A Proof of Theorem 4.1

Let a stationary point $\boldsymbol{\pi} \in \mathbb{R}_+^n$ of the Augmented Lagrangian (7) be such that:

$$\frac{\partial L_\rho(\boldsymbol{\pi}, \boldsymbol{W}^k, \boldsymbol{y}^k)}{\partial \pi_i} = 0 \ \ \forall i \in [n] \tag{20a}$$

$$\Leftrightarrow \frac{\partial \mathcal{L}(\mathcal{D}|\boldsymbol{\pi})}{\partial \pi_i} + y_i^k + \rho \frac{\partial D_p(\boldsymbol{\pi}||\tilde{\boldsymbol{\pi}}^k)}{\partial \pi_i} = 0. \ \ \forall i \in [n] \tag{20b}$$

Let $\sigma_i(\boldsymbol{\pi}) = \rho \frac{\partial D_p(\boldsymbol{\pi}||\tilde{\boldsymbol{\pi}}^k)}{\partial \pi_i} + y_i^k$, for all $i \in [n]$. Then, Eq.(20a) is equivalent to:

$$\frac{\partial \mathcal{L}(\mathcal{D}|\boldsymbol{\pi})}{\partial \pi_i} + \sigma_i(\boldsymbol{\pi}) = 0 \ \ \forall i \in [n]\,. \tag{21}$$

Partial derivatives of the negative log-likelihood $\mathcal{L}(\mathcal{D}|\boldsymbol{\pi})$ are given by:

$$\frac{\partial \mathcal{L}(\mathcal{D}|\boldsymbol{\pi})}{\partial \pi_i} = \sum_{\ell \in W_i} \left(\frac{1}{\sum_{t \in A_\ell} \pi_t} - \frac{1}{\pi_i} \right) + \sum_{\ell \in L_i} \frac{1}{\sum_{t \in A_\ell} \pi_t}, \tag{22}$$

for all $i \in [n]$, where $W_i = \{\ell \,|\, i \in A_\ell, c_\ell = i\}$ is the set of observations where sample $i \in [n]$ is chosen and $L_i = \{\ell \,|\, i \in A_\ell, c_\ell \neq i\}$ is the set of observations where sample $i \in [n]$ is not chosen. Setting $\frac{\partial \mathcal{L}(\mathcal{D}|\boldsymbol{\pi})}{\partial \pi_i}$ from Eq. (22) to Eq. (21), we have:

$$\frac{\partial L_\rho(\boldsymbol{\pi}, \boldsymbol{W}^k, \boldsymbol{y}^k)}{\partial \pi_i} = \sum_{\ell \in W_i} \left(\frac{1}{\sum_{t \in A_\ell} \pi_t} - \frac{1}{\pi_i} \right) + \sum_{\ell \in L_i} \frac{1}{\sum_{t \in A_\ell} \pi_t} + \sigma_i(\boldsymbol{\pi}) = 0, \tag{23}$$

for all $i \in [n]$. Multiplying both sides of Eq. (23) with $-\pi_i$, $i \in [n]$, we have:

$$\sum_{\ell \in W_i} \left(\frac{\sum_{j \neq i \in A_\ell} \pi_j}{\sum_{t \in A_\ell} \pi_t} \right) - \sum_{\ell \in L_i} \left(\frac{\pi_i}{\sum_{t \in A_\ell} \pi_t} \right) - \pi_i \sigma_i(\boldsymbol{\pi}) = 0, \tag{24}$$

for all $i \in [n]$. Note that $\sum_{\ell \in W_i} \sum_{j \neq i \in A_\ell} \cdot = \sum_{j \neq i} \sum_{\ell \in W_i \cap L_j} \cdot$ and $\sum_{\ell \in L_i} \cdot = \sum_{j \neq i} \sum_{\ell \in W_j \cap L_i} \cdot$. Accordingly, we rewrite Eq. (24) as:

$$\sum_{j \neq i} \sum_{\ell \in W_i \cap L_j} \left(\frac{\pi_j}{\sum_{t \in A_\ell} \pi_t} \right) - \sum_{j \neq i} \sum_{\ell \in W_j \cap L_i} \left(\frac{\pi_i}{\sum_{t \in A_\ell} \pi_t} \right) - \pi_i \sigma_i(\boldsymbol{\pi}) = 0, \tag{25}$$

for all $i \in [n]$. Then, the stationarity condition given by Eq.(20a) is equivalent to:

$$\sum_{j \neq i} \pi_j \lambda_{ji}(\boldsymbol{\pi}) - \sum_{j \neq i} \pi_i \lambda_{ij}(\boldsymbol{\pi}) = \pi_i \sigma_i(\boldsymbol{\pi}) \ \ \forall i \in [n]\,, \tag{26}$$

where $\lambda_{ji}(\boldsymbol{\pi})$, $i, j \in [n]\,, i \neq j$ are given by Eq. (12).

It is not evident that Eq.(26) corresponds to the balance equations of an MC as, in general, $\boldsymbol{\sigma}(\boldsymbol{\pi}) = [\sigma_i(\boldsymbol{\pi})]_{i \in [n]} \neq \mathbf{0}$. Nevertheless, for $\sigma_i(\boldsymbol{\pi}) = \rho \frac{\partial D_p(\boldsymbol{\pi}||\tilde{\boldsymbol{\pi}}^k)}{\partial \pi_i} + y_i^k$, $i \in [n]$, Eq.(26) has the same form as the balance equations in Theorem 4. 2 established by Yıldız et al. (2020). By this theorem, a stationary $\boldsymbol{\pi} \in \mathbb{R}_+^n$ satisfying (20a) is also the stationary distribution of the continuous-time MC with transition rates given by Eq. (10). □

B Datasets

Retinopathy of Prematurity (ROP). The Retinopathy of Prematurity (ROP) dataset contains $n = 100$ vessel-segmented retina images with dimensions $d = 224 \times 224$ (Ataer-Cansızoğlu, 2015). Experts are provided with two images and are asked to choose the image with higher severity of the ROP disease. Five experts independently label 5941 image pairs; the resulting dataset contains $m = 29705$ pairwise comparisons. Note that some pairs are labelled more than once by different experts.

International Conference on Learning Representations (ICLR). The ICLR Dataset contains abstracts and reviewer ratings of 2561 papers that are submitted to ICLR 2020 conference and are available on OpenReview website (Sun, 2020). We choose the top $n = 100$ papers, and extract $d = 768$ numerical features from each abstract using the Deep Bidirectional Transformers (BERT) (Devlin et al., 2019) architecture, pre-trained on the Books Corpus dataset (Zhu et al., 2015) and English Wikipedia. We normalize $\boldsymbol{X}$ to have 0 mean and unit variance over samples $[n]$. We generate all possible $m = 120,324(2,248,524)$ $K = 3(4)$-way rankings w.r.t. the relative order of the average reviewer ratings. We add noise to the resulting rankings following the same process as Movehub-Cost.

Movehub-Cost. The Movehub-Cost dataset contains the total ranking of 216 cities w.r.t. cost of living (Blitzer, 2017). Each city is associated with $d = 6$ numerical features, which are average costs for cappuccino, cinema, wine, gasoline, rent, and disposable income. We normalize $\boldsymbol{X}$ to have 0 mean and unit variance over samples $[n]$. We select $n = 50$ cities and generate all $m = 230,298(2,118,756)$ $K = 4(5)$-way rankings w.r.t. the relative order of the queried cities in the total ranking. To mimic the real-life noise introduced by human labelling, we apply the following post-processing to the resulting rankings: For each ranking, we sample a value uniformly at random in $[0, 1]$. If the value is less than 0.1, we add noise to the ranking by a cyclic permutation of the ranked samples.

Movehub-Quality. The Movehub-Quality dataset contains total ranking of the same 216 cities as Movehub-Cost, this time w.r.t. quality of life. Each city is associated with $d = 5$ numerical features, including overall scores for purchase power, healthcare, pollution, quality of life, and crime. We normalize $\boldsymbol{X}$ to have 0 mean and unit variance over samples $[n]$. We select $n = 50$ cities and generate all $m = 230,298(2,118,756)$ $K = 4(5)$-way rankings w.r.t. the relative order of the queried cities in the total ranking. We add noise to the rankings following the same process as Movehub-Cost.

IMDB. The IMDB Movies Dataset contains IMDB ratings of 14,762 movies, each of which is associated with $d = 36$ numerical features (Leka, 2016). We normalize $\boldsymbol{X}$ to have 0 mean and unit variance over samples $[n]$. We select $n = 50$ movies and generate all possible $m = 85,583$ $K = 4$-way rankings w.r.t. the relative order of the ratings of queried movies. We add noise to the resulting rankings following the same process as Movehub-Cost.