A Arora-Hazan-Kale result

For completeness, we quote here verbatim (except for the numbering) the relevant definitions and results from (Arora et al., 2012, Sec. 3.3.1, p. 137).

Imagine that we have the following feasibility problem:

$$\exists ?\mathbf{x} \in \mathcal{P} : \ \forall i \in [m] : \ f_i(\mathbf{x}) \geq 0 \tag{5}$$

where $\mathcal{P} \in \mathbb{R}^n$ is a convex domain, and for $i \in [m]$, $f_i : \mathcal{P} \to \mathbb{R}$ are concave functions. We wish to satisfy this system approximately, up to an additive error of ε. We assume the existence of an Oracle, which, when given a probability distribution $\mathbf{p} = (p_1, p_2, \ldots, , p_m)$ solves the following feasibility problem:

$$\exists ?\mathbf{x} \in \mathcal{P} : \ \sum_i p_i f_i(\mathbf{x}) \geq 0 \tag{6}$$

An Oracle would be called (ℓ, ρ)-bounded if there is a fixed subset of constraints $I \subseteq [m]$ such that whenever it returns a feasible solution $\mathbf{x}$ to (6), all constraints $i \in I$ take values i the range $[-\ell, \rho]$ on the point $\mathbf{x}$, and all the rest take values in $[-\rho, \ell]$.

Theorem A.1 (Theorem 3.4 in Arora et al. (2012))**.** *Let $\varepsilon > 0$ be a given error parameter. Suppose there exists an (ℓ, ρ)-bounded* Oracle *for the feasibility problem (5). Assume the $\ell \geq \varepsilon/2$. Then there is an algorithms which either solves the problem up to an additive error of ε, or correctly concludes that the system s infeasible, making only $O(\ell\rho \log(m)/\varepsilon^2)$ calls to the* Oracle*, with an additional processing time of $O(m)$ per call.*

B Approximate oracle

To use the MWU method (see Theorem 3.5 in Arora et al. (2012)), we design an approximate oracle for the following problem.

Problem B.1. *Given non-negative edge weights w_Φ and w_{ij}, which add up to 1, find $\widetilde{\mathbf{Y}}$ such that*

$$w_\Phi h_\Phi(\widetilde{\mathbf{Y}}) + \sum_{(i,j)\in E} w_{(i,j)} h_{(i,j)}(\widetilde{\mathbf{Y}}) \geq -\varepsilon. \tag{7}$$

Let $\mu_{ij} = \frac{w_{ij}+\varepsilon/(m+1)}{M^2\|x_i - x_j\|^2}$ and $\lambda_i = \lambda = (w_\Phi + \varepsilon/(m+1))/\Phi_0$. We solve Laplace's problem with parameters μ_{ij} and λ_i (see Section 3.2 and Line 9 of the algorithm). We get a matrix $\widetilde{\mathbf{Y}} = (\tilde{y}_1, \ldots, \tilde{y}_n)$ minimizing

$$\lambda \sum_{i=1}^n \|y_i - \tilde{y}_i\|^2 + \sum_{(i,j)\in E} \mu_{ij} \|\tilde{y}_i - \tilde{y}_j\|^2.$$

Consider the optimal solution $\tilde{y}_1^*, \ldots, \tilde{y}_n^*$ for Lipschitz Smoothing. We have,

$$\begin{aligned}
&\lambda \sum_{i=1}^n \|y_i - \tilde{y}_i\|^2 + \sum_{(i,j)\in E} \mu_{ij} \|\tilde{y}_i - \tilde{y}_j\|^2 && (8)\\
&\leq \lambda \sum_{i=1}^n \|y_i - \tilde{y}_i^*\|^2 + \sum_{(i,j)\in E} \mu_{ij} \|\tilde{y}_i^* - \tilde{y}_j^*\|^2 \\
&\leq (w_\Phi + \varepsilon/(m+1)) + \\
&\sum_{(i,j)\in E} \left(w_{ij} + \frac{1}{m+1}\right) \frac{\|\tilde{y}_i^* - \tilde{y}_j^*\|}{M^2\|x_i - x_j\|^2} \\
&\leq \left(w_\Phi + \sum_{(i,j)\in E} w_{ij}\right) + \varepsilon = 1 + \varepsilon && (9)
\end{aligned}$$

We verify that $\widetilde{\mathbf{Y}}$ is a feasible solution for Problem B.1. We have,

$$\begin{aligned}
1 &- \left(w_\Phi h_\Phi(\widetilde{\mathbf{Y}}) + \sum_{(i,j)\in E} w_{(i,j)} h_{(i,j)}(y)\right) \\
&= w_\Phi(1 - h_\Phi) + \sum_{(i,j)\in E} w_{(i,j)}(1 - h_{(i,j)}(y)) \\
&\leq \sqrt{w_\Phi(1 - h_\Phi)^2 + \sum_{(i,j)\in E} w_{(i,j)}(1 - h_{(i,j)}(y))^2} \\
&= \sqrt{w_\Phi \frac{\Phi(\mathbf{Y}, \widetilde{\mathbf{Y}})}{\Phi_0} + \sum_{(i,j)\in E} w_{(i,j)} \frac{\|\tilde{y}_i - \tilde{y}_j\|^2}{M^2\|x_i - x_j\|^2}} \\
&\leq \sqrt{\lambda\Phi(\mathbf{Y}, \widetilde{\mathbf{Y}}) + \sum_{(i,j)\in E} w_{(i,j)} \mu_{ij} \|\tilde{y}_i - \tilde{y}_j\|^2} \\
&\leq \sqrt{1 + \varepsilon} \leq 1 + \varepsilon,
\end{aligned}$$

as required.

Finally, we bound the width of the problem. We have $h_\Phi(\widetilde{\mathbf{Y}}) \leq 1$ and $h_{ij}(\widetilde{\mathbf{Y}}) \leq 1$. Then, using (8), we get

$$\begin{aligned}
(1 - h_\Phi(\widetilde{\mathbf{Y}}))^2 &= \frac{1}{\Phi_0} \sum_{i=1}^n \|y_i - \tilde{y}_i\|^2 \leq \frac{1+\varepsilon}{\lambda\Phi_0} \\
&\leq \frac{(1+\varepsilon)(m+1)}{\varepsilon}
\end{aligned}$$

Therefore, $-h_\Phi(\widetilde{\mathbf{Y}}) \leq O(\sqrt{m/\varepsilon})$.

Similarly,

$$\begin{aligned}
(1 - h_{ij}(\widetilde{\mathbf{Y}}))^2 &= \frac{\|y_i - \tilde{y}_i\|^2}{M^2\|x_i - x_j\|^2} \leq \frac{1+\varepsilon}{\mu_{ij} \cdot M^2\|x_i - x_j\|^2} \\
&\leq \frac{(1+\varepsilon)(m+1)}{\varepsilon}
\end{aligned}$$

Therefore, $-h_{ij}(\widetilde{\mathbf{Y}}) \leq O(\sqrt{m/\varepsilon})$.