A **DISTRIBUTION-AGNOSTIC** AND **CORRELATION-AWARE** ANALYSIS OF PERIODIC TASKS

European Research Council
Established by the European Commission

RTSS 2024
December 12

RTSS * Artifact * Consistent * Complete * Well Documented * Easy to Reuse * Evaluated

Filip Marković, Georg von der Brüggen, Mario Günzel,
Jian-Jia Chen, and Björn B. Brandenburg

MAX-PLANCK-GESELLSCHAFT
MAX PLANCK INSTITUTE **FOR SOFTWARE SYSTEMS**
technische universität dortmund

THE RISE OF PROBABILISTIC ANALYSIS

Many modern systems ***fail*** *to meet their timing requirements*

Expert Opinion on Tesla Model S with Autopilot, 2016

NHTSA Recall notice 1

NHTSA Recall notice 2

NHTSA Recall notice 3

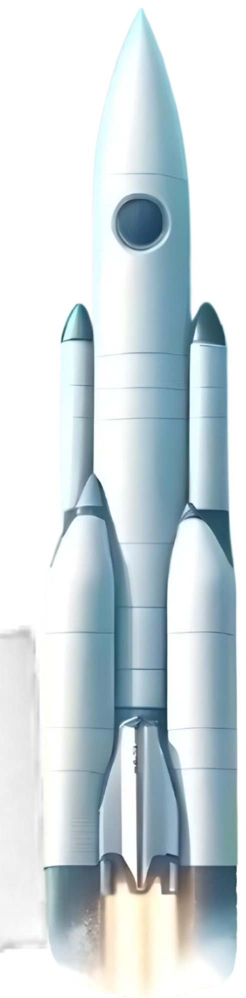
NHTSA Recall notice 4

VW/Audi / Nov. 2022

NHTSA
NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION

Surviving an In-Flight Anomaly: What Happened on Ingenuity's Sixth Flight,

NASA, Håvard Grip. 2021

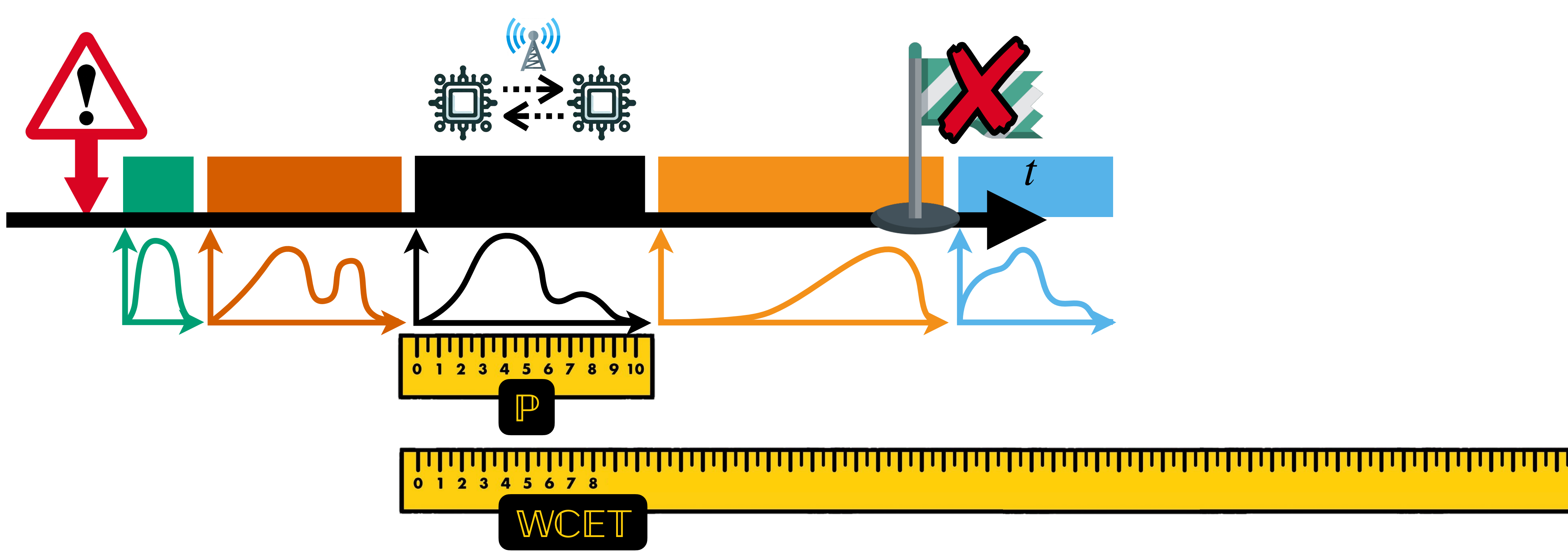


Many safety standards are defined in terms of ***failure probability***

SIL	Low demand mode prob. failure on demand	Continuous/High demand mode prob. failure per hour
1	$\geq 10^{-2}$ to $< 10^{-1}$	$\geq 10^{-6}$ to $< 10^{-5}$
2	$\geq 10^{-3}$ to $< 10^{-2}$	$\geq 10^{-7}$ to $< 10^{-6}$
3	$\geq 10^{-4}$ to $< 10^{-3}$	$\geq 10^{-8}$ to $< 10^{-7}$
4	$\geq 10^{-5}$ to $< 10^{-4}$	$\geq 10^{-9}$ to $< 10^{-8}$

Table 1: IEC 61508: Permitted Failure Probabilities [1]

Many modern systems are ***not statically*** *analyzable but rather* ***statistically***



[1] "The safe and effective application of probabilistic techniques in safety-critical systems", Agrawal et al. *ICCAD* (2020)

DEPENDENCE – A MAJOR OBSTACLE

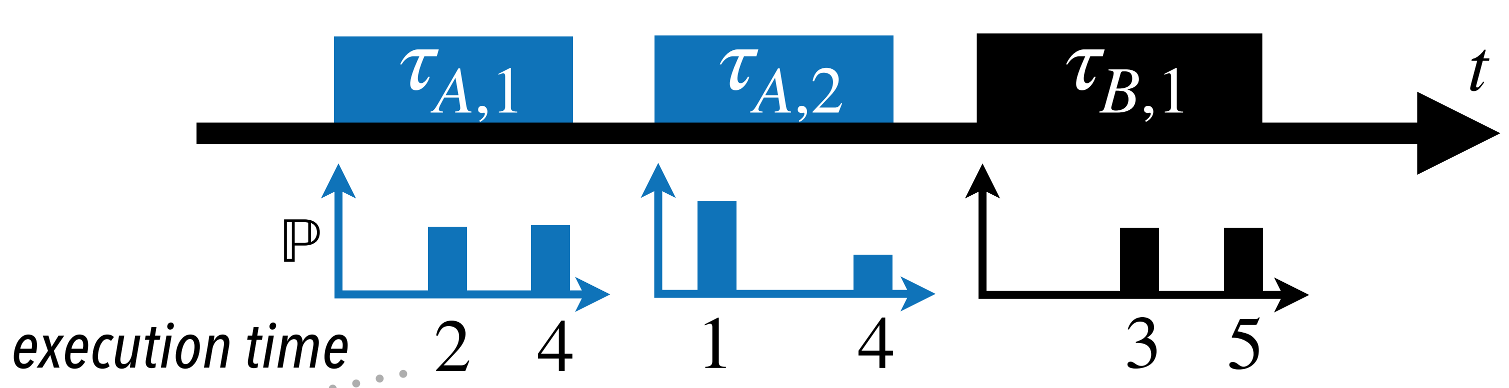
There are several ***open problems*** *in the field of probabilistic analysis*

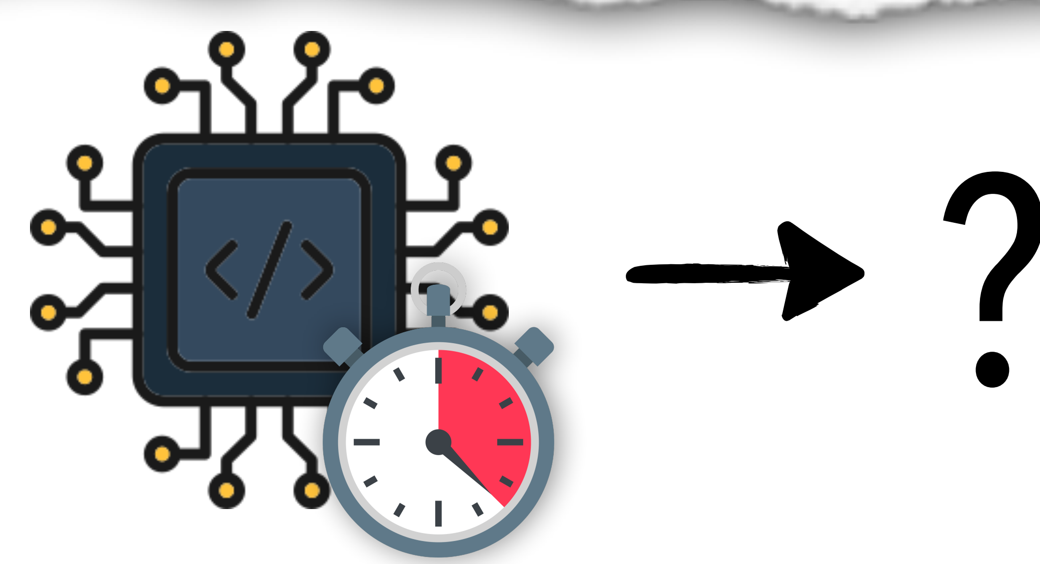
HOW TO MODEL DEPENDENCE?

"How to handle … dependences between the execution times of
(i) jobs of ***the same task****, and*
(ii) jobs ***of different tasks****?"* [2]

HOW TO STATISTICALLY INFER DEPENDENCE?

"Appropriate statistical studies are needed to investigate the ***types of dependences*** *and their impact on probabilistic schedulability analysis"* [2]



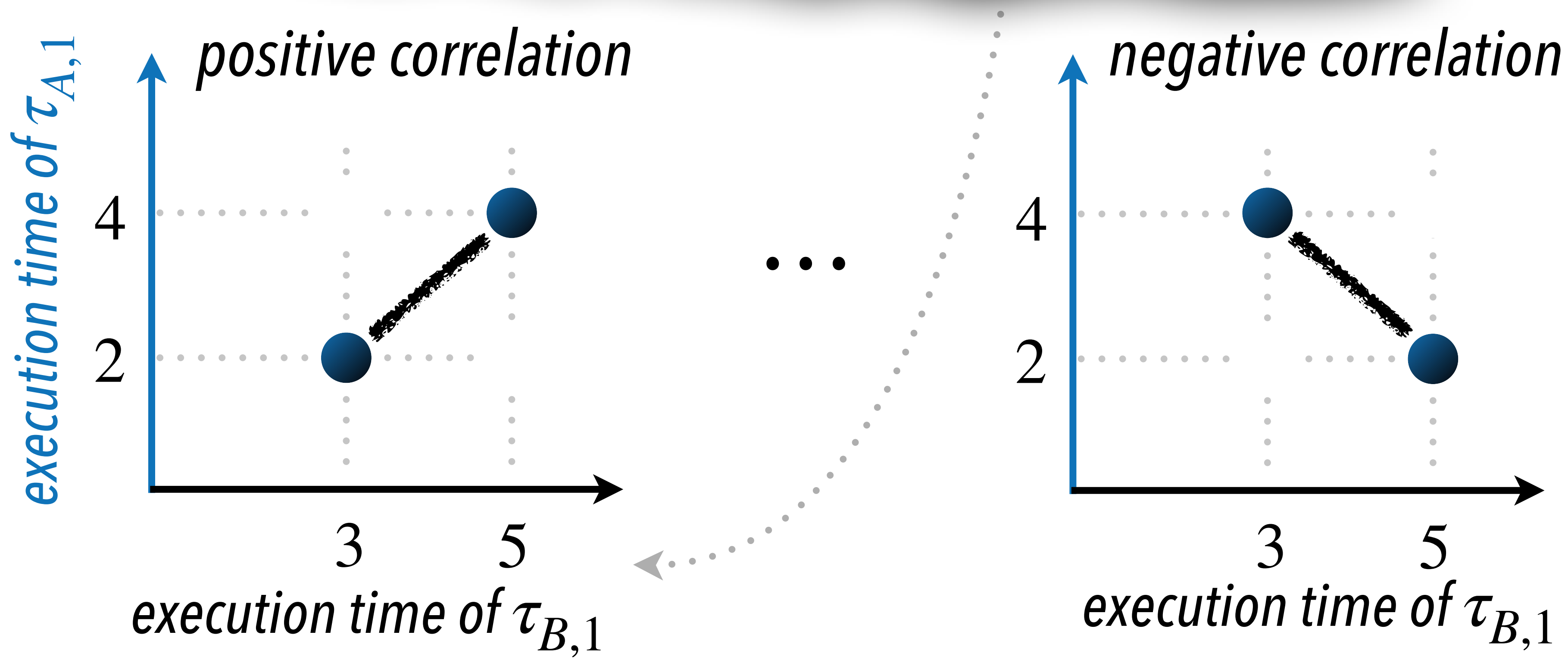


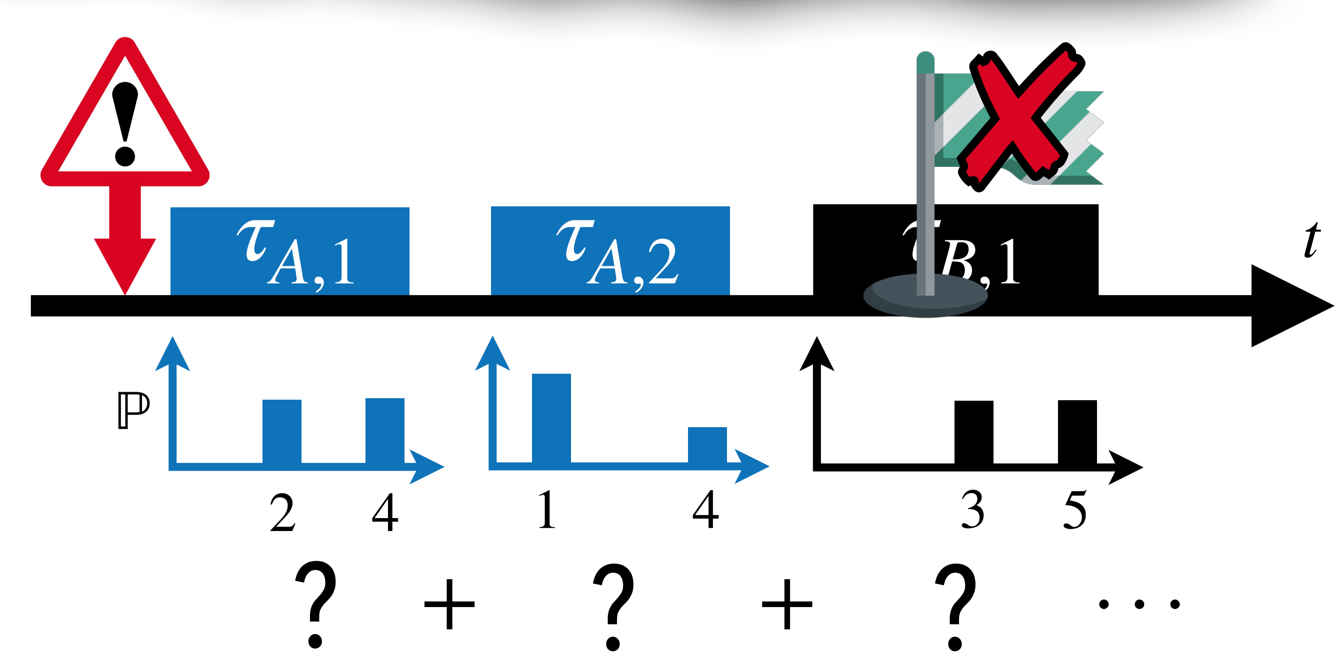
HOW TO QUANTIFY DEPENDENCE?

"The impact of these dependences may vary based on how strong they are." [2]

HOW TO USE IT IN RT ANALYSIS?

"Analyses are needed that can address dependencies." [2]





[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean. *LITES* (2019)

CORRELATION-TOLERANT ANALYSIS [3]

*Sound, irrespective of **any** potential dependence*

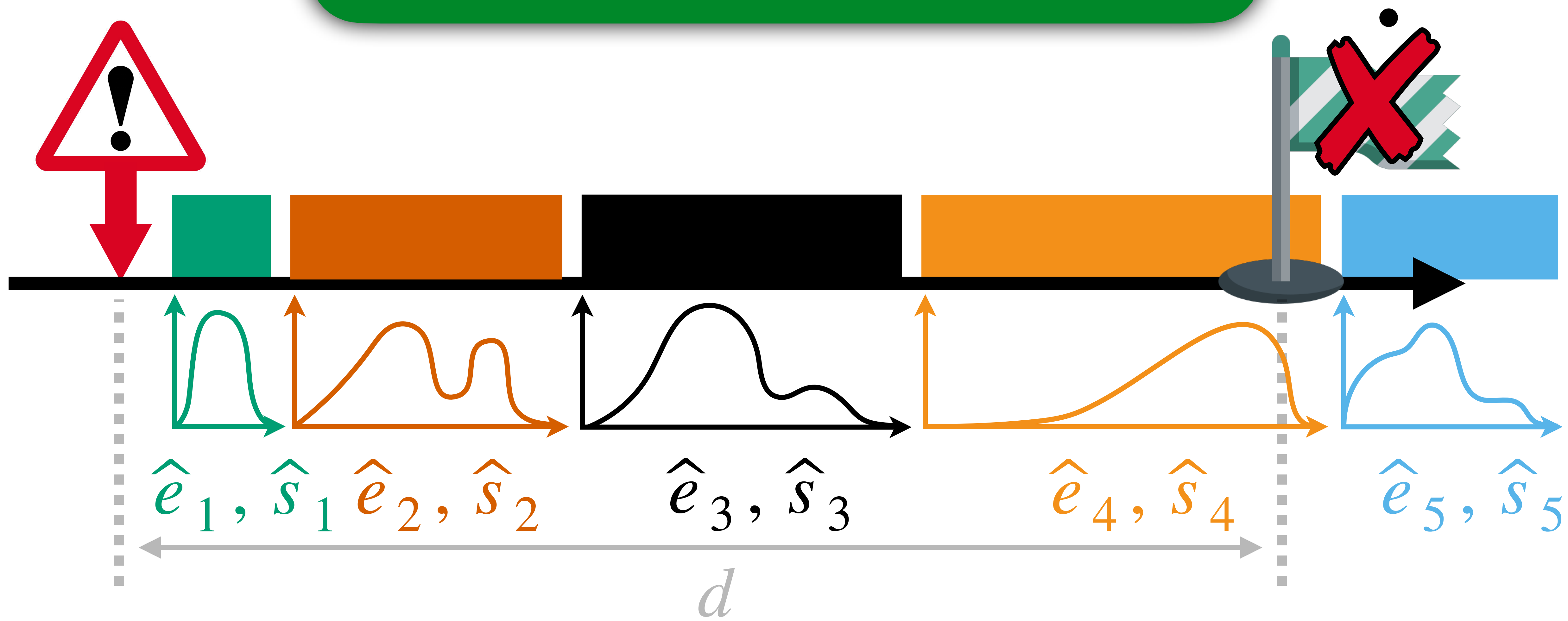
Inputs:

- $\hat{e}_i$: *upper bound on the **mean** execution time (ET) of any job of task* τ_i
- $\hat{s}_i$: *upper bound on the **standard deviation** of the ET of any job of* τ_i

$$\mathbb{P}[\sum_{i=1}^{5} X_i \geq d] \leq \frac{\left(\sum_{i=1}^{5} \hat{s}_i\right)^2}{\left(\sum_{i=1}^{5} \hat{s}_i\right)^2 + \left(d - \sum_{i=1}^{5} \hat{e}_i\right)^2}$$

deadline-failure probability of τ_5

deadline-failure of τ_5



[3] "CTA: A Correlation-Tolerant Analysis of the Deadline-Failure Probability of Dependent Tasks" Marković et al. *RTSS* (2023)

DEPENDENCE – A MAJOR OBSTACLE

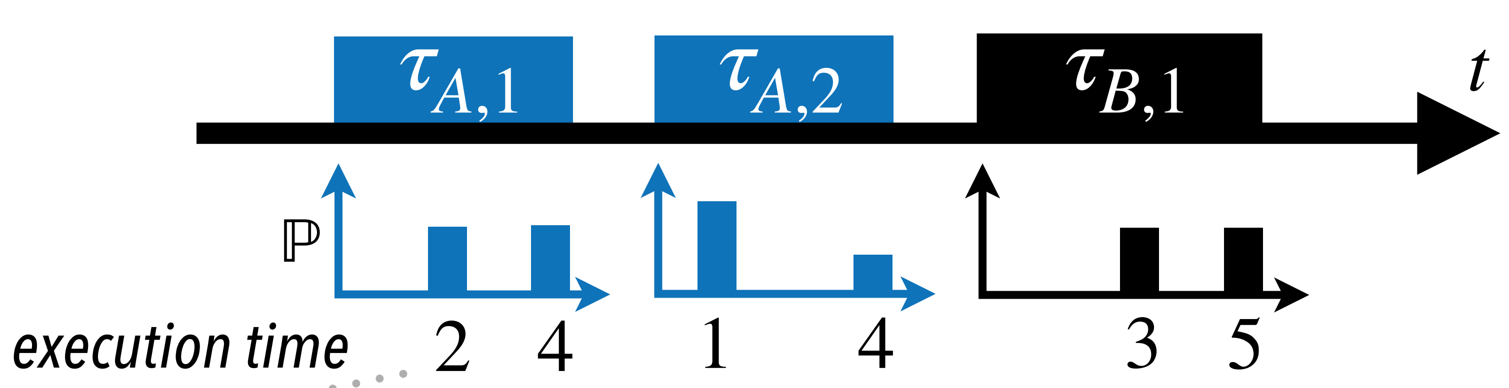
However, these open problems remain ***unaddressed***

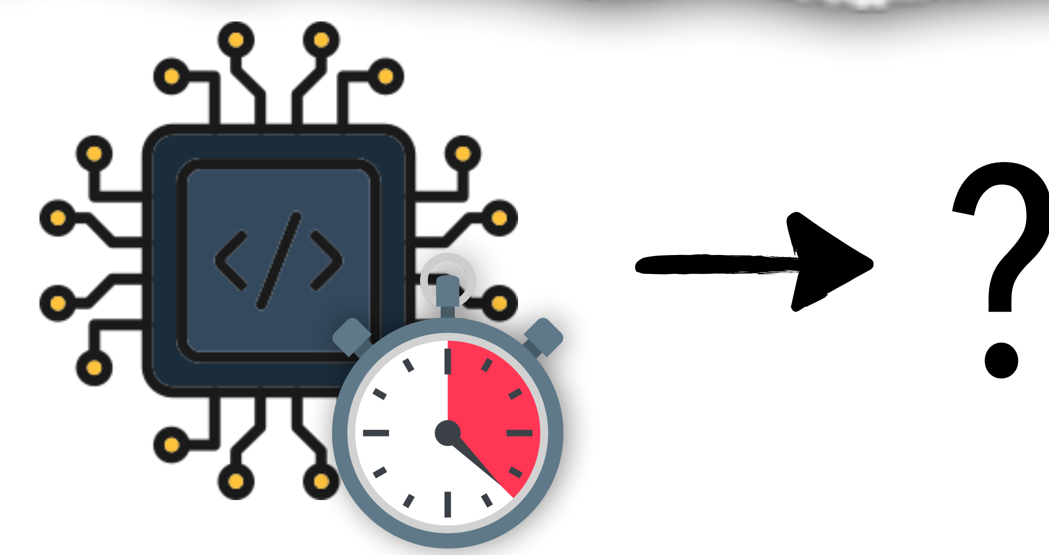
HOW TO MODEL DEPENDENCE?

*"How to handle … dependences between the execution times of (i) jobs of **the same task**, and (ii) jobs **of different tasks**?"* [2]

HOW TO STATISTICALLY INFER DEPENDENCE?

*"Appropriate statistical studies are needed to investigate the **types of dependences** and their impact on probabilistic schedulability analysis"* [2]



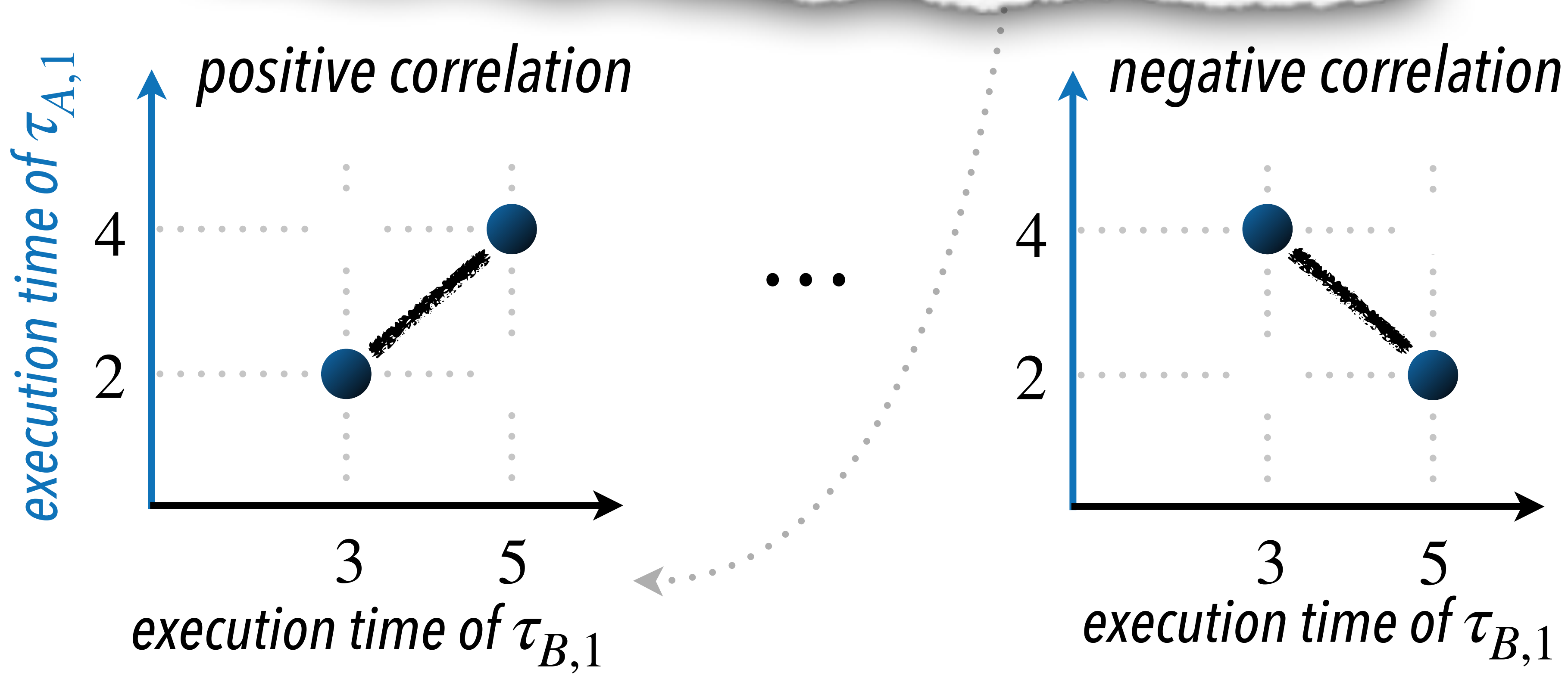


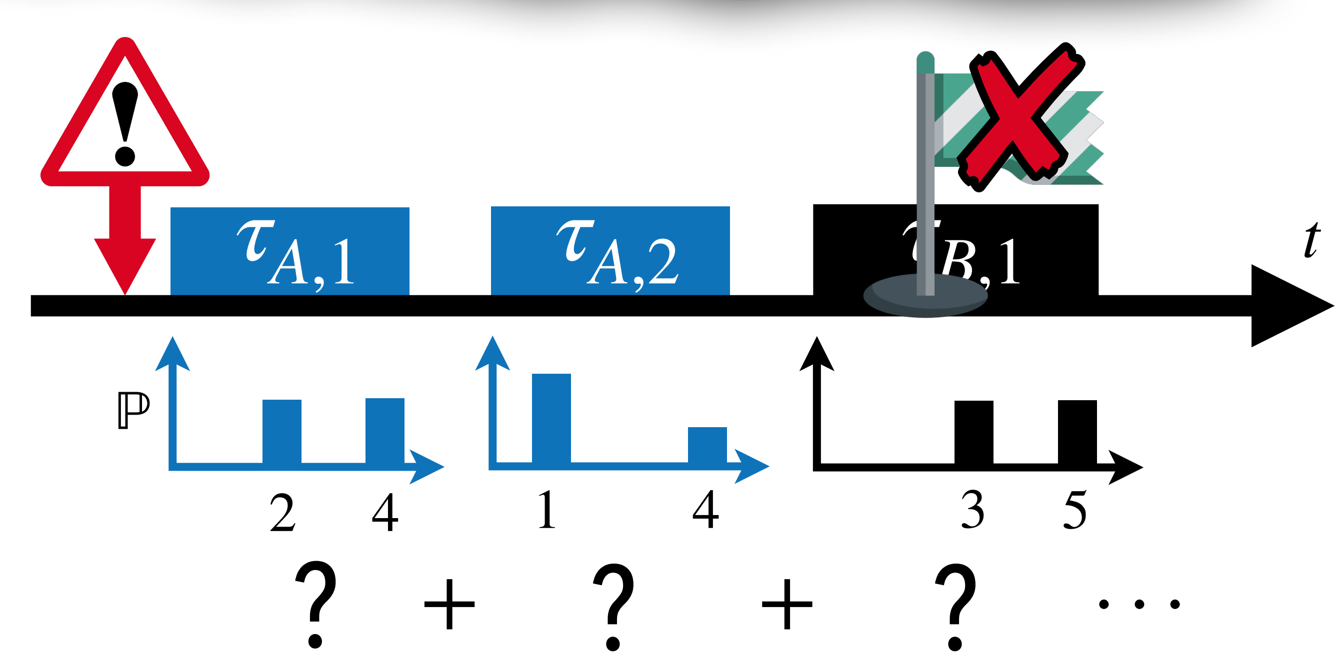
HOW TO QUANTIFY DEPENDENCE?

"The impact of these dependences may vary based on how strong they are." [2]

HOW TO USE IT IN RT ANALYSIS?

"Analyses are needed that can address dependencies." [2]





[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean. *LITES* (2019)

THIS PAPER (PART 1)
A CORRELATION-AWARE ANALYSIS

HOW TO MODEL DEPENDENCE?

*"How to handle … dependences between the execution times of (i) jobs of **the same task**, and (ii) jobs **of different tasks**?"* [2]

HOW TO QUANTIFY DEPENDENCE?

"The impact of these dependences may vary based on how strong they are." [2]

HOW TO USE IT IN RT ANALYSIS?

"Analyses are needed that can address dependencies." [2]

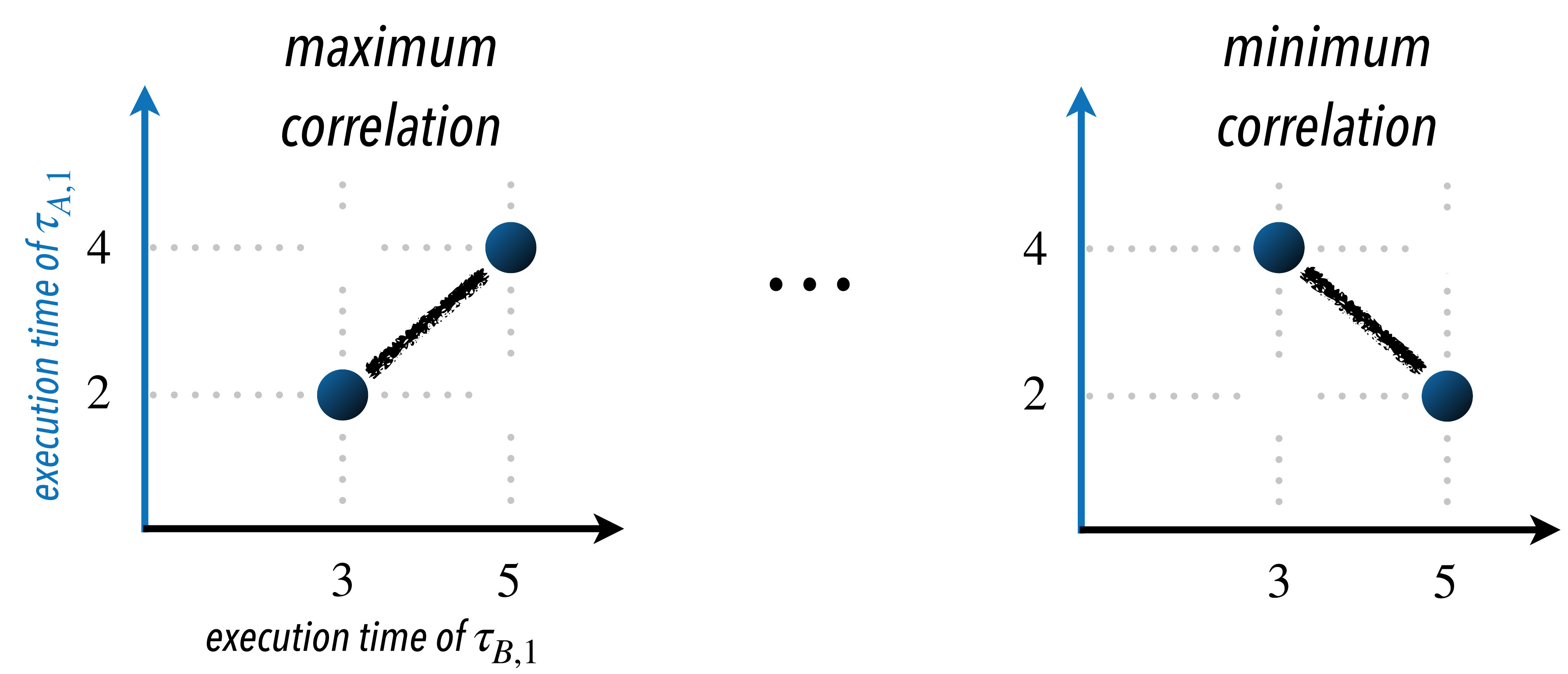
[2] "*A survey of probabilistic schedulability analysis techniques for real-time systems*" Davis and Cucu-Grosjean *LITES* (2019)

CORRELATION AWARENESS

How do we model various types of dependence?

HOW TO QUANTIFY DEPENDENCE?

"The impact of these dependences may vary based on how strong they are." [2]



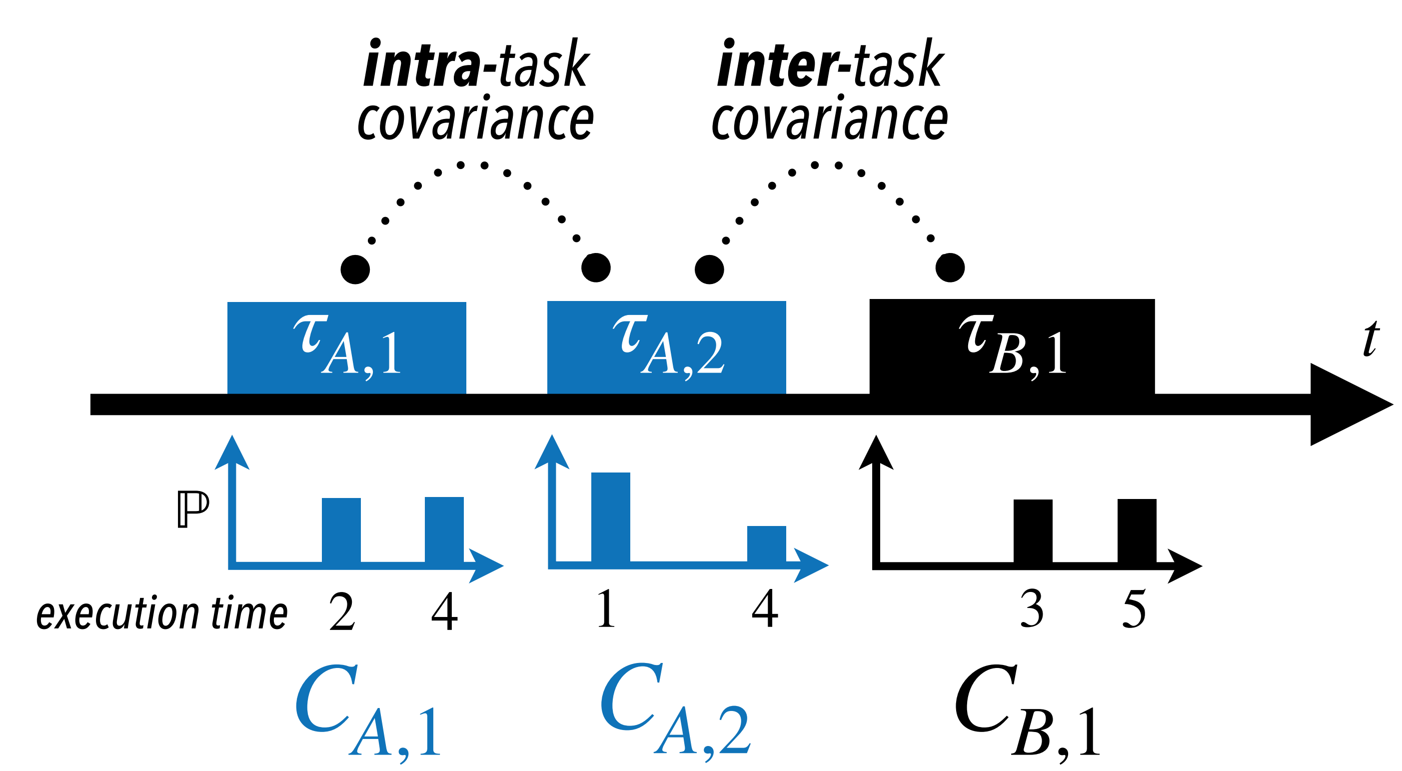
Covariance *– a measure of the degree to which two random variables fluctuate together.*

Covariance of $C_{A,1}$ and $C_{B,1}$

$Cov[C_{A,1}, C_{B,1}] = 1$ … $Cov[C_{A,1}, C_{B,1}] = -1$

HOW TO MODEL DEPENDENCE?

*"How to handle … dependences between the execution times of (i) jobs of **the same task**, and (ii) jobs **of different tasks?**"* [2]



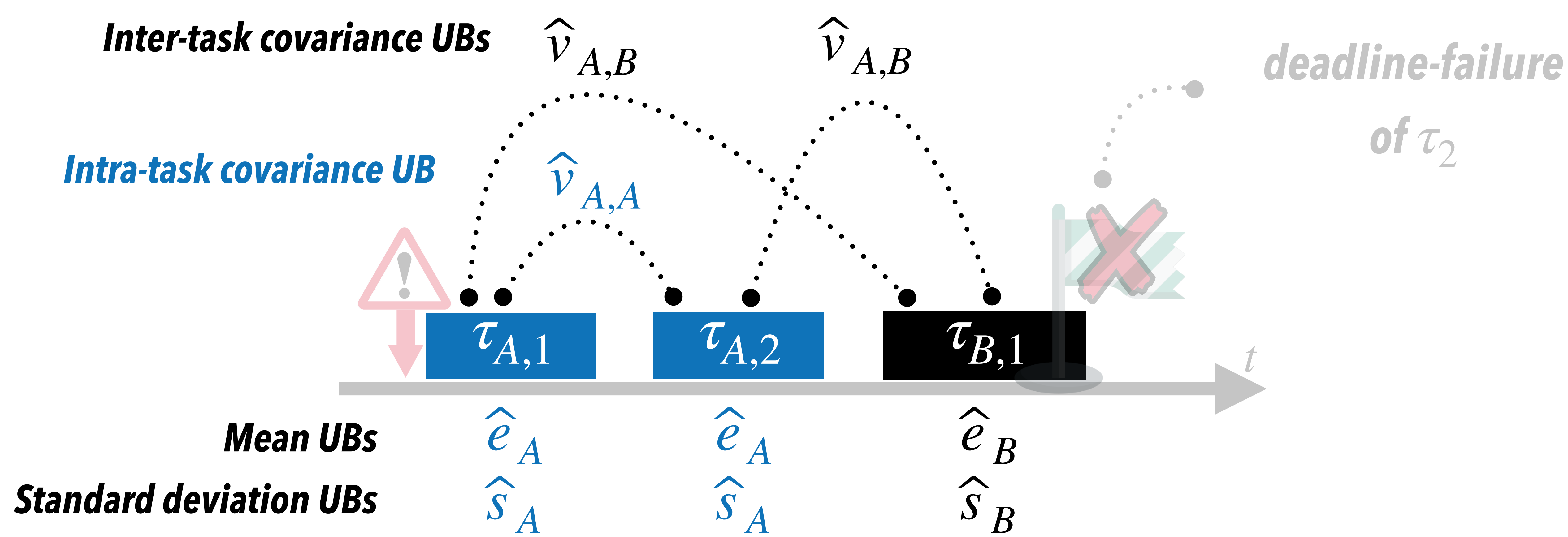
[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean *LITES* (2019)

CORRELATION-AWARE ANALYSIS

*An **efficient** way to analyse deadline-failure probability using **covariances***

***Inputs**:*

- $\widehat{e}_i$: *upper bound on the **mean** execution time (ET) of any job of task* τ_i
- $\widehat{s}_i$: *upper bound on the **standard deviation** of ET of any job of* τ_i
- $\widehat{v}_{i,i}$: *upper bound on the ET **intra**-task **covariance** between any two jobs of* τ_i
- $\widehat{v}_{i,k}$: *upper bound on the ET **inter**-task **covariance** between any two jobs of two distinct tasks* τ_i *and* τ_k



HOW TO USE IT IN RT ANALYSIS?

"Analyses are needed that can address dependencies." [2]

[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean *LITES* (2019)

CAA IN A NUTSHELL

*The **goal** is to express everything in terms of CAA inputs*

deadline-failure probability of τ_B

$$\mathbb{P}[R_B \geq d] \leq \boxed{?}$$

unknown response-time distribution of τ_B

relative deadline

CAA IN A NUTSHELL

CAA rests on …

deadline-failure probability of τ_B

unknown response-time distribution of τ_B

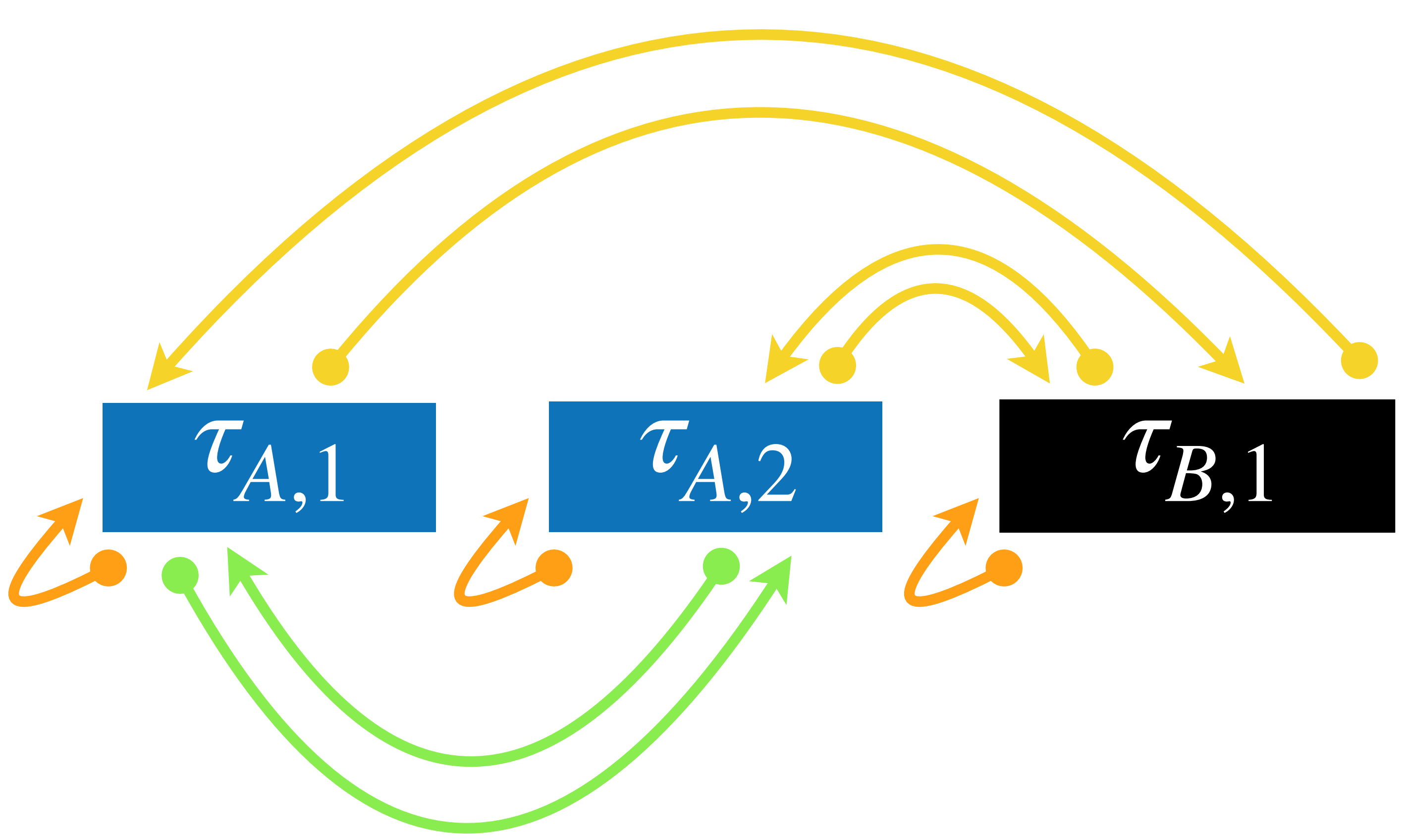
relative deadline

$$\mathbb{P}[R_B \geq d] \leq \frac{\widehat{\mathbb{V}}[R_B]}{\widehat{\mathbb{V}}[R_B] + \left(d - \widehat{\mathbb{E}}[R_B]\right)^2}$$

upper bound on the variance of R_B

relative deadline of τ_B

upper bound on the mean of R_B



all covariance pairs

$$\widehat{\mathbb{E}}[R_B] = 2 \cdot \hat{e}_A + \hat{e}_B$$

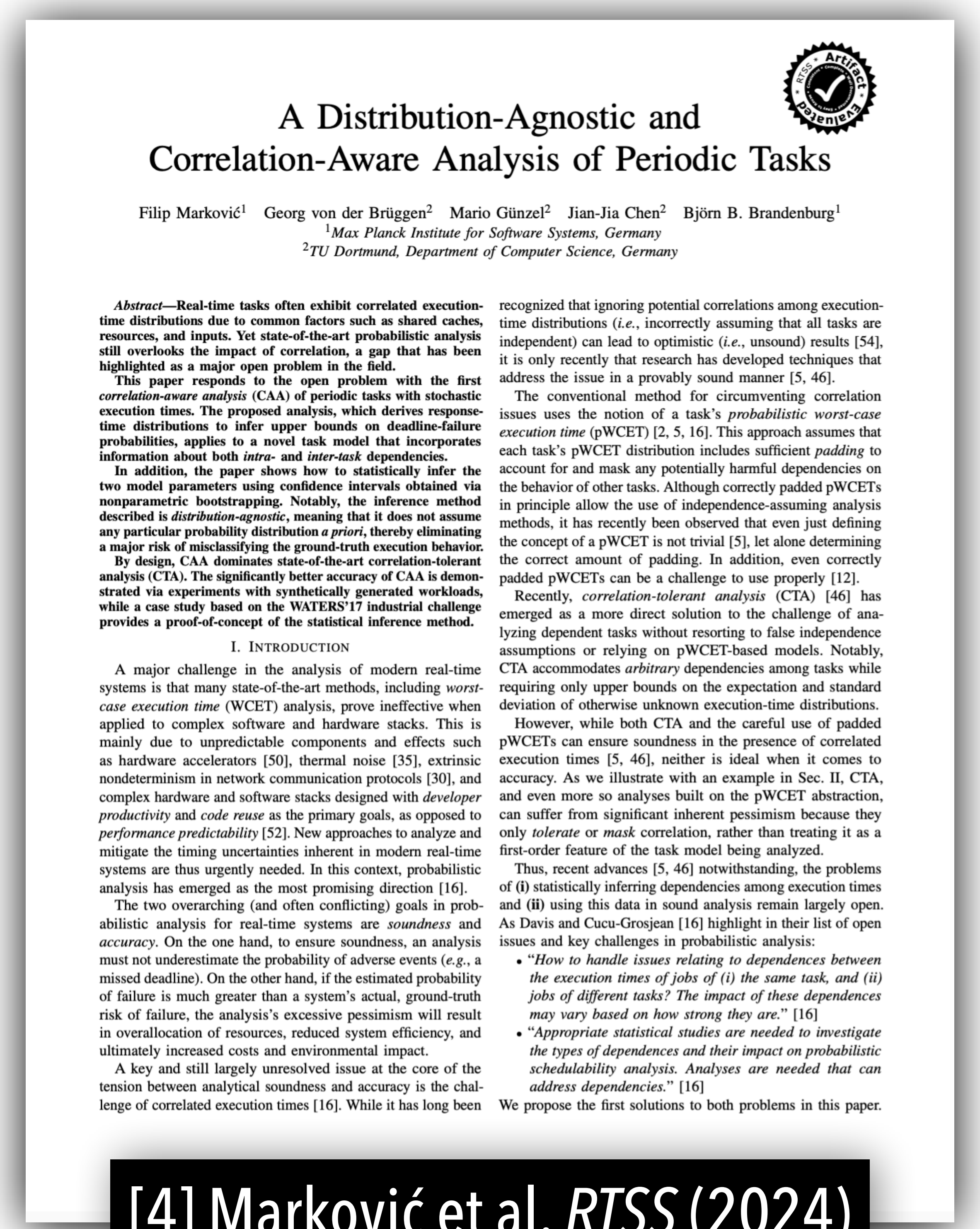
$$\widehat{\mathbb{V}}[R_B] = 2 \cdot \hat{s}_A^2 + \hat{s}_B^2 + 2 \cdot \hat{v}_{A,A} + 4 \cdot \hat{v}_{A,B}$$

variance UBs **Intra-task cov UB** **Inter-task cov UB**

EVALUATION

*How do **CAA** results compare to **CTA** in general?*

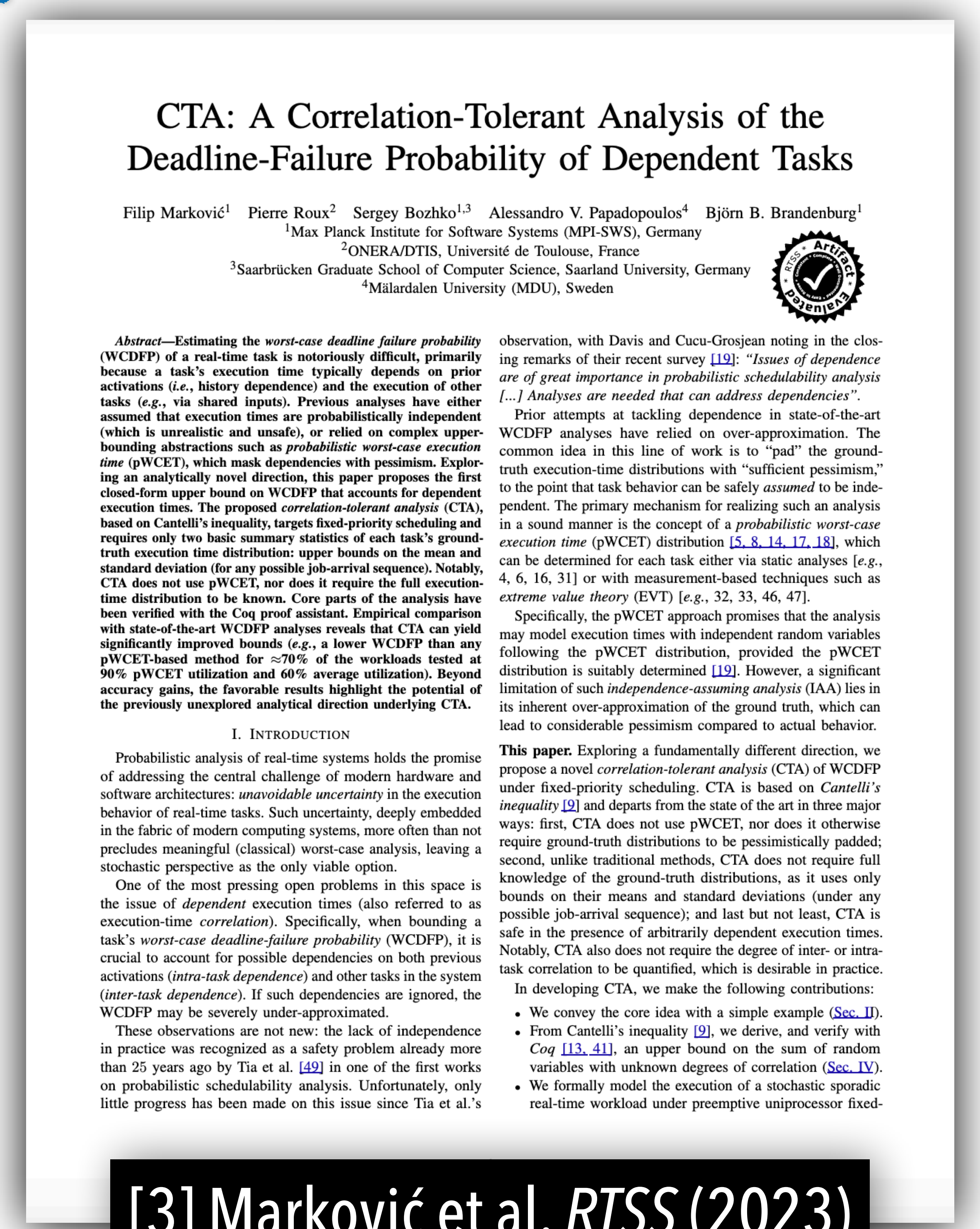
We compared **CAA** and **CAA**.



A Distribution-Agnostic and Correlation-Aware Analysis of Periodic Tasks

Filip Marković[1] Georg von der Brüggen[2] Mario Günzel[2] Jian-Jia Chen[2] Björn B. Brandenburg[1]
[1]Max Planck Institute for Software Systems, Germany
[2]TU Dortmund, Department of Computer Science, Germany

[4] Marković et al. *RTSS* (2024)



CTA: A Correlation-Tolerant Analysis of the Deadline-Failure Probability of Dependent Tasks

Filip Marković[1] Pierre Roux[2] Sergey Bozhko[1,3] Alessandro V. Papadopoulos[4] Björn B. Brandenburg[1]
[1]Max Planck Institute for Software Systems (MPI-SWS), Germany
[2]ONERA/DTIS, Université de Toulouse, France
[3]Saarbrücken Graduate School of Computer Science, Saarland University, Germany
[4]Mälardalen University (MDU), Sweden

[3] Marković et al. *RTSS* (2023)

EVALUATION

*Synthetic task sets were randomly generated to highlight differences between **CTA** and **CAA**.*

Four experiments were conducted to investigate:

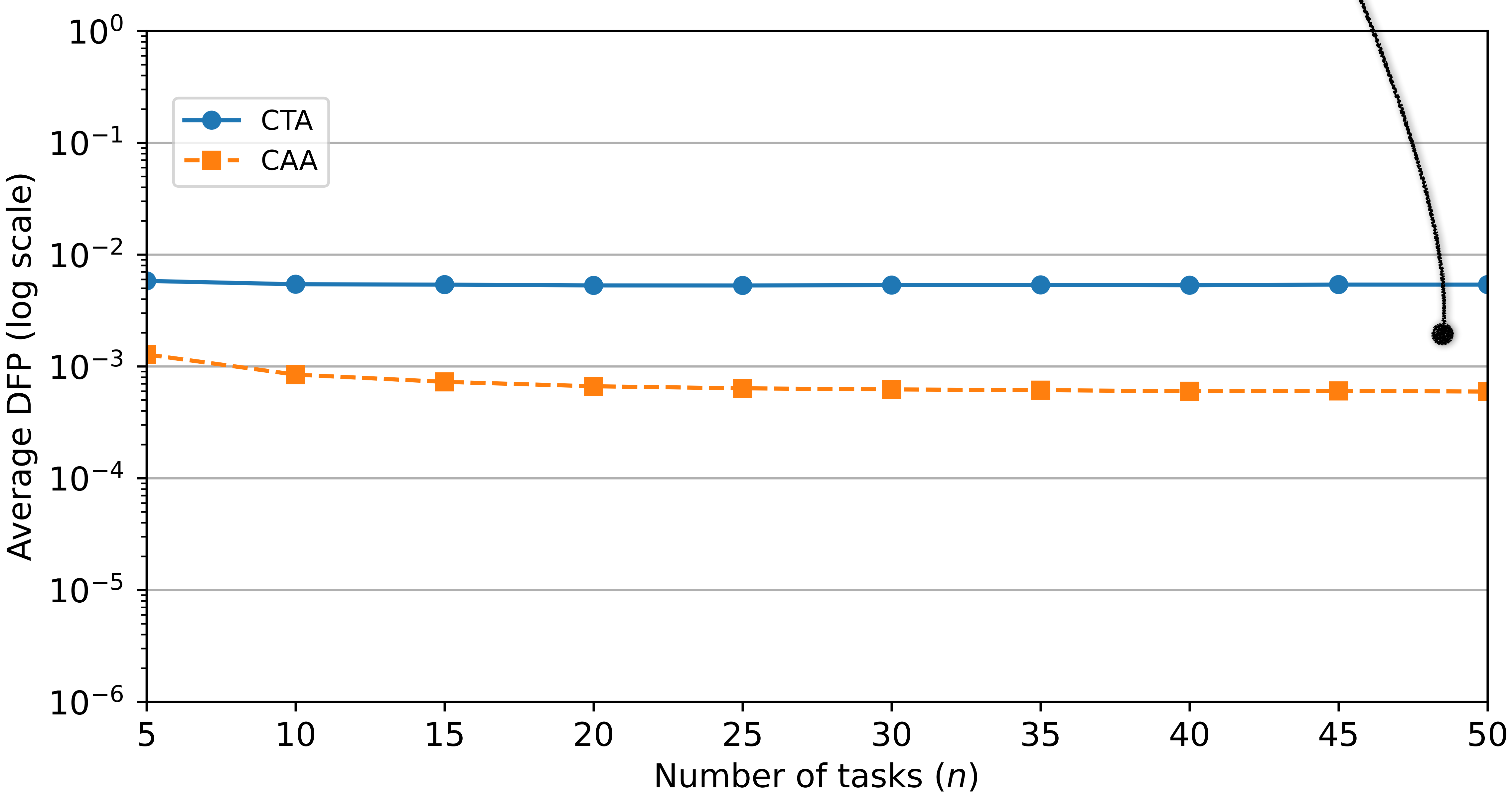
1. Influence of the **task set size** on DFP,
2. The influence of the **total mean utilization** on DFP,
3. The influence of the **maximum standard deviation** on DFP,
4. The influence of the **maximum correlation** on DFP.

In this talk, we focus on (1).

EVALUATION

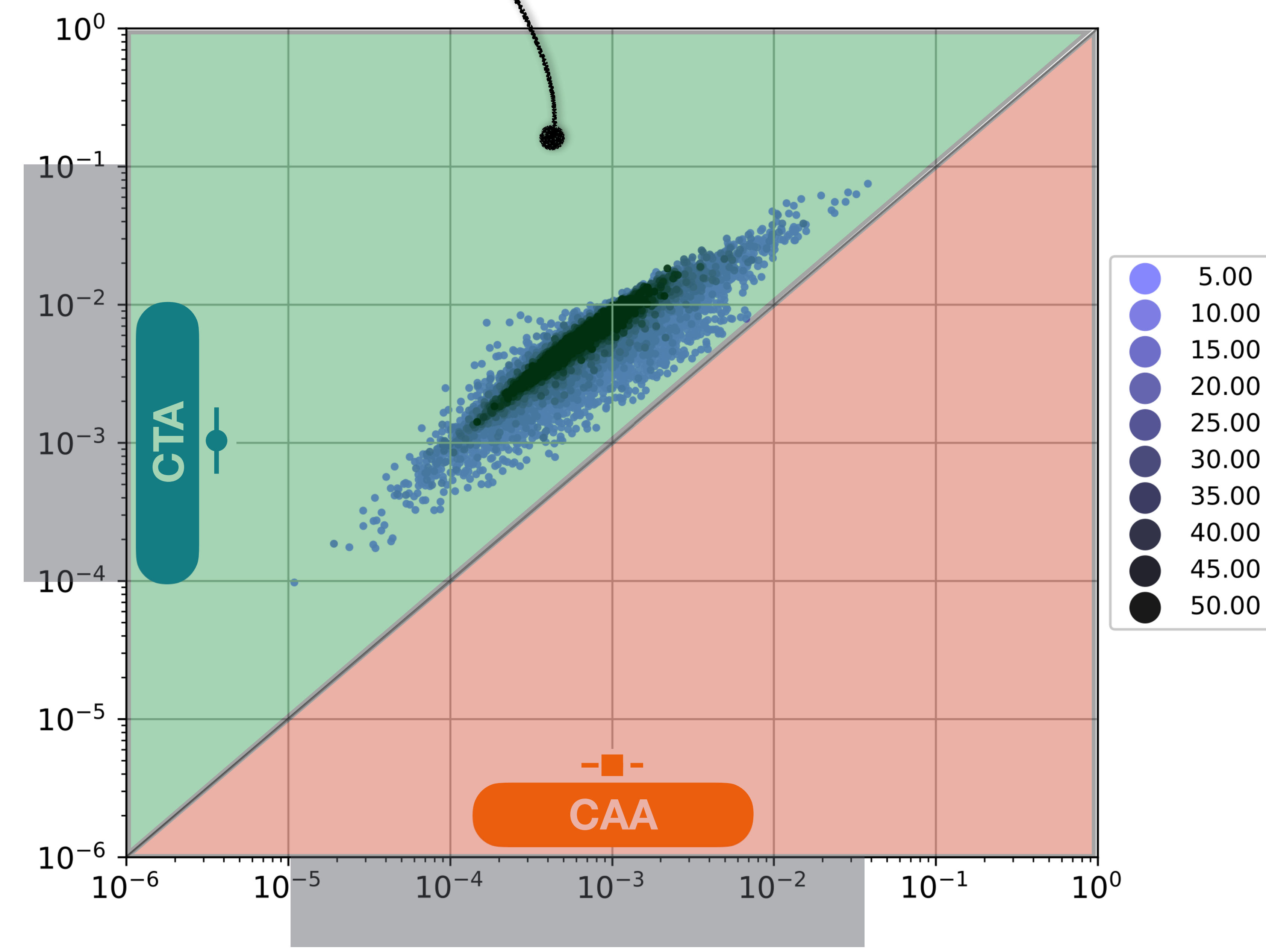
Investigating the influence of the task-set size

***log scale** ***lower is better**



CAA derives **better** bounds than ***CTA***

CAA ***always*** derives **better** bounds than ***CTA***



CAA dominates **CTA**, i.e., it always computes better or, at worst, equal DFP estimates

The gap becomes slightly more pronounced as the number of tasks increases due to **Bienaymé's** identity

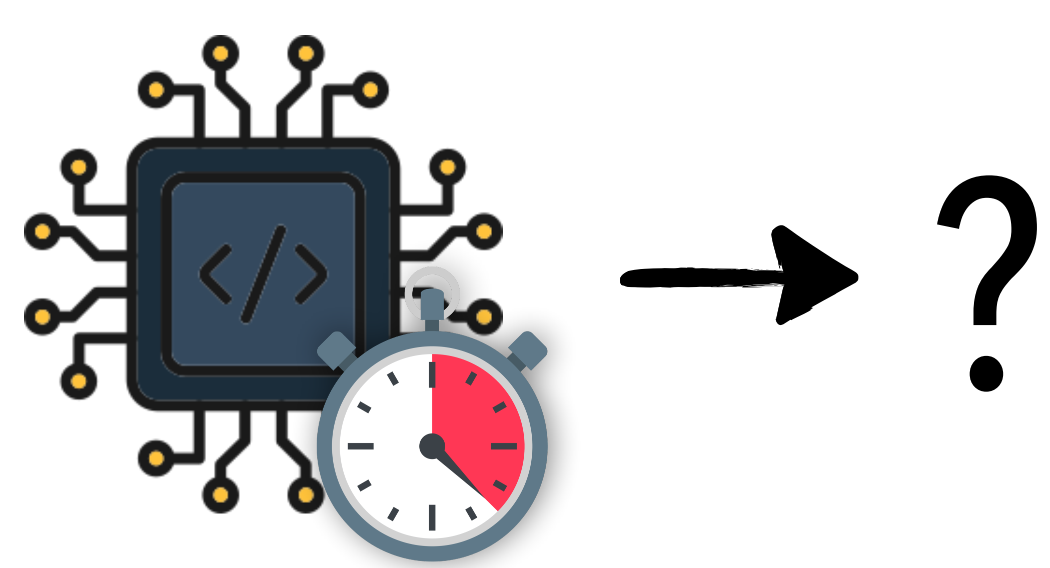
$$\widehat{\mathbb{V}}[R_B] = 2 \cdot \hat{s}_A^2 + \hat{s}_B^2 + 2 \cdot \hat{v}_{A,A} + 4 \cdot \hat{v}_{A,B}$$

$$\widehat{\mathbb{V}}^{CTA}[R_B] = 2 \cdot \hat{s}_A^2 + \hat{s}_B^2 + 2 \cdot \hat{s}_A^2 + 4 \cdot \hat{s}_A \cdot \hat{s}_B$$

BUT, HOW DO WE DERIVE CAA (AND CTA) INPUTS? THIS PAPER (PART 2)

HOW TO STATISTICALLY INFER DEPENDENCE?

*"Appropriate statistical studies are needed to investigate the **types of dependences** and their impact on probabilistic schedulability analysis"* [2]



[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean *LITES* (2019)

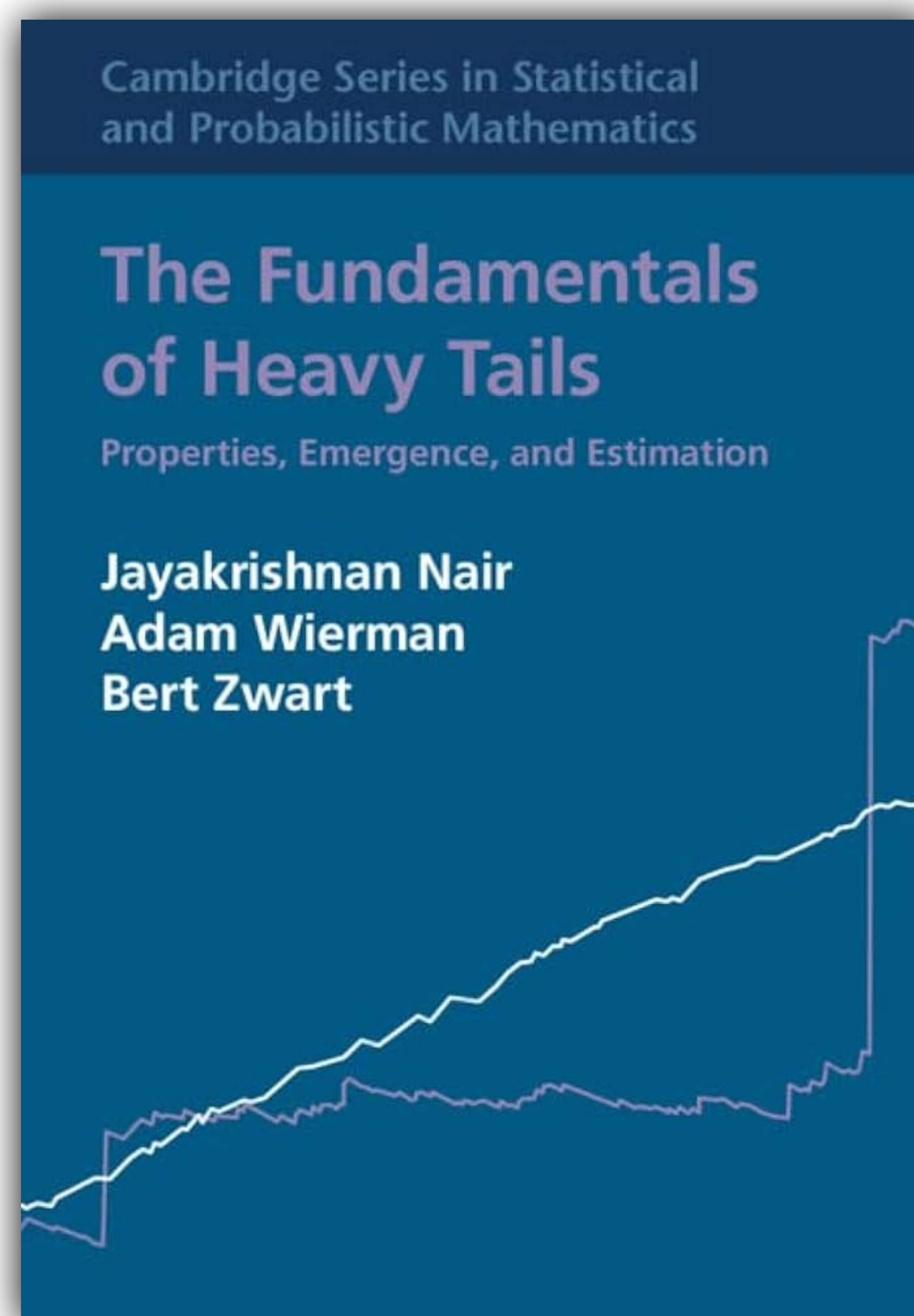
ANOTHER PROBLEM: DISTRIBUTION MISCLASSIFICATION

***Incorrectly** assuming the **wrong** underlying distribution can lead to **unsound** results*

Many **refuted** results!

- **heavy-tailed distributions**
  As explained in The Fundamentals of Heavy Tails, 2022, Jayakrishnan et al.

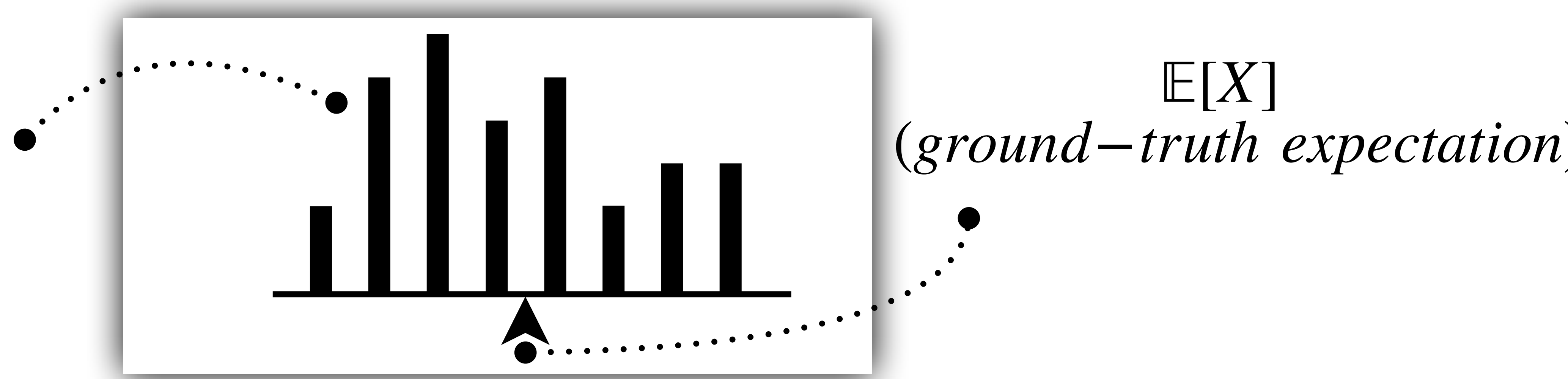
Since our research focuses on safety properties, it is crucial to eliminate the risk of distribution misestimation.



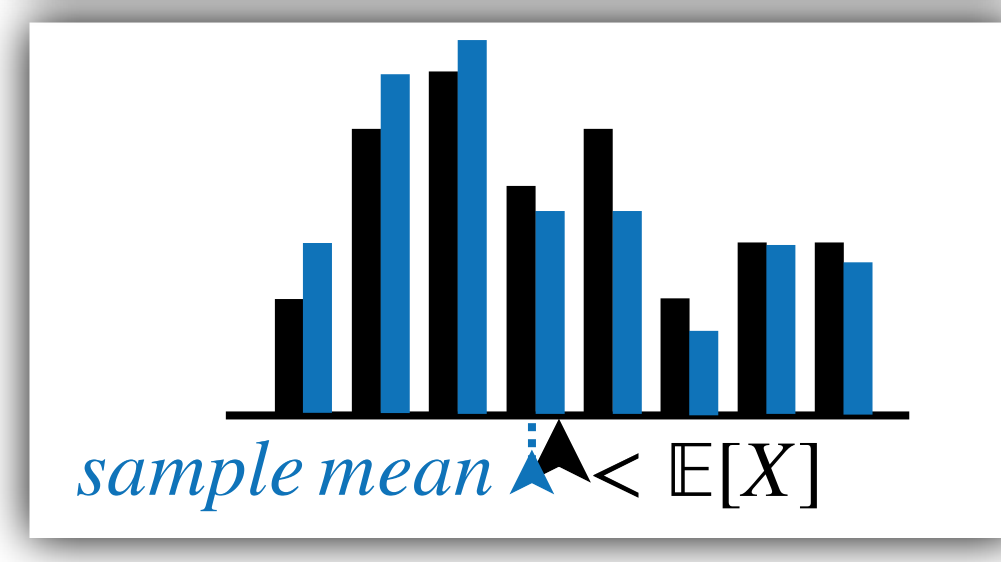
GOAL: DISTRIBUTION-AGNOSTIC STATISTICAL INFERENCE

We use Nonparametric Bootstrap

*A random variable with an **unknown** distribution and **unknown** expectation*

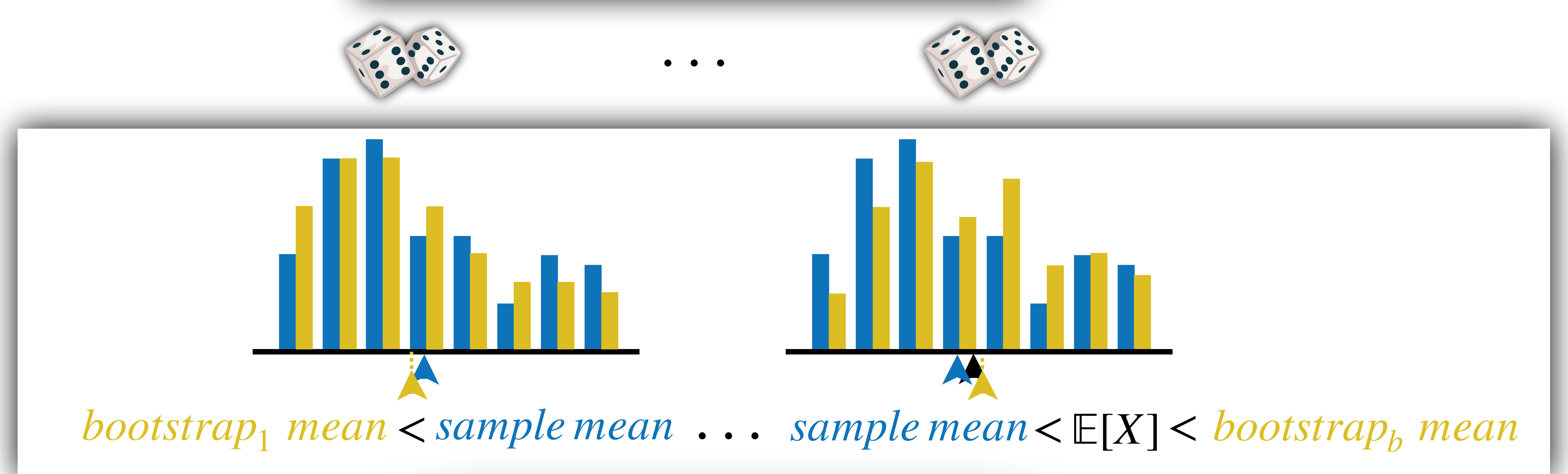


Step 1. *Draw an initial sample I of n independent observations of X*

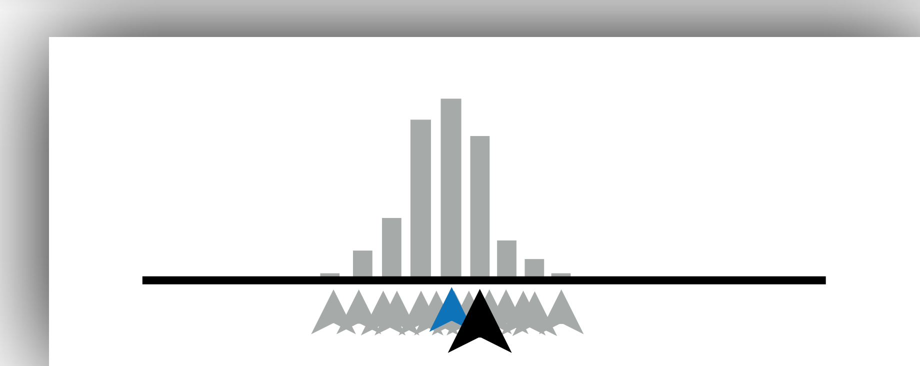


Step 2. *Generate b bootstrap samples (by randomly resampling with replacement from the initial sample I)*

Step 3. *Compute the bootstrap statistic on each bootstrap sample B*



Output: *Bootstrap distribution of the expectation*



But how do we obtain sound upper bounds from the bootstrap distribution?

CONFIDENCE INTERVALS

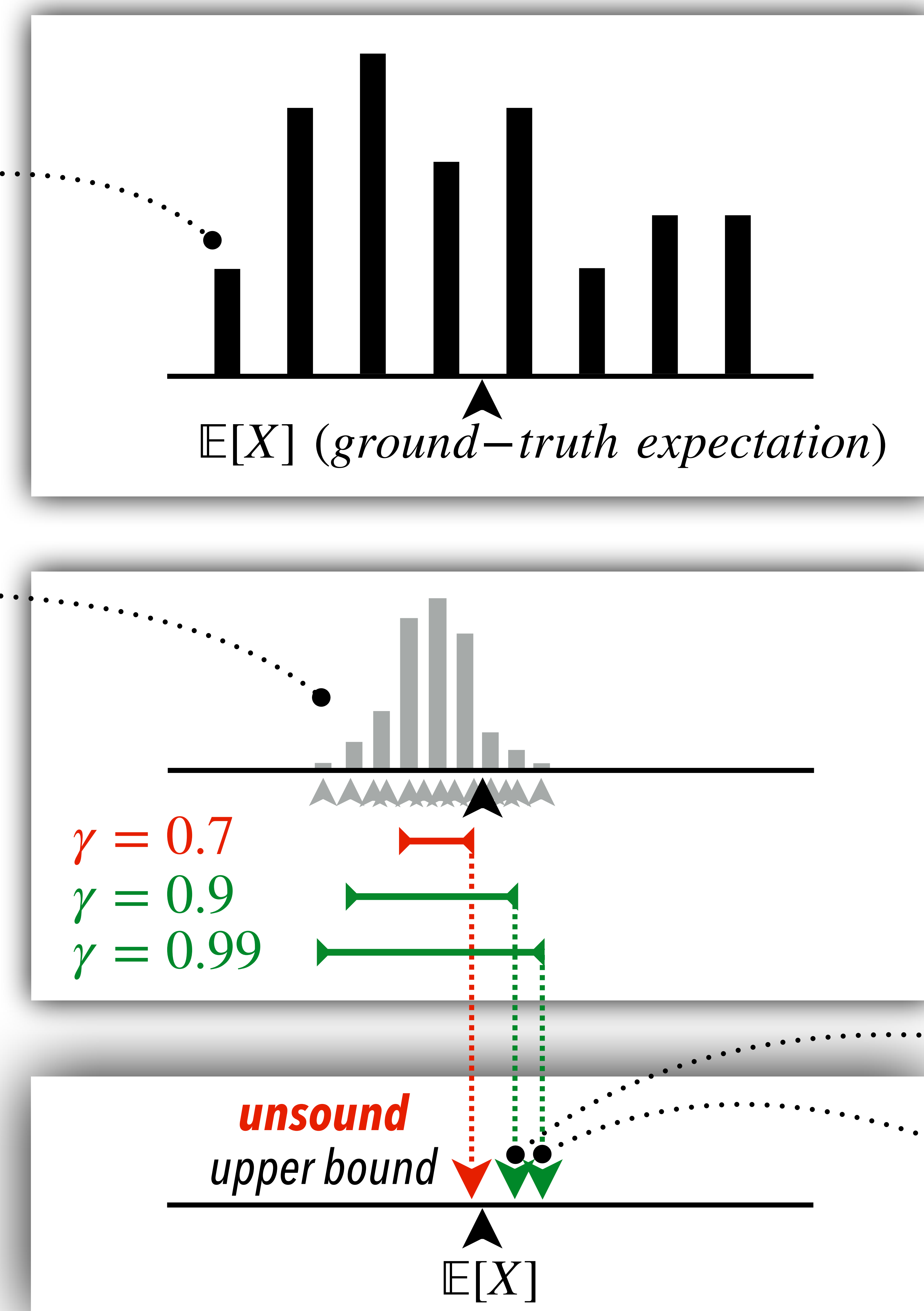
*The key to deriving **upper bounds** on parameters from an unknown distribution*

*A random variable X with an **unknown** distribution and **unknown** expectation*

*Bootstrap distribution of the **expectation***

***Step 4.** Compute the confidence interval with a given level of confidence γ*

***Output:** The **upper bound** of the confidence interval.*



***sound** upper bound*

***sound but overly pessimistic** upper bound*

WHAT ARE THE GUARANTEES?

How should we interpret the derived results?

Wrong confidence interval interpretation

- *"A common misunderstanding about CIs is that for say a 95% CI (A to B), there is a 95% probability that the true population mean lies between A and B."*

[5]

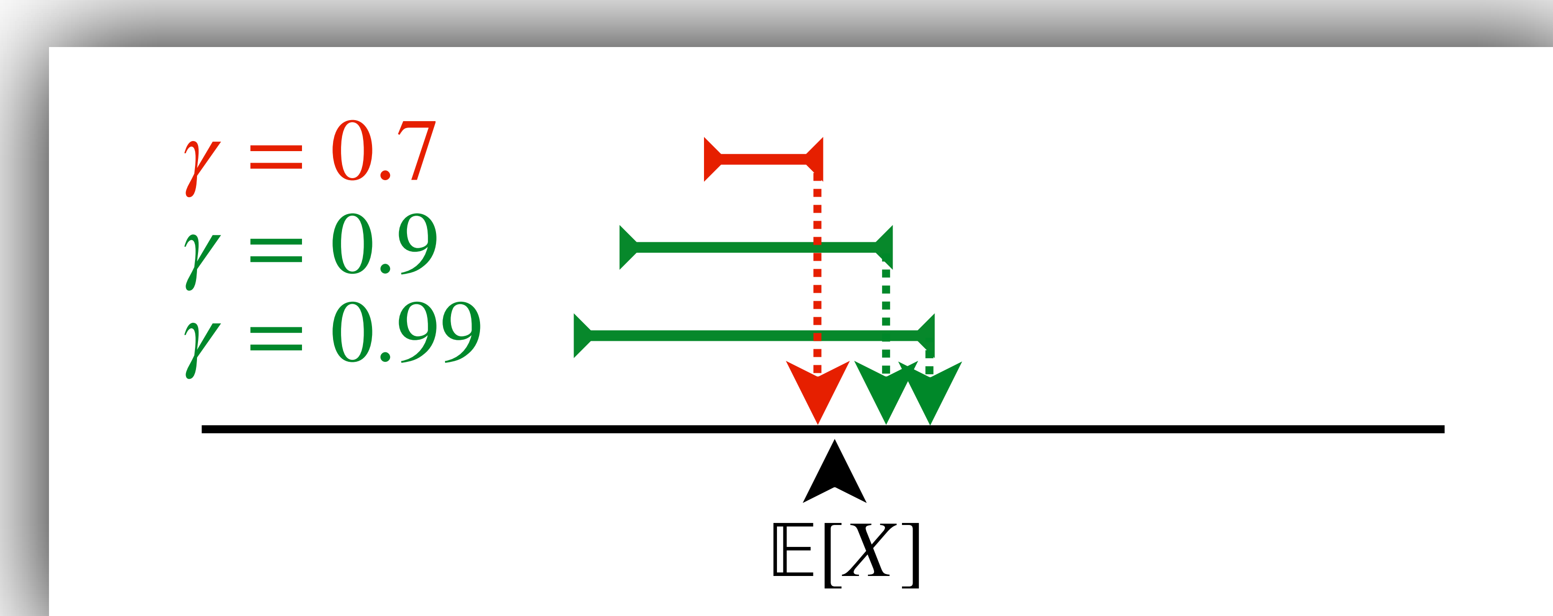
Correct confidence interval intepretation

- *"A 95% CI simply means that if the study is conducted multiple times (multiple sampling from the same population) with corresponding 95% CI for the mean constructed, we expect 95% of these CIs to contain the true population mean"*

[5]

- ***No*** *statistical inference method can provide* ***absolute certainty***
- *There is always a minuscule but* ***non-zero chance*** *that a ground-truth parameter* ***lies outside*** *the statistically estimated range*

- *Bootstrapped CIs provide an excellent means of estimating ground truth that is*
  - ***statistically rigorous***
  - ***distribution-agnostic***
  - ***sample-efficient***
  - ***mathematically well-understood***



[5] *"The correct interpretation of confidence intervals"* Tan and Tan. *Proceedings of Singapore Healthcare* (2010)

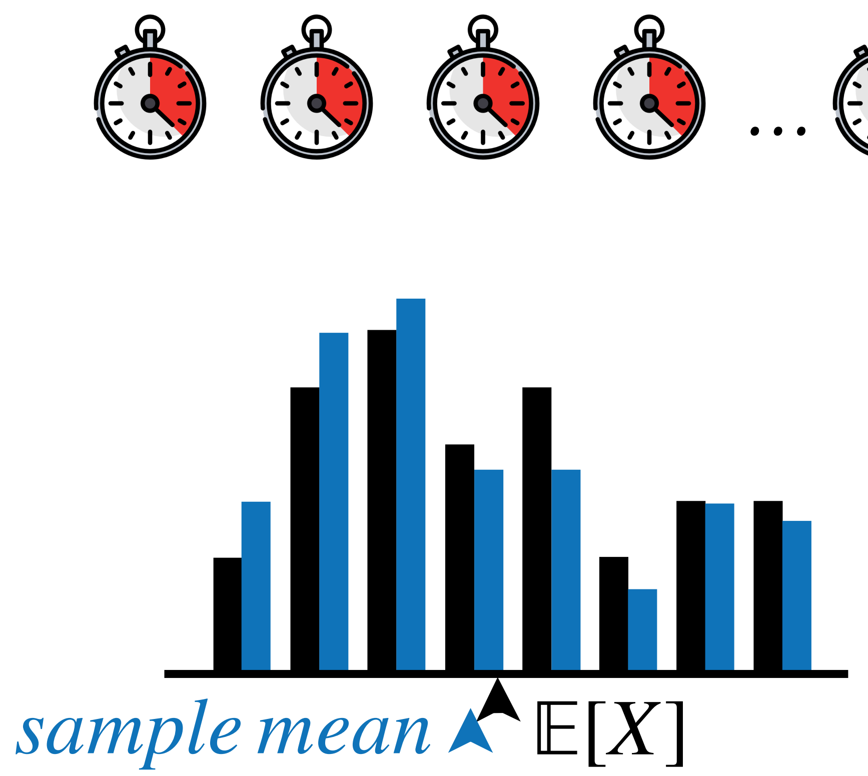
CONFIDENCE-BASED TASK PARAMETERS

All CAA and CTA inputs are obtained with bootstrapping from bounds on confidence intervals

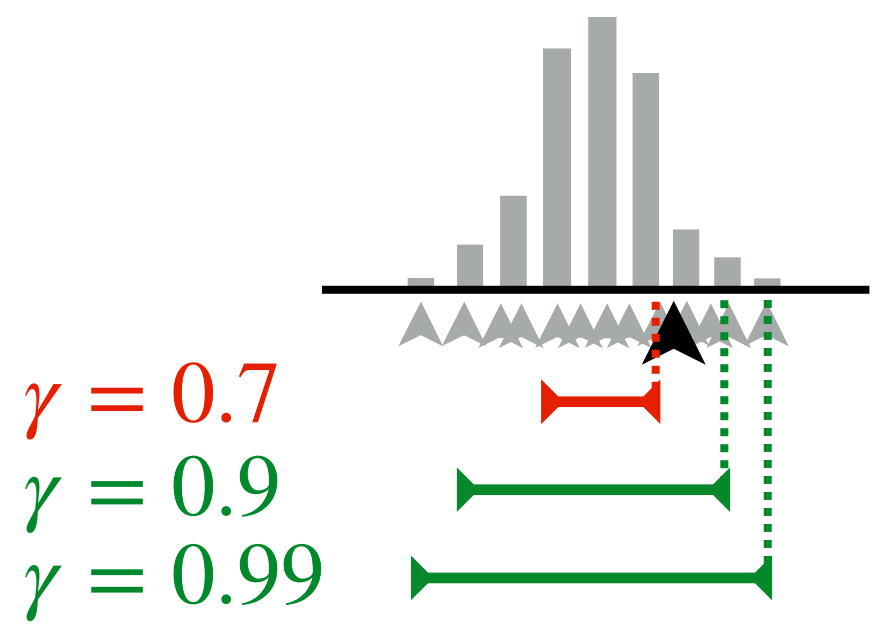
- $\widehat{e}_i$: *upper bound on the* ***mean*** *execution time (ET)*
- $\widehat{s}_i$: *upper bound on the* ***standard deviation*** *of ET*
- $\widehat{v}_{i,i}$: *upper bound on the ET* ***intra****-task* ***covariance***
- $\widehat{v}_{i,k}$: *upper bound on the ET* ***inter****-task* ***covariance***

Important tuning knobs

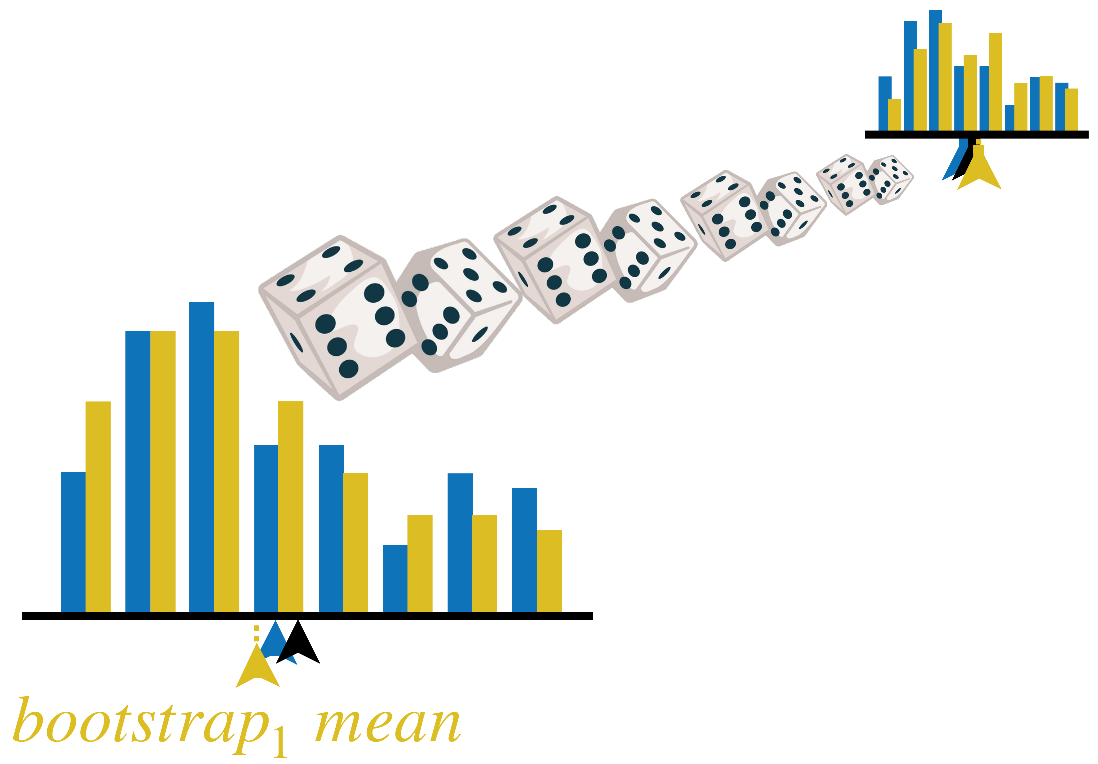
Initial sample *size* n



Level of ***confidence*** γ



Number of ***bootstrap*** *samples* b



CASE STUDY

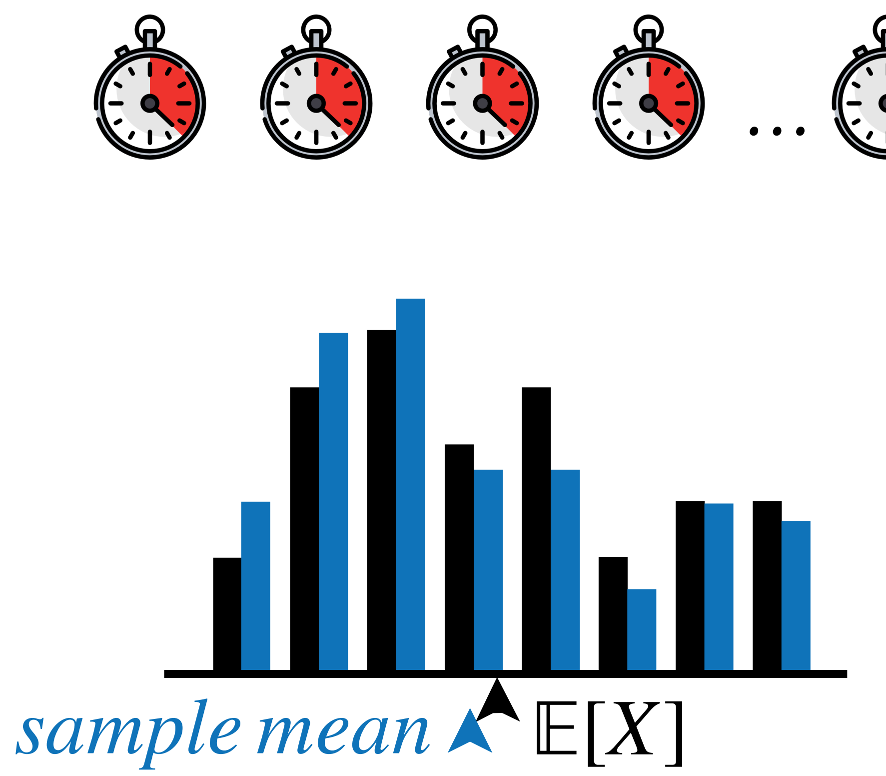
How do CAA results compare to CTA ***when using statistical inference****?*

Three experiments on a **proof-of-concept** case study from WATERS 2017 workloads were conducted to investigate:

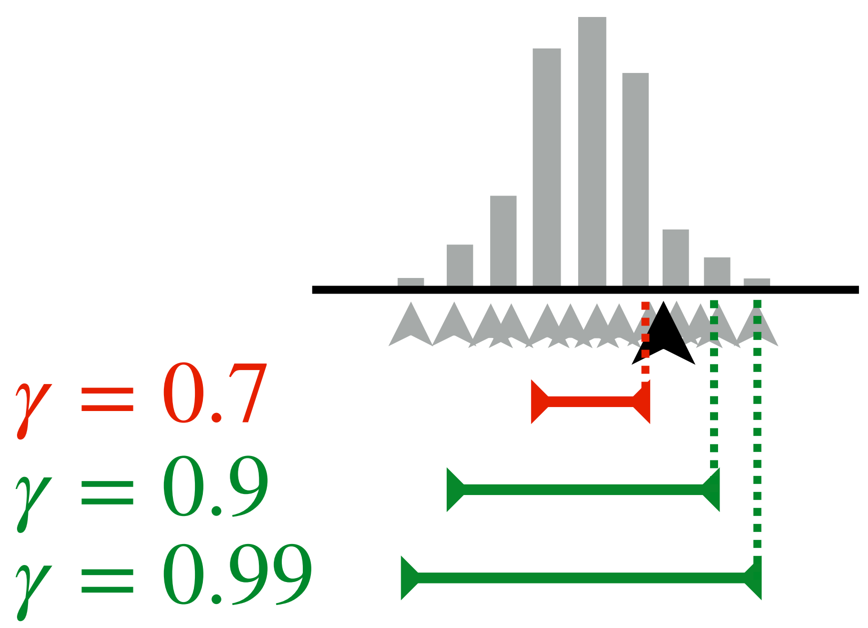
1. Influence of the **initial sample size** on DFP,
2. Influence of the **level of confidence** on DFP,
3. Influence of the **number of bootstrap samples** on DFP.

Important tuning knobs

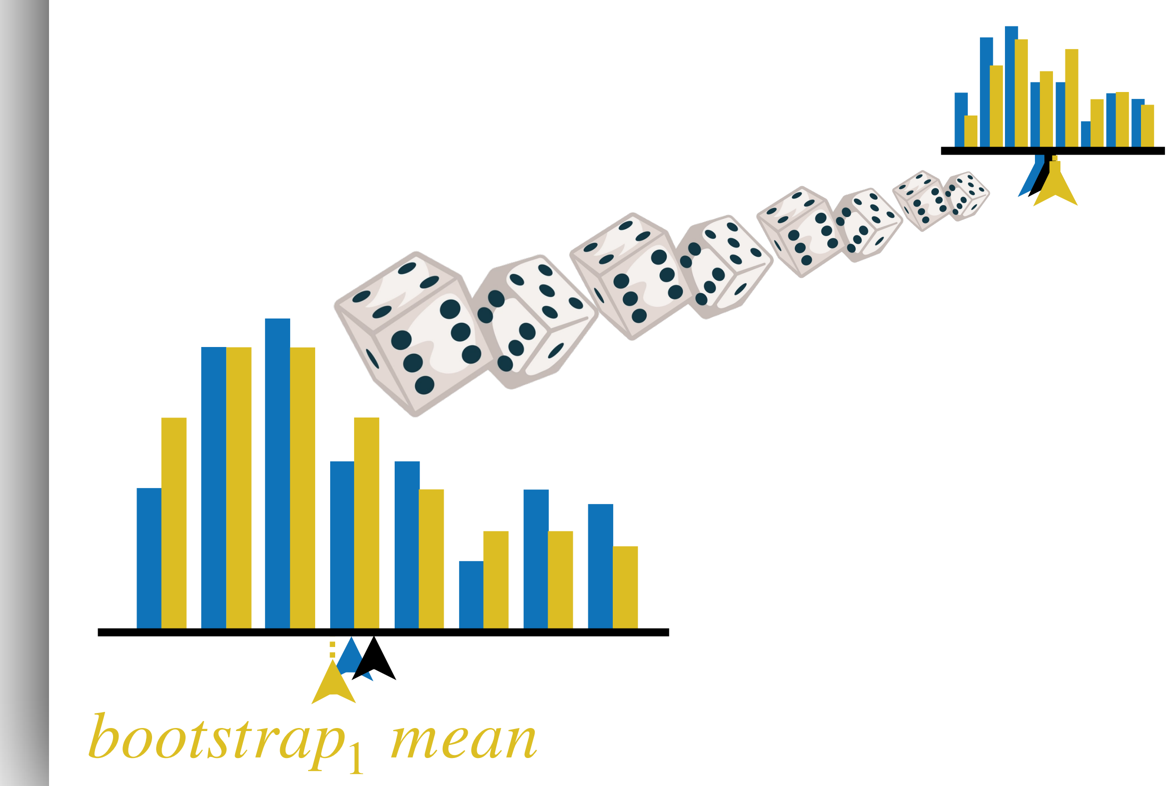
Initial sample *size* n



Level of ***confidence*** γ

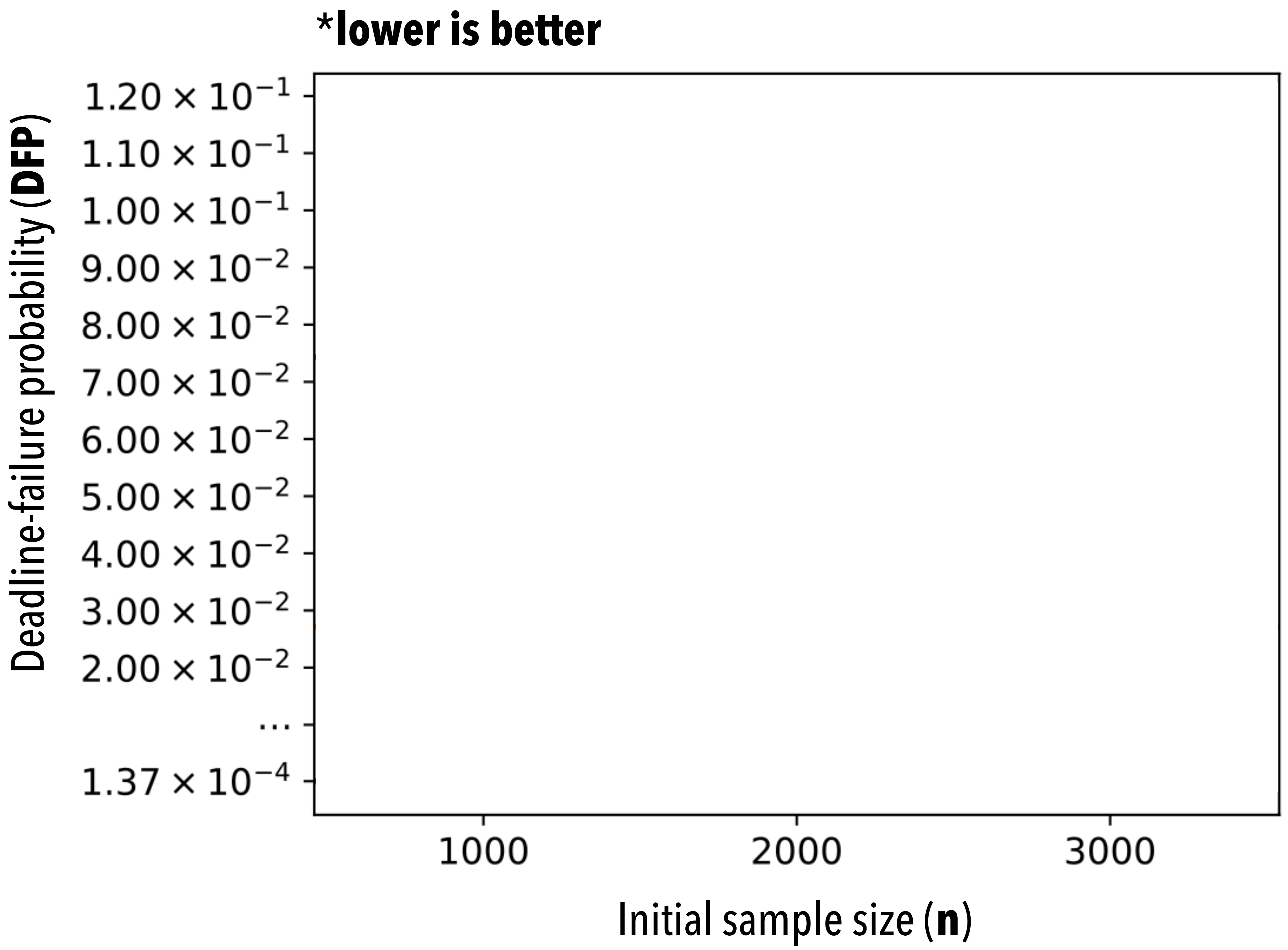


Number of ***bootstrap*** *samples* b



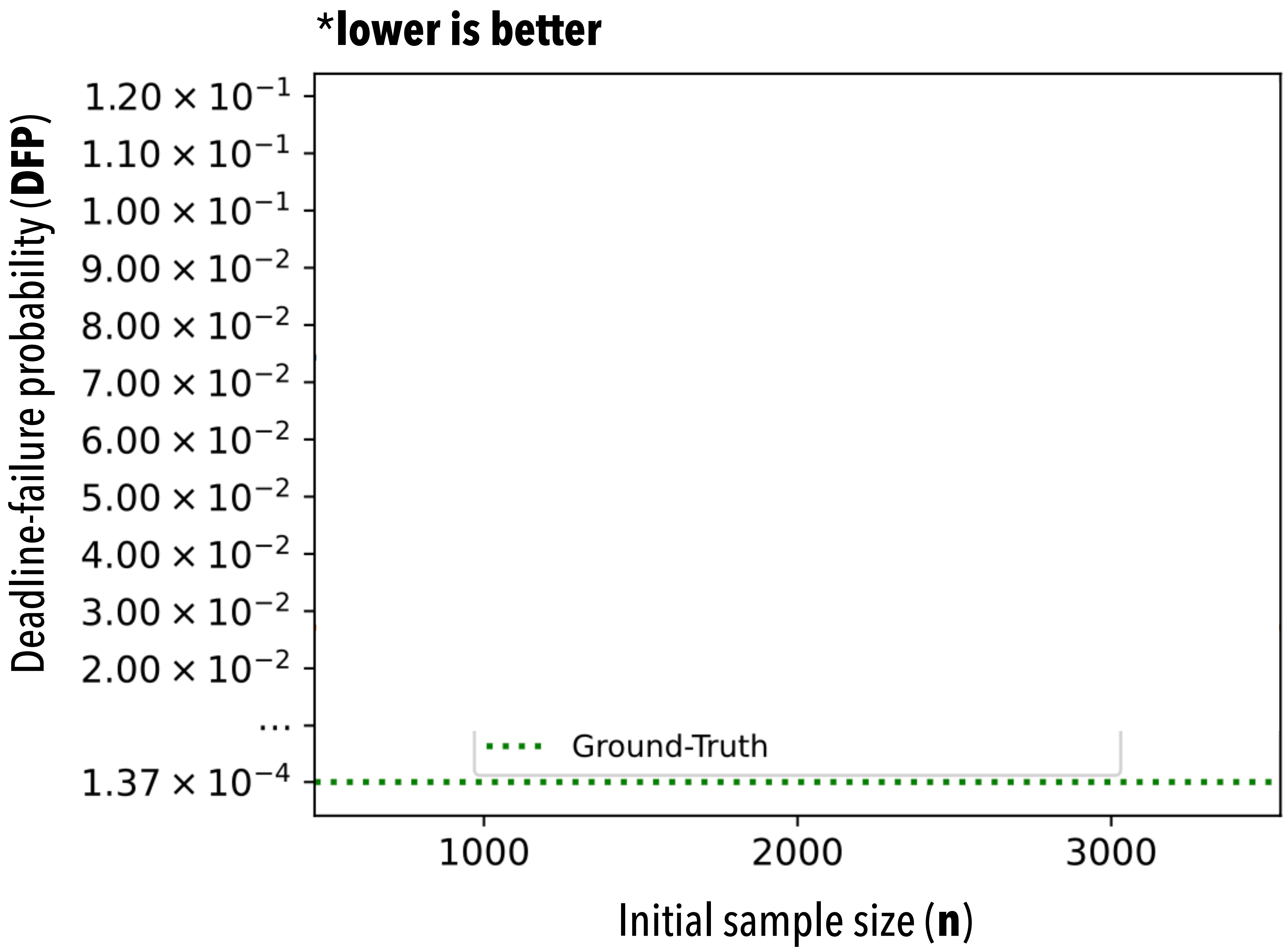
CASE STUDY

Investigating the influence of the ***initial sample size***



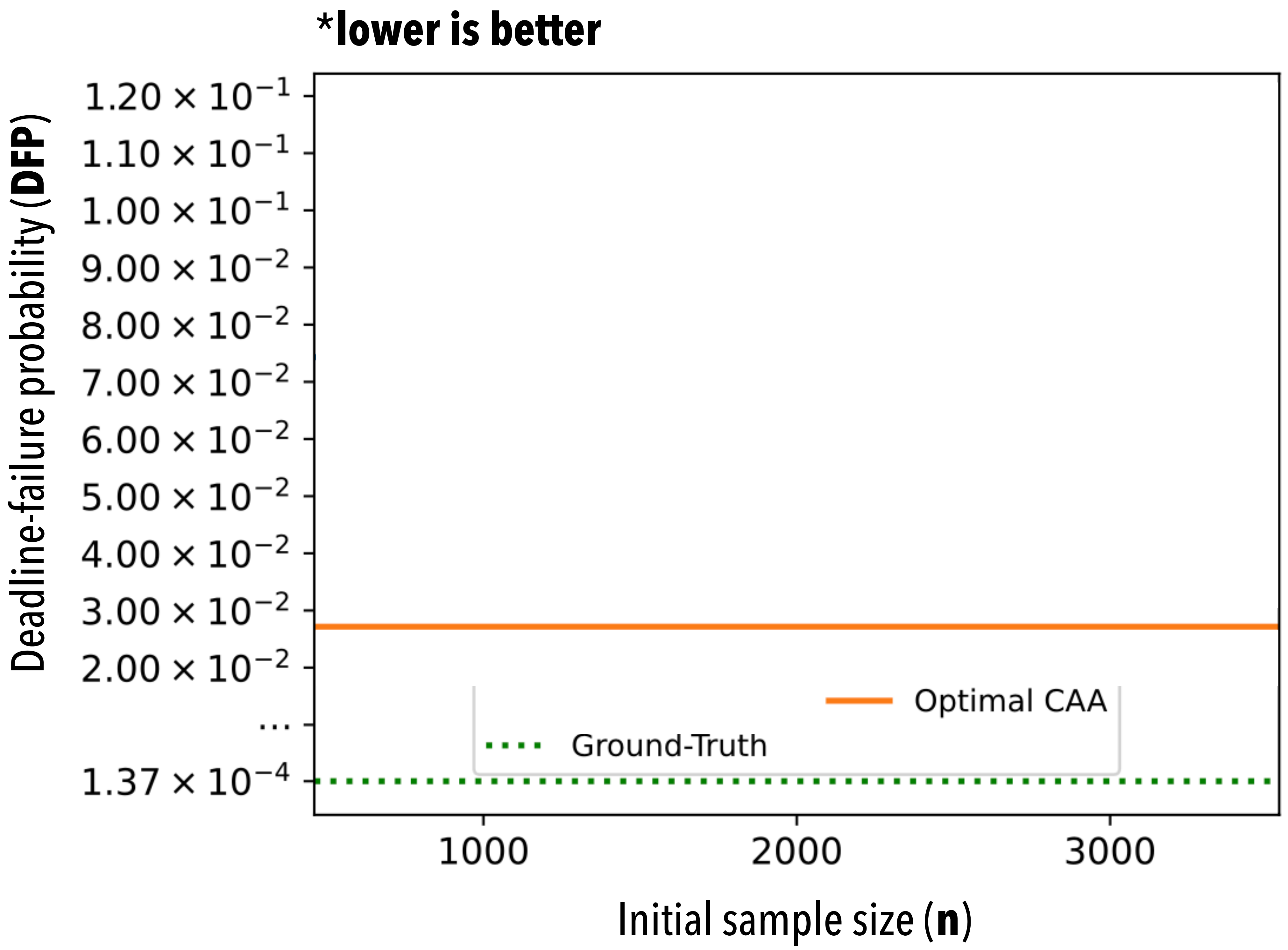
CASE STUDY

*Investigating the influence of the **initial sample size***



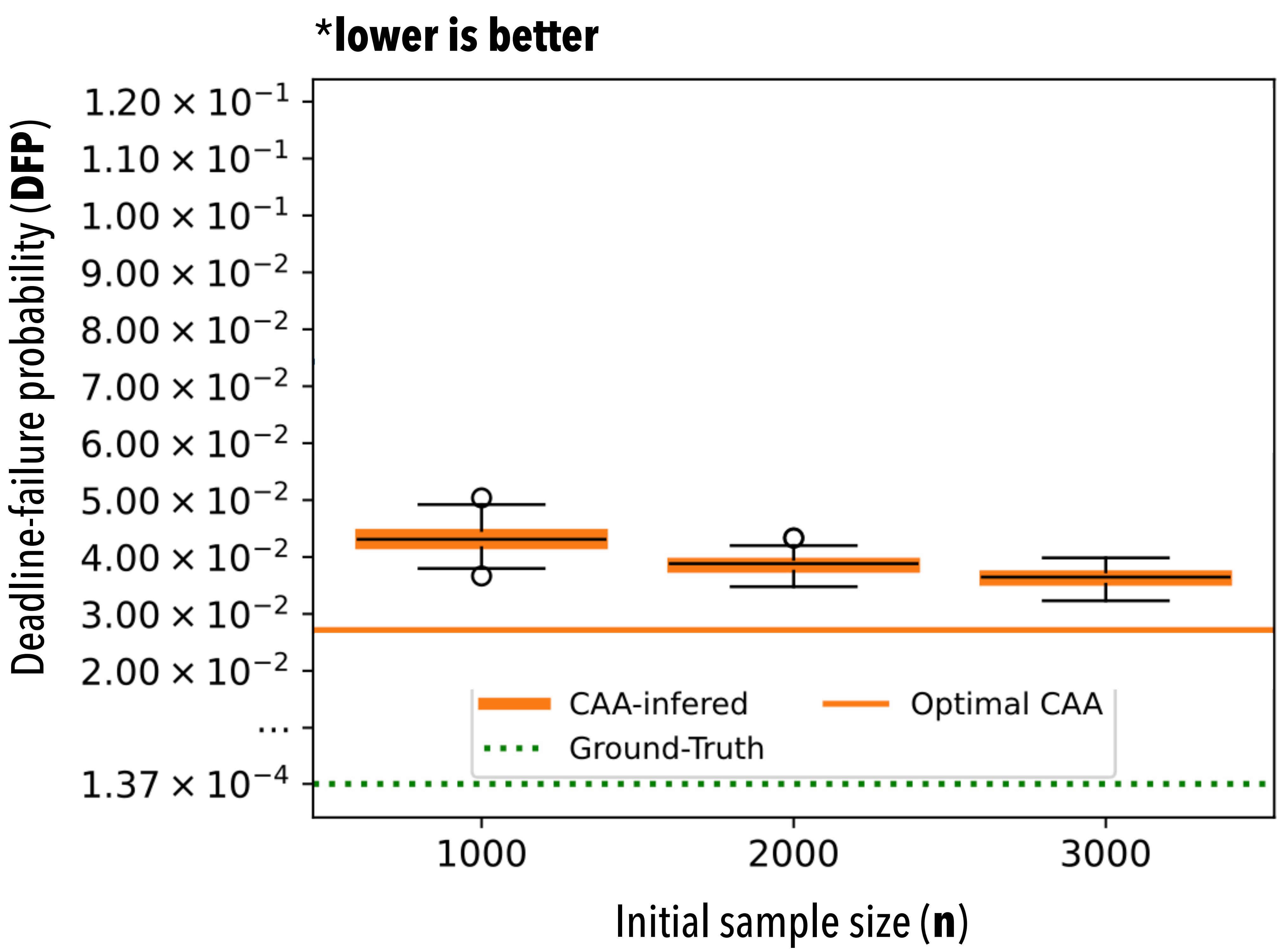
CASE STUDY

Investigating the influence of the ***initial sample size***



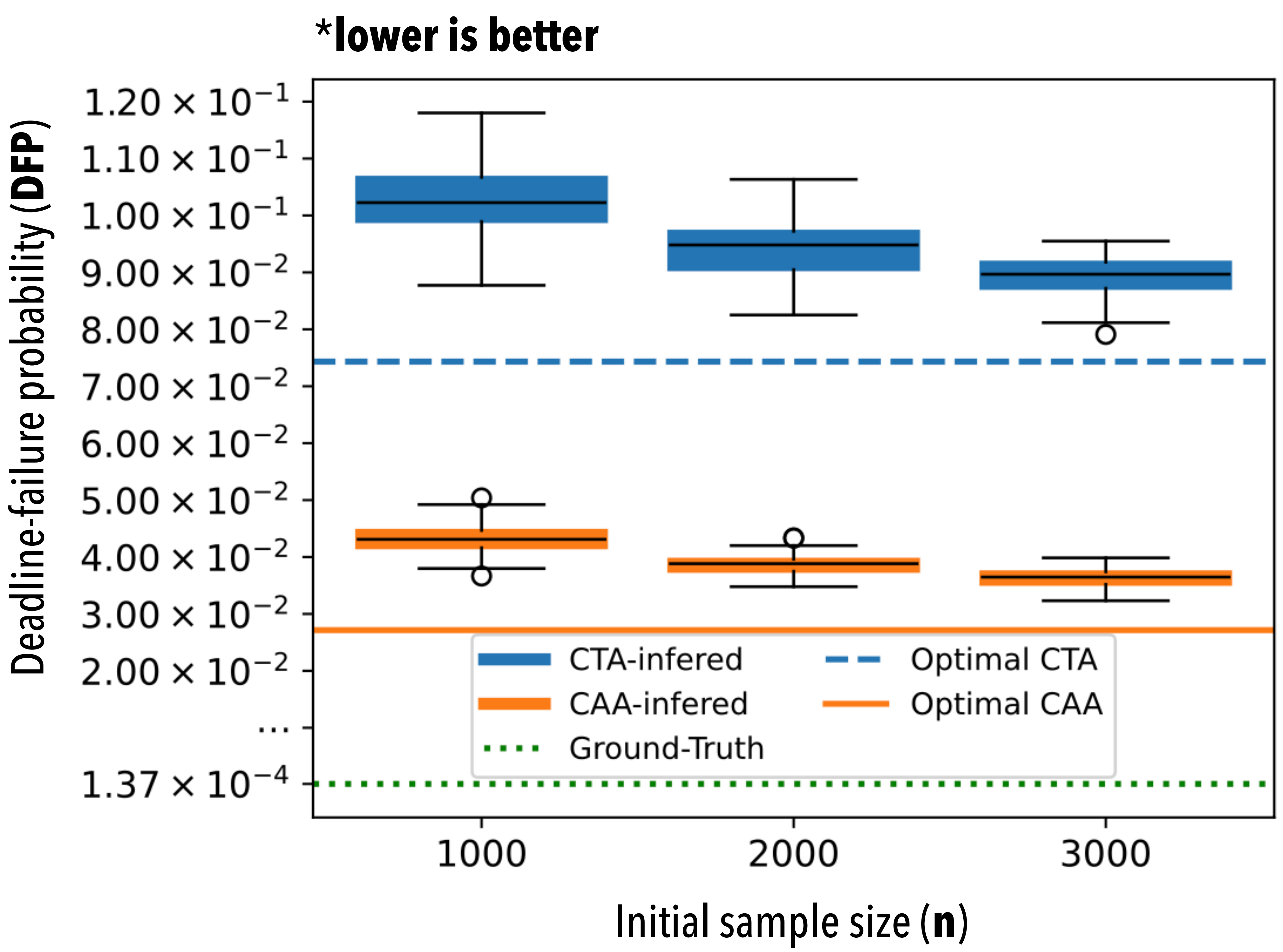
CASE STUDY

*Investigating the influence of the **initial sample size***



CASE STUDY

Investigating the influence of the ***initial sample size***

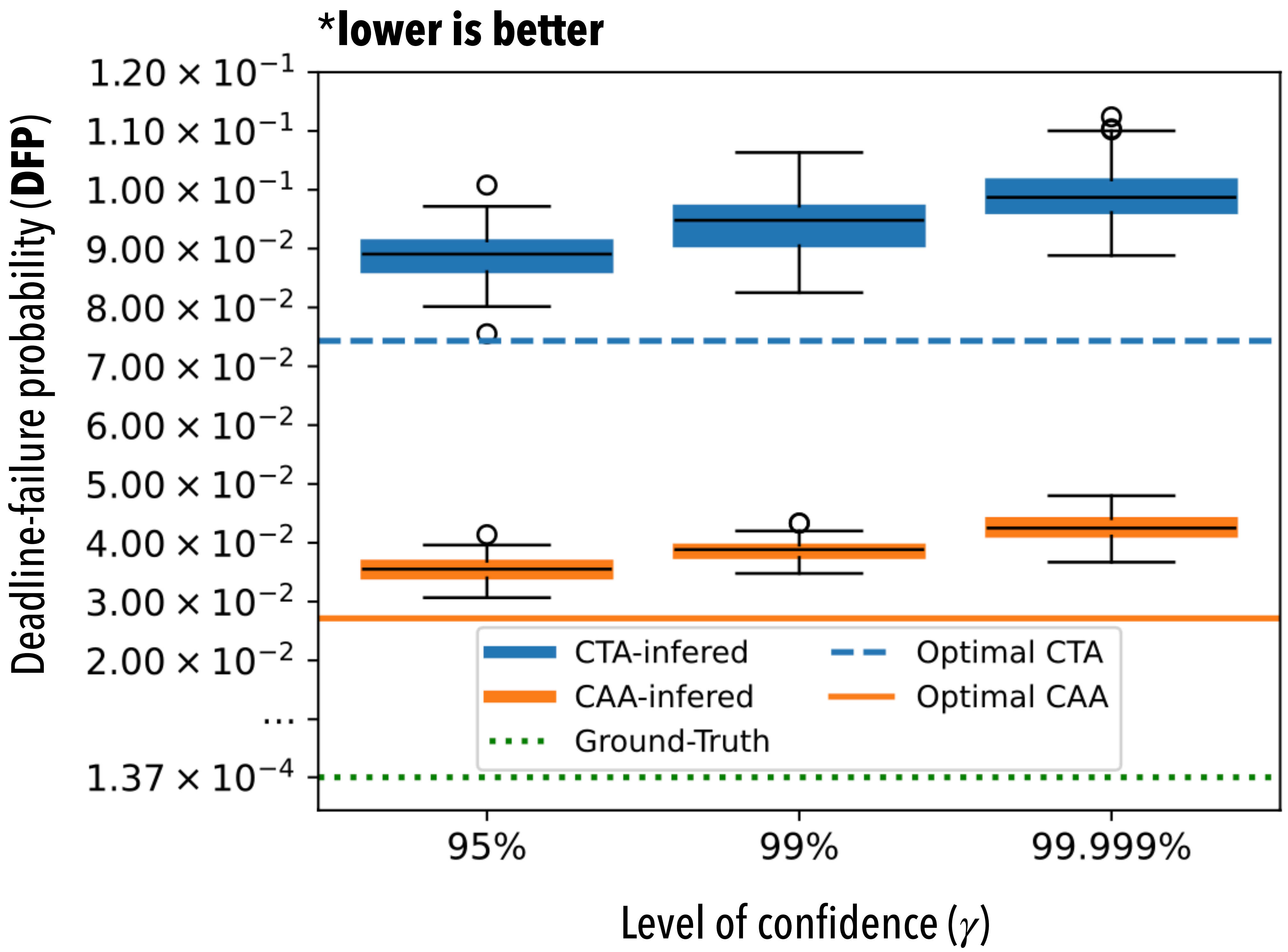


All DFP estimates **are sound**.

Both **CAA** and **CTA** tend towards their respective optimal DFP estimates as the initial sample size increases.

CASE STUDY

*Investigating the influence of the **level of confidence***



All DFP estimates **are sound**.

Increasing the level of **confidence** leads to **more conservative** results for both analyses.

SUMMARY

CONTRIBUTIONS

HOW TO MODEL DEPENDENCE?

*"How to handle … dependences between the execution times of (i) jobs of **the same task**, and (ii) jobs **of different tasks**?"*

HOW TO QUANTIFY DEPENDENCE?

*"The impact of these dependences may vary based on **how strong** they are."*

HOW TO USE THESE IN RT ANALYSIS?

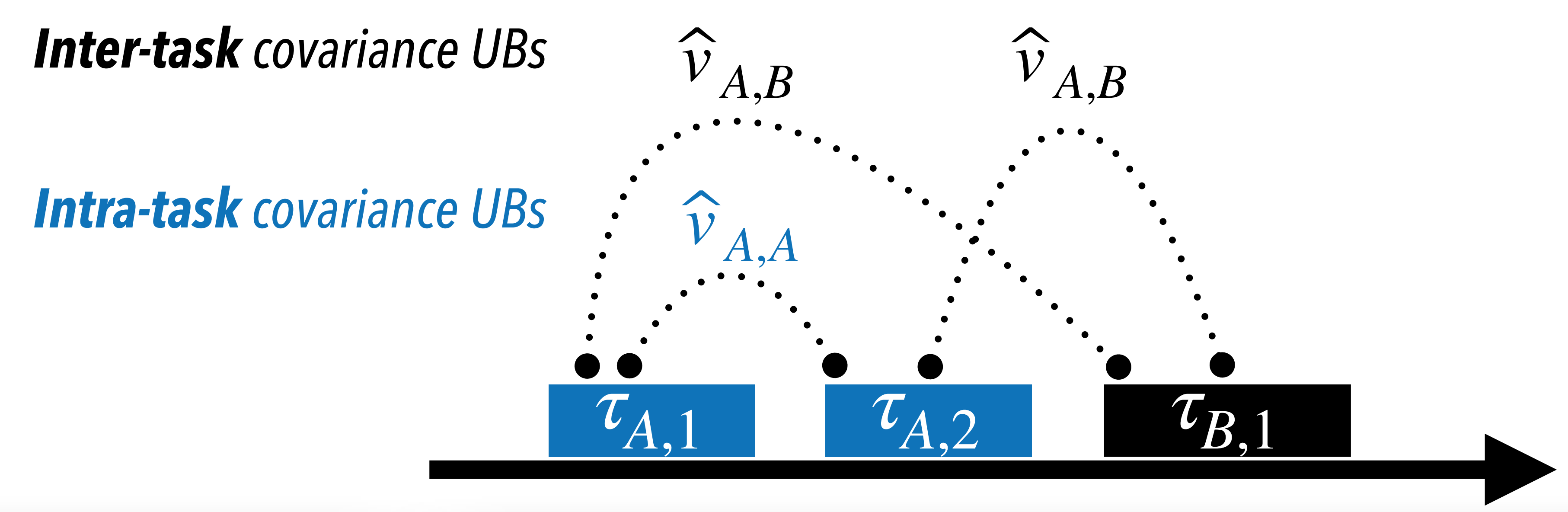
*"Analyses are needed that can **address dependencies**."*

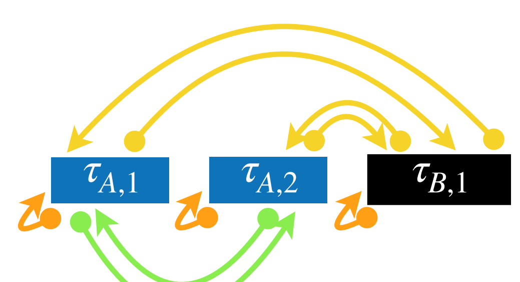
HOW TO STATISTICALLY INFER DEPENDENCE?

*"Appropriate **statistical studies** are needed to investigate the **types of dependences** and their impact on probabilistic schedulability analysis"*

BONUS: COMPUTATION EFFICIENCY!

*"ensuring that they [analyses] can be applied to **problems of a practical size**."*





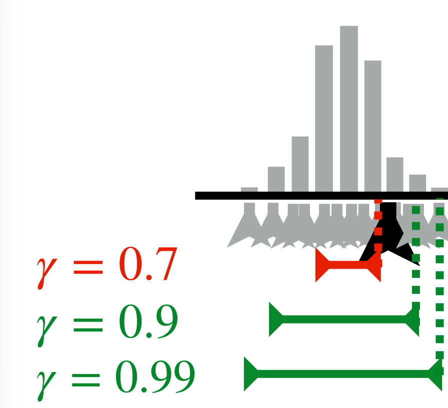
all covariance pairs

$$\mathbb{P}[R_2 \ge d] \le \frac{2\cdot \hat{s}_A^2 + \hat{s}_B^2 + 2\cdot \hat{v}_{A,A} + 4\cdot \hat{v}_{A,B} \;\; (\neq \widehat{\mathbb{V}}[R_B])}{\widehat{\mathbb{V}}[R_B] + \left(d - \widehat{\mathbb{E}}[R_B]\right)^2}$$

deadline-failure probability

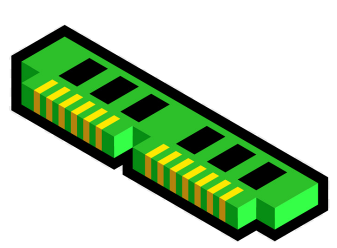
CAA (dominates CTA)

Nonparametric *bootstrapping*



Confidence-based *task parameters*

Closed-form solution



[2] *"A survey of probabilistic schedulability analysis techniques for real-time systems"* Davis and Cucu-Grosjean *LITES* (2019)

GOAL: DISTRIBUTION-AGNOSTIC STATISTICAL INFERENCE

We use Nonparametric Bootstrap

A random variable with an ***unknown*** *distribution*

1. *Draw an initial sample I of n independent observations of X*

2. *Generate b bootstrap samples (by randomly resampling with replacement from the initial sample I)*

3. *Compute the bootstrap statistic on each bootstrap sample B*

Output: *Bootstrap distribution of θ*