A. Appendix

Proof of Proposition 1. Cycle consistency amounts to the following property: Whenever there is a path $p_1, \ldots, p_k$ with $p_i \in [d]$ and nodes $s_1, \ldots, s_k$ with $s_i \in [m_{p_i}]$ such that $X^{[p_i p_{i+1}]}_{s_i s_{i+1}} = 1$ then it must hold that $X^{[p_1 p_k]}_{s_1 s_k} = 1$ as well. The constraints $X^{[pq]} X^{[qr]} \leq X^{[pr]}$ enforce the above constraints for paths of length three. By triangulation we can extend this equation to paths of arbitrary length. We use the path $p_1, \ldots, p_k$ with $p_i \in [d]$ as above. Then

$$
\begin{aligned}
& X^{[p_1 p_2]} \cdot X^{[p_2 p_3]} \cdot \ldots \cdot X^{[p_{k-1} p_k]} \\
\leq & X^{[p_1 p_3]} \cdot \ldots \cdot X^{[p_{k-1} p_k]} \\
& \ldots \\
\leq & X^{[p_1 p_{k-1}]} \cdot X^{[p_{k-1} p_k]} \\
\leq & X^{[p_1 p_k]} \, .
\end{aligned}
$$

□

Example 1 (A minimal non-cycle consistent problem)**.** *Consider the following multi-graph matching instance* (1) *with* $d = 3$ *and* $m_p = 2 \; \forall p \in [d]$. *Let*

$$W^{[12]} = W^{[13]} = W^{[23]} = \mathrm{diag}(-1, -10, -10, -1) \, . \tag{24}$$

Then without cycle consistency constraints the optimal assignment will be

$$X^{[12]} = X^{[13]} = X^{[23]} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{25}$$

with objective value -60. *After adding cycle-consistency constraints an optimal solution is*

$$X^{[12]} = X^{[13]} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad X^{[23]} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{26}$$

with objective value -42.

A.1. Cutting planes with cycle consistency subproblems

Algorithm 2 proceeds by first computing the dual lower bound of all subproblems that are connected to a given cycle consistency subproblem $x^{[pqr],st}$ (line 1). Then messages are sent to $x^{[pqr],st}$ from all factors that are connected to it (lines 2-9). Finally, the lower bound after these operations is computed (line 10) and the initial reparametrization is restored (lines 11-14).

Algorithm 2: Dual lower bound increase for $x^{[pqr],st}$

/* lower bound without cycle consistency subproblem $x^{[pqr],st}$ */
1 $lb^0 = \min_{x \in \overline{Y}^{[pr],s}} \langle \overline{\theta}^{[pq],s}, x \rangle + \min_{x \in Y^{[qr],t}} \langle \theta^{[qr],t}, x \rangle + \min_{x \in Y^{[pr],s}} \langle \theta^{[pr],s}, x \rangle + \min_{x \in \overline{Y}^{[pr],t}} \langle \overline{\theta}^{[pr],t}, x \rangle$;
/* send messages to cycle consistency subproblem $x^{[pqr],st}$ */
2 $\overline{\Delta}^{[pq],s} = \mathrm{msg}(\overline{x}^{[pq],s}, x^{[pqr],st})$;
3 $\Delta^{[pq],t} = \mathrm{msg}(x^{[qr],t}, x^{[pqr],st})$;
4 $\Delta^{[pr],t} = \mathrm{msg}(x^{[pr],t}, x^{[pqr],st})$;
5 $\overline{\Delta}^{[qr],s} = \mathrm{msg}(\overline{x}^{[qr],s}, x^{[pqr],st})$;
6 $\mathrm{repam}(\overline{\Delta}^{[pq],s}, \overline{x}^{[pq],s}, x^{[pqr],st})$;
7 $\mathrm{repam}(\Delta^{[pq],t}, \overline{x}^{[pq],t}, x^{[pqr],st})$;
8 $\mathrm{repam}(\Delta^{[pr],t}, \overline{x}^{[pr],t}, x^{[pqr],st})$;
9 $\mathrm{repam}(\overline{\Delta}^{[qr],s}, \overline{x}^{[qr],s}, x^{[pqr],st})$;
/* lower bound after adding cycle consistency subproblem $x^{[pqr],st}$ */
10 $lb^1 = \min_{x \in \overline{Y}^{[pr],s}} \langle \overline{\theta}^{[pq],s}, x \rangle + \min_{x \in Y^{[qr],t}} \langle \theta^{[qr],t}, x \rangle + \min_{x \in Y^{[pr],s}} \langle \theta^{[pr],s}, x \rangle + \min_{x \in \overline{Y}^{[pr],t}} \langle \overline{\theta}^{[pr],t}, x \rangle + \min_{(a,b,c) \in Y^{[pqr],st}} \langle \theta^{[pqr],st}, (a,b,c) \rangle$;
/* restore original reparametrization */
11 $\mathrm{repam}(-\overline{\Delta}^{[pq],s}, \overline{x}^{[pq],s}, x^{[pqr],st})$;
12 $\mathrm{repam}(-\Delta^{[pq],t}, \overline{x}^{[pq],t}, x^{[pqr],st})$;
13 $\mathrm{repam}(-\Delta^{[pr],t}, \overline{x}^{[pr],t}, x^{[pqr],st})$;
14 $\mathrm{repam}(-\overline{\Delta}^{[qr],s}, \overline{x}^{[qr],s}, x^{[pqr],st})$;
15 return $lb^1 - lb^0$;