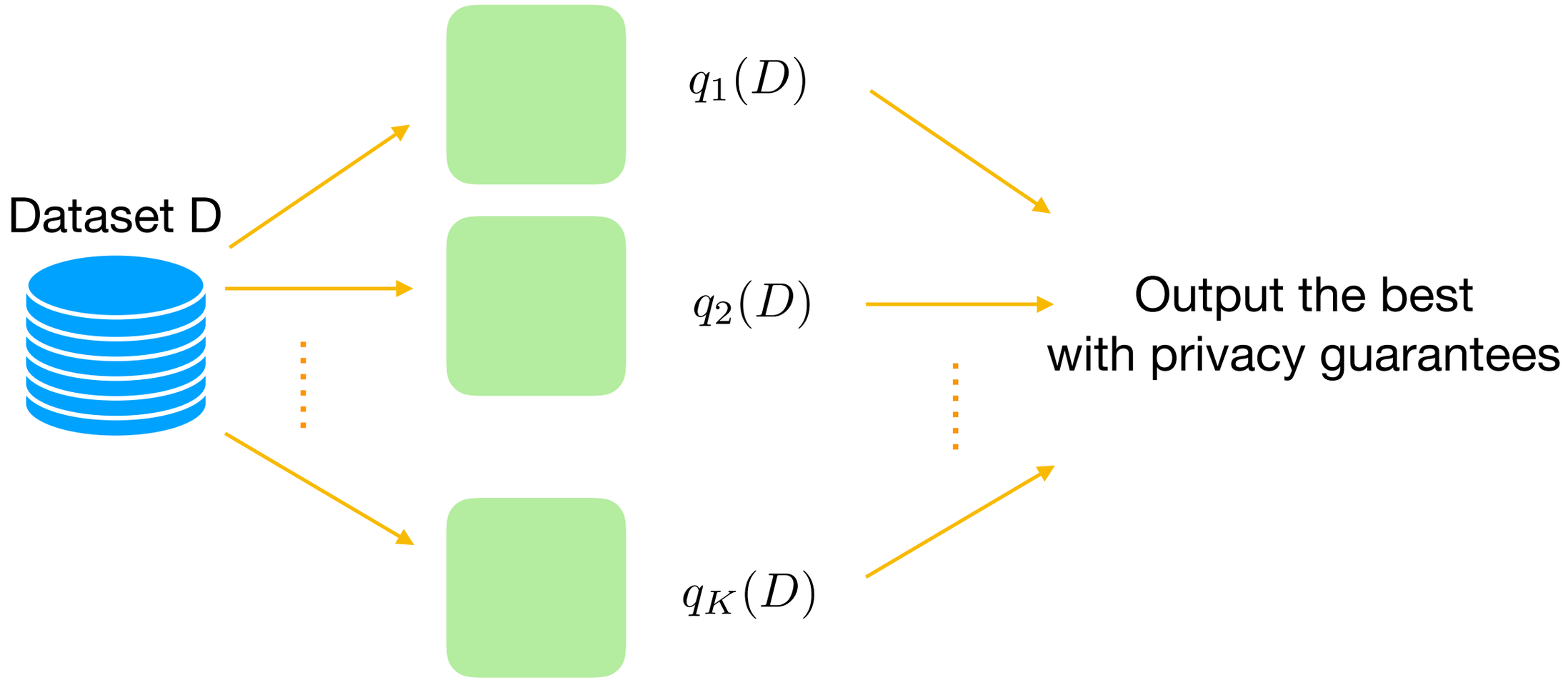
Private Selection from Private Candidates

Jingcheng Liu (UC Berkeley ⟶ Caltech)

Joint work with Kunal Talwar (Google AI)

Private Selection



Life is all about choices!

Differential privacy (DP)

(by Dwork, McSherry, Nissim, Smith)

Whether you are in the data set or not, it makes little difference

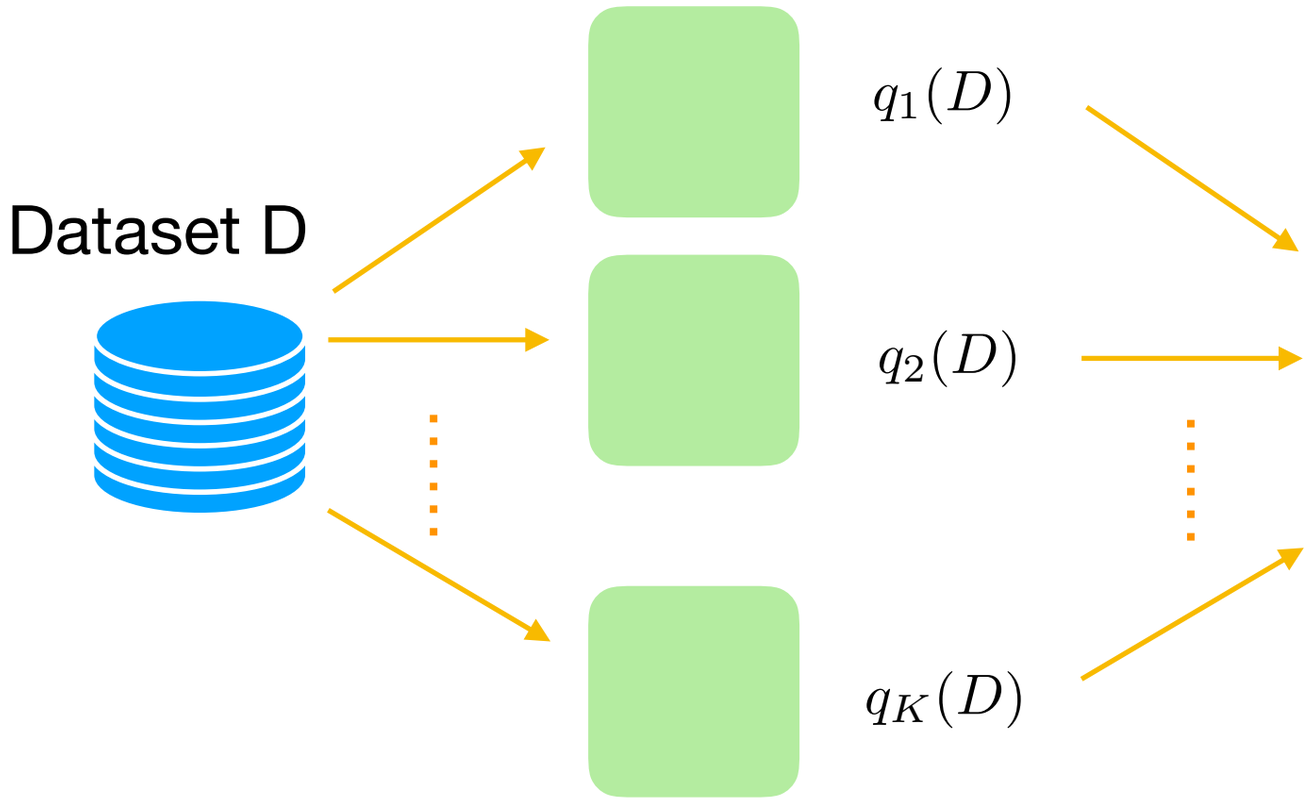
Let $\mathcal{M} : \mathcal{D}^n \to \mathcal{R}$ be a randomized algorithm

We say that $\mathcal{M}$ satisfies $(\varepsilon, \delta) - \mathrm{DP}$ if $\forall D, D' \; s.t. \; |D - D'| \leq 1, \forall S$

$$\Pr[\mathcal{M}(D) \in S] \leq \exp(\varepsilon) \cdot \Pr[\mathcal{M}(D') \in S] + \delta.$$

If $\delta = 0$, we say that $\mathcal{M}$ satisfies $\varepsilon - \mathrm{DP}$

Private Selection



$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$

Exponential mechanism [McSherry, Talwar]

If $\forall i, q_i(D)$ is $1-$Lipschitz: $|q_i(D) - q_i(D')| \leq 1$, then

$$\max\{q_i(D) + \mathrm{Lap}(1/\varepsilon)\} \text{ is } \varepsilon - \mathrm{DP}$$

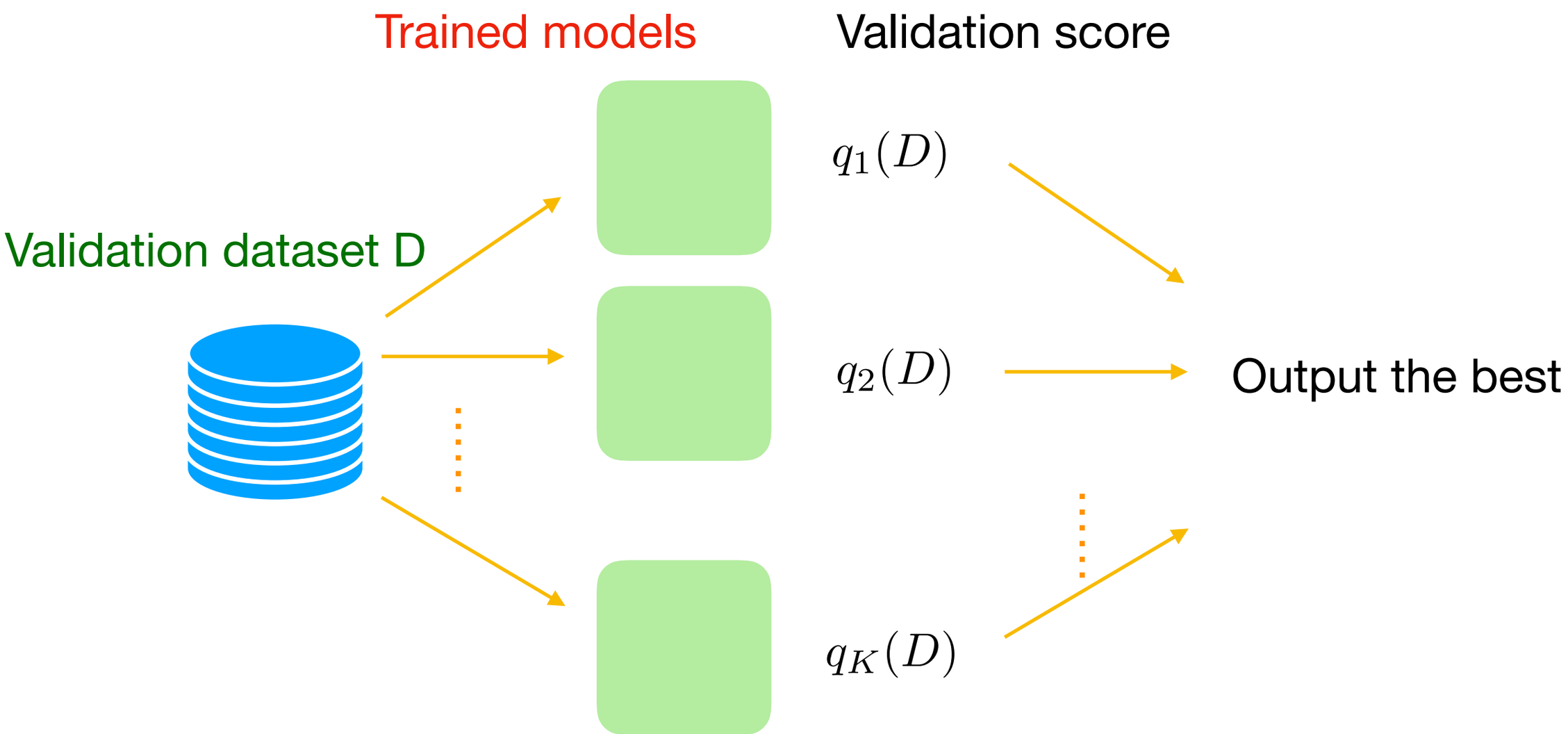
Privacy

Moreover, if the index J is the output of the algorithm, then

$$\Pr\left[q_J(D) \leq \max_i \{q_i(D)\} - \frac{1}{\varepsilon} \ln \frac{K}{\delta}\right] \leq \delta$$

Utility

Private Selection Example



$$q_i(D) := (\text{classification accuracy of the } i\text{-th model})$$

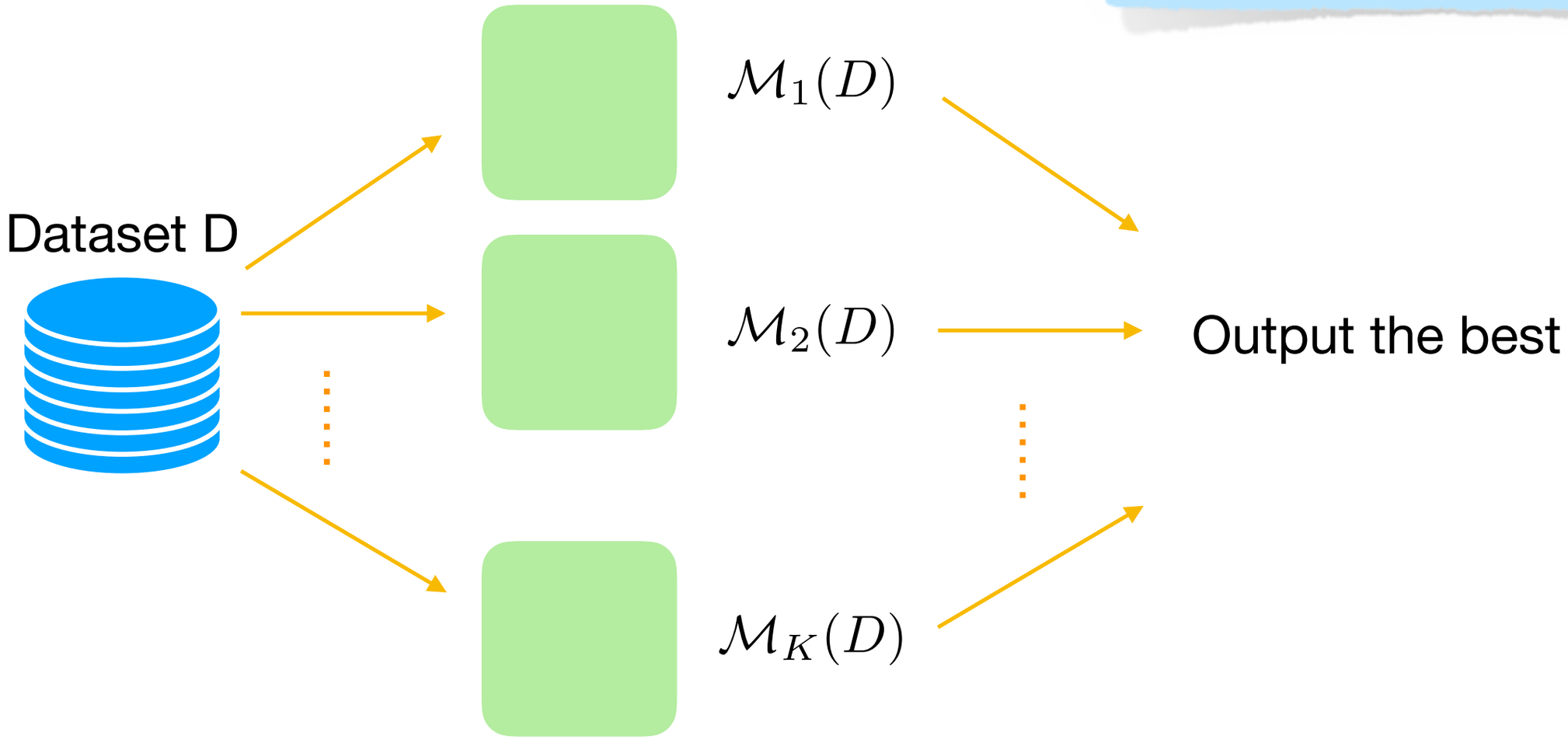
$$\forall i, q_i(D) \text{ is } \frac{1}{n} - \text{Lipschitz: } |q_i(D) - q_i(D')| \leq 1/n$$

Validation dataset ✓

Training dataset ✗

Private Candidates

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$



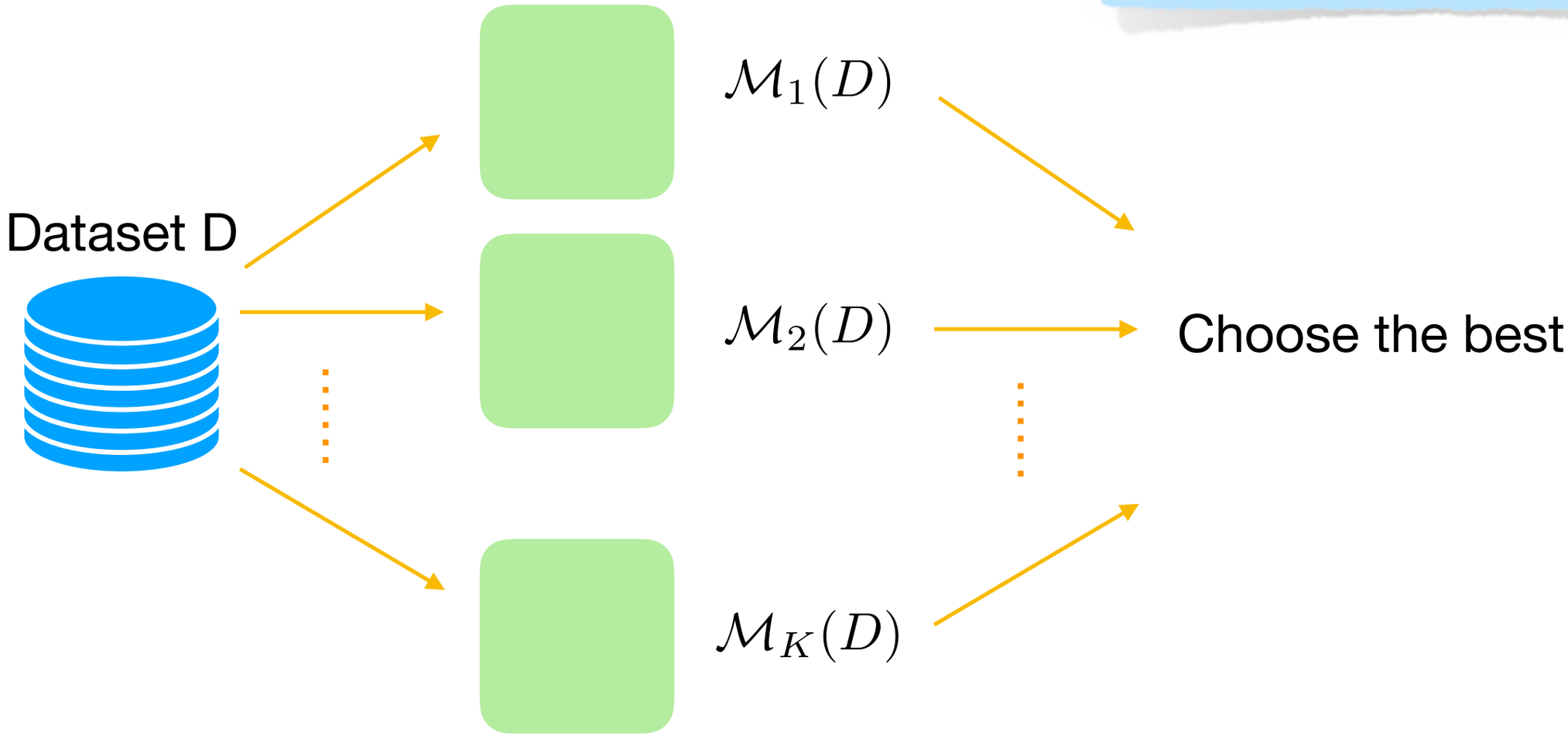
$(x_i, q_i) \leftarrow \mathcal{M}_i(D)$, which are $\varepsilon -$ DP randomized algorithms

output
(e.g. trained model)

score

Private Candidates Examples

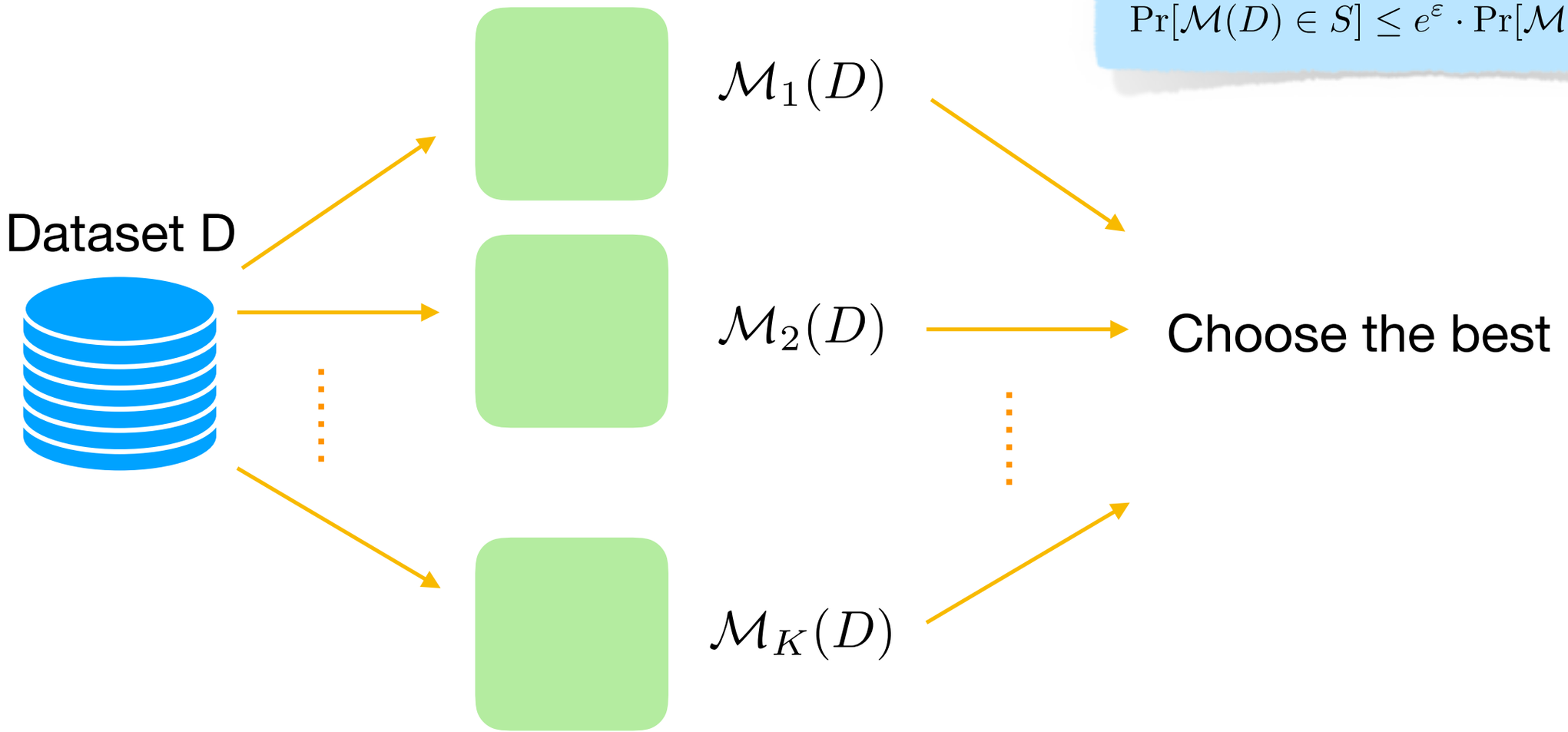
$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$



- Algorithm selection
- Model selection
- Neural network architecture
- Hyperparameters selection
- ....

Private Selection: Naive Attempt #1

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$



$(x_i, q_i) \leftarrow \mathcal{M}_i(D)$, which are $\varepsilon - \mathrm{DP}$ randomized algorithms

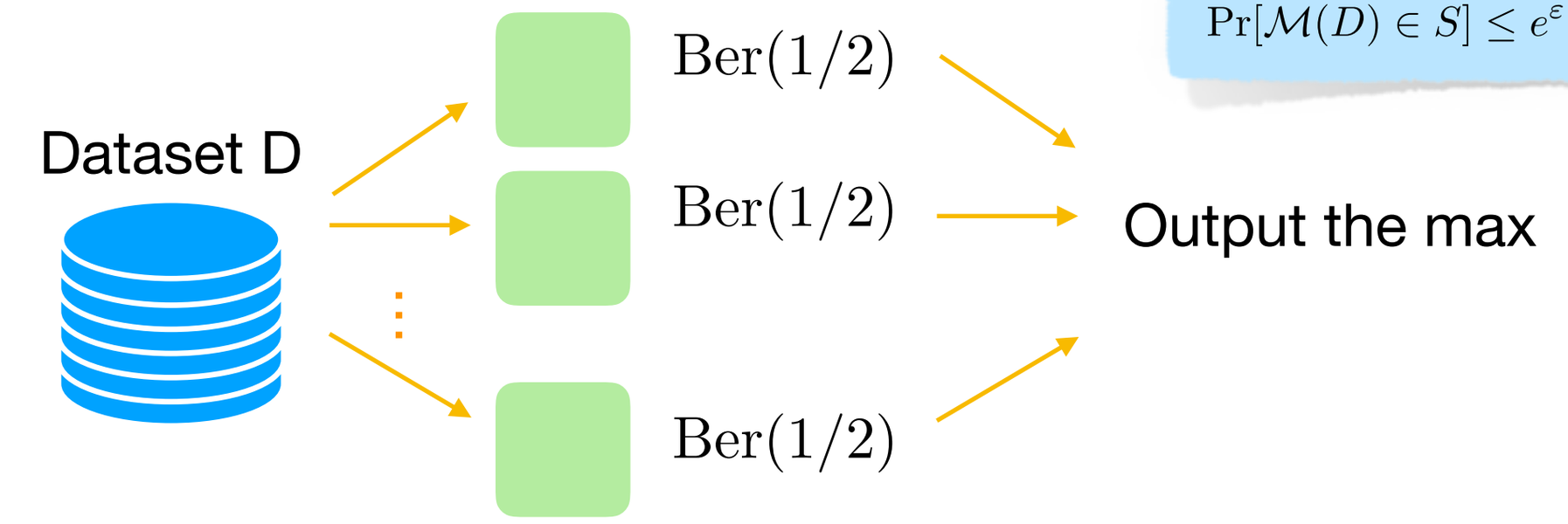
output score

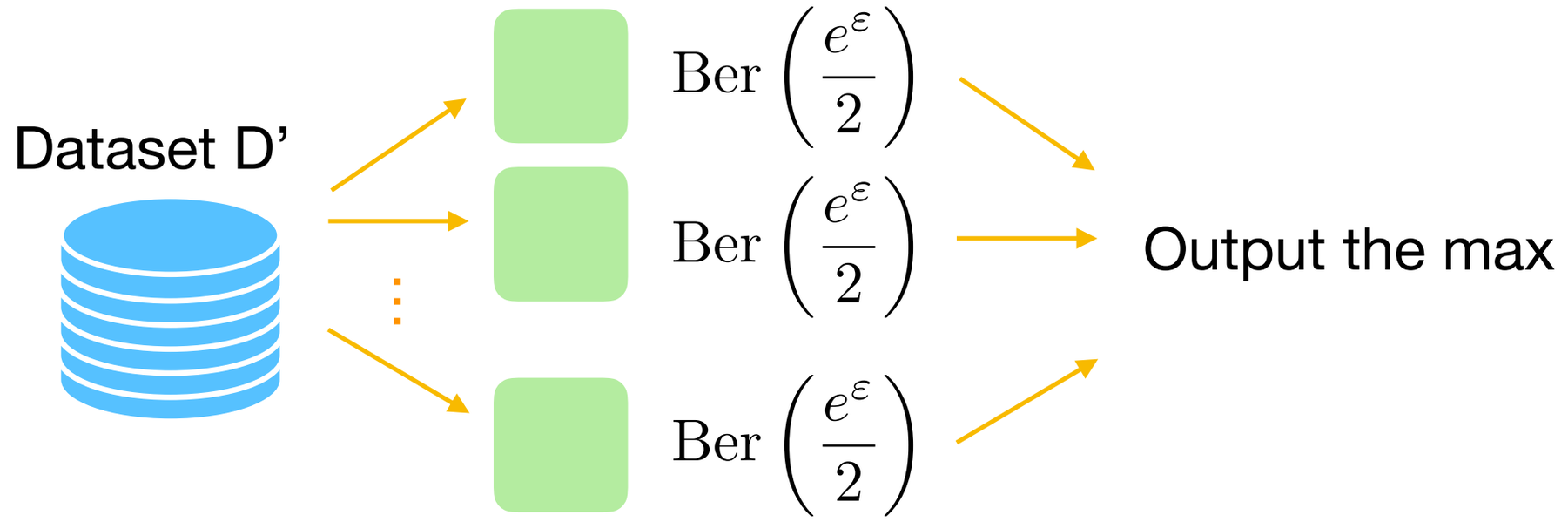
Since $\forall i, \mathcal{M}_i(D)$ is $\varepsilon - \mathrm{DP}$, what if we choose $\max_i \mathcal{M}_i$?

Basic composition: $(K\varepsilon) - \mathrm{DP}$

Tight Example for Attempt #1

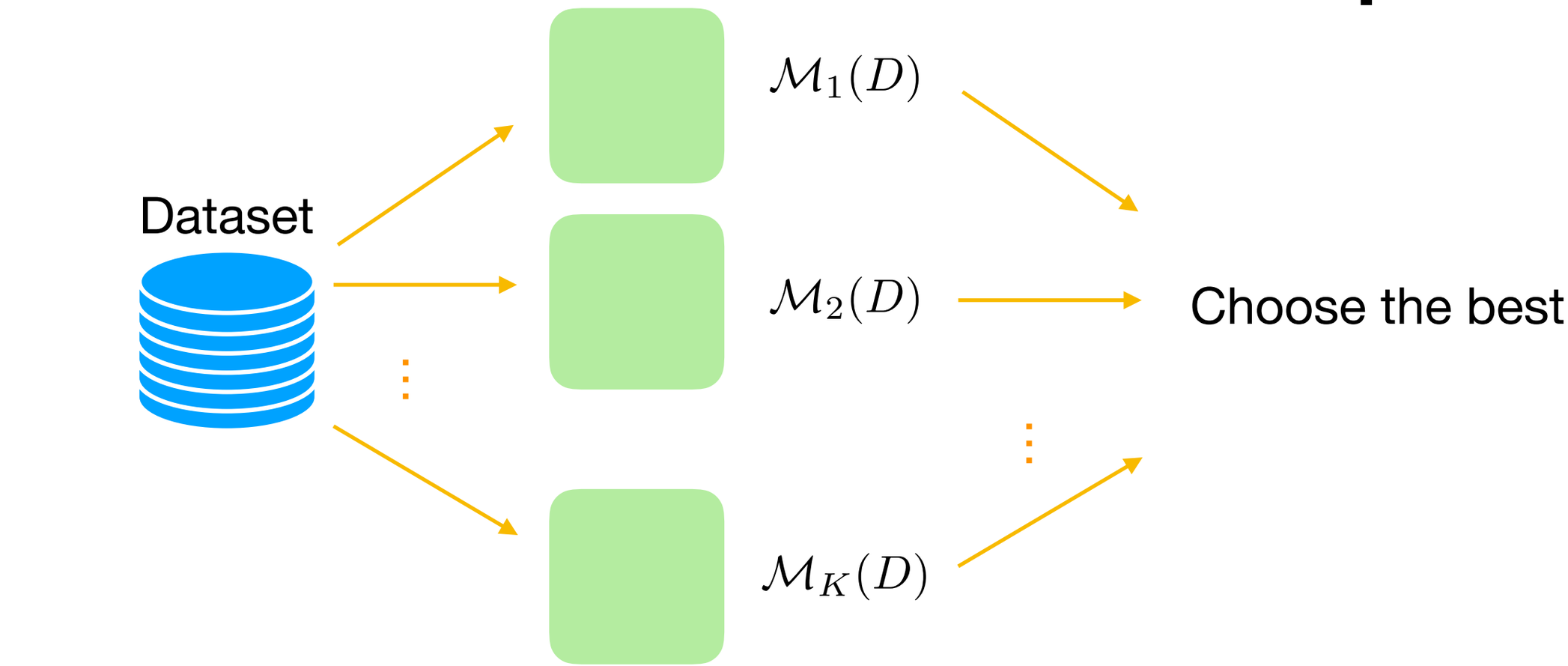
$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$





$$\Pr[\text{output}(D) = 0] = \frac{1}{2^K}, \qquad \Pr[\text{output}(D') = 0] = \left(1 - \frac{e^{\varepsilon}}{2}\right)^K \approx \frac{e^{-K\varepsilon}}{2^K}$$

Private Selection: Naive Attempt #2



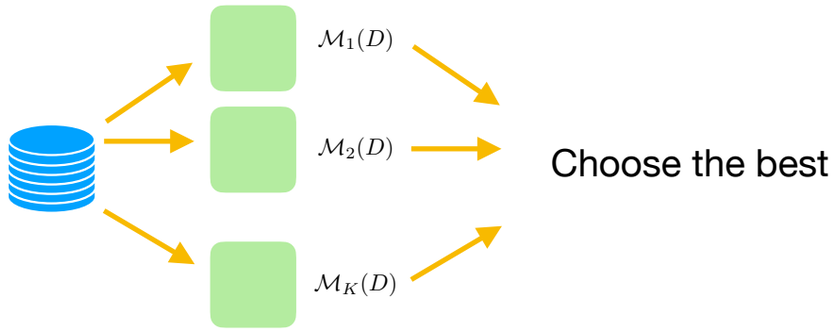
$(x_i, q_i) \leftarrow \mathcal{M}_i(D)$, which are $\varepsilon - \mathrm{DP}$ randomized algorithms

output score

Since $\forall i, \mathcal{M}_i(D)$ is $\varepsilon - \mathrm{DP}$, what if we choose $i \sim [K]$ u.a.r?

Indeed we do get $\varepsilon - \mathrm{DP}$, but the probability of getting the best can be $\frac{1}{K}$

Repetition with Random Stopping



Algorithm 1:

Repeat for T rounds:

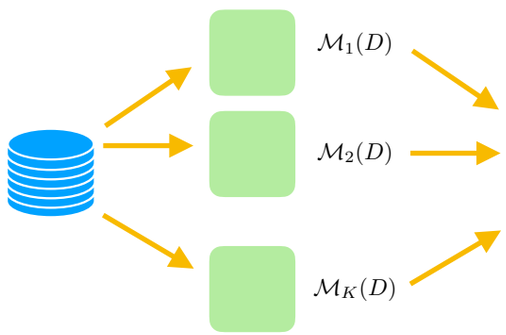
1. Choose $i \sim [K]$ u.a.r
2. Get $(x_i, q_i) \leftarrow \mathcal{M}_i(D)$

$\varepsilon - \mathrm{DP}$

Output the best of seen-so-far

- For every fixed T, Algorithm 1 is $(T\varepsilon) - \mathrm{DP}$
- Our result: if $T \sim \mathrm{Geom}(\cdot)$, then Algorithm 1 is $(3\varepsilon) - \mathrm{DP}$

Repetition with Random Stopping



Choose the best

Algorithm 1:

Repeat for $T \sim \mathrm{Geom}(\cdot)$ rounds:

1. Choose $i \sim [K]$ u.a.r
2. Get $(x_i, q_i) \leftarrow \mathcal{M}_i(D)$

$\varepsilon - \mathrm{DP}$

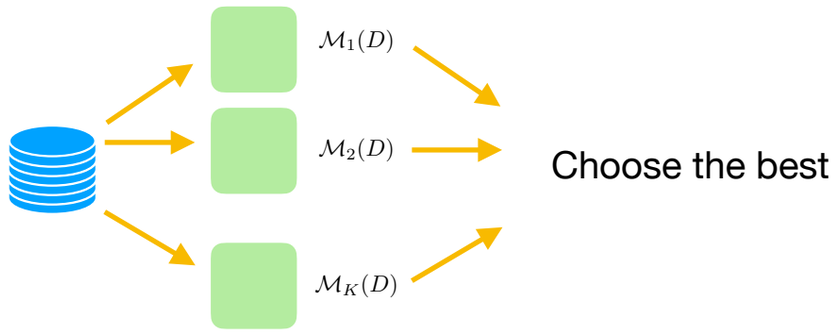
Output the best of seen-so-far

- For every fixed T, Algorithm 1 is $(T\varepsilon) - \mathrm{DP}$
- Our result: if $T \sim \mathrm{Geom}(\cdot)$, then Algorithm 1 is $(3\varepsilon) - \mathrm{DP}$

Remarks:

- Generalizes 1-Lipschitz queries
- For a suitable choice of Geom(.), can match the utility of the Exp. Mech.
- Can be improved to $(2\varepsilon + \varepsilon_0, \delta) - \mathrm{DP}$ with a more sophisticated algorithm

Prior work



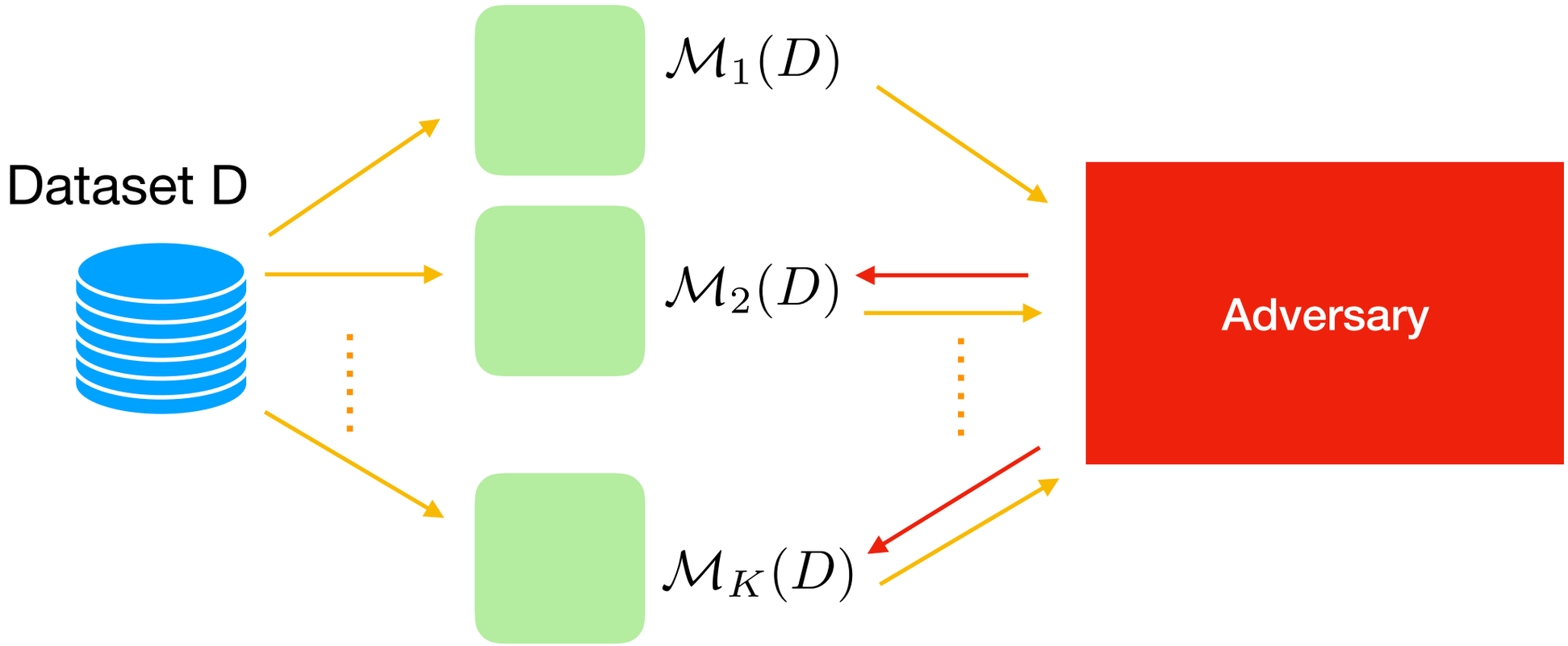
- Assume Lipschitzness: Exponential mechanism [McSherry, Talwar' 07]
- Assume "local" Lipschitzness [Raskhodnikova, Smith' 16]
- Assume a known "good" target [Gupta, Ligett, McSherry, Roth, Talwar' 10]

We are able to improve along

- Privacy: $(3\varepsilon) - \mathrm{DP}$
- Utility: allow unknown target
- Computational (sample) efficiency: nearly linear time (in expectation)

Adaptive and Online

Online Private Selection

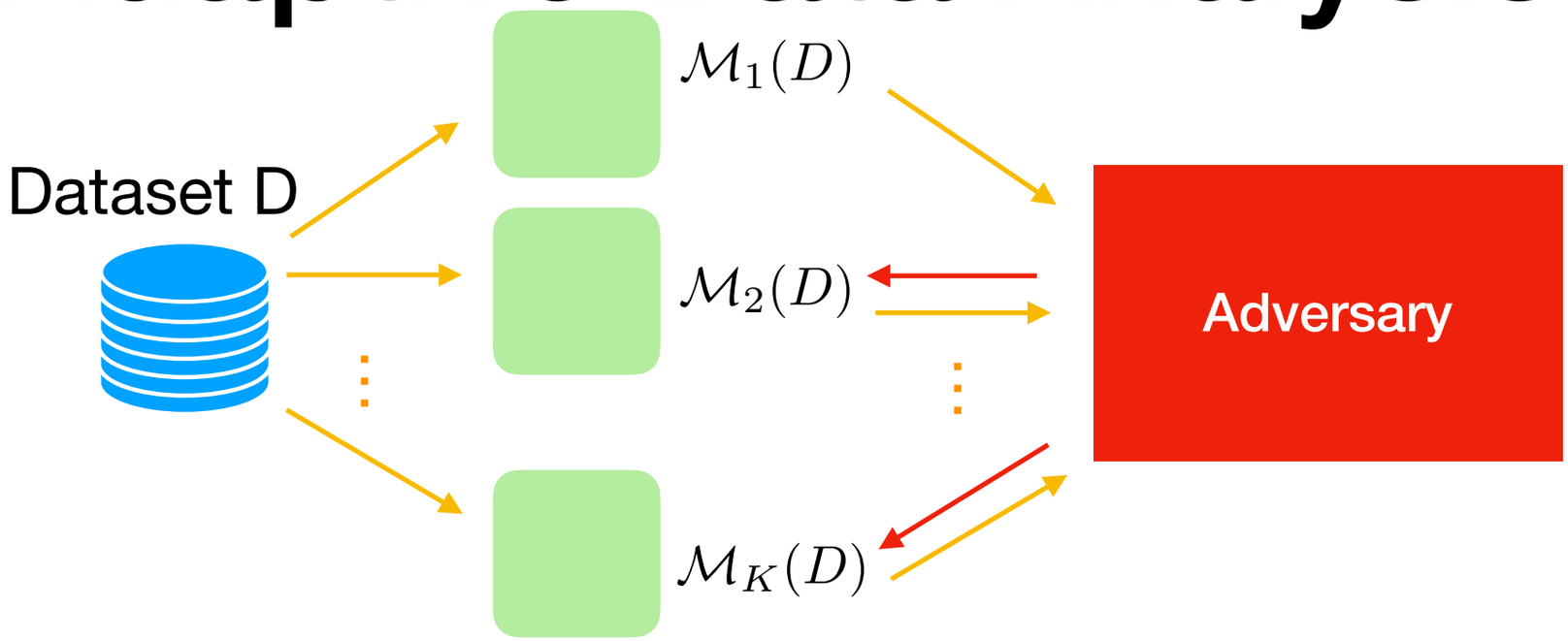


Actually, life is about **online (adaptive)** choices….

Indeed, there is an adaptive version of the **Exp. Mech**.:

- **sparse vector algorithm**
  - due to Dwork, Naor, Reingold, Rothblum, and Vadhan' 09

Adaptive Data Analysis



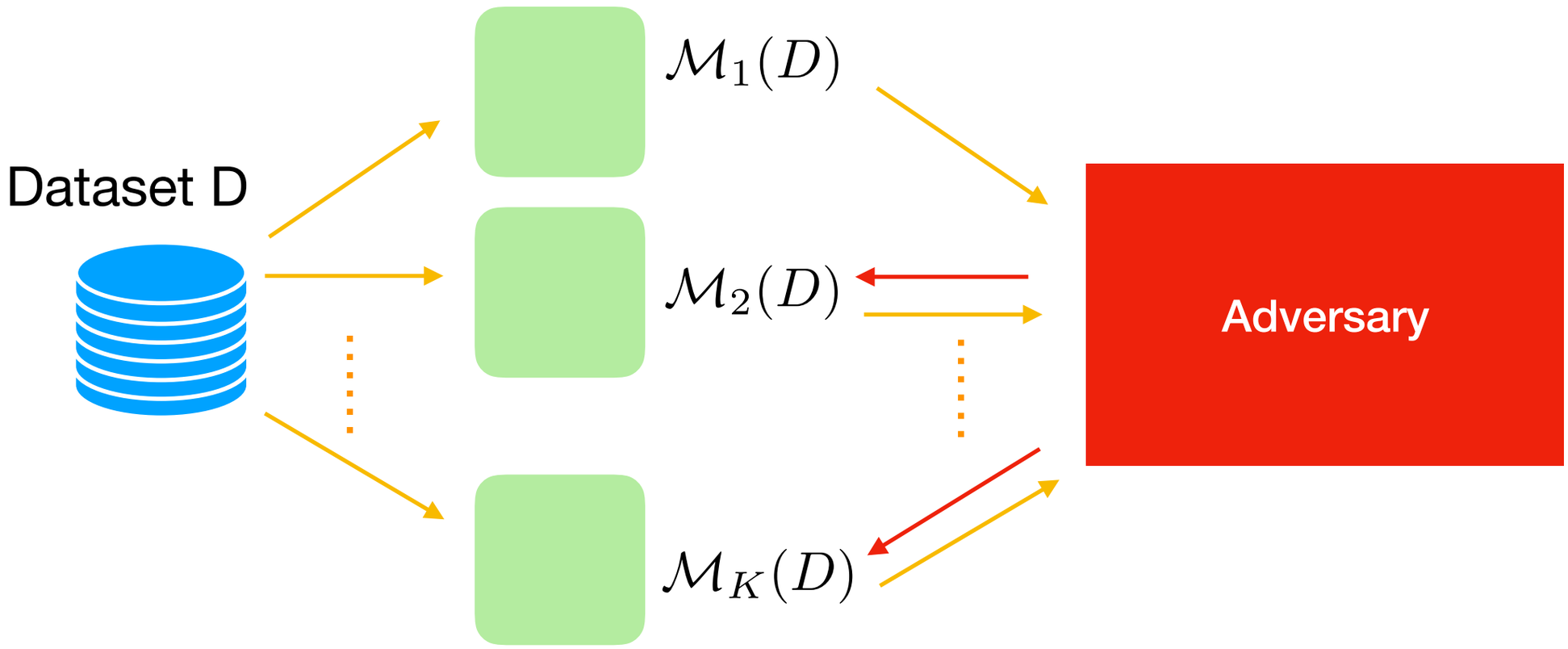
Motivations:

- Adaptive tuning in machine learning
- Adaptive data analysis: garden of forking paths.
  [Dwork, Feldman, Hardt, Pitassi, Reingold, and Roth' 15]

- ...

Sparse vector algorithm only works for Lipschitz queries.
Can we go beyond Lipschitz assumption, e.g. private candidates?

Online Private Selection



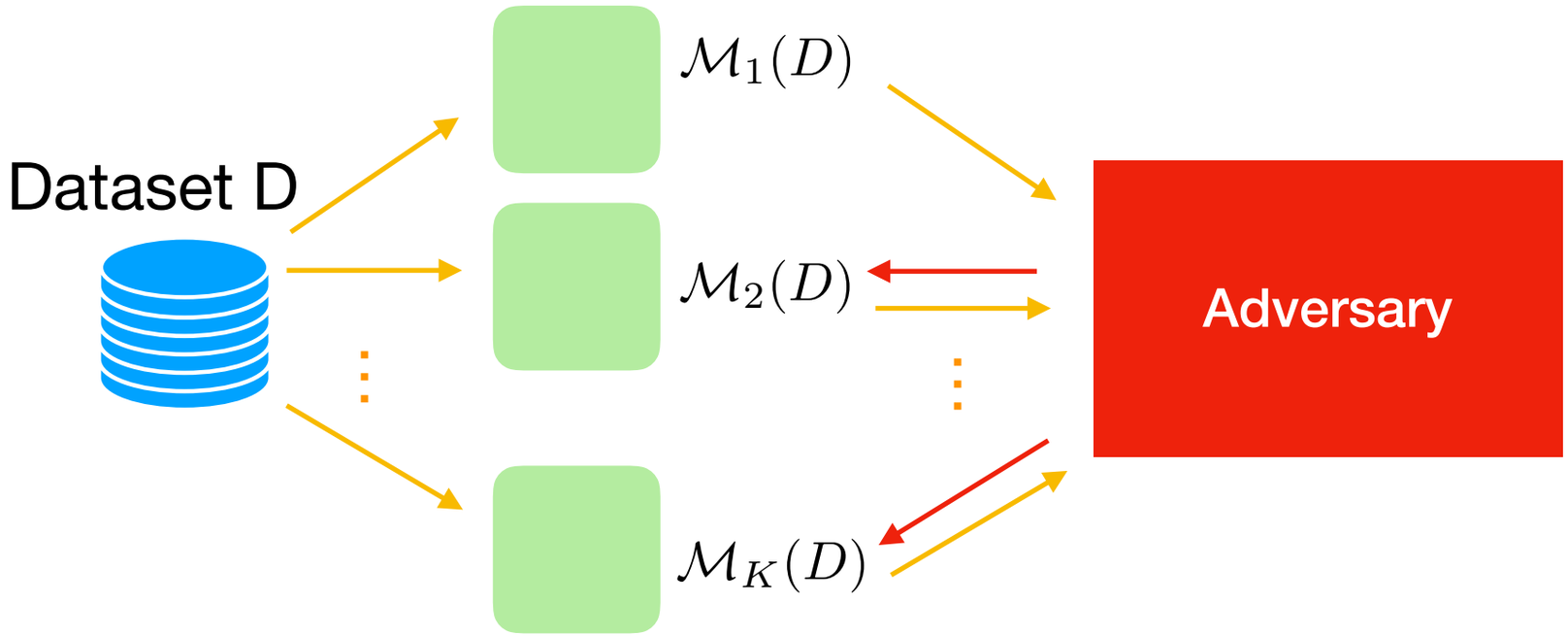
- Allow online queries: $\mathcal{M}_i(D)$ are $\varepsilon-$DP randomized algorithms
- But only threshold queries

Median score := Median(score)

$$\text{Median}\left(\mathcal{M}_i(D)\right) := \sup\left\{\tau \ : \ \Pr_{(\widetilde{x},\widetilde{q})\sim\mathcal{M}_i(D)}[\widetilde{q}\geq\tau]\geq\frac{1}{2}\right\}.$$

Given a threshold τ
Goal: output the first i such that $\text{Median}\left(\mathcal{M}_i(D)\right)\geq\tau$.

Online Private Selection



Median score

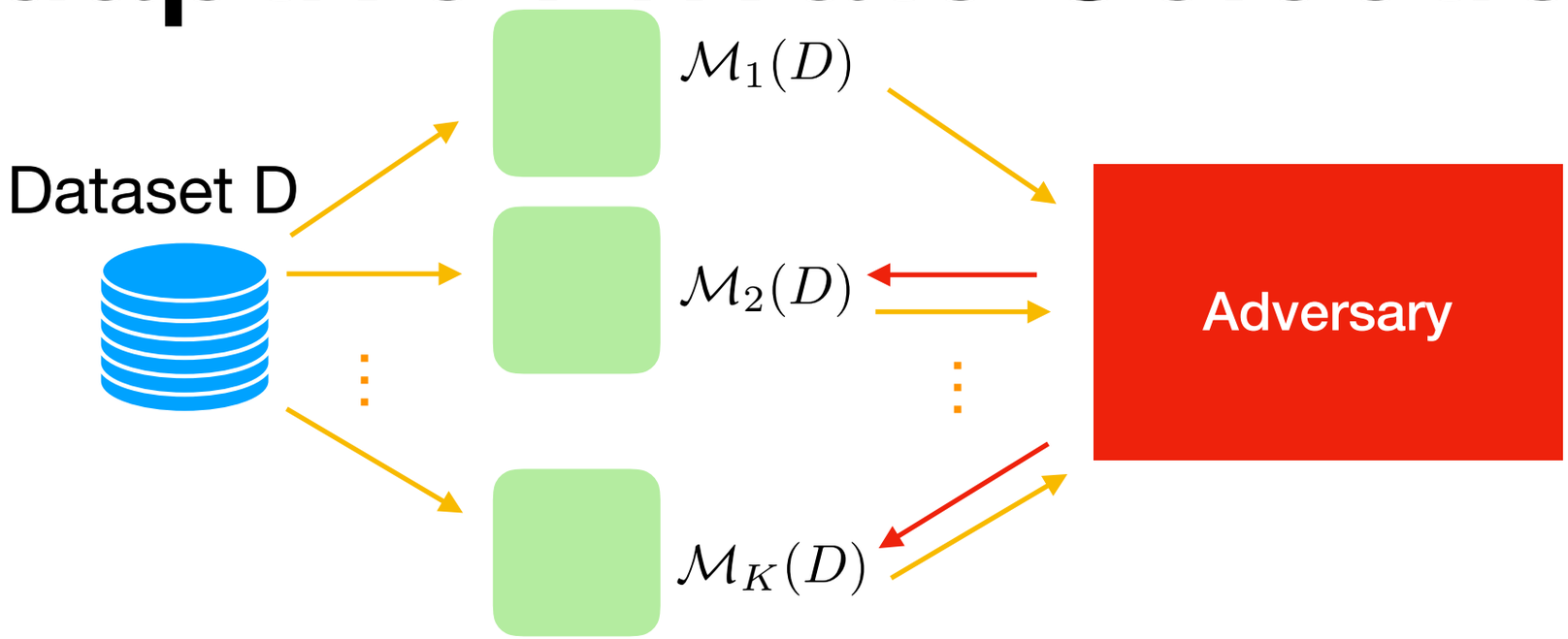
$$\mathrm{Median}\left(\mathcal{M}_i(D)\right) := \sup\left\{\tau \ : \ \Pr_{(\widetilde{x},\widetilde{q})\sim\mathcal{M}_i(D)}[\widetilde{q} \geq \tau] \geq \frac{1}{2}\right\}.$$

Given a threshold τ
Goal: output the first i such that $\mathrm{Median}\left(\mathcal{M}_i(D)\right) \geq \tau$.

Our result: there is an $(\varepsilon, \delta) - \mathrm{DP}$ algorithm such that w.h.p.

- **If a query is above the threshold, then Alg. reports "AboveThreshold"**
- **If Alg. reports i-th query is "AboveThreshold", then the i-th query is not "too much below the threshold" (w.r.t percentile)**

Adaptive Private Selection



Median score

$$\text{Median}\left(\mathcal{M}_i(D)\right) := \sup\left\{\tau \ : \ \Pr_{(\widetilde{x},\widetilde{q})\sim\mathcal{M}_i(D)}[\widetilde{q} \geq \tau] \geq \frac{1}{2}\right\}.$$

Given a threshold τ
Goal: output the first i such that $\text{Median}\left(\mathcal{M}_i(D)\right) \geq \tau$.

Our approach:

- Work with “percentile-score”, which generalizes the median score
- Estimating & testing the “percentile-score” differentially privately
- Bound a variant of Earth mover’s distance between sums of Bernoullis

Open problems & future works:

- Other random stopping time distribution?
- DP preserving mechanism for more sophisticated (hyperparameter) optimization algorithm

Thanks & QA