IEIIT-CNR

# On the Sample Complexity of Probabilistic Analysis and Design Methods


Roberto Tempo

IEIIT-CNR

Politecnico di Torino

tempo@polito.it

IEIIT-CNR


bined 48th IEEE Conference on Decision and Contro
and 28th Chinese Control Conference
Shanghai Internation
Decem
Sha

IEIIT-CNR

# Buon Compleanno Yutaka!


To Yamamoto Sensei on the occasion of his 60th birthday!

IEIIT-CNR

# Randomized Algorithms

IEIIT-CNR

# Randomized Algorithm

- **Randomized Algorithm:** An algorithm that makes random choices during its execution to produce a result

IEIIT-CNR

# Randomized Algorithm

- **Randomized Algorithm:** An algorithm that makes random choices during its execution to produce a result
- Example of a *random choice* is a coin toss

*heads* or *tails*


IEIIT-CNR

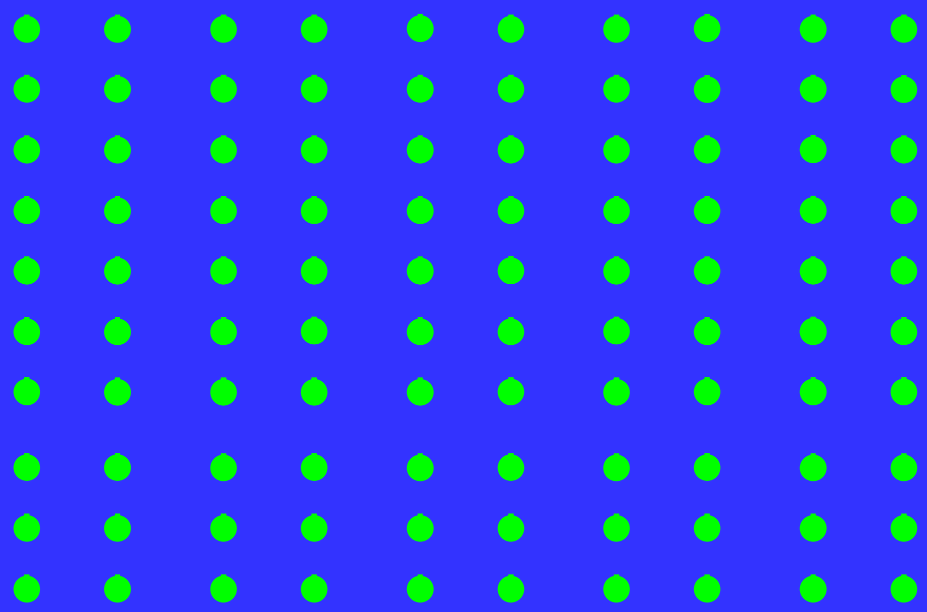
# Las Vegas Randomized Algorithm

- **Las Vegas Randomized Algorithm:** A randomized algorithm that always (i.e. w.p.1) produces correct results, the only variation from one run to another is the running time

IEIIT-CNR

# Example of Las Vegas Randomization

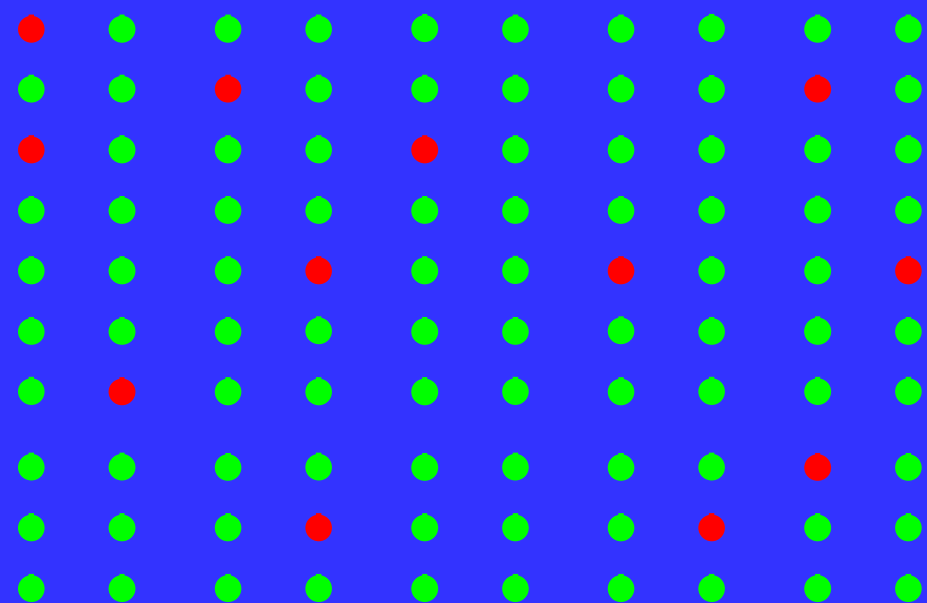
*we may explore the entire discrete sample space*



IEIIT-CNR

# Example of Las Vegas Randomization

*we may explore the entire discrete sample space*



IEIIT-CNR

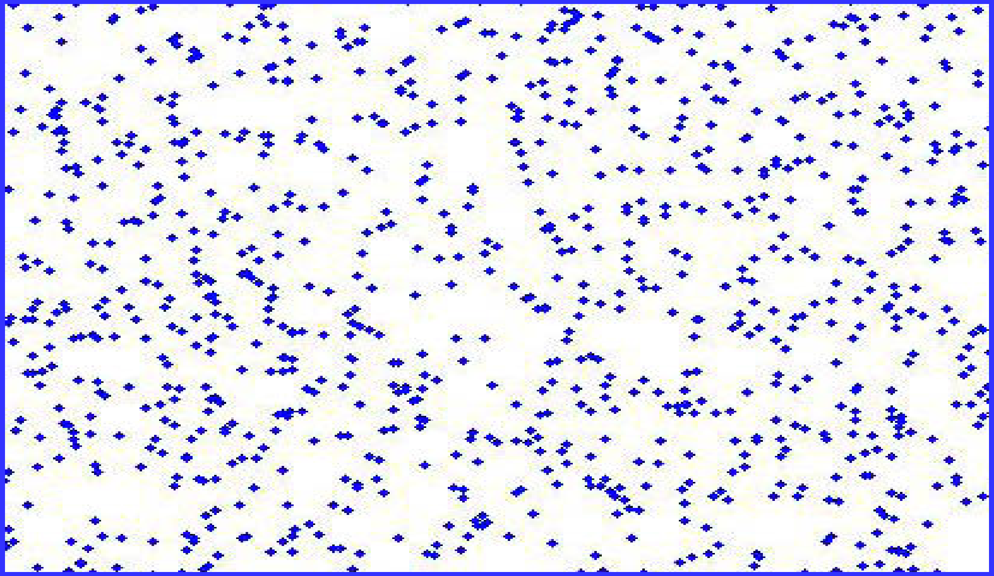
# Monte Carlo Randomized Algorithm

- **Monte Carlo Randomized Algorithm:** A randomized algorithm that may produce incorrect results, but with bounded error probability

IEIIT-CNR

# Example of Monte Carlo Randomization

*non-zero probability of error (which can be bounded) due to finite sample size*



IEIIT-CNR

# This Talk

- Address good news and bad news
  - bad news: we may produce incorrect results
  - good news: the probability of error is bounded

- Tradeoff between good and bad news is the sample complexity

IEIIT-CNR

# Estimating the Probability of Violation

# Notation

- Uncertainty $\Delta$ (random matrix) is bounded in a set $\mathcal{B}$
- Controller (design) parameters $\theta$ varying in a set $\Theta$
- $g: \Theta \times \mathcal{B} \rightarrow \{0,1\}$ is a binary violation function
- $N$ is the sample complexity

IEIIT-CNR

# Binary Violation Function

- Binary violation function

  $g: \Theta \times \mathcal{B} \rightarrow \{0,1\}$

  is defined as follows

  $g(\theta, \Delta) = 0$ *if* $\theta$ *meets control specs for* $\Delta$

  $g(\theta, \Delta) = 1$ *otherwise*

IEIIT-CNR

# Binary Violation Function

- Binary violation function

  $g: \Theta \times \mathcal{B} \rightarrow \{0,1\}$

  is defined as follows

  $g(\theta, \Delta) = 0$ *if* $\theta$ *meets control specs for* $\Delta$

  $g(\theta, \Delta) = 1$ *otherwise*

- Examples of control specs

  - closed loop poles into the open left half plane

  - $\mathcal{H}_\infty$ bound on the sensitivity function

# Probability of Violation

- Given $\theta \in \Theta$, the probability of violation for the function $g(\theta, \Delta)$ is defined as

$$E_g(\theta) = \mathrm{Prob}\{\Delta \in \mathcal{B}\colon g(\theta, \Delta) = 1\}$$

# Empirical Mean of Violation

- Draw $N$ i.i.d. samples (multisample)

$$\Delta^{1,\ldots,N} = \{\Delta^1, \Delta^2, \ldots, \Delta^N\} \in \mathcal{B}$$

according to a given probability measure

- Given $\theta \in \Theta$, the empirical mean of $g(\theta, \Delta)$ is defined as

$$\hat{E}_N(\theta) = \frac{1}{N}\sum_{i=1}^{N} g(\theta, \Delta^i)$$

- Since $g$ is a binary function $\hat{E}_N(\theta) \in [0,1]$

# Accuracy and Confidence

- We introduce two probabilistic parameters

  - accuracy $\varepsilon \in (0,1)$

  - confidence $\delta \in (0,1)$

IEIIT-CNR

# Estimating the Probability of Violation

- For fixed $\theta \in \Theta$, given accuracy $\varepsilon \in (0,1)$, from the Hoeffding inequality we obtain

$$\mathrm{Prob}\{\Delta^{1,\ldots,N} : | E_g(\theta) - \hat{E}_N(\theta) | \geq \varepsilon\} \leq 2e^{(-2N\varepsilon^2)}$$

where e denotes the Euler number

- To guarantee confidence $\delta \in (0,1)$, we set

$$2e^{(-2N\varepsilon^2)} \leq \delta$$

- We obtain immediately the (additive) Chernoff bound

$$N \geq \frac{1}{2\varepsilon^2} \log \frac{2}{\delta}$$

IEIIT-CNR

# Sample Complexity and (additive) Chernoff Bound

- Hoeffding inequality provides a bound on the tail distribution

$$2e^{(-2N\varepsilon^2)}$$

- Computationally, evaluating the minimum $N$ such that

$$2e^{(-2N\varepsilon^2)} \leq \delta$$

  is immediate (one-parameter problem given $\varepsilon$ and $\delta$)

- Chernoff bound provides a fundamental relation (sample complexity) $N = N(\varepsilon, \delta)$ where $1/\varepsilon$ is *quadratic* and $1/\delta$ logarithmic

# Estimating the Probability of Failure

# Randomized Feasibility Problem

- Ideally, we would like to find θ (if it exists) such that

$$g(\theta, \Delta) = 0 \text{ for all } \Delta \in \mathcal{B}$$

- This problem is too hard (semi-infinite feasibility problem)
- We study its randomized version

$$g(\theta, \Delta^i) = 0 \text{ for all } i = 1, 2, \ldots, N$$

# Probabilistic Relaxation

- We allow (at most) $m$ violations of the $N$ constraints

$$g(\theta, \Delta^i) = 0 \text{ for all } i = 1, 2, \ldots, N$$

- That is

$$\sum_{i=1}^{N} g(\theta, \Delta^i) \leq m$$

which can be written using the empirical violation

$$\hat{E}_N(\theta) \leq \frac{m}{N}$$

- **Motivation:** Dealing with data outliers

# Probability of Failure

- Given $N$, $m$, $\varepsilon \in (0,1)$ and $g$, the probability of failure is defined as

$$p_g(N,\varepsilon,m) =$$

$$\mathrm{Prob}\left\{\Delta^{1,\dots,N} : \text{there exists } \theta \in \Theta \text{ such that } \hat{E}_N(\theta) \le \frac{m}{N},\ E_g(\theta) > \varepsilon\right\}$$

IEIIT-CNR

# Sample Complexity for Finite Families

# Finite Families of Controllers

- Suppose that the controller set has finite cardinality $n_C$

$$\Theta = \{\theta^1, \theta^2, \ldots, \theta^{n_C}\}$$

- The probability of violation may be computed by a repeated application of the Hoeffding inequality
- The obtained sample complexity is quadratic in $1/\varepsilon$

IEIIT-CNR

# Probability of Failure and Binomial Distribution

- The binomial distribution provides a bound on the probability of failure

$$p_g(N, \varepsilon, m) \leq n_C \, \mathrm{B}(N, \varepsilon, m)$$

where

$$\mathrm{B}(N, \varepsilon, m) = \sum_{i=0}^{m} \binom{N}{i} \varepsilon^{i} \left(1-\varepsilon\right)^{N-i}$$

- The constant is the cardinality of the controller family

IEIIT-CNR

# Bounding the Probability of Failure and Sample Complexity

- **Theorem**[1]: Suppose that card(Θ) ≤ $n_C$ . Given ε, δ ∈(0,1) and $m \geq 0$, if

$$N \geq \inf_{a>1} \frac{1}{\varepsilon}\left(\frac{a}{a-1}\right)\left(\log\frac{n_C}{\delta} + m\log(a)\right)$$

then

$$p_g(N, \varepsilon, m) \leq n_C \, \mathrm{B}(N, \varepsilon, m) \leq \delta$$

- **Remark**: This result provides the sample complexity so that the probability of failure is bounded by δ

[1] T. Alamo, R. Tempo and A. Luque (2010)

# Bounding the Probability of Failure and Sample Complexity

IEIIT-CNR

- Suboptimal value of *a* is the Euler number e
- Sample complexity is given by

$$N \geq \frac{1}{\varepsilon}\left(\frac{\mathrm{e}}{\mathrm{e}-1}\right)\left(\log\frac{n_C}{\delta}+m\right)$$

- Sample complexity is *linear* in
  - $1/\varepsilon$ (not quadratic!)
  - $m$
  - $\log\frac{n_C}{\delta}$

# Optimization of Convex Problems

# Semi-Infinite Convex Optimization[1]

- **Semi-Infinite Convex Optimization:** Find the optimal solution of the problem

$$\min_{\theta \in \Theta} \; c^{T}\theta \quad \text{subject to } f(\theta,\Delta) \le 0 \text{ for all } \Delta \in \mathcal{B}$$

where $f(\theta,\Delta)$ is convex in $\theta$ for all $\Delta \in \mathcal{B}$

- Study the randomized optimization problem

$$\min_{\theta \in \Theta} \; c^{T}\theta \quad \text{subject to } f(\theta,\Delta^{i}) \le 0 \text{ for all } i = 1, 2, \ldots, N$$

[1] G. Calafiore and M.C. Campi (2006), M.C. Campi and S. Garatti (2008)

IEIIT-CNR

# Sample Complexity for Convex Optimization[1]

$$N \geq \inf_{a>1} \frac{1}{\varepsilon}\left(\frac{a}{a-1}\right)\left(\log\frac{1}{\delta} + (n-1)\log(a)\right)$$

$$N \geq \frac{1}{\varepsilon}\left(\frac{\mathrm{e}}{\mathrm{e}-1}\right)\left(\log\frac{1}{\delta} + (n-1)\right)$$

where *n* is the number of controller parameters

[1] T. Alamo, R. Tempo and A. Luque (2010)

# Statistical Learning Theory

- Statistical learning theory comes to rescue for
  - non-convex optimization problems
  - feasibility for infinite families of controllers

[1] T. Alamo, R. Tempo and E.F. Camacho (2009)

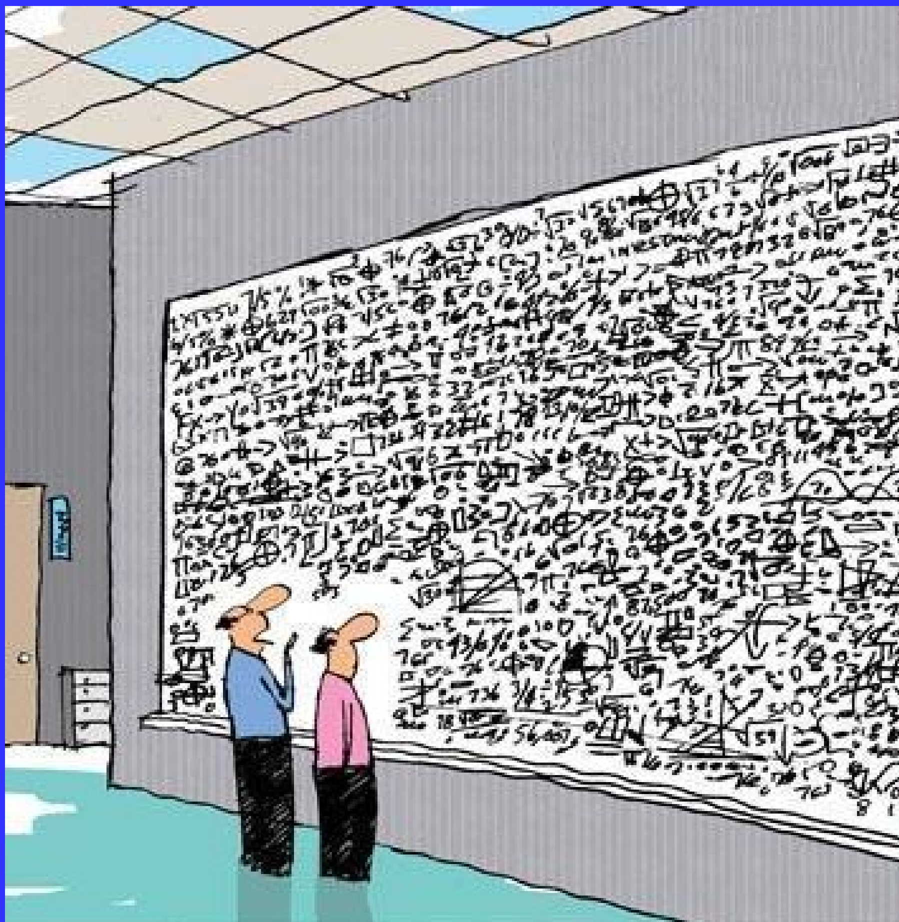
IEIIT-CNR

# Acknowledgment

- The results presented here are joint work with Teodoro Alamo and Amalia Luque

IEIIT-CNR

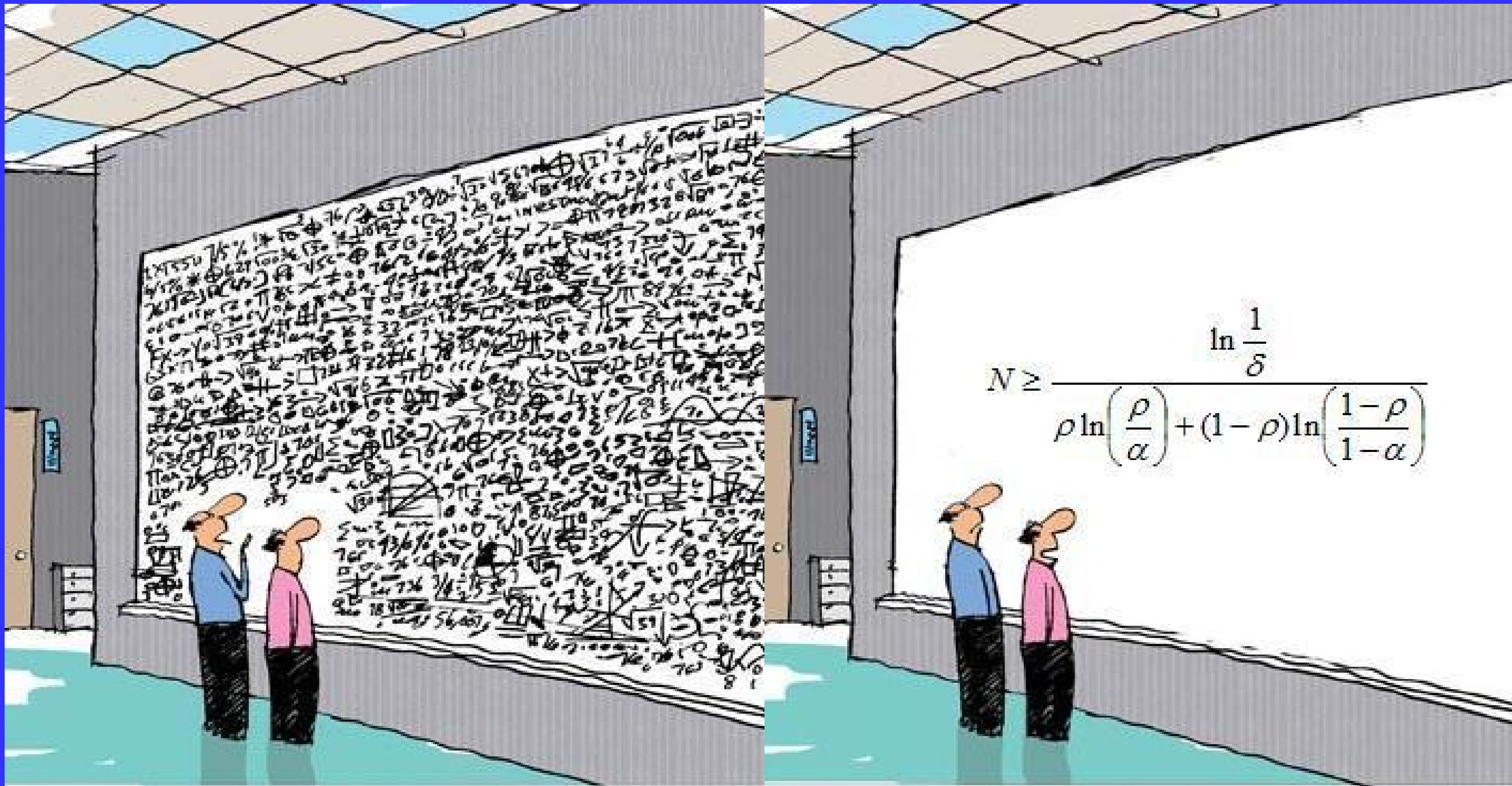
# Conclusion



Other approaches

IEIIT-CNR

# Conclusion




Other approaches | Randomized algorithms