Supplement to ”On Unbalanced Optimal Transport: An Analysis of Sinkhorn Algorithm”

In this appendix, we provide proofs for the remaining results in the paper.

6. Proofs of Remaining Results

Before proceeding with the proofs, we state the following simple inequalities:

Lemma 6. *The following inequalities are true for all positive* x_i, y_i, x, y *and* $0 \le z < \frac{1}{2}$*:*

$$\begin{aligned}
&(a) \quad \min_{1\le i\le n} \frac{x_i}{y_i} \le \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i} \le \max_{1\le i\le n} \frac{x_i}{y_i},\\
&(b) \quad \exp(z) \le 1 + |z| + |z|^2,\\
&(c) \quad \textit{If} \max\left\{\frac{x}{y}, \frac{y}{x}\right\} \le 1+\delta, \textit{ then } |x-y| \le \delta \min\{x, y\},\\
&(d) \quad \left(1+\frac{1}{x}\right)^{x+1} \ge e.
\end{aligned}$$

Proof of Lemma 6.
(a) Given x_i and y_i positive, we have

$$\begin{gathered}
\min_{1\le i\le n} \frac{x_i}{y_i} \le \frac{x_j}{y_j} \le \max_{1\le i\le n} \frac{x_i}{y_i},\\
y_j\left(\min_{1\le i\le n} \frac{x_i}{y_i}\right) \le x_j \le y_j\left(\max_{1\le i\le n} \frac{x_i}{y_i}\right).
\end{gathered}$$

Taking the sum over j, we get

$$\begin{gathered}
\left(\sum_{j=1}^n y_j\right)\left(\min_{1\le i\le n} \frac{x_i}{y_i}\right) \le \sum_{j=1}^n x_j \le \left(\sum_{j=1}^n y_j\right)\left(\max_{1\le i\le n} \frac{x_i}{y_i} y_j\right),\\
\min_{1\le i\le n} \frac{x_i}{y_i} \le \frac{\sum_{j=1}^n x_j}{\sum_{j=1}^n y_j} \le \max_{1\le i\le n} \frac{x_i}{y_i}.
\end{gathered}$$

(b) For the second inequality, $\exp(x) \le 1 + |x| + |x|^2$, we have to deal with the case $x > 0$. Since $x \le \frac{1}{2}$,

$$\begin{aligned}
\exp(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!} &= 1 + x + x^2 - \frac{x^2}{2} + \sum_{n=3}^{\infty} \frac{x^n}{n!} \le 1 + x + x^2 - \frac{x^2}{2} + \frac{x^3}{6}\sum_{n=3}^{\infty} x^{n-3},\\
&\le 1 + x + x^2 - \frac{x^2}{2} + \frac{x^3}{6}\frac{1}{1-x} \le 1 + x + x^2 - \frac{x^2}{2} + \frac{x^3}{3} \le 1 + x + x^2.
\end{aligned}$$

(c) For the third inequality, WLOG assume $x > y$. Then, we have

$$\frac{x}{y} \le 1 + \delta \Rightarrow x \le y + y\delta \Rightarrow |x-y| \le y\delta.$$

(d) For the fourth inequality, taking the $\log$ of both sides, it is equivalent to $(x+1)\big[\log(x+1) - \log(x)\big] \ge 1$. By the mean value theorem, there exists a number y between x and $x+1$ such that $\log(x+1) - \log(x) = 1/y$, then $(x+1)/y \ge 1$. $\square$

By the choice of $\eta = \frac{\varepsilon}{U}$ and the definition of U, we also have the following conditions on η:

$$\eta \leq \frac{1}{2}; \quad \frac{\eta}{\tau} \leq \frac{1}{4\log(n)\max\{1, \alpha+\beta\}}. \tag{19}$$

Now we come to the proofs of lemmas and the corollary in the main text.

6.1. Proof of Lemma 2

(a) + (b): From the definitions of a_i^k and a_i^*, we have

$$\log\left(\frac{a_i^*}{a_i^k}\right) = \left(\frac{u_i^* - u_i^k}{\eta}\right) + \log\left(\frac{\sum_{j=1}^n \exp(\frac{v_j^* - C_{ij}}{\eta})}{\sum_{j=1}^n \exp(\frac{v_j^k - C_{ij}}{\eta})}\right).$$

The required inequalities are equivalent to an upper bound and a lower bound for the second term of the RHS. Apply part (a) of Lemma 6, we obtain

$$\min_{1\leq j\leq n} \frac{v_j^* - v_j^k}{\eta} \leq \log\left(\frac{a_i^*}{a_i^k}\right) - \frac{u_i^* - u_i^k}{\eta} \leq \max_{1\leq j\leq n} \frac{v_j^* - v_j^k}{\eta}.$$

Part (b) follows similarly. Therefore, we obtain the conclusion of Lemma 2.

6.2. Proof of Corollary 2

Recall that we have proved in Lemma 4:

$$\begin{aligned} g(X^*) + (2\tau+\eta)x^* &= \tau(\alpha+\beta), \\ f(\widehat{X}) + 2\tau\widehat{x} &= \tau(\alpha+\beta). \end{aligned}$$

From the second equality and the fact that $f(\widehat{X}) \geq 0$ (it is easy to see that for X that $X_{ij} \geq 0$, the **KL** terms and $\langle C, X\rangle$ are all non-negative), we immediately have $\widehat{x} \leq \frac{\alpha+\beta}{2}$, proving the second inequality. For the first inequality, we have $g(X^*) \geq -\eta H(X^*) \geq -2\eta x^* \log(n) - \eta x^* + \eta x^* \log(x^*)$. Therefore, we find that

$$\begin{aligned} -2\eta x^* \log(n) - \eta x^* + \eta x^* \log(x^*) &\leq \tau(\alpha+\beta) - (2\tau+\eta)x^*, \\ \eta x^* \log(x^*) + 2\Big(\tau - \eta\log(n)\Big)x^* &\leq \tau(\alpha+\beta). \end{aligned}$$

It follows from the inequality $z\log(z) \geq z - 1$ for all $z > 0$ that

$$\begin{aligned} \eta(x^* - 1) + 2(\tau - \eta\log(n))x^* &\leq \tau(\alpha+\beta), \\ x^*(2\tau - 2\eta\log(n) + \eta) &\leq \tau(\alpha+\beta) + \eta. \end{aligned}$$

By inequality (19), $4\eta\log(n) \leq \tau$. Then

$$\begin{aligned} x^* &\leq \frac{\tau(\alpha+\beta)+\eta}{2\tau - 2\eta\log(n) + \eta} \leq \frac{\tau(\alpha+\beta) - (\alpha+\beta)\eta\log(n)}{2\tau - 2\eta\log(n)} + \frac{(\alpha+\beta)\eta\log(n) + \eta}{2\tau - 2\eta\log(n)}, \\ &\leq \frac{\alpha+\beta}{2} + (\alpha+\beta)\frac{\eta\log(n)}{2\tau - 2\eta\log(n)} + \frac{\frac{\tau}{4\log(n)}}{\frac{3}{2}\tau} \leq \Big(\frac{1}{2} + \frac{\eta\log(n)}{2\tau - 2\eta\log(n)}\Big)(\alpha+\beta) + \frac{1}{6\log(n)}. \end{aligned}$$

As a consequence, we obtain the conclusion of the corollary.

6.3. Proof of Lemma 5

(a) We prove that $\Lambda_k \leq \frac{\eta^2}{8(\tau+1)}$ for $\eta = \frac{\varepsilon}{U}$ and $k \geq \left(\frac{\tau}{\eta} + 1\right)\left[\log(8\eta R) + \log(\tau(\tau+1)) + 3\log(\frac{1}{\eta})\right]$ (note that the stated bound can be obtained by replacing k with $k-1$).

Denote $\frac{8\eta R(\tau+1)}{\tau^2} = D$ and $\frac{\eta}{\tau} = s > 0$. From inequality (19), we have $s < 1$. The required inequality is equivalent to

$$\frac{\eta^2}{8(\tau+1)} \geq \Big(\frac{\tau}{\tau+\eta}\Big)^k \tau R \iff \Big(\frac{\tau+\eta}{\tau}\Big)^k \frac{\eta^3}{\tau^3} \geq \frac{8\eta R(\tau+1)}{\tau^2} \iff \big(1+s\big)^k s^3 \geq D.$$

Let $t = 1 + \frac{\log(D)}{3\log(\frac{1}{s})}$. By definition (5), $R \geq \log(n)$, thus $D \geq \frac{8\eta\log(n)(\tau+1)}{\tau^2} > \frac{\eta^3}{\tau^3} = s^3$ and $t > 1 + \frac{3\log(s)}{3\log(\frac{1}{s})} = 0$. We claim the following chain of inequalities

$$\begin{aligned} s^3(1+s)^k &\geq s^3(1+s)^{(\frac{1}{s}+1)3\log(\frac{1}{s})t} \\ &\geq s^3 e^{3\log(\frac{1}{s})t}. \end{aligned}$$

The first inequality results from $k \geq \big(\frac{\tau U}{\epsilon}+1\big)\big[\log(8\eta R) + \log(\tau(\tau+1)) + 3\log\big(\frac{U}{\epsilon}\big)\big] = \big(1+\frac{1}{s}\big)3\log\big(\frac{1}{s}\big)t > 0$ (using the definitions of D, s, the choice of t and $\eta = \frac{\varepsilon}{U}$). The second inequality is due to part (d) of Lemma 6. The last equality is

$$s^3 e^{3\log(\frac{1}{s})t} = \frac{1}{s^{3t-3}} = \frac{1}{s^{\log(D)/\log(1/s)}} = \frac{1}{s^{-\log_s(D)}} = D.$$

We have thus proved our claim of part (a).

(b) We need to prove $|x^k - x^*| \leq \frac{3}{\eta}\min\{x^*, x^k\}\Delta_k$. From the definition of x^k and x^* and note that they are non-negative:

$$x^k = \sum_{i,j=1}^{n} \exp\left(\frac{u_i^k + v_j^k - C_{ij}}{\eta}\right) \quad \text{and} \quad x^* = \sum_{i,j=1}^{n} \exp\left(\frac{u_i^* + v_j^* - C_{ij}}{\eta}\right).$$

Now, we have

$$\frac{\exp\left(\frac{u_i^k+v_j^k-C_{ij}}{\eta}\right)}{\exp\left(\frac{u_i^*+v_j^*-C_{ij}}{\eta}\right)} = \exp\left(\frac{u_i^k - u_i^*}{\eta}\right)\exp\left(\frac{v_j^k - v_j^*}{\eta}\right) \leq \left[\max_{1\leq i\leq n}\exp\left(\frac{|u_i^k - u_i^*|}{\eta}\right)\right]\left[\max_{1\leq j\leq n}\exp\left(\frac{|v_j^k - v_j^*|}{\eta}\right)\right].$$

Note that each of x^k and x^* is the sum of n^2 elements and the ratio between $\exp\left(\frac{u_i^k+v_j^k-C_{ij}}{\eta}\right)$ and $\exp\left(\frac{u_i^*+v_j^*-C_{ij}}{\eta}\right)$ is bounded by $\left[\max_{1\leq i\leq n}\exp\left(\frac{|u_i^k - u_i^*|}{\eta}\right)\right]\left[\max_{1\leq j\leq n}\exp\left(\frac{|v_j^k - v_j^*|}{\eta}\right)\right]$ for all pairs i, j. Apply part (a) of Lemma 6, we find that

$$\max\left\{\frac{x^*}{x^k}, \frac{x^k}{x^*}\right\} \leq \left[\max_{1\leq i\leq n}\exp\left(\frac{|u_i^k - u_i^*|}{\eta}\right)\right]\left[\max_{1\leq j\leq n}\exp\left(\frac{|v_j^k - v_j^*|}{\eta}\right)\right].$$

We have proved from part (a) that $\Lambda_{k-1} \leq \frac{\eta^2}{8(\tau+1)} \leq \frac{\eta^2}{8}$. From Theorem 1 we get $\Delta_k \leq \Lambda_{k-1}$. It means that

$$\max_{i,j}\left\{\left|\frac{u_i^k - u_i^*}{\eta}\right|, \left|\frac{v_j^k - v_j^*}{\eta}\right|\right\} = \frac{\Delta_k}{\eta} \leq \frac{\Lambda_{k-1}}{\eta} \leq \frac{\eta}{8} \leq \frac{1}{8}.$$

Apply part (b) of Lemma 6,

$$\exp\frac{|u_i^k - u_i^*|}{\eta} \leq 1 + \frac{|u_i^k - u_i^*|}{\eta} + \Big(\frac{|u_i^k - u_i^*|}{\eta}\Big)^2, \quad \text{and} \quad \exp\frac{|v_j^k - v_j^*|}{\eta} \leq 1 + \frac{|v_j^k - v_j^*|}{\eta} + \Big(\frac{|v_j^k - v_j^*|}{\eta}\Big)^2.$$

Then, we find that

$$\begin{aligned} \max\left\{\frac{x^*}{x^k}, \frac{x^k}{x^*}\right\} &\leq \left(1 + \frac{1}{\eta}\Delta_k + \frac{1}{\eta^2}\Delta_k^2\right)\left(1 + \frac{1}{\eta}\Delta_k + \frac{1}{\eta^2}\Delta_k^2\right) = 1 + 2\frac{\Delta_k}{\eta} + 3\frac{\Delta_k^2}{\eta^2} + 2\frac{\Delta_k^3}{\eta^3} + \frac{\Delta_k^4}{\eta^4} \\ &\leq 1 + \frac{\Delta_k}{\eta}\Big(2 + 3\frac{\Delta_k}{\eta} + 2\frac{\Delta_k^2}{\eta^2} + \frac{\Delta_k^3}{\eta^3}\Big) \leq 1 + \frac{\Delta_k}{\eta}\Big(2 + 3\frac{1}{8} + 2\frac{1}{8^2} + \frac{1}{8^3}\Big) \\ &\leq 1 + 3\frac{\Delta_k}{\eta}. \end{aligned}$$

Apply part (c) of Lemma 6, we get

$$|x^k - x^*| \le \frac{3}{\eta} \Delta_k \min\{x^k, x^*\}.$$

Therefore, we obtain the conclusion of part (b).

(c) From Lemma 5(a) and Theorem 1 we have $\frac{\Delta_k}{\eta} \le \frac{\Lambda_k}{\eta} \le \frac{\eta}{8} \le \frac{1}{12}$. By part (b) of Lemma 5, we have $x^k \le x^* + \frac{3}{\eta}\Delta_k x^* \le \frac{3}{2}x^*$. Then, we obtain that

$$\begin{aligned}
x^k \le x^* + \frac{3}{\eta}\Delta_k x^* &\le \left[(\alpha+\beta)\Big(\frac{1}{2} + \frac{\eta\log(n)}{2\tau - 2\eta\log(n)}\Big) + \frac{1}{6\log(n)}\right]\big(1 + 3\frac{\Delta_k}{\eta}\big) \\
&\le (\alpha+\beta)\Big(\frac{1}{2} + \frac{\eta\log(n)}{2\tau - 2\eta\log(n)}\Big)\big(1 + 3\frac{\Delta_k}{\eta}\big) + \frac{1}{4\log(n)} \\
&\le \frac{1}{2}(\alpha+\beta) + (\alpha+\beta)\frac{3}{2}\frac{\Delta_k}{\eta} + (\alpha+\beta)\frac{\eta\log(n)}{\tau} + \frac{1}{4\log(n)} \\
&\le \frac{1}{2}\big(\alpha+\beta\big) + \frac{1}{4} + (\alpha+\beta)\frac{3\eta}{12\tau} + \frac{1}{4\log(n)} \\
&\le \frac{1}{2}\big(\alpha+\beta\big) + \frac{1}{2} + \frac{1}{4\log(n)}.
\end{aligned}$$

As a consequence, we reach the conclusion of part (c).