Online Covering with Sum of $\ell_q$-Norm Objectives

Viswanath Nagarajan$^1$ and Xiangkun Shen$^2$

1 University of Michigan, Ann Arbor, MI, USA
viswa@umich.edu
2 University of Michigan, Ann Arbor, MI, USA
xkshen@umich.edu

Abstract

We consider fractional online covering problems with $\ell_q$-norm objectives. The problem of interest is of the form

$$\min \{ f(x) : Ax \preceq 1, x \geq 0 \}$$

where $f(x) = \sum_{e} c_e \| x(S_e) \|_{q_e}$ is the weighted sum of $\ell_q$-norms and $A$ is a non-negative matrix. The rows of $A$ (i.e. covering constraints) arrive online over time. We provide an online $O(\log d + \log \rho)$-competitive algorithm where $\rho = \max_{a_{ij}} \min_{a_{ij}}$ and $d$ is the maximum of the row sparsity of $A$ and $\max |S_e|$. This is based on the online primal-dual framework where we use the dual of the above convex program. Our result expands the class of convex objectives that admit good online algorithms: prior results required a monotonicity condition on the objective $f$ which is not satisfied here. This result is nearly tight even for the linear special case. As direct applications we obtain (i) improved online algorithms for non-uniform buy-at-bulk network design and (ii) the first online algorithm for throughput maximization under $\ell_p$-norm edge capacities.

1998 ACM Subject Classification F.1.2 Modes of Computation

Keywords and phrases online algorithm, covering/packing problem, convex, buy-at-bulk, throughput maximization


1 Introduction

The online primal-dual approach is a widely used approach for online problems. This involves solving a discrete optimization problem online as follows (i) formulate a linear programming relaxation and obtain a primal-dual online algorithm for it; (ii) obtain an online rounding algorithm for the resulting fractional solution. While this is similar to a linear programming (LP) based approach for offline optimization problems, a key difference is that solving the LP relaxation in the online setting is highly non-trivial. (Recall that there are general polynomial time algorithms for solving LPs offline.) So there has been a lot of effort in obtaining good online algorithms for various classes of LPs: see [1, 13, 23] for pure covering LPs, [13] for pure packing LPs and [5] for certain mixed packing/covering LPs. Such online LP solvers have been useful in obtaining online algorithms for various problems, eg. set cover [2], facility location [1], machine scheduling [5], caching [8] and buy-at-bulk network design [21].

Recently, [6] initiated a systematic study of online fractional covering and packing with convex objectives; see also the full versions [7, 11, 15]. These papers obtained good online algorithms for a large class of fractional convex covering problems. They also demonstrated the utility of this approach via many applications that could not be solved using just online LPs. However these results were limited to convex objectives $f : \mathbb{R}_+^n \to \mathbb{R}_+$ satisfying a

* A full version of the paper is available at https://arxiv.org/abs/1705.02194.
monotone gradient property, i.e. $\nabla f(z) \geq \nabla f(y)$ pointwise for all $z, y \in \mathbb{R}^n$ with $z \geq y$. There are however many natural convex functions that do not satisfy such a gradient monotonicity condition. Note that this condition requires the Hessian $\nabla^2 f(x)$ to be pointwise non-negative in addition to convexity which only requires $\nabla^2 f(x)$ to be positive semidefinite.

In this paper, we focus on convex functions $f$ that are sums of different $\ell_q$-norms. This is a canonical class of convex functions with non-monotone gradients and prior results are not applicable; see Section 1.1 for a more detailed comparison. We show that sum of $\ell_q$-norm functions admit a logarithmic competitive online algorithm. This result is nearly tight because there is a logarithmic lower bound even for online covering LPs (which corresponds to an $\ell_1$ norm objective). We also provide two applications of our result (i) improved competitive ratios (by two logarithmic factors) for some online non-uniform buy-at-bulk problems studied in [21], and (ii) the first online algorithm for throughput maximization with $\ell_p$-norm edge capacities (the competitive ratio is logarithmic which is known to be best possible even in the special case of individual edge capacities).

Given that we achieve log-competitive online algorithms for sums of $\ell_q$-norms, a natural question is whether such a result holds for all norms. Recall that any norm is a convex function. It turns out that a log-competitive algorithm is not possible for general norms. This follows from a result in [7] which shows an $\Omega(q \log d)$ lower bound for minimizing the objective $\|Bx\|_q$ under covering constraints (where $B$ is a non-negative matrix). It is still an interesting open question to identify the correct competitive ratio for general norm functions.

### 1.1 Our Results and Techniques

We consider the online covering problem

$$
\min \left\{ \sum_{e=1}^r c_e \|x(S_e)\|_{q_e} : Ax \geq 1, x \in \mathbb{R}^n_+ \right\},
$$

(1)

where each $S_e \subseteq [n] := \{1, 2, \cdots, n\}$, $q_e \geq 1$, $c_e \geq 0$ and $A$ is a non-negative $m \times n$ matrix. For any $S \subseteq [n]$ and $q \geq 1$ we use the standard notation $\|x(S)\|_q = (\sum_{i \in S} x_i^q)^{1/q}$. We also consider the dual of this convex program, which is the following packing problem:

$$
\max \left\{ \sum_{k=1}^m y_k : A^Ty = \mu, \sum_{e=1}^r \nu_e = \mu, \|\mu_e(S_e)\|_{p_e} \leq c_e \forall e \in [r], y \geq 0 \right\}.
$$

(2)

The values $p_e$ above satisfy $\frac{1}{p_e} + \frac{1}{\bar{p}_e} = 1$; so $\|\cdot\|_{\bar{p}_e}$ is the dual norm of $\|\cdot\|_{p_e}$. This dual can be derived from (1) using Lagrangian duality.

Our framework captures the classic setting of packing/covering LPs when $r = n$ and for each $e \in [n]$ we have $S_e = \{e\}$ and $q_e = 1$. Our main result is:

**Theorem 1.** There is an $O(\log d + \log \rho)$-competitive online algorithm for (1) and (2) where the covering constraints in (1) and variables $y$ in (2) arrive over time. Here $d$ is the maximum of the row-sparsity of $A$ and $\max_{e=1}^r |S_e|$ and $\rho = \frac{\max_{e=1}^r |S_e|}{\min_{e=1}^r |S_e|}$.

We note that this bound is also the best possible, even in the linear case [13]. For just the covering problem, a better $O(\log d)$ bound is known in the linear case [23] as well as for convex functions with monotone gradients [6].

The algorithm in Theorem 1 is the natural extension of the primal-dual approach for online LPs [13]. We use the gradient $\nabla f(x)$ at the current primal solution $x$ as the cost function, and use this to define a multiplicative update for the primal. Simultaneously, the
dual solution $y$ is increased additively. This algorithm is in fact identical to the one in [6] for convex functions with monotone gradients. See Algorithm 1 for the formal description. The contribution of this paper is in the analysis of this algorithm, which requires new ideas to deal with non-monotone gradients.

**Limitations of previous approaches [6]**

Recall that the general convex covering problem is

$$
\min \left\{ f(x) : Ax \geq 1, x \in \mathbb{R}^n_+ \right\},
$$

where $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ is a convex function. Its dual is:

$$
\max \left\{ m \sum_{k=1}^m y_k - f^* (\mu) : A^T y = \mu, y \geq 0 \right\},
$$

where $f^* (\mu) = \max_{x \in \mathbb{R}^n_+} \{ \mu^T x - f(x) \}$ is the Fenchel conjugate of $f$. When $f$ is the sum of $\ell_q$-norms, these primal-dual convex programs reduce to (1) and (2).

We restrict the discussion of prior techniques to functions $f$ with monotone gradients. See Algorithm 1 for the formal description. The dual solution

$$
(\bar{x}, y) = \arg\min_{x, y} f^*(y) + A^T y
$$

is the Fenchel conjugate of $f$. When $f$ is the sum of $\ell_q$-norms, these primal-dual convex programs reduce to (1) and (2).

Consider an instance with objective function $f(x) = \|x\|_2 = \sqrt{\sum_{i=1}^m x_i^2}$. There are $m = \sqrt{n}$ covering constraints, where the $k$th constraint is $\sum_{i=k(m-1)+1}^{km} x_i \geq 1$. Note that each variable appears in only one constraint. Let $P$ be the value of primal objective and $D$ be the value of dual objective at any time. Suppose that the rate of increase of the primal objective is at most $\alpha$ times that of the dual; $\alpha$ corresponds to the competitive ratio in the online primal-dual algorithm. Upon arrival of any constraint $k$, it follows from the primal updates that all the variables $\{x_i\}_{i=k(m-1)+1}^{km}$ increase from 0 to $\frac{1}{m}$. So the increase in $P$ due to constraint $k$ is $(\sqrt{k} - \sqrt{k-1}) \frac{1}{\sqrt{m}}$ for iteration $k$. This means that the increase in $D$ is at least $\frac{1}{2} (\sqrt{k} - \sqrt{k-1}) \frac{1}{\sqrt{m}}$, and so $y_k \geq \frac{1}{\alpha} (\sqrt{k} - \sqrt{k-1}) \frac{1}{\sqrt{m}}$. Finally, since $\bar{x} = \frac{m}{n} \bar{1}$, we know that $\nabla f(\bar{x}) = \frac{1}{m} \bar{1}$ (recall $n = m^2$). On the other hand, $(A^T y)_1 = y_1 \geq \frac{1}{\alpha \sqrt{m}}$. Therefore, in order to guarantee $A^T y \leq \nabla f(\bar{x})$ we must have $\alpha \geq \sqrt{m} = n^{1/4}$.

**Our approach**

First, we show that by duplicating variables and using an online separation oracle approach (as in [1]) one can ensure that the sets $\{S_e\}_{e=1}^E$ are disjoint. This allows for a simple

---

1 The result in [6] also applies to other convex functions with monotone gradients, but the competitive ratio depends exponentially on $\max_{x \in \mathbb{R}^n_+} \frac{x^T \nabla f(x)}{f(x)}$. 
Online Covering with Sum of $\ell_q$-Norm Objectives

expression for $\nabla f$ which is useful in the later analysis. Then we utilize the specific form of the primal-dual convex programs (1) and (2) and an explicit expression for $\nabla f$ to show that the dual $y$ is approximately feasible. In particular we show that $\|y^T A(S_r)\|_{p_e} \leq O(\log d \rho) \cdot c_e$ for each $e \in [r]$; here $A(S_r)$ denotes the submatrix of $A$ with columns from $S_r$. Note that this is a weaker requirement than upper bounding $A^T y$ pointwise by $\nabla f(\bar{x})$.

In order to bound $\|y^T A(S_r)\|_{p_e}$, we analyze each $e \in [r]$ separately. We partition the steps of the algorithm into phases where phase $j$ corresponds to steps where $\Phi_e = \sum_{i \in S_\tau} x_{e_i} \approx \theta^j$; here $\theta > 1$ is a parameter that depends on $q_e$. The number of phases can be bounded using the fact that $\Phi_e$ is monotonically increasing. By triangle inequality we upper bound $\|y^T A(S_r)\|_{p_e}$ by $\sum_j \|y^T_{(j)} A(S_r)\|_{p_e}$, where $y_{(j)}$ denotes the dual variables that arrive in phase $j$. And in each phase $j$, we can upper bound $\|y^T_{(j)} A(S_r)\|_{p_e}$ using the differential equations for the primal and dual updates.

Applications

We also provide two applications of Theorem 1.

Non-uniform multicommodity buy-at-bulk. This is a well-studied network design problem in the offline setting [16, 17]. For its online version, the first poly-logarithmic competitive ratio was obtained recently in [21]. A key step in this result was a fractional online algorithm for a certain mixed packing-covering LP. We improve the competitive ratio of this step from $O(\log^3 n)$ to $O(\log n)$ which leads to a corresponding improvement in the final result of [21]. See Theorem 6.

Throughput maximization with $\ell_p$-norm capacities. The online problem of maximizing throughput subject to edge capacities is well studied and a tight logarithmic competitive ratio is known [4, 13]. We consider the generalization where instead of individual edge capacities, we can have capacity constraints on subsets as follows. A $\ell_p$-norm capacity of $c$ for some subset $S$ of edges means that the $\ell_p$-norm of the loads on edges of $S$ must be at most $c$. We show that one can obtain a randomized log-competitive algorithm even in this setting, which generalizes the case with edge-capacities. See Theorem 7.

1.2 Related Work

The online primal-dual framework for linear programs [14] is fairly well understood. Tight results are known for the class of packing and covering LPs [13, 23], with competitive ratio $O(\log d)$ for covering LPs and $O(\log d \rho)$ for packing LPs; here $d$ is the row-sparsity and $\rho$ is the ratio of the maximum to minimum entries in the constraint matrix. Such LPs are very useful because they correspond to the LP relaxations of many combinatorial optimization problems. Combining the online LP solver with suitable online rounding schemes, good online algorithms have been obtained for many problems, eg. set cover [2], group Steiner tree [1], caching [8] and ad-auctions [12]. Online algorithms for LPs with mixed packing and covering constraints were obtained in [5]; the competitive ratio was improved in [6]. Such mixed packing/covering LPs were also used to obtain an online algorithm for capacitated facility location [5]. A more complex mixed packing/covering LP was used recently in [21] to obtain online algorithms for non-uniform buy-at-bulk network design: as an application of our result, we obtain a simpler and better (by two log-factors) online algorithm for this problem.
There have also been a number of results utilizing the online primal-dual framework with convex objectives for specific problems, e.g. matching [19], caching [25], energy-efficient scheduling [18, 22] and welfare maximization [10, 24]. All of these results involve separable convex/concave functions. Recently, [6] considered packing/covering problems with general (non-separable) convex objectives, but (as discussed previously) this result requires a monotone gradient assumption on the convex function. The sum of $\ell_q$-norm objectives considered in this paper does not satisfy this condition. While our primal-dual algorithm is identical to [6], we need new techniques in the analysis.

All the results above (as well as ours) involve convex objectives and linear constraints. We note that [20] obtained online primal-dual algorithms for certain semidefinite programs (i.e. involving non-linear constraints). While both our result and [20] generalize packing/covering LPs, they are not directly comparable.

We also note that online algorithms with $\ell_q$-norm objectives have been studied previously for many scheduling problems, e.g. [3, 9]. These results use different approaches and are not directly comparable to ours. More recently [7] used ideas from the online primal-dual approach in an online algorithm for unrelated machine scheduling with $\ell_p$-norm objectives as well as startup costs. However, the algorithm in [7] was tailored to their scheduling setting and we do not currently see a connection between our result and [7].

## 2 Preliminaries

Recall the primal covering problem (1) and its dual packing problem (2). In the online setting, the constraints in the primal and variables in the dual arrive over time. We need to maintain monotonically increasing primal ($x$) and dual ($y$) solutions. The following lemma shows that one can assume that the sets $\{S_e\}_{e=1}^r$ are disjoint without loss of generality. We defer the proof of the lemma to the full version. This leads to a much simpler expression for $\nabla f$ that will be used in Section 3.

**Lemma 2.** If there is a poly-time $\alpha$-competitive algorithm for instances with disjoint $S_e$, then there is a poly-time $O(1)$-competitive algorithm for general instances.

Henceforth we will assume that the sets $\{S_e\}_{e=1}^r$ are disjoint. The dual program (2) in this case reduces to:

$$\max \left\{ \sum_{k=1}^m y_k : A^T y = \mu, \|\mu(S_e)\|_{p_e} \leq c_e \forall e \in [r], y \geq 0 \right\}.$$  

(3)

It is easy to see that weak duality holds (Lemma 3). Strong duality also holds because (1) satisfies Slater’s condition; however we do not use this fact.

**Lemma 3.** For any pair of feasible solutions $x$ to (1) and $(y, \mu)$ to (3), we have

$$\sum_{e=1}^r c_e \|x(S_e)\|_{q_e} \geq \sum_{k=1}^m y_k.$$

**Proof.** This follows from the following inequalities:

$$\sum_{k=1}^m y_k = y^T 1 \leq y^T A x = \mu^T x \leq \sum_{e=1}^r \mu_e x_i \leq \sum_{e=1}^r \|\mu(S_e)\|_{p_e} \cdot \|x(S_e)\|_{q_e} \leq \sum_{e=1}^r c_e \|x(S_e)\|_{q_e}.$$

The first inequality is by primal feasibility; the second and last are by dual feasibility; the fourth is by Hölder’s inequality.
When the $k^{th}$ request $\sum_{i=1}^{n} a_{ki} x_i \geq 1$ arrives
Let $\tau$ be a continuous variable denoting the current time;
while the constraint is unsatisfied, i.e., $\sum_{i=1}^{n} a_{ki} x_i < 1$ do
    For each $i$ with $a_{ki} > 0$, increase $x_i$ at rate $\frac{\partial x_i}{\partial \tau} = \frac{a_{ki} x_i + \frac{1}{d}}{c_e x_i^{q_c - 1}} \|x(S_e)\|_{q_c}^{-1}$;
    Increase $y_k$ at rate $\frac{\partial y_k}{\partial \tau} = 1$;
    Set $\mu = A^T y$;
end

Algorithm 1: Algorithm for $\ell_q$-norm packing/covering.

### 3 Algorithm and analysis

Let $f(x) = \sum_{e=1}^{r} c_e \|x(S_e)\|_{q_e}$ denote the primal objective in (1).

In order to ensure that the gradient $\nabla f$ is positive, the primal solution $x$ starts off as $\delta \cdot 1$ where $\delta > 0$ is arbitrarily small. So we assume that the initial primal value is zero.

It is clear that the algorithm maintains a feasible and monotonically non-decreasing primal solution $x$. The dual solution $(y, \mu)$ is also monotonically non-decreasing, but not necessarily feasible. We will show that $(y, \mu)$ is $O(\log \rho d)$-approximately feasible, i.e. the packing constraints in (3) are violated by at most an $O(\log \rho d)$ factor.

#### Lemma 4. The primal objective $f(x)$ is at most twice the dual objective $\sum_{k=1}^{m} y_k$.

**Proof.** We will show that the rate of increase of the primal is at most twice that of the dual. Consider the algorithm upon the arrival of some constraint $k$. Then

$$\frac{df(x)}{d\tau} = \sum_{i:a_{ki} > 0} \nabla_i f(x) \cdot \frac{\partial x_i}{\partial \tau} = \sum_{i:a_{ki} > 0} (a_{ki} x_i + \frac{1}{d}) \leq 2.$$

The inequality comes from the fact that (i) the process for the $k$th constraint is terminated when $\sum_{i} a_{ki} x_i = 1$ and (ii) the number of non-zeroes in constraint $k$ is at most $d$. Also it is clear that the dual objective increases at rate one, which finishes the proof. 

#### Lemma 5. The dual solution $(y, \mu)$ is $O(\log \rho d)$-approximately feasible, i.e.

$$\|\mu(S_e)\|_{q_e} \leq O(\log \rho d) \cdot c_e, \quad \forall e \in [r].$$

**Proof.** Fix any $e \in [r]$. When $q_e = 1$ the analysis is simple (details are deferred to the full version). Here we only consider $q_e > 1$ which is the main part of the analysis. In order to prove the desired upper bound on $\|\mu(S_e)\|_{q_e}$ we use a potential function $\Phi = \sum_{i \in S_e} (x_i^{q_e})$. Let phase zero denote the period where $\Phi \leq \zeta := (\frac{1}{\max(a_{ij}) d})^{q_e}$, and for each $j \geq 1$, phase $j$ is the period where $\theta^{j-1} \cdot \zeta \leq \Phi < \theta \cdot \zeta$. Here $\theta > 1$ is a parameter depending on $q_e$ that will be determined later. Note that $\Phi \leq d(\frac{1}{\min(a_{ij})})^{q_e}$ as variable $x_i$ will never be increased beyond $1/\min(a_{ij})$. So the number of phases is at most $3q_e \cdot \log(d \rho)/\log \theta$. Next, we bound the increase in $\|\mu(S_e)\|_{q_e}$ for each phase separately.

For any phase, we have the following equalities

$$\frac{\partial x_i}{\partial \tau} = \frac{a_{ki} x_i + \frac{1}{d}}{c_e x_i^{q_e - 1}} \|x(S_e)\|_{q_e}^{-1}, \quad \frac{\partial y_k}{\partial \tau} = 1, \quad \frac{\partial \mu_i}{\partial \tau} = a_{ki}$$

$$\Rightarrow d \mu_i = \frac{c_e a_{ki} x_i^{q_e - 1}}{(\sum_{j \in S_e} x_j^{q_e})^{1 - \frac{1}{q_e}}} \left( a_{ki} x_i + \frac{1}{d} \right) dx_i$$

(4)
Phase zero. Suppose that each \( x_i \) increases to \( \alpha_i \) in phase zero. From (4) we have

\[
d\mu_i \leq \frac{c e a_{ki} x_i^{q_e-1}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i \quad \Rightarrow \quad \frac{1}{d c e a_{ki}} \, d\mu_i \leq \frac{x_i^{q_e-1}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i.
\]

This means that the increase \( \Delta \mu_i \) in \( \mu_i \) can be bounded as:

\[
\frac{1}{d c e a_{ki}} \Delta \mu_i \leq \int_{\delta}^{\alpha_i} \frac{x_i^{q_e-1}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i \leq \int_0^{\alpha_i} 1 \, dx_i \leq \alpha_i.
\]

Since in phase zero, \( \Phi \leq \left( \frac{1}{\max(a_{ij})} \right)^{q_e} \), we know that each \( \alpha_i \leq \frac{1}{\max(a_{ij})} \). So \( \Delta \mu_i \leq \frac{c e}{d} \)

and at the end of phase zero, we have \( \|\mu(S_e)\|_{p_e} \leq \|\mu(S_e)\|_1 \leq c e \). The last inequality is because \( d \geq \max_i |S_e| \).

Phase \( j \geq 1 \). Let \( \Phi_0 \) and \( \Phi_1 \) be the value of \( \Phi \) at the beginning and end of this phase respectively. In phase \( j \), suppose that each \( x_i \) increases from \( s_i \) to \( t_i \). Then,

\[
d\mu_i = \frac{c e a_{ki} x_i^{q_e-1}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i \leq \frac{c e x_i^{q_e-2}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i.
\]

So the increase \( \Delta \mu_i \) in \( \mu_i \) during this phase is:

\[
\Delta \mu_i \leq \int_{s_i}^{t_i} \frac{c e x_i^{q_e-2}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i.
\]

Note that variables \( x_{i'} \) for \( i' \neq i \) can also increase in this phase; so we cannot directly bound the above integral. This is precisely where the potential \( \Phi \) is useful. We know that throughout this phase, \( \sum_{i \in S_e} x_i^{p_e} \geq \Phi_0 \). So,

\[
\Delta \mu_i \leq c e \int_{s_i}^{t_i} \frac{x_i^{q_e-2}}{(\sum_{j \in S_e} x_j^{q_e})^{1-\frac{1}{q_e}}} \, dx_i = c e \frac{t_i^{q_e-1} - s_i^{q_e-1}}{(q_e - 1)q_0^{1-\frac{1}{q_e}}} = c e \frac{t_i^{q_e-1} - s_i^{q_e-1}}{(q_e - 1)q_0^{1-\frac{1}{q_e}}}.
\]

Above we used the assumption that \( q_e > 1 \) in evaluating the integral. Now,

\[
(\Delta \mu_i)^{p_e} \leq \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} \cdot (t_i^{q_e-1} - s_i^{q_e-1})^{p_e} \leq \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} \cdot (t_i^{1-\frac{p_e}{q_e}})^{p_e} \cdot (t_i^{q_e-1} - s_i^{q_e-1})
\]

\[
= \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} \cdot (t_i^{q_e-1} - s_i^{q_e-1})
\]

The first inequality above uses the fact that \( (z_1 + z_2)^{p_e} \geq z_1^{p_e} + z_2^{p_e} \) for any \( p_e > 1 \) and \( z_1, z_2 \geq 0 \), with \( z_1 = s_i^{q_e-1} \) and \( z_2 = t_i^{q_e-1} - s_i^{q_e-1} \). The last equality uses \( \frac{1}{p_e} + \frac{1}{q_e} = 1 \).

We can now bound

\[
\sum_{i \in S_e} (\Delta \mu_i)^{p_e} \leq \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} \sum_{i \in S_e} (t_i^{q_e-1} - s_i^{q_e-1}) = \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} (\Phi_1 - \Phi_0) \leq \frac{c e^{p_e}}{(q_e - 1)^{p_e} \Phi_0} (\theta - 1).
\]

Let \( \mu_j \) denote the increase in \( \mu(S_e) \) during phase \( j \). It follows from the above that \( \|\mu_j\|_{p_e} \leq \frac{c e}{q_e-1} (\theta - 1)^{1/p_e} \).
Combining across phases. Note that \( \mu = \sum_{j \geq 0} \mu_j \). By triangle inequality, we have

\[
\|\mu\|_{p_e} \leq \sum_{j \geq 0} \|\mu_j\|_{p_e} \leq \epsilon_c + \sum_{j \geq 1} \|\mu_j\|_{p_e} \leq \epsilon_c \left( 1 + \frac{3\epsilon_c (\theta - 1)^{1/p_e}}{(q_e - 1) \log \theta \cdot \log(d\rho)} \right).
\]

(5)

To complete the proof we show next that for any \( q_e > 1 \), there is some choice of \( \theta > 1 \) such that the right-hand-side above is \( O(\log(d\rho)) \cdot \epsilon_c \).

- If \( q_e \geq 2 \) then setting \( \theta = 2 \), we have \( \frac{3\epsilon_c}{(q_e - 1)}(\theta - 1)^{1/p_e} / \log \theta \leq 6 \).
- If \( 1 < q_e < 2 \) then set \( \theta = 1 + (q_e - 1)^{-\epsilon/p_e} \), where \( \epsilon = \frac{1}{\log(q_e - 1)} > 0 \). We have

\[
\frac{(\theta - 1)^{1/p_e}}{\log \theta} \leq \frac{(q_e - 1)^{-\epsilon}}{\log(q_e - 1)^{-\epsilon/p_e}} = \frac{(q_e - 1)^{-\epsilon}}{-\epsilon/p_e \cdot \log(q_e - 1)} = \frac{(q_e - 1)^{-\epsilon}}{p_e} = 2.
\]

The first inequality above uses that \( \theta - 1 = (q_e - 1)^{-\epsilon/p_e} > 1 \). Thus we have

\[
\frac{3\epsilon_c (\theta - 1)^{1/p_e}}{(q_e - 1) \log \theta} \leq \frac{6\epsilon_c}{(q_e - 1)p_e} = 6,
\]

where the last equality uses \( \frac{1}{p_e} + \frac{1}{q_e} = 1 \).

So in either case we have that the right-hand-side of (5) is at most \( (1 + 6 \log(d\rho)) \cdot \epsilon_c \). ▶

4 Applications

4.1 Online Buy-at-Bulk Network Design

In the non-uniform buy-at-bulk problem, we are given a directed graph \( G = (V,E) \) with a monotone subadditive cost function \( g_e : \mathbb{R}_+ \to \mathbb{R}_+ \) on each edge \( e \in E \) and a collection \( \{(s_i,t_i)\}_{i=1}^m \) of \( m \) source/destination pairs. The goal is to find an \( s_i - t_i \) path \( P_i \) for each \( i \in [m] \) such that the objective \( \sum_{e \in E} g_e(\text{load}_e) \) is minimized; here \( \text{load}_e \) is the number of paths using \( e \). An equivalent view of this problem involves two costs \( c_e \) and \( \ell_e \) for each edge \( e \in E \) and the objective \( \sum_{e \in P_i} c_e + \sum_{e \in E} \ell_e \cdot \text{load}_e \). In the online setting, the pairs \( (s_i,t_i) \) arrive over time and we need to decide on the path \( P_i \) immediately after the \( i^{th} \) pair arrives. Recently, [21] gave a modular online algorithm for non-uniform buy-at-bulk with competitive ratio \( O(\alpha \beta \gamma \cdot \log^3 n) \) where:

- \( \alpha \) is the “junction tree” approximation ratio,
- \( \beta \) is the integrality gap of the natural LP for single-sink instances,
- \( \gamma \) is the competitive ratio of an online algorithm for single-sink instances.

See [21] for more details. One of the main components in this result was an \( O(\log^3 n) \)-competitive fractional online algorithm for a certain mixed packing/covering LP. Here we show that Theorem 1 can be used to provide a better (and tight) \( O(\log n) \)-competitive ratio. This leads to the following improvement:

▶ Theorem 6. There is an \( O(\alpha \beta \gamma \cdot \log^3 n) \)-competitive ratio for non-uniform buy-at-bulk, where \( \alpha, \beta, \gamma \) are as above.
The LP relaxation. Let $\mathcal{T} = \{s_i, t_i : i \in [m]\}$ denote the set of all sources/destinations. For each $i \in [m]$ and root $r \in V$ variable $z_{ir}$ denotes the extent to which both $s_i$ and $t_i$ route to/from $r$. For each $r \in V$ and $e \in E$, variable $x_{e,r}$ denotes the extent to which edge $e$ is used in the routing to root $r$. For each $r \in V$ and $u \in \mathcal{T}$, variables $\{f_{r,u,e} : e \in E\}$ represent a flow between $r$ and $u$. \cite{21} relied on solving the following LP:

$$\begin{align*}
\min & \sum_{r \in V} \sum_{e \in E} c_e \cdot x_{e,r} + \sum_{r \in V} \sum_{e \in E} \ell_e \cdot \sum_{u \in \mathcal{T}} f_{r,u,e} \\
\text{s.t.} & \sum_{r \in V} z_{ir} \geq 1, \quad \forall i \in [m] \\
& \{f_{r,s_i,e} : e \in E\} \text{ is a flow from } s_i \text{ to } r \text{ of } z_{ir} \text{ units}, \quad \forall r \in V, i \in [m] \\
& \{f_{r,t_i,e} : e \in E\} \text{ is a flow from } r \text{ to } t_i \text{ of } z_{ir} \text{ units}, \quad \forall r \in V, i \in [m] \\
& f_{r,u,e} \leq x_{e,r}, \quad \forall u \in \mathcal{T}, e \in E \\
& x, f, z \geq 0
\end{align*}$$

The online algorithm in \cite{21} for this LP has competitive ratio $O(D \cdot \log n)$ w.r.t. the optimal integral solution; here $D$ is an upper bound on the length of any $s_i - t_i$ path (note that $D$ can be as large as $n$). Using a height reduction operation, they could ensure that $D = O(\log n)$ while incurring an additional $O(\log n)$ factor loss in the objective. This lead to the $O(\log^3 n)$ factor for the fractional online algorithm. Here we provide an improved $O(\log n)$-competitive algorithm for this LP which does not require any bound on the path-lengths.

For any $r \in V$ and $u \in \mathcal{T}$, let $MC(r, u)$ denote the $u - r$ (resp. $r - u$) minimum cut in the graph with edge capacities $\{f_{r,u,e} : e \in E\}$ if $u$ is a source (resp. destination). By the max-flow min-cut theorem, it follows that $z_{ir} \leq \min\{MC(r, s_i), MC(r, t_i)\}$. Using this, we can combine the first three constraints of the above LP into the following:

$$\sum_{r \in V} \min\{MC(r, s_i), MC(r, t_i)\} \geq 1, \quad \forall i \in [m].$$

For a fixed $i \in [m]$, this constraint is equivalent to the following. For each $r \in V$, pick either an $s_i - r$ cut (under capacities $f_{r,s_i,*}$) or an $r - t_i$ cut (under capacities $f_{r,t_i,*}$), and check if the total cost of these cuts is at least 1. This leads to the following reformulation that eliminates the $x$ and $z$ variables.

$$\begin{align*}
\min & \sum_{r \in V} \sum_{e \in E} c_e \cdot \left( \max_{u \in \mathcal{T}} f_{r,u,e} \right) + \sum_{r \in V} \sum_{e \in E} \ell_e \cdot \sum_{u \in \mathcal{T}} f_{r,u,e} \\
\text{s.t.} & \sum_{r \in R_s} f_{r,s_i}(S_r) + \sum_{r \in R_t} f_{r,t_i}(T_r) \geq 1, \quad \forall i \in [m], \forall (R_s, R_t) \text{ partition of } V, \\
& \forall S_r : s_i - r \text{ cut, } \forall r \in R_s, \forall T_r : r - t_i \text{ cut, } \forall r \in R_t \\
& f \geq 0.
\end{align*}$$

Note that $\ell_{\log(n)}$-norm is a constant approximation for $\ell_\infty$. Therefore we can reformulate the above objective function (at the loss of a constant factor) as the sum of $\ell_{\log(n)}$ and $\ell_1$ norms. Our fractional solver applies to this convex covering problem, and yields an $O(\log n)$-competitive ratio (note that $\rho = 1$ for this instance). In order to get a polynomal running time, we can use the natural “separation oracle” approach to produce violated covering constraints.

Each iteration above runs in polynomial time since the minimum cuts can be computed in polynomial time. In order to bound the number of iterations, consider the potential $\psi = \sum_{e \in E} (f_{r,s_i,e} + f_{r,t_i,e})$. Note that $0 \leq \psi \leq 2|E|$ and each iteration increases $\psi$ by at least $\frac{1}{2}$. So the number of iterations is at most $4|E|$.  

ICALP 2017
When the $i^{th}$ request $(s_i, t_i)$ arrives
repeat
    For each $r \in V$, compute $MC(r, s_i)$ and $MC(r, t_i)$ and the respective cuts $S_r$ and $T_r$;
    Let $R_s = \{ r \in V : MC(r, s_i) \leq MC(r, t_i) \}$ and $R_t = V \setminus R_s$;
    Run Algorithm 1 with constraint $\sum_{r \in R_s} f_{r, s_i}(S_r) + \sum_{r \in R_t} f_{r, t_i}(T_r) \geq 1$;
until $\sum_{r \in V} \min \{MC(r, s_i), MC(r, t_i)\} \geq \frac{1}{2}$.

Algorithm 2: Separation Oracle Based Algorithm for Buy-at-Bulk.

4.2 Throughput Maximization with $\ell_p$-norm Capacities

The online problem of maximizing multicommodity flow was studied in [4, 13]. In this problem, we are given a directed graph with edge capacities $u(e)$. Requests $(s_i, t_i)$ arrive in an online fashion. The algorithm should choose a path between $s_i$ and $t_i$ and allocate a bandwidth of 1 on the path to serve request $i$. The total bandwidth allocated on any edge is not allowed to exceed its capacity. This is the simplest version of the multicommodity routing problem. Here we consider an extension with $\ell_p$-norm capacity constraints on subsets of edges. This can be used to model situations where edges are provided by multiple agents. Each agent $j$ owns a subset $S_j$ of edges and it requires the $\ell_p$-norm of the bandwidths of these edges to be at most $c_j$. In this section we assume the $S_j$ are disjoint. Our result also applies to general $S_j$ via a reduction to disjoint instances.

\textbf{Theorem 7.} Assume that $c_j = \Omega(\log m) \cdot |S_j|^{1/p}$ for each $j$. Then there is a randomized $O(\log m)$-competitive online algorithm for throughput maximization with $\ell_p$-norm capacities, where $m$ is the number of edges in the graph.

We note that a similar “high capacity” assumption is also needed in the linear special case [4, 13] where each $|S_j| = 1$.

In a fractional version of the problem, a request can be satisfied by several paths and the allocation of bandwidth can be in range $[0, 1]$ instead of being restricted from $\{0, 1\}$. For request $(s_i, t_i)$, let $P_i$ be the set of simple paths between $s_i$ and $t_i$. Variable $f_{i, P}$ is defined to be the amount of flow on the path $P$ for request $(s_i, t_i)$. The total profit of algorithm is the (fractional) number of requests served and the performance is measured with respect to the maximum number of requests that could be served if the requests are known beforehand. We describe the problem as a packing problem:

$$\max \sum_i \sum_{P \in P_i} f_{i, P} \quad (6)$$

s.t. $\sum_{P \in P_i} f_{i, P} \leq 1, \quad \forall i$ \hspace{1cm} (7)

$$\sum_i \sum_{P \in P, e \in P} f_{i, P} = u_e, \quad \forall e$$ \hspace{1cm} (8)

$$\|\mu(S_j)\|_{\ell_p} \leq c_j, \quad \forall j$$ \hspace{1cm} (9)

$f \in \mathbb{R}_+^{P_i \times P}$
Note that single edge capacity is a special case of (9) with \(|S_j| = 1\) and any \(p_j\). The corresponding primal problem is the following.

\[
\begin{align*}
\min & \quad \sum_j c_j \|x(S_j)\|_{q_j} + \sum_i z_i \\
\text{s.t.} & \quad z_i + \sum_{e \in P} x_e \geq 1, \quad \forall i, P \in \mathcal{P}_i \\
& \quad x, z \in \mathbb{R}^+_n
\end{align*}
\]

where \(z\) is the dual variable of constraints (7) and \(x\) is the dual variable of constraints (8) to (9). This formulation falls in the form of (1). Combining our algorithm with online rounding techniques, we can prove Theorem 7. A formal proof appears in the full version.

References

Online Covering with Sum of $\ell_q$-Norm Objectives