New Directions in Secure Multi-Party Computation: Techniques and Information Disclosure Analysis

Dissertation Defense

Alessandro N. Baccarini

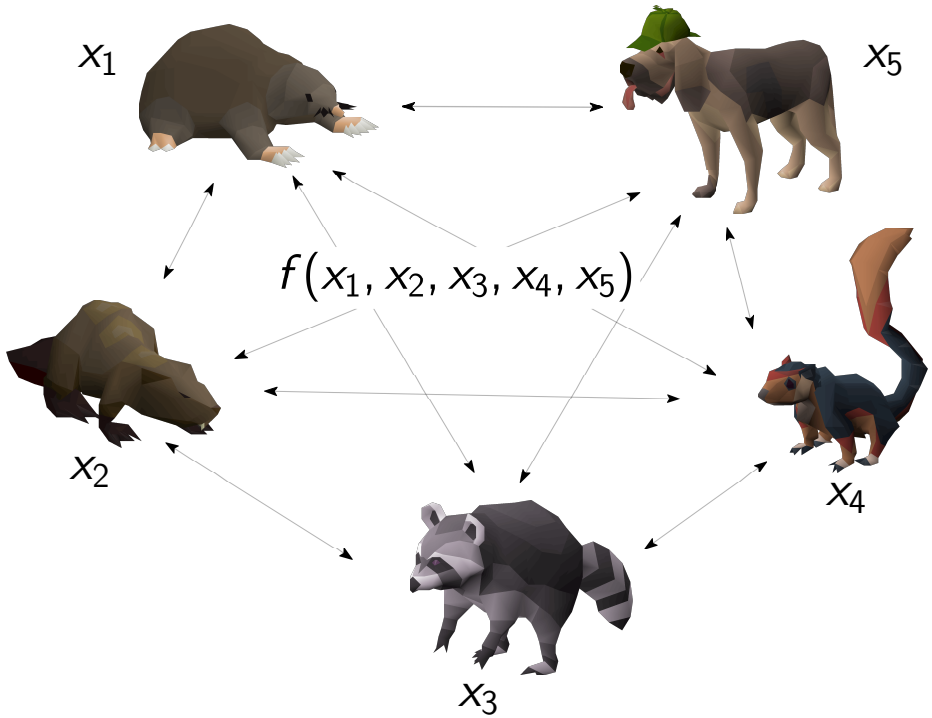
Department of Computer Science and Engineering
University at Buffalo

July 31, 2024

Table of Contents

Background and Motivation

What *is* secure multi-party computation?

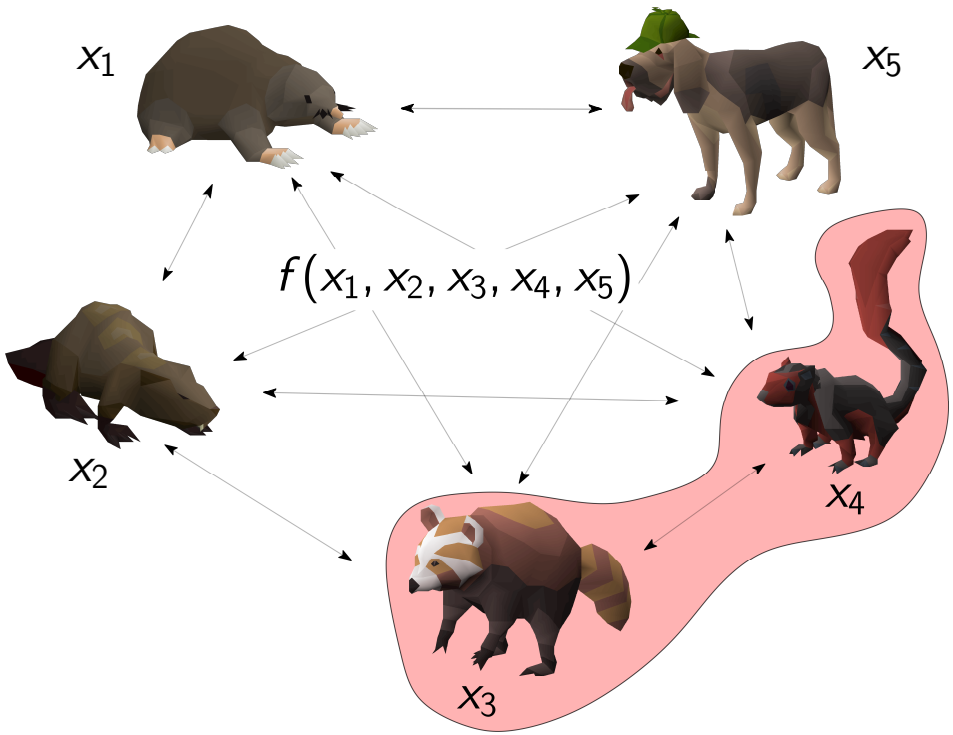


Secure multi-party computation (SMC)

Multiple participants jointly evaluating a function on secret inputs

- **No information is disclosed other than the output**
- Applications in healthcare, ML, **data analytics**
- Homomorphic encryption, garbled circuits, **secret sharing**

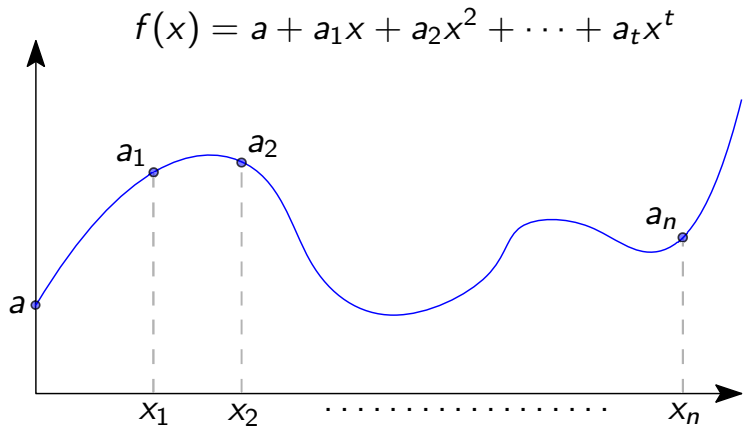
Secret sharing



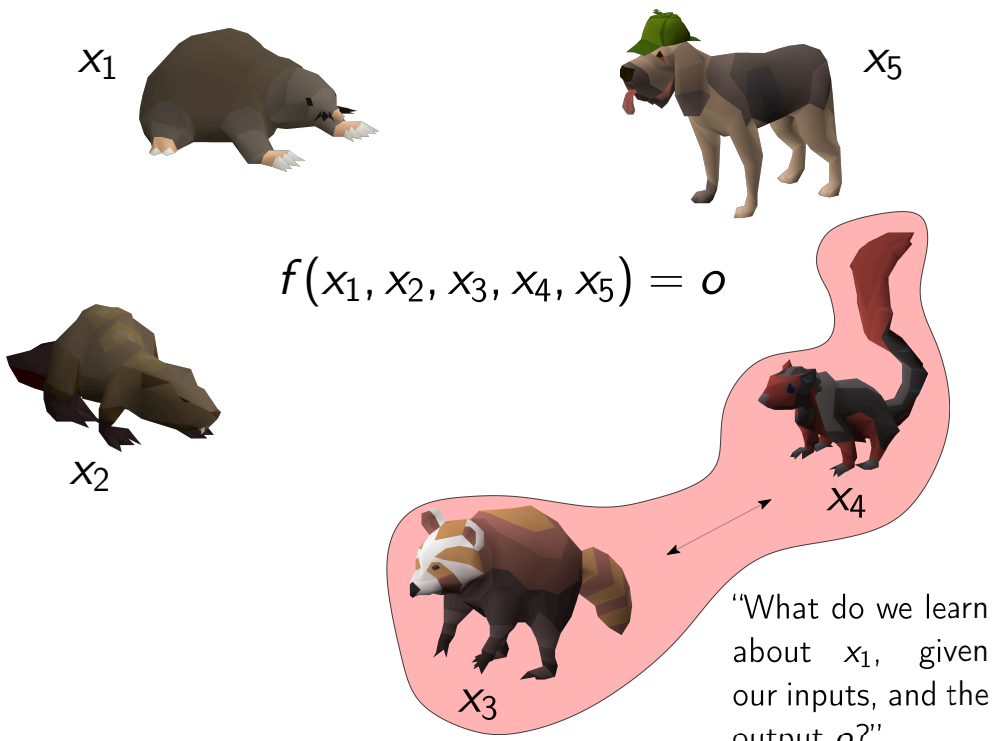
- (n, t)-threshold scheme, $n \geq 2$ and $t < n$
  - $\leq t$ **cannot** recover a secret
  - $> t$ **can** reconstruct a secret
- Setting: **semi-honest, honest majority** $(n = 2t + 1)$

Secret sharing (SS) techniques

- Fields $\mathbb{F}_p$
  - Shamir SS [Sha79]
  - **Modular reduction**
  - **Multiplicative inverses** needed for interpolation
  - **Large-number libraries** (e.g., GMP)
- Rings $\mathcal{R}$
  - **Replicated** SS (RSS) [ISN87]
  - Compatible with **native CPU instructions**
  - Existing frameworks limited to **three or four parties**
  - No comprehensive ring-based framework for **integer** and **floating-point** computation



Broader (unanswered) questions?



- Guaranteed to not leak any private information throughout the computation
- What about **after** the function is evaluated?
- Does the **output itself** leak any sensitive information?
- Can we **measure** this leakage in a meaningful way?

Dissertation overview

Part I. Complete RSS framework over a ring

- Comprehensive RSS framework for n parties over a ring
- Integer and floating-point compatibility
- Extensive benchmarks against field equivalent and state-of-the-art

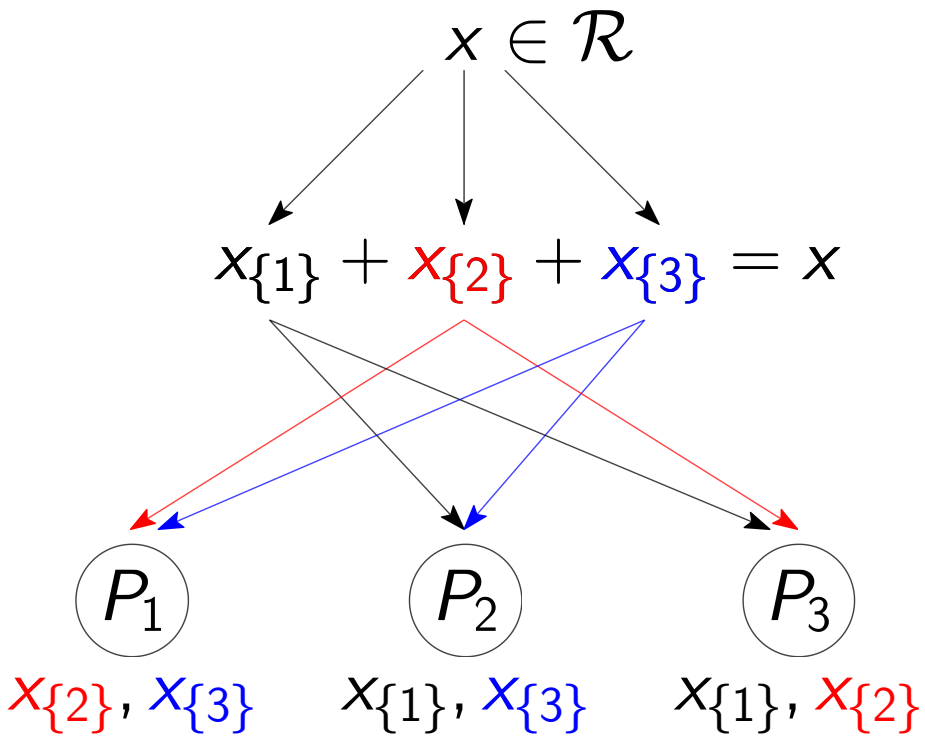
Part II. Information disclosure analysis of statistical functions

- Develop an information-theoretic approach to measure disclosure
- Study a practically significant function (average salary)
- Extend our analysis to complex statistical functions

An n-party RSS Framework

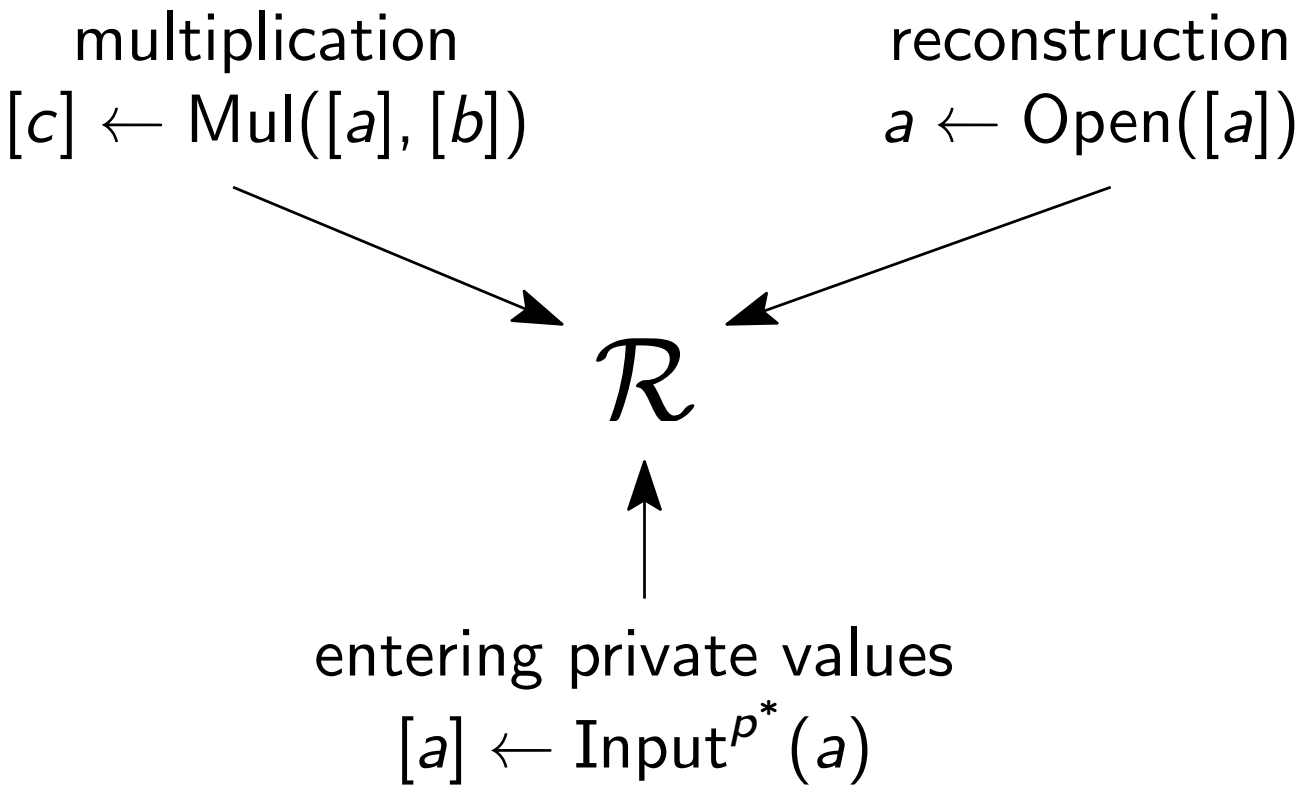
Replicated secret sharing (RSS)

- Shamir $\implies$ one share per party
- RSS:
  - Split x into $\binom{n}{t}$ shares
  - Distribute $\binom{n-1}{t}$ shares per party
- $t+1$ parties have access to **all shares**



Building blocks

– Addition, multiplication by a **public** value are non-interactive (local)



– All single-round operations

Binary-to-arithmetic conversion (B2A)

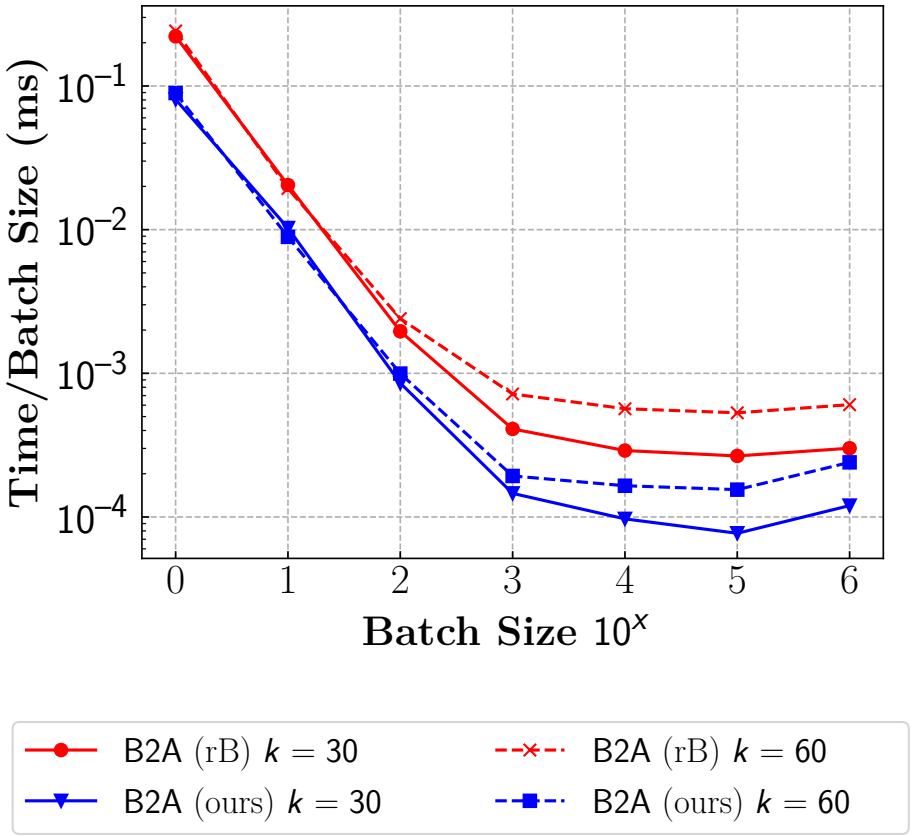
- Convert shares from $\mathbb{Z}_2 \to \mathbb{Z}_{2^k}$
- Prior works use **randBit** [Dam+19] $\implies$ temporarily in $\mathbb{Z}_{2^{k+2}}$ 👎
- Blanton et al. [BGY23] **eliminated this requirement** for 3PC RSS

Generalization to *any n*

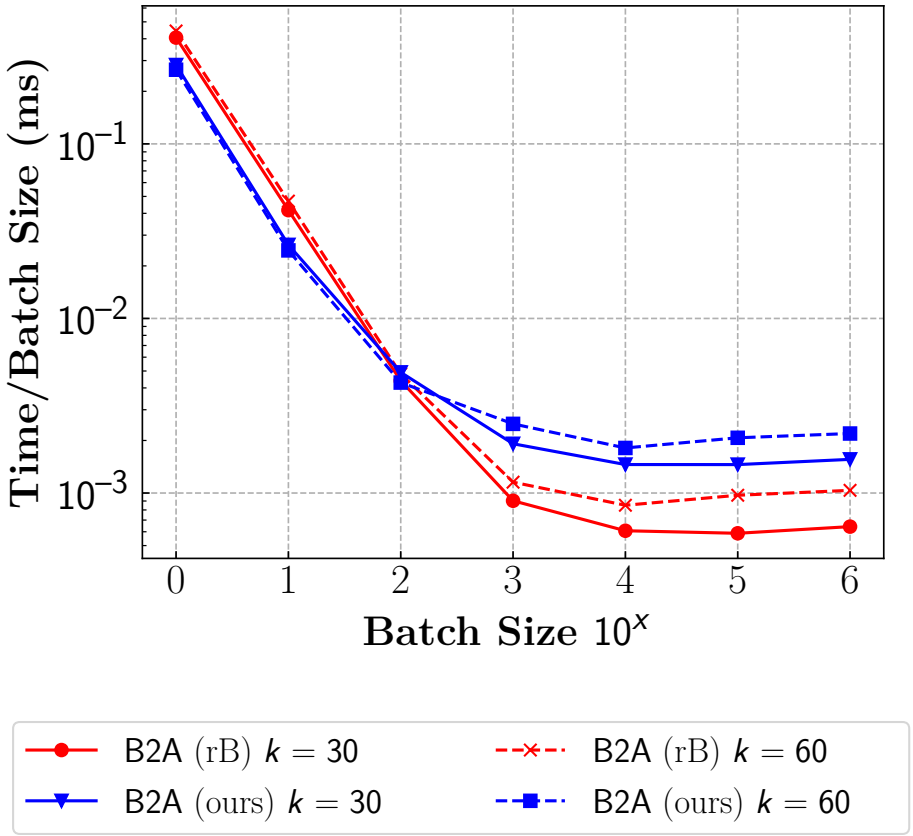
1. t parties locally XOR some (or all) of their shares, Input result into computation
2. Remaining $t+1$ parties locally reshare the "last" share (all but one share is nonzero)
3. Compute XOR (in $\mathbb{Z}_{2^k}$) of inputted secrets(s) and the last share as a tree

- **Optimization**: one XOR is with a secret with one **nonzero share**
  - Can use a cheaper multiplication!

B2A Performance



3-party B2A



5-party B2A

- Inferior performance for 7 parties
- ... but don't have to switch rings 🤷

Experimental evaluation

- Many more protocols in the dissertation (comparisons, truncation, division, ...)
- RSS superior to field-based equivalents for 3, 5, 7 parties (10-33× improvement)

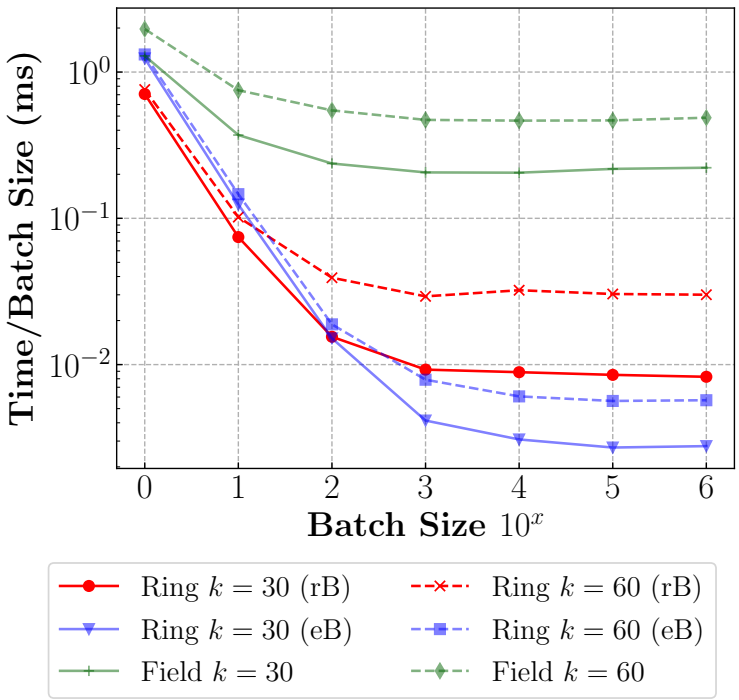
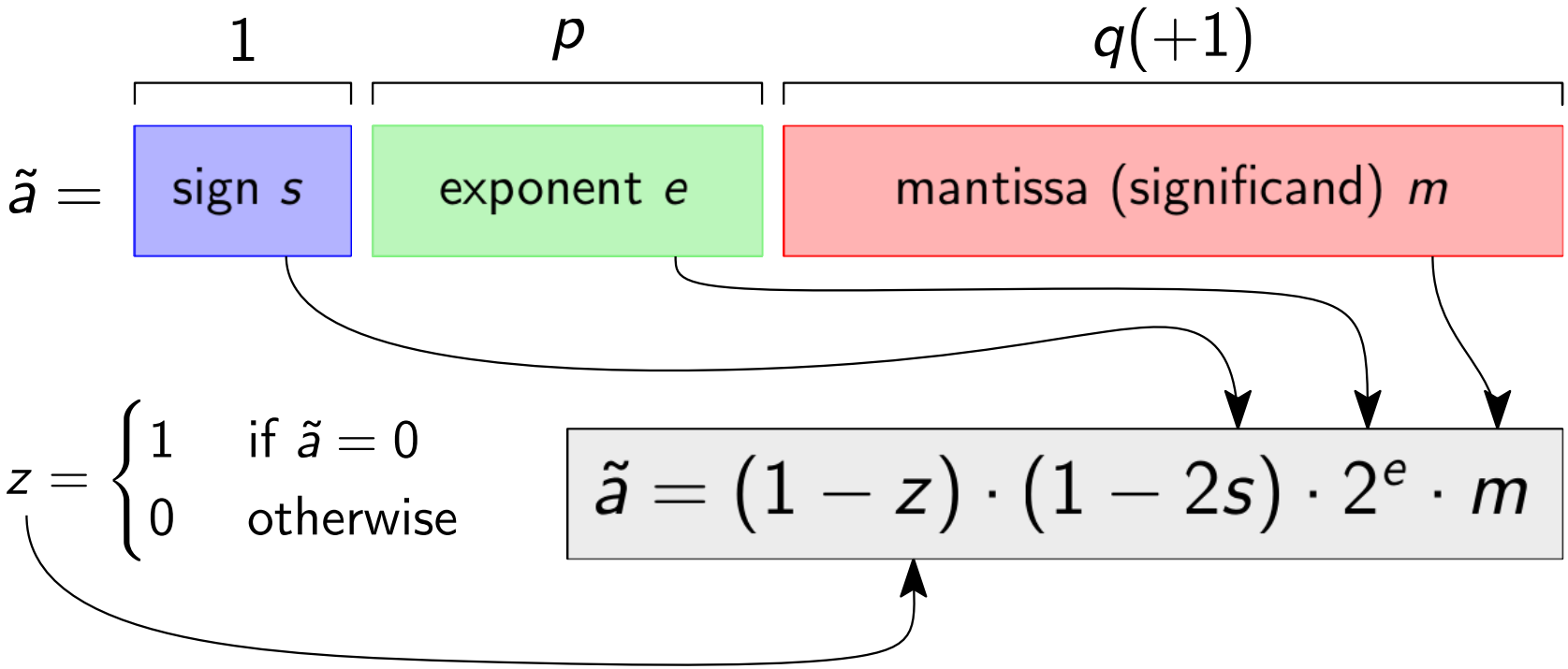


Figure: 3PC comparisons ($a \overset{?}{<} b$)

Floating-point protocols

- Prior protocols designed for **integers**
- But what about **floating-point**?



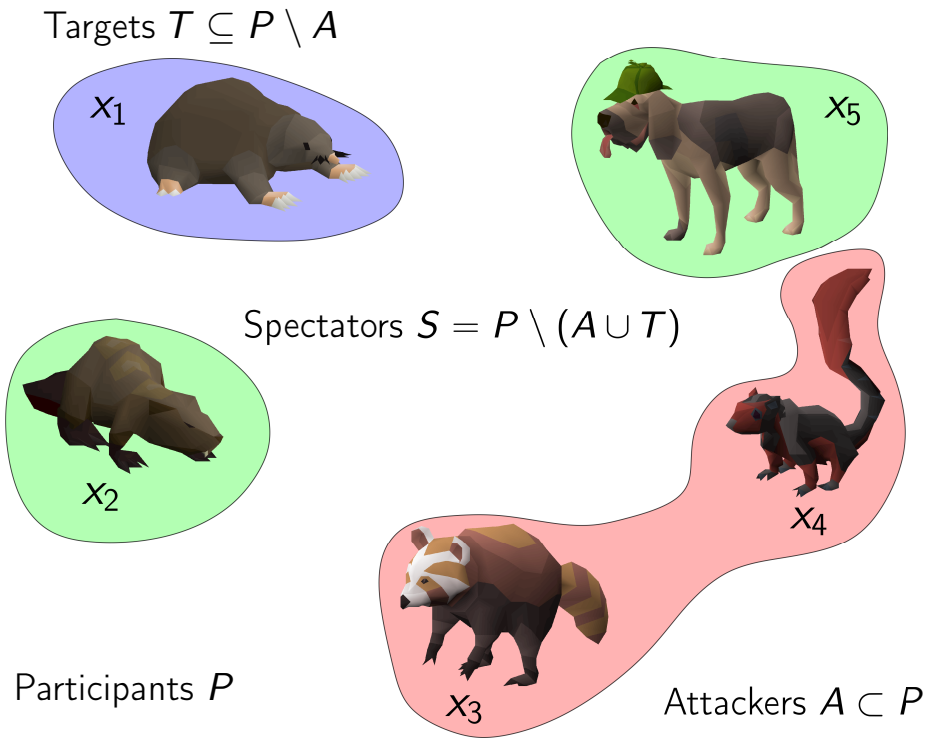
Operations

- Arithmetic protocols from [Ali+13; Rat+22; Cat20]
- Most operations conceptually "similar" to integer equivalents...
  - Comparisons $\tilde{a} \overset{?}{<} \tilde{b}$ (LT, EQ)
  - Multiplication $\tilde{a} \cdot \tilde{b}$ (LT, Trunc)
  - Division $\tilde{a}/\tilde{b}$ (LT, Trunc)
- ... except for Addition $\tilde{a} + \tilde{b}$ (LT, EQ, Trunc, Pow2, PrefixOR, ...)
  - Exponents must be **aligned**
  - **Left shift** larger input's mantissa
  - Truncate and round the result
- **Rounding** rules
  - **Directed** rounding to $-\infty$ $\mathrm{Trunc}([a], m)$
  - Rounding to **nearest**, ties to **even** $\mathrm{RNTE}([a], m)$
- Implementation: WIP 👷

Information disclosure analysis

Formal setting

– How do we differentiate participants from one another?



Metric?

- Model participant i's input by a random variable X_{P_i}
- How should we **measure** the **information** disclosed from the output?

Entropy!

Shannon (discrete)	Differential (continuous)
$H(X)$	$h(X)$

Putting it together

- Attackers $\vec{X}_A$, targets $\vec{X}_T$, and spectators $\vec{X}_S$
- Treat the **output** as a random variable: $f(\vec{X}_A, \vec{X}_T, \vec{X}_S) = O$

Attackers' weighted average entropy [AH17]

$$H(\vec{X}_T \mid \vec{X}_A = \vec{x}_A, O) \implies$$ how much information is learned about the target, given $\vec{x}_A$ and O

- Equivalent expression for differential entropy

Where to begin?

- 2016 Boston gender pay gap survey [Cou17]
- Analyzed the **private** wages based on gender and race **using SMC**
- **Average salary computation**

Mayor Walsh & Boston Women's Workforce Council Release 2016 Gender Wage Gap Report; New Partnership with BU

Thursday, January 5, 2017 – Mayor Martin J. Walsh and the Boston Women's Workforce Council (BWWC) released the 2016 gender wage report and announced a new academic partnership with Boston University, where the BWWC will now be hosted within the BU Hariri Institute for Computing.

Source: Boston University, 2017

- Average $\implies$ reduces to a **sum**:

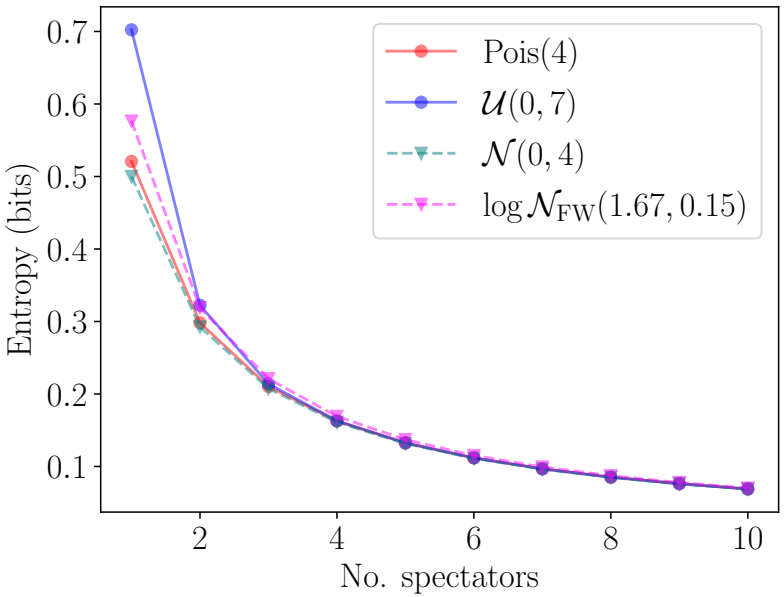
$$f_\mu(\vec{x}) = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \implies \boxed{x_1 + x_2 + \cdots + x_n}$$

Single evaluation

- Information disclosure is **independent** of:
  - the attackers' input(s) $\implies H(\vec{X}_T \mid \vec{X}_A = \vec{x}_A, O) = H(\vec{X}_T \mid O)$

$$O = X_T + X_S + X_A \text{ “ = ” } X_T + X_S$$

  - the input distribution (Poisson, uniform, **Gaussian**, log-normal [CG05]) and its parameters



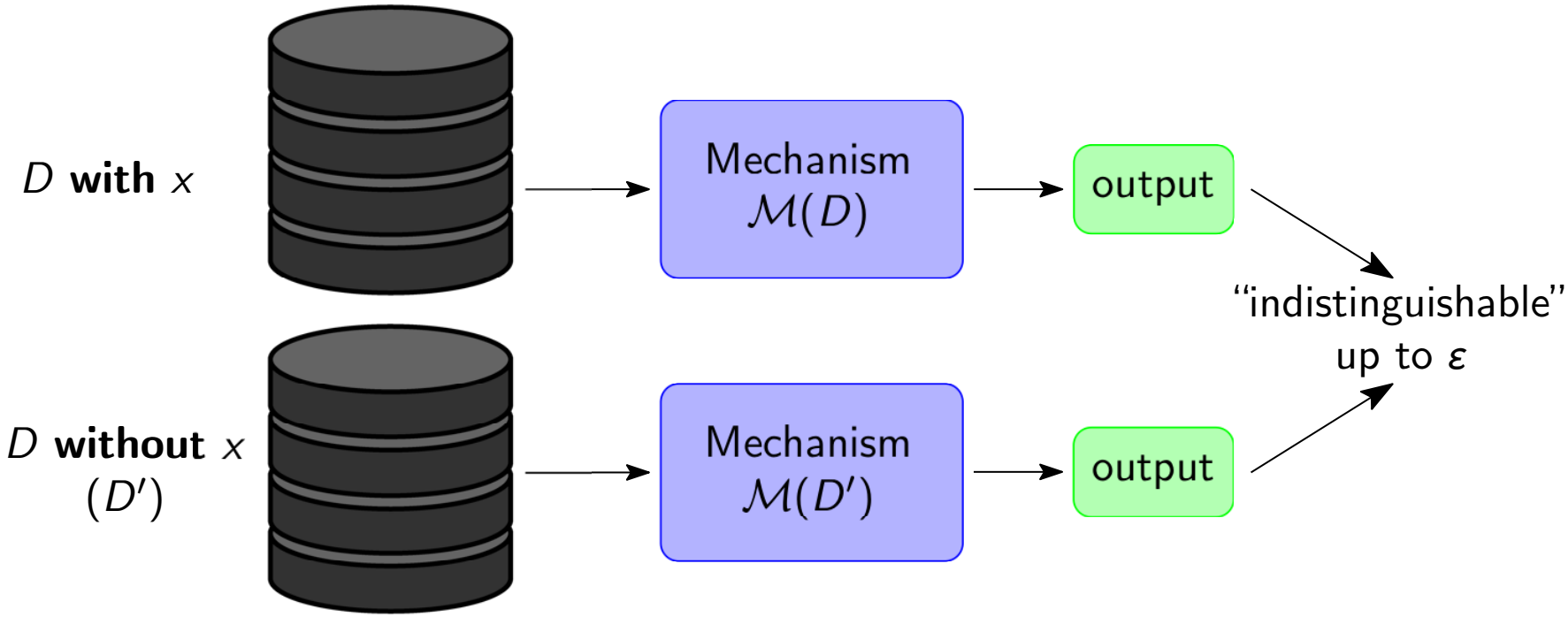
Absolute loss
= Target's initial entropy
– remaining entropy

- Disclosure is **proportional** to the number of spectators

Elephant in the room...

"Have you considered using **DIFFERENTIAL PRIVACY**???"

Differential privacy 🙄



- Useful for large databases (think $n \geq 10,000$)...
- ... but **absolutely destroys** the utility of the result (up to 100% error!)
- Our goal: **determine** if a function discloses too much information

Two evaluations

- Does the wage gap change over time?
- That's what the BWWC wanted to find out!

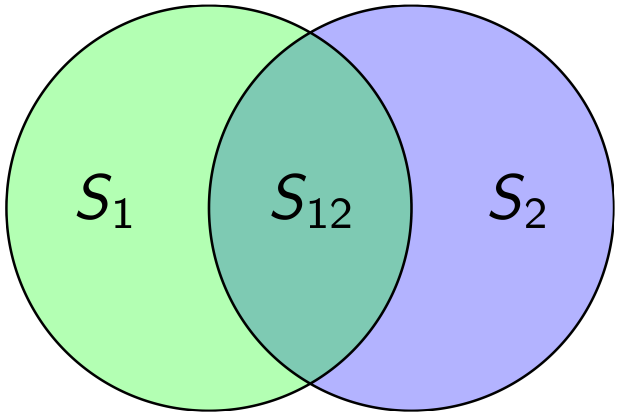
Mayor Walsh & BWWC Release 2017 Wage Gap Report

The Boston Women's Workforce Council released its 2017 report this morning, which uses real employer wage information to assess the pay gap in Boston. The 2017 report

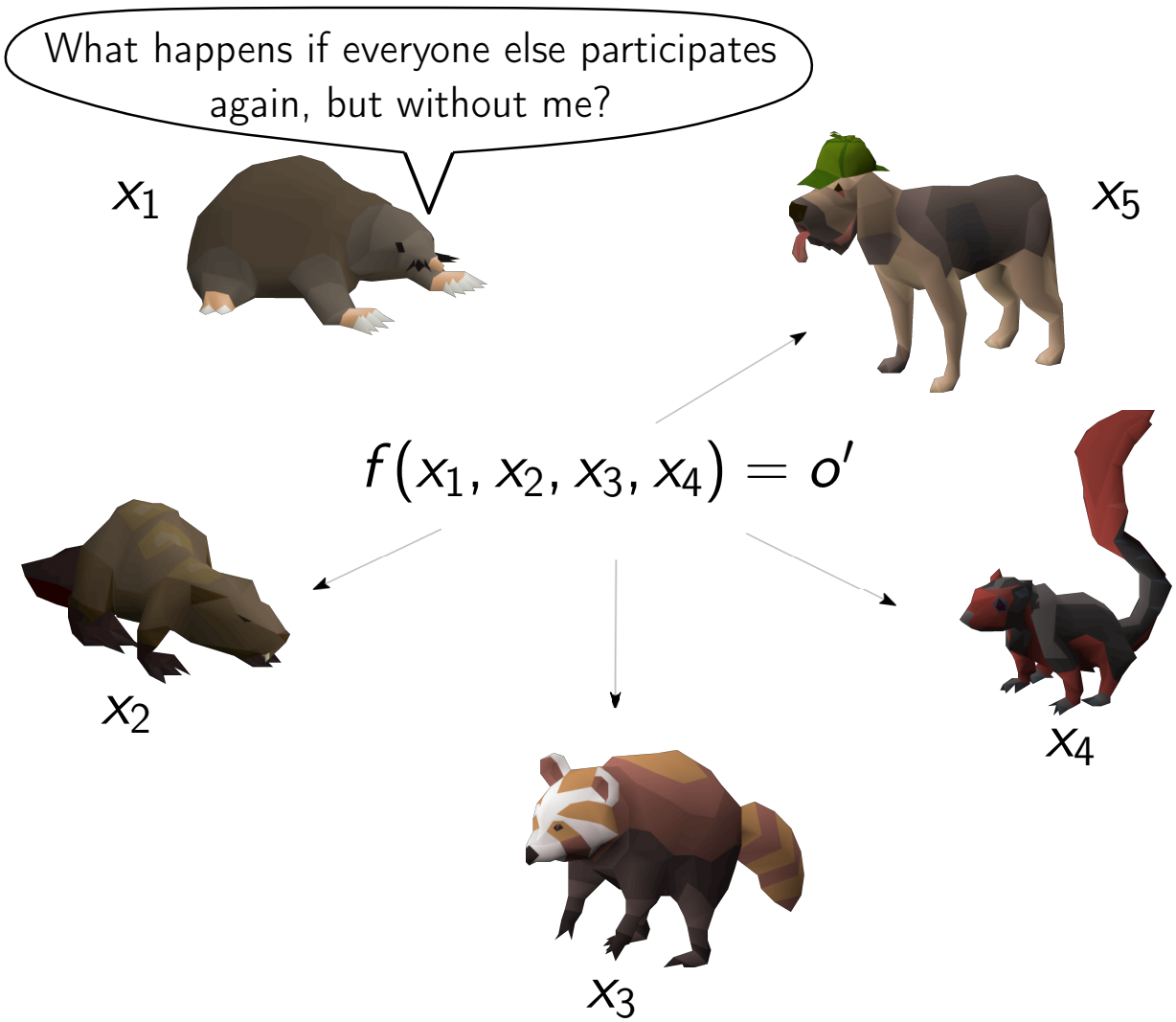
BOSTON WOMEN'S WORKFORCE COUNCIL REPORT 2017

Source: Boston University, 2018

- Spectators present in the first, second, and both evaluation(s)

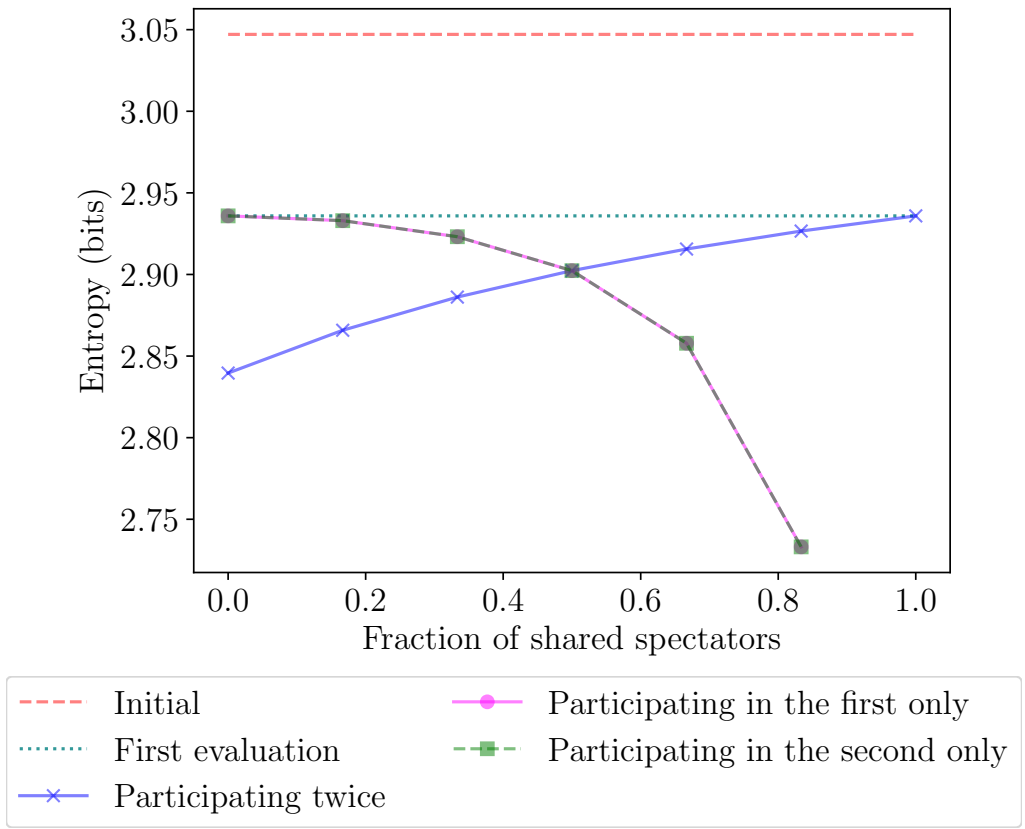


Two evaluations



Two evaluations

- Participating **once** vs. **twice**
- Largest protection at **50% overlap**
- Undesirable disclosure at extrema



Next steps

- Advanced statistical functions:

maximum $\quad f_{\max}(\vec{x}) = \max_{i \in [1,n]} x_i \quad$ (equiv. for min)

median* $\quad f_{\text{med}}(\vec{x}) = \begin{cases} x_{\frac{n+1}{2}} & \text{odd } n \\ \min\left(x_{n/2}, x_{n/2+1}\right) & \text{even } n \end{cases}$

variance $\quad f_{\sigma^2}(\vec{x}) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - f_{\mu}(\vec{x})\right)^2 \implies$ mean **and** var $f_{(\mu,\sigma^2)}(\vec{x})$

New functions $\Longrightarrow$ new challenges

- **No closed-form expressions** of the entropy anymore 😢
- Output RV could be **discrete**, while the input RVs are **continuous**
- Look to **data-driven techniques** [Gao+17]

Estimating Mutual Information for Discrete-Continuous Mixtures

---

Weihao Gao
Department of ECE
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
wgao9@illinois.edu

Sreeram Kannan
Department of Electrical Engineering
University of Washington
ksreeram@uw.edu

Sewoong Oh
Department of IESE
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
swoh@illinois.edu

Pramod Viswanath
Department of ECE
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
pramodv@illinois.edu

mutual information

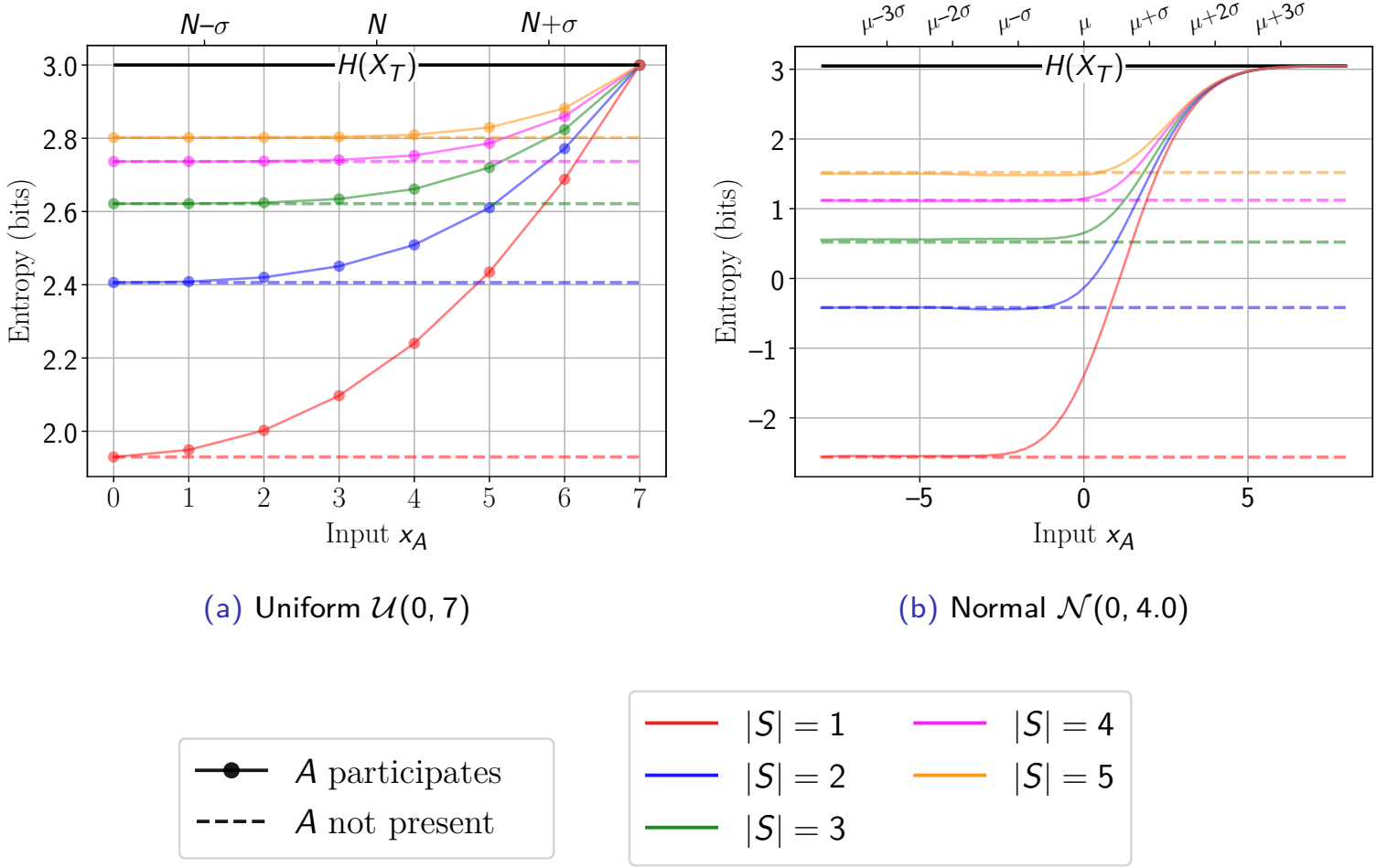
$\Leftrightarrow$

absolute loss

Some intuitive observations...

Maximum

Adversary **maximizes** information learned by **minimizing** their influence.



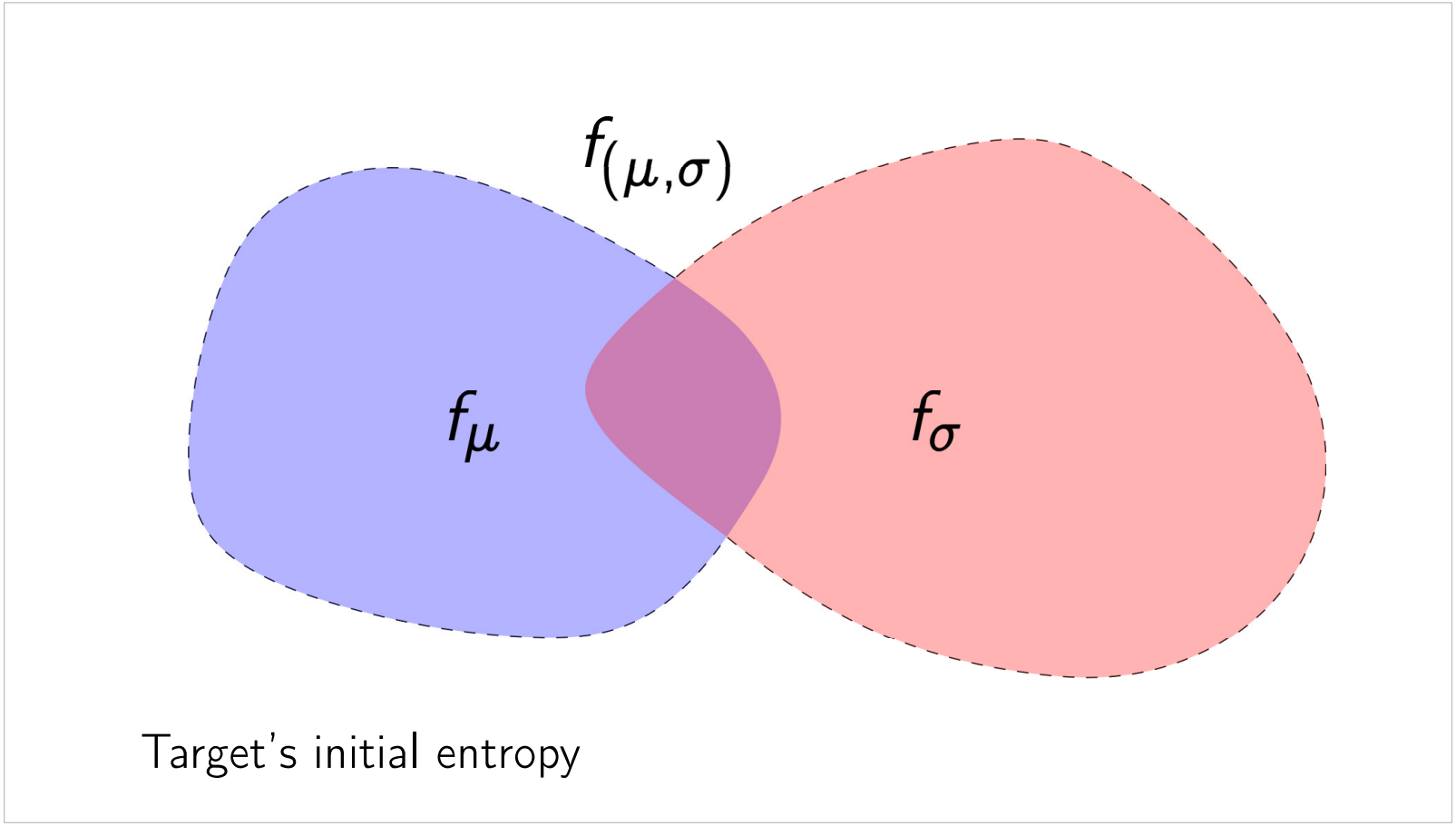
(a) Uniform $\mathcal{U}(0, 7)$

(b) Normal $\mathcal{N}(0, 4.0)$

... and not so intuitive ones

Variance **and** mean release

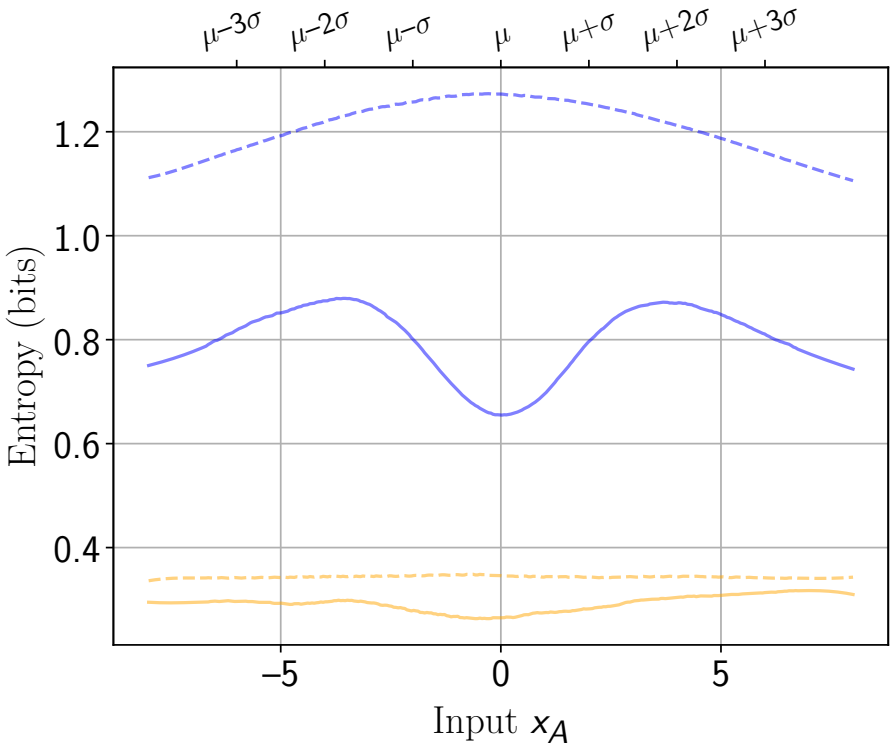
The total disclosure from **individual** function outputs f_μ and f_σ is **at least** the amount of information disclosed from a **joint release** $f_{(\mu,\sigma)}$?



... and not so intuitive ones

Variance **and** mean release

The total disclosure from **individual** function outputs f_μ and f_σ is **at least** the amount of information disclosed from a **joint release** $f_{(\mu,\sigma)}$?



(a) Abs. entropy loss, Normal $\mathcal{N}(0.0, 2.0)$ (lower is better)

- **Gap** between the curves
- Possible to learn **more** information about the target

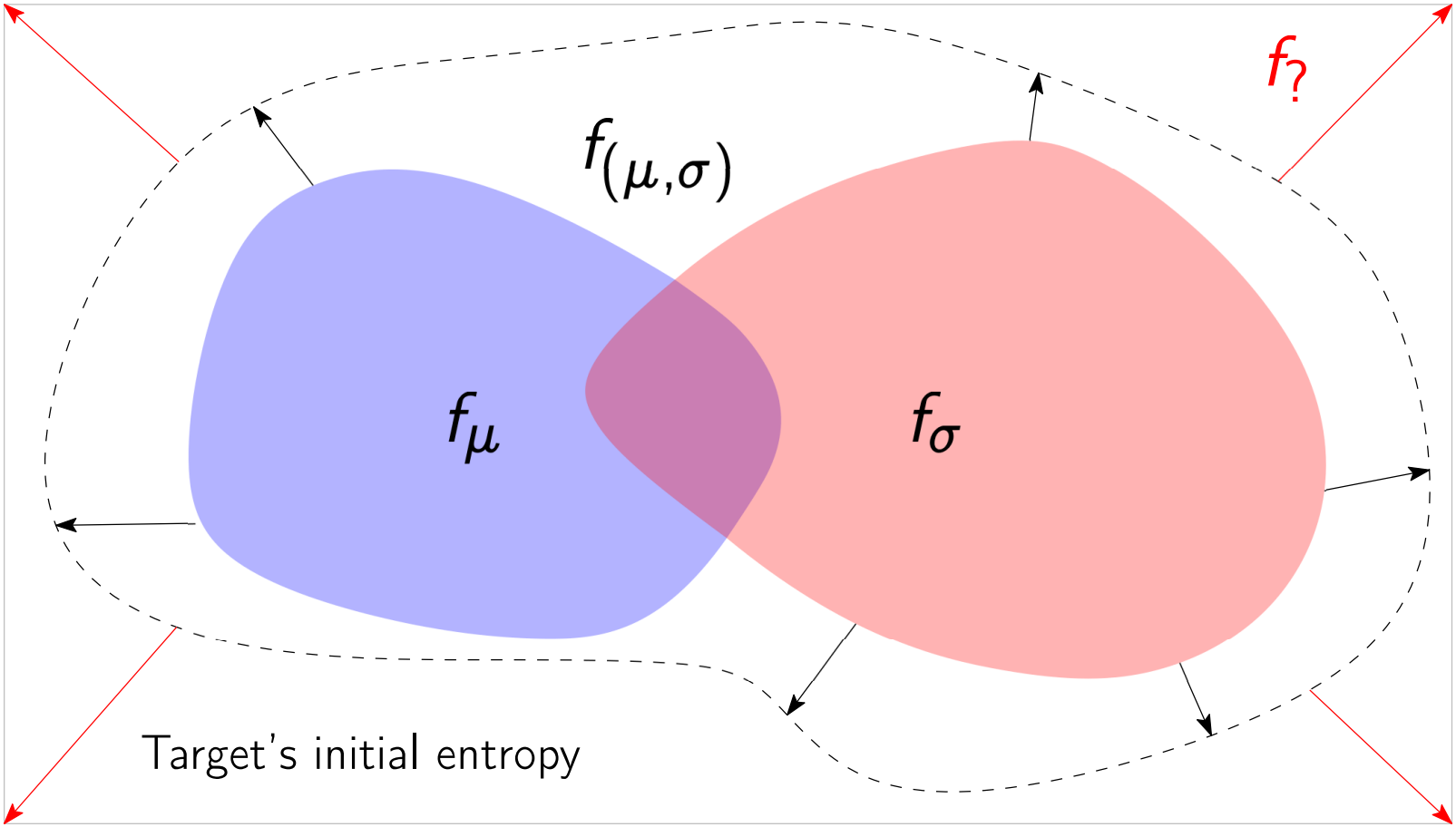
$H_{f_\mu} + H_{f_{\sigma^2}}$ $H_{f_{(\mu,\sigma^2)}}$

$|S| = 2$
$|S| = 5$

... and not so intuitive ones

Variance **and** mean release

More information is revealed from the **joint release** $f_{(\mu,\sigma)}$ than from **individual** function outputs f_μ and f_σ (under summation).



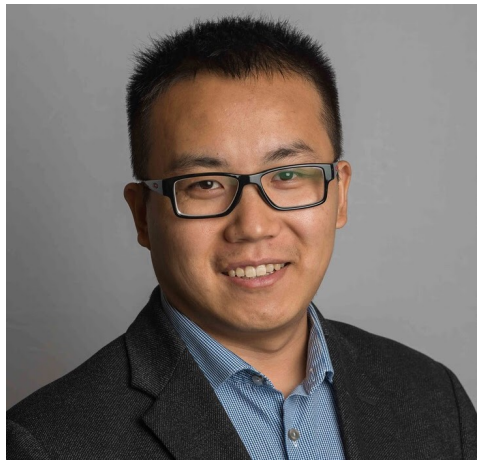
Conclusions

Conclusions and future directions

- RSS framework
  - Constructed a comprehensive n-party framework
  - Integer and floating-point operations
  - Additional operations (e.g., square-root, logarithm)
  - Weaker security assumption
- Information Disclosure analysis
  - Comprehensively analyzed disclosure from the output of the average salary
  - Recommendations for computation designers
  - Studied advanced descriptive statistics
  - Derive closed-form expressions for complex functions

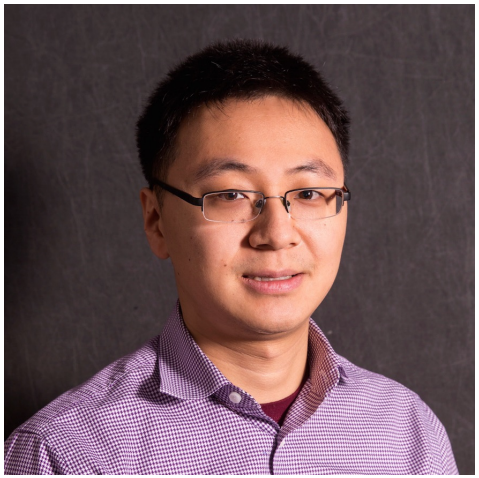
A few acknowledgments...

Marina Blanton



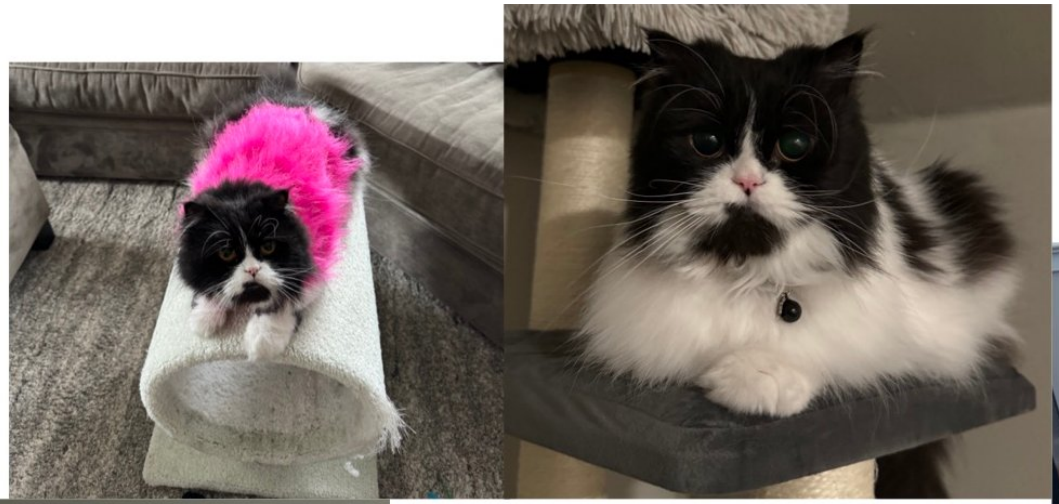
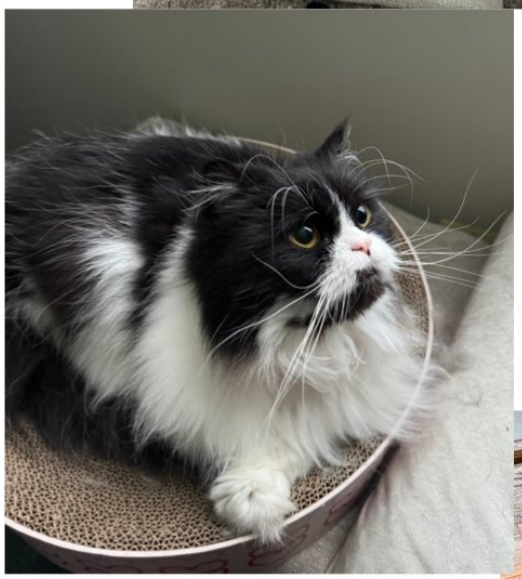
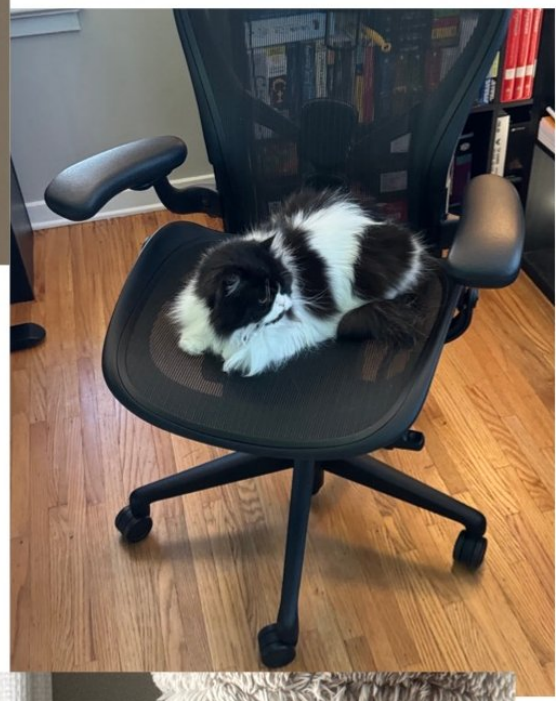
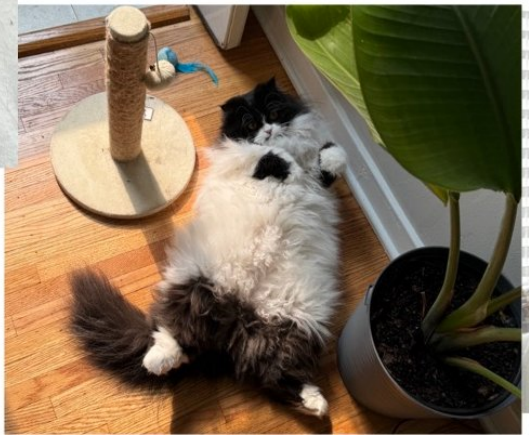
Shaofeng Zou

Good luck at ASU!



Ziming Zhao

Good luck at Northeastern!


Thank you!

Questions?

References

[AH17] P. Ah-Fat and M. Huth. "Secure Multi-party Computation: Information Flow of Outputs and Game Theory". In: *International Conference on Principles of Security and Trust*. 2017, pp. 71–92.

[Ali+13] M. Aliasgari et al. "Secure computation on floating point numbers". In: *Network and Distributed System Security Symposium (NDSS)*. 2013.

[BGY23] M. Blanton et al. "Secure and Accurate Summation of Many Floating-Point Numbers". In: *Proceedings on Privacy Enhancing Technologies (PoPETs)* 2023.3 (2023), pp. 432–445.

[Cat20] O. Catrina. "Evaluation of floating-point arithmetic protocols based on Shamir secret sharing". In: *International Joint Conference on e-Business and Telecommunications (ICETE)*. 2020, pp. 108–131.

[CG05] F. Clementi and M. Gallegati. "Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States". In: *Econophysics of Wealth Distributions*. 2005, pp. 3–14. DOI: 10.1007/88-470-0389-X_1.

[Cou17] Boston Women's Workforce Council. *Boston Women's Workforce Council Report 2016*. https://htv-prod-media.s3.amazonaws.com/files/bwwc-report-final-january-4-2017-1483635889.pdf. 2017.

[Dam+19] I. Damgård et al. "New Primitives for Actively-Secure MPC over Rings with Applications to Private Machine Learning". In: *IEEE Symposium on Security and Privacy (S&P)*. 2019, pp. 1102–1120.

[Gao+17] W. Gao et al. "Estimating mutual information for discrete-continuous mixtures". In: *Proceedings on Advances in Neural Information Processing Systems (NeurIPS)* 30 (2017), pp. 5988–5999.

[ISN87] M. Ito et al. "Secret sharing schemes realizing general access structures". In: *IEEE Global Telecommunication Conference (GLOBECOM)*. 1987, pp. 99–102.

[Rat+22] D. Rathee et al. "SecFloat: Accurate Floating-Point meets Secure 2-Party Computation". In: *IEEE Symposium on Security and Privacy (S&P)*. 2022, pp. 1553–1553.

[Sha79] A. Shamir. "How to share a secret". In: *Communications of the ACM* 22.11 (1979), pp. 612–613.