

Appendix

A1 Two Algorithms in Literature

The two algorithm discussed in Section 5.1 are Algorithm 2 and Algorithm 3.

Algorithm 2 Dynamic Learning Algorithm (first proposed in [3] and then refined in [20])

- 1: Input: d
- 2: Let $p_1 = \mathbf{0}$
- 3: **for** $t = 1, 2, \dots, n$ **do**
- 4:

$$x_t = \begin{cases} 1, & \text{if } r_t > \mathbf{a}_t^\top \mathbf{p}_t \\ 0, & \text{if } r_t \leq \mathbf{a}_t^\top \mathbf{p}_t \end{cases}$$

- 5: The scaled primal LP is

$$\begin{aligned} \max \quad & \sum_{j=1}^t r_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^t a_{ij} x_j \leq t d_i, \quad i = 1, \dots, m \\ & 0 \leq x_j \leq 1, \quad j = 1, \dots, t \end{aligned}$$

- 6: Solve its dual problem and obtain the optimal dual variable \mathbf{p}_{t+1}

$$\mathbf{p}_{t+1} = \arg \min_{\mathbf{p} \geq 0} \sum_{i=1}^m d_i p_i + \frac{1}{t} \sum_{j=1}^t \left(r_j - \sum_{i=1}^m a_{ij} p_i \right)^+$$

- 7: **end for**
-

Algorithm 3 Primal-beats-dual Algorithm [18]

- 1: Input: d
- 2: Let $p_1 = \mathbf{0}$
- 3: **for** $t = 1, 2, \dots, n$ **do**
- 4: Solve the scaled LP

$$\begin{aligned} \max \quad & \sum_{j=1}^t r_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^t a_{ij} x_j \leq t d_i, \quad i = 1, \dots, m \\ & 0 \leq x_j \leq 1, \quad j = 1, \dots, t \end{aligned}$$

- 5: Denote the optimal solution as $\tilde{\mathbf{x}}^{(t)} = (\tilde{x}_1^{(t)}, \dots, \tilde{x}_t^{(t)})$

- 6:

$$x_t = \begin{cases} 1, & \text{with probability } \tilde{x}_t^{(t)} \\ 0, & \text{with probability } 1 - \tilde{x}_t^{(t)} \end{cases}$$

- 7: **end for**
-

B1 Concentration Inequalities under Random Permutation

Lemma 4. Let U_1, \dots, U_n be a random sample without replacement from the real numbers $\{c_1, \dots, c_N\}$. Then for every $s > 0$,

$$\mathbb{P}(|\bar{U}_n - \bar{c}_N| \geq s) \leq \begin{cases} 2 \exp\left(-\frac{2ns^2}{\Delta_N^2}\right) & \text{(Hoeffding),} \\ 2 \exp\left(-\frac{2ns^2}{(1-(n-1)/N)\Delta_N^2}\right) & \text{(Serfling),} \\ 2 \exp\left(-\frac{ns^2}{2\sigma_N^2 + s\Delta_N}\right) & \text{(Hoeffding-Bernstein),} \\ 2 \exp\left(-\frac{ns^2}{m\sigma_N^2}\right) & \text{if } N = mn \text{ (Massart),} \end{cases}$$

where $\bar{c}_N = \frac{1}{N} \sum_{i=1}^N c_i$, $\sigma_N^2 = \frac{1}{N} \sum_{i=1}^N (c_i - \bar{c}_N)^2$ and $\Delta_N = \max_{1 \leq i \leq N} c_i - \min_{1 \leq i \leq N} c_i$.

Proof. See Theorem 2.14.19 in [28]. □

B2 Proof of Lemma 1

Proof. By taking expectation with respect to the elements in (5), the function $f_n(\mathbf{p})$ in (5) can be viewed as a *sample average approximation* of the following stochastic program

$$\begin{aligned} \min_{\mathbf{p}} f(\mathbf{p}) &= \mathbf{d}^\top \mathbf{p} + \mathbb{E}_{(r, \mathbf{a}) \sim \mathcal{P}} \left[\left(r - \mathbf{a}^\top \mathbf{p} \right)^+ \right] & \text{(SP)} \\ \text{s.t. } & \mathbf{p} \geq \mathbf{0}. \end{aligned}$$

Denote the optimal solution to (SP) as \mathbf{p}^* . For \mathbf{p}^* , we have

$$\underline{d} \|\mathbf{p}^*\|_1 \leq \mathbf{d}^\top \mathbf{p}^* \stackrel{(a)}{\leq} \mathbb{E}r \leq \bar{r},$$

where inequality (a) is due to that if otherwise, \mathbf{p}^* cannot be the optimal solution because it will give a larger objective value of $f(\mathbf{p})$ than setting $\mathbf{p} = \mathbf{0}$. Given the non-negativeness of \mathbf{p}^* , we have $\|\mathbf{p}^*\|_2 \leq \|\mathbf{p}^*\|_1$. So we obtain the first inequality in the lemma.

For \mathbf{p}_t specified by Algorithm 1, we have,

$$\begin{aligned} \|\mathbf{p}_{t+1}\|_2^2 &\leq \|\mathbf{p}_t + \gamma_t (\mathbf{a}_t x_t - \mathbf{d})\|_2^2 \\ &= \|\mathbf{p}_t\|_2^2 + \gamma_t^2 \|\mathbf{a}_t x_t - \mathbf{d}\|_2^2 + 2\gamma_t (\mathbf{a}_t x_t - \mathbf{d})^\top \mathbf{p}_t \\ &\leq \|\mathbf{p}_t\|_2^2 + \gamma_t^2 m(\bar{a} + \bar{d})^2 + 2\gamma_t \mathbf{a}_t^\top \mathbf{p}_t x_t - 2\gamma_t \mathbf{d}^\top \mathbf{p}_t \end{aligned}$$

where the first inequality comes from the projection (into the non-negative orthant) step in the algorithm. Note that

$$\mathbf{a}_t^\top \mathbf{p}_t x_t = \mathbf{a}_t^\top \mathbf{p}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \leq r_t \leq \bar{r}.$$

Therefore,

$$\|\mathbf{p}_{t+1}\|_2^2 \leq \|\mathbf{p}_t\|_2^2 + \gamma_t^2 m(\bar{a} + \bar{d})^2 + 2\gamma_t \bar{r} - 2\gamma_t \mathbf{d}^\top \mathbf{p}_t,$$

and it holds with probability 1.

Next, we establish that when $\|\mathbf{p}_t\|_2$ is large enough, then it must hold that $\|\mathbf{p}_{t+1}\|_2 \leq \|\mathbf{p}_t\|_2$. Specifically, when $\|\mathbf{p}_t\|_2 \geq \frac{2\bar{r} + m(\bar{a} + \bar{d})^2}{\underline{d}}$, we have

$$\begin{aligned} \|\mathbf{p}_{t+1}\|_2^2 - \|\mathbf{p}_t\|_2^2 &\leq \gamma_t^2 m(\bar{a} + \bar{d})^2 + 2\gamma_t \bar{r} - 2\gamma_t \mathbf{d}^\top \mathbf{p}_t \\ &\leq \gamma_t^2 m(\bar{a} + \bar{d})^2 + 2\gamma_t \bar{r} - 2\gamma_t \underline{d} \|\mathbf{p}_t\|_1 \\ &\leq \gamma_t^2 m(\bar{a} + \bar{d})^2 + 2\gamma_t \bar{r} - 2\gamma_t \underline{d} \|\mathbf{p}_t\|_2 \\ &\leq 0 \end{aligned}$$

when $\gamma_t \leq 1$. On the other hand, when $\|\mathbf{p}_t\|_2 \leq \frac{2\bar{r} + m(\bar{a} + \bar{d})^2}{\underline{d}}$,

$$\begin{aligned} \|\mathbf{p}_{t+1}\|_2 &\leq \|\mathbf{p}_t + \gamma_t (\mathbf{a}_t x_t - \mathbf{d})\|_2 \\ &\stackrel{(b)}{\leq} \|\mathbf{p}_t\|_2 + \gamma_t \|\mathbf{a}_t x_t - \mathbf{d}\|_2 \\ &\leq \frac{2\bar{r} + m(\bar{a} + \bar{d})^2}{\underline{d}} + m(\bar{a} + \bar{d}) \end{aligned}$$

where (b) comes from the triangle inequality of the norm.

Combining these two scenarios with the fact that $\mathbf{p}_1 = \mathbf{0}$, we have

$$\|\mathbf{p}_t\|_2 \leq \frac{2\bar{r} + m(\bar{a} + \bar{d})^2}{\underline{d}} + m(\bar{a} + \bar{d})$$

for $t = 1, \dots, n$ with probability 1. □

B3 Proof of Theorem 1

Proof. First, the primal optimal objective value is no greater than the dual objective with $\mathbf{p} = \mathbf{p}^*$. Specifically,

$$\begin{aligned} R_n^* &= \text{P-LP} = \text{D-LP} \\ &\leq n\mathbf{d}^\top \mathbf{p}^* + \sum_{j=1}^n (r_j - \mathbf{a}_j^\top \mathbf{p}^*)^+. \end{aligned}$$

Taking expectation on both sides,

$$\begin{aligned} \mathbb{E}[R_n^*] &\leq \mathbb{E}\left[n\mathbf{d}^\top \mathbf{p}^* + \sum_{t=1}^n (r_t - \mathbf{a}_t^\top \mathbf{p}^*)^+\right] \\ &\leq nf(\mathbf{p}^*). \end{aligned}$$

Thus, the optimal objective value of the stochastic program (by a factor of n) is an upper bound for the expected value of the primal optimal objective. Hence

$$\begin{aligned} \mathbb{E}[R_n^* - R_n] &\leq nf(\mathbf{p}^*) - \sum_{j=1}^n \mathbb{E}\left[r_j I(r_j > \mathbf{a}_j^\top \mathbf{p}_t)\right] \\ &\leq \sum_{t=1}^n \mathbb{E}[f(\mathbf{p}_t)] - \sum_{t=1}^n \mathbb{E}\left[r_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right] \\ &\leq \sum_{t=1}^n \mathbb{E}\left[\mathbf{d}^\top \mathbf{p}_t + (r_t - \mathbf{a}_t^\top \mathbf{p}_t)^+ - r_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right] \\ &= \sum_{t=1}^n \mathbb{E}\left[\left(\mathbf{d}^\top - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right)^\top \mathbf{p}_t\right]. \end{aligned}$$

where the expectation is taken with respect to (r_t, \mathbf{a}_t) 's. In above, the second line comes from the optimality of \mathbf{p}^* , while the third line is valid because of the independence between \mathbf{p}_t and (r_t, \mathbf{a}_t) .

The importance of the above inequality lies in that it relates and represents the primal optimality gap in the dual prices \mathbf{p}_t – which is the core of Algorithm 1. From the updating formula in Algorithm 1, we know

$$\begin{aligned} \|\mathbf{p}_{t+1}\|_2^2 &\leq \|\mathbf{p}_t\|_2^2 - \frac{2}{\sqrt{n}} \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right)^\top \mathbf{p}_t + \frac{1}{n} \left\|\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right\|_2^2 \\ &\leq \|\mathbf{p}_t\|_2^2 - \frac{2}{\sqrt{n}} \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right)^\top \mathbf{p}_t + \frac{m(\bar{a} + \bar{d})^2}{n}. \end{aligned}$$

Moving the cross-term to the right-hand-side, we have

$$\begin{aligned} \sum_{t=1}^n \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t)\right)^\top \mathbf{p}_t &\leq \sum_{t=1}^n \left(\sqrt{n}\|\mathbf{p}_t\|_2^2 - \sqrt{n}\|\mathbf{p}_{t+1}\|_2^2 + \frac{m(\bar{a} + \bar{d})^2}{\sqrt{n}}\right) \\ &\leq m(\bar{a} + \bar{d})^2 \sqrt{n}. \end{aligned}$$

Consequently,

$$\mathbb{E}[R_n^* - R_n] \leq m(\bar{a} + \bar{d})^2 \sqrt{n}$$

hold for all n and distribution $\mathcal{P} \in \Xi$.

For the constraint violation, if we revisit the updating formula, we have

$$\mathbf{p}_{t+1} \geq \mathbf{p}_t + \frac{1}{\sqrt{n}} (\mathbf{a}_t x_t - \mathbf{d})$$

where the inequality is elementwise. Therefore,

$$\begin{aligned} \sum_{t=1}^n \mathbf{a}_t x_t &\leq n\mathbf{d} + \sum_{t=1}^n \sqrt{n}(\mathbf{p}_{t+1} - \mathbf{p}_t) \\ &\leq \mathbf{b} + \sqrt{n}\mathbf{p}_{n+1} \end{aligned}$$

In the second line, we remove the term involve \mathbf{p}_1 with the algorithm specifying $\mathbf{p}_1 = \mathbf{0}$. Then with Lemma 1, we have

$$\mathbb{E}[v(\mathbf{x})] = \mathbb{E}[\|(\mathbf{A}\mathbf{x} - \mathbf{b})^+\|_2] \leq \sqrt{n}\mathbb{E}\|\mathbf{p}_{n+1}\|_2 \leq \left(\frac{2\bar{r} + m(\bar{a} + \bar{d})^2}{\underline{d}} + m(\bar{a} + \bar{d}) \right) \sqrt{n}.$$

□

B4 Proof of Proposition 1

Proof. Define $\text{SLP}(s, \mathbf{b}_0)$ as the following LP

$$\begin{aligned} \max \quad & \sum_{j=1}^s r_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^s a_{ij} x_j \leq \frac{s b_i}{n} + b_{0i} \\ & 0 \leq x_j \leq 1 \text{ for } j = 1, \dots, s. \end{aligned}$$

where $\mathbf{b}_0 = (b_{01}, \dots, b_{0m})$ denotes the constraint relaxation quantity for the scaled LP. Denote the optimal objective value of $\text{SLP}(s, \mathbf{b}_0)$ as $R^*(s, \mathbf{b}_0)$. Also, denote $\mathbf{x}(\mathbf{p}) = (x_1(\mathbf{p}), \dots, x_n(\mathbf{p}))$ and $x_j(\mathbf{p}) = I(r_j > \mathbf{a}_j^\top \mathbf{p})$. It denotes the decision variables we obtain with a dual price \mathbf{p} .

We prove the following three results:

(i) The following bounds hold for R_n^* ,

$$\sum_{j=1}^n r_j x_j(\mathbf{p}_n^*) \leq R_n^* \leq \sum_{j=1}^n r_j x_j(\mathbf{p}_n^*) + m\bar{r}.$$

(ii) When $s \geq \max\{16\bar{a}^2, \exp\{16\bar{a}^2\}, e\}$, then the optimal dual solution \mathbf{p}_n^* is a feasible solution to $\text{SLP}\left(s, \frac{\log s}{\sqrt{s}} \mathbf{1}\right)$ with probability no less than $1 - \frac{m}{s}$.

(iii) The following inequality holds for the optimal objective values to the scaled LP and its relaxation

$$R_s^* \geq R^*\left(s, \frac{\log s}{\sqrt{s}} \mathbf{1}\right) - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}}.$$

For part (i), this inequality replace the optimal value with bounds containing the objective values obtained by adopting optimal dual solution. The left hand side of the inequality comes from the complementarity condition while the right hand side can be shown from Lemma 2.

For part (ii), the motivation to introduce a relaxed form of the scaled LP is to ensure the feasibility of \mathbf{p}_n^* . The key idea for the proof is to utilize the feasibility of \mathbf{p}_n^* for (2). To see that, let $\alpha_{ij} = a_{ij} I(r_j > \mathbf{a}_j^\top \mathbf{p}_n^*)$ and

$$\begin{aligned} c_\alpha &= \max_{i,j} \alpha_{ij} - \min_{i,j} \alpha_{ij} \leq 2\bar{a}, \\ \bar{\alpha}_i &= \frac{1}{n} \sum_{j=1}^n \alpha_{ij} = \frac{1}{n} \sum_{j=1}^n a_{ij} x_j(\mathbf{p}_n^*) \leq d_i, \\ \sigma_i^2 &= \frac{1}{n} \sum_{j=1}^n (\alpha_{ij} - \bar{\alpha}_i)^2 \leq 4\bar{a}^2. \end{aligned} \tag{7}$$

Here the first and third inequality comes from the bounds on a_{ij} 's while the second one comes from the feasibility of the optimal solution for (2).

Then, when $k > \max\{16\bar{a}^2, \exp\{16\bar{a}^2\}, e\}$, by applying Hoeffding-Bernstein's Inequality

$$\begin{aligned} \mathbb{P}\left(\sum_{j=1}^k \alpha_{ij} - kd_i \geq \sqrt{k} \log k\right) &\stackrel{(e)}{\leq} \mathbb{P}\left(\sum_{j=1}^k \alpha_{ij} - k\bar{\alpha}_i \geq \sqrt{k} \log k\right) \\ &\stackrel{(f)}{\leq} \exp\left(-\frac{k \log^2 k}{8k\bar{a}^2 + 2\bar{a}\sqrt{k} \log k}\right) \\ &\stackrel{(g)}{\leq} \frac{1}{k} \end{aligned}$$

for $i = 1, \dots, m$. Here inequality (e) comes from (7), (f) comes from applying Lemma 4, and (g) holds when $s > \max\{16\bar{a}^2, \exp\{16\bar{a}^2\}, e\}$.

Let event

$$E_i = \left\{ \sum_{j=1}^s \alpha_{ij} - sd_i < \sqrt{s} \log s \right\}$$

and $E = \bigcap_{i=1}^m E_i$. The above derivation tells $\mathbb{P}(E_i) \geq 1 - \frac{1}{s}$. By applying union bound, we obtain $\mathbb{P}(E) \geq 1 - \frac{m}{s}$ and it completes the proof of part (ii).

For part(iii), denote the optimal solution to $\text{SLP}\left(s, \frac{\log s}{\sqrt{s}} \mathbf{1}\right)$ as $\tilde{\mathbf{p}}_s$.

$$\begin{aligned} R^* \left(s, \frac{\log s}{\sqrt{s}} \mathbf{1} \right) &= s \left(\mathbf{d} + \frac{\log s}{\sqrt{s}} \mathbf{1} \right)^\top \tilde{\mathbf{p}}_s^* + \sum_{j=1}^s (r_j - \mathbf{a}_j^\top \tilde{\mathbf{p}}_s^*)^+ \\ &\leq s \left(\mathbf{d} + \frac{\log s}{\sqrt{s}} \mathbf{1} \right)^\top \mathbf{p}_s^* + \sum_{j=1}^s (r_j - \mathbf{a}_j^\top \mathbf{p}_s^*)^+ \\ &\leq \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} + R_s^*. \end{aligned}$$

where the first inequality comes from dual optimality of $\tilde{\mathbf{p}}_s^*$ and the second inequality comes from the upper bound of $\|\mathbf{p}_s^*\|$ and the duality of the scaled LP R_s^* . Therefore,

$$R_s^* \geq R^* \left(s, \frac{\log s}{\sqrt{s}} \mathbf{1} \right) - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}}.$$

Finally, we complete the proof with the help of the above three results.

$$\begin{aligned} \frac{1}{s} \mathbb{E} [\mathbb{I}_E R_s^*] &\geq \frac{1}{s} \mathbb{E} \left[\mathbb{I}_E R^* \left(s, \frac{\log s}{\sqrt{s}} \mathbf{1} \right) \right] - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} \\ &\geq \frac{1}{s} \mathbb{E} \left[\mathbb{I}_E \sum_{j=1}^s r_j x_j(\mathbf{p}^*) \right] - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} \end{aligned}$$

where \mathbb{I}_E denotes an indicator function for event E . The first line comes from applying part (iii) while the second line comes from the feasibility of \mathbf{p}^* on event E . Then,

$$\begin{aligned} \frac{1}{s} \mathbb{E} [R_s^*] &\geq \frac{1}{s} \mathbb{E} \left[\sum_{j=1}^s r_j x_j(\mathbf{p}^*) \right] - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} - \frac{m\bar{r}}{s} \\ &= \frac{1}{n} \mathbb{E} \left[\sum_{j=1}^n r_j x_j(\mathbf{p}^*) \right] - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} - \frac{m\bar{r}}{s} \\ &\geq \frac{1}{n} R_n^* - \frac{\bar{r}\sqrt{s} \log s}{\underline{d}} - \frac{m\bar{r}}{s} - \frac{m\bar{r}}{n} \end{aligned}$$

where the first line comes from part (ii) – the probability bound on event E , the second line comes from the symmetry of the random permutation probability space, and the third line comes from part (i). We complete the proof. \square

B5 Proof of Theorem 2

Proof. For the regret bound,

$$R_n^* - \mathbb{E} [R_n] = R_n^* - \sum_{t=1}^n \mathbb{E} [r_t x_t]$$

where x_t 's are specified according to Algorithm 1. Then

$$\begin{aligned} R_n^* - \mathbb{E}[R_n] &= R_n^* - \sum_{t=1}^n \frac{1}{t} \mathbb{E}[R_t^*] + \sum_{t=1}^n \frac{1}{t} \mathbb{E}[R_t^*] - \sum_{t=1}^n \mathbb{E}[r_t x_t] \\ &= \sum_{t=1}^n \left(\frac{1}{n} R_n^* - \frac{1}{t} \mathbb{E}[R_t^*] \right) + \sum_{t=1}^n \mathbb{E} \left[\frac{1}{n+1-t} \tilde{R}_{n-t+1}^* - r_t x_t \right] \end{aligned} \quad (8)$$

where \tilde{R}_{n-t+1}^* is defined as the optimal value of the following LP

$$\begin{aligned} \max \quad & \sum_{j=t}^n r_j x_j \\ \text{s.t.} \quad & \sum_{j=t}^n a_{ij} x_j \leq \frac{(n-t+1)b_i}{n} \\ & 0 \leq x_j \leq 1 \text{ for } j = 1, \dots, m. \end{aligned}$$

For the first part of (8), we can apply Proposition 1. Meanwhile, the analyses of the second part takes a similar form as the previous stochastic input model. Specifically,

$$\mathbb{E} \left[\frac{1}{n+1-t} \tilde{R}_{n-t+1}^* - r_t x_t \right] \leq \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \right)^\top \mathbf{p}_t.$$

Similar to the stochastic input model,

$$\begin{aligned} \|\mathbf{p}_{t+1}\|_2^2 &\leq \|\mathbf{p}_t\|_2^2 - \frac{2}{\sqrt{n}} \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \right)^\top \mathbf{p}_t + \frac{1}{n} \left\| \mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \right\|_2^2 \\ &\leq \|\mathbf{p}_t\|_2^2 - \frac{2}{\sqrt{n}} \left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \right)^\top \mathbf{p}_t + \frac{m(\bar{a} + \bar{d})^2}{n}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \sum_{t=1}^n \mathbb{E} \left[\left(\mathbf{d} - \mathbf{a}_t I(r_t > \mathbf{a}_t^\top \mathbf{p}_t) \right)^\top \mathbf{p}_t \right] &\leq \sum_{t=1}^n \mathbb{E} [\sqrt{n}(\|\mathbf{p}_t\|_2^2 - \|\mathbf{p}_{t+1}\|_2^2)] + \sum_{t=1}^n \frac{m(\bar{a} + \bar{d})^2}{\sqrt{n}} \\ &\leq m(\bar{a} + \bar{d})^2 \sqrt{n}. \end{aligned}$$

Combine two parts above, finally we have

$$\begin{aligned} R_n^* - \mathbb{E}[R_n(\pi)] &\leq m\bar{r} + \frac{\bar{r} \log n \sqrt{n}}{\underline{d}} + m\bar{r} \log n + \frac{\max\{16\bar{a}^2, \exp\{16\bar{a}^2\}, e\} \bar{r}}{n} + m(\bar{a} + \bar{d})^2 \sqrt{n} \\ &= O((m + \log n) \sqrt{n}) \end{aligned}$$

Thus, we complete the proof for the regret. The proof for the constraint violation part follows exactly the same way as the stochastic input model. \square

B6 Proof of Proposition 2

First, we define the notion of permutational Rademacher complexity. Consider set $\mathcal{Z}_n = \{z_1, \dots, z_n\}$ where $z_j \in \mathbb{R}^k, j = 1, \dots, n$ and a family of functions $\mathcal{F} = \{f : \mathbb{R}^k \rightarrow \mathbb{R}\}$ (to be specified later). Throughout this section, we use the subscript to indicate the cardinality of a set. For function $f \in \mathcal{F}$ and $\mathcal{S} \subset \mathcal{Z}$, denote $\bar{f}(\mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} f(x)$ as the mean function value on the set \mathcal{S} . The definition of the permutational Rademacher complexity and its analysis largely mimic the analyses of the transductive learning problem in [27].

Definition 1 (Permutational Rademacher Complexity and Conditional Rademacher Complexity (See Definition 3 in [27])). *For any $1 \leq s \leq t-1$, permutational Rademacher complexity (PRC) is defined as follows:*

$$Q_{t,s}(\mathcal{F}, \mathcal{Z}_t) = \mathbb{E} \sup_{f \in \mathcal{F}} \left| \bar{f}(\mathcal{Z}_s) - \bar{f}(\tilde{\mathcal{Z}}_t) \right|,$$

where \mathcal{Z}_s is subset of \mathcal{Z}_t with s elements sampled uniformly without replacement and $\tilde{\mathcal{Z}}_t = \mathcal{Z}_t \setminus \mathcal{Z}_s, l = t - s$. The expectation is taken with respect to the random sampling of \mathcal{Z}_s .

Conditional Rademacher complexity (CRC) is defined as follows:

$$R_t(\mathcal{F}, \mathcal{Z}_t) = \mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{t} \sum_{j=1}^t \epsilon_j f(z_j) \right|,$$

where $\mathcal{Z}_t = \{z_1, \dots, z_t\}$ and ϵ_j 's are i.i.d. random variables following Rademacher distribution ($P(\epsilon_j = 1) = P(\epsilon_j = -1) = 1/2$). The expectation here is taken with respect to ϵ_j 's.

Both the above two quantities are dependent on the set \mathcal{Z}_t because for the PRC, the two subsets \mathcal{Z}_s and $\tilde{\mathcal{Z}}_t$ are sampled from \mathcal{Z}_t and for CRC, it is computed based on the function values of the elements in \mathcal{Z}_t . Both PRC and CRC are deterministic with a given function class \mathcal{F} and conditional on \mathcal{Z}_t . However, they could be random variables if the set \mathcal{Z}_t is random. The following lemma explains the motivation for the definition of permutational Rademacher complexity and it is inspired from Theorem 2 in [27].

Lemma 5. \mathcal{Z}_t is a subset of \mathcal{Z}_n obtained by uniform sampling without replacement, and $\tilde{\mathcal{Z}}_{t'} = \mathcal{Z}_n \setminus \mathcal{Z}_t$, $t' = n - t$. Without loss of generality, assume $t \geq t'$, then the following inequality holds for all $s < t'$,

$$\mathbb{E} \left[\sup_{f \in \mathcal{F}} \left| \bar{f}(\mathcal{Z}_t) - \bar{f}(\tilde{\mathcal{Z}}_{t'}) \right| \middle| \mathcal{Z}_n \right] \leq \mathbb{E} \left[Q_{t,s}(\mathcal{F}, \mathcal{Z}_t) \middle| \mathcal{Z}_n \right]$$

where the expectation is taken with respect to the random sampling of \mathcal{Z}_t from \mathcal{Z}_n

Proof. We have

$$\begin{aligned} \mathbb{E} \left[\sup_{f \in \mathcal{F}} \left| \bar{f}(\mathcal{Z}_t) - \bar{f}(\tilde{\mathcal{Z}}_{t'}) \right| \middle| \mathcal{Z}_n \right] &= \mathbb{E} \left[\sup_{f \in \mathcal{F}} \left| \mathbb{E} [\bar{f}(\mathcal{Z}_{t-s}) | \mathcal{Z}_t] - \mathbb{E} [\bar{f}(\tilde{\mathcal{Z}}_s) | \tilde{\mathcal{Z}}_{t'}] \right| \middle| \mathcal{Z}_n \right] \\ &\leq \mathbb{E} \left[\sup_{f \in \mathcal{F}} \left| \bar{f}(\mathcal{Z}_{t-s}) - \bar{f}(\tilde{\mathcal{Z}}_s) \right| \middle| \mathcal{Z}_n \right] = \mathbb{E} [Q_{t,s}(\mathcal{F}, \mathcal{Z}_t) | \mathcal{Z}_n]. \end{aligned}$$

For the first line, on the right hand side, the two inner expectations are taken with respect to a uniform random sampling on \mathcal{Z}_t and $\tilde{\mathcal{Z}}_{t'}$ respectively. Specifically, \mathcal{Z}_{t-s} (or $\tilde{\mathcal{Z}}_s$) can be viewed as a random sampled subset from \mathcal{Z}_t (or $\tilde{\mathcal{Z}}_{t'}$). For the second line, the first part comes from Jensen's inequality and the expectation in the second part is taken with respect to the random sampling of \mathcal{Z}_t from \mathcal{Z}_n . \square

Let $\mathcal{F}_{\mathbf{p}} = \{f_{\mathbf{p}} : f_{\mathbf{p}}(r, \mathbf{a}) = rI(r > \mathbf{a}^T \mathbf{p})\}$ denote a family of functions $f : \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ indexed by the parameter \mathbf{p} , and let $(r_1, \mathbf{a}_1), \dots, (r_n, \mathbf{a}_n)$ be a random permutation of the dataset \mathcal{D} . Denote $\mathcal{Z}_t = \{(r_1, \mathbf{a}_1), \dots, (r_t, \mathbf{a}_t)\}$, and then \mathcal{Z}_t can be viewed as a randomly sampled subset of \mathcal{D} . The following lemma relates the PRC with the classic notion of CRC, and the benefit is that the CRC under random permutation model possesses a natural upper bound.

Lemma 6. The following inequalities hold for PRC and CRC of the family $\mathcal{F}_{\mathbf{p}}$ and dataset \mathcal{D} that satisfies Assumption 2,

$$\begin{aligned} |Q_{t, \lfloor t/2 \rfloor}(\mathcal{F}_{\mathbf{p}}, \mathcal{Z}_t) - R_t(\mathcal{F}_{\mathbf{p}}, \mathcal{Z}_t)| &\leq \frac{4\bar{r}}{\sqrt{t}}. \\ R_t(\mathcal{F}, \mathcal{Z}_t) &\leq \frac{\sqrt{2\bar{r}^2 m \log n}}{\sqrt{t}} \end{aligned}$$

Proof. For the first inequality, we refer to Theorem 3 in [27]. For the second inequality, it is a direct application of Massart's Lemma (See Lemma 26.8 of [26]). \square

Proof of Proposition 2. Let $\mathcal{F}_{\mathbf{p}} = \{f_{\mathbf{p}} : f_{\mathbf{p}}(r, \mathbf{a}) = rI(r > \mathbf{a}^T \mathbf{p})\}$, $\mathcal{Z}_t = \{(r_1, \mathbf{a}_1), \dots, (r_t, \mathbf{a}_t)\}$, and $\tilde{\mathcal{Z}}_{n-t} = \{(r_{t+1}, \mathbf{a}_{t+1}), \dots, (r_n, \mathbf{a}_n)\}$. Also, we assume $n - t > t$ without loss of generality. Then,

$$\begin{aligned} \mathbb{E} \left[\left| \frac{1}{n-t} \sum_{j=t+1}^n r_j I(r_j > \mathbf{a}_j^T \mathbf{p}) - \frac{1}{t} \sum_{j=1}^t r_j I(r_j > \mathbf{a}_j^T \mathbf{p}) \right| \right] &\leq \mathbb{E} \left[\sup_{f \in \mathcal{F}_{\mathbf{p}}} \left| \bar{f}(\mathcal{Z}_t) - \bar{f}(\tilde{\mathcal{Z}}_{n-t}) \right| \middle| \mathcal{Z}_n = \mathcal{D} \right] \\ &\leq \mathbb{E} [Q_{t, \lfloor t/2 \rfloor}(\mathcal{F}, \mathcal{Z}_t) | \mathcal{Z}_n] \\ &\leq \frac{4\bar{r}}{t} + \frac{2\sqrt{2\bar{r}^2 m \log n}}{\sqrt{t}}. \end{aligned}$$

Here the first line comes from taking maximum over $\mathcal{F}_{\mathbf{p}}$, the second line comes from lemma 5 and the third line comes from lemma 6.

Similarly, we can show that the inequality on a_{ij} 's holds. Thus the proof is completed. \square

B6.1 Proof for Theorem 3

Proof. At time $t + 1$,

$$\begin{aligned}\mathbb{E}[r_{t+1}x_{t+1}] &= \mathbb{E}\left[r_{t+1}I(r_{t+1} > \mathbf{a}_{t+1}^\top \mathbf{p}_{t+1})\right] \\ &= \frac{1}{n-t}\mathbb{E}\left[\sum_{j=t+1}^n r_j I(r_j > \mathbf{a}_j \mathbf{p}_{t+1})\right] \\ &\geq \frac{1}{t}\mathbb{E}\left[\sum_{j=1}^t r_j I(r_j > \mathbf{a}_j \mathbf{p}_{t+1})\right] - \frac{4\bar{r}}{\sqrt{\min\{t, n-t\}}} - \frac{2\sqrt{2\bar{r}^2 m \log n}}{\sqrt{\min\{t, n-t\}}},\end{aligned}$$

where the expectation is taken with respect to the random permutation. The first line comes from the algorithm design, the second line comes from the symmetry over the last $n - t$ terms, and the last line comes from the application of Proposition 2. To relate the first term in the last line with the offline optimal R_n^* , we utilize Proposition 1. Then the optimality gap of Algorithm 2 is as follows,

$$\begin{aligned}R_n^* - \mathbb{E}\left[\sum_{t=1}^n r_t x_t\right] &= R_n^* - \sum_{t=1}^n \mathbb{E}[r_t x_t] \\ &\leq R_n^* - \sum_{t=2}^n \left(\frac{1}{t}\mathbb{E}\left[\sum_{j=1}^t r_j I(r_j > \mathbf{a}_j \mathbf{p}_t)\right] - \frac{4\bar{r} + 2\sqrt{2\bar{r}^2 m \log n}}{\sqrt{\min\{t, n-t\}}}\right) \\ &\leq m\bar{r} + \frac{\bar{r}}{d}\sqrt{n} \log n + m\bar{r} \log n + \bar{r} \max\{16\bar{a}^2, e^{16\bar{a}^2}, e\} + (8\bar{r} + 4\sqrt{2\bar{r}^2 m \log n})\sqrt{n} \\ &= O(\sqrt{mn} \log n)\end{aligned}$$

where the last line comes from an application of Proposition 1. Next, we analyze the constraint; again, from Proposition 2, we know

$$\begin{aligned}\frac{1}{n-t}\mathbb{E}\left[\sum_{j=t+1}^n a_{ij} I(r_j > \mathbf{a}_j^\top \mathbf{p}_t)\right] &\leq \mathbb{E}\left[\frac{1}{t}\sum_{j=1}^t a_{ij} I(r_j > \mathbf{a}_j^\top \mathbf{p}_t)\right] + \frac{4\bar{a}}{\sqrt{\min\{t, n-t\}}} + \frac{2\sqrt{2\bar{a}^2 m \log n}}{\sqrt{\min\{t, n-t\}}} \\ &\leq d_i + \frac{6\sqrt{2\bar{a}^2 m \log n}}{\sqrt{\min\{t, n-t\}}}\end{aligned}$$

where the second line comes from the feasibility of the scaled LP solved at time t . Due to the symmetry of the random permutation,

$$\mathbb{E}\left[a_{i,t+1} I(r_{t+1} > \mathbf{a}_{t+1}^\top \mathbf{p}_{t+1})\right] \leq d_i + \frac{6\sqrt{2\bar{a}^2 m \log n}}{\sqrt{\min\{t, n-t\}}}.$$

Summing up the inequality, we have

$$\mathbb{E}[\mathbf{A}\mathbf{x} - \mathbf{b}] \leq O(\sqrt{mn} \log n).$$

□

B7 Proof of Theorem 4

Proof. At time t , the optimal solution to the scaled LP is $\tilde{\mathbf{x}}^{(t)} = (\tilde{x}_1^{(t)}, \dots, \tilde{x}_t^{(t)})$. We have

$$\begin{aligned}\mathbb{E}[r_t x_t] &= \mathbb{E}\left[r_t \tilde{x}_t^{(t)}\right] \\ &= \frac{1}{t}\mathbb{E}\left[\sum_{j=1}^t r_s \tilde{x}_j^{(t)}\right].\end{aligned}$$

Then, for the objective,

$$\begin{aligned}R_n^* - \sum_{t=1}^n \mathbb{E}[r_t x_t] &= R_n^* - \sum_{t=1}^n \frac{1}{t}\mathbb{E}\left[\sum_{j=1}^t r_s \tilde{x}_j^{(t)}\right] \\ &\leq m\bar{r} + \frac{\bar{r}}{d}\log n \sqrt{n} + m\bar{r} \log n + \bar{r} \max\{16\bar{a}^2, e^{16\bar{a}^2}, e\}.\end{aligned}$$

where the second line comes from an application of Proposition 1. Then, we analyze the constraint violation. From the construction of the algorithm, we have that $\mathbb{E}[a_{it}x_t] \leq d_i$. Let

$$A_{it} = a_{it}x_t - d_i$$

and then we know

$$M_{it} = \sum_{j=n-t+1}^n A_{ij}$$

is a supermartingale with $|A_{ij}| \leq \bar{a} + \bar{d}$. Then if we apply Hoeffding's lemma for supermartingale, we have

$$\begin{aligned} \mathbb{P}(M_{in} \geq 2(\bar{a} + \bar{d})\sqrt{n} \log n) &\leq \exp\left\{-\frac{2(\bar{a} + \bar{d})^2 n \log^2 n}{n(\bar{a} + \bar{d})^2}\right\} \\ &\leq \exp\{-2 \log^2 n\} \leq \frac{1}{n}, \end{aligned}$$

when $n > 3$. Thus,

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{t=1}^n a_{it}x_t - d_i\right)^+\right] &= \mathbb{E}[(M_{in})^+] \\ &\leq 2(\bar{a} + \bar{d})\sqrt{n} \log n \mathbb{P}(M_{in} < 2(\bar{a} + \bar{d})\sqrt{n} \log n) \\ &\quad + \bar{a}n \mathbb{P}(M_{in} \geq 2(\bar{a} + \bar{d})\sqrt{n} \log n) \\ &\leq 2(\bar{a} + \bar{d})\sqrt{n} \log n + \bar{a} \\ \mathbb{E}[v(\mathbf{x})] &\leq 2(\bar{a} + \bar{d})\sqrt{mn} \log n + \bar{a}\sqrt{m}. \end{aligned}$$

□

C1 Algorithm Discussion

Obtaining Feasible Solution

We present a simple approach to convert the solution obtained from Algorithm 1 into a feasible solution. Let $\mathbf{x} = (x_1, \dots, x_n)$ be a solution by Algorithm 1, and $S_+ = \{t : x_t = 1, t = 1, \dots, n\}$ be the index set of nonzero x_t 's and $n_+ = |S_+|$ be the cardinality of S_+ . The idea is to randomly select a subset of S_+ and force $x_t = 0$ for indices in this subset. Note that the expected total constraint violation is sublinear in n , we only need to select a small proportion of x_t 's and force them to be 0. Specifically, define the maximum constraint violation quantity over all constraints:

$$v = \frac{1}{\sqrt{n} \log n} \max_{i=1, \dots, m} \left\{ \left(\sum_{t=1}^n a_{it}x_t - b_i \right)^+ \right\}.$$

Moreover, we require $v \geq 1$. We choose a set $S_0 \subset S_+$ uniformly with $|S_0| = \min\left\{\left\lceil \frac{2v n_+ \log n}{d\sqrt{n}} \right\rceil + 1, n_+\right\}$, and let

$$\hat{x}_t = \begin{cases} 0, & t \in S_0 \\ x_t, & t \notin S_0 \end{cases}$$

for $t = 1, \dots, n$. The following theorem characterizes the properties of \hat{x}_t .

Theorem 5. *If $n > \max\left\{16, d^2, \left(\frac{6\bar{a}}{d}\right)^4\right\}$ and $\sqrt{n} > \frac{12\bar{a}(\bar{r} + (\bar{a} + \bar{d})^2 m) \log n}{d^2}$, then $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n)$ is a feasible solution with probability at least $1 - \frac{2}{n}$. Also, a feasible solution \mathbf{x} can be constructed based on $\hat{\mathbf{x}}$ s.t.,*

$$\mathbb{E}\left[R_n^* - \mathbf{r}^\top \mathbf{x}\right] \leq O((m + \log n)\sqrt{n})$$

for all $m, n \in \mathbb{N}_+$. The results hold under both the stochastic input model and the random permutation model, and the expectation is taken with respect to \mathcal{P} or the random permutation accordingly.

Theorem 5 tells that in a large- n -small- m regime, precisely when $n \geq O(m^2 \log n)$, we can easily obtain a feasible solution with high probability based on the output of Algorithm 1 by randomly selecting $O(\sqrt{n} \log n)$ number of x_t and forcing them to be 0. Furthermore, the newly obtained solution does not change the regret much. The theorem provides a guideline of the implementation of Algorithm 1 for the binary LP setting when a feasible solution is desired.

Feasible Online Algorithm

Algorithm 4 is another natural variant of Algorithm 1 that outputs feasible solutions. Compared with Algorithm 1, Algorithm 4 sets $x_t = 1$ only when the constraints permit. This design is more aligned with the online LP algorithms that guarantees feasibility. [20] provided a regret analysis framework for this type of feasible

algorithms, and the key is to analyze the stopping time of constraint violation and the remaining resources for binding constraints. In this paper, the assumptions on (r_j, \mathbf{a}_j) are parsimonious and they might be not sufficient to derive an upper bound on these two key quantities. Numerically, we observe that this feasible algorithm, in comparison with Algorithm 1, does not compromise the performance in terms of the regret. We will elaborate more on its numerical performance in the next section and leave the regret analysis of this algorithm as an open question.

Algorithm 4 Simple Feasible Algorithm

- 1: Input: d
- 2: Initialize $\mathbf{p}_1 = \mathbf{0}$
- 3: **for** $t = 1, \dots, n$ **do**
- 4: Set

$$\tilde{x}_t = \begin{cases} 1, & r_t > \mathbf{a}_t^\top \mathbf{p}_t \\ 0, & r_t \leq \mathbf{a}_t^\top \mathbf{p}_t \end{cases}$$

- 5: Compute

$$\begin{aligned} \mathbf{p}_{t+1} &= \mathbf{p}_t + \gamma_t (\mathbf{a}_t \tilde{x}_t - \mathbf{d}) \\ \mathbf{p}_{t+1} &= \mathbf{p}_{t+1} \vee \mathbf{0} \end{aligned}$$

- 6: If constraints permit, set $x_t = \tilde{x}_t$; otherwise set $x_t = 0$.
 - 7: **end for**
 - 8: Output: $\mathbf{x} = (x_1, \dots, x_n)$
-

Nonstationary Algorithm

We consider another variant of the algorithm that takes into account the resource consumption while doing the subgradient descent. The intuition is similar to the action-history-dependent algorithm in [20]. If excessive resources are consumed in the early periods, the remaining resource \mathbf{b}_t will shrink, and this nonstationary algorithm will accordingly push up the dual price and be more inclined to reject an order. On the contrary, if we happen to reject a lot orders at the beginning and it results in too much remaining resources, the algorithm will lower down the dual price so as to accept more orders in the future. In numerical experiments, this nonstationary algorithm performs better, but it is still on the same order of regret and constraint violation as Algorithm 1. The open question is if there exists a first-order algorithm that is free of re-solving any linear programs and could achieve $O(\log n)$ regret, possibly under stronger statistical assumptions.

Algorithm 5 Simple Nonstationary Algorithm

- 1: Input: d
- 2: Initialize $\mathbf{p}_1 = \mathbf{0}, \mathbf{b}_0 = \mathbf{b}$
- 3: **for** $t = 1, \dots, n$ **do**
- 4: Set

$$x_t = \begin{cases} 1, & r_t > \mathbf{a}_t^\top \mathbf{p}_t \\ 0, & r_t \leq \mathbf{a}_t^\top \mathbf{p}_t \end{cases}$$

- 5: Update

$$\mathbf{b}_t = \mathbf{b}_{t-1} - \mathbf{a}_t x_t$$

- 6: Compute

$$\begin{aligned} \mathbf{p}_{t+1} &= \mathbf{p}_t + \gamma_t \left(\mathbf{a}_t x_t - \frac{\mathbf{b}_t}{n-t} \right) \\ \mathbf{p}_{t+1} &= \mathbf{p}_{t+1} \vee \mathbf{0} \end{aligned}$$

- 7: **end for**
 - 8: Output: $\mathbf{x} = (x_1, \dots, x_n)$
-

D1 Numerical Experiments

The first experiment compares the performance of Algorithm 1, Algorithm 4, and Algorithm 5 in terms of regret and constraint violation. Algorithm 1 is implemented with two different choices of step size γ_t . In

this experiment, $m = 10$, a_{ij} 's and r_j 's are sampled i.i.d. from $\text{Unif}[0, 2]$. For each value of n , we run 100 simulation trials and in each trial, d_i 's are sampled i.i.d. from $\text{Unif}[1/3, 2/3]$. The average regret and constraint violation over all the simulation trials are shown in Figure 1. We plotted normalized regret and constraint violation, which is absolute regret and constraint violation divided the optimal objective value and the L_2 norm of the constraint capacity, respectively. We observe that the step size of $1/\sqrt{n}$ results in larger constraint violation but smaller regret compared with the step size of $1/\sqrt{t}$. This is because for the step size of $1/\sqrt{n}$, the updating of the dual vector \mathbf{p}_t is slower. Consequently, more requests will be accepted at early stage and the constraint violation is larger in the end. The non-stationary algorithm (Algorithm 5) performs better than the simple algorithm (Algorithm 1) with $\gamma_t = 1/\sqrt{t}$. Also, the feasible algorithm (Algorithm 4) guarantees feasibility, i.e. zero constraint violation; it produces slightly larger regret, but the regret is still on the order of \sqrt{n} .

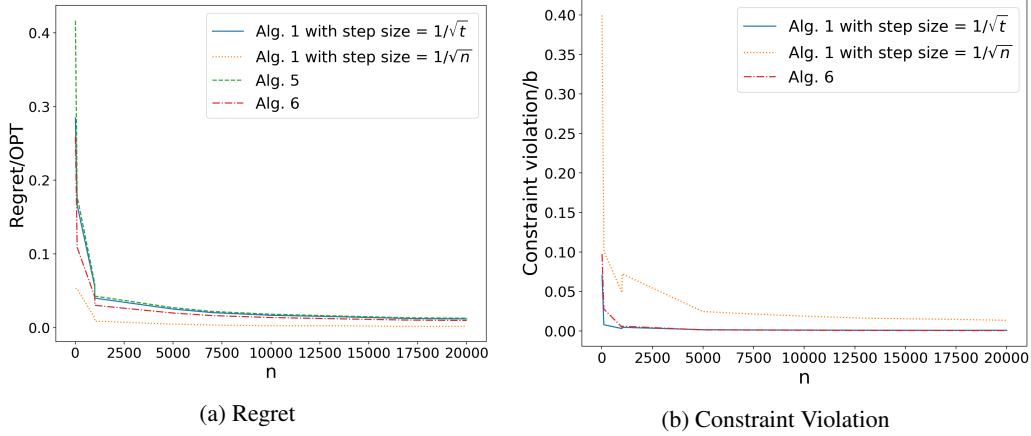


Figure 1: Experiment with Uniform i.i.d. input

In the second experiment (Figure 2), we compare the three algorithms in a setting where the boundedness of the support of distribution \mathcal{P} is violated. Specifically, $m = 10$, a_{ij} 's are generated i.i.d. from $\mathcal{N}(1, 1)$ and $r_j = \sum_{i=1}^m a_{ij} - \epsilon_j$ where $\epsilon_j \sim \text{Unif}(0, m)$. For each value of n , we run 100 simulation trials, and in each trial, d_i 's are sampled i.i.d. from $\text{Unif}[1/3, 2/3]$. In this experiment, the regret performances of Algorithm 1 (with step size of $1/\sqrt{t}$) and Algorithm 5 are quite close to each other, while Algorithm 5 still performs better in respect with constraint violation. The feasible algorithm (Algorithm 4) still achieves regret on the order of \sqrt{n} . Note that our theoretical results, also all the previous literature on online LP problem, rely on the boundedness assumption for the LP input. An open question is how to generalize these analyses to the case when the distribution \mathcal{P} has an unbounded support.

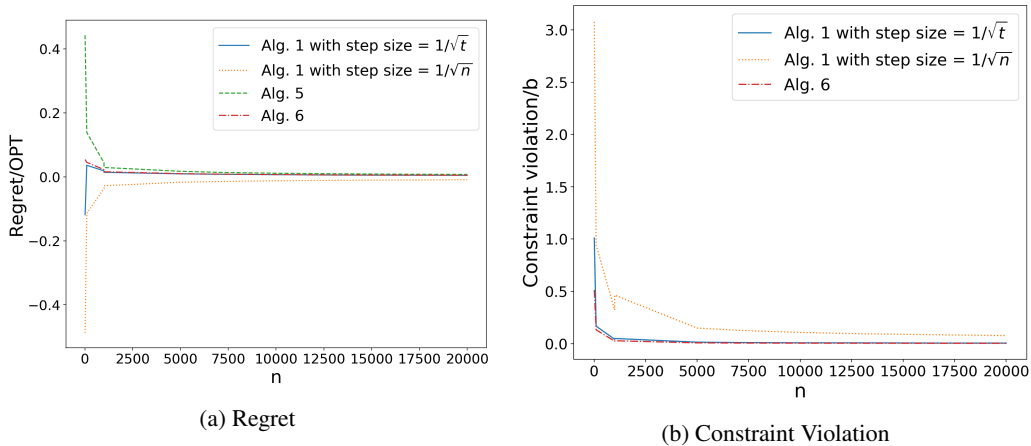


Figure 2: Experiment with Gaussian i.i.d. input

The third experiment (Figure 3) presents a negative result on all three algorithms. Specifically, a_{ij} 's are generated i.i.d. from truncated Cauchy(1, 1) (with different thresholds) and $r_j = \sum_{i=1}^m a_{ij} - \epsilon_j$ where $\epsilon_j \sim \text{Unif}(0, m)$. As before, for each value of n , we run 100 simulation trials, and in each trial, d_i 's are sampled i.i.d. from $\text{Unif}[1/3, 2/3]$. We observe that the performance becomes unstable as the truncation threshold goes up. The

phenomenon is consistent with the previous analysis that the algorithm regret is positively affected by the upper bound on a_{ij} 's and r_j 's. The empirical finding suggests that a light-tail distribution is probably necessary for an online LP algorithm to succeed.

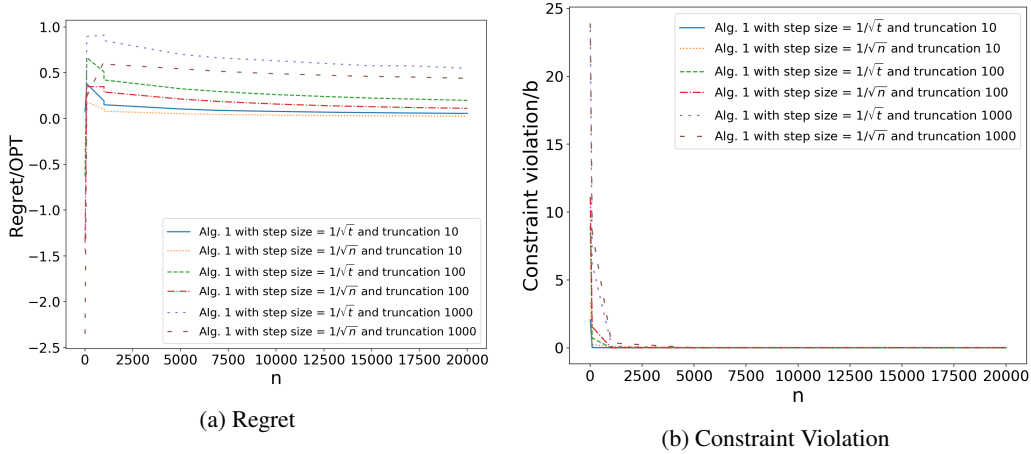


Figure 3: Experiment with Cauchy i.i.d. input

Figure 4 presents the algorithm performance under the random permutation model. We first generate four groups of data with equal size from four different distributions and then combine these groups as the LP input: (i) a_{ij} 's are generated from $\text{Unif}[0, 2]$; (ii) a_{ij} are generated from $\mathcal{N}(1, 1)$; (iii) a_{ij} are generated from $\mathcal{N}(0, 1)$; (iv) a_{ij} are generated from uniform distribution on $\{-1, 1, 3\}$. r_j 's for all four groups are generated from $\text{Unif}[0, 1]$. Note this data set can not be generated from any distribution in the stochastic input model. For each value of n , we run 100 simulation trials, and in each trial, d_i 's are sampled i.i.d. from $\text{Unif}[1/3, 2/3]$. In this experiment, we observe that the algorithms all achieves $O(\sqrt{n})$ regret except for Algorithm 1 with step size $1/\sqrt{n}$. The step size results in a negative regret but much larger constraint violation than the other algorithms. All the presented algorithms achieve $O(\sqrt{n})$ constraint violation.

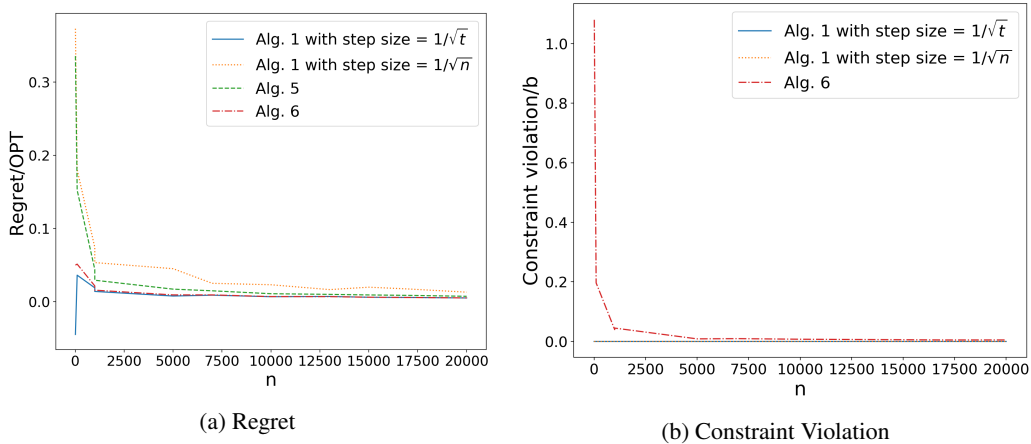


Figure 4: Experiment with randomly permuted input

In addition, we conduct two groups of experiments to illustrate the computational aspect of our algorithms. The algorithms are implemented on a PC with Intel Core i7-9700K Processor. We use the Gurobi solver - one of the state-of-the-art LP and integer LP solvers - as benchmark and to compute the optimal solution. We emphasize that the code for our algorithms, unlike the Gurobi solver, is not fully optimized in the experiment, so there is still great room for improving the computational efficiency of our algorithms. Table 1 presents the experiments under the worst-case example for online LP problem in [3]. The example is constructed under the random permutation model, and it provides a lower bound on the right-hand-side of the constraint for the existence of an $(1 - \epsilon)$ -competitiveness online LP algorithm. In this sense, it represents one of the most challenging problem instances under the random permutation model. Gurobi solves the binary LP problem with 1% MIPGap and all competitiveness ratios in the table are reported against the optimal objective value of the relaxed LP's optimal objective. Gurobi computes the optimal solution in an offline fashion while the other three algorithms

are online. The numbers are computed based on an average over 100 different problem instances. Algorithm 1 is implemented with both step size of $\frac{1}{\sqrt{t}}$ and $\frac{1}{\sqrt{n}}$. We stop the algorithm and set the rest decision variables to be zero when any of the constraint is exhausted. Though Algorithm 2 and Algorithm 3 have provably smaller regret bounds, our algorithm provides better empirical performance. For small value of m , our algorithm outputs a near-optimal solution within much less time than Gurobi. Also, the competitiveness ratio decreases as m goes larger, which is consistent with the $O(m)$ term in the regret bound of our algorithm.

		Gurobi	Alg. 1 ($\frac{1}{\sqrt{t}}$)	Alg. 1 ($\frac{1}{\sqrt{n}}$)	Alg. 2	Alg. 3
$m = 8, n = 10^3$	CPU time	0.082	0.023	0.015	126.684	126.696
	Cmpt. Ratio	100.0%	99.1%	99.7%	89.8%	97.9%
$m = 128, n = 10^4$	CPU time	0.408	0.141	0.138	2338.149	2338.285
	Cmpt. Ratio	100.0%	99.3%	98.8%	94.8%	97.7%
$m = 1024, n = 10^5$	CPU time	52.496	3.479	3.270	>3000	>3000
	Cmpt. Ratio	100.0%	94.6%	98.7%		
$m = 4096, n = 10^5$	CPU time	114.96	27.093	32.020	>3000	>3000
	Cmpt. Ratio	99.6%	73.1%	83.3%		
$m = 4096, n = 2 \times 10^5$	CPU time	254.243	57.541	53.887	>3000	>3000
	Cmpt. Ratio	99.7%	80.0%	89.1%		

Table 1: Performance under worst-case example in [3]

In Table 2, we test the performance of our algorithm on some Multi-knapsack benchmark problems [7, 12]. As the last experiment, algorithm 1 is implemented with both step size of $\frac{1}{\sqrt{t}}$ and $\frac{1}{\sqrt{n}}$, and the experiment setup is the same as the last experiment. The competitiveness ratios are reported against the relaxed LP's optimal objective value. The computational advantage of our simple online algorithm is significant. Like the last experiment, although Algorithm 2 and Algorithm 3 reduces the regret upper bound by a factor of \sqrt{m} , the advantage of these two algorithms is not evident in practice. It deserves more efforts to understand whether the current regret bound for Algorithm 1 is tight. Also, it merits more study on how to design online algorithms that works effectively for a large- m -and-small- n regime.

		Gurobi	Alg. 1 ($\frac{1}{\sqrt{t}}$)	Alg. 1 ($\frac{1}{\sqrt{n}}$)	Alg. 2	Alg. 3
$m = 5, n = 500$	time	0.116	0.006	0.006	70	70
	Cmpt. Ratio	99.6%	92.3%	75.05%	91.8%	91.5%
$m = 10, n = 500$	time	0.136	0.006	0.006	132	132
	Cmpt. Ratio	99.6%	91.8%	80.9%	91.6%	90.7%
$m = 30, n = 500$	time	95.2	0.006	0.005	134	133
	Cmpt. Ratio	99.4%	91.5%	89.4%	89.1%	90.4%
$m = 10^3, n = 10^5$	time	857	2.711	2.679	>3000	>3000
	Cmpt. Ratio	99.8%	94.9%	97.3%		
$m = 3 \times 10^3, n = 10^5$	time	1069	16.63	16.61	>3000	>3000
	Cmpt. Ratio	94.0%	82.3%	86.8%		
$m = 3 \times 10^3, n = 2 \times 10^5$	time	2799	40.21	34.84	>3000	>3000
	Cmpt. Ratio	93.8%	84.7%	88.4%		

Table 2: Multi-knapsack benchmark problem