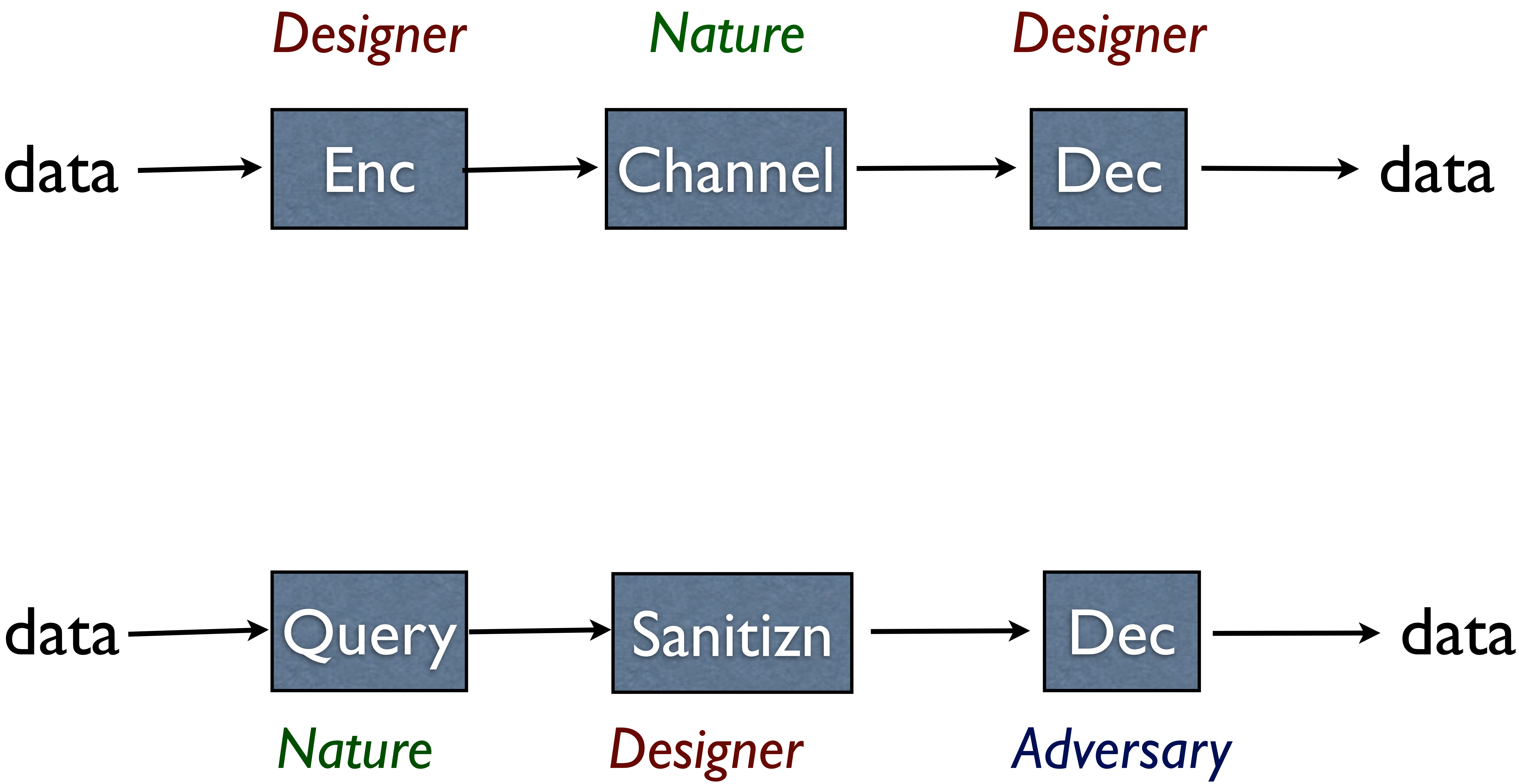
Extremal Mechanisms in Differential Privacy

Quan Geng, Sewoong Oh, Pramod Viswanath,

University of Illinois

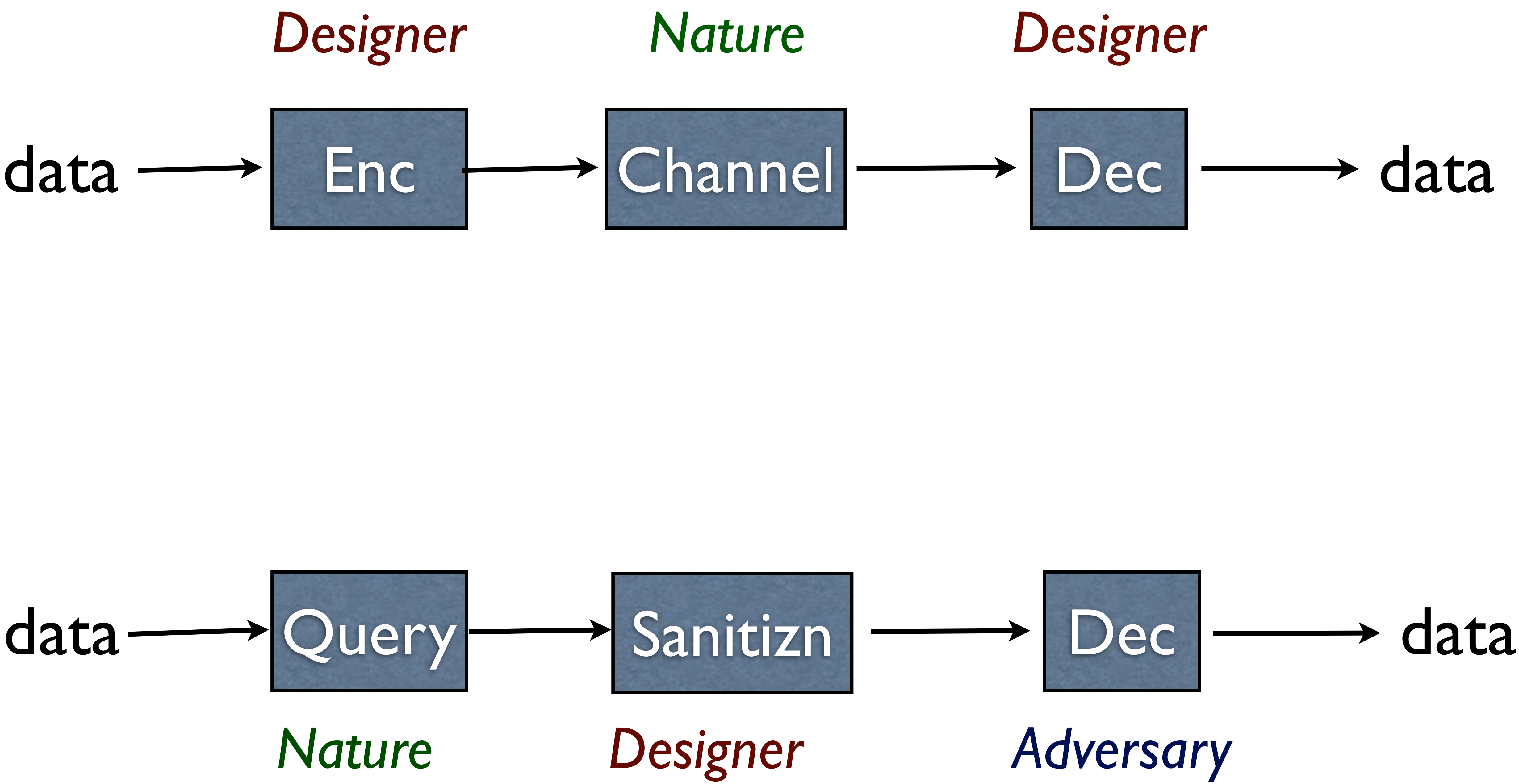
December 12, 2013

Information Theory and Differential Privacy



- Communication -- small error probability
- Privacy -- large error probability

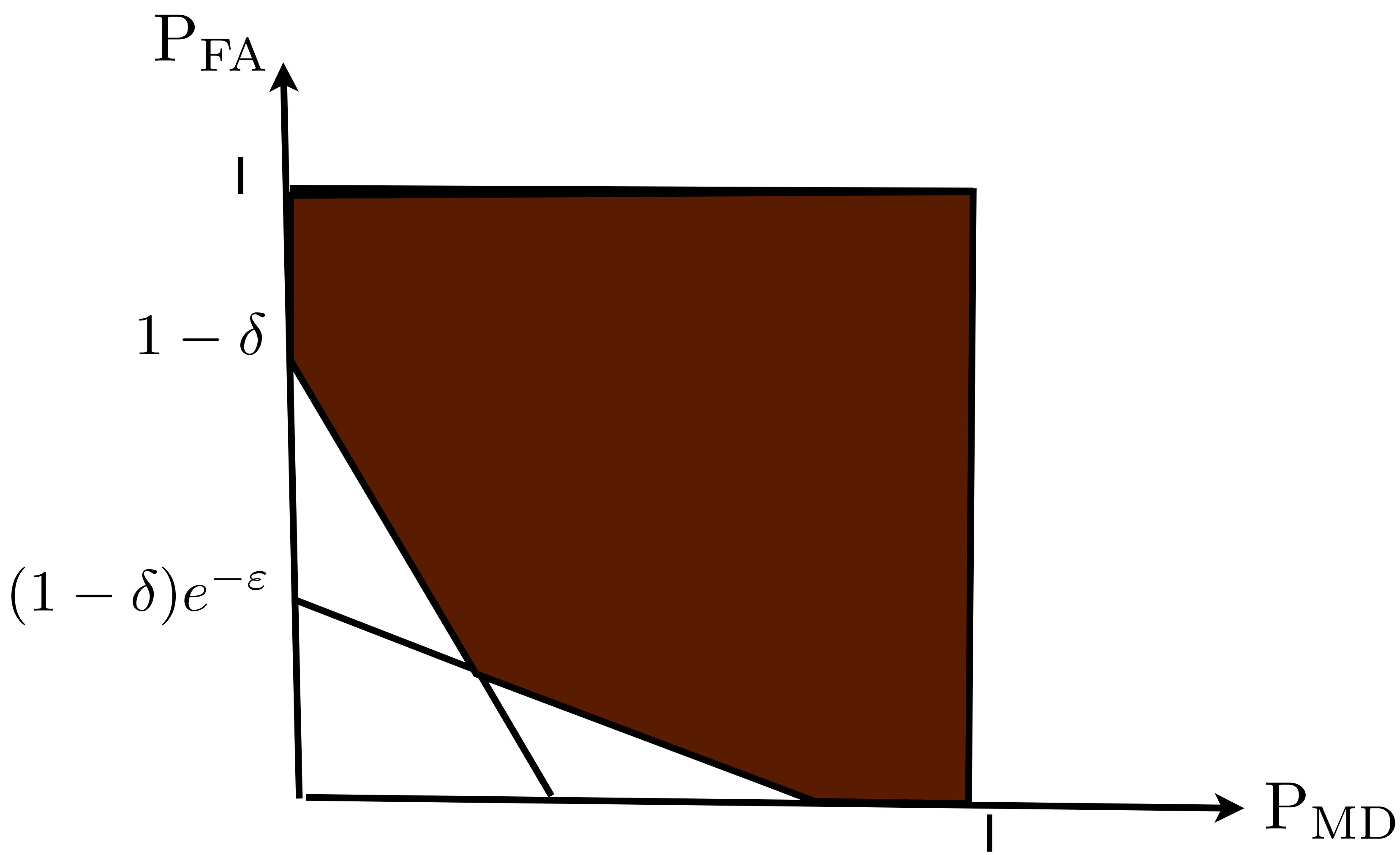
Information Theory and Differential Privacy



- Communication -- multi hypothesis testing
- Privacy -- binary hypothesis testing

Binary Inference Errors

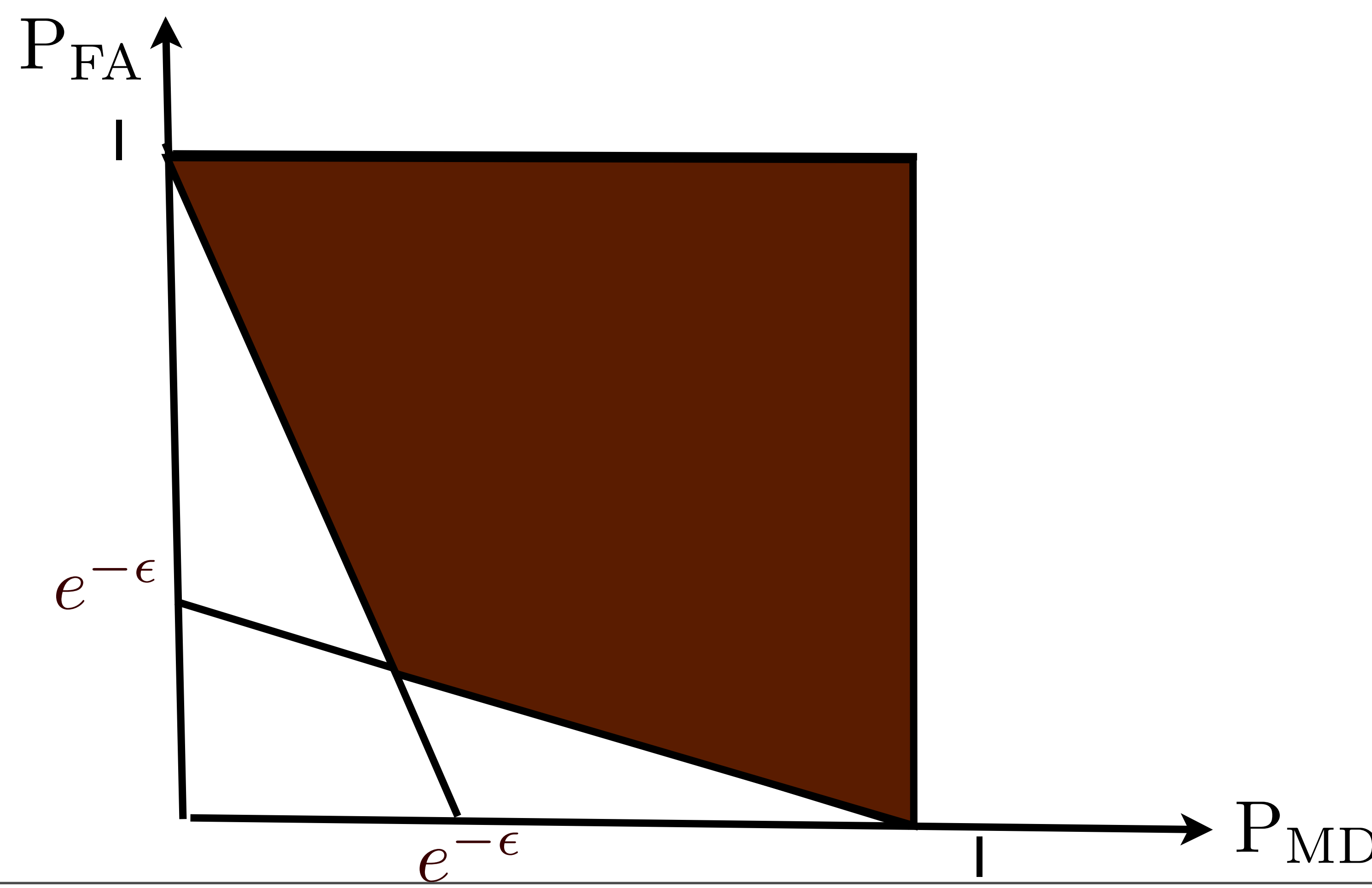
- Two error types
  - False Alarm and Missed Detection
- Privacy: guarantee enough error



-

Differential Privacy

- A specific way of enforcing inference errors
  - WZ11
- Original formulation involves likelihood ratios
  - DKMNS05
- ϵ controls privacy level



-

Differential Privacy

- For competing hypotheses D1 and D2

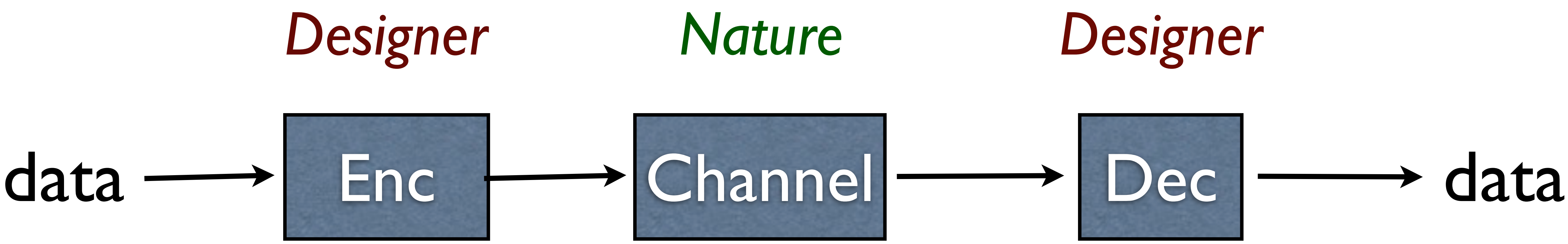
$$e^{-\epsilon} \leq \frac{\Pr(K(D_1) \in S)}{\Pr(K(D_2) \in S)} \leq e^{\epsilon}$$

- Equivalently:

$$P_{\mathrm{MD}} + e^{-\epsilon} P_{\mathrm{FA}} \geq e^{-\epsilon}$$
$$P_{\mathrm{FA}} + e^{-\epsilon} P_{\mathrm{MD}} \geq e^{-\epsilon}$$

- Likelihood ratios in a bounded interval
- ϵ small is high privacy
- ϵ large is low privacy

Information Theory is Mature



- Shannon, 1948
  - A mathematical theory of communication
- Success
  - extremal limits
    - capacity, single-letter expressions
    - fundamental benchmarks
    - practical schemes
  - operational interpretation
    - data processing inequalities

This Talk

- Similar program for differential privacy
  - extremal mechanisms
  - fundamental limits
  - operational interpretation

- Results
  - Staircase mechanism
    - universally optimal noise adding mechanism
  - Optimal Composition theorems
  - Abstract Staircase mechanism
    - dominates every other privacy mechanism

State of the Art

- Noise adding mechanisms
- Real valued query
  - Laplacian noise
    - regular differential privacy
  - Gaussian noise
    - approximate differential privacy
- No exact optimality results

State of the Art

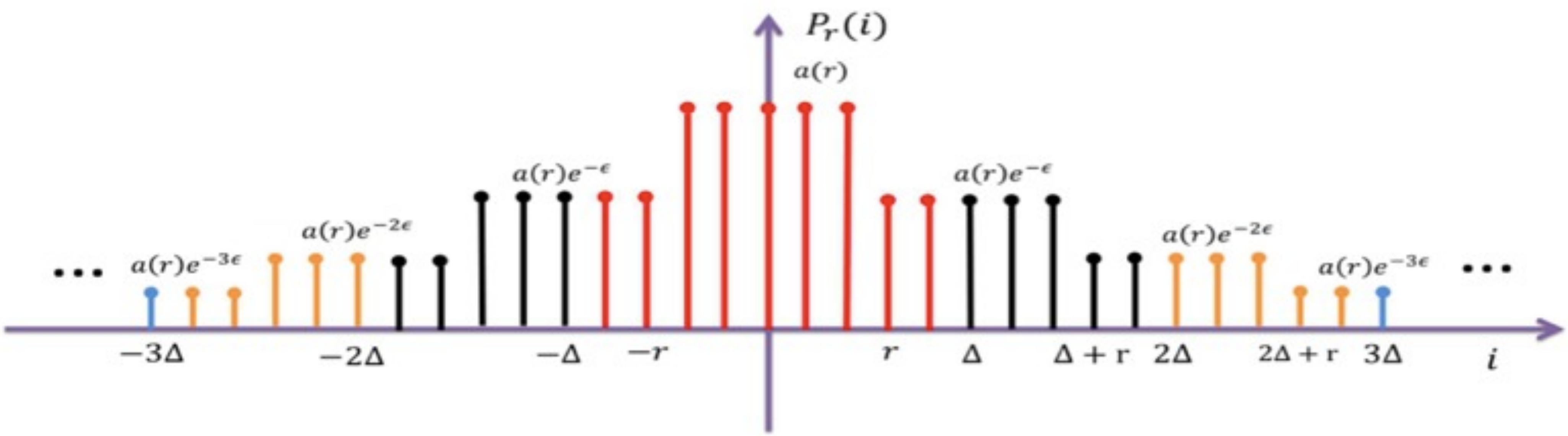
- Integer valued query
- Count queries (sensitivity is one)

- Geometric noise added
  - universal optimality in Bayesian cost minimization framework [GRS09]
  - no natural generalization
    - larger sensitivity [GS10]

- No operational interpretation
  - Hint: Log Likelihood ratio $\in \{-\varepsilon, +\varepsilon\}$

Staircase Mechanism

- Universally optimal noise adding mechanism
  - worst case setting
  - generalization of GRS09 $(\Delta = 1)$



  - no operational interpretation
    - Log Likelihood ratio $\in \{-\varepsilon, 0, +\varepsilon\}$

Example Cost Functions

- Privacy mechanism involves adding noise

$$K(D) = q(D) + X$$

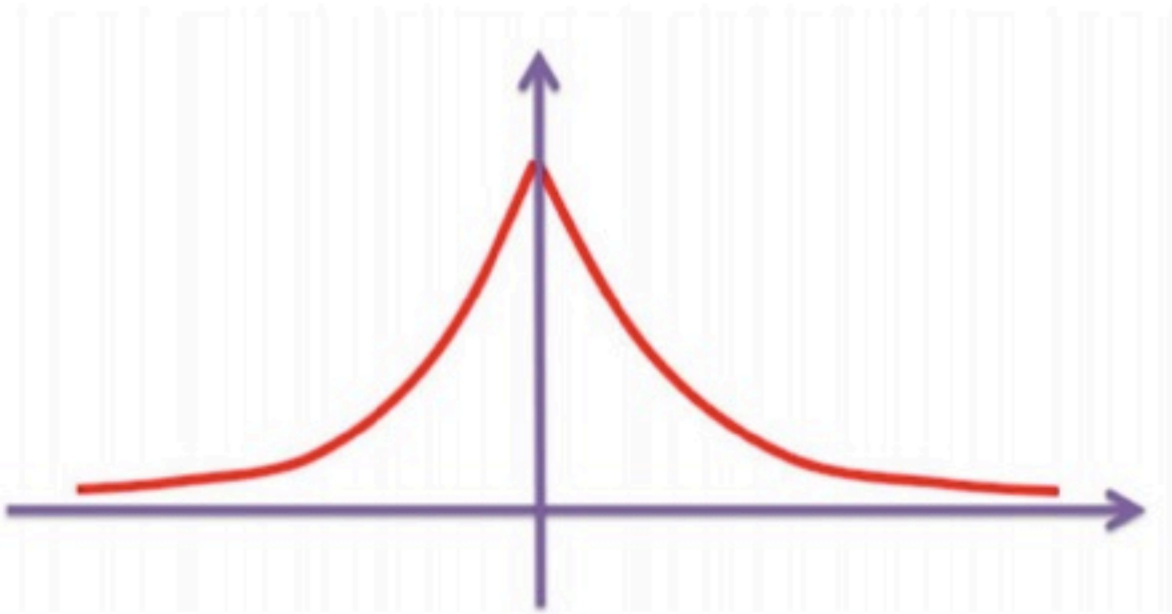
  - amplitude of noise $E[|X|]$ $L(x) = |x|$
  - variance of noise $E[X^2]$ $L(x) = x^2$

- In general any cost function
  - monotonically increasing
  - symmetric around origin

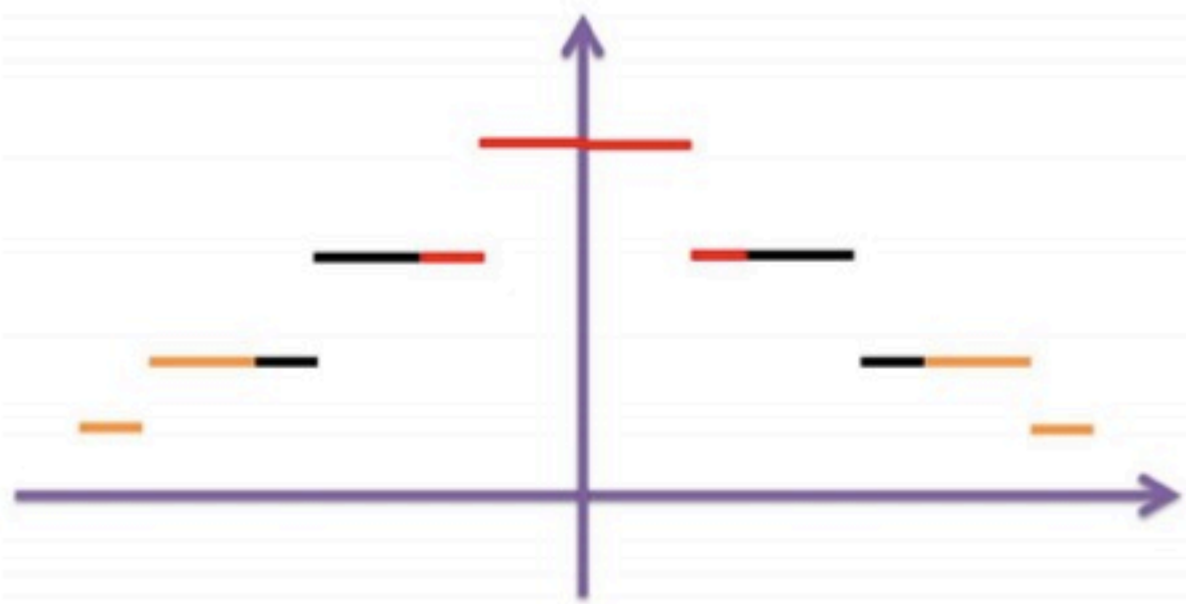
- $\min \quad E[L(X)]$

Universal Optimality

- Theorem: Optimal Noise is Staircase shaped



(a) Laplace Mechanism

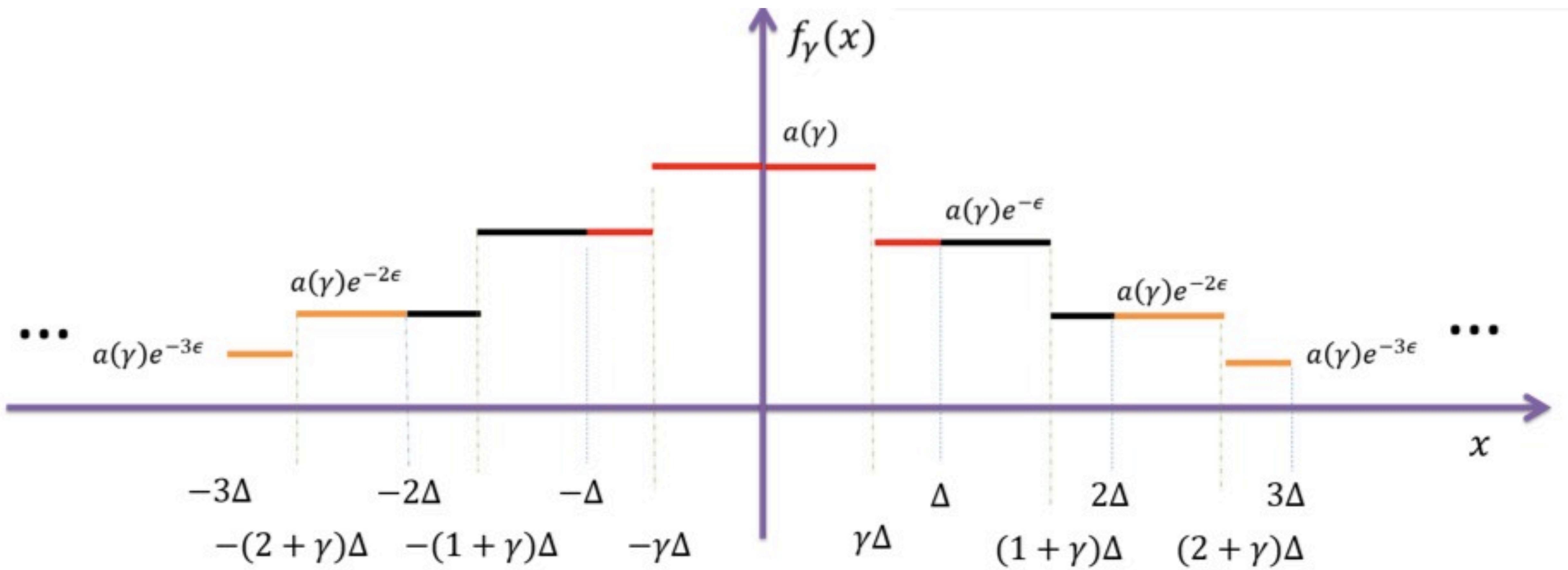


(b) Staircase Mechanism

- Geometric mixture of uniform random variables

Staircase Mechanism

- Theorem: Optimal Noise is universally Staircase shaped



- Geometric decaying
  - $\gamma \in [0, 1]$ depends on cost function

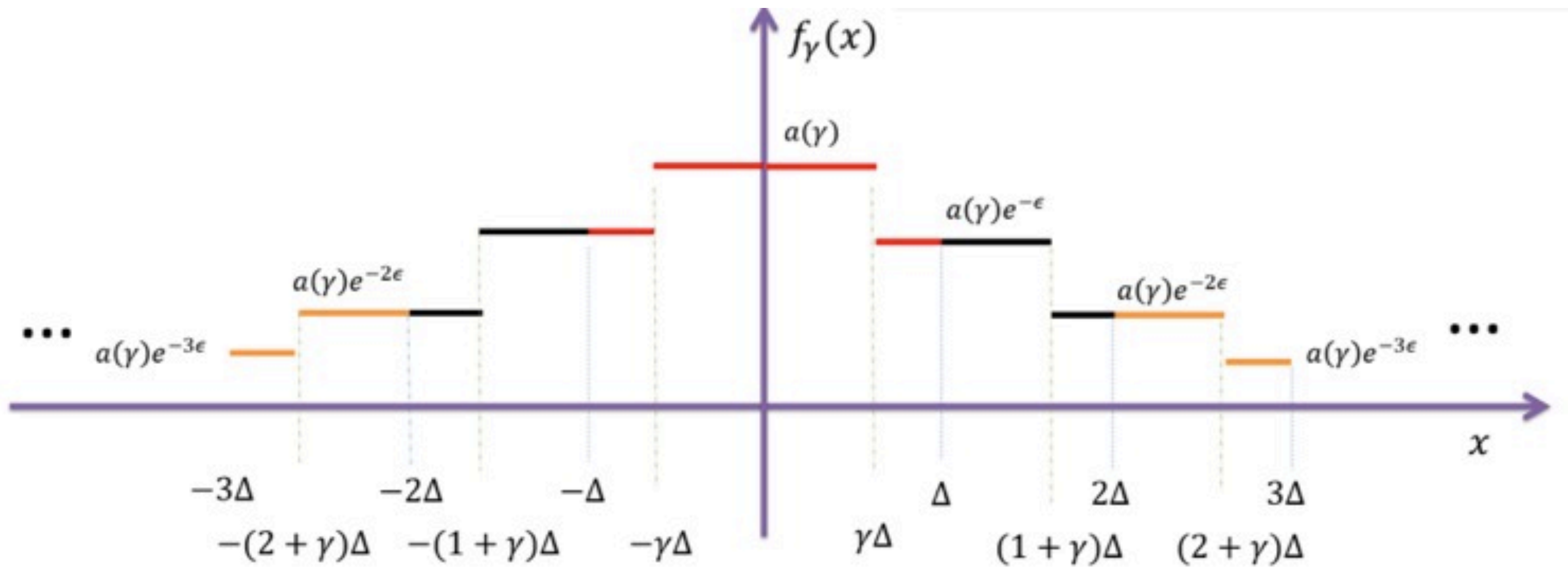
Price of Privacy

- For $L(x) = |x|$
- Minimum noise magnitude $\frac{\Delta e^{-\varepsilon/2}}{1-e^{-\varepsilon/2}}$
- Laplace noise magnitude $\frac{\Delta}{\varepsilon}$
- High privacy
  - gap is small
- Low privacy
  - exponential improvement
- Low privacy costs exponentially less

Price of Privacy

- For $L(x) = x^2$
- Minimum noise variance $\Theta(\frac{\Delta^2 e^{-2\varepsilon/3}}{(1-e^{-\varepsilon})^2})$
- Laplace noise variance $\frac{\Delta^2}{\varepsilon^2}$
- High privacy
    - gap is small
- Low privacy
    - exponential improvement
- Low privacy costs exponentially less

Properties of γ^*



- Need to pick γ^*; depends on cost function
- General Properties:

$$\gamma^* \to \frac{1}{2} \qquad \epsilon \to 0$$

$$\gamma^* \to 0 \qquad \epsilon \to \infty$$

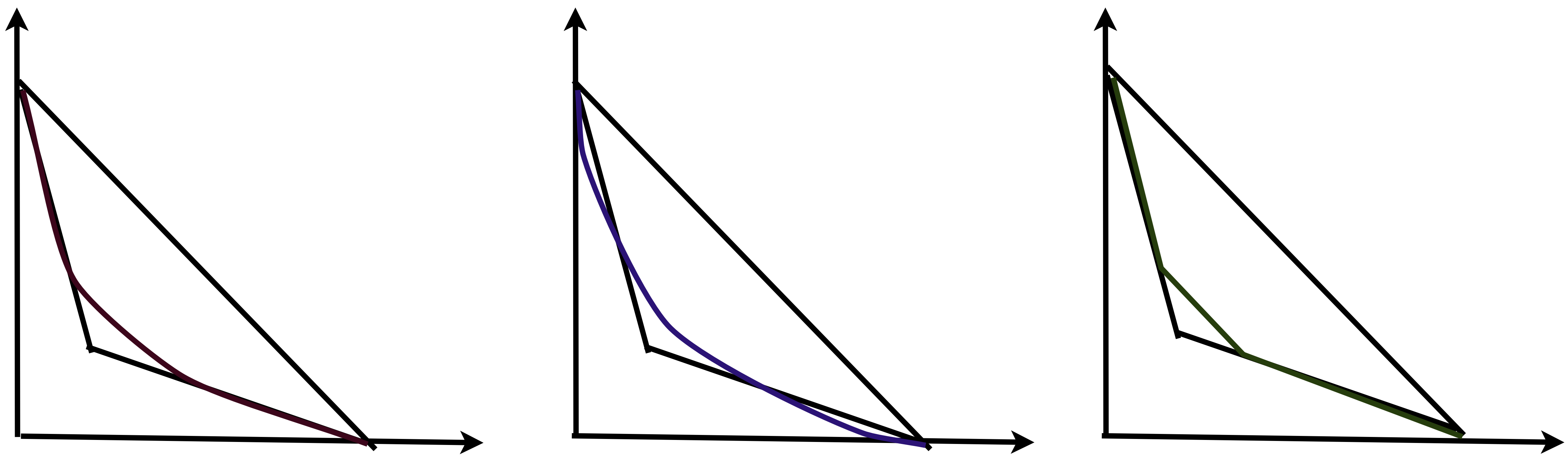
- Log Likelihood ratio $\in \{-\varepsilon, 0, +\varepsilon\}$

Canonical Result

- Laplacian mechanism (and variants) widely used
  - many papers on differential privacy
- Staircase mechanism applies
  - in nearly each case
  - improves performance nearly each time
  - pronounced improvement in moderate/low privacy regimes
- Two limitations
  - intuition missing
  - generalization hard
    - data/query dependent mechanisms

FA-MD Tradeoff Curves

- Operational setting
    - binary hypothesis testing

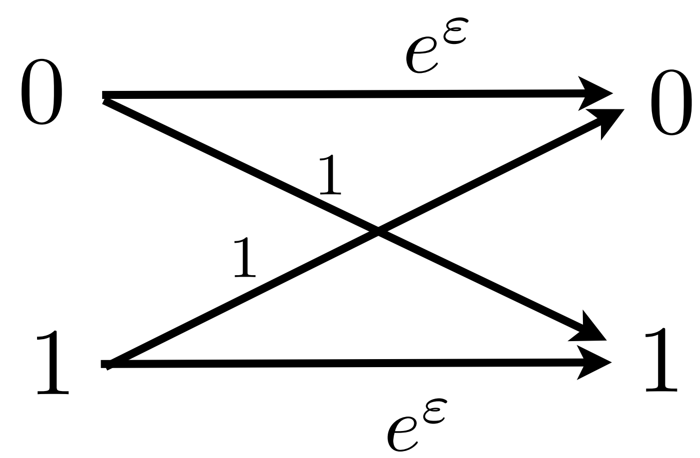


Laplacian Gaussian Staircase

- too complicated
    - multiple query output values

Binary Query

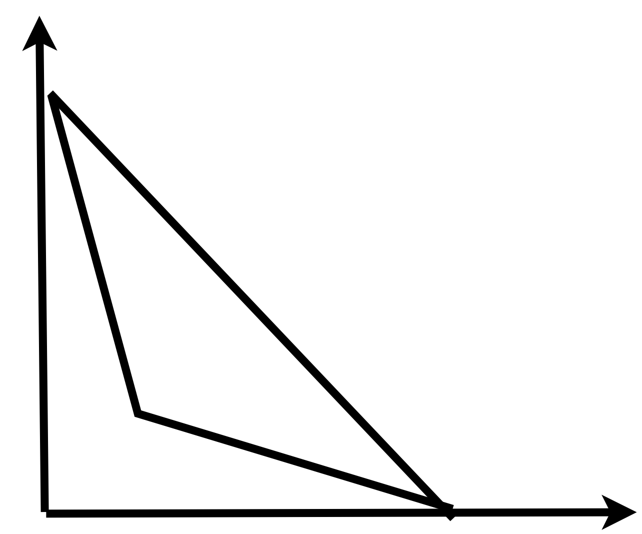
- Binary output
  - Yes or No answer
- Natural mechanism
  - randomized response; W59



- Potentially suboptimal in general
  - more complicated outputs
  - 2-party distributed AND computation GMPS13

Operational Look

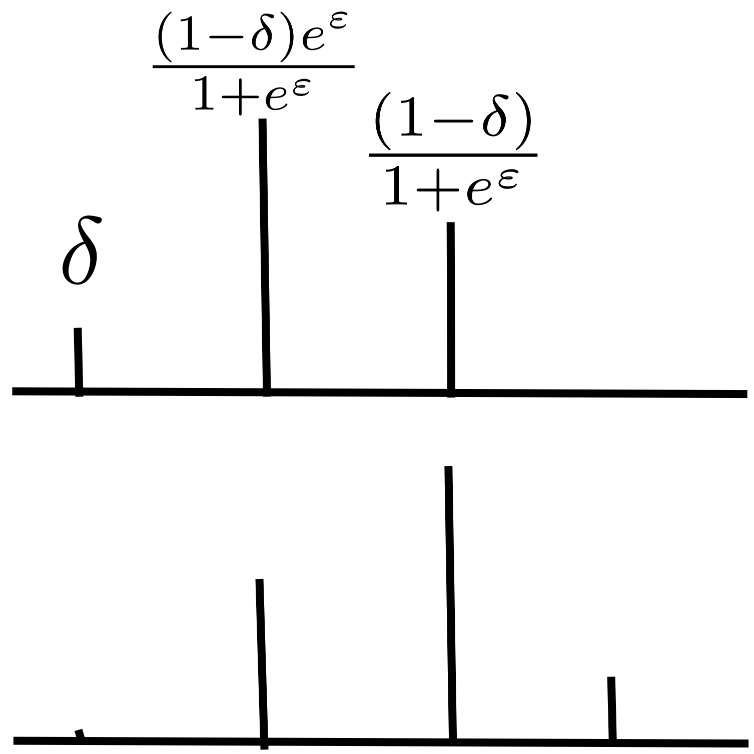
- Binary output
  - randomized response X
  - likelihood ratio $\in \{-\varepsilon, +\varepsilon\}$



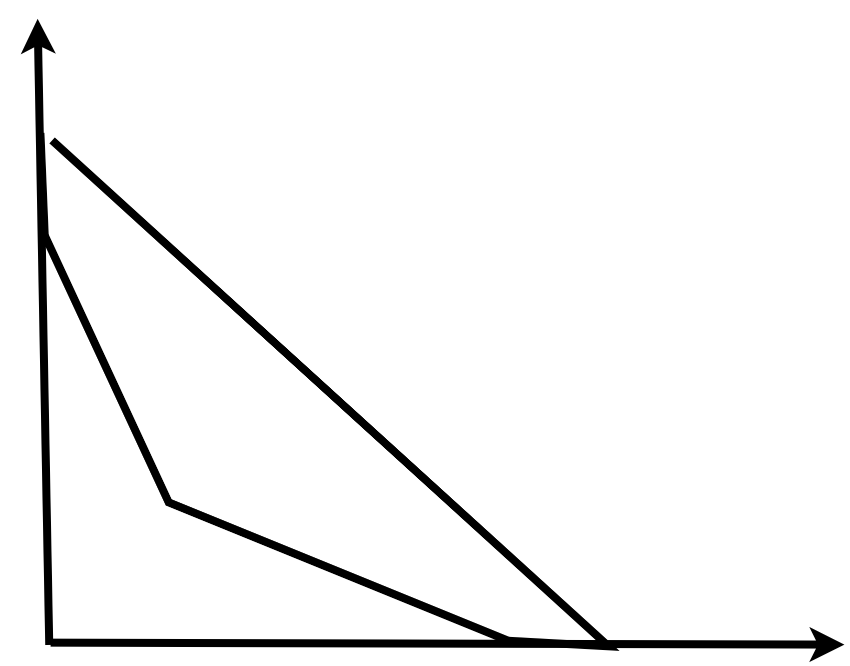
- Exactly meets the privacy region
- Any other mechanism Y
  - only inside the triangular region
- Reverse Data Processing Theorem: B53
  - $D - X - Y$ -- Y can be simulated from X
  - Implications for GMPS13 -- distributed AND computation

Approximate Differential Privacy

- Privatized response has four output letters

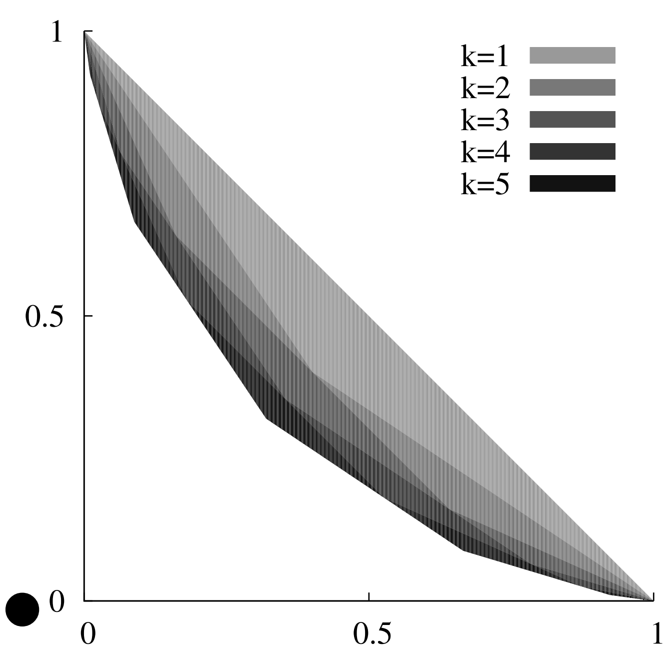


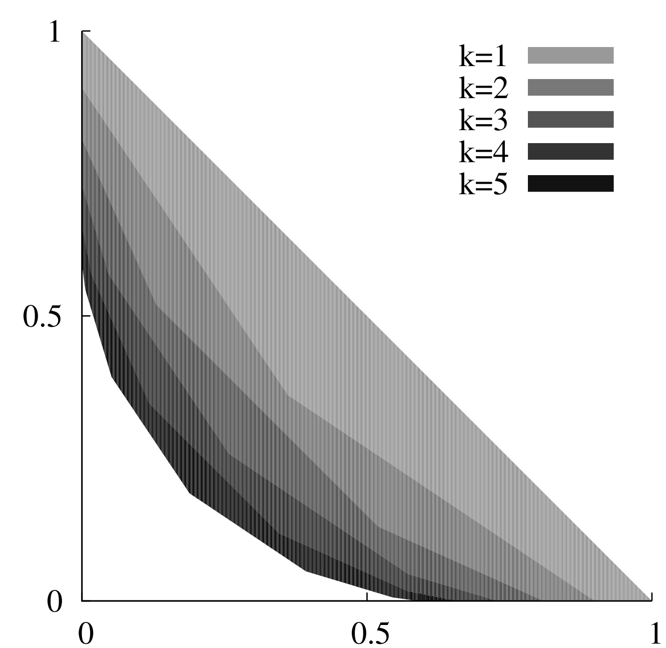
- Exactly meets the privacy region
- Any other mechanism Y
  - only inside the privacy region
  - $D - X - Y$



Composition Theorem

- Privacy region met exactly
  - every other mechanism can be simulated
- Optimal Composition Theorem
  - Composing k queries
  - privacy region is intersection
  - of $((k-2i)\varepsilon, \delta_i)$ privacy regions for i=1..k





Composition Theorem Simplified

- Optimal Composition Theorem
  - conceptually straightforward

- Can be expressed as $(\tilde{\varepsilon}, \delta)$ privacy
  - k-fold composition, each $(\varepsilon, 0)$ private

$$\tilde{\varepsilon} \approx k\varepsilon^2 + \varepsilon\sqrt{2k\log(e + (\sqrt{k\varepsilon^2}/\delta))}$$

  - contrast with state of the art [DRV10]

$$\tilde{\varepsilon} \approx k\varepsilon^2 + \varepsilon\sqrt{2k\log(1/\delta)}$$

  - saving of log factor

Applications of the Composition Theorem

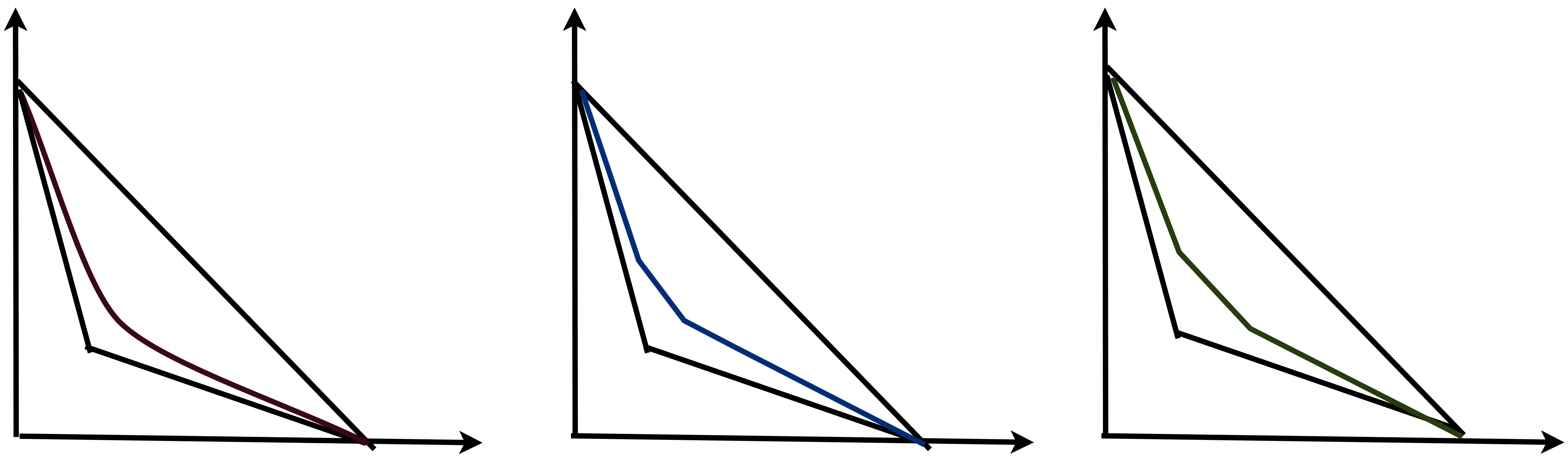
- Order optimality
  - for many mechanisms
  - Laplace
  - Staircase
  - Gaussian

- Direct composition improves performance of Gaussian mechanism
  - sharper concentration analysis
  - chernoff bound
  - direct expression for privacy region

- Immediate applications
  - each intermediate step has less noise

Back to the Staircase Mechanism

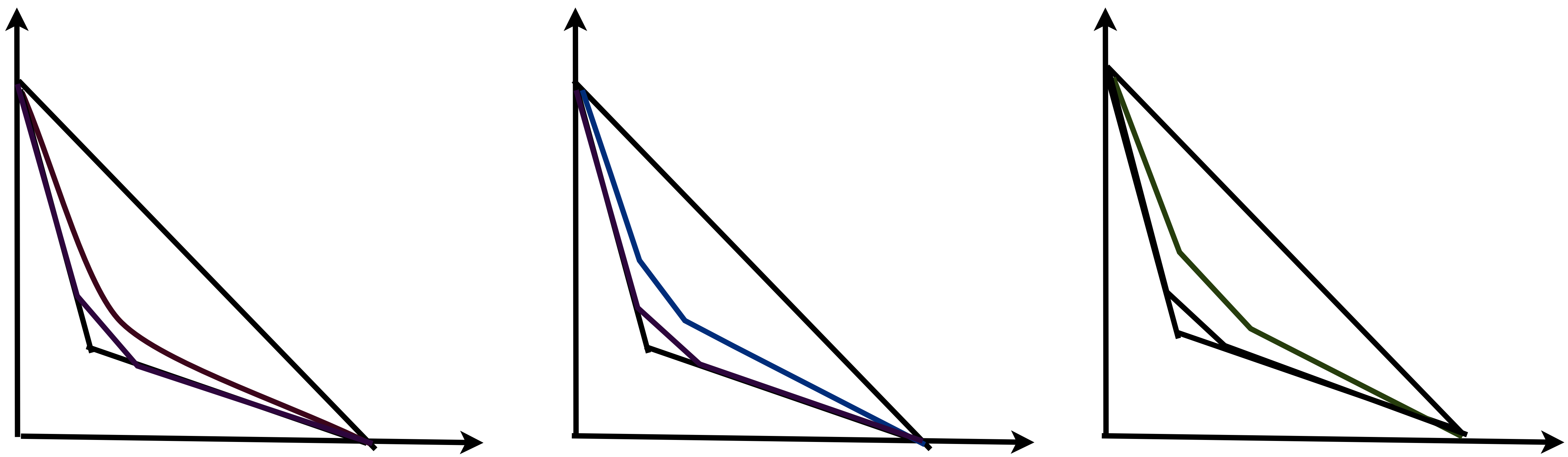
- Ternary query output
  - each pair is neighboring
- View through the operational lens
  - three FA-MD diagrams, one for each pair



- tradeoff among the privacy regions
  - all three regions cannot meet the full triangular region

Back to the Staircase Mechanism

- Ternary query output
  - each pair is neighboring
- Tradeoff among the privacy regions



- Staircase mechanism universally dominates
- Theorem: Every mechanism can be simulated from the staircase mechanism
  - Special reverse data processing inequality

Summary

- Fundamental Mechanisms
  - Staircase mechanism
- Universality
  - cost framework
  - Markov chain framework
- Operational Lens
  - data processing inequalities
- Connections to statistics
  - Blackwell, LeCam
  - converse results to Neyman-Pearson

- Q. Geng and P. Viswanath,
- The Optimal Mechanism for Differential Privacy
- arxiv.org/1212.1186

- Q. Geng and P. Viswanath,
- The Optimal Mechanism for Differential Privacy: Multidimensional Setting
- arxiv.org/1312.0655

- S. Oh and P. Viswanath
- The Composition Theorem for Differential Privacy
- arxiv.org/1311.0776

- Q. Geng and P. Viswanath
- Optimal Mechanisms for Approximate Differential Privacy
- arxiv.org/1305.1330

- Acknowledgement: K Chaudhury, M Hardt and A Smith