Learning and Strongly Truthful Multi-Task Peer Prediction

Grant Schoenebeck, University of Michigan
Fang-Yi Yu, Harvard University

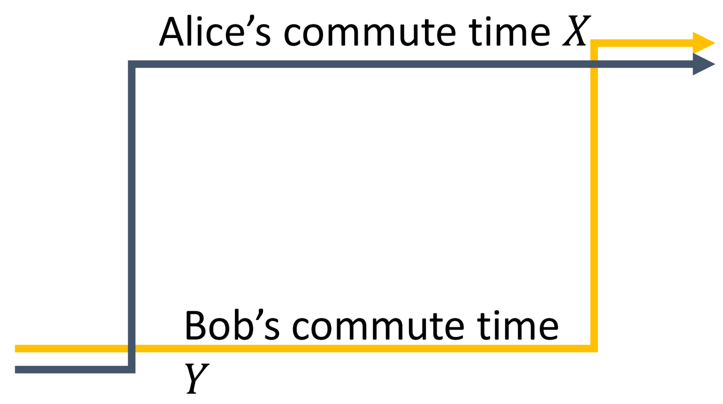
Motivation Questions

How can we design mechanisms to collect agents' truthful report without verification?

- Continuous signal
- Minimal assumption on agent's prior

Peer Prediction without Verification

People are connected and their signals are dependent



Can we design strongly truthful mechanisms?

- Bayesian Nash equilibrium
- Highest payment

Multi-task setting peer prediction (PP)

	signals	strategy	reports	payment
$P_{X,Y}$	$x_1, \dots, x_m$	$\xrightarrow{\theta_A}$	$\hat{\mathbf{x}} = \hat{x}_1, \dots, \hat{x}_m$	$M(\hat{\mathbf{x}}, \hat{\mathbf{y}})$
	$y_1, \dots, y_m$	$\xrightarrow{\theta_B}$	$\hat{\mathbf{y}} = \hat{y}_1, \dots, \hat{y}_m$	

- a prior similar tasks
- task independent strategy

Data processing inequality and PP

Strategy as a noisy channel

- $Y \xrightarrow{P_{X|Y}} X \xrightarrow{\theta_A} \hat{X}$ Is a Markov chain

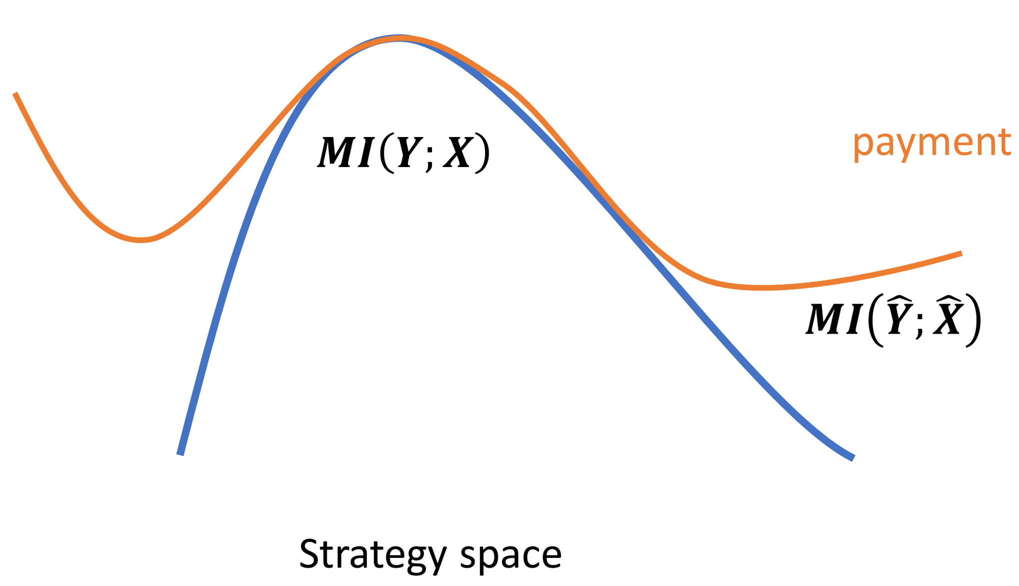
Mutual information

- Shannan: $MI(X;Y) = \int \log \frac{dP_{XY}}{dP_X P_Y} dP_{XY}$
- Φ-divergence:

$$MI^{\Phi}(X;Y) = \int \Phi\left(\frac{dP_{XY}}{dP_X P_Y}\right) dP_X P_Y$$

- Data processing inequality

$$\underbrace{MI^{\Phi}(Y;X)}_{\text{Truth-telling}} \geq MI^{\Phi}(Y;\hat{X}) \geq \underbrace{MI^{\Phi}(\hat{Y};\hat{X})}_{\text{Non-truthful}}$$



Estimate mutual information

Given Alice and Bob reports on m tasks, how can we estimate mutual information between their reports?

$\hat{x}_1$	$\hat{x}_2$	…	$\hat{x}_b$	…	$\hat{x}_p$	…	$\hat{x}_q$	…	$\hat{x}_m$
$\hat{y}_1$	$\hat{y}_2$	…	$\hat{y}_b$	…	$\hat{y}_p$	…	$\hat{y}_q$	…	$\hat{y}_m$

Plug-in estimation

- $\hat{P}_{X,Y}$ from a pair of reports on a common task.
- $\hat{P}_X \hat{P}_Y$ from a pair of reports on distinct tasks.

Accuracy and error

- No unbiased estimator
- Uniform error upper bound

Variational Statistics to Mechanisms

Convex conjugate

- $\Phi(a) = \sup_b ab - \Phi^*(b)$ with maximum at $b = \Phi'(a)$.
- $MI^{\Phi} = \sup_K \left\{\mathbb{E}_{P_{X,Y}}[K(x,y)] - \mathbb{E}_{P_X P_Y}[\Phi^* K(x,y)]\right\}$ with maximum at $K^*(x,y) = \Phi'\left(\frac{dP_{XY}}{dP_X P_Y}(x,y)\right)$

Ideal Scoring function

Pairing mechanism

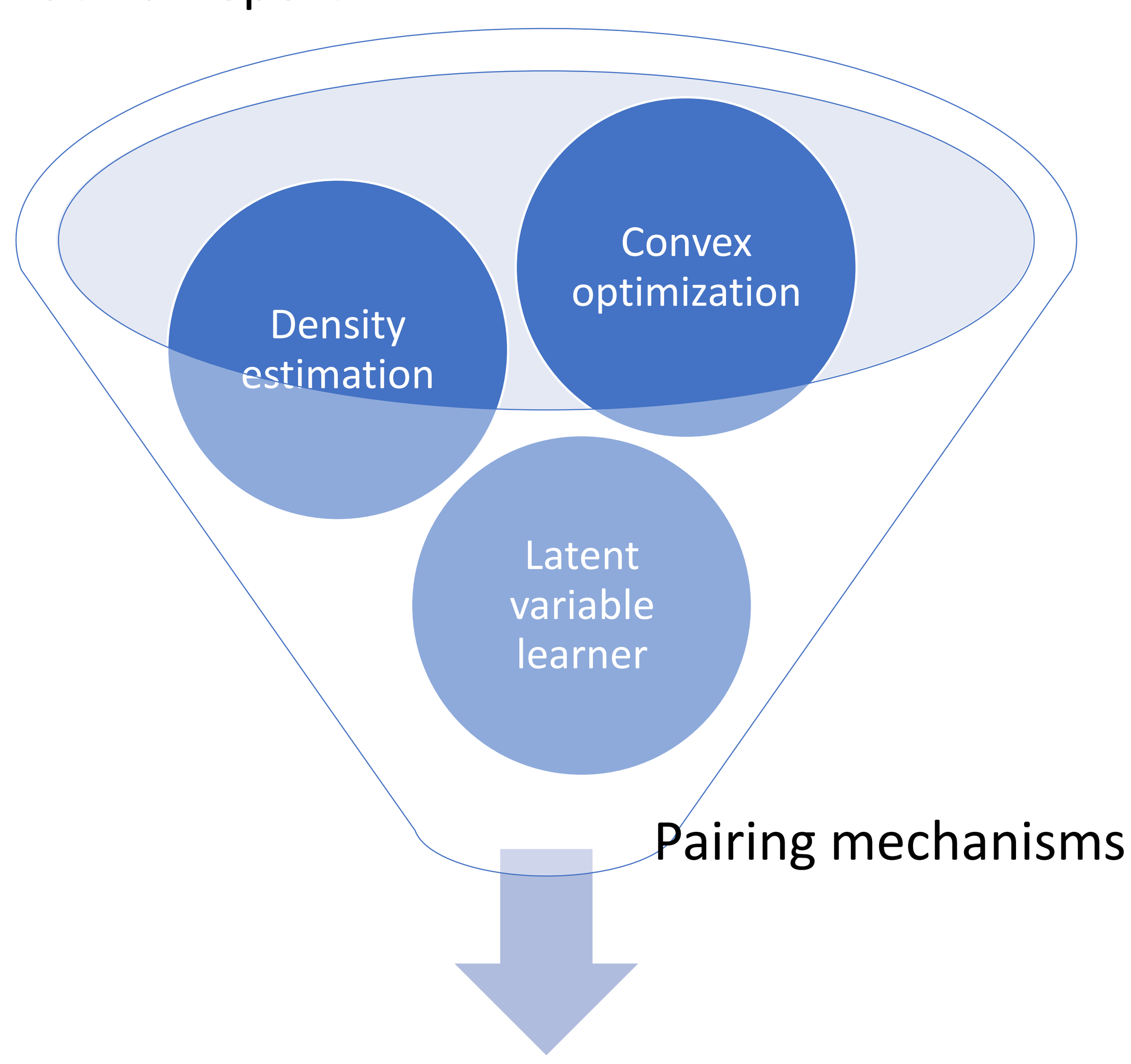
1. Estimate K^* from learning tasks.
2. Sample $P_{X,Y}$ from a pair of reports on a common task.
3. Sample $P_X P_Y$ from a pair of reports on distinct tasks.
4. Pay $K^*(x_b, y_b) - \Phi^*\left(K^*(x_p, y_q)\right)$

x_1	x_2	…	x_b	…	x_p	…	x_q	…	…	…	x_m
y_1	y_2	…	y_b	…	y_p	…	y_q	…	…	…	y_m

Tasks for payment — Tasks for learning

Maximum happens only if both

- Ideal scoring function $\hat{K} = K^* = \Phi'\left(\frac{dP_{XY}}{dP_X P_Y}\right)$
- Truthful report



Approximated Strongly truthful