such that $z_c^{\star}(u^{\star}) = 1$. Note that $u^{\star}$ is unique since $z_c^{\star} \in \mathcal{X}_c$, cf. (8). In particular, $z_c^{\star}(u) = 0$ for all $u \in c$, $u \neq u^{\star}$. Furthermore, it is $\langle \theta_c, z_c^{\star} \rangle = \theta_c(u^{\star}) \leq B_c \leq 0$. We now obtain:

$$\begin{aligned}
D(\Delta_c^{\uparrow}) - D(0) &= \min\{0, B_c\} - \langle \theta_c, z_c^{\star} \rangle + \sum_{\substack{u \in c: \\ z_c^{\star}(u) = 0}} \left[\min_{x \in \mathcal{X}_u} \left(\langle \theta_u, x \rangle + [\theta_c(u) - B_c] \cdot x_{\mathsf{det}} \right) - \min_{x \in \mathcal{X}_u} \langle \theta_u, x \rangle \right] \\
&\quad + \min_{x \in \mathcal{X}_{u^{\star}}} \left(\langle \theta_{u^{\star}}, x \rangle + [\theta_c(u^{\star}) - B_c] \cdot x_{\mathsf{det}} \right) - \min_{x \in \mathcal{X}_{u^{\star}}} \langle \theta_{u^{\star}}, x \rangle \\
&\geq B_c - \langle \theta_c, z_c^{\star} \rangle + \min_{x \in \mathcal{X}_{u^{\star}}} \langle \theta_{u^{\star}}, x \rangle + \langle \theta_c, z_c^{\star} \rangle - B_c - \min_{x \in \mathcal{X}_{u^{\star}}} \langle \theta_{u^{\star}}, x \rangle = 0
\end{aligned}$$

Hence, $D(\Delta_c^{\uparrow}) - D(0) \geq 0$.

Case 2: Let $\Delta = \Delta_u^{\uparrow}$, $u \in \mathcal{V}_{\mathsf{det}}$ arbitrary but fixed. Recall that for all $e \in \mathsf{conf}(u)$:

$$\Delta_u^{\uparrow}(e) := \min_{x \in \mathcal{X}_u \colon x_{\mathsf{det}} = 1} \frac{\langle \theta_u, x \rangle}{|\mathsf{conf}(u)|} = \frac{1}{|\mathsf{conf}(u)|} \min_{x \in \mathcal{X}_u \colon x_{\mathsf{det}} = 1} \langle \theta_u, x \rangle .$$

Now, rewriting the difference $D(\Delta_u^{\uparrow}) - D(0)$ yields:

$$\begin{aligned}
&D(\Delta_u^{\uparrow}) - D(0) \\
&= \sum_{v \in \mathcal{V}_{\mathsf{det}}} \min_{x \in \mathcal{X}_v} \langle \theta_v^{\Delta_u^{\uparrow}}, x \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \min_{x \in \mathcal{X}_c} \langle \theta_c^{\Delta_u^{\uparrow}}, x \rangle - \left[\sum_{v \in \mathcal{V}_{\mathsf{det}}} \min_{x \in \mathcal{X}_v} \langle \theta_v, x \rangle + \sum_{c \in \mathcal{V}_{\mathsf{conf}}} \min_{x \in \mathcal{X}_c} \langle \theta_c, x \rangle \right] \\
&= \min_{x \in \mathcal{X}_u} \langle \theta_u^{\Delta_u^{\uparrow}}, x \rangle - \min_{x \in \mathcal{X}_u} \langle \theta_u, x \rangle + \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}: \\ u \in c}} \min_{x \in \mathcal{X}_c} \langle \theta_c^{\Delta_u^{\uparrow}}, x \rangle - \sum_{\substack{c \in \mathcal{V}_{\mathsf{conf}}: \\ u \in c}} \min_{x \in \mathcal{X}_c} \langle \theta_c, x \rangle
\end{aligned}$$

proof continues on next page

$$
\begin{aligned}
&D(\Delta_u^{\uparrow}) - D(0)\\
&= \min_{x\in\mathcal{X}_u}\Big[\langle\theta_u,x\rangle - x_{\mathsf{det}}\cdot\sum_{e\in\mathsf{conf}(u)}\Delta_u^{\uparrow}(e)\Big] - \min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\Big[\min_{x\in\mathcal{X}_c}\big(\langle\theta_c,x\rangle + \Delta_u^{\uparrow}(u\not\sim c)\cdot x(u)\big) - \min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle\Big]\\
&= \min_{x\in\mathcal{X}_u}\Big[\langle\theta_u,x\rangle - x_{\mathsf{det}}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\Big] - \min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\Big[\min_{x\in\mathcal{X}_c}\big(\langle\theta_c,x\rangle + \tfrac{x(u)}{|\mathsf{conf}(u)|}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\big) - \min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle\Big]\\
&= \min\Big\{0,\min_{\substack{x\in\mathcal{X}_u:\\ x_{\mathsf{det}}=1}}\langle\theta_u,x\rangle - \min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\Big\} - \min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\Big[\min_{x\in\mathcal{X}_c}\big(\langle\theta_c,x\rangle + \tfrac{x(u)}{|\mathsf{conf}(u)|}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\big) - \min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle\Big]\\
&\geq 0 - \min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\Big[\min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle + \min_{x\in\mathcal{X}_c}\Big(\tfrac{x(u)}{|\mathsf{conf}(u)|}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\Big) - \min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle\Big]\\
&= -\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\min_{x\in\mathcal{X}_c}\Big(\tfrac{x(u)}{|\mathsf{conf}(u)|}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\Big) = -\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{c\in\mathcal{V}_{\mathsf{conf}}:\\ u\in c}}\min\Big\{0,\tfrac{1}{|\mathsf{conf}(u)|}\cdot\min_{\substack{y\in\mathcal{X}_u:\\ y_{\mathsf{det}}=1}}\langle\theta_u,y\rangle\Big\}\\
&= -\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \min\Big\{0,\min_{x\in\mathcal{X}_u,x_{\mathsf{det}}=1}\langle\theta_u,x\rangle\Big\} = -\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle = 0
\end{aligned}
$$

Hence, $D(\Delta_u^{\uparrow}) - D(0) \geq 0$.

Case 3: Let $\Delta = \Delta_u^{\rightarrow}$, $u\in\mathcal{V}_{\mathsf{det}}$ arbitrary but fixed. Recall that for all $e\in\mathsf{out}(u)$:

$$
\Delta_u^{\rightarrow}(e) := \min_{\substack{x\in\mathcal{X}_u:\\ x_{\mathsf{out}}(e)=1}}\langle\theta_u,x\rangle - \Theta_{u,\mathsf{out}},\ \text{if } e\in\mathcal{E}_{\mathsf{move}},\qquad (\Delta_u^{\rightarrow})_v(e) := \tfrac{1}{2}\Big[\min_{\substack{x\in\mathcal{X}_u:\\ x_{\mathsf{out}}(e)=1}}\langle\theta_u,x\rangle - \Theta_{u,\mathsf{out}}\Big],\ \text{if } e = u\rightrightarrows v/w
$$

where $\Theta_{u,\mathsf{out}} := \min\Big\{0,\tfrac{1}{2}\big[\langle\theta_u,x_u^\star\rangle + \langle\theta_u,(1,x_{u,\mathsf{in}}^\star,y_{u,\mathsf{out}}^\star)\rangle\big]\Big\}$, $x_u^\star := \operatorname*{arg\,min}_{x\in\mathcal{X}_u:\ x_{\mathsf{det}}=1}\langle\theta_u,x\rangle$, and $y_u^\star := \operatorname*{arg\,min}_{\substack{x\in\mathcal{X}_u:\ x_{\mathsf{det}}=1,\\ x_{\mathsf{in}}\neq x_{u,\mathsf{in}}^\star,\ x_{\mathsf{out}}\neq x_{u,\mathsf{out}}^\star}}\langle\theta_u,x\rangle$.

Using similar techniques as above we can rewrite the difference $D(\Delta_u^{\rightarrow}) - D(0)$ as follows:

$$
\begin{aligned}
&D(\Delta_u^{\rightarrow}) - D(0)\\
&= \sum_{v\in\mathcal{V}_{\mathsf{det}}}\min_{x\in\mathcal{X}_v}\langle\theta_v^{\Delta_u^{\rightarrow}},x\rangle + \sum_{c\in\mathcal{V}_{\mathsf{conf}}}\min_{x\in\mathcal{X}_c}\langle\theta_c^{\Delta_u^{\rightarrow}},x\rangle - \Big[\sum_{v\in\mathcal{V}_{\mathsf{det}}}\min_{x\in\mathcal{X}_v}\langle\theta_v,x\rangle + \sum_{c\in\mathcal{V}_{\mathsf{conf}}}\min_{x\in\mathcal{X}_c}\langle\theta_c,x\rangle\Big]\\
&= \min_{x\in\mathcal{X}_u}\langle\theta_u^{\Delta_u^{\rightarrow}},x\rangle + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\mathsf{move}}\cap\mathsf{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v^{\Delta_u^{\rightarrow}},x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\mathsf{div}}\cap\mathsf{out}(u)}}\Big[\min_{x\in\mathcal{X}_v}\langle\theta_v^{\Delta_u^{\rightarrow}},x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w^{\Delta_u^{\rightarrow}},x\rangle\Big]\\
&\quad - \Bigg(\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\mathsf{move}}\cap\mathsf{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v,x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\mathsf{div}}\cap\mathsf{out}(u)}}\Big[\min_{x\in\mathcal{X}_v}\langle\theta_v,x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w,x\rangle\Big]\Bigg)\\
&= \min_{x\in\mathcal{X}_u}\Big[\langle\theta_u,x\rangle - \sum_{e\in\mathsf{out}(u)\cap\mathcal{E}_{\mathsf{move}}}\Delta_u^{\rightarrow}(e)\cdot x_{\mathsf{out}}(e) - \sum_{\substack{e=u\rightrightarrows v/w\in\\ \mathsf{out}(u)\cap\mathcal{E}_{\mathsf{div}}}}\big[(\Delta_u^{\rightarrow})_v(e) + (\Delta_u^{\rightarrow})_w(e)\big]\cdot x_{\mathsf{out}}(e)\Big] + \sum_{\substack{e=u\to v\in\\ \mathcal{E}_{\mathsf{move}}\cap\mathsf{out}(u)}}\min_{x\in\mathcal{X}_v}\Big[\langle\theta_v,x\rangle + \Delta_u^{\rightarrow}(e)\cdot x_{\mathsf{in}}(e)\Big]\\
&\quad + \sum_{\substack{e=u\rightrightarrows v/w\in\\ \mathcal{E}_{\mathsf{div}}\cap\mathsf{out}(u)}}\Bigg[\min_{x\in\mathcal{X}_v}\Big(\langle\theta_v,x\rangle + (\Delta_u^{\rightarrow})_v(e)\cdot x_{\mathsf{in}}(e)\Big) + \min_{x\in\mathcal{X}_w}\Big(\langle\theta_w,x\rangle + (\Delta_u^{\rightarrow})_w(e)\cdot x_{\mathsf{in}}(e)\Big)\Bigg]\\
&\quad - \Bigg(\min_{x\in\mathcal{X}_u}\langle\theta_u,x\rangle + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\mathsf{move}}\cap\mathsf{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v,x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\mathsf{div}}\cap\mathsf{out}(u)}}\Big[\min_{x\in\mathcal{X}_v}\langle\theta_v,x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w,x\rangle\Big]\Bigg)
\end{aligned}
$$

proof continues on next page

For convenience, we set $B_{\text{out}} := \frac{1}{2}\left[\langle\theta_u, x_u^\star\rangle + \theta_u(1, x^\star_{u,\text{in}}, y^\star_{u,\text{out}})\right]$. Observe $B_{\text{out}} \geq \langle\theta_u, x_u^\star\rangle$. With this we get:

$$
\begin{aligned}
&D(\Delta_u^{\rightarrow}) - D(0) \\
&= \min_{x\in\mathcal{X}_u}\left[\langle\theta_u, x\rangle - \sum_{e\in\text{out}(u)}\left(\min_{\substack{y\in\mathcal{X}_u:\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right)\cdot x_{\text{out}}(e)\right] \\
&\quad + \sum_{\substack{e=u\to v\in\\ \mathcal{E}_{\text{move}}\cap\,\text{out}(u)}}\min_{x\in\mathcal{X}_v}\left[\langle\theta_v, x\rangle + \left(\min_{\substack{y\in\mathcal{X}_u:\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right)\cdot x_{\text{in}}(e)\right] \\
&\quad + \sum_{\substack{e=u\rightrightarrows v/w\in\\ \mathcal{E}_{\text{div}}\cap\,\text{out}(u)}}\left[\min_{x\in\mathcal{X}_v}\left(\langle\theta_v, x\rangle + \frac{1}{2}\left[\min_{\substack{y\in\mathcal{X}_u:\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right]\cdot x_{\text{in}}(e)\right)\right. \\
&\qquad\qquad \left. + \min_{x\in\mathcal{X}_w}\left(\langle\theta_w, x\rangle + \frac{1}{2}\left[\min_{\substack{y\in\mathcal{X}_u:\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right]\cdot x_{\text{in}}(e)\right)\right] \\
&\quad - \left(\min_{x\in\mathcal{X}_u}\langle\theta_u, x\rangle + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\text{move}}\cap\,\text{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\text{div}}\cap\,\text{out}(u)}}\left[\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w, x\rangle\right]\right) \\
&\geq \min\left\{0, \min_{e\in\text{out}(u)}\left[\min_{\substack{x\in\mathcal{X}_u:\\ x_{\text{out}}(e)=1}}\langle\theta_u, x\rangle - \min_{\substack{y\in\mathcal{X}_u:\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle + \min\{0, B_{\text{out}}\}\right]\right\} \\
&\quad + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\text{move}}\cap\,\text{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\text{div}}\cap\,\text{out}(u)}}\left[\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w, x\rangle\right] + \sum_{e\in\text{out}(u)}\min\left\{0, \min_{\substack{y\in\mathcal{X}_u,\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right\} \\
&\quad - \left(\min_{x\in\mathcal{X}_u}\langle\theta_u, x\rangle + \sum_{\substack{u\to v\in\\ \mathcal{E}_{\text{move}}\cap\,\text{out}(u)}}\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \sum_{\substack{u\rightrightarrows v/w\in\\ \mathcal{E}_{\text{div}}\cap\,\text{out}(u)}}\left[\min_{x\in\mathcal{X}_v}\langle\theta_v, x\rangle + \min_{x\in\mathcal{X}_w}\langle\theta_w, x\rangle\right]\right) \\
&= \min\{0, B_{\text{out}}\} + \sum_{e\in\text{out}(u)}\min\left\{0, \min_{\substack{y\in\mathcal{X}_u,\\ y_{\text{out}}(e)=1}}\langle\theta_u, y\rangle - \min\{0, B_{\text{out}}\}\right\} - \min_{x\in\mathcal{X}_u}\langle\theta_u, x\rangle \\
&= \min\{0, B_{\text{out}}\} + \min\left\{0, \langle\theta_u, x_u^\star\rangle - \min\{0, B_{\text{out}}\}\right\} - \min\{0, \langle\theta_u, x_u^\star\rangle\} = 0
\end{aligned}
$$

Hence, $D(\Delta_u^{\rightarrow}) - D(0) \geq 0$.

Case 4: Let $\Delta = \Delta_u^{\leftarrow}$, $u \in \mathcal{V}_{\text{det}}$ arbitrary but fixed. In this case the argument is completely analogous to 3. □

Proposition 4. *The maximization of the dual (11) yields the same value as the natural LP relaxation of (10), more precisely*

$$
\max_{\lambda\in\Lambda}\left[D(\lambda) = \sum_{v\in\mathcal{V}_{\text{det}}}\min_{x_v\in\mathcal{X}_v}\langle\theta_v^\lambda, x_v\rangle + \sum_{c\in\mathcal{V}_{\text{conf}}}\min_{x_c\in\mathcal{X}_c}\langle\theta_c^\lambda, x_c\rangle\right] = \min_{\substack{x\in[0,1]\\ st.\ (9)\ hold}}\left[E(\theta, x) = \sum_{v\in\mathcal{V}_{\text{det}}}\langle\theta_v, x_v\rangle + \sum_{c\in\mathcal{V}_{\text{conf}}}\langle\theta_c, x_c\rangle\right]. \tag{34}
$$

Proof. Instead of showing this result directly we will reference the corresponding general results in the literature as this property is not special to the Lagrange decomposition at hand. We refer to the excellent survey by Guignard (2003) that summarizes the Lagrange decomposition technique and gives a number of mathematical and applied insights. Generally, it is known that the Lagrange decomposition is always at at least as good as the LP relaxation, i. e. using "$\leq$" instead of "$=$" in (34). If the relaxed solutions for all subproblems of the Lagrange decomposition are integer (i. e. the LP relaxation of all subproblems is tight) then the Lagrange decomposition dual is not stronger than the LP relaxation, i. e. they have the same optimal value (Guignard, 2003, Corallary 5.1).

In our decomposition we have dualized all coupling constraints (9) which leads us to the dual function

$$
D(\lambda) = \sum_{v\in\mathcal{V}_{\text{det}}}\min_{x_v\in\mathcal{X}_v}\langle\theta_v^\lambda, x_v\rangle + \sum_{c\in\mathcal{V}_{\text{conf}}}\min_{x_c\in\mathcal{X}_c}\langle\theta_c^\lambda, x_c\rangle\,. \tag{11}
$$

All subproblems in our dual $D(\lambda)$ consists of minimizing simple inner products. Hence it is trivial to see that the the LP relaxation of all subproblems are tight. □