Online Integer Covering in Random Order

Anupam Gupta
Carnegie Mellon University

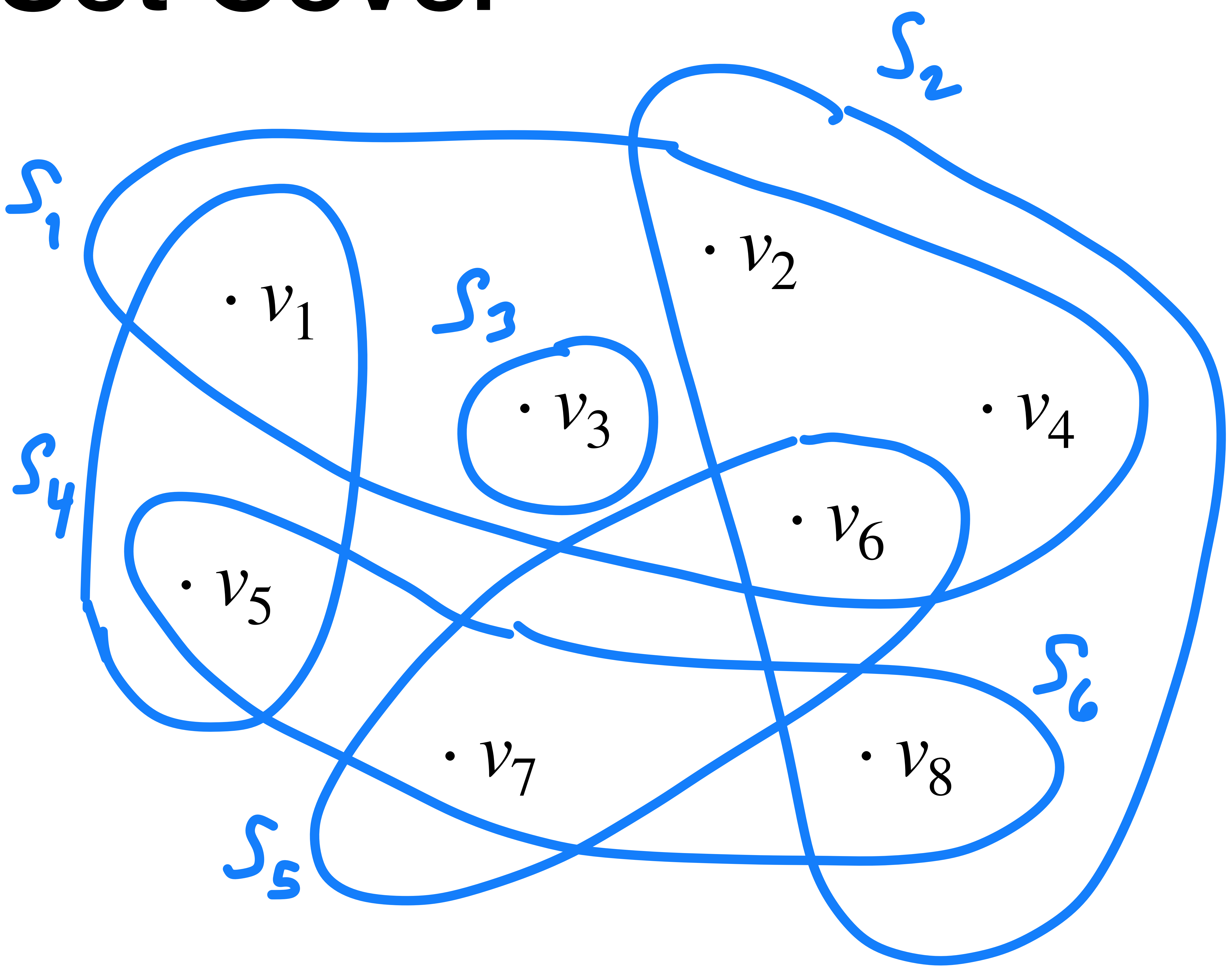
Gregory Kehne
HARVARD UNIVERSITY

Roie Levin
TEL AVIV UNIVERSITY אוניברסיטת תל אביב

November 10, 2022

Duke University

Set Cover



Goal: find the smallest (cheapest) cover of all of $\mathscr{U} = \{v_1, \ldots, v_8\}$ using sets from $\mathscr{S} = \{S_1, \ldots, S_6\}$

e.g. $\{S_2, S_3, S_4, S_5\}$, $\{S_1, S_6\}$

Set Cover:

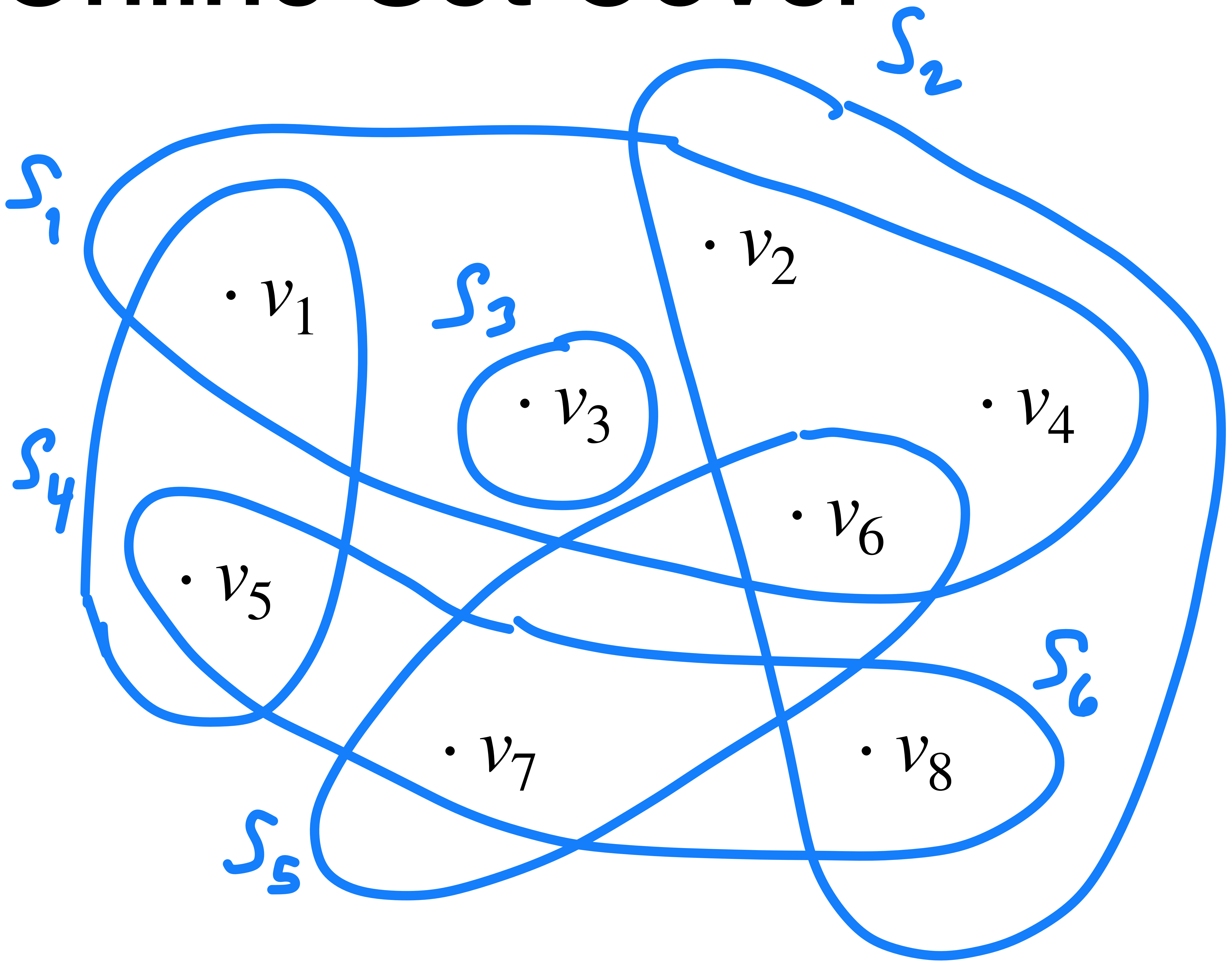
$$\min \ c^T x$$
$$Ax \geq \mathbf{1}$$
$$x \in \{0,1\}$$

$$c \geq \mathbf{0}$$
$$A \in \{0,1\}^{m \times n}$$

$|\mathscr{U}| = n$ elements, $|\mathscr{S}| = m$ sets

v_1: $\quad x_1 + x_4 \geq 1$

Online Set Cover



Online Set Cover:

$$\begin{aligned} \min\ & c^T x \\ Ax &\geq \mathbf{1} \\ x &\in \{0,1\} \end{aligned}$$

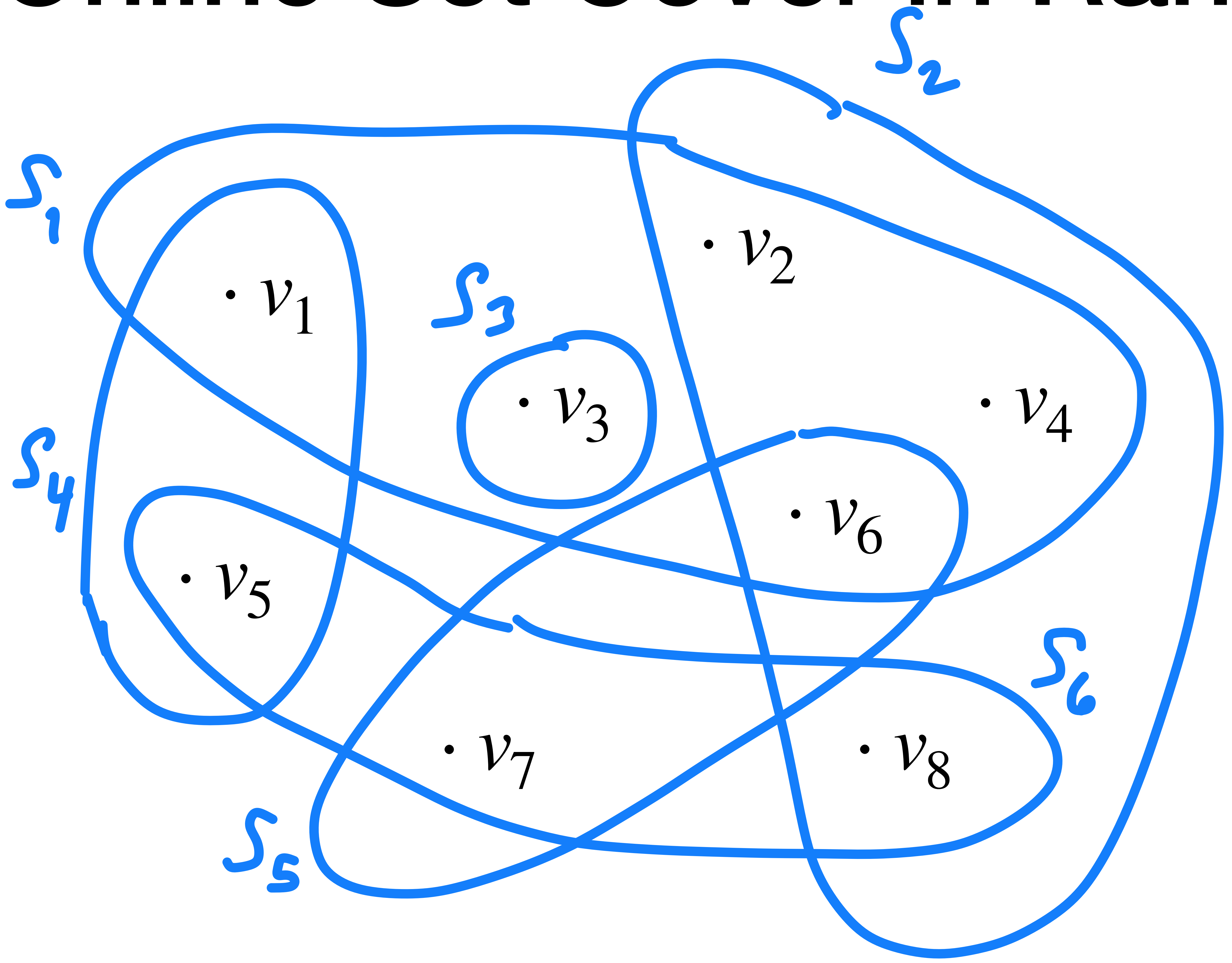
constraints arrive one-at-a-time

Goals:
- Satisfy each constraint upon arrival
- Maintain a solution which is monotone increasing
- Compete with the best solution in retrospect

Arrivals	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
Sets Bought	S_4	S_1					S_5	S_6

$\{S_1, S_6\}$

Online Set Cover in Random Order (RO)



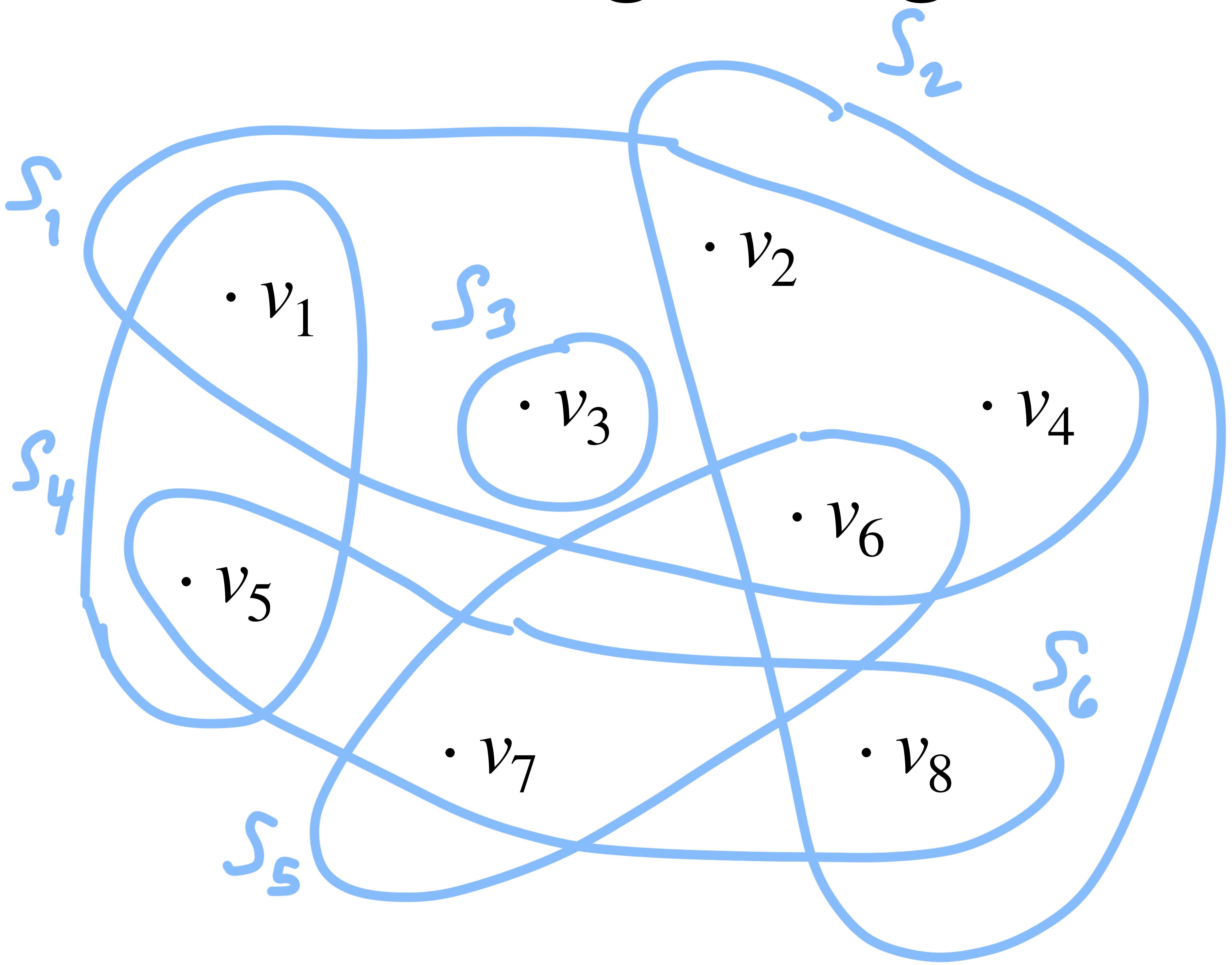
Arrivals	v_5	v_8	v_3	v_1	v_2	v_6	v_7	v_4
Sets Bought	S_4	S_6	S_1					

RO Set Cover:

$$\min \ c^T x$$
$$\left.\begin{aligned} a_1^T x &\geq 1 \\ a_2^T x &\geq 1 \\ &\vdots \\ a_m^T x &\geq 1 \end{aligned}\right\} \text{constraints arrive in a random order}$$
$$x \in \{0,1\}$$

Goal: Same as Online Set Cover, but compete *in expectation* over the randomness of the arrival order

RO Covering Integer Programs (IPs)



Arrivals $\cdot v_5 \quad \cdot v_8 \quad \cdot v_3 \quad \cdot v_1 \quad \cdot v_2 \quad \cdot v_6 \quad \cdot v_7 \quad \cdot v_4$

Vars Bought $S_4 S_4 \quad S_6 \quad S_1 \quad\quad S_2 \quad S_1 \quad\quad S_1 S_2$

RO Covering IP:

$$\begin{aligned} &\min \ c^T x \\ &a_1^T x \geq 1 \\ &a_2^T x \geq 1 \\ &\quad \vdots \\ &a_m^T x \geq 1 \\ &\quad x \in \{0,1,2,\ldots\} \end{aligned}$$

constraints arrive in a random order

$$c \geq \mathbf{0}$$

$$a_i \in [0,1]^m$$

Outline

- Introduction
- Prior Work
- LearnOrCover (Warmup)
- LearnOrCover in Polynomial Time
- Some Intuition
- Lower Bounds
- More Adversaries Beaten!

Outline

Introduction

Prior Work

LearnOrCover (Warmup)

LearnOrCover in Polynomial Time

Some Intuition

Lower Bounds

More Adversaries Beaten!

The Landscape

Offline	$\log n + 1$ [Johnson '74], [Lovasz '75], [Chvatal '79]
Online Adversarial	$O(\log m \ \log n)$ [Alon Awerbuch Azar Buchbinder Naor '03] $\Omega(\log m \ \log n)$ [Korman '04]
Online Stochastic	$\Theta(\log mn)$ [Grandoni Gupta Leonardi Miettinen Sankowski Singh '08]
Online RO	$\Theta(\log mn)$ [Gupta **K.** Levin '21]

<u>Theorem:</u> (Gupta **K.** Levin): There is a randomized poly-time algorithm for RO covering IPs with an expected competitive ratio of $O(\log mn)$

The Landscape from a different view

Instance

Arrival Order		Stochastic	Adversarial
	RO	$\Theta(\log mn)$ [Grandoni Gupta Leonardi Miettinen Sankowski Singh ’08] constraints are i.i.d. samples from known $\mathscr{D}$	$\Theta(\log mn)$ [this talk] (secretary setting)
	Adversarial	$\boxed{\Theta(\log mn)\ ?}$ constraints are independent samples from known $\mathscr{D}_i$ (prophet setting)	$\Theta(\log m\ \log n)$ [Alon Awerbuch Azar Buchbinder Naor ’03] [Korman ’04]

What makes online integer covering (online set cover) harder than offline?

Outline

- Introduction
- Prior Work
- ▶ LearnOrCover (Warmup)
- LearnOrCover in Polynomial Time
- Some Intuition(?!)
- Lower Bounds
- More Adversaries Beaten!

Warmup: LearnOrCover (proof of concept)

How can we get a (randomized) $O(\log mn)$-approximation to ROSC online, supposing

- all sets have unit cost, and
- we are allowed exponential time?

$|\mathscr{U}| = n$ elements

$|\mathscr{S}| = m$ sets

Unit-Cost Set Cover:

$$\begin{aligned} &\min \ \mathbf{1}^T x \\ &a_1^T x \geq 1 \\ &a_2^T x \geq 1 \\ &\vdots \\ &a_n^T x \geq 1 \\ &x \in \{0,1\} \end{aligned}$$

$a_i \in \{0,1\}^m$

Warmup: LearnOrCover (proof of concept)

guess $k = |OPT|$ and set $\mathscr{P}^0$

for v arriving uncovered:

(Cover)

choose $T \sim \mathscr{P}^t$

buy $S \sim T$

(Learn)

$\mathscr{P}^{t+1} \leftarrow \mathscr{P}^t \setminus \{T \not\ni v\}$

(Backup)

buy arbitrary $S \ni v$

Case 1: $\geq 1/2$ of $T \in \mathscr{P}^t$ cover $\geq 1/2$ of $\mathscr{U}^t$

S covers at least $\dfrac{|\mathscr{U}^t|}{4k}$ elements in expectation

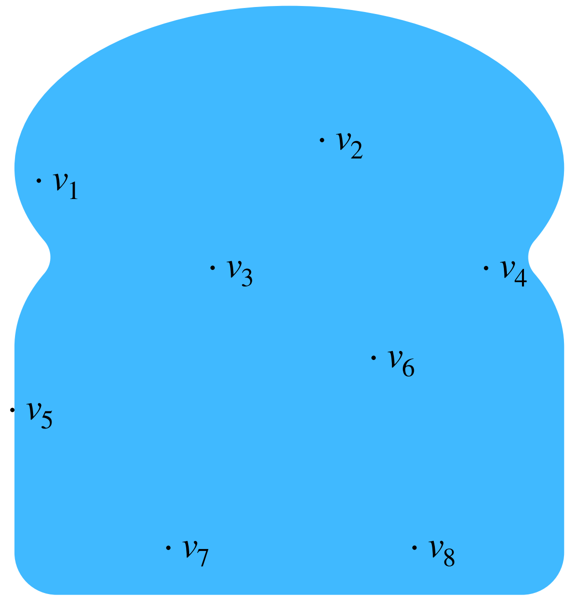
$|\mathscr{U}^{t+1}| \leq \left(1 - \dfrac{1}{4k}\right)|\mathscr{U}^t|$ in expectation.

Case 2: $> 1/2$ of $T \in \mathscr{P}^t$ cover $< 1/2$ of $\mathscr{U}^t$

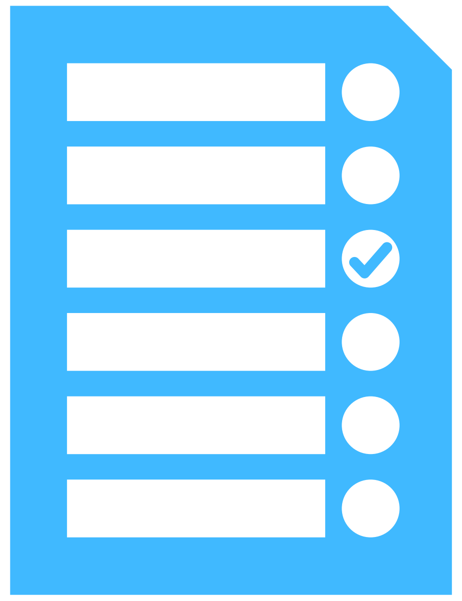
at least $1/4$ of $T \in \mathscr{P}^t$ pruned in expectation

$|\mathscr{P}^{t+1}| \leq \dfrac{3}{4}|\mathscr{P}^t|$ in expectation.

$\mathscr{U}^t$



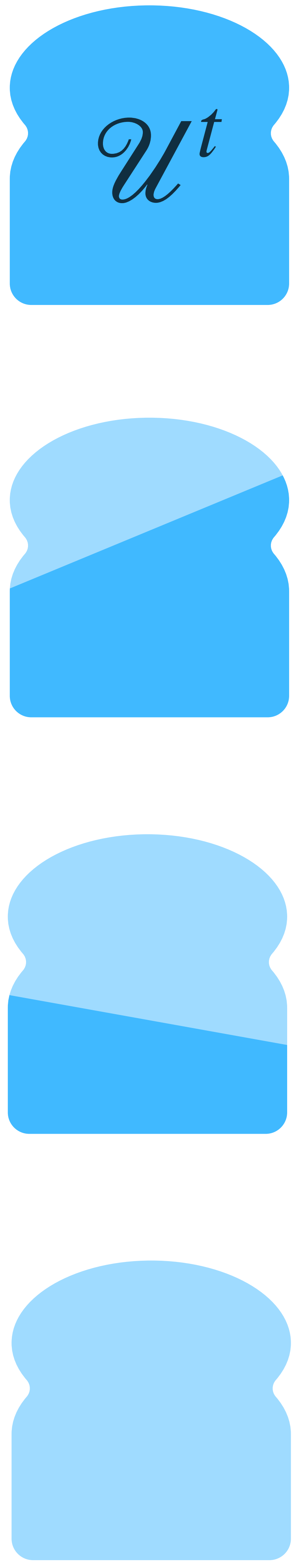
$\mathscr{P}^t$



$\mathscr{U}^0 = \mathscr{U}$

$\mathscr{P}^0 = \binom{\mathscr{S}}{k}$

Warmup: LearnOrCover (proof of concept)



Case 1: $\geq 1/2$ of $T \in \mathscr{P}^t$ cover $\geq 1/2$ of $\mathscr{U}^t$

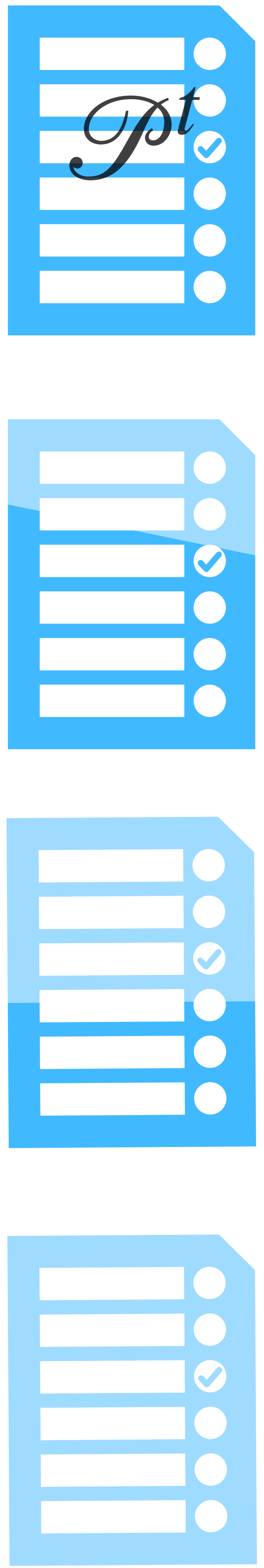
$$\mathbb{E}\,|\mathscr{U}^{t+1}| \leq \left(1 - \frac{1}{4k}\right)|\mathscr{U}^t|$$

$|\mathscr{U}^0| = n$, so $O(k \log n)$ **Cover** steps suffice

Case 2: $> 1/2$ of $T \in \mathscr{P}^t$ cover $< 1/2$ of $\mathscr{U}^t$

$$\mathbb{E}\,|\mathscr{P}^{t+1}| \leq \frac{3}{4}|\mathscr{P}^t|$$

$|\mathscr{P}^0| \leq m^k$, so $O(k \log m)$ **Learn** steps suffice



Once $\mathbb{E}\,|\mathscr{U}^t| = 1$ or $\mathbb{E}\,|\mathscr{P}^t| = 1$ we are done, so $OPT \cdot O(\log mn)$ steps suffice!

In other words:

Potential:
$$\Phi(t) = k \log|\mathscr{U}^t| + \log|\mathscr{P}^t|$$
$$0 \leq \Phi(0) \leq k \log n + k \log m$$

In expectation over the randomness of the arrival order + the algorithm, its solution will cost $O(\log mn)$ times the cost of the optimal offline solution.

every step decreases $\Phi(t)$ by $\Omega(1)$ in expectation, so $OPT \cdot O(\log mn)$ steps suffice!

Outline

Introduction

Prior Work

LearnOrCover (Warmup)

▷LearnOrCover in Polynomial Time

Some Intuition

Lower Bounds

More Adversaries Beaten!

LearnOrCover in Polynomial Time (unit cost: $c = 1$)

(estimate $k = |OPT|$)

initialize $x \leftarrow k/m$

for v arriving uncovered (round t)

(Cover)

buy random $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow e \cdot x_S$ for all $S \ni v$

$x \leftarrow k\dfrac{x}{\|x\|}$

(Backup)

buy arbitrary $S \ni v$

$x_v = \sum_{S \ni v} x_S$ is coverage of v by x

$$KL(x^* \| x^t) = \sum_S x^*_S \log\left(\frac{x^*_S}{x^t_S}\right)$$

Potential: $\Phi(t) = c_1 \cdot KL(x^* \| x^t) + c_2 \cdot k \log |\mathcal{U}^t|$

Claim 1: $\Phi(0) = O(k \log mn)$ and $\Phi(t) \geq 0$

Claim 2: $\mathbb{E}[\Delta\Phi] \leq -1$ whenever v arrives uncovered

$KL(x^* \| x^t)$ and $k \log |\mathcal{U}^t|$ are nonincreasing

Case 1: $\mathbb{E}_v[x_v] > 1/4$

expected change to $k \log |\mathcal{U}^t|$ is $-\Omega(1)$

Case 2: $\mathbb{E}_v[x_v] \leq 1/4$

expected change to $KL(x^* \| x^t)$ is $-\Omega(1)$

Claim 1 ∧ Claim 2 ⇒ LearnOrCover has $O(\log mn)$ CR

LearnOrCover in Polynomial Time (unit cost: $c = 1$)

(estimate $k = |OPT|$)

initialize $x \leftarrow k/m$

for v arriving uncovered (round t)

(Cover)

buy random $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow e \cdot x_S$ for all $S \ni v$

$x \leftarrow k\dfrac{x}{\|x\|}$

(Backup)

buy arbitrary $S \ni v$

$x_v = \sum_{S \ni v} x_S^{t-1}$ is coverage of v by x^{t-1}

Lemma 1: if $\mathbb{E}_v[x_v] > 1/4$, then the expected change to $k \log |\mathcal{U}^t|$ is $-\Omega(1)$.

Proof:

$$\log|\mathcal{U}^t| - \log|\mathcal{U}^{t-1}| = \log\left(1 - \frac{|\mathcal{U}^{t-1}| - |\mathcal{U}^t|}{|\mathcal{U}^{t-1}|}\right)$$

log(1-z) < -z

$$\leq \frac{-1}{|\mathcal{U}^{t-1}|}\sum_{v \in \mathcal{U}^{t-1}} \mathbb{I}\{S \ni v\}$$

sample rate

$$k\mathbb{E}_S[\Delta \log|\mathcal{U}^t|] \leq \frac{-k}{|\mathcal{U}^{t-1}|}\sum_S \frac{x_S^{t-1}}{k}\sum_{v \in \mathcal{U}^{t-1}} \mathbb{I}\{S \ni v\}$$

$$= \frac{-1}{|\mathcal{U}^{t-1}|}\sum_{v \in \mathcal{U}^{t-1}}\sum_{S \ni v} x_S^{t-1}$$

switch sum

$$= -\mathbb{E}_v[x_v] \quad \blacksquare$$

LearnOrCover in Polynomial Time (unit cost: $c = 1$)

(estimate $k = |OPT|$)

initialize $x \leftarrow k/m$

for v arriving uncovered (round t)

(Cover)

buy random $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow e \cdot x_S$ for all $S \ni v$

$x \leftarrow k\dfrac{x}{\|x\|}$

(Backup)

buy arbitrary $S \ni v$

$x_v = \sum_{S \ni v} x_S^{t-1}$ is coverage of v by x^{t-1}

Lemma 2: if $\mathbb{E}_v[x_v] \leq 1/4$, then the expected change to $KL(x^* \| x^t)$ is $-\Omega(1)$.

Proof:

$$\sum_S x_S^* \log\left(\frac{x_S^*}{x_S^t}\right) - \sum_S x_S^* \log\left(\frac{x_S^*}{x_S^{t-1}}\right) = \sum_S x_S^* \log\left(\frac{x_S^{t-1}}{x_S^t}\right)$$

$$= \sum_S x_S^* \log\left(x_S^{t-1}\frac{\sum_T x_T^{t-1} \cdot e^{\mathbb{1}\{T \ni v\}}}{k \cdot x_S^{t-1} \cdot e^{\mathbb{1}\{S \ni v\}}}\right)$$

defn. of x_S^t

$$= \sum_S x_S^* \log\left(\frac{1}{k}\sum_T x_T^{t-1} \cdot e^{\mathbb{1}\{T \ni v\}}\right) - \sum_{S \ni v} x_S^* \log e \quad \} \geq 1$$

$k \log(1+\frac{z}{k}) \leq z$

$$\leq k \log\left(\frac{1}{k}\underbrace{\sum_S x_S^{t-1}}_{1} + \frac{1}{k}\sum_{S \ni v}(e-1)\cdot x_S^{t-1}\right) - 1$$

$$\leq (e-1)\cdot x_v - 1$$

$$\mathbb{E}_v[\Delta KL] \leq -\Omega(1)$$

by Markov

LearnOrCover in Polynomial Time (unit cost: $c = 1$)

(estimate $k = |OPT|$)

initialize $x \leftarrow k/m$

for v arriving uncovered (round t)

(Cover)

buy random $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow e \cdot x_S$ for all $S \ni v$

$$x \leftarrow k\frac{x}{\|x\|}$$

(Backup)

buy arbitrary $S \ni v$

$x_v = \sum_{S \ni v} x_S$ is coverage of v by x

Potential: $\Phi(t) = c_1 \cdot KL(x^* \| x^t) + c_2 \cdot k \log |\mathcal{U}^t|$

Claim 1: $\Phi(0) = O(k \log mn)$ and $\Phi(t) \geq 0$

Claim 2: $\mathbb{E}[\Delta\Phi] \leq -1$ whenever v arrives uncovered

$KL(x^* \| x^t)$ and $k \log |\mathcal{U}^t|$ are nonincreasing

expected change to $k \log |\mathcal{U}^t|$ is $-\Omega(1)$

or expected change to $KL(x^* \| x^t)$ is $-\Omega(1)$

punchline: $0 \leq \mathbb{E}[\Phi(t)] \leq \Phi(0) - \Omega(t)$

Theorem (Gupta **K.** Levin): LearnOrCover has a competitive ratio of $O(\log mn)$ for unit-cost RO set cover.

LearnOrCover OSC with General Costs

(estimate $k = c(OPT)$)

initialize $x \leftarrow \mathrm{k}/(c_S \cdot m)$

for v arriving uncovered (round t)

$c_v \leftarrow$ cost of cheapest $S \ni v$

(Cover)

sample c_v worth of $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow x_S \cdot e^{\mathbb{I}\{S \ni v\} \frac{c_v}{c_S}}$

$x \leftarrow k \dfrac{x}{\|x\|}$

(Backup)

buy $S \ni v$ costing c_v

$KL_w(x^* \| x^t) = \sum_S c_S \cdot x_S^* \log\left(\frac{x_S^*}{x_S^t}\right)$

$\rho^t = \sum_{v \in \mathcal{U}^t} c_v$

Potential: $\Phi(t) = c_1 \cdot KL_w(x^* \| x^t) + c_2 \cdot k \log |\rho^t / k|$

Claim 1: $\Phi(0) = O(k \log mn)$ and $\Phi(t) \geq 0$

Claim 2: $\mathbb{E}[\Delta\Phi] \leq -c_v$ whenever v arrives uncovered

$KL(x^* \| x^t)$ and $k \log |\mathcal{U}^t|$ are nonincreasing

expected change to $k \log |\mathcal{U}^t|$ is $-\Omega(c_v)$

or expected change to $KL(x^* \| x^t)$ is $-\Omega(c_v)$

spend $O(c_v)$ per step, so $\mathbb{E}[c(LoC) + \Phi]$ nonincreasing.

Theorem (Gupta **K.** Levin): LearnOrCover has a competitive ratio of $O(\log mn)$ for RO set cover.

LearnOrCover for RO Covering IPs

(estimate $k = c(OPT)$)

initialize $x \leftarrow$ k$/(c_S \cdot m)$

for v arriving uncovered (round t)

$c_v \leftarrow$ cost of cheapest cover

(Cover)

sample c_v worth of $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow x_S \cdot e^{\mathbb{I}\{S \ni v\} \frac{c_v}{c_S} a_{vS}}$

$x \leftarrow k \dfrac{x}{||x||}$

(Backup)

buy $S \ni v$ costing c_v

…very similar!

Major changes are:

- incorporating partial coverage: measure ρ^t according to remaining uncoverage, sample "sets" according to a_{vS}
- analysis of $\mathbb{E}[\Delta \log |\mathcal{U}|]$ is more involved (independent sampling with partial coverage)

RO Covering IP:

$$\min \ c^T x$$
$$Ax \geq \mathbf{1}$$
$$x \in \{0,1,2,\ldots\}$$
$$c \geq 0$$
$$A \in [0,1]^{m \times n}$$

<u>Theorem</u> (Gupta **K**. Levin): LearnOrCover has a competitive ratio of $O(\log mn)$ for RO CIP.

Outline

Introduction

Prior Work

LearnOrCover (Warmup)

LearnOrCover in Polynomial Time

Some Intuition

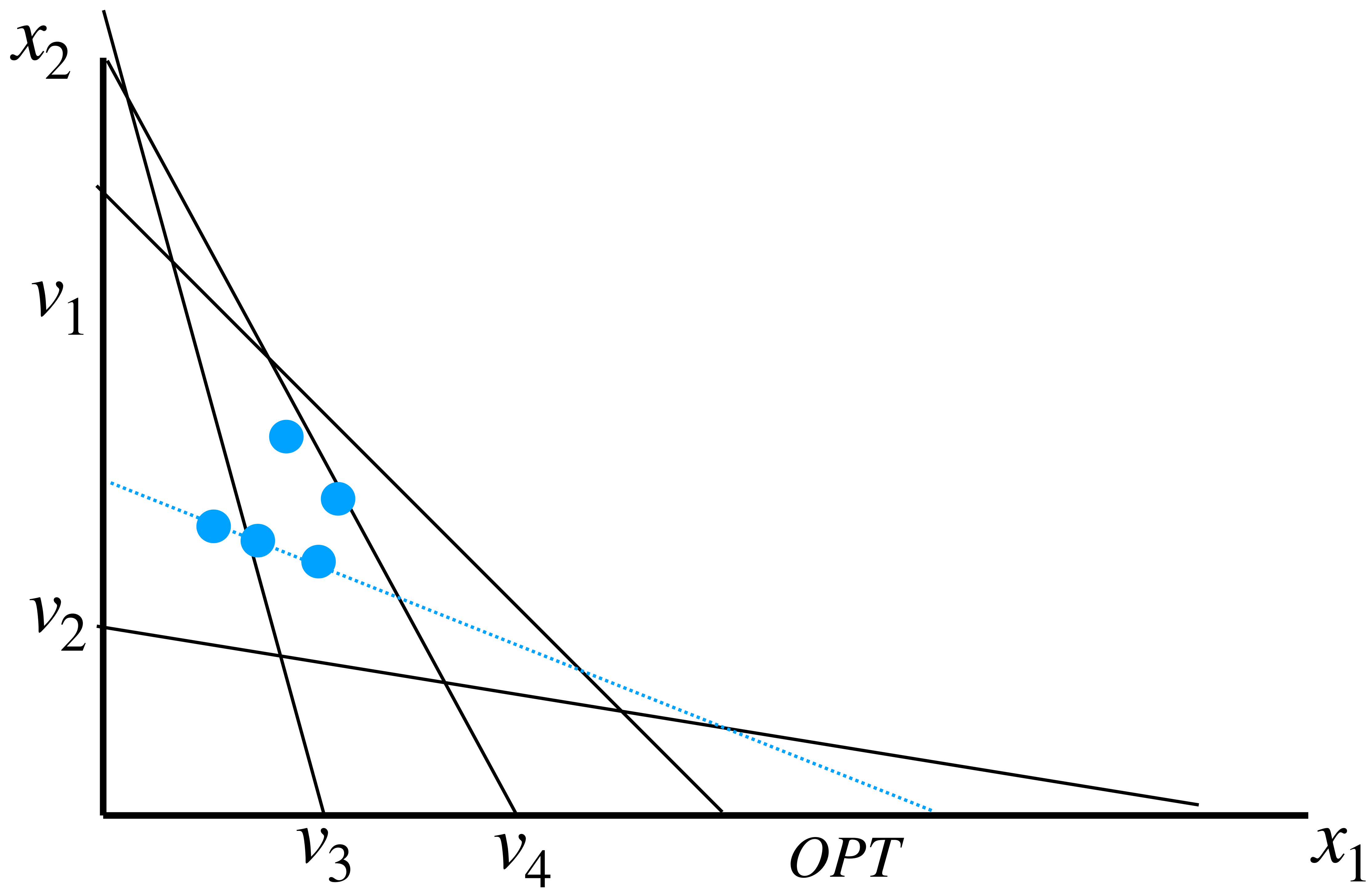
Lower Bounds

More Adversaries Beaten!

LearnOrCover: Two Informal Views

KL Projection

- One analysis of the primal-dual algorithm for adversarial order OSC casts it as iteratively performing a KL projection onto the feasible region.
- LoC does something similar, but renormalizs the weight of x^t.
- Is there a primal-dual interpretation of LoC?



LoC as Sample-Efficient Greedy

LoC (unit cost)

(estimate $k = |OPT|$)

initialize $x \leftarrow$ k$/m$

for v arriving uncovered (round t)

(Cover)

buy random $S \sim x$

(Learn)

if $x_v \leq (e-1)^{-1}$:

$x_S \leftarrow e \cdot x_S$ for all $S \ni v$

$$x \leftarrow k\frac{x}{\|x\|}$$

(Backup)

buy arbitrary $S \ni v$

Sample from dist. over current good coverage sets

learn which sets provide good coverage going forward

Greedy (offline)

$\mathscr{U}^0 \leftarrow [n]$

while there are v uncovered

buy $S \in \mathscr{S}$ maximizing $S \cap \mathscr{U}^t$

$\mathscr{U}^{t+1} \leftarrow \mathscr{U}^t \backslash S$

Can the distribution x be seen as maintaining a noisy estimate of which set provides the most marginal coverage?

Outline

Introduction

Prior Work

LearnOrCover (Warmup)

LearnOrCover in Polynomial Time

Some Intuition

Lower Bounds

More Adversaries Beaten!

Lower Bounds for ROSC

Two natural ways to relax random Order assumption:

Allow constraints to arrive in randomly-ordered batches:

Theorem (Gupta **K.** Levin): Batched RO set cover is $\Omega(\log b \log s)$ for b batches of size s

Relax the entropy of the distribution over arrival orders?

It quickly becomes easy to embed hard instance in the arrival sequence

…so what else is LearnOrCover good for?

The Landscape (again)

Instance

Arrival Order		Stochastic	Adversarial
	RO	$\Theta(\log mn)$ [Grandoni Gupta Leonardi Miettinen Sankowski Singh '08] constraints are i.i.d. samples from known $\mathscr{D}$	$\Theta(\log mn)$ [this talk] (secretary setting)
	Adversarial	$\Theta(\log mn)$? constraints are independent samples from known $\mathscr{D}_i$ (prophet setting)	$\Theta(\log m \ \log n)$ [Alon Awerbuch Azar Buchbinder Naor '03] [Korman '04]

> Q: What makes online integer covering (online set cover) harder than offline?

Outline

Introduction

Prior Work

LearnOrCover (Warmup)

LearnOrCover in Polynomial Time

Some Intuition

Lower Bounds

More Adversaries Beaten!

Online Set Cover **With a Sample**

Setting: online set cover/covering IPs, with the advantage that the algorithm observes a uniformly random fraction of the constraints at the outset.

<u>Theorem:</u> LearnOrCover can solve online set cover with an α sample with an expected competitive ratio of $O(1/\alpha\ \log mn)$.

Idea: run LearnOrCover on the sampled constraints in a random order. The potential Φ permits the cost of the adversarial portion to be charged to the sampled portion, in expectation.

1

Samples $\cdot\, v_8 \quad \cdot\, v_3 \quad \cdot\, v_2 \quad \cdot\, v_6 \quad \cdot\, v_7$

Arrivals $\cdot\, v_5 \quad \cdot\, v_8 \quad \cdot\, v_3 \quad \cdot\, v_1 \quad \cdot\, v_2 \quad \cdot\, v_6 \quad \cdot\, v_7 \quad \cdot\, v_4$

Sets Bought 2

Prophet Online Set Cover

Setting: online set cover/covering IPs, where each arriving constraint v_i is an independent sample from some known distribution $\mathscr{D}_i$

Theorem: LearnOrCover can solve prophet online integer covering with an expected competitive ratio of $O(\log mn)$.

Idea: Sample constraints from the $\mathscr{D}_i$ and run LearnOrCover on them. We can again make a coupling argument that charges the online constraints to the sampled ones, despite their arbitrary arrival order.

Draws $\cdot v_1' \quad \cdot v_2' \quad \cdot v_3' \quad \cdot v_4' \quad \cdot v_5' \quad \cdot v_6' \quad \cdot v_7' \quad \cdot v_8'$

1

$v_i, v_i' \sim \mathcal{D}_i$

Arrivals $\cdot v_1 \quad \cdot v_2 \quad \cdot v_3 \quad \cdot v_4 \quad \cdot v_5 \quad \cdot v_6 \quad \cdot v_7 \quad \cdot v_8$

Sets Bought

2

In conclusion

Arrival Order \ Instance	Stochastic	Adversarial
RO	$\Theta(\log mn)$ [Grandoni Gupta Leonardi Miettinen Sankowski Singh '08] constraints are i.i.d. samples from known $\mathscr{D}$	$\Theta(\log mn)$ [this talk] (secretary setting)
Adversarial	$\Theta(\log mn)$ [in prep] constraints are independent samples from known $\mathscr{D}_i$ (prophet setting)	$\Theta(\log m \ \log n)$ [Alon Awerbuch Azar Buchbinder Naor '03] [Korman '04]

Q: What makes online integer covering (online set cover) harder than offline?

A: Both having no foreknowledge of the instance, **and** facing it in adversarial order!

Thank you!

Questions welcome now or later: gkehne@g.harvard.edu