Appendix

A. Proof of Theorem 1

Problem 2 ($k-\mathrm{CenterUniform}$)**.** *Given a uniform metric space $d : V \times V \mapsto \mathbb{R}_{\geq 0}$ ($d(u,v) = 1$ in case $(u \neq v)$) and a set of requests $R_1, \ldots, R_m \subseteq V$. Select $F \subseteq V$ such as $|F| = k$ and $\sum_{s=1}^{m} C_{R_s}(F)$ is minimized where p is ∞.*

Lemma 6. *Any c-approximation algorithm for $k-\mathrm{CenterUniform}$ implies a c-approximation algorithm for $\mathrm{Min}-p-\mathrm{Union}$.*

Proof. Given the collection $U = \{S_1, \ldots, S_m\}$ of the $\mathrm{Min}-p-\mathrm{Union}$, we construct a uniform metric space V of size m, where each node of V corresponds to a set S_i.

For each elements $e \in E$ of the $\mathrm{Min}-p-\mathrm{Union}$ we construct a request $R_e \subseteq V$ for the $k-\mathrm{CenterUniform}$ that is composed by the nodes corresponding to the sets S_i that containt e. Observe that due to the uniform metric and the fact that $p = \infty$, for any $V' \subseteq V$

$$\sum_{e \in E} C_{R_e}(V') = |\cup_{S_i \notin V'} S_i|$$

□

Lemma 7. *Any polynomial time c-regret algorithm for the online k-Center implies a $(c+1)$-approximation algorithm (offline) for the $k-\mathrm{CenterUniform}$.*

Proof. Let assume that that there exists a polynomial-time online learning algorithm such that for any request sequence $R_1, \ldots, R_T$,

$$\sum_{t=1}^{T} \mathbb{E}[C_{R_t}(F_t)] \leq c \min_{|F^*|=k} \sum_{t=1}^{T} \mathbb{E}[C_{R_t}(F^*)] + \Theta(\mathrm{poly}(n, D) \cdot T^{\alpha})$$

for some $\alpha < 1$.

Now let the requests lie on the uniform metric, $p = \infty$ and that the adversary at each round t selects uniformly at random one of the requests $R_1, \ldots, R_m$ that are given by the instance of $k-\mathrm{CenterUniform}$. In this case the above equation takes the following form,

$$\sum_{t=1}^{T} \frac{1}{m} \sum_{s=1}^{m} \mathbb{E}[C_{R_s}(F_t)] \leq c\, \frac{T}{m} \sum_{s=1}^{m} \mathbb{E}[C_{R_s}(\mathrm{OPT}^*)] + \Theta(n^{\beta} \cdot T^{\alpha})$$

where OPT^* is the optimal solution for the instance of $k-\mathrm{CenterUniform}$ and F_t is the random set that the online algorithm selects at round t.

Now consider the following randomized algorithm for the $k-\mathrm{CenterUniform}$.

1. Select uniformly at random a t from $\{1, \ldots, T\}$.
2. Select a set $F \subseteq V$ according to the probability distribution F_t.

The expected cost of the above algorithm, denoted by $\mathbb{E}[\mathrm{ALG}]$, is

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{m} \mathbb{E}_{F \sim F_t}\left[C_{R_i}(F)\right] &= m \cdot \left(\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{m} \frac{1}{m} \mathbb{E}_{F \sim F_t}\left[C_{R_i}(F)\right] \right) \\
&\leq \frac{c \cdot m}{T} \cdot \frac{T}{m} \sum_{i=1}^{m} C_{R_i}(\mathrm{OPT}^*) \\
&+ \Theta\left(\frac{m \cdot n^{\beta}}{T^{1-\alpha}} \right)
\end{aligned}$$

By selecting $T = \Theta(m^{\frac{1}{1-\alpha}} \cdot n^{\frac{\beta}{1-\alpha}})$ we get that $\mathbb{E}[\mathrm{ALG}] \leq (c+1) \cdot \mathrm{OPT}^*$. □

B. Omitted Proof of Section 3

B.1. Proof of Lemma 2

The Langragian of the convex program of Definition 2 is,

$$\begin{aligned} L(\beta, y, x, A, k, \lambda) &= \left(\sum_{j \in R} \beta_j^p\right)^{1/p} \\ &+ \sum_{j \in R} \lambda_j \cdot (d_{ij} x_{ij} - \beta_j) + \sum_{j \in R} A_j \cdot \left(1 - \sum_{i \in V} x_{ij}\right) \\ &+ \sum_{i \in V} \sum_{j \in V} k_{ij} \cdot (x_{ij} - y_i) - \sum_{i \in V} \sum_{j \in R} \mu_{ij} \cdot x_{ij} \end{aligned}$$

Rearranging the terms we get,

$$\begin{aligned} L(\beta, y, x, A, k, \lambda) &= \sum_{j \in R} A_j - \sum_{i \in V} \sum_{j \in R} k_{ij} \cdot y_i \\ &+ \sum_{i \in V} \sum_{j \in R} x_{ij} \cdot (k_{ij} - \mu_{ij} + d_{ij} \cdot \lambda_j - A_j) \\ &+ \left(\sum_{j \in R} \beta_j^p\right)^{1/p} - \sum_{j \in R} \lambda_j \cdot \beta_j \end{aligned}$$

In order for the function $g(A, k, \lambda) = \min_{\beta, y, x, M^+, M^-} L(\beta, y, x, A, k, \lambda)$ to get a finite value the following constraints must be satisfied,

- $k_{ij} + d_{ij} \cdot \lambda_j - A_j = \mu_{ij}$
- $||\lambda||_p^* \leq 1$ since otherwise $\left(\sum_{j \in R} \beta_j^p\right)^{1/p} - \sum_{j \in R} \lambda_j \cdot \beta_j$ can become $-\infty$.

Using the fact that the Lagragian multipliers $\mu_{ij} \geq 0$, we get the constraints of the convex program of Lemma 2. The objective comes from the fact that once $g(A, k, \lambda)$ admits a finite value then $g(A, k, \lambda) = \sum_{j \in R} A_j - \sum_{i \in V} \sum_{j \in R} k_{ij} \cdot y_i$.

B.2. Proof of Lemma 3

Let $\lambda_j^*, A_j^*, k_{ij}^*$ denote the values of the respective variables in the optimal solution of the convex program of Lemma 2 formulated with respect to the vector $y = (y_1, \ldots, y_n)$. Respectively consider $\lambda_j', A_j', k_{ij}'$ denote the values of the respective variables in the optimal solutions of the convex program of Lemma 2 formulated with respect to the vector $y' = (y_1', \ldots, y_n')$.

$$\begin{aligned} \mathrm{FC}_R(y') &= \sum_{j \in R} A_j' - \sum_{i \in V} \sum_{j \in R} k_{ij}' \cdot y_i' && (3) \\ &\geq \sum_{j \in R} A_j^* - \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot y_i' && (4) \\ &= \sum_{j \in R} A_j^* - \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot y_i' + \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot y_i - \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot y_i && (5) \\ &= \mathrm{FC}_R(y) + \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot (y_i - y_i') && (6) \end{aligned}$$

Equations 5 and 6 follow by strong duality, more precisely $\mathrm{FC}_R(y) = \sum_{j \in R} A_j^* - \sum_{i \in V} \sum_{j \in R} k_{ij}^* \cdot y_i$ since the convex program of Lemma 2 is the dual of the convex program the solution of which defines $\mathrm{FC}_R(y')$ (respectively for

$\text{FC}_R(y') = \sum_{j\in R} A'_j - \sum_{i\in V}\sum_{j\in R} k'_{ij}\cdot y'_i)$. Equation 4 is implied by the fact that the solution (λ', k', A') is optimal when the objective function is $\sum_{j\in R} A_j - \sum_{i\in V}\sum_{j\in R} k_{ij}\cdot y'_i$. Notice that the constraints of the convex program in Lemma 2 do not depend on the y-values. As a result, the solution (λ^*, k^*, A^*) (that is optimal for the dual convex program formulated for y) is feasible for the dual program formulated for the values y'. Thus Equation 4 follows by the optimality of (λ', k', A').

Up next we prove the correctness of Algorithm 1. Notice that the the solution β, x that Algorithm 1 constructs is feasible for the primal convex program of Definition 2. We will prove that the dual solution that Algorithm 1 constructs is feasible for the dual of Lemma 2 while the exact same value is obtained.

- $||\lambda||_p^* = 1$: It directly follows by the fact that $\lambda_j = \left[\frac{\beta_j}{||\beta||_p}\right]^{p-1}$ and $||\lambda||_p^* = \left[\sum_{j\in R}\lambda_j^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$.
- $d_{ij}\cdot\lambda_j + k_{ij} \geq A_j$: In case $d_{ij} < D_j^*$, Algorithm 1 implies that $x_{ij} = y_i$ and the inequality directly follows. In case $d_{ij} \leq D_j^*$ the inequality holds trivially since $k_{ij} = 0$.

Now consider the objective function,

$$\begin{aligned}
\sum_{j\in R} A_j - \sum_{i\in V}\sum_{j\in R} y_i\cdot k_{ij} &= \sum_{j\in R} A_j - \sum_{j\in R}\sum_{i\in V_j^+} y_i\cdot k_{ij} \\
&= \sum_{j\in R}\lambda_j\cdot D_j - \sum_{j\in R}\sum_{i\in V_j^+} y_i\left[\lambda_j\cdot\frac{x_{ij}}{y_i}(D_j - d_{ij})\right] \\
&= \sum_{j\in R}\lambda_j\sum_{i\in V_j^+} d_{ij}\cdot x_{ij} \\
&= \sum_{j\in R}\lambda_j\cdot\beta_j \\
&= \left(\sum_{j\in R}\beta_j^p\right)^{1/p}
\end{aligned} \tag{7}$$

where Equation 8 follows by the fact that $x_{ij} = 0$ for all $j \notin V_j^+$ and thus $\sum_{j\in V_j^+} x_{ij} = 1$. Finally notice that $|\lambda_j| \leq 1$ and thus $k_{ij} \leq D$ where D is the diameter of the metric space.

B.3. Proof of Theorem 6

By Lemma 3, $|g_i^t| = |-\sum_{j\in R^t} k_{ij}^{t*}| \leq Dr$ since $|R^t| \leq r$. Applying Theorem 1.5 of (Hazan, 2016) we get that

$$\sum_{t=1}^{T}\sum_{i\in V} g_i^t(y_i^t - y_i^*) \leq \Theta\left(kDr\sqrt{\log nT}\right)$$

Applying Lemma 3 for $y' = y^*$,

$$\sum_{t=1}^{T}\left(\text{FC}_{R_t}(y^t) - \text{FC}_{R_t}(y^*)\right) \leq \sum_{t=1}^{T}\sum_{i\in V} g_i^t(y_i^t - y_i^*) \leq \Theta\left(kDr\sqrt{\log nT}\right)$$

C. Omitted Proof of Section 4

C.1. Proof of Lemma 4

The following claim trivially follows by Step 10 of Algorithm 4.

Claim 1. *For any node $j \in V$, $d(j, F_y) \leq 6k\cdot\beta_j^*$.*

We are now ready to prove the first item of Lemma 4. Let a request $R \subseteq V$,

$$\mathrm{C}_R(F_y) = \left(\sum_{j\in R} d(j,F_y)^p\right)^{1/p} \leq \left(\sum_{j\in R}(6k)^p \cdot \beta_j^{*p}\right)^{1/p} = 6k \cdot \left(\sum_{j\in R}\beta_j^{*p}\right)^{1/p}$$

We proceed with the second item of Lemma 4. For a given node $j \in S$, let $B_j = \{i \in V : d_{ij} \leq 3k \cdot \beta_j^*\}$. It is not hard to see that for any $j \in F_y$,

$$\sum_{i\in B_j} y_i \geq 1 - \frac{1}{3k}$$

Observe that in case the latter is not true then $\sum_{i\notin B_j} x_{ij}^* \geq \frac{1}{3k}$, which would imply that $\beta_j^* > \beta_j^*$.

The second important step of the proof is that for any $j, j' \in F_y$,

$$B_j \cap B_{j'} = \emptyset.$$

Observe that in case there was $m \in B_j \cap B_{j'}$ would imply $d(j,m) \leq 3k \cdot \beta_j^*$ and $d(j',m) \leq 3k \cdot \beta_{j'}^*$. By the triangle inequality we get $d(j,j') \leq 6k \cdot \beta_{j'}^*$ (without loss of generality $\beta_j^* \leq \beta_{j'}^*$). The latter contradicts with the fact that both j and j' belong in set F_y.

Now assume that $|F_y| \geq k+1$. Then $\sum_{i\in F_y} y_i \geq |F_y| \cdot (1 - \frac{1}{3k}) \geq (k+1)\cdot(1-\frac{1}{3k}) > k$. But the latter contradicts with the fact that $\sum_{i\in V} y_i = k$. As a result, $|F_y| \leq k$.

D. Omitted Proofs of Section 5

Proof of Theorem 4. To simplify notation the quantity $\mathbb{E}_{F\sim \mathrm{CL}(y_t)}[C_{R_t}(F)]$ is denoted as $\mathbb{E}[C_{R_t}(F_t)]$. At first notice that by the first case of Lemma 5, Algorithm 5 ensures that exactly k facilities are opened at each round t.

Concerning its overall expected connection cost we get,

$$\mathbb{E}\left[C_{R_t}(F_{y_t})\right] \leq \sum_{j\in R_t} \mathbb{E}[C_{\{j\}}(F_{y_t})] \leq 4\sum_{j\in R_t} \mathrm{FC}_{\{j\}}(y_t)$$

where the fist inequality is due to the fact that $\sum_{j\in R_t} \mathrm{d}(j,F)^p \leq \left(\sum_{j\in R_t}\mathrm{d}(j,F)\right)^p$ and the second is derived by applying the second case of Lemma 5. We overall get,

$$\begin{aligned}
\sum_{t=1}^T \mathbb{E}[C_{R_t}(F_{y_t})] &\leq 4\sum_{t=1}^T\sum_{j\in R_t}\mathrm{FC}_{\{j\}}(y_t) \\
&\leq 4\sum_{t=1}^T |R_t|\cdot \mathrm{FC}_{R_t}(y_t) \\
&\leq 4r\min_{y^*}\sum_{t=1}^T \mathrm{FC}_{R_t}(y^*) \\
&+ \Theta\left(kDr\sqrt{\log nT}\right) \\
&\leq 4r\min_{|F^*|=k}\sum_{t=1}^T \mathbb{E}[\mathrm{C}_{R_t}(F^*)] \\
&+ \Theta\left(kDr\sqrt{\log nT}\right)
\end{aligned} \tag{8}$$

where inequality 3 follows by the fact that $\mathrm{FC}_{\{j\}}(y) \leq \mathrm{FC}_{\{R\}}(y)$ for all $j \in R$ and the last two inequalities follow by Theorem 6 and Lemma 1 respectively. ◻