Robust Traceability of Trace Amounts

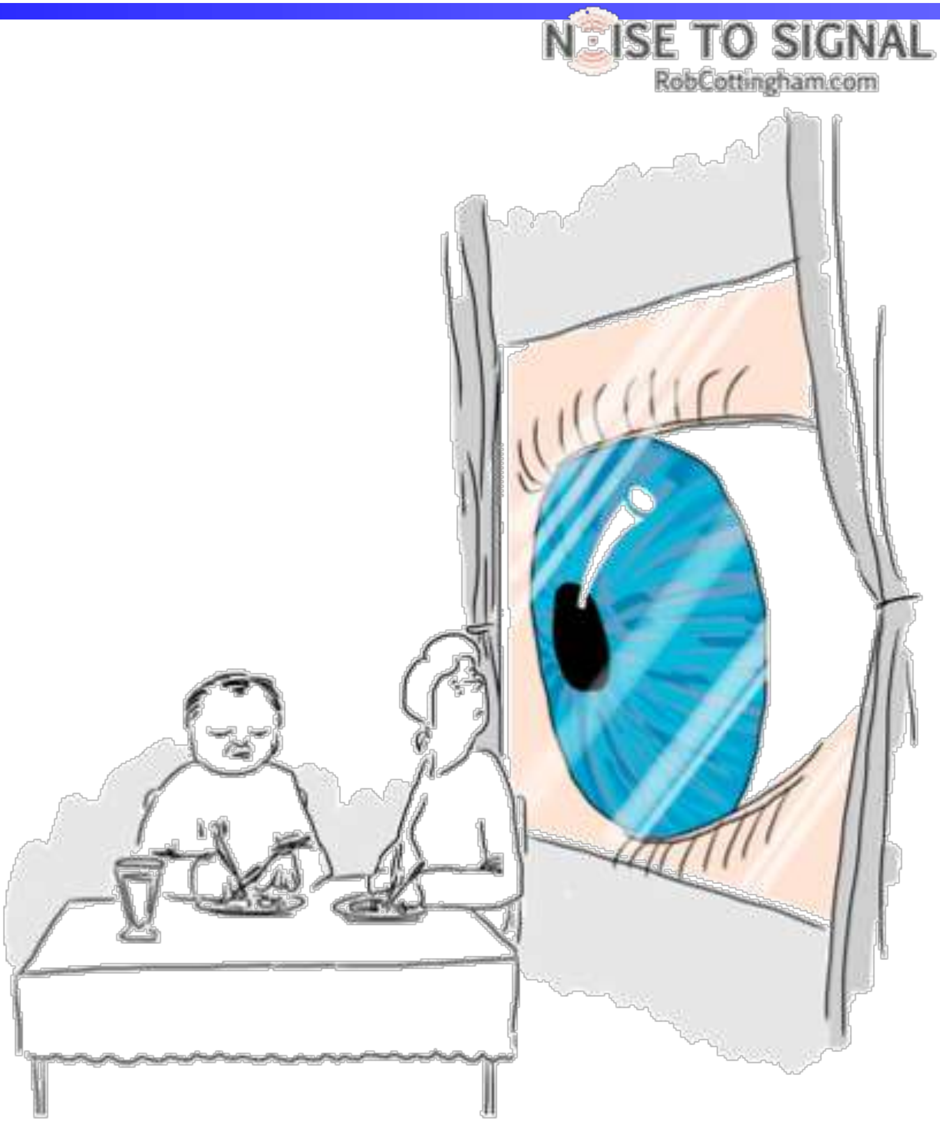
Cynthia Dwork · Adam Smith · Thomas Steinke · Salil Vadhan · Jonathan Ullman

Microsoft · Penn State · Harvard · Northeastern

PennState
College of Engineering

Thanks to Chen Sun
for help with data

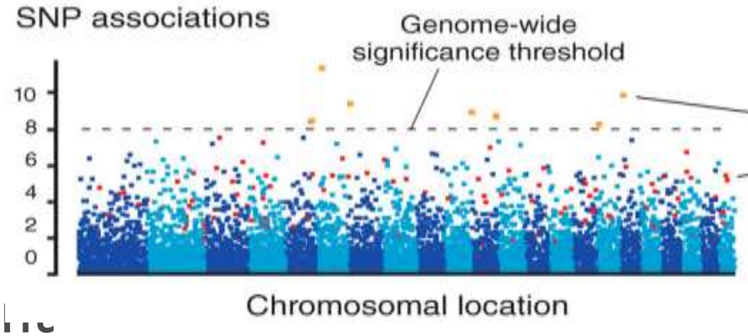
Sometimes summaries reveal a lot



Relax — it can only see metadata.

Sometimes summaries reveal a lot

- [Homer et al. (2008)] showed exact high-dimensional summaries allow an attacker with knowledge of population to test membership in a data set
  - ➢ Can also find out whether participant was case or control, or…
  - ➢ Not specific to genetic data



- This paper: strengthened membership tests
  - ➢ Approximate statistics
  - ➢ Less side information

This talk

- Background
- An abstract setting
- Results

Abstract setting

Two applications

- Deanonymization
- Forensics

- Data X : $x_1, x_2, \ldots, x_n \in \{0,1\}^d$
  - d binary attributes for each person
  - Think: d big and n moderate
- Summary statistcs
  - Column averages $\bar{x}(j) = \sum_i x_i(j)$, for $j = 1, \ldots, d$.
- Actual output
  - Estimates $q(j) \in \bar{x}(j) \pm \alpha$
- Goal:

given q and a "target person" $z \in \{0,1\}^d$,
determine if $z \in X$

Impossible without some assumptions

- Assumptions:
  - $x_1, \ldots, x_n$ i.i.d. from distribution P
  - Attributes are independent
    - $P = P_\mu$ is a described by vector $\mu_1, \ldots, \mu_d$

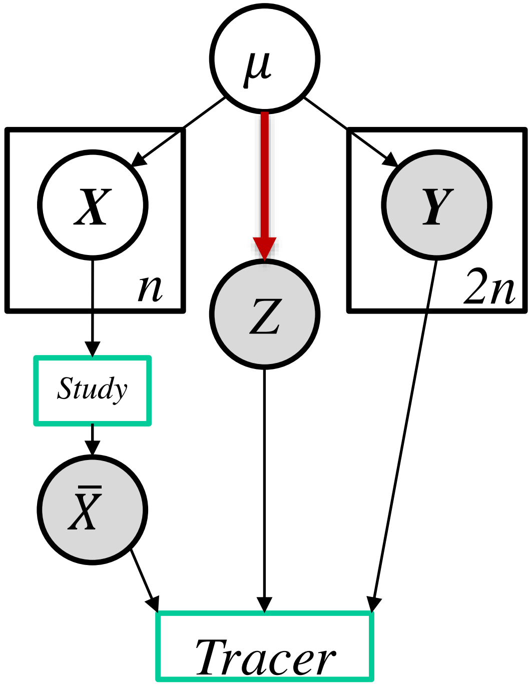
$$E_{X \sim P_\mu}(X) = \vec{\mu}$$

  - Z either uniform in sample $\boldsymbol{X}$ or fresh from P

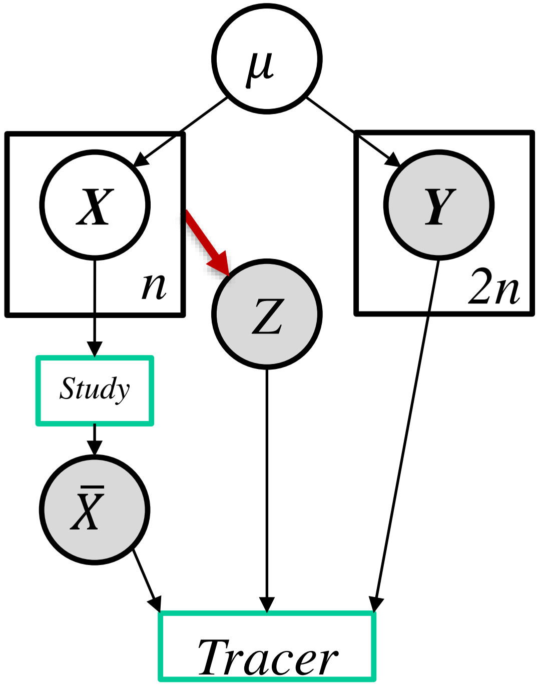
Relation to Previous work

- Membership tests [Sankararaman et al., Nature Genomics 2009] assume
  - ➢ Exact statistics are published ($\alpha = 0$)
  - ➢ Nearly-exact knowledge of distribution
- Fingerprinting codes [Tardos 2003, Bun, Ullman, Vadhan 2014, Steinke, Ullman 2015] assume
  - ➢ Robust to perturbed statistics ($\alpha < 1/2$)
  - ➢ Artificial distribution, exactly known
- This work
  - ➢ Robust to perturbation: analysis for arbitrary $\alpha < 1/2$
    - Same test works for all perturbation mechanisms
    - Mathematically, very different from "normal" hypothesis testing
  - ➢ Limited side information
    - Reference sample of size $m \geq 1$ from the population
- Related: Heuristic attacks using more complex statistics [Wang, Li, Wang, Tang, Zhou 2009]

Graphical Model

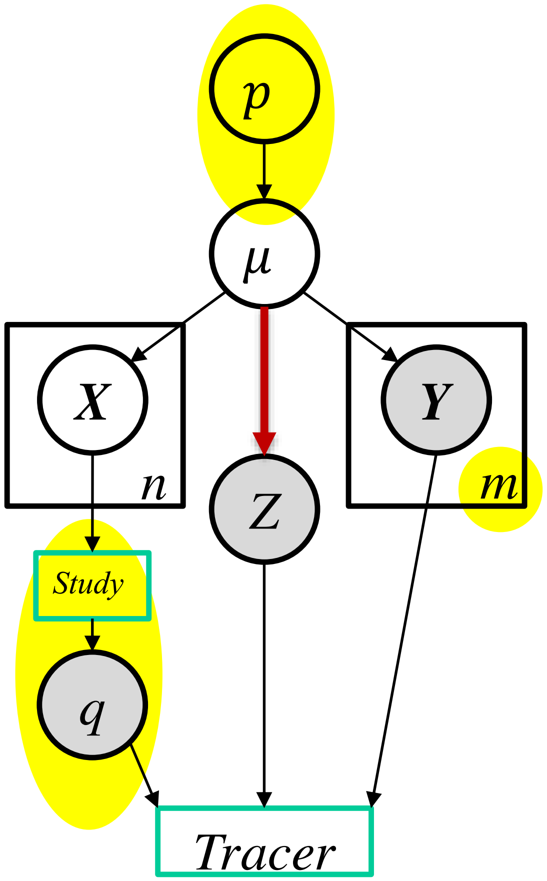


"Out"

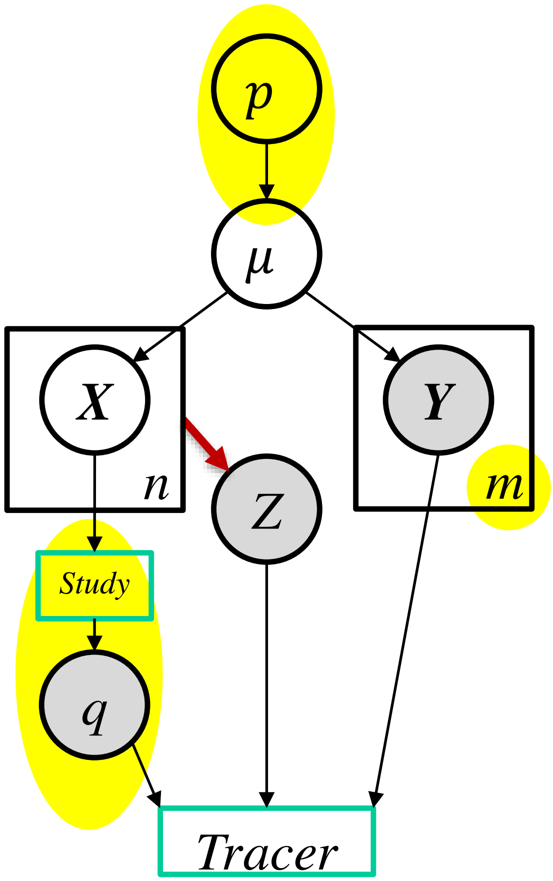


"In"

Graphical Model: This Work



"Out"



"In"

Comparison

	Previous work	This work
Tracer knowledge about P_p	Exact parameters or large sample from P ($2n$ points)	$m \geq 1$ fresh samples from P
Mechanism	$q(\boldsymbol{X}) = \bar{X}$	$q(j) \in \bar{X}(j) \pm \alpha$ (for α constant) and $\mu_i \sim p_i$ where the p_i are "smooth" (e.g. uniform, Lipschitz differentiable density)
Dimension of released data	$d > n$	$d > n + \alpha^2 n^2 + n^2/m$
Success probability (max of FP and FN rates)	$1 - \exp\left(C\frac{d}{n}\right)$	$1 - exp\left(-C\frac{d}{n+\alpha^2 n^2+n^2/m}\right)$ if we assume q depends only on $\bar{x}$
		$\Omega(\alpha^2)$ in general

- *Simple test; same test works in many settings*
- *Matches asymptotic accuracy of differentially private release: $\alpha \approx \sqrt{d}/(\epsilon n)$ so $d \approx \alpha^2(\epsilon n)^2$*

Tracing algorithm

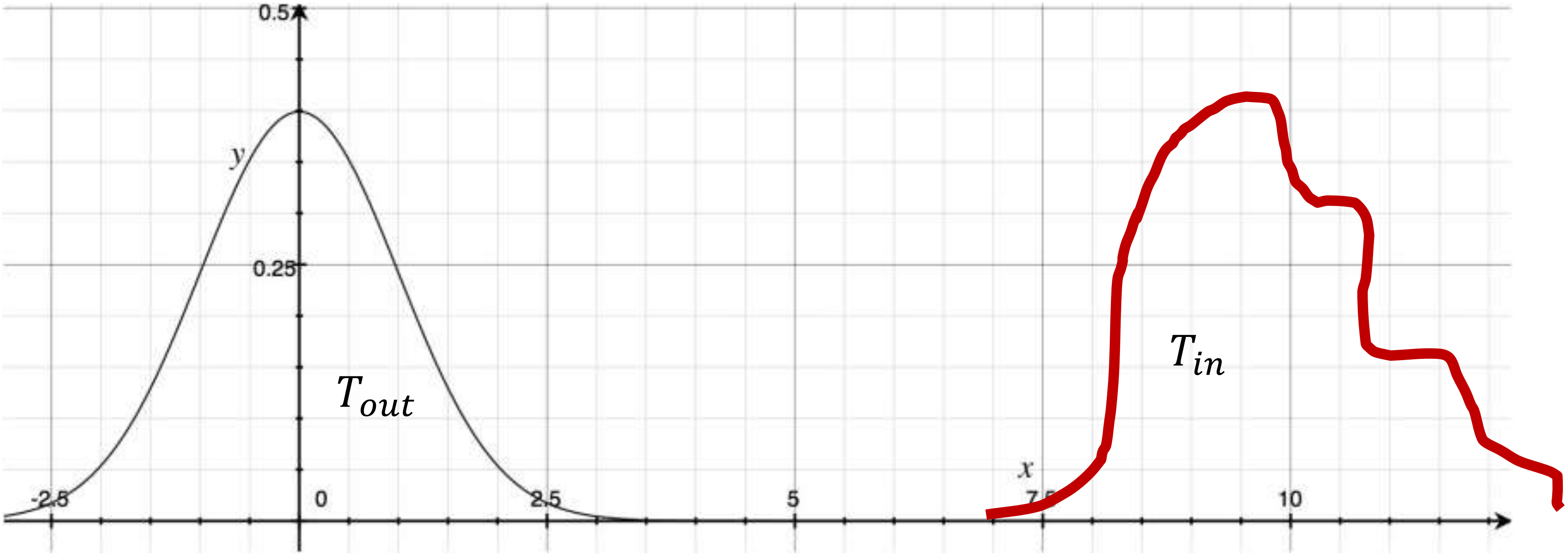
- Given $q \in [0,1]^d$ and $z, y_1, \ldots, y_m \in \{0,1\}^d$ and $\delta > 0$
  - ➢ Compute
    $$T = \langle z - y_1, q - \overline{y_{-1}} \rangle$$
  - ➢ If $T > 3\alpha\sqrt{d \log(1/\delta)}$, return "In"
    Else return "Out"

- **Theorems** [see paper]: Under various conditions,
  $\Pr(\mathit{Tracer\ says\ ``In"} \mid \mathit{OUT}) < \delta$, and
  $\Pr(\mathit{Tracer\ says\ ``In"} \mid \mathit{IN}) > 1 - \exp(\ldots)$.

Previous work: Likelihood ratio test

$$T \approx \left\langle z \ , \log\left(\frac{q_j}{1-q_j}\right) - \log\left(\frac{p_j}{1-p_j}\right) \right\rangle$$

Proof Idea

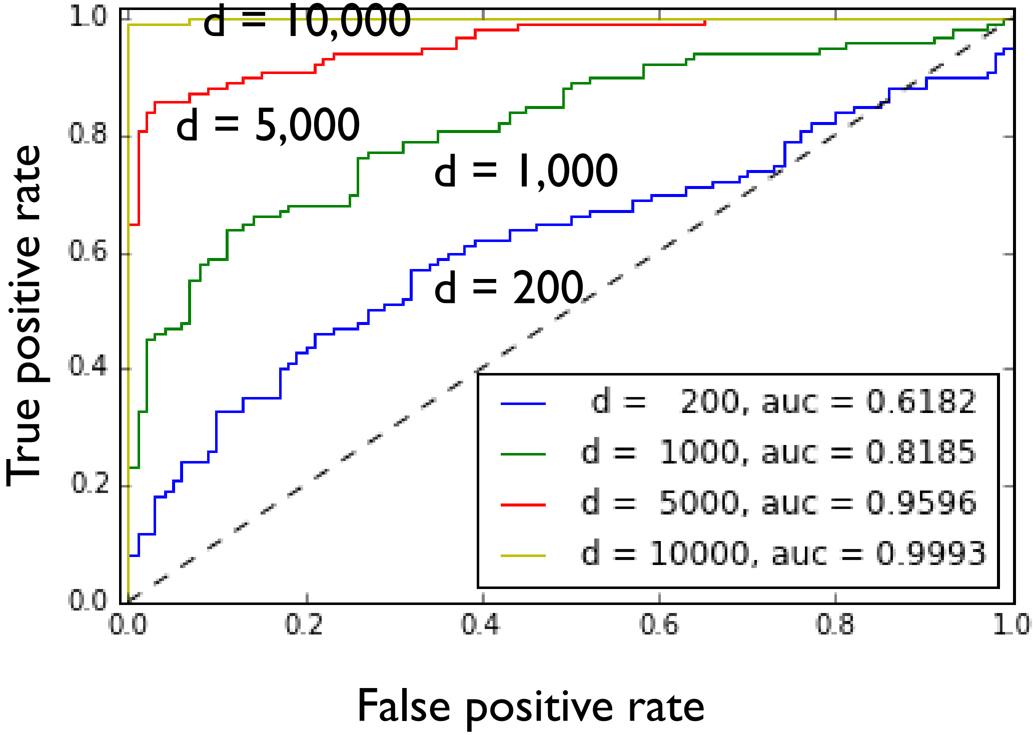


This talk

- Background
- An abstract setting
- Results

Increasing the dimension

- Simulated data
  - Independent columns ("linkage equilibrium")
- Means drawn from actual distribution on allele frequencies (Hapmap CEU)
  - Following set up from Sankararaman et al.

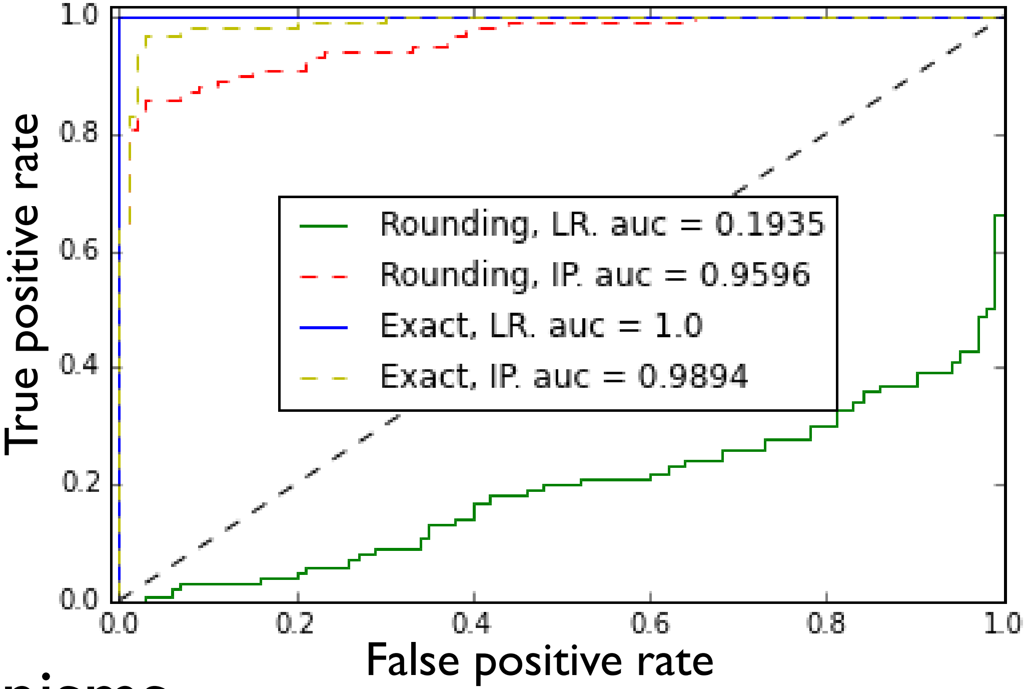


- $n = 100$
- $m = 200$
- Published statistics rounded down to multiple of 0.1

Conclusion: Results fit roughly to theory

Robustness to perturbation

- $n = 100$
- $m = 200$
- $d = 5{,}000$
- Two tests
  - ➢ LR [Sankararam et al]
  - ➢ IP [this work]

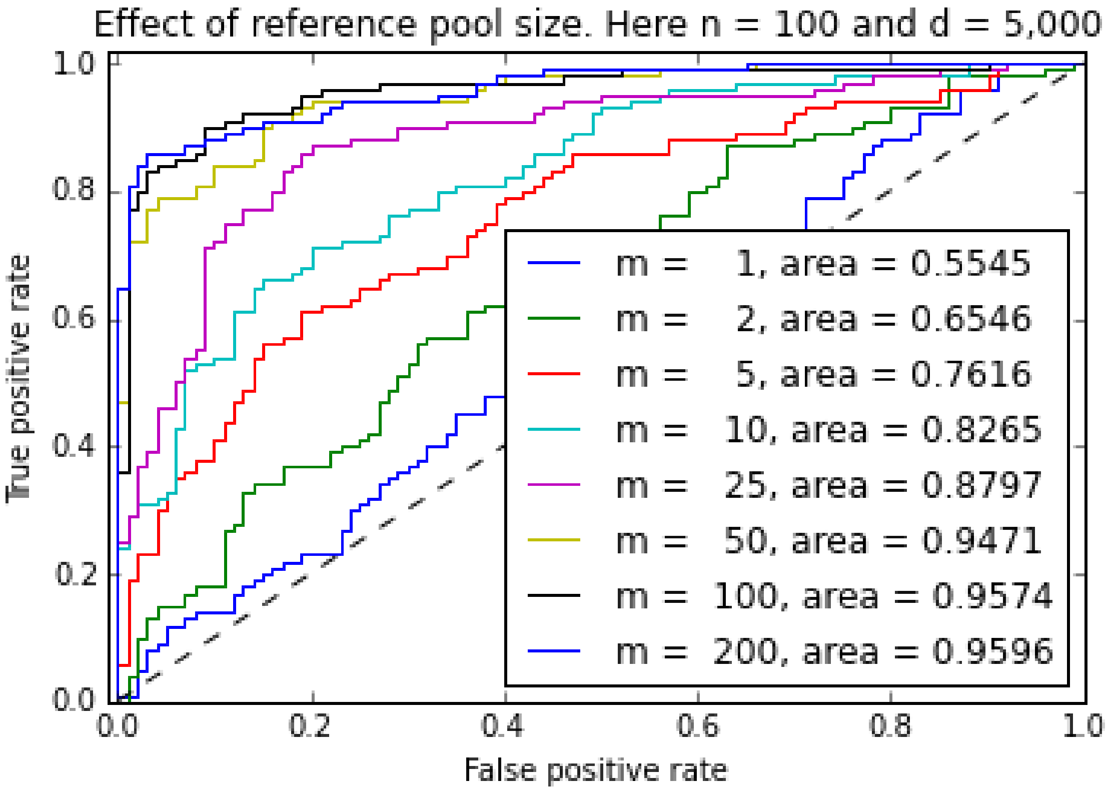


- Two publication mechanisms
  - ➢ Rounded to nearest multiple of 0.1 (red / green)
  - ➢ Exact statistics (yellow / blue)

Conclusion: IP test is robust.
Calibrating LR test seems difficult

Shrinking the reference pool

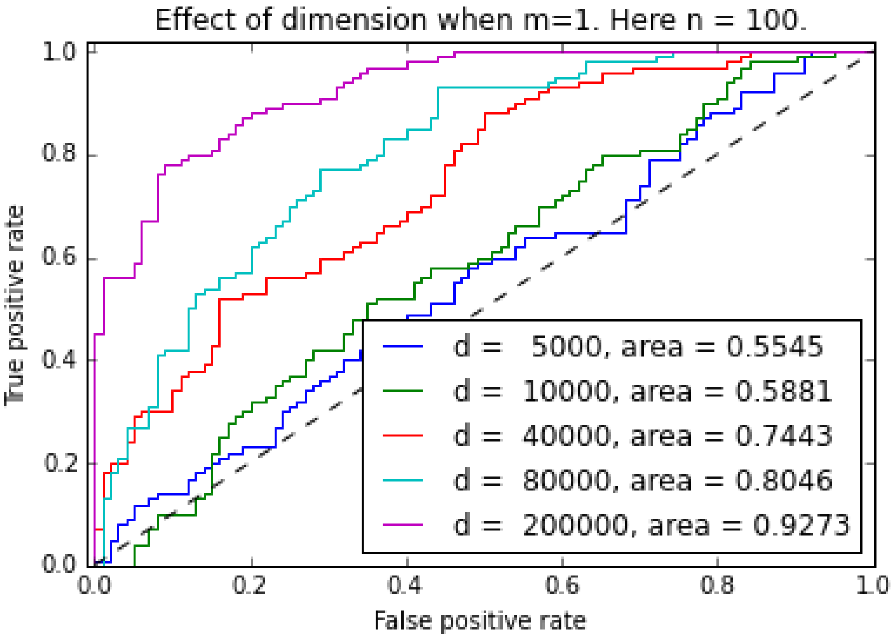
- Rounding to 0.1
- $n = 100$ and $d = 5{,}000$
- Get reliable signal for m above about 25

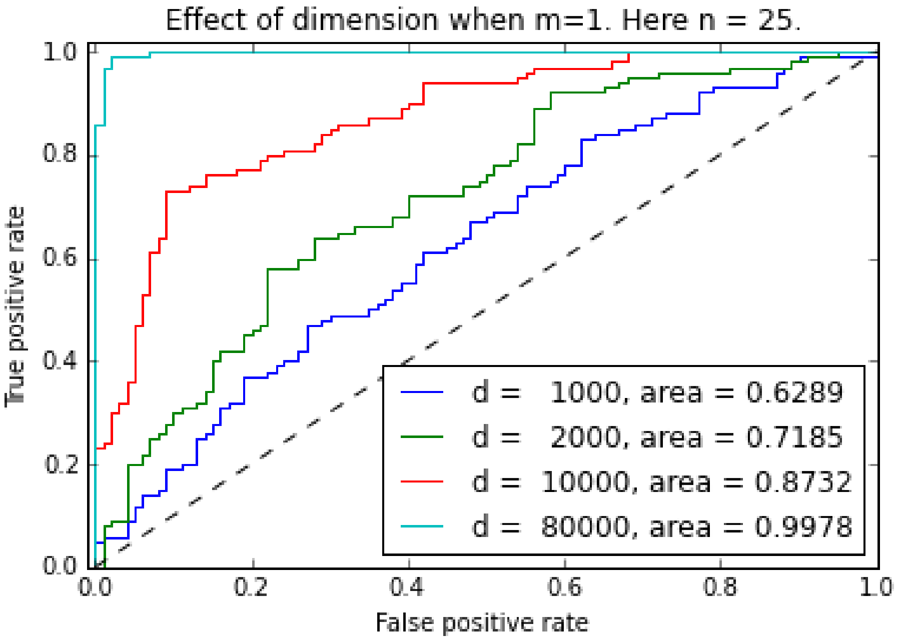


Effect of reference pool size. Here n = 100 and d = 5,000

What happens when m = 1?

- Here $n = 100$ and $m = 1$
- Mechanism rounds down to multiples of 0.1
- Still get a reliable signal for individual’s presence
  - ➢ As predicted, much larger dimension is necessary





Future Work

- Real data
- Optimal test
  - Application: calibrating competitions
- Other types of statistics
  - Preliminary results on pairwise frequencies

Bigger questions

- How common are these problems “in the wild”?
- How should policies adjust?