On Quantification of Accuracy Loss in Approximate Computing

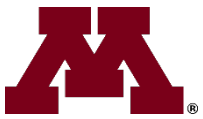
Ismail Akturk, **Karen Khatamifard**, Ulya R. Karpuzcu
{akturo02, khata006, ukarpuzc}@umn.edu

06/14/2015

UNIVERSITY OF MINNESOTA

Outline

- **Background**
- Pitfalls & Fallacies
- Practical Guidelines



Background: Approximate Computing

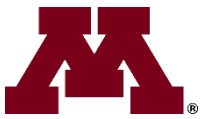
- Goal: trading off accuracy for energy efficiency
- Bag of tricks
  - Precision reduction
  - Computation perforation
  - Approximate consistency
  - Hardware simplification
  - Embrace of errors
- Robust **means** required to quantify accuracy loss under approximation

Background: Accuracy Metrics

- Numeric
  - Measure **deviation** from non-approximated value per output element
    - Metrics differ by the definition of **deviation**
  - Multi-dimensional output: **distortion**
    - Average **deviation** across all output elements
- Clustering, sorting, searching
  - Mismatch-based
- Multi-media
  - PSNR, SSIM

Outline

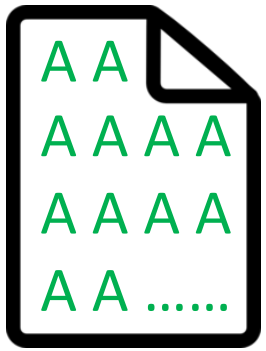
- Background
- **Pitfalls & Fallacies**
- Practical Guidelines



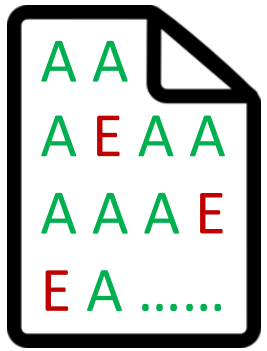
Pitfalls & Fallacies: Accuracy ≠ Validity

Compression
Accuracy Metric: file size

exact

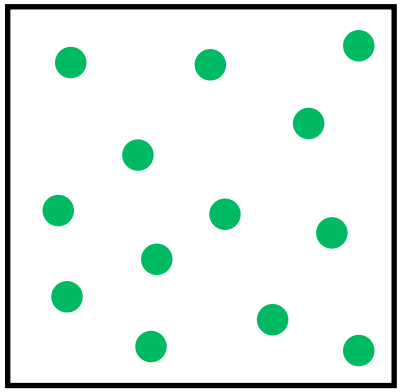


approximate

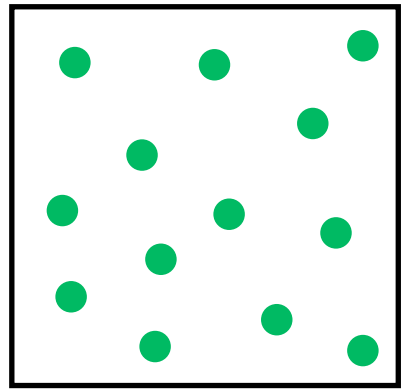


N-body simulation
Accuracy metric: average displacement

exact



approximate



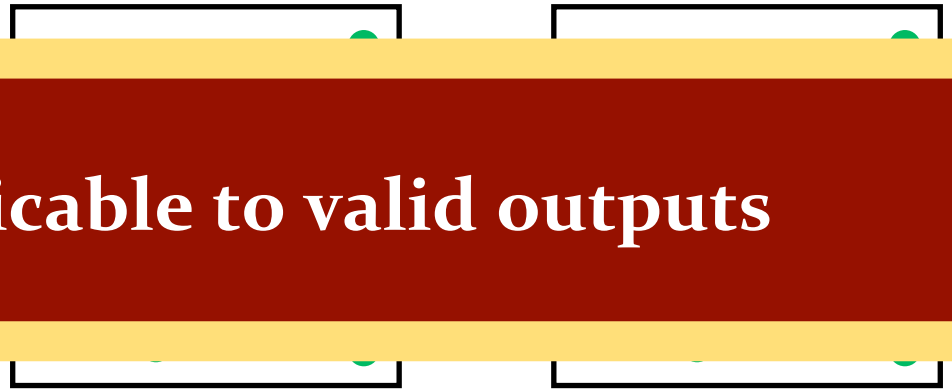
Pitfalls & Fallacies: Accuracy ≠ Validity

Compression
Accuracy Metric: file size

exact approximate

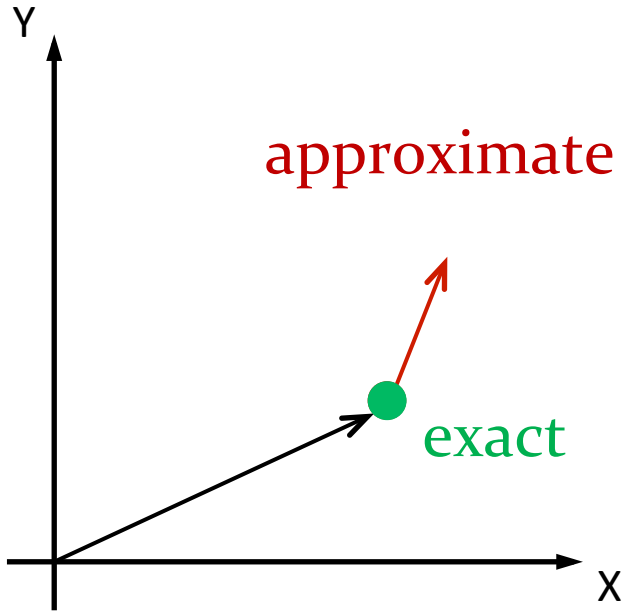
N-body simulation
Accuracy metric: average displacement

exact approximate



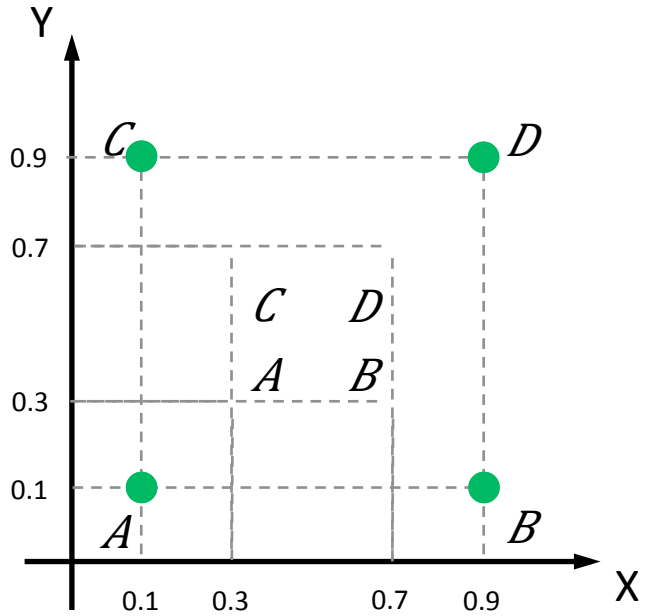
Accuracy metrics are only applicable to valid outputs

Pitfalls & Fallacies: Baseline for Normalization



$$\frac{\text{displacement from exact position}}{\text{distance from origin (exact position)}}$$

Pitfalls & Fallacies: Baseline for Normalization



$$\frac{\text{displacement from exact position}}{\text{distance from origin (exact position)}}$$

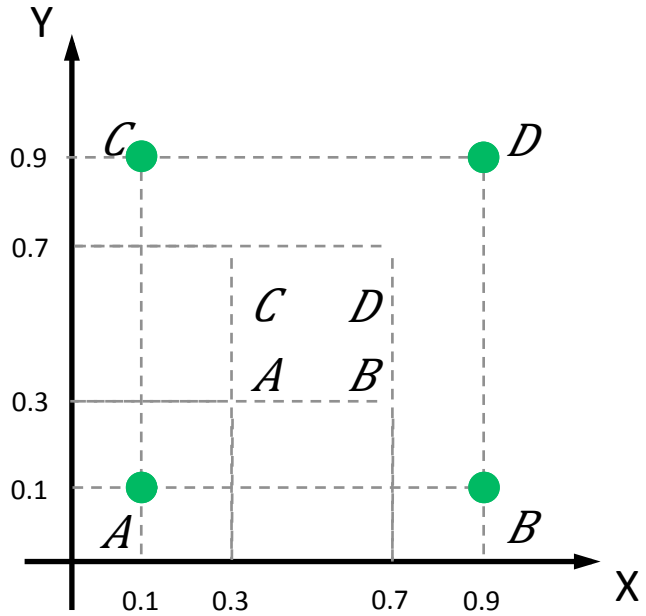
$$deviation\downarrow\{A\} = \sqrt{0.2\uparrow 2 + 0.2\uparrow 2} \; / \sqrt{0.1\uparrow 2 + 0.1\uparrow 2} = 200\%$$

$$deviation\downarrow\{B\} = \sqrt{0.2\uparrow 2 + 0.2\uparrow 2} \; / \sqrt{0.9\uparrow 2 + 0.1\uparrow 2} = 31.2\%$$

$$deviation\downarrow\{C\} = \sqrt{0.2\uparrow 2 + 0.2\uparrow 2} \; / \sqrt{0.9\uparrow 2 + 0.1\uparrow 2} = 31.2\%$$

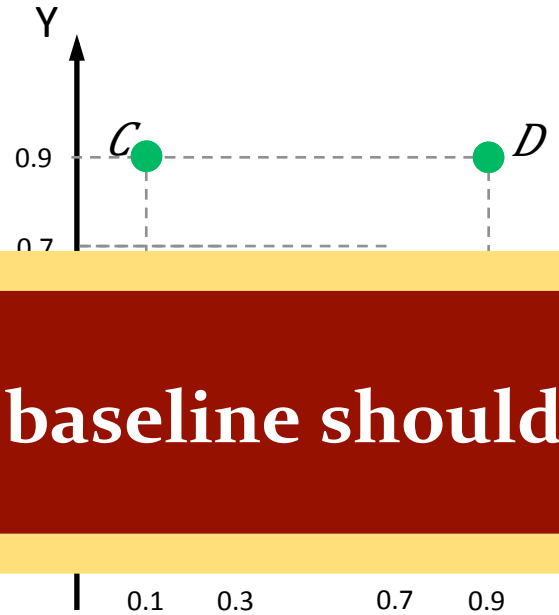
$$deviation\downarrow\{D\} = \sqrt{0.2\uparrow 2 + 0.2\uparrow 2} \; / \sqrt{0.9\uparrow 2 + 0.9\uparrow 2} = 22.3\%$$

Pitfalls & Fallacies: Baseline for Normalization



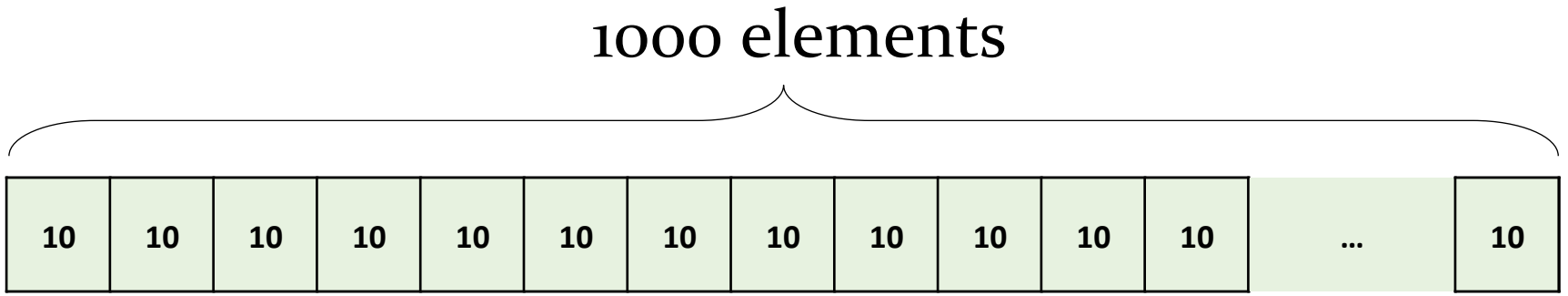
$$\frac{\text{displacement from exact position}}{\text{diagonal of square}}$$

Pitfalls & Fallacies: Baseline for Normalization

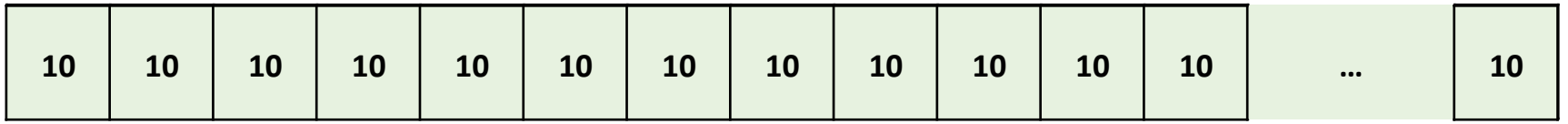


Normalization baseline should not introduce any bias on deviation

Pitfalls & Fallacies: Averaging Effects

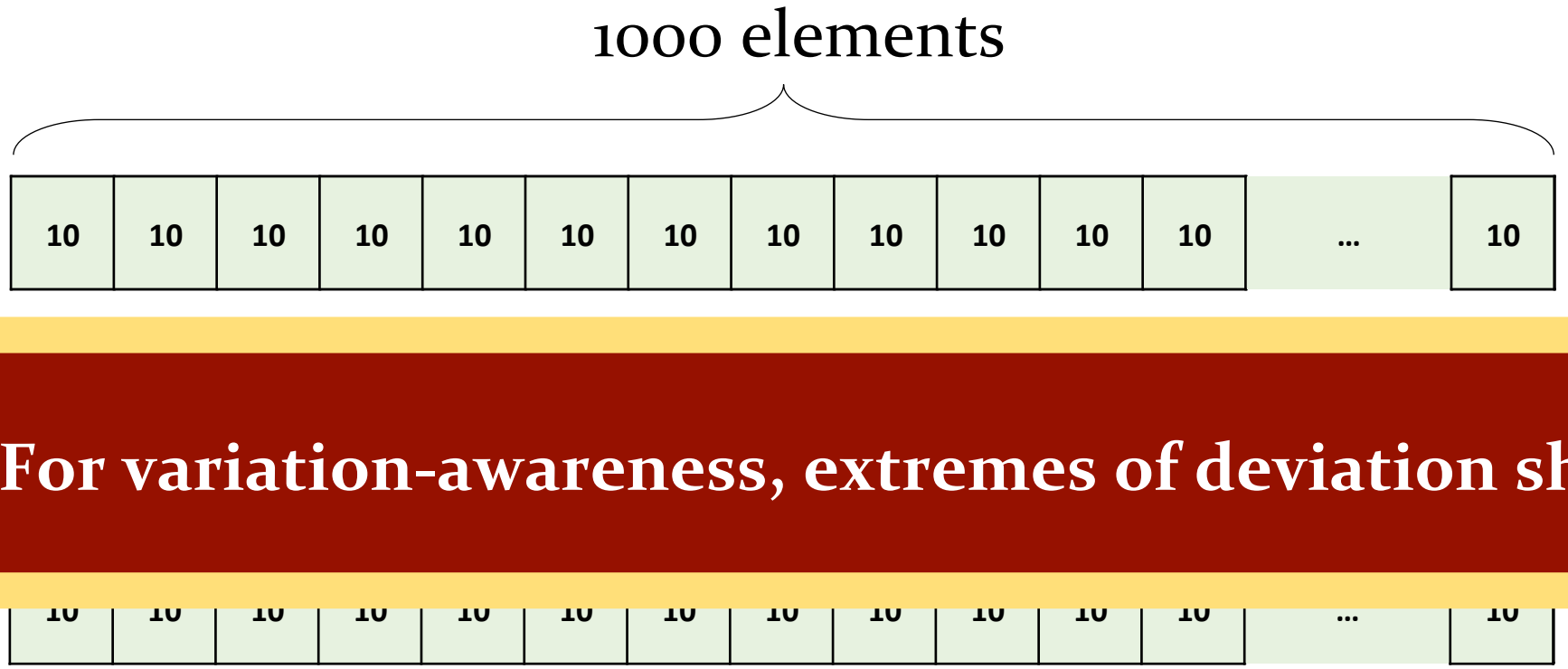


distortion=1%



distortion=1%

Pitfalls & Fallacies: Averaging Effects



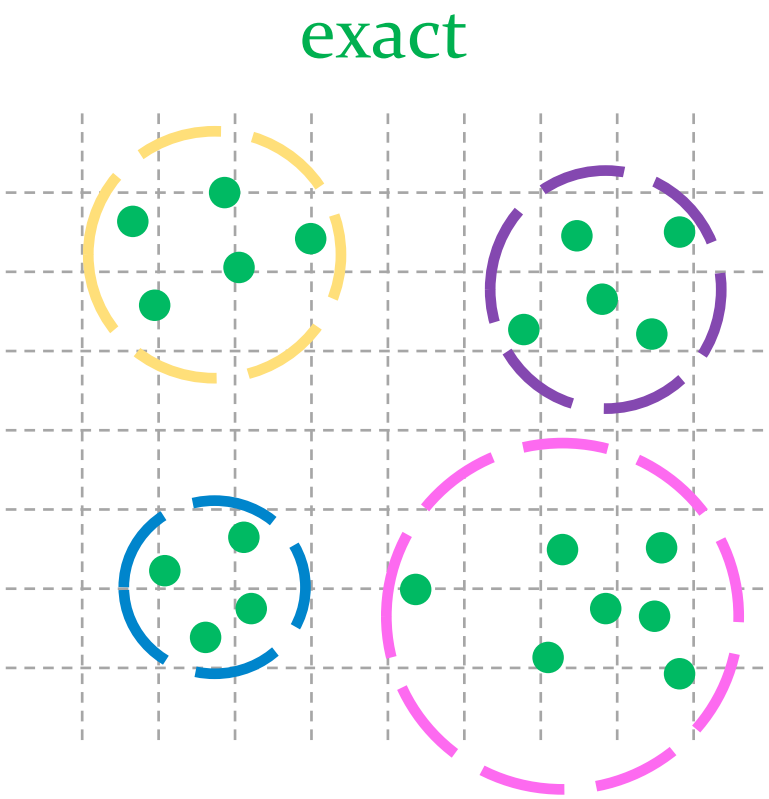
For variation-awareness, extremes of deviation should be reported

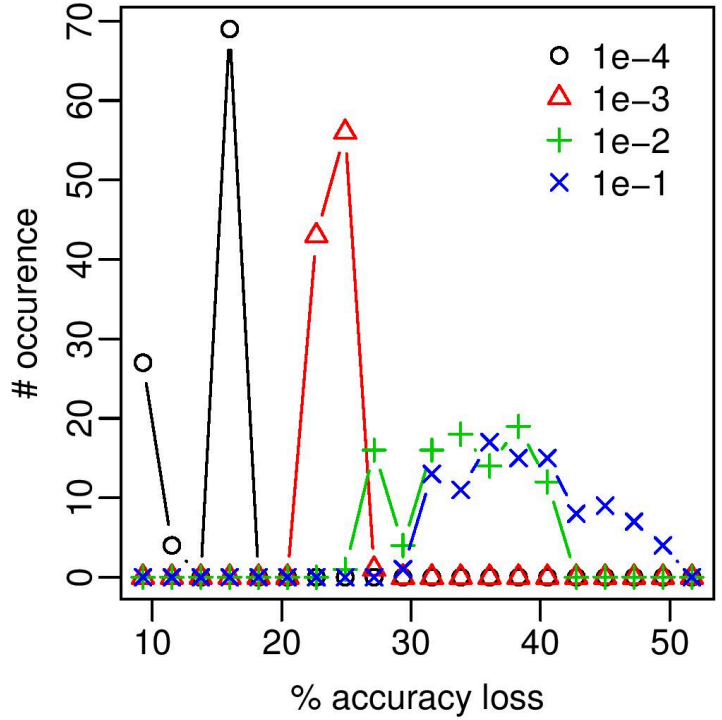
distortion=1%

Pitfalls & Fallacies: Non-determinism

1. Approximation induced
2. Application induced

Non-determinism: Approximation induced

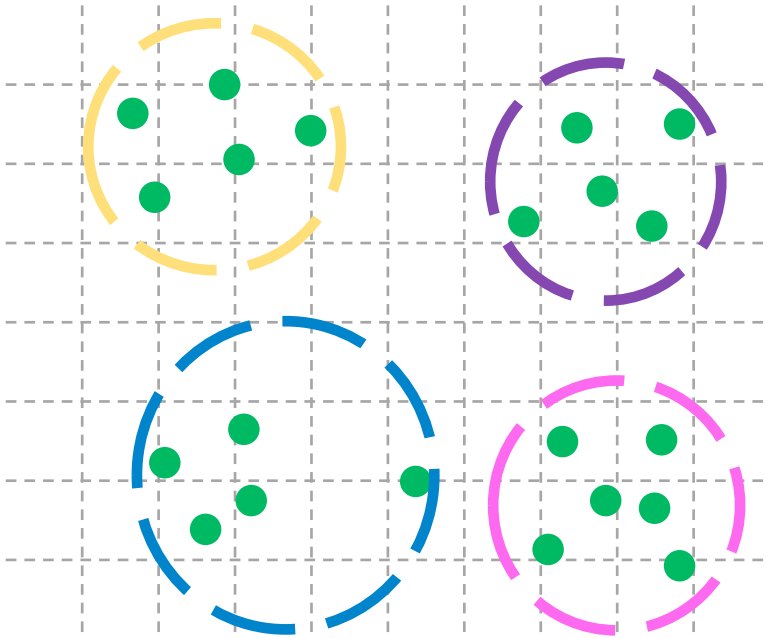


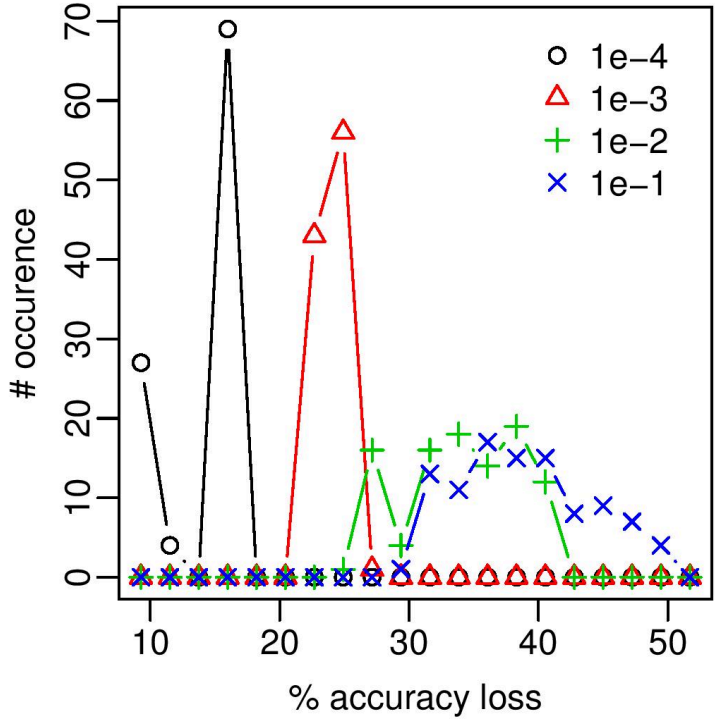


kmeans (clustering)

Non-determinism: Approximation induced

approximate



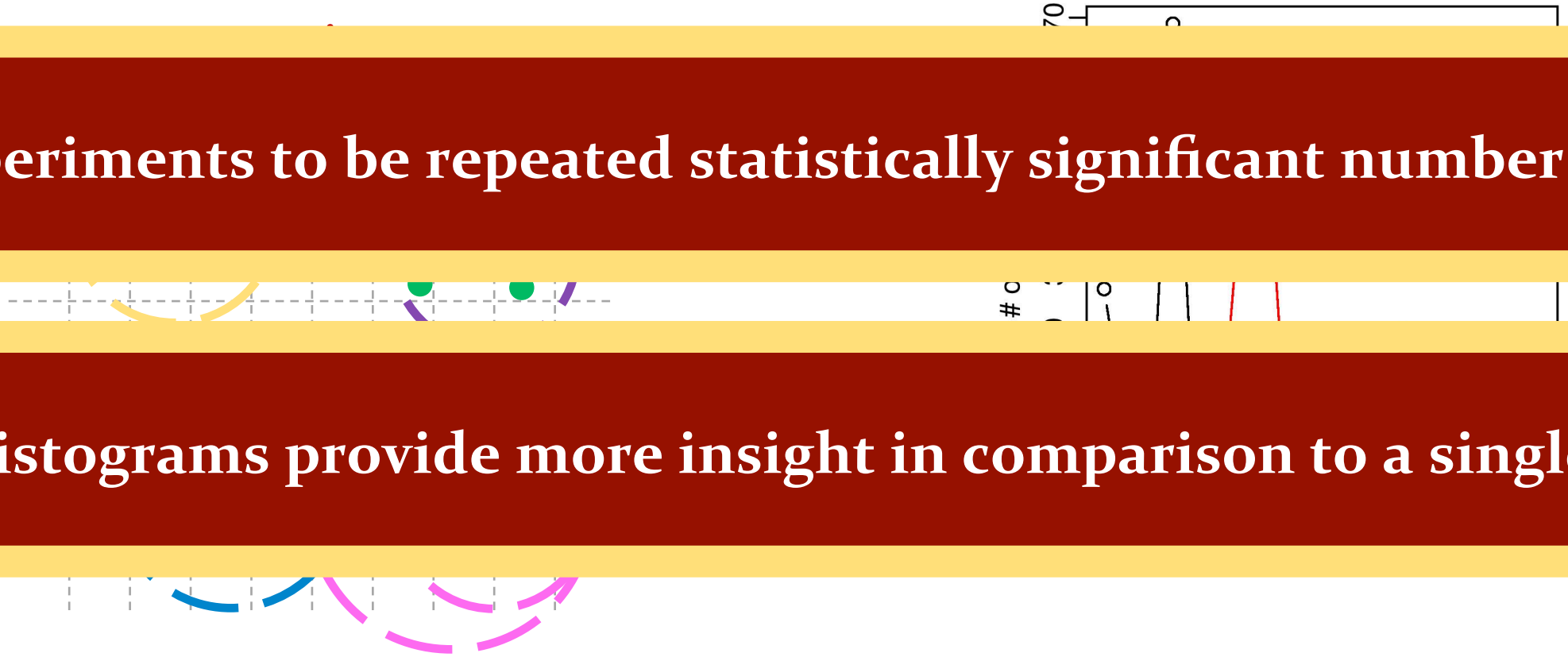


kmeans (clustering)

Non-determinism: Approximation induced

Experiments to be repeated statistically significant number of times

Histograms provide more insight in comparison to a single value



kmeans (clustering)

Non-determinism: Application induced

Scenario 1

F += 5e-8
F += 5e-8
F += 1.0

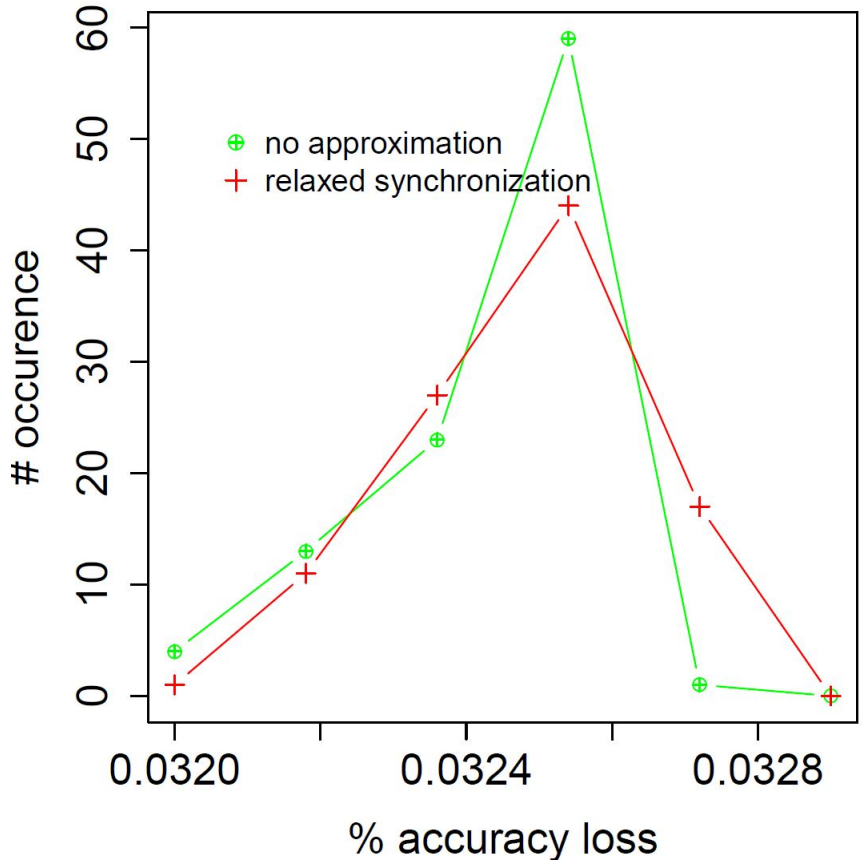
F: 1.0000001

Scenario 2

F += 1.0
F += 5e-8
F += 5e-8

F: 1.0000000

fluidanimate (n-body)



Non-determinism: Application induced

Scenario 1

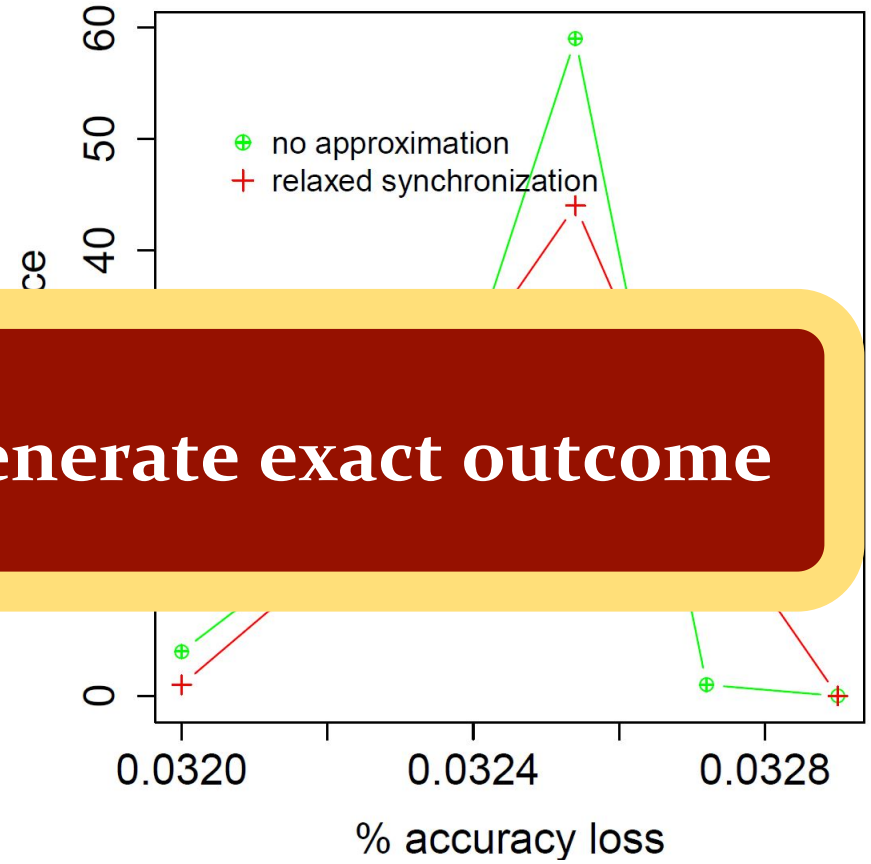
F += 5e-8
F += 5e-8

Scenario 2

F += 1.0
F += 5e-8

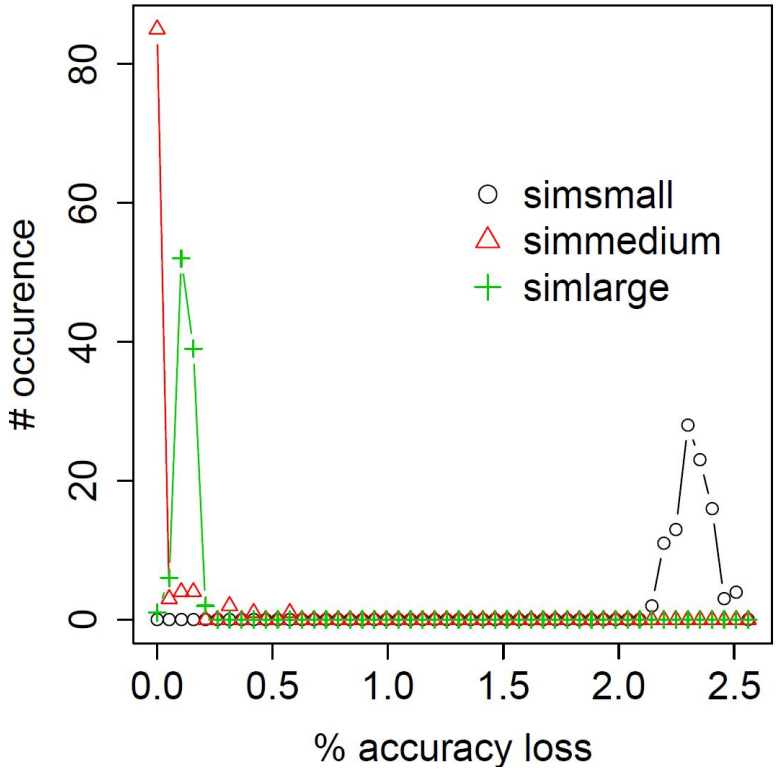
Only a single threaded run is safe to generate exact outcome

fluidanimate (n-body)

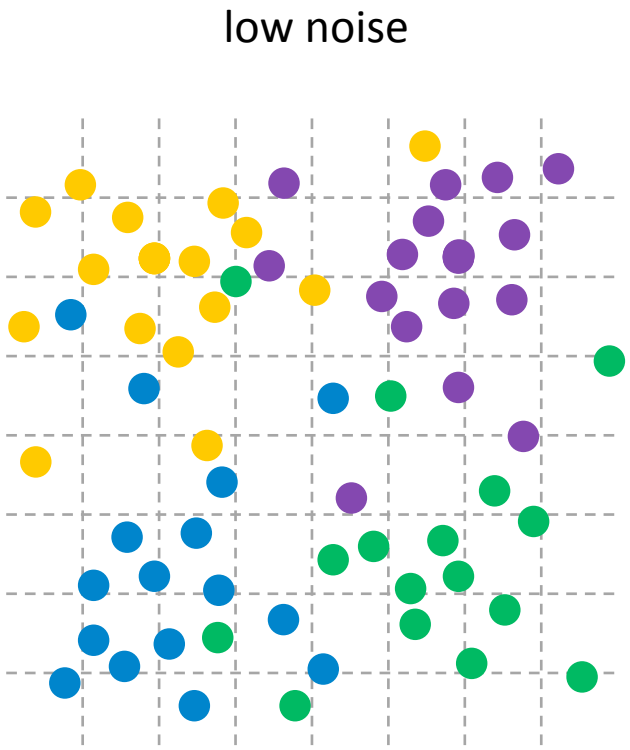


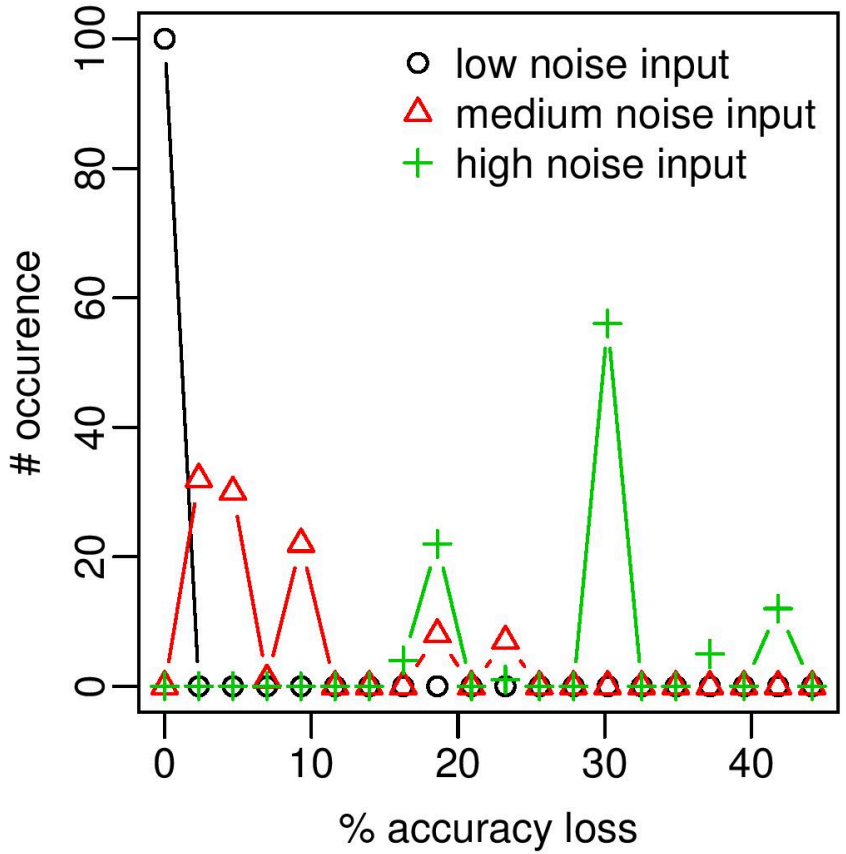
Impact of Input Data: Size

fluidanimate (n-body)



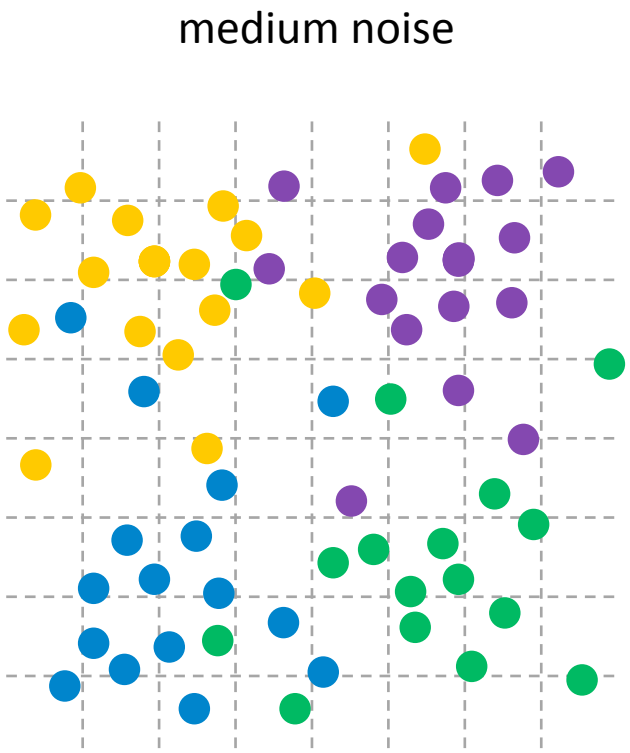
Impact of Input Data: Value

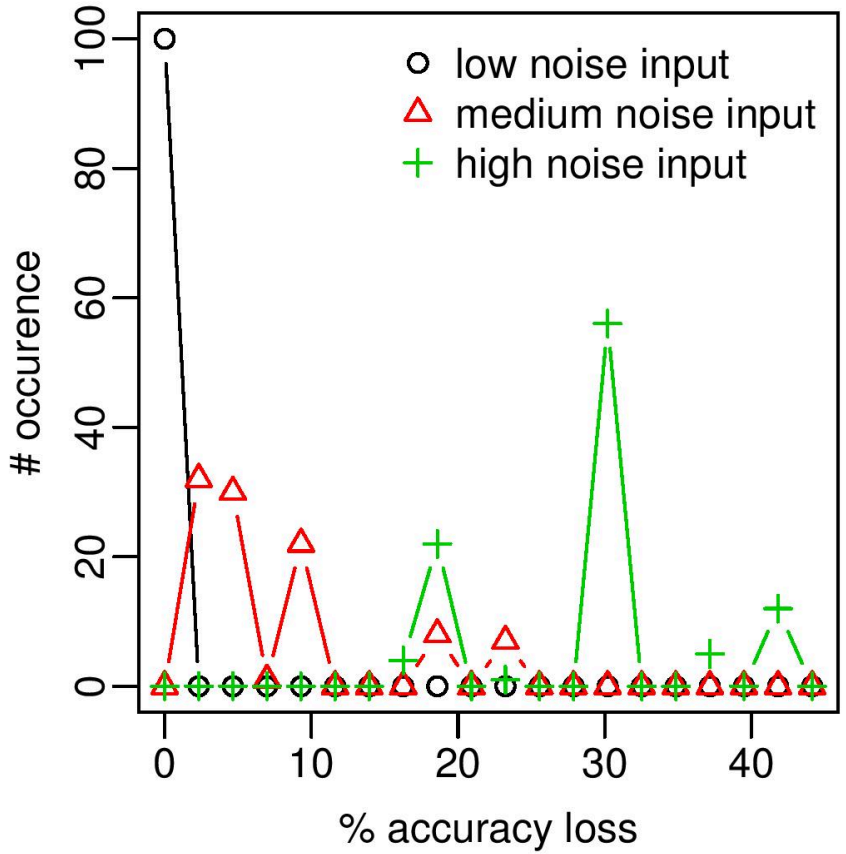




kmeans (clustering)

Impact of Input Data: Value

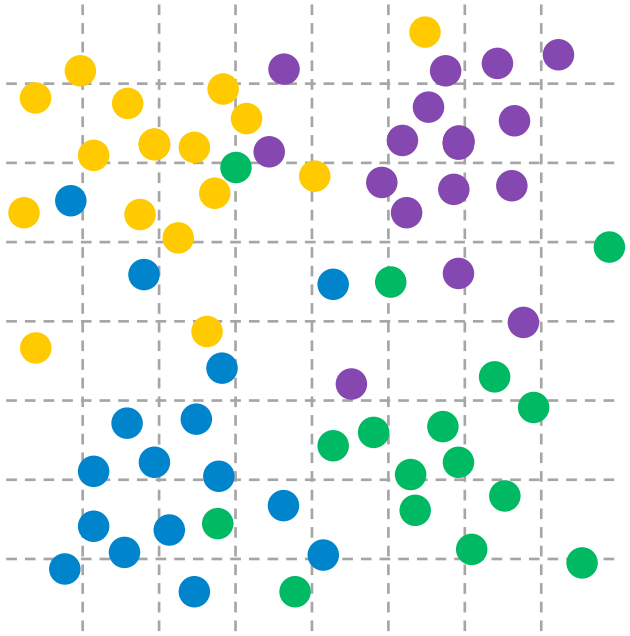


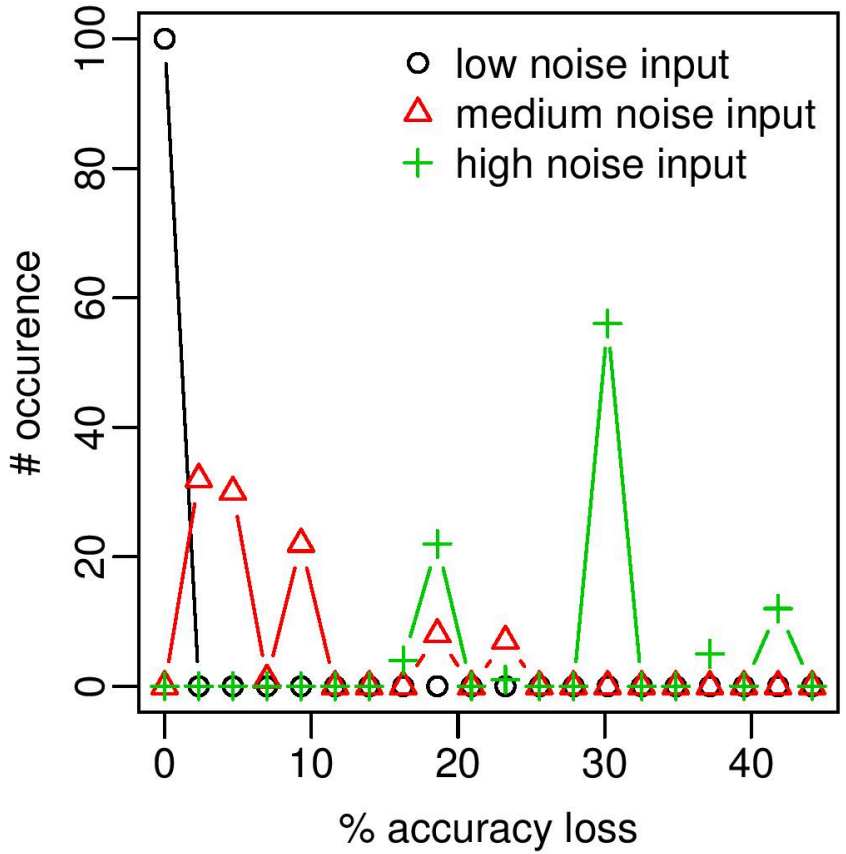


kmeans (clustering)

Impact of Input Data: Value

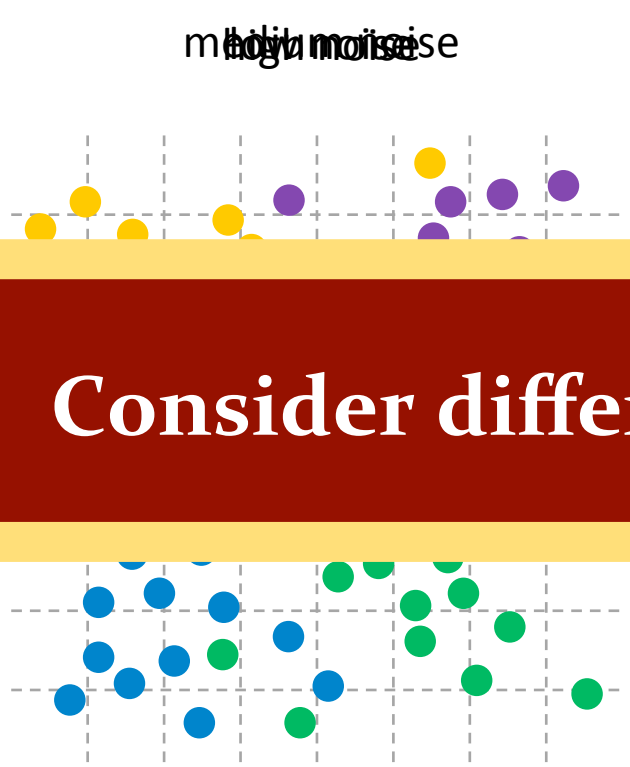
high noise

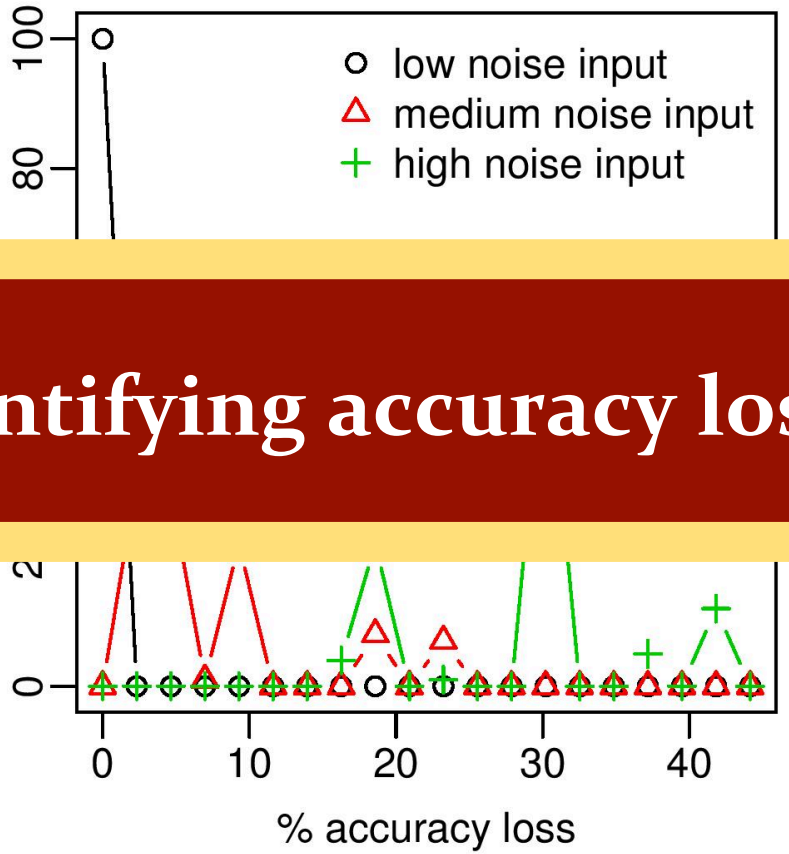




kmeans (clustering)

Impact of Input Data: Value



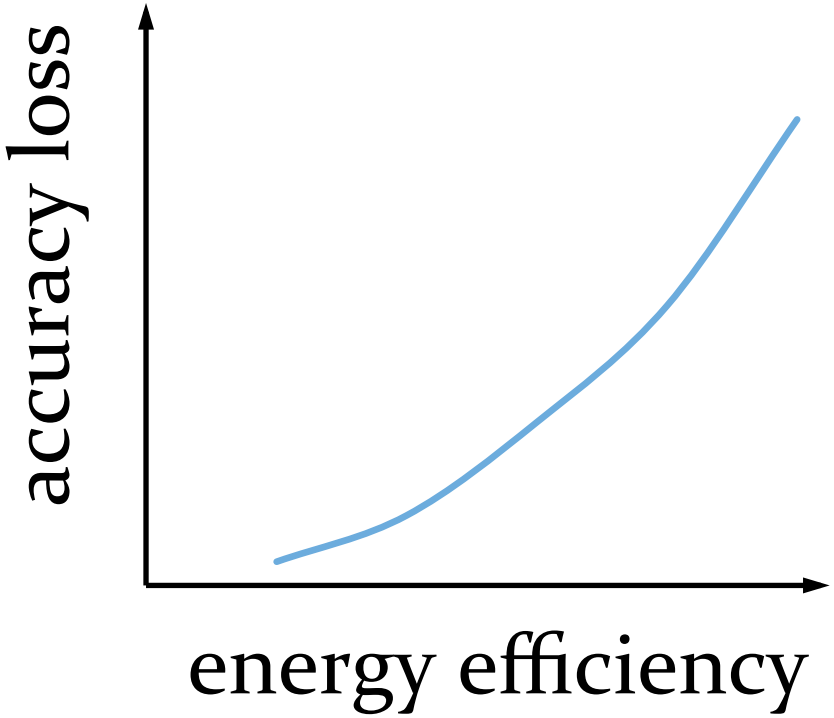


Consider different inputs in quantifying accuracy loss

kmeans (clustering)

Pitfalls & Fallacies: Accuracy ≠ Acceptability

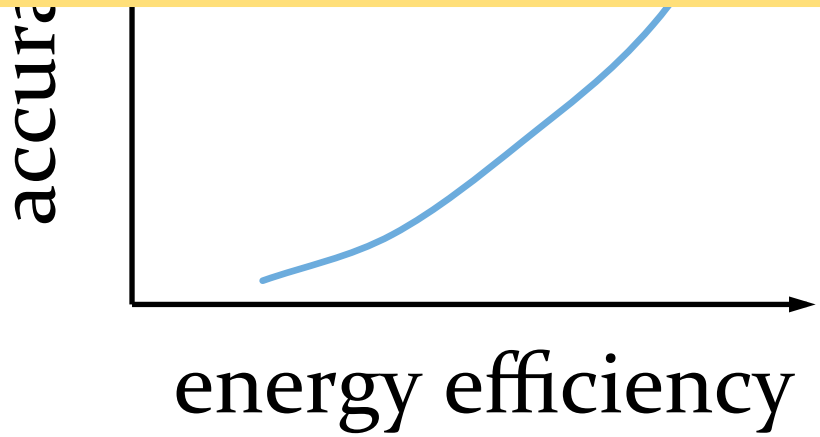
- Acceptability depends on the context
- Reporting trade-off spaces, pareto fronts, … is more meaningful



Pitfalls & Fallacies: Accuracy ≠ Acceptability

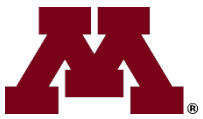
- Acceptability depends on the context
- Reporting trade-off spaces, pareto fronts, … is more meaningful

Context-oblivious studies should report trade-off spaces



Outline

- Background
- Pitfalls & Fallacies
- **Practical Guidelines**



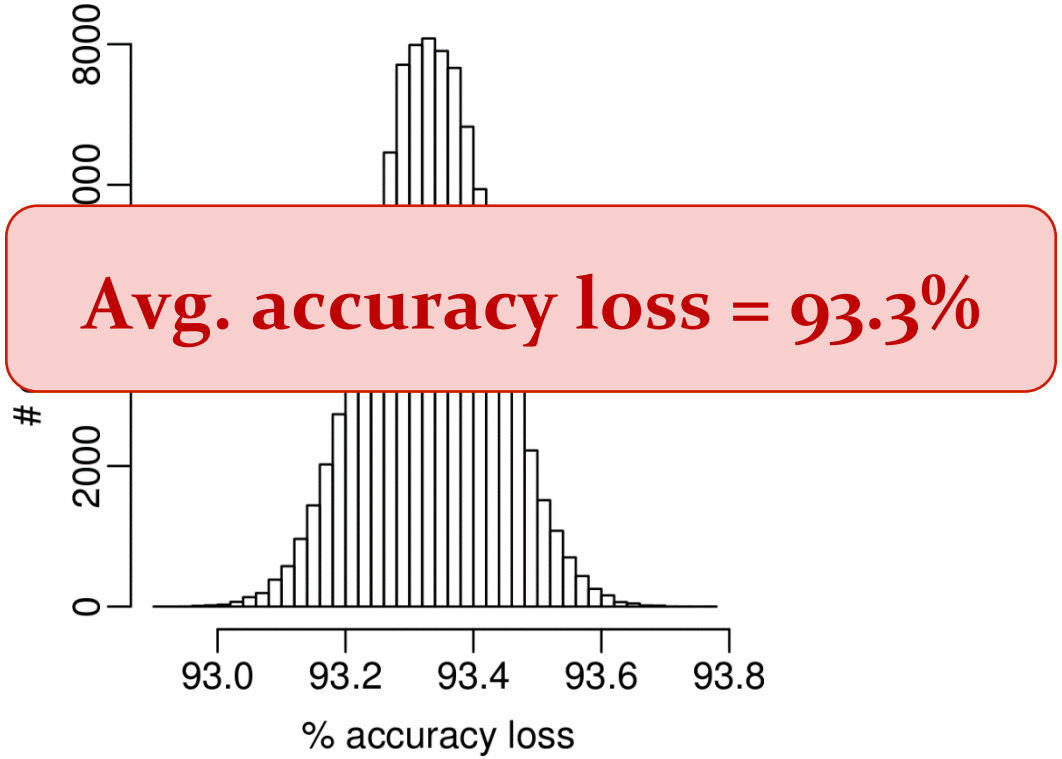
Output Randomization

- How “acceptable” is 10% of accuracy loss?
- Compare to worst case?
  - Randomize application outputs
  - Calculate the accuracy range of totally randomized
  - Compare to accuracy loss under approximation

Output Randomization

kmeans

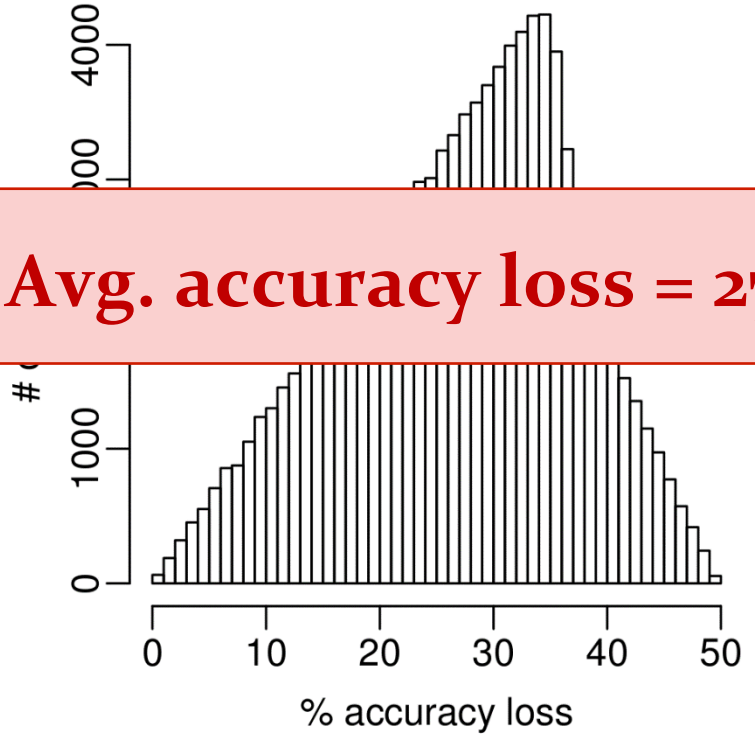
Output: 65K assignments to 15 clusters



Avg. accuracy loss = 93.3%

particlefilter

Output: particle position (X,Y)

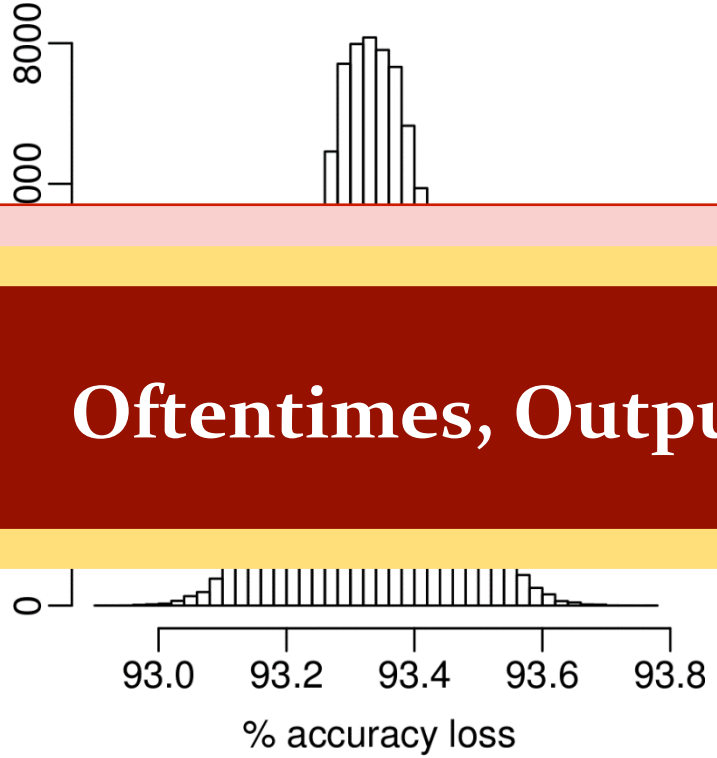


Avg. accuracy loss = 27.1%

Output Randomization

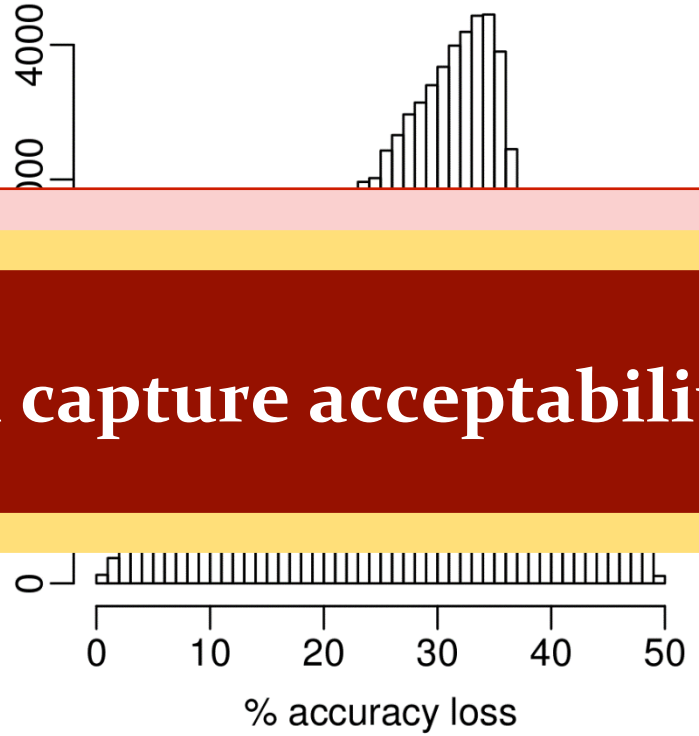
kmeans

Output: 65K assignments to 15 clusters



particlefilter

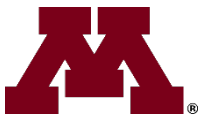
Output: particle position (X,Y)



Oftentimes, Output Randomization can capture acceptability

Putting It All Together

- Metric selection
- Design of experiments
- Reading the outcome of experiments



Metric Selection

Class	Output Data Type	Accuracy Metric	Domain
I	Numeric: scalar	Relative deviation in output value	Optimization, Compression
II	Numeric: multi-dimensional	Relative displacement Avg. Noise to Peak Signal (ANPS)	N-body Simulation, Computer Vision Linear Algebra, Histogram
III	Compound	Relative mismatch Positional error	Clustering, Similarity, Search Sort
IV	Multi-media	SSIM (Structural Similarity Index)	Image Processing, Video Encoding

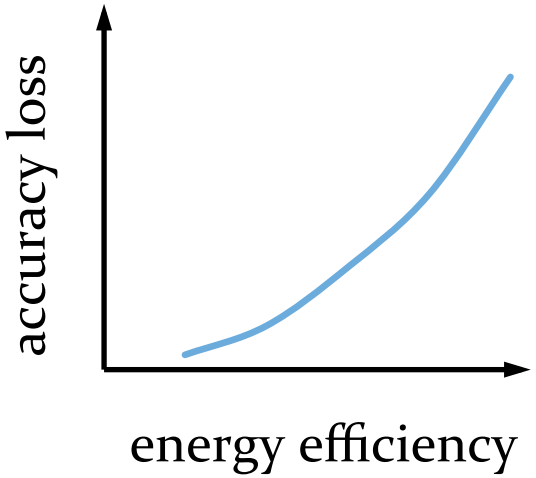
altai.ece.umn.edu/accurax

Design of Experiments

- Tame the sources of non-determinism
- Report accuracy loss in the form of histograms or trade-off spaces.
- Invalid outcome: reflect the overhead of safety nets to the trade-off space
- Metrics at application phase boundaries?
  - Only safe for selective approximation

Reading the Outcome

- Acceptance depends on the context
  - Context determines which points of the trade-off space are feasible
- Oftentimes, *Output Randomization* can capture acceptability
- Do not average across different inputs
- Do not average across different applications



On Quantification of Accuracy Loss in Approximate Computing

Ismail Akturk, **Karen Khatamifard**, Ulya R. Karpuzcu
{akturo02, khata006, ukarpuzc}@umn.edu

altai.ece.umn.edu/accurax

06/14/2015

ALTAI

UNIVERSITY OF MINNESOTA