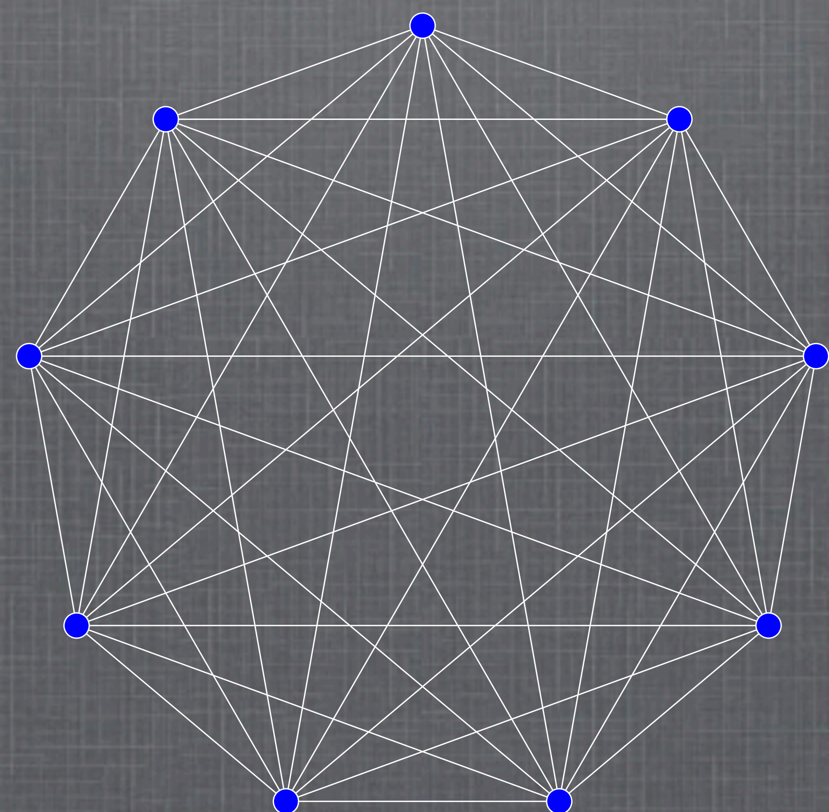
SINGLE-BALLOT RISK-LIMITING AUDITS USING CONVEX OPTIMIZATION

Stephen Checkoway, Anand Sarwate, Hovav Shacham



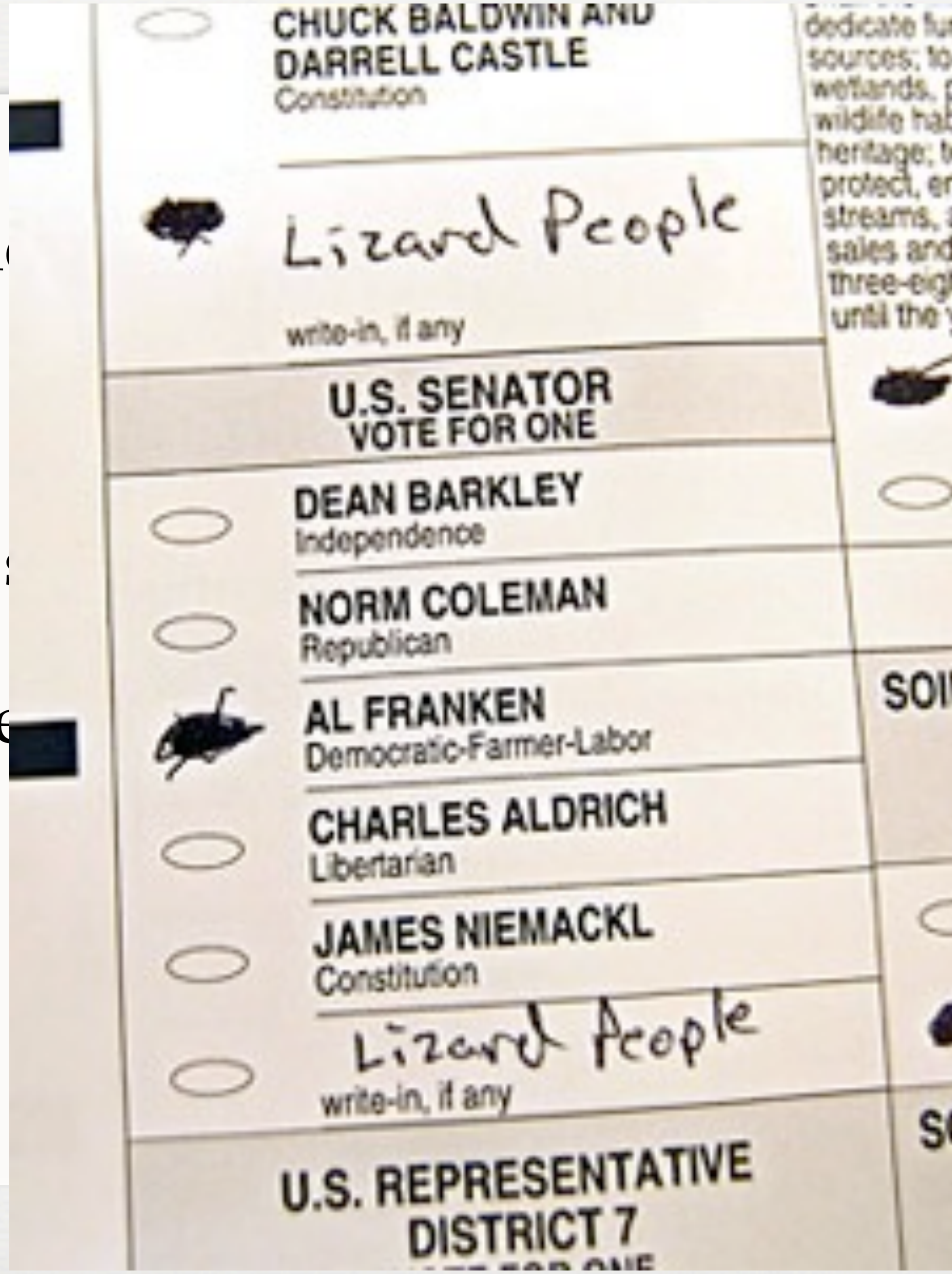
OVERVIEW

- New model of elections
- Simple, ballot-based auditing algorithm

ASSUMPTIONS

- We have electronic Cast Vote Records (CVRs)
- Inspecting ballot reveals voter's intent
- Efficiently sample ballots uniformly at random
- 2 candidates and 1 "no vote" candidate (for this talk)

ASSUMPTIONS

- We have el


Rs)
- Inspecting
- Efficiently dom
- 2 candidate or this talk)

ASSUMPTIONS

- We have electronic Cast Vote Records (CVRs)
- Inspecting ballot reveals voter's intent
- Efficiently sample ballots uniformly at random
- 2 candidates and 1 "no vote" candidate (for this talk)

BALLOTS AS PAIRS

- $\boldsymbol{X}$ = {No vote, Candidate 1, Candidate 2} = {0, 1, 2}
- ith ballot $Z_i = (X_i, Y_i) \in \boldsymbol{X} \times \boldsymbol{X}$
- X_i : actual vote
- Y_i : reported vote (CVR)
- Examples: (2, 0); (1, 2)

ELECTION RESULTS

Reported Votes

Actual Votes

	No vote	Candidate 1	Candidate 2	Actual total
No vote	9,500	80	75	9,655
Candidate 1	130	50,000	40	50,170
Candidate 2	145	30	40,000	40,175
Reported total	9,775	50,110	40,115	

PROBABILITY DISTRIBUTION

- Empirical joint probability distribution

$$M = \begin{pmatrix} 0.09500 & 0.00080 & 0.00075 \\ 0.00130 & 0.50000 & 0.00040 \\ 0.00145 & 0.00030 & 0.40000 \end{pmatrix} \qquad \text{Unknown}$$

$$q = \begin{pmatrix} 0.09775 & 0.50110 & 0.40115 \end{pmatrix} \qquad \text{Known}$$

- Margin 9.995%

AUDITING TASK

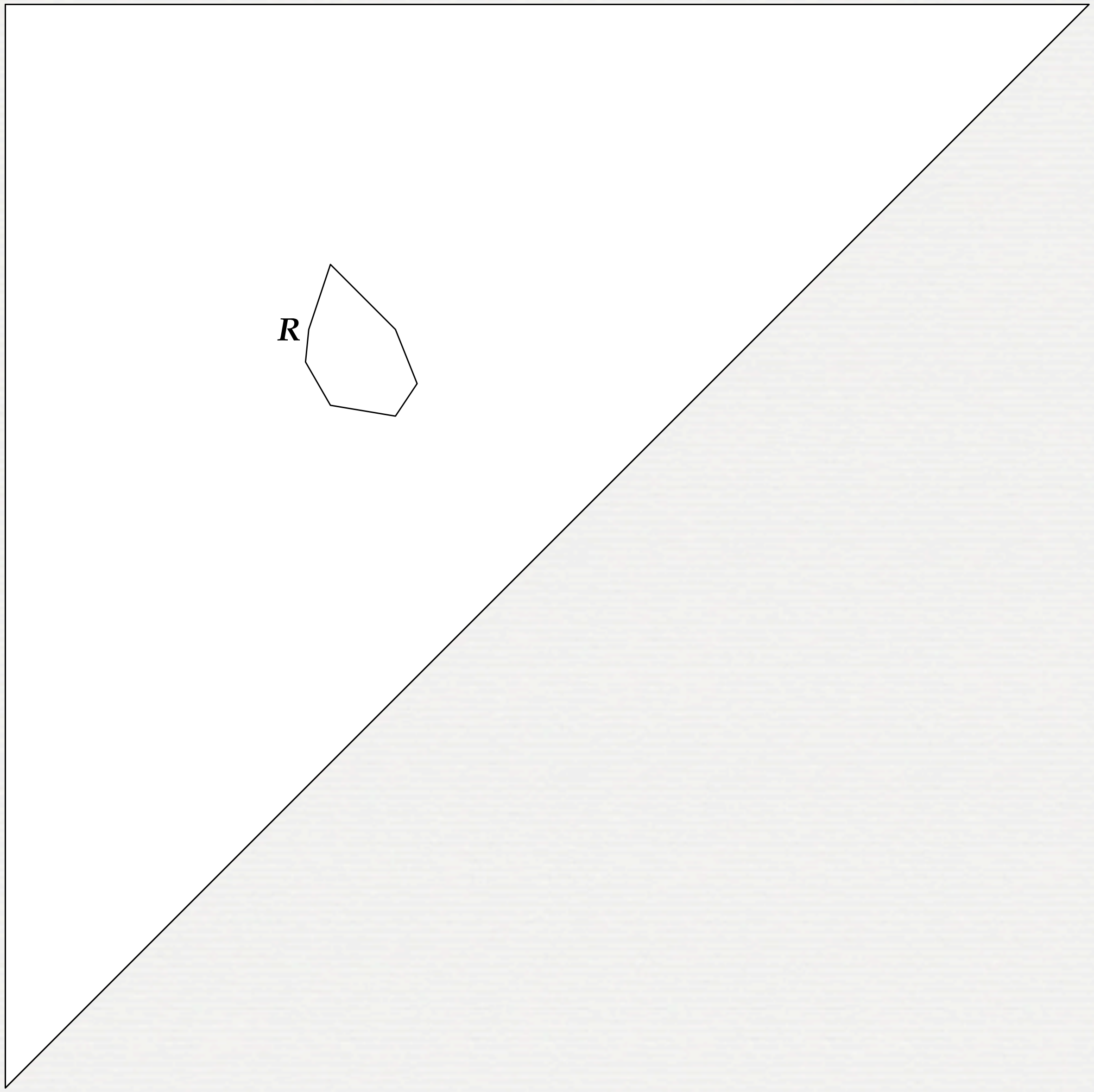
- Sample ballots $Z_1, Z_2, \ldots, Z_K \sim M$
- Task: Decide if the reported winner is actual winner
  - Risk-limiting procedure
- Form estimate $\hat{M}$ of M and compute some function of K and $\hat{M}$

ROTTEN REGION

- Rotten region: a set ***R*** of
  - Joint probability distribution
  - Actual winner and reported winner differ
  - Y-marginal agrees with reported outcome q

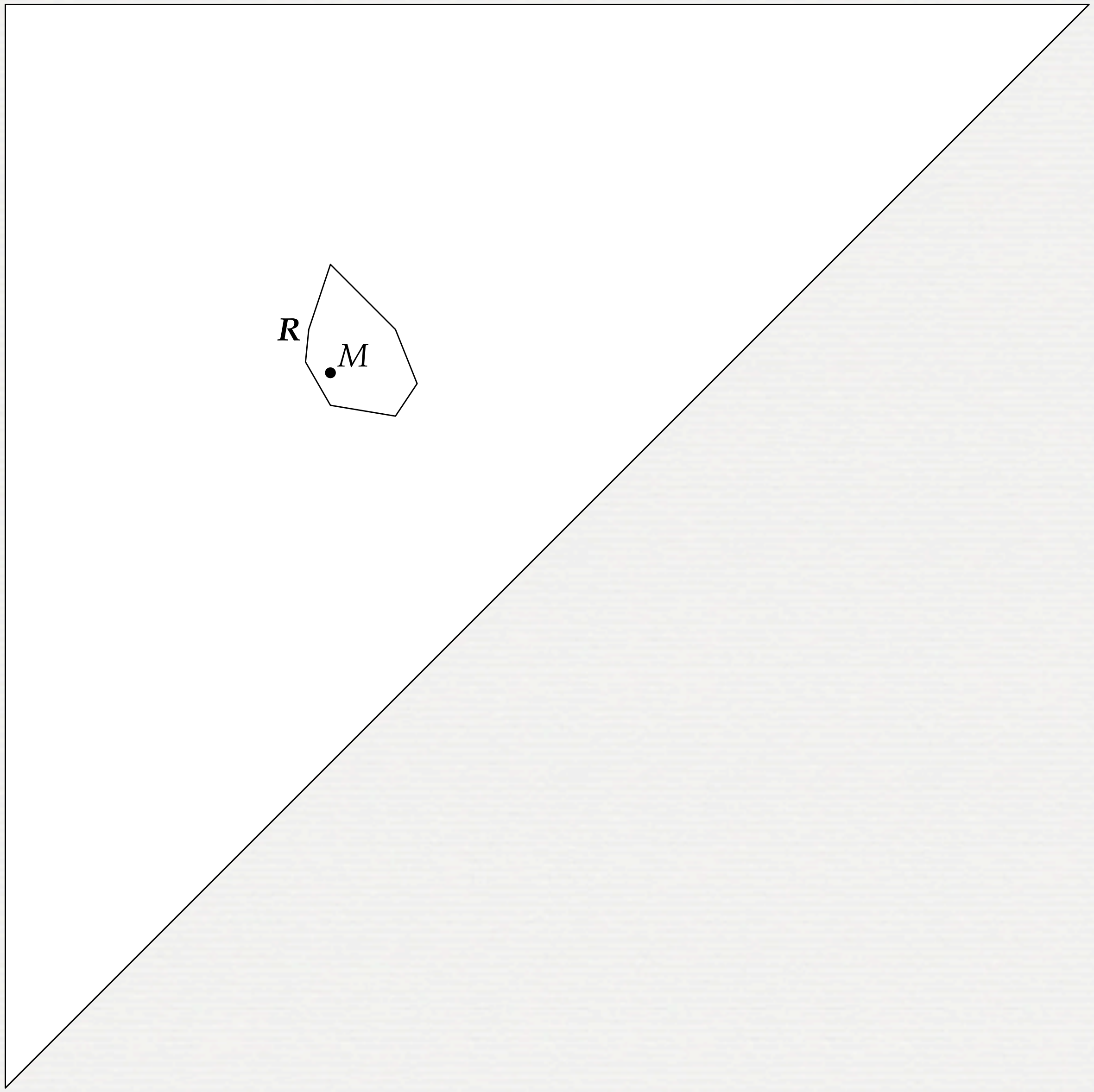
HIGH-LEVEL IDEA

Probability space



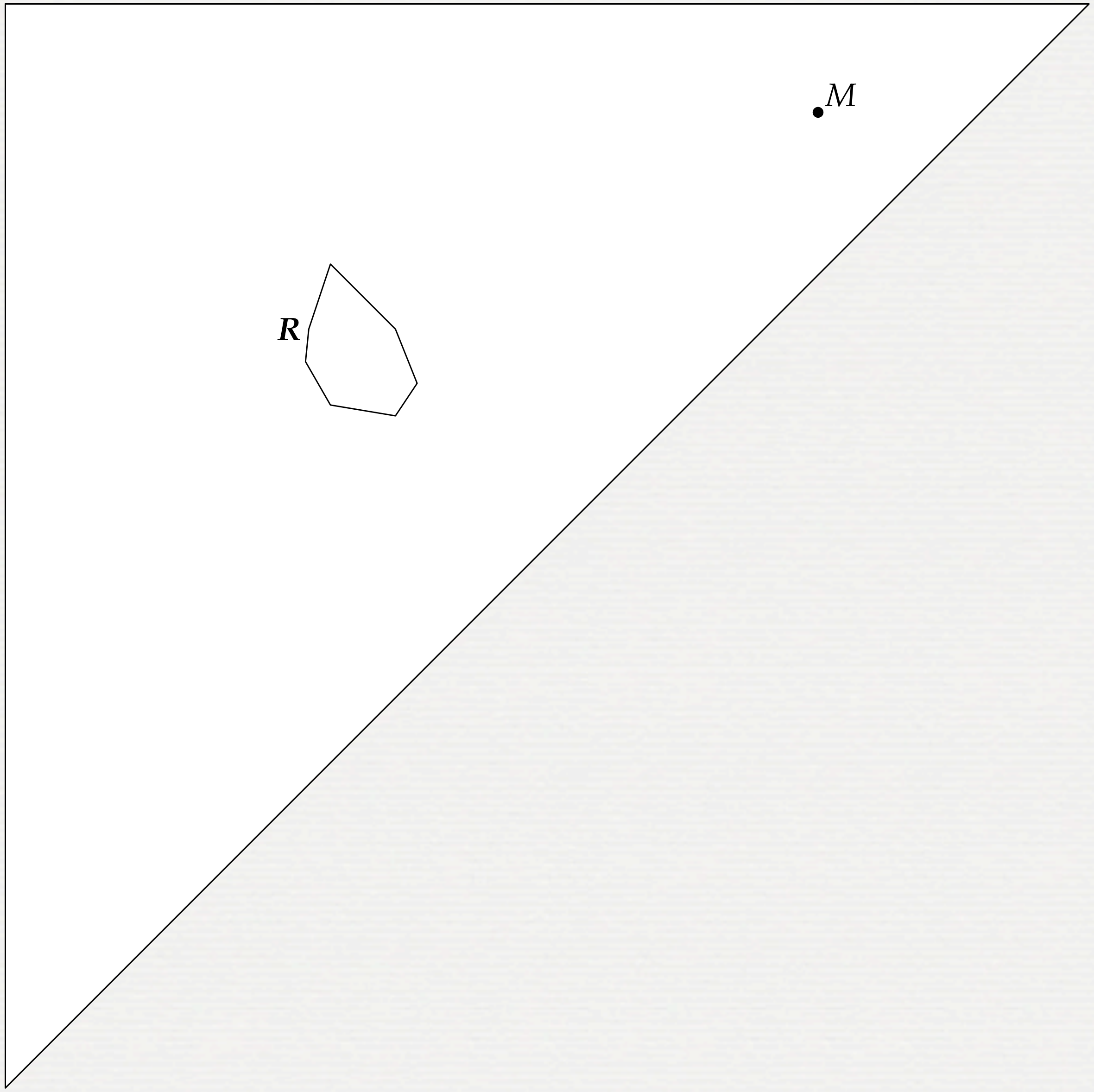
HIGH-LEVEL IDEA

Probability space



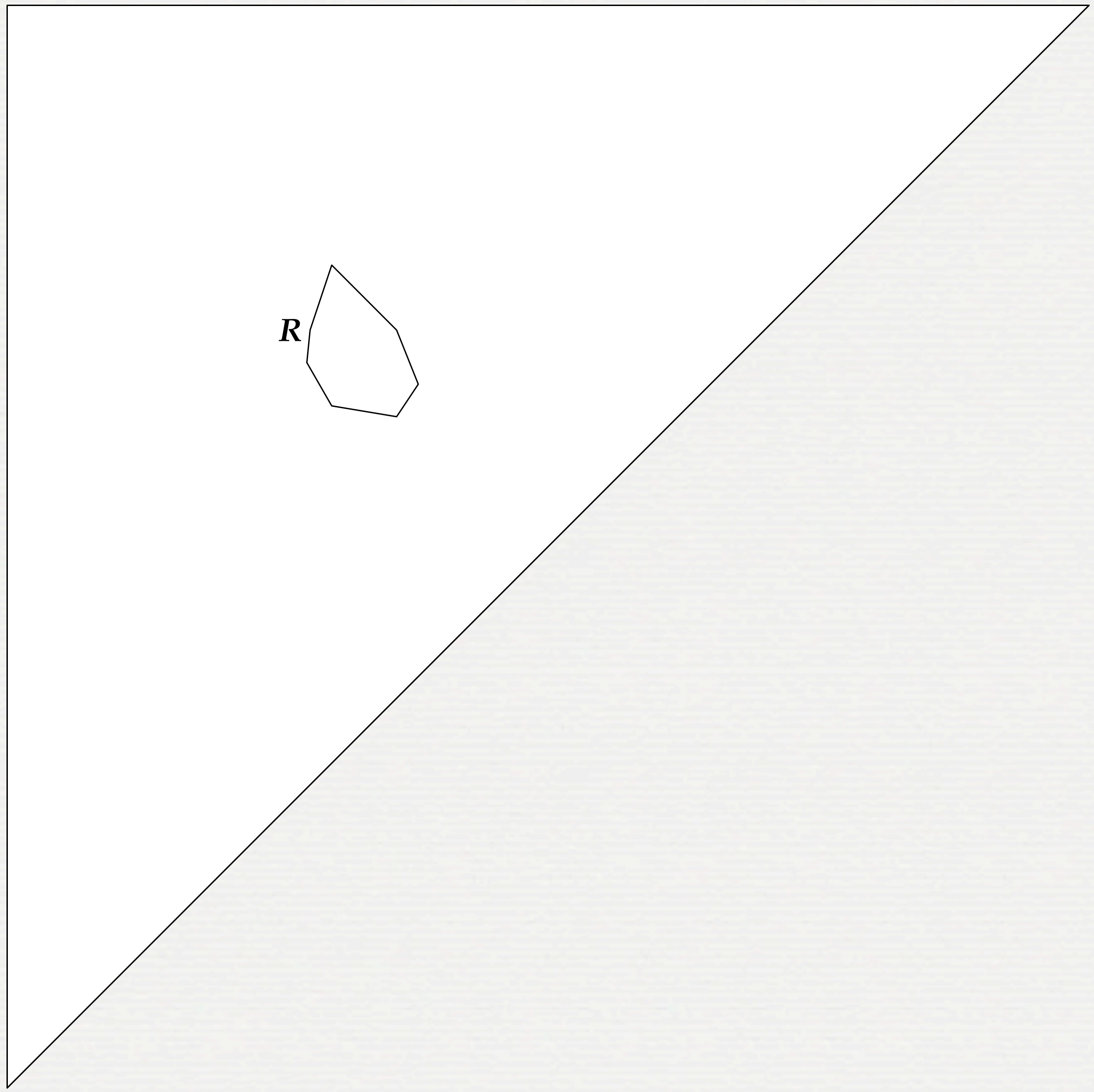
HIGH-LEVEL IDEA

Probability space



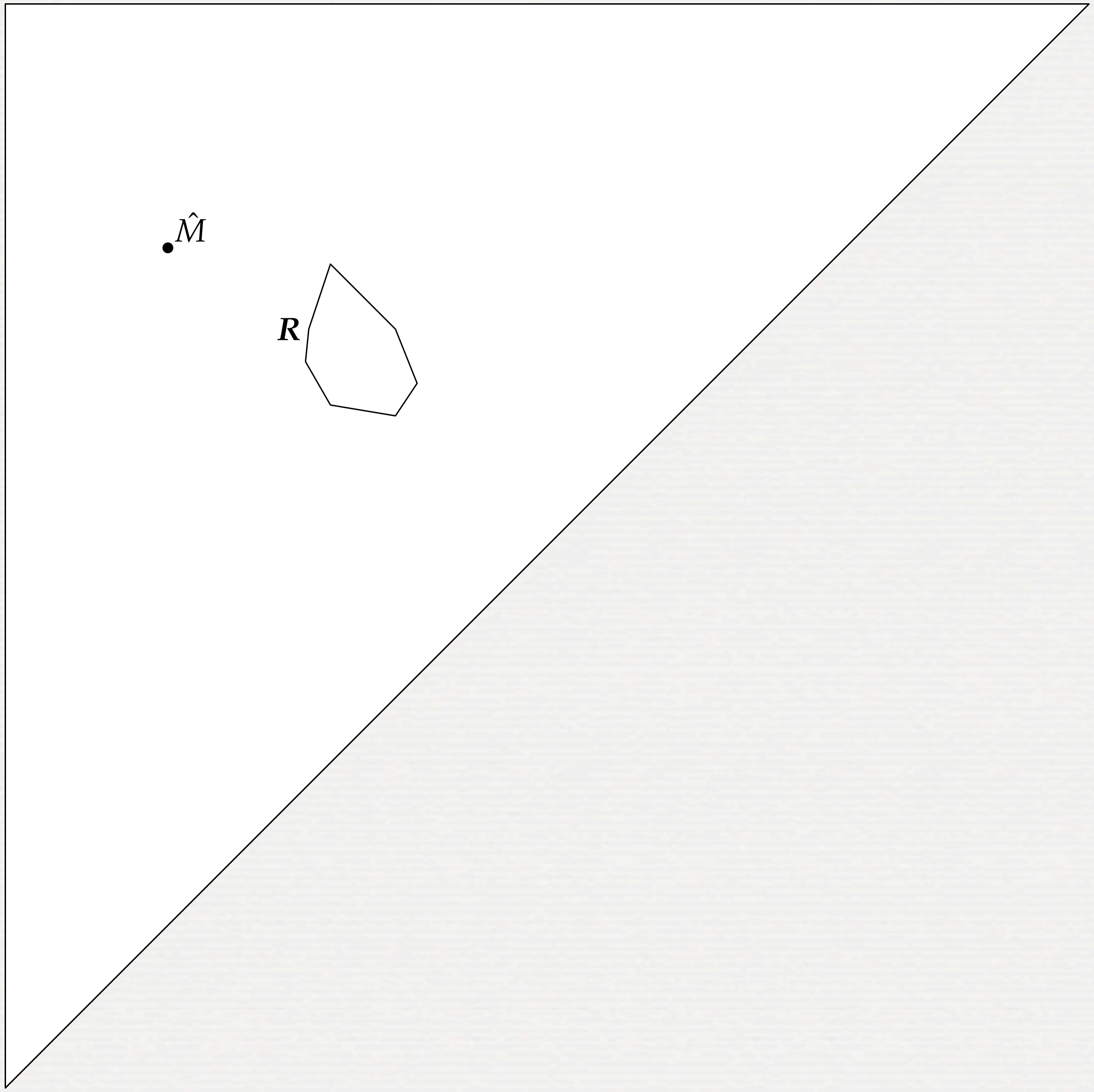
HIGH-LEVEL IDEA

Probability space



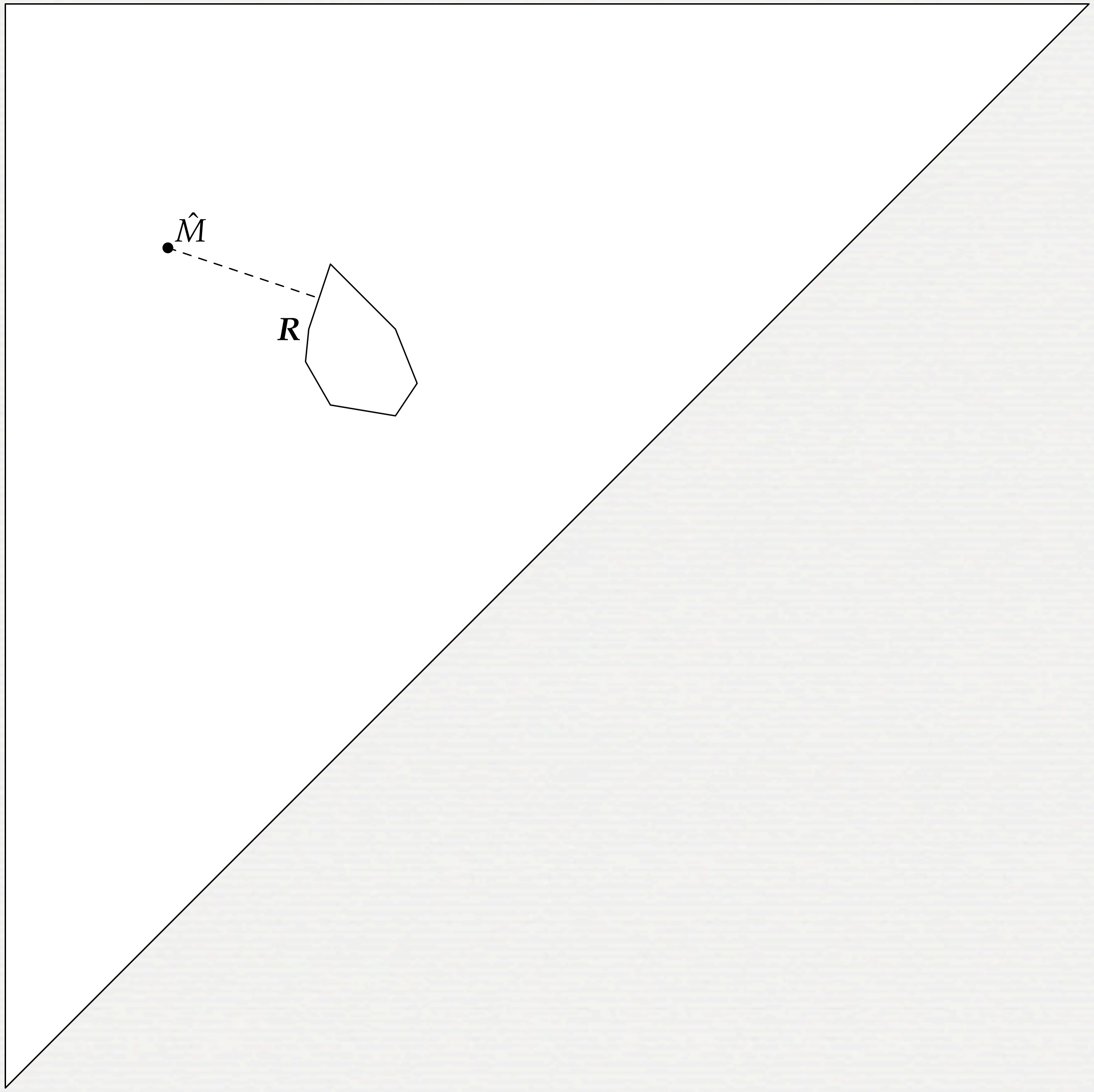
HIGH-LEVEL IDEA

Probability space



HIGH-LEVEL IDEA

Probability space



KULLBACK-LEIBLER DIVERGENCE

- Measure of discrepancy between distributions $\hat{M}$ and R

$$D(\hat{M} \parallel R) = \sum_{z \in \boldsymbol{X} \times \boldsymbol{X}} \hat{M}(z) \log \frac{\hat{M}(z)}{R(z)}$$

THE ALGORITHM (FOR ONE ROUND)

- Count K ballots and form approximation $\hat{M}$
- Certify if $\min_{R \in \boldsymbol{R}} D(\hat{M} \parallel R) > \frac{1}{K} \log \frac{f(\hat{M})}{\xi}$
- Minimize using convex optimization

BOUNDING THE RISK

- For any M,

$$\mathbb{P}\left(\begin{array}{c}\text{Drawing } K \text{ ballots from} \\ M \text{ with distribution } \hat{M}\end{array}\right) = f(\hat{M})e^{-K \cdot D(\hat{M} \parallel M)}$$

- Risk-limiting

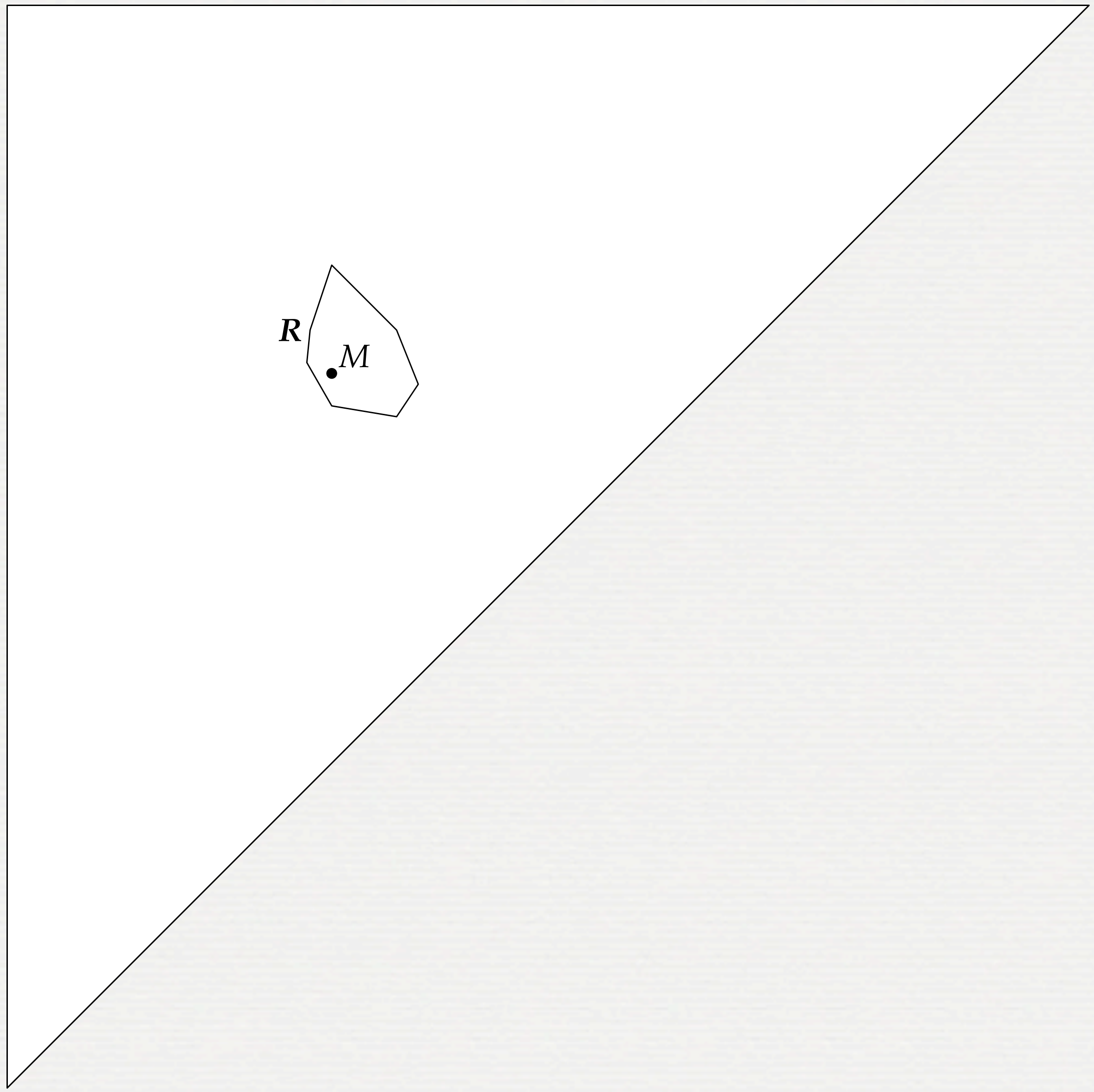
$$\mathbb{P}(\text{certify} \mid M \in \boldsymbol{R}) = \sum_{\substack{\hat{M} \text{ that} \\ \text{are certified}}} f(\hat{M})e^{-K \cdot D(\hat{M} \parallel M)} < \alpha$$

PICKING ξ

- Pick ξ to satisfy risk-level α
- What do we know about $f(\hat{M})e^{-K \cdot D(\hat{M} \;\|\; M)}$
  - For each $\hat{M}$ that is certified, $f(\hat{M})e^{-K \cdot D(\hat{M} \;\|\; M)} < \xi$
  - For most $\hat{M}$, it is significantly smaller

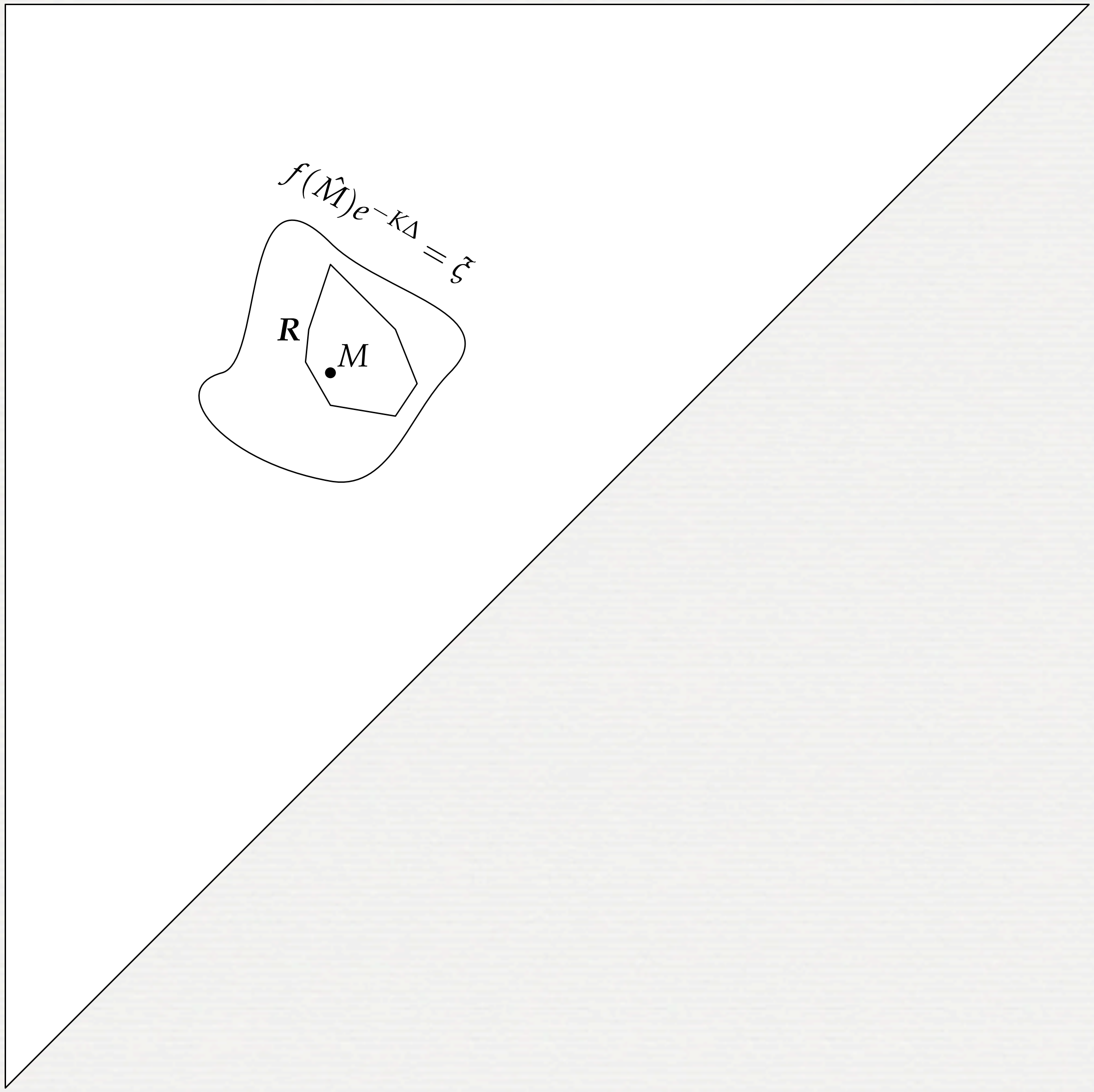
PICKING ξ PICTORIALLY

Probability space



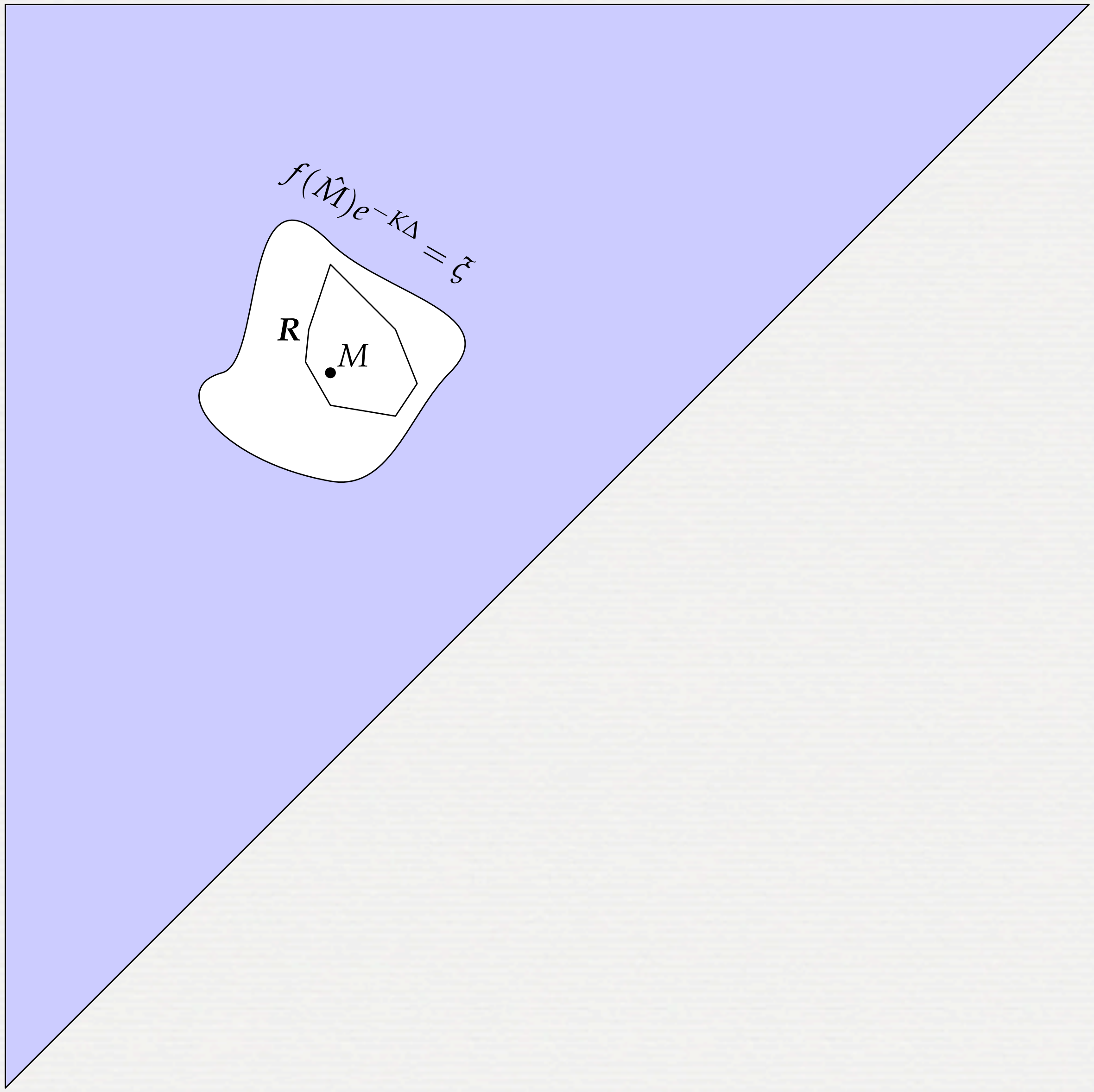
PICKING ξ PICTORIALLY

Probability space



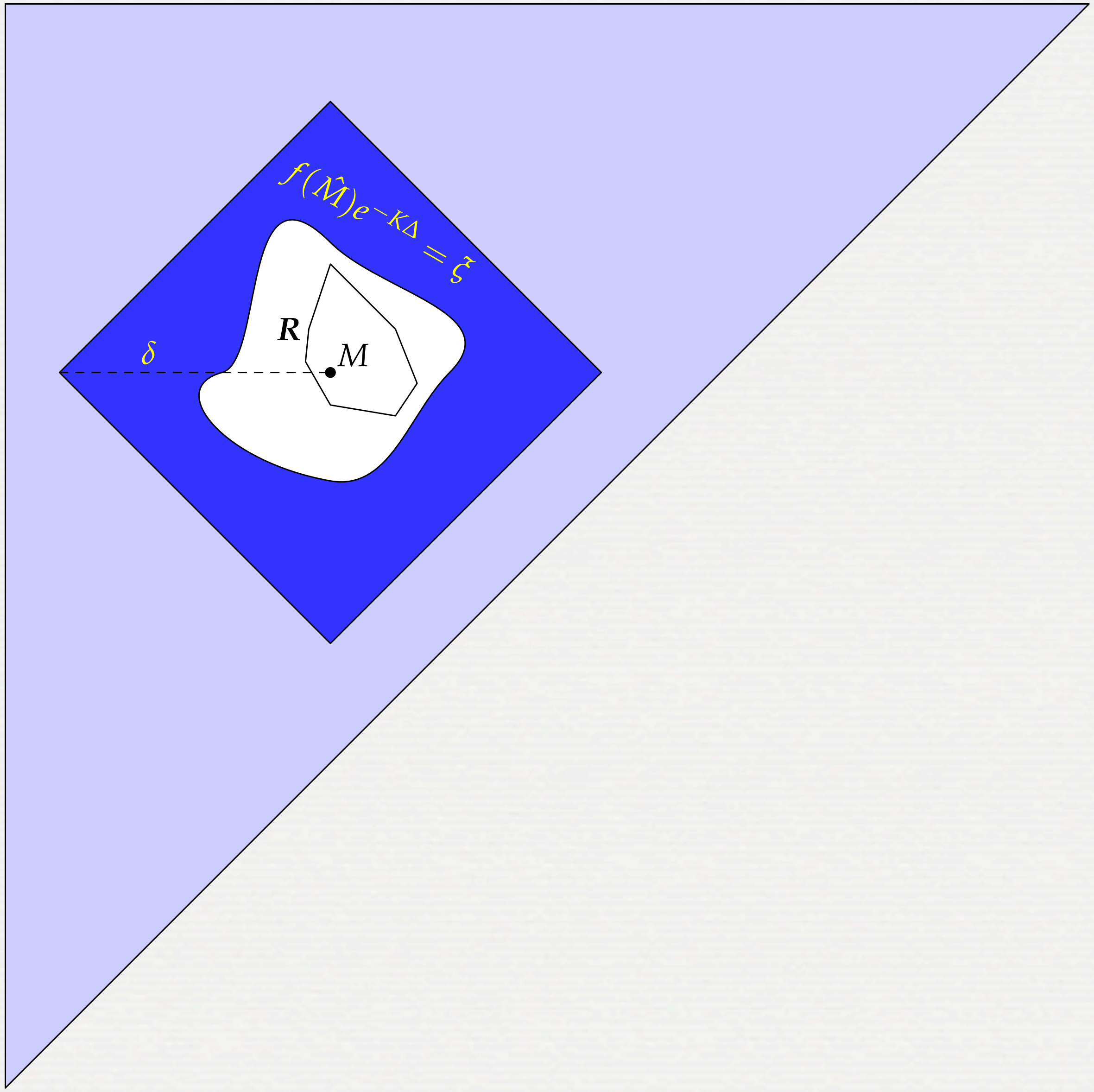
PICKING ξ PICTORIALLY

Probability space



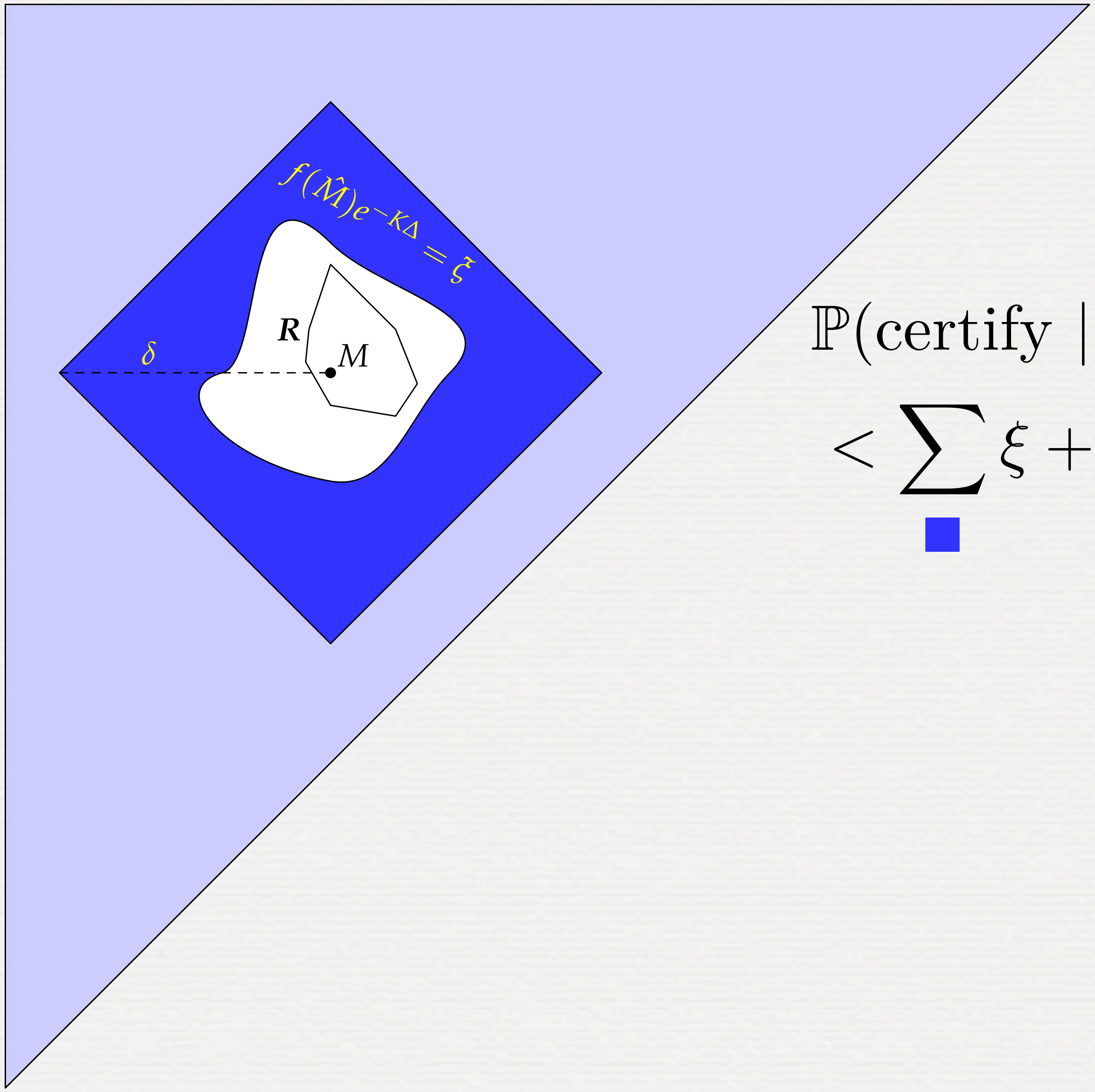
PICKING ξ PICTORIALLY

Probability space



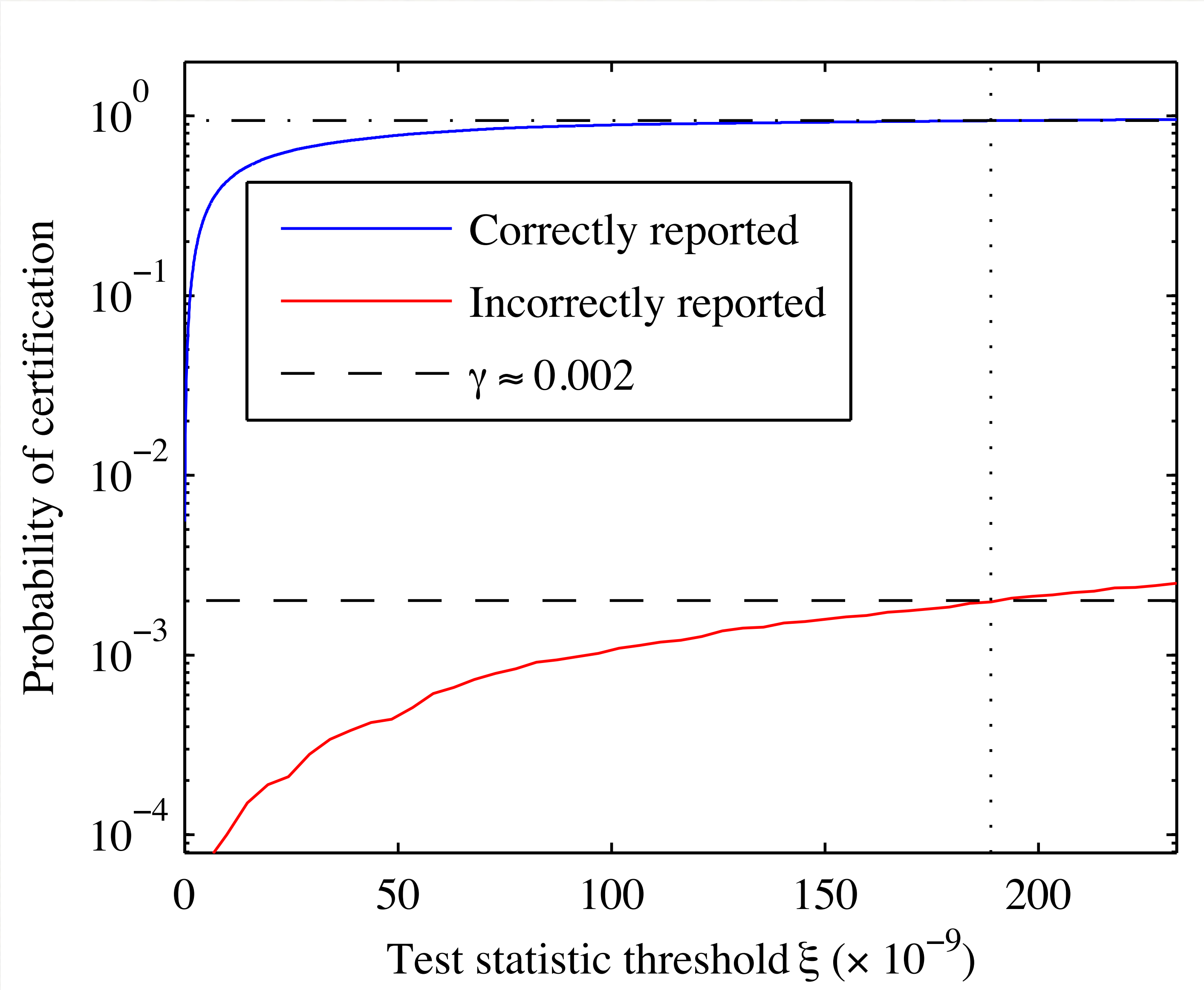
PICKING ξ PICTORIALLY

Probability space



$$\mathbb{P}(\text{certify} \mid M \in R) < \sum \xi + \sum p(\delta) = \alpha$$

CAN ONE DO BETTER?



CONCLUSIONS

- New way to view ballots and their selection
- Simple auditing algorithm that doesn't throw away any information
- More clever analysis = more powerful algorithm!