Appendix A: Proof of lemma 1

Let us write down the linear program corresponding to the QPBO method. The roof duality relaxation for function E is given by equation (17). Adding pairwise terms $Cx_a x_b$ for $a, b \in A(p)$, $a \neq b$ to function (13) will affect the relaxation (17) as follows: linear terms Cx_{ab} will be added to function (17), and corresponding constraints will be imposed (see (17)). Since C is a large constant, new variables x_{ab} will be forced to 0. Therefore, we arrive at the following linear program:

$$\min \sum_{a \in A} \bar{\theta}_a x_a + \sum_{(a,b) \in N} \bar{\theta}_{ab} x_{ab} \tag{19}$$
$$\text{subject to} \begin{cases} x_a + x_b \leq 1 & \forall\, a, b \in N(p), p \in P, a \neq b \\ 0 \leq x_a \leq 1 & \forall\, a \in A \\ x_{ab} \leq x_a, x_{ab} \leq x_b & \forall\, (a,b) \in N \\ x_{ab} \geq x_a + x_b - 1 & \forall\, (a,b) \in N \\ x_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

Let us now derive the relaxation solved by the decomposition approach with the linear and maxflow subproblems, i.e. with $I = \{L, F\}$. It is known [2] that the optimal value of the linear matching problem $\Phi_L(\boldsymbol{\theta}^L)$ is equal to the optimal value of the following linear program:

$$\min \sum_{a \in A} \theta^L_a x_a \quad \text{sb.t.} \begin{cases} \sum_{a \in A(p)} x_a \leq 1 & \forall\, p \in P \\ x_a \geq 0 & \forall\, a \in A \end{cases}$$
$$= \max \sum_{p \in P} -\mu_p \quad \text{sb.t.} \begin{cases} -\mu_p - \mu_q \leq \theta^L_a & \forall\, a = (p,q) \in A \\ \mu_p \geq 0 & \forall\, p \in P \end{cases} \tag{20}$$

Similarly, the lower bound for the maxflow subproblem $\Phi_F(\boldsymbol{\theta}^F)$ can be written as the dual problem to (17):

$$\max \sum_{a \in A} -\lambda_a + \sum_{(a,b) \in N} -\lambda_{ab} \tag{21}$$
$$\text{subject to} \begin{cases} -\lambda_a + \sum_{(a,b) \in N} \left[\bar{\lambda}_{ab} - \lambda_{ab}\right] \leq \theta^F_a & \forall\, a \in A \\ -\bar{\lambda}_{ab} - \bar{\lambda}_{ba} + \lambda_{ab} \leq \theta^F_{ab} & \forall\, (a,b) \in N \\ \lambda_a \geq 0 & \forall\, a \in A \\ \bar{\lambda}_{ab} \geq 0, \bar{\lambda}_{ba} \geq 0, \lambda_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

Here we denoted $\bar{\lambda}_{ab}$ and λ_{ab} to be the dual variables for the constraints $x_{ab} \leq x_a$ and $x_{ab} \geq x_a + x_b - 1$, respectively. Note that $\bar{\lambda}_{ab}$ and $\bar{\lambda}_{ba}$ are distinct variables, while λ_{ab} and λ_{ba} denote the same variable.

Using (20) and (21), we can write the optimal lower bound of the decomposition approach as follows:

$$\max_{\substack{\boldsymbol{\theta}^L \in \Omega_L, \\ \boldsymbol{\theta}^L + \boldsymbol{\theta}^F = \bar{\boldsymbol{\theta}}}} \Phi_L(\boldsymbol{\theta}^L) + \Phi_F(\boldsymbol{\theta}^F) = \max_{\boldsymbol{\theta} : \theta_{ab} = 0} \Phi_L(-\boldsymbol{\theta}) + \Phi_F(\bar{\boldsymbol{\theta}} + \boldsymbol{\theta})$$

which yields

$$\max_{\boldsymbol{\theta}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}} \sum_{p \in P} -\mu_p + \sum_{a \in A} -\lambda_a + \sum_{(a,b) \in N} -\lambda_{ab}$$
$$\text{subject to} \begin{cases} -\mu_p - \mu_q \leq -\theta_a & \forall\, a = (p,q) \in A \\ -\lambda_a + \sum_{(a,b) \in N} \left[\bar{\lambda}_{ab} - \lambda_{ab}\right] \leq \bar{\theta}_a + \theta_a & \forall\, a \in A \\ -\bar{\lambda}_{ab} - \bar{\lambda}_{ba} + \lambda_{ab} \leq \bar{\theta}_{ab} & \forall\, (a,b) \in N \\ \mu_p \geq 0 & \forall\, p \in P \\ \lambda_a \geq 0 & \forall\, a \in A \\ \bar{\lambda}_{ab} \geq 0, \bar{\lambda}_{ba} \geq 0, \lambda_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

We can eliminate θ_a from the first and the second constraint and combine them into one constraint, then we obtain

$$\max_{\boldsymbol{\mu}, \boldsymbol{\lambda}, \bar{\boldsymbol{\lambda}}} \sum_{p \in P} -\mu_p + \sum_{a \in A} -\lambda_a + \sum_{(a,b) \in N} -\lambda_{ab}$$
$$\text{sb.t.} \begin{cases} -\mu_p - \mu_q - \lambda_a + \sum_{(a,b) \in N} \left[\bar{\lambda}_{ab} - \lambda_{ab}\right] \leq \bar{\theta}_a \\ \qquad\qquad\qquad\qquad\qquad\qquad \forall\, a = (p,q) \in A \\ -\bar{\lambda}_{ab} - \bar{\lambda}_{ba} + \lambda_{ab} \leq \bar{\theta}_{ab} & \forall\, (a,b) \in N \\ \mu_p \geq 0 & \forall\, p \in P \\ \lambda_a \geq 0 & \forall\, a \in A \\ \bar{\lambda}_{ab} \geq 0, \bar{\lambda}_{ba} \geq 0, \lambda_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

The dual to this linear program is given by

$$\min \sum_{a \in A} \bar{\theta}_a x_a + \sum_{(a,b) \in N} \bar{\theta}_{ab} x_{ab}$$
$$\text{subject to} \begin{cases} \sum_{a \in A(p)} -x_p \geq -1 & \forall\, p \in P \\ -x_a \geq -1 & \forall\, a \in A \\ -x_a - x_b + x_{ab} \geq -1 & \forall\, (a,b) \in N \\ x_a - x_{ab} \geq 0, x_b - x_{ab} \geq 0 & \forall\, (a,b) \in N \\ x_a \geq 0 & \forall\, a \in A \\ x_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

Thus, the optimal value of the lower bound equals

$$\min \sum_{a \in A} \bar{\theta}_a x_a + \sum_{(a,b) \in N} \bar{\theta}_{ab} x_{ab} \tag{22}$$
$$\text{subject to} \begin{cases} \sum_{a \in A(p)} x_p \leq 1 & \forall\, p \in P \\ 0 \leq x_a \leq 1 & \forall\, a \in A \\ x_{ab} \leq x_a, x_{ab} \leq x_b & \forall\, (a,b) \in N \\ x_{ab} \geq x_a + x_b - 1 & \forall\, (a,b) \in N \\ x_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

It is easy to see that the optimal value of (22) is the same or larger than the optimal value of (19). Indeed, the only difference between (19) and (22) is that the first constraint in (22) is tighter than the corresponding constraint in (19): $\sum_{a \in A(p)} x_a \leq 1$ implies $x_a + x_b \leq 1$ for $a, b \in A(p)$, $a \neq b$, but not the other way around. (Note that the labeling $x_a = 0.5$ for $a \in A(p)$ satisfies the latter constraint but not the former, if $|A(p)| > 2$.)

APPENDIX B: PROOF OF LEMMA 2

Consider a local subproblem $\sigma \in I$. Let σ' be a subproblem of σ, i.e. the feasibility set of σ' is contained in the feasibility set of σ: $\Omega_{\sigma'} \subseteq \Omega_\sigma$. It can be seen that adding σ' to I as another local subproblem does not affect the optimal lower bound. Indeed, it is clear that adding σ' cannot decrease the optimal bound. The optimal bound also cannot increase since for any vector $\boldsymbol{\theta}' = (\ldots, \boldsymbol{\theta}^\sigma, \boldsymbol{\theta}^{\sigma'}, \ldots) \in \Omega'$, where Ω' is the constraint set for the new problem, there exists vector $\boldsymbol{\theta} = (\ldots, \boldsymbol{\theta}^\sigma + \boldsymbol{\theta}^{\sigma'}, \ldots) \in \Omega$ whose bound is not worse since

$$\Phi_\sigma(\boldsymbol{\theta}^\sigma + \boldsymbol{\theta}^{\sigma'}) = 2\Phi_\sigma\left(\frac{\boldsymbol{\theta}^\sigma + \boldsymbol{\theta}^{\sigma'}}{2}\right) \geq \Phi_\sigma(\boldsymbol{\theta}^\sigma) + \Phi_{\sigma'}(\boldsymbol{\theta}^{\sigma'}).$$

(The inequality holds since $\Phi_{\sigma'}(\cdot)$ is the same function as $\Phi_\sigma(\cdot)$, and it is concave.)

Let us prove part (a). Let I be a set of subproblem indexes which does not include the linear problem L. We need to show that adding L to I cannot increase the optimal lower bound. Instead of L, let us add a new subproblem p to I for each point $p \in P$ which includes only assignments in $A(p)$ (and does not include any edges), i.e. the feasibility set Ω_p for this subproblem is defined by $\theta^p_a = 0$ for all assignments $a \in A - A(p)$ and $\theta^p_{ab} = 0$ for all edges $(a,b) \in N$. As follows from the argument above and conditions of part (a), this operation cannot improve the best lower bound. Thus, it suffices to prove that replacing the new set of subproblems with L would not improve optimal bound. In other words, we need to show that for any $\boldsymbol{\theta}^L$

$$\Phi_L(\boldsymbol{\theta}^L) \;\leq \max \sum_{p\in P} \Phi_p(\boldsymbol{\theta}^p) \qquad \text{sb.t.} \sum_{p\in P} \boldsymbol{\theta}^p = \boldsymbol{\theta}^L \tag{23}$$

Using LP duality, it is easy to show that in fact an equality holds in (23). Indeed, the optimal solution for vector $\boldsymbol{\theta}^p$ can be obtained as follows: $\Phi_p(\boldsymbol{\theta}^p) = \min\{0, \min_{a\in A(p)} \theta^p_a\}$. Thus, the maximization problem in (23) can be written as

$$\max \sum_{p\in P} -\mu_p \qquad \text{sb.t.} \begin{cases} \theta^p_a + \theta^q_a = \theta^L_a & \forall\, a = (p,q) \in A \\ -\mu_p \leq \theta^p_a & \forall\, p \in P, a \in A(p) \\ -\mu_p \leq 0 & \forall\, p \in P \end{cases}$$

Constraints $\{\theta^p_a + \theta^q_a = \theta^L_a, -\mu_p \leq \theta^p_a, -\mu_q \leq \theta^q_a\}$ for $a = (p,q) \in A$ can be replaced with a single constraint $-\mu_p - \mu_q \leq \theta^L_a$ since variables θ^p_a and θ^q_a are not involved in any other constraints. Then we arrive at the linear program (20) which equals $\Phi_L(\boldsymbol{\theta}^L)$.

Let us now prove part (b). Using a similar argumentation, we conclude that it suffices to prove that

$$\Phi_F(\boldsymbol{\theta}^F) \leq \max \sum_{a\in A} \Phi_a(\boldsymbol{\theta}^a) + \sum_{(a,b)\in N} \Phi_{ab}(\boldsymbol{\theta}^{ab}) \tag{24}$$
$$\text{subject to} \sum_{a\in A} \boldsymbol{\theta}^a + \sum_{(a,b)\in N} \boldsymbol{\theta}^{ab} = \boldsymbol{\theta}^F$$

where $\sigma = a$ is a local subproblem in which only the element θ^a_a is allowed to be non-zero and $\sigma = (a,b)$ is a local subproblem in which only the elements θ^{ab}_a, θ^{ab}_b, θ^{ab}_{ab} are allowed to be non-zero.

It can be shown that if we take $\Phi_{ab}(\boldsymbol{\theta}^{ab})$ to be a lower bound $\min_{\boldsymbol{x}\in\{0,1\}^A} E(\boldsymbol{x}|\boldsymbol{\theta}^{ab})$ rather than the global minimum $\min_{\boldsymbol{x}\in M} E(\boldsymbol{x} \mid \boldsymbol{\theta}^{ab})$ then we get an equality in (24). (An equivalent fact was proved in [21].) This implies (24) since using the global minimum instead of a lower bound can only increase the RHS.

For completeness, let us prove this equality. We have

$$\begin{aligned} \Phi_a(\boldsymbol{\theta}^a) &= \min\{0, \theta^a_a\} \\ \Phi_{ab}(\boldsymbol{\theta}^{ab}) &= \min\{0, \theta^{ab}_a, \theta^{ab}_b, \theta^{ab}_a + \theta^{ab}_b + \theta^{ab}_{ab}\} \end{aligned}$$

Thus, the maximization in (24) can be written as

$$\max \sum_{a\in A} -\lambda_a + \sum_{(a,b)\in N} -\lambda_{ab}$$
$$\text{subject to} \begin{cases} \theta^a_a + \sum_{(a,b)\in N} \theta^{ab}_a = \theta^F_a & \forall\, a \in A \\ -\lambda_a \leq \theta^a_a & \forall\, a \in A \\ -\lambda_a \leq 0 & \forall\, a \in A \\ -\lambda_{ab} \leq \theta^{ab}_a, -\lambda_{ab} \leq \theta^{ab}_b & \forall\, (a,b) \in N \\ -\lambda_{ab} \leq \theta^{ab}_a + \theta^{ab}_b + \theta^F_{ab} & \forall\, (a,b) \in N \\ -\lambda_{ab} \leq 0 & \forall\, (a,b) \in N \end{cases}$$

We can eliminate θ^a_a from the first and the second constraint and combine them into one constraint:

$$\max \sum_{a\in A} -\lambda_a + \sum_{(a,b)\in N} -\lambda_{ab}$$
$$\text{subject to} \begin{cases} -\lambda_a + \sum_{(a,b)\in N} \theta^{ab}_a \leq \theta^F_a & \forall\, a \in A \\ \lambda_a \geq 0 & \forall\, a \in A \\ \theta^{ab}_a + \lambda_{ab} \geq 0, \theta^{ab}_b + \lambda_{ab} \geq 0 & \forall\, (a,b) \in N \\ -\lambda_{ab} - \theta^{ab}_a - \theta^{ab}_b \leq \theta^F_{ab} & \forall\, (a,b) \in N \\ \lambda_{ab} \geq 0 & \forall\, (a,b) \in N \end{cases}$$

Let us use variables $\bar{\lambda}_{ab}$ instead of θ^{ab}_a such that $\theta^{ab}_a = \bar{\lambda}_{ab} - \lambda_{ab}$, or $\bar{\lambda}_{ab} = \theta^{ab}_a + \lambda_{ab}$. It is straightforward to see that then we arrive at the linear program (21) whose optimal value equals $\Phi_F(\boldsymbol{\theta}^F)$.