Minimax Optimal Two-Sample Testing under Local Differential Privacy

(Journal of Machine Learning Research)

Jongmin Mun

University of Southern California

December 16, 2025

Outline

Motivation: Federated Analytics

Federated Learning (FL)

- **Privacy concern:** Neural nets should be trained on sensitive user data (e.g., messages, photos) saved in local devices
- **Solution:** Data remains local; **only weight updates** are transmitted

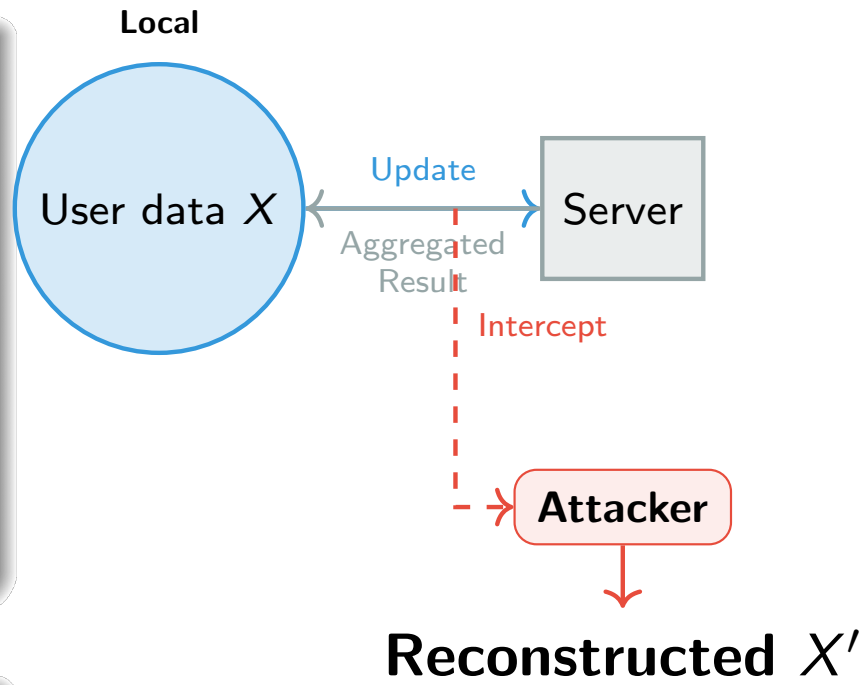
Federated Analytics (FA)

- **Goal:** Leverage FL infrastructure for **data science** (counts, rates, distributions, testing) rather than neural net training.
- **Mechanism:** Instead of gradient, devices perform local computation of data summary; server receives them and *aggregates*
- **Benefit:** (1) supports FL (aggregate model evaluation metrics, check local dataset size), (2) Generates product insights under privacy

Privacy in Federated Systems

Compute Locally, Aggregate Centrally

- **Federated Learning (FL):** Transmits **gradients**
- **Federated Analytics (FA):** Transmits **data summary**
- **Key:** Raw data never leaves the device; Only summaries transmitted



Aggregation $\neq$ Privacy

Attackers can intercept the aggregation and summary and reconstructs user data.

Noises Give More privacy in FL and FA

The Strategy: Differential Privacy (DP)

- **Mechanism:** Devices transmit *noisy* versions of data.
- **Intuition:** Local noise cancels out during aggregation, preserving population trends while masking individuals during transmission
- **Formalism:** DP mathematically determines the precise noise magnitude required for a given privacy budget ϵ.

Industry Implementation: RAPPOR (Google)

- Federated Analytics tool for aggregating noisy multicategory counts
- **Use Case:** Analyzing Chrome homepage settings (e.g., malware detection) while preserving privacy (local differential privacy)
- **Local DP (LDP)**: given privacy budget ε, the noisy output satisfies:

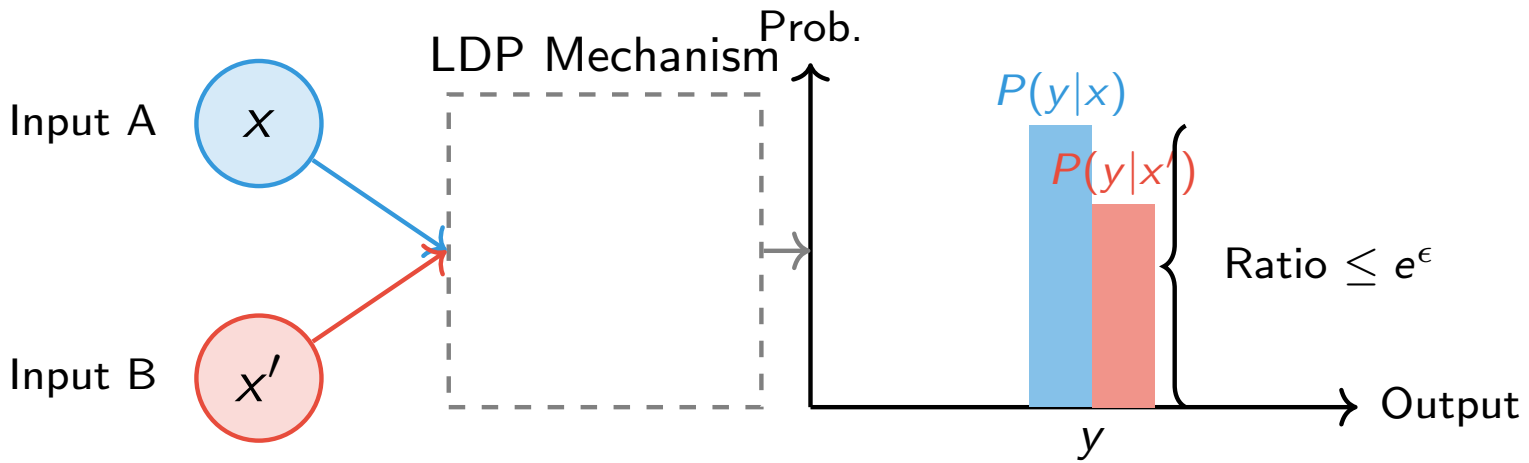
$$\frac{P(\text{Output } y \mid \text{Input } x)}{P(\text{Output } y \mid \text{Input } x')} \leq e^{\epsilon}$$

Local Differential Privacy: The Definition

Local differential privacy constraint for randomized mechanism

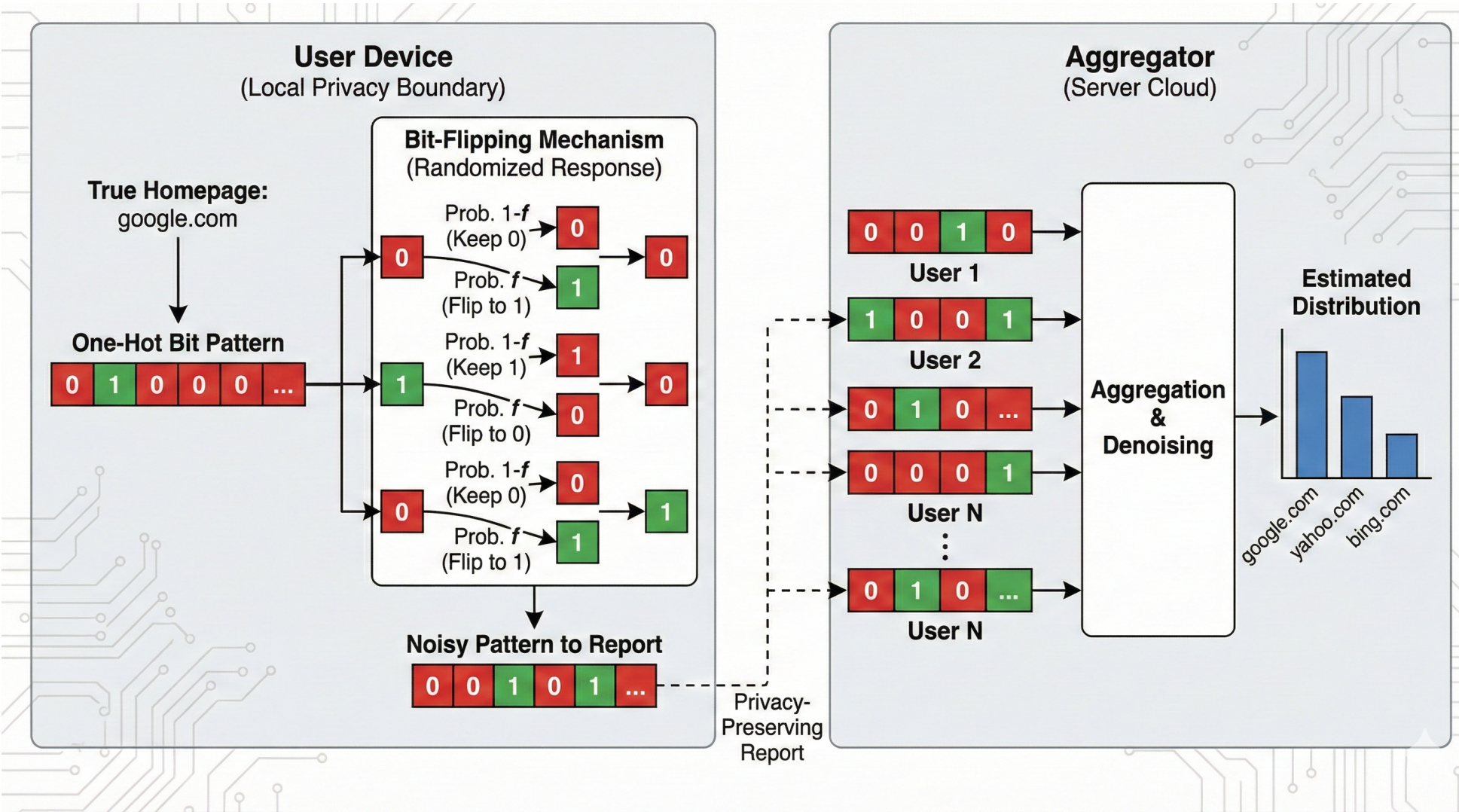
LDP requires that for any two distinct inputs x and x', and any output y:

$$\frac{\mathbb{P}(\text{Output } y \mid \text{Input } x)}{\mathbb{P}(\text{Output } y \mid \text{Input } x')} \leq e^{\epsilon}$$



Intuition: If ε is very small, the probability of seeing output y is almost the same, regardless of whether the user said "Yes" (x) or "No" (x').

Google RAPPOR: bit-flipping



Research Question: From Estimation to Testing

- **Motivation: Malware Detection**
  - Goal: Detect distributional shifts in Chrome homepage settings between consecutive timestamps, under privacy-preserving noises.
  - This is **Two-Sample Testing**; Given two sets of samples drawn from two distributions P and Q, decide whetheher $P = Q$.

- **Strategy**
  - Two-sample testing = Distance estimation + Decision thresholding

- **Key Inquiries**
  1. Is RAPPOR optimal for testing (as it is for estimation)?
  2. How can we efficiently estimate the distance under privacy noise?
  3. How to rigorously define decision thresholds given privacy noise?
  4. Can we quantify the fundamental accuracy-privacy trade-off?
  5. Do other algorithms compete with RAPPOR?
  6. Can RAPPOR be adapted to test continuous distributions?

Direct Distance Estimation by U-statistic

- **Goal:** Estimate the squared distance between distribution P and Q.
- **Mechanism:** Compute similarities for every combination of data points and aggregate them.

U-statistic: Unbiased Estimator of Distance

$$\text{Distance} = \begin{pmatrix}\text{Average Similarity}\\ \text{within P}\end{pmatrix} + \begin{pmatrix}\text{Average Similarity}\\ \text{within Q}\end{pmatrix} - 2 \times \begin{pmatrix}\text{Average Similarity}\\ \text{between P and Q}\end{pmatrix}$$

- **Interpretation:**
  - If $P = Q$, the "Between" similarity equals the "Within" similarity. The terms cancel out, yielding an expected value of **0**.
  - If $P \neq Q$ are different, the "Between" similarity drops, and the total value becomes **positive**.
  - **Directly** estimating the distance avoids error amplification induced by squaring the differences between noisy distribution estimates

Threshold Determination: Why Permutation?

The Challenge: Asymptotics Fail

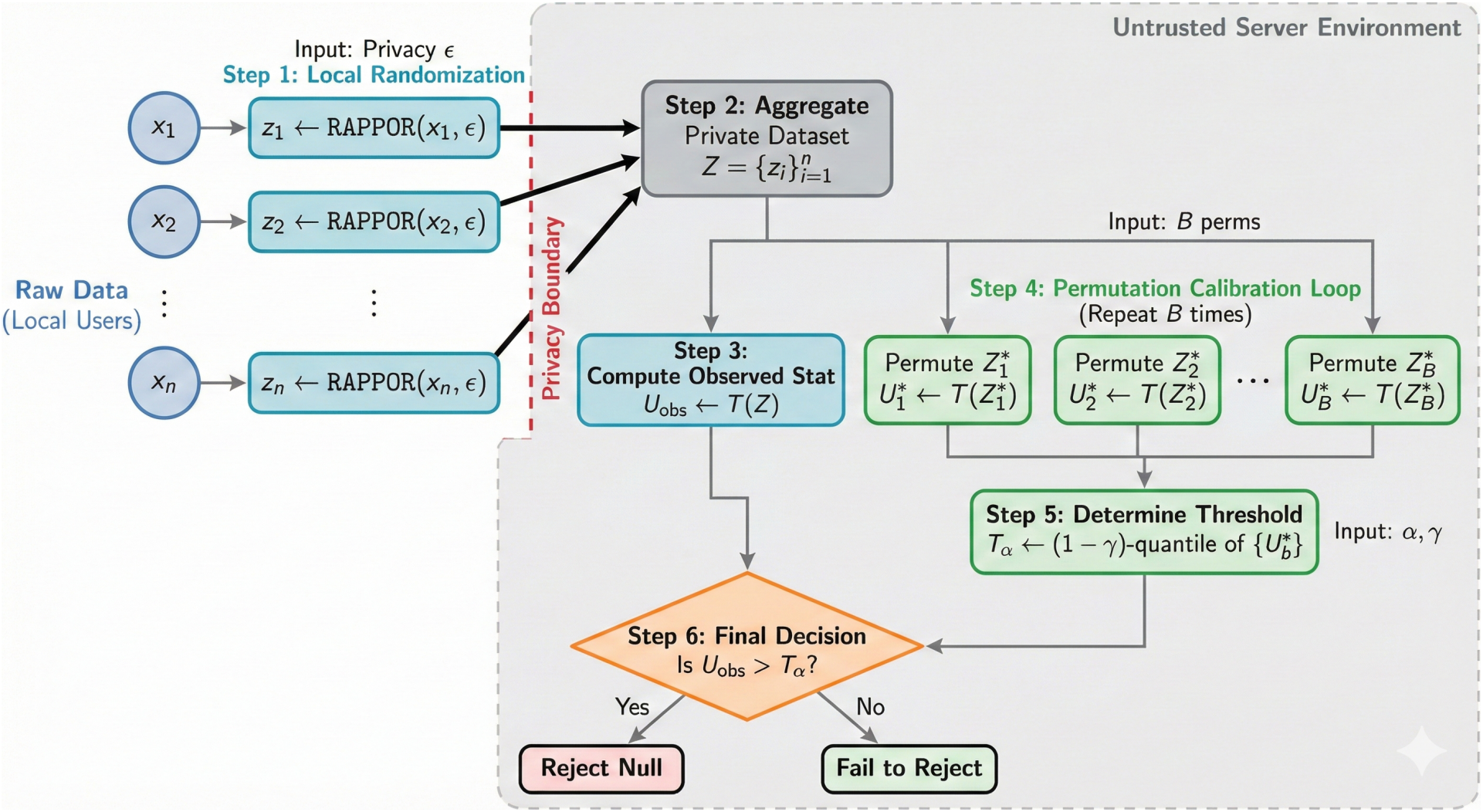
- **Standard Approach:** Relies on asymptotic limit ($n \to \infty$) distribution (e.g., Chi-square) to derive rejection threshold using tail probability
- **Privacy Problem:** RAPPOR injects massive non-standard noise, causing the statistic to converge too slowly to these theoretical limits.
- **Result:** Theoretical thresholds yield invalid Type I errors (false positives) on finite data.

The Solution: Permutation Tests

- **Method:** Shuffle group labels to empirically reconstruct the *exact* null distribution from the data itself.
- **Guarantee:** Provides rigorous Type I error control for **any** sample size n and **any** privacy budget ϵ.

Full algorithm

Implemented as python package: https://pypi.org/project/privateAB/



Theoretical Guarantees and Trade-offs

1. Optimality

Given a fixed privacy budget ϵ and desired target error rates (Type I and Type II), our algorithm reuqires the minimum sample size required to detect a distributional shift, among all possible algorithms

2. Quantifying the Privacy-Accuracy Trade-off

We provide a precise quantification of the "cost" of Local Differential Privacy.

- In the high-privacy regime ($\epsilon < 1$), to match the testing error of a non-private test with n samples, an LDP test typically needs about n/ϵ^2 samples.

Additional Results: Flexibility and Optimality

- **Extension to Continuous Data**
  - *Strategy:* Bin continuous values into discrete cells ("discretize-then-test").
  - *Result:* This approach preserves **minimax optimality** for density testing under standard smoothness assumptions.

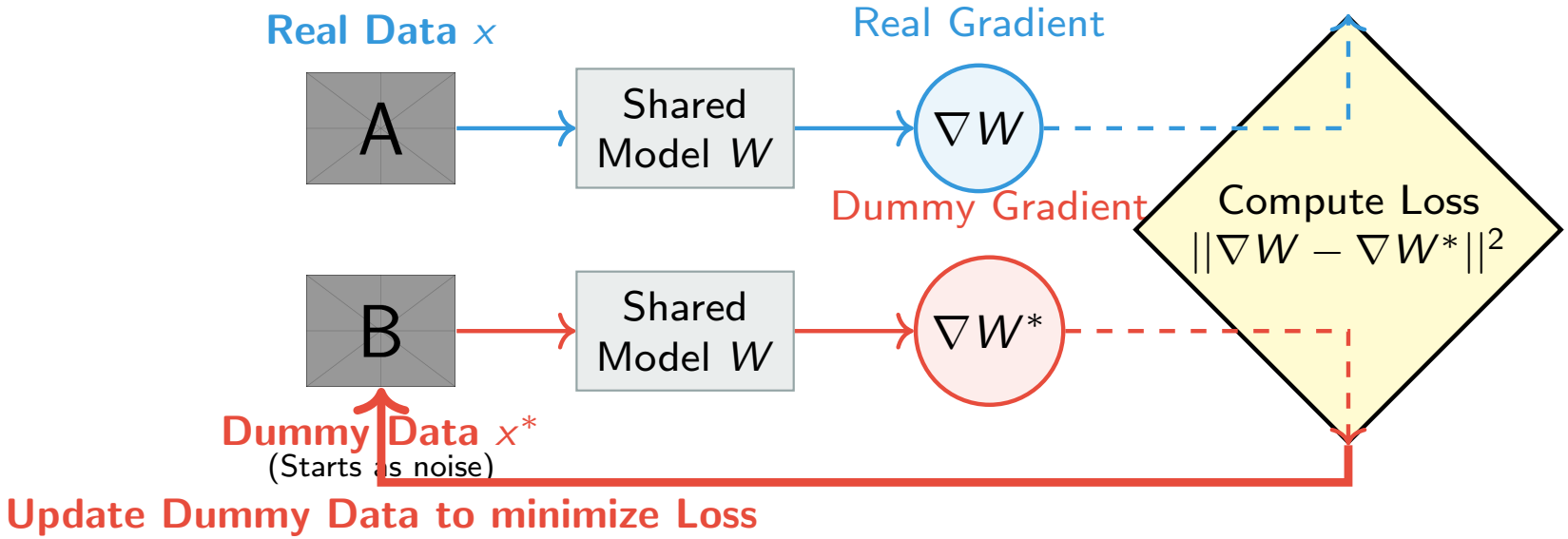
- **Flexibility of Mechanisms**
  - The framework is not limited to RAPPOR.
  - Other ϵ-LDP mechanisms (e.g., adding **Laplace noise** to one-hot encodings) also achieve **optimal utility** within this testing framework.

Summary: Optimal Distribution Testing under LDP

- **Problem Formulation**
  - *Context:* Federated Analytics for malware detection (e.g., Chrome homepage shifts), using Google's RAPPOR mechanism
  - *Task:* Two-sample testing ($H_0 : P = Q$) under Local Differential Privacy constraints.
- **Methodological Framework**
  - **Estimation:** Uses an **unbiased U-statistic** to directly estimate distribution distance under privacy
  - **Calibration:** Employs a **Permutation Test** to determine rejection thresholds, guaranteeing valid Type I error control for finite samples where asymptotics fail under privacy
  - **Implemented as Python package** at pypi.org/project/privateAB/
- **Theoretical Guarantees**
  - **Optimality:** The proposed algorithm achieves the theoretical minimum sample complexity for detecting shifts.
  - **Trade-off:** The cost of privacy is quantified; an LDP test with n/ϵ^2 users has the statistical power of a non-private test with $\approx n$ users.

(Backup) Example: Gradient Inversion Attack

The Intuition: The gradient is calculated using the data. Therefore, the gradient encodes the data.



The Mechanism

The attacker starts with random noise and iteratively tweaks it until the "Dummy Gradient" matches the "Real Gradient." When they match, the noise has transformed into the user's private image.