

# **Oracle-Efficient Differentially Private Learning with Public Data**

**Adam Block, Mark Bun, Rathin Desai, Abhishek Shetty, and Steven Wu**

# Warm-up: (Non-Private) Binary Classification

**PAC Model** [Valiant84]

**Known:**

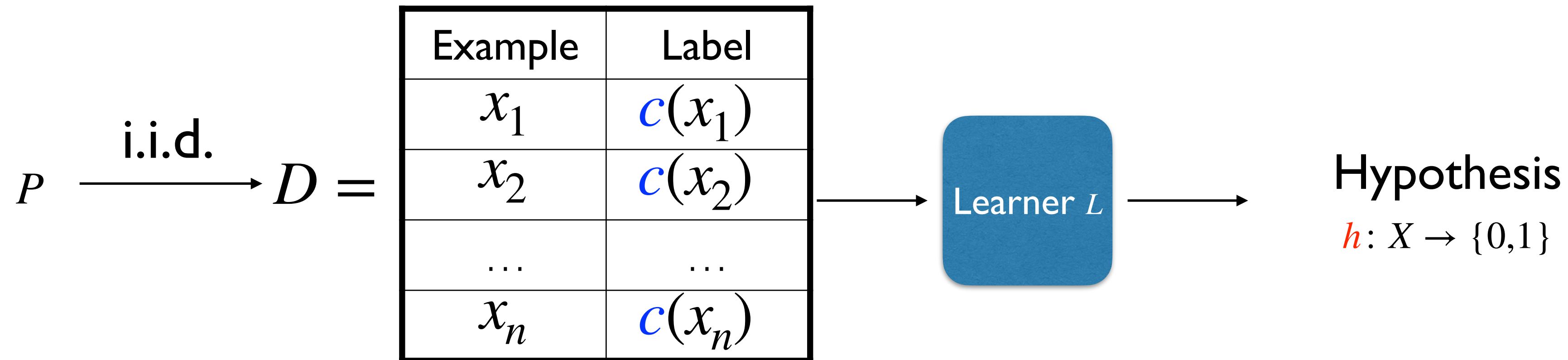
Space of examples  $X$

Concept class  $C = \{f: X \rightarrow \{0,1\}\}$

**Unknown:**

Distribution  $P$  over  $X$

Target concept  $c \in C$



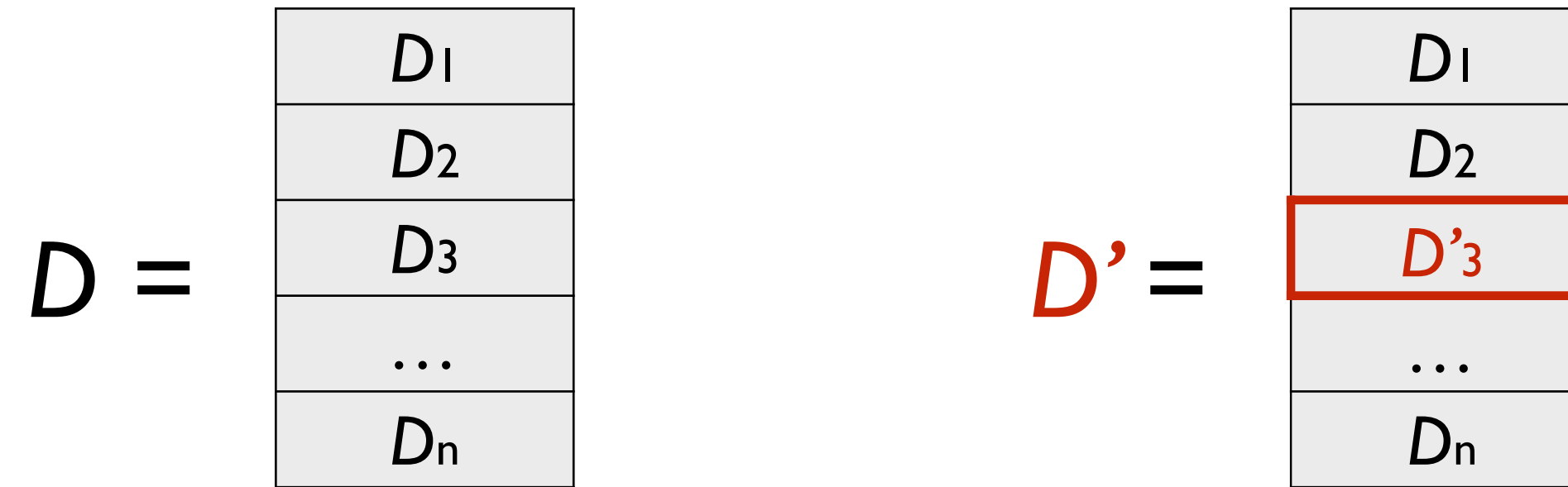
$C$  is PAC learnable if there is a  $L$  such that for all  $P$  and all  $c$

$$\Pr[h(x) \neq c(x)] \leq 0.01$$

with randomness over  $D$  and  $L$

# Differential Privacy

[DN03, DMNS06]



$D$  and  $D'$  are *neighbors* if they differ by at most one row

A private algorithm needs to have close output distributions on any pair of neighbors

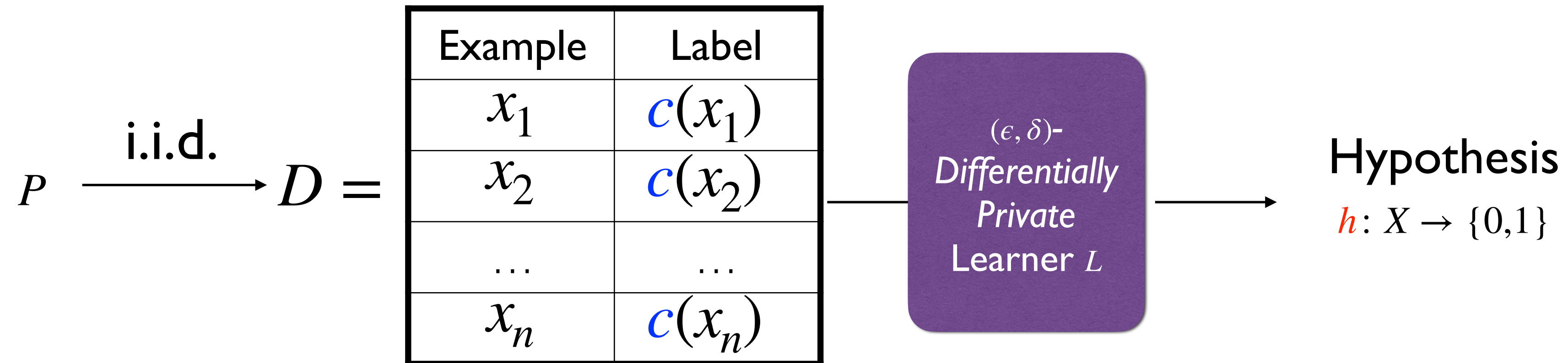
Definition: A (randomized) algorithm  $A$  is  $(\epsilon, \delta)$ -differentially private if for all neighbors  $D, D'$  and every  $S \subseteq \text{Range}(A)$

$$\Pr[A(D) \in S] \leq e^\epsilon \Pr[A(D') \in S] + \delta$$

# Private Binary Classification

## Differentially Private PAC Model

[Kasiviswanathan-Lee-Raskhodnikova-Nissim-Smith08]



Definition: An algorithm  $L$  is  $(\epsilon, \delta)$ -differentially private if for all pairs of  $D, D'$  that differ by one example and every  $S \subseteq \text{Range}(L)$

$$\Pr[L(D) \in S] \leq \exp(\epsilon) \Pr[L(D') \in S] + \delta$$

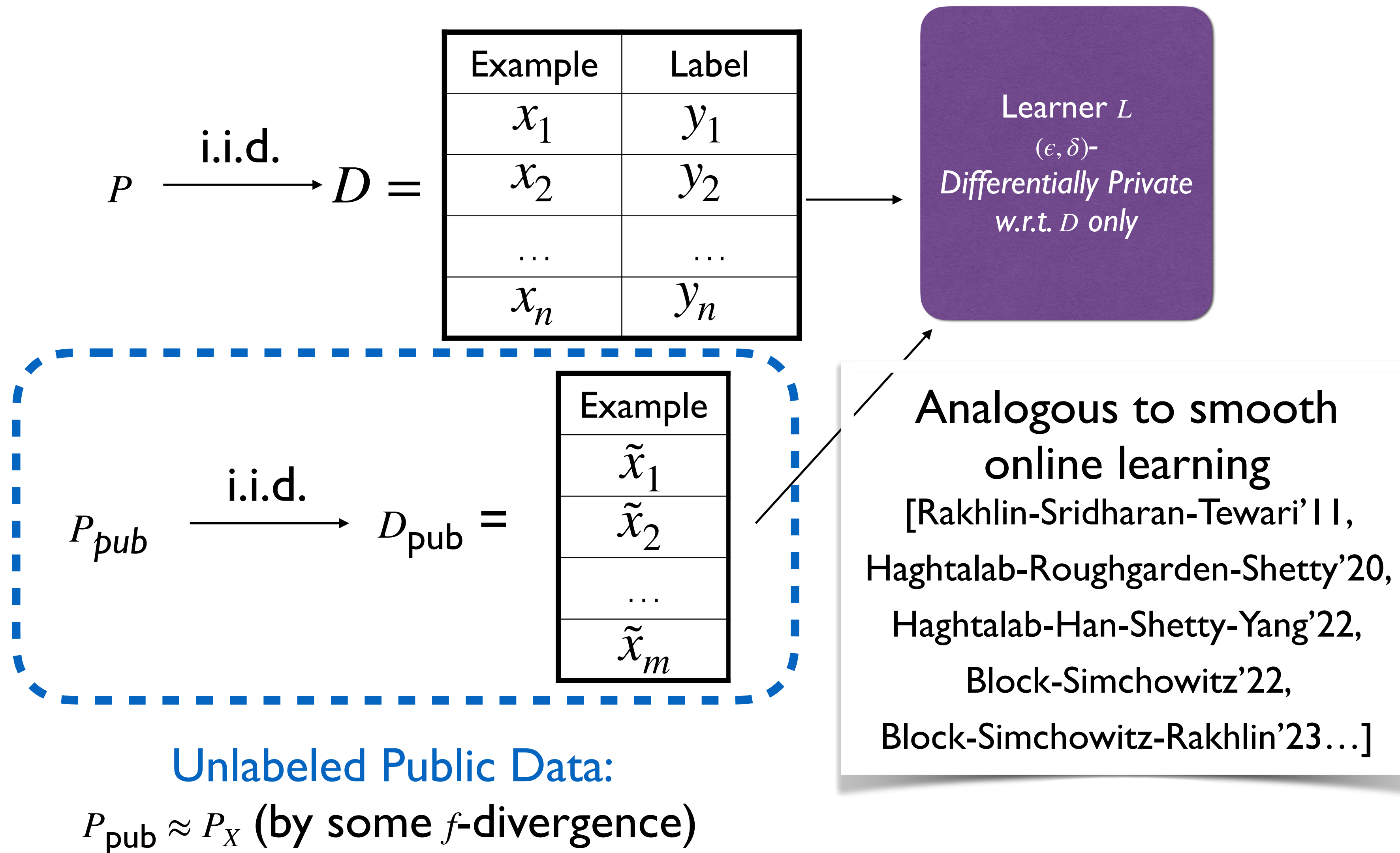
# But...

- **Statistical Feasibility:**
  - Littlestone dimension is a pessimistic *worst-case measure*
  - Rules out simple functions (e.g., thresholds, half-spaces)
  - Does not reflect recent empirical advances in DP ML

Can we leverage external information to sidestep Littlestone dimension lower bound?

# Private Learning with Unlabeled Public Data

[Beimel-Nissim-Stemmer'14, Alon-Bassily-Moran'19...]



# Formulation of Computational Efficiency

Oracle Efficiency:

Abstraction for powerful solvers for non-convex optimization (e.g., SGD, integer program solvers...)

Assume access to an *oracle* that can solve (non-private) empirical risk minimization problem of the form:

$$\arg \min_{h \in C} \sum_{i=1}^n \ell(h(x_i), y_i)$$

Solve the private learning problem efficiently (in polynomial time)

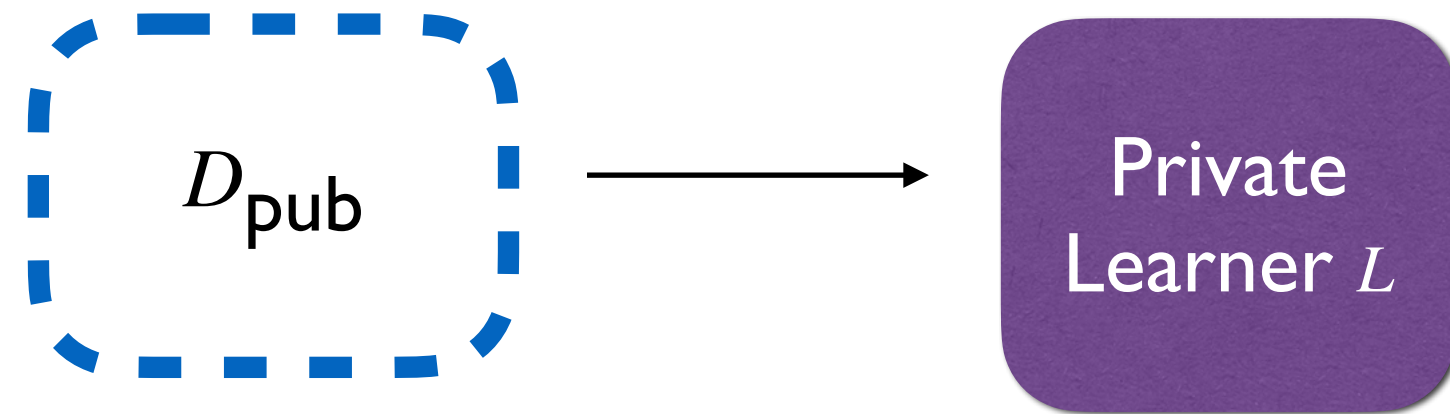
Prime oracle-efficient in online learning:

Follow-the-perturbed-leader (FTPL)

[Kalai-Vempala'02]

# DP Oracle-Efficient Learner with Unlabeled Public Data

[Block-Bun-Desai-Shetty-Wu24]



Unlabeled Public Data for  
Stability

$D_{\text{pub}}$ -FTPL:

$$\tilde{f} = \arg \min_{h \in C} \sum_{i=1}^n \ell(h(x_i), y_i) + \underbrace{\sum_{\tilde{x}_j \in D_{\text{pub}}} w_j \ell(h(\tilde{x}_j), \tilde{y}_j)}_{\text{"Perturbation"}}$$

$w_j$ : Laplace noise;  $\tilde{y}_j$ : random label in  $\{0,1\}$

Beyond Classification:  
Also extends regression with  
convex, Lipschitz loss function  $\ell$

Label  $D_{\text{pub}}$  with  
 $\tilde{f}$

Example	Label
$\tilde{x}_1$	$\tilde{f}(\tilde{x}_1)$
$\tilde{x}_2$	$\tilde{f}(\tilde{x}_2)$
...	...
$\tilde{x}_m$	$\tilde{f}(\tilde{x}_m)$

Solve ERM

Output accurate predictor  $h$   
with sample size  $\approx \text{VC-Dim}(C)^2$



# Summary

- Designed algorithms based on FTPL, FTRL from online learning that ensure stability to get private learning algorithms.
- Improved previous set of results by giving
  - Oracle efficient algorithms for more general function classes
  - Using public unlabelled data as opposed to public labelled data
  - Minimizing number of calls to the oracle
- First to design learning algorithms for real valued functions.