TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Mechanical Science and Engineering| Institute of Fluid Mechanics | Chair of Turbomachinery and Flight Propulsion

Meta-model-based Quality Assessment of Sample Estimates

Andriy Prots, Lars Högner, Matthias Voigt, Ronald Mailach

Andriy.Prots@tu-dresden.de

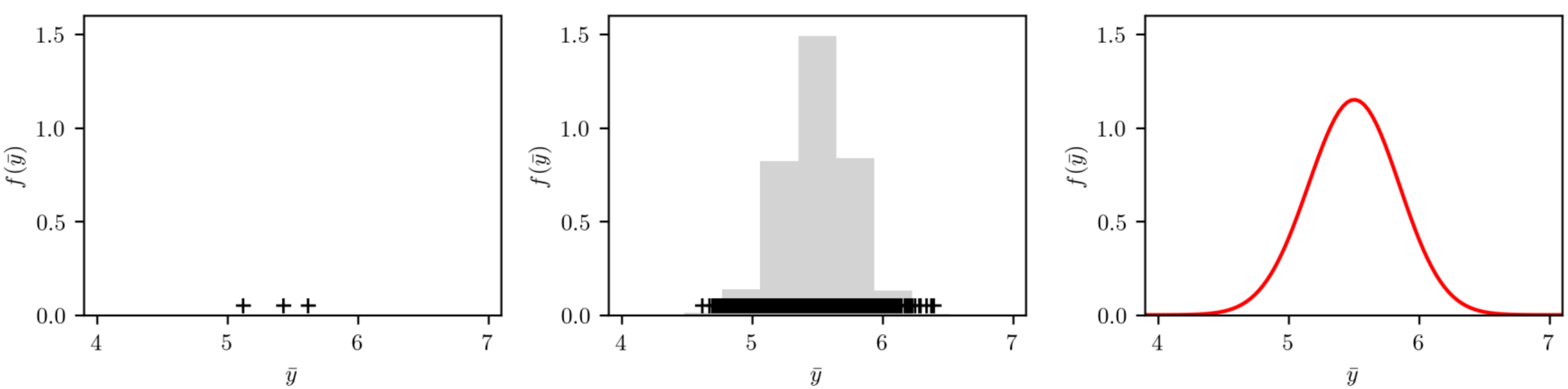
DRESDEN concept
Exzellenz aus Wissenschaft und Kultur

Dresden, October 10, 2019

- Probabilistic Methods are gaining in importance

- Monte Carlo simulation (MCS):
  - Sample of representative realizations (e. g. with Simple Random Sampling (SRS))
  - Calculation of the sample with deterministic methods
  - Statistical evaluation of the results (Mean, Variance, Quantile value, Correlation Coefficient)

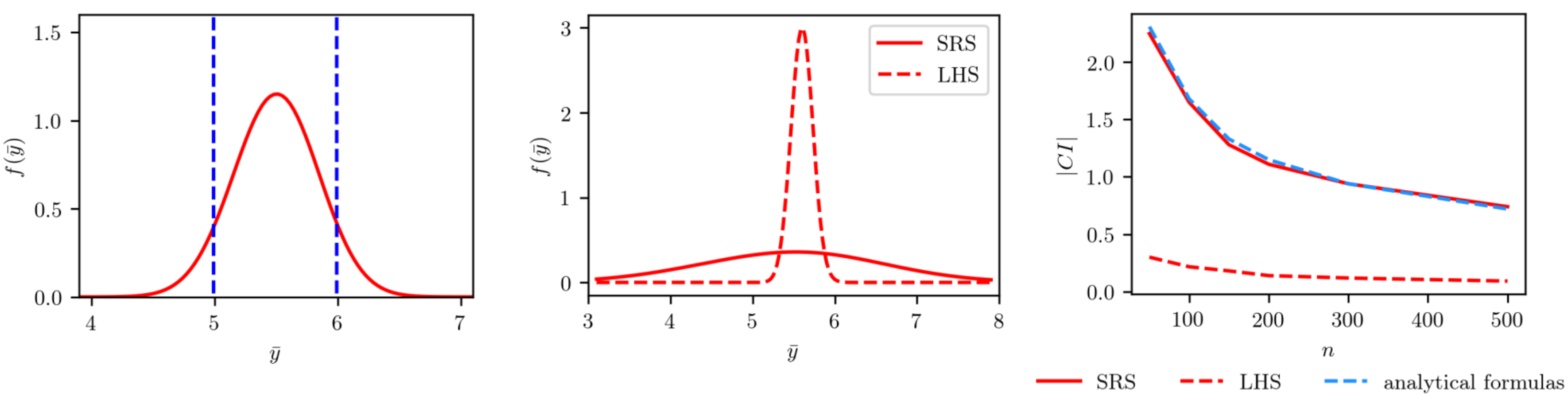
- Sample is generated randomly → Result of a MCS is also random



- Value to describe the variance of the result: Confidence Interval (CI)
- Real CI: repetitions of MCS required
- Determinable by analytical formulas or bootstrapping

- Latin Hypercube Sampling is used to reduce variance of result

- Problem 1: Known methods cannot describe the variance reduction
- Problem 2: Factor of variance reduction is unknown

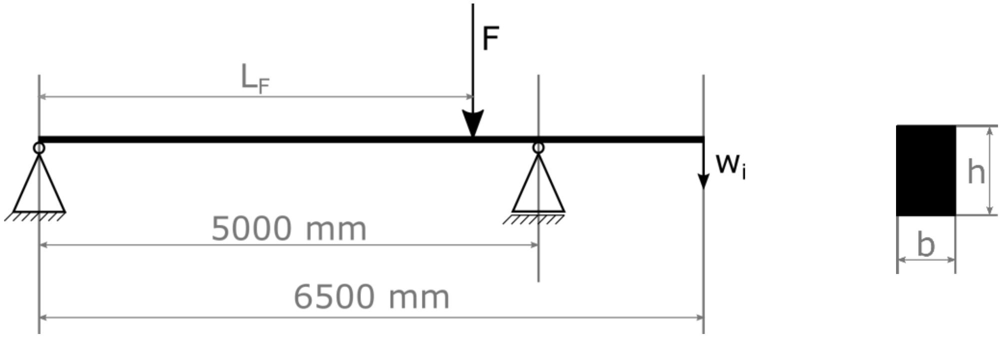


- Goal:
  - Predicting the size of CI more precisely, when using LHS
  - Reduce required sample size to reach a target size of CI

- Idea:
  - Approximate system behavior with meta model (MM)
  - Simulate MCS with help of MM
  - Evaluate virtual MCS

- Assumption:
  - Good description of system behavior with MM
  - Real MCS can be approximated with virtual MCS

- Beam system:



- Input values:
  - Height h (uniform in $[95\ mm; 105\ mm]$)
  - Width b (uniform in $[45\ mm, 55\ mm]$)
  - Young's module E (normal with $\mu = 210\,000\ GPa$ und $\sigma = 10\,000\ GPa$)
  - Force F (normal with $\mu = 2\,500\ N$ und $\sigma = 300\ N$)
  - Position of the Force L_F (uniform in $[0\ mm,\ 6\,500\ mm]$)

- Output value:
  - Deflection w_i at the end (calculated by beam theory)

- MCS: description of system behavior based on a random sample

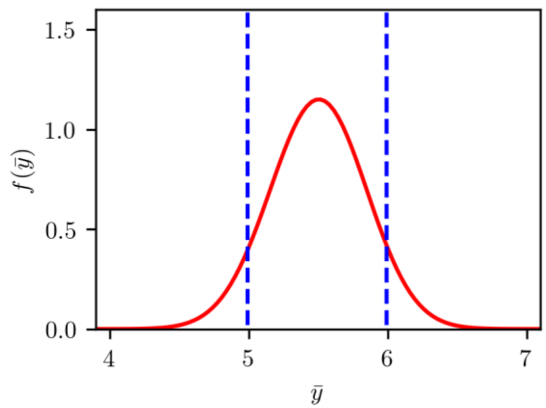
- Required steps:
  - Generate samples
  - Evaluate sample with deterministic models
  - Evaluate the results statistically

- Possible result values:
  - Mean
  - Variance / Standard deviation
  - Quantile values
  - Correlation coefficient

- Confidence Intervals (CI): describes the variance of a result value of MCS
- Significance level α: CI contains the real value in $(1-\alpha)*100\%$ of the cases

- Smaller CI → result can be more trusted

- Possibilities of determining the CI
  - Repetition of MCS (not practical)
  - Analytical formulas
    - E. g. for mean: $\bar{y} \pm z_{1-\frac{\alpha}{2}}\, \sigma/\sqrt{n}$
  - Bootstrapping

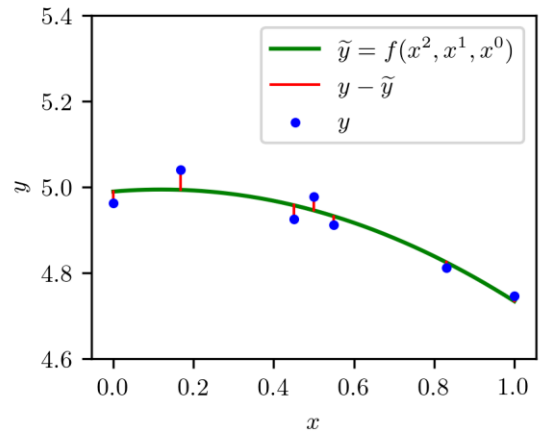


- Assumption: realization within sample are independent
- Reduction of variance from LHS is not considered!

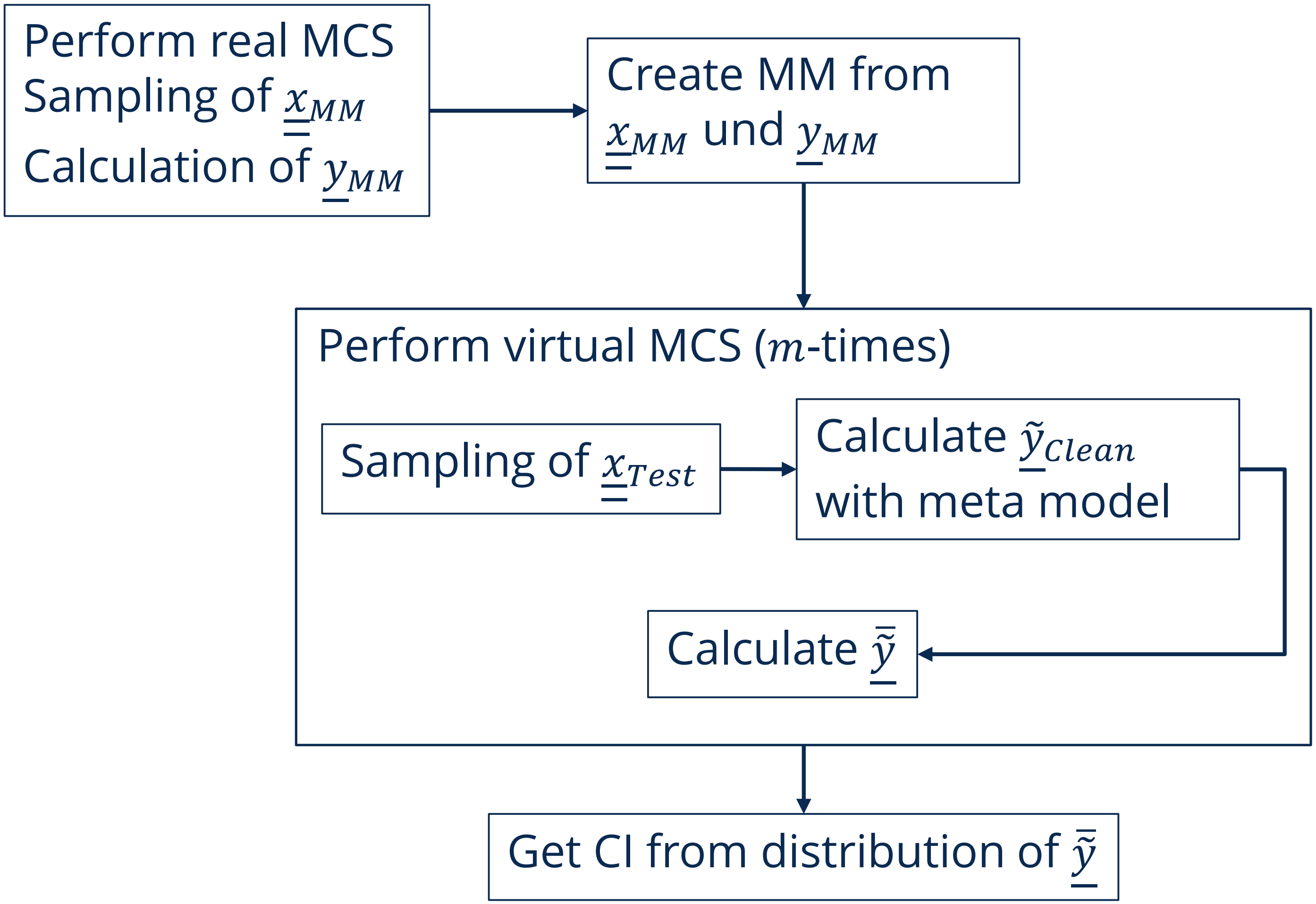
- Meta model (MM): description of system behavior with simple approach

- E. g. polynomial meta model:
  - $\tilde{y} = c_0 + c_1\, x_1 + c_2\, x_2 + c_3\, x_1^2 + c_4\, x_2^2$

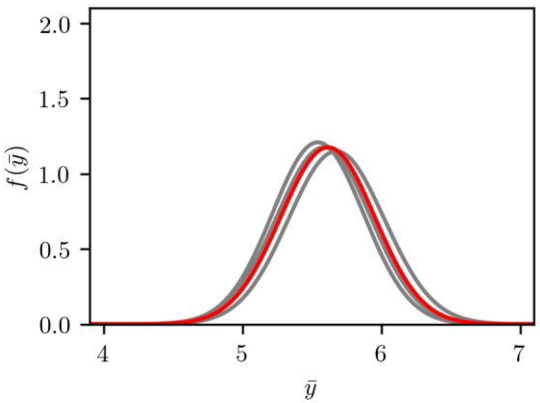
- Determination of coefficients: least square approach



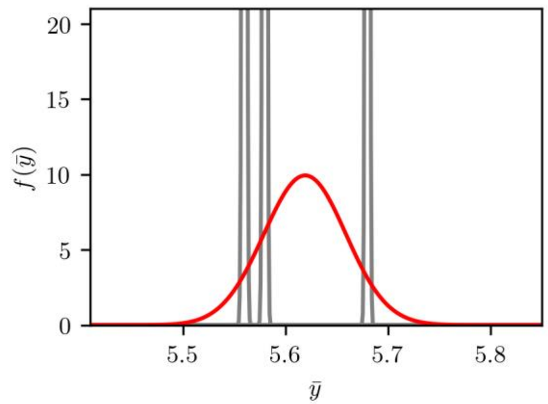
- Quality Assessment:
  - Coefficient of Determination $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{\sum_{i=i}^{n}(y_i - \bar{y})}$
  - Cross validation


Perform real MCS
Sampling of $\underline{x}_{MM}$
Calculation of $\underline{y}_{MM}$
Create MM from $\underline{x}_{MM}$ und $\underline{y}_{MM}$
Perform virtual MCS (m-times)
Sampling of $\underline{x}_{Test}$
Calculate $\underline{\tilde{y}}_{Clean}$ with meta model
Calculate $\underline{\bar{\tilde{y}}}$
Get CI from distribution of $\underline{\bar{\tilde{y}}}$

- SRS:



LHS:



- Works only for SRS
- Variance of mean heavily underestimated for LHS

- Reason: Error of the meta model $\varepsilon = y - \tilde{y}$

- Impact of $\varepsilon = y - \tilde{y}$:

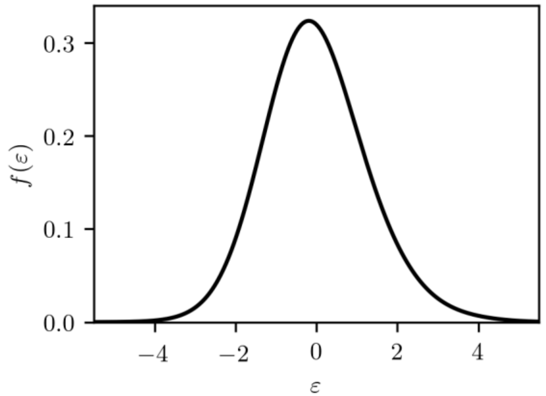
$$y = \tilde{y} + \varepsilon$$
$$\bar{y} = \bar{\tilde{y}} + \bar{\varepsilon}$$
$$Var(\bar{y}) = Var(\bar{\tilde{y}} + \bar{\varepsilon}) = Var(\bar{\tilde{y}}) + 2Cov(\bar{\tilde{y}}; \bar{\varepsilon}) + Var(\bar{\varepsilon})$$

- Assumption: Independence between $\tilde{y}$ and ε => $Cov(\bar{\tilde{y}}; \bar{\varepsilon}) = 0$

$$Var(\bar{y}) = Var(\bar{\tilde{y}}) + Var(\bar{\varepsilon})$$

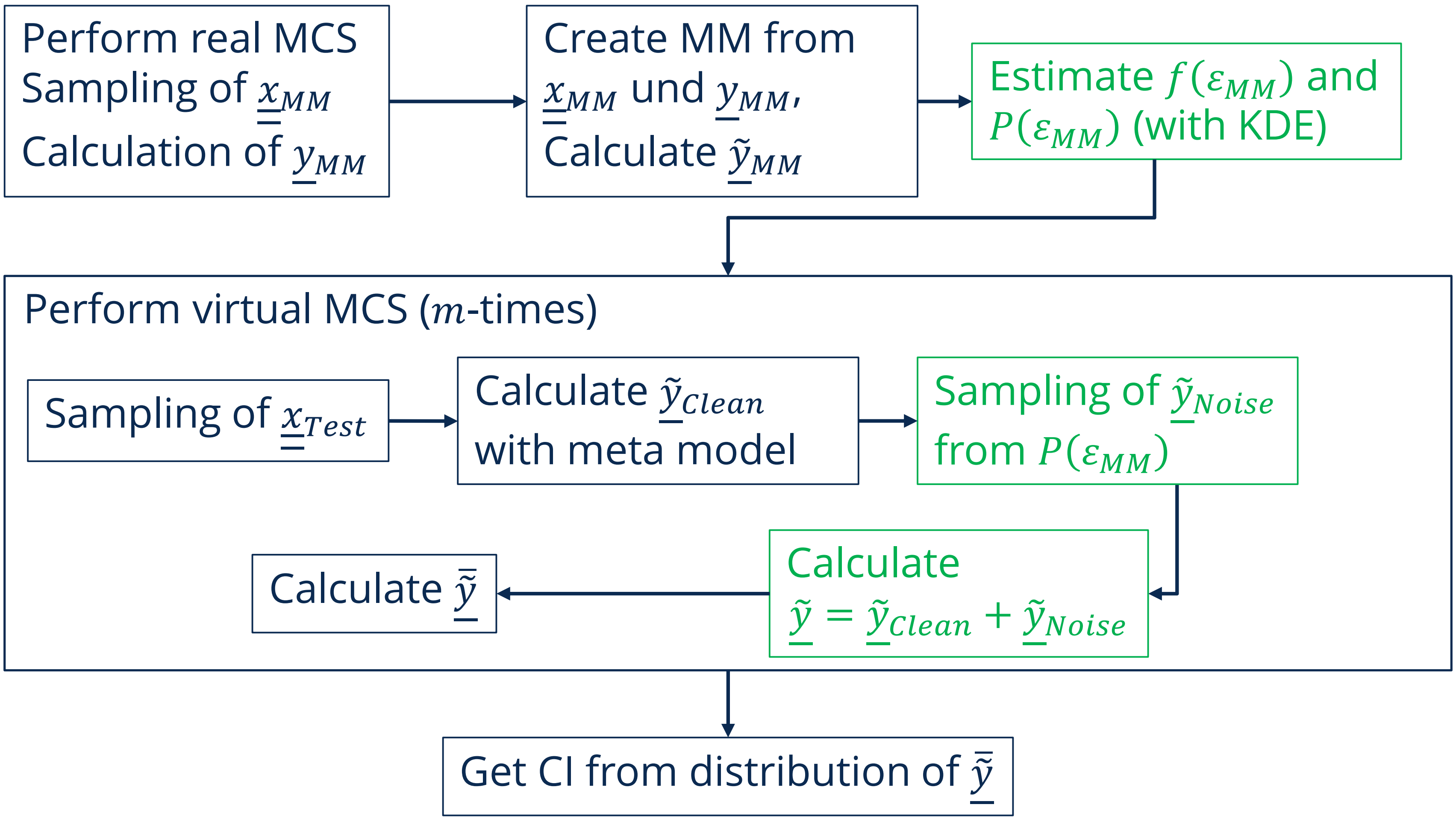
- Problem: for LHS often $Var(\bar{\varepsilon}) \gg Var(\bar{\tilde{y}})$

- Solution: generate additional sample for ε
- Since independence between $\tilde{y}$ and ε is assumed: only $\mathrm{P}(\varepsilon)$ is required

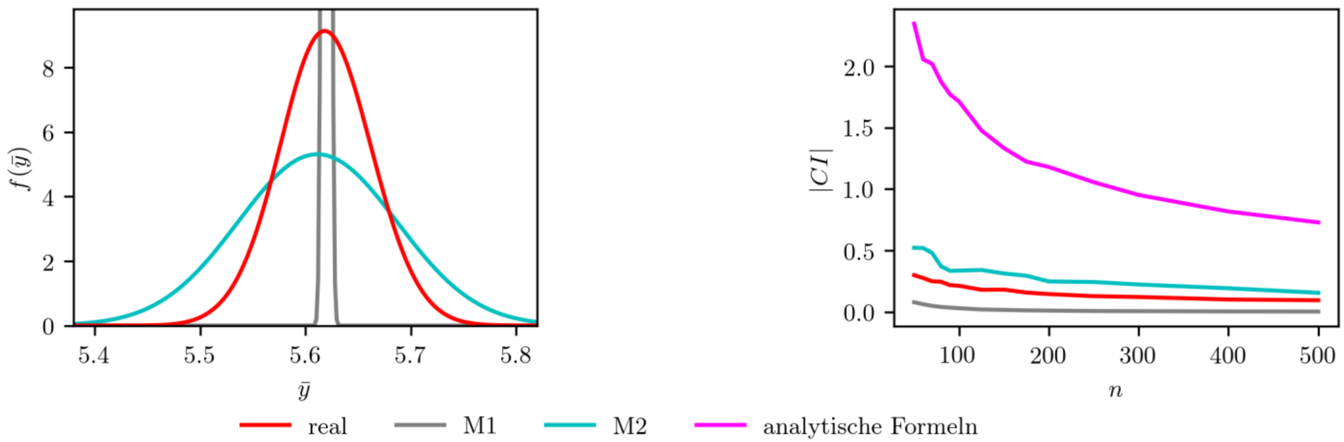


- Approximation of $f(\varepsilon)$ with kernel density estimation (KDE), Approximation of $\mathrm{P}(\varepsilon)$ with numerical integration

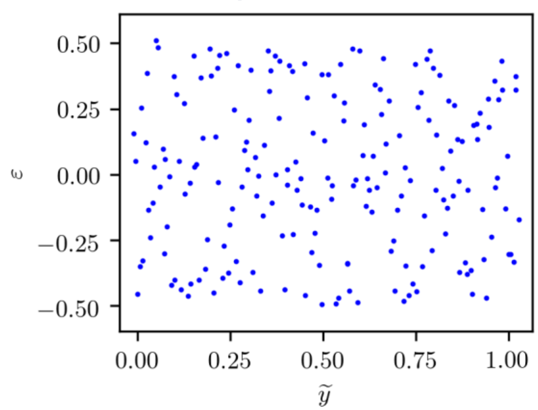
TECHNISCHE UNIVERSITÄT DRESDEN
Faculty of Mechanical Science and Engineering | Institute of Fluid Mechanics | Chair of Turbomachinery and Flight Propulsion

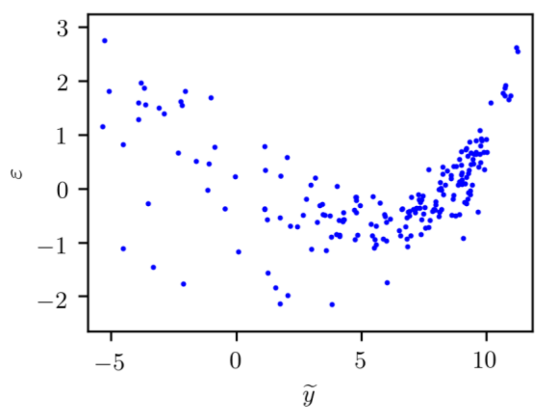


- LHS

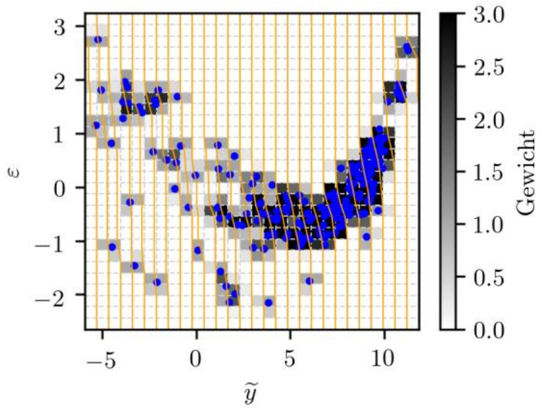


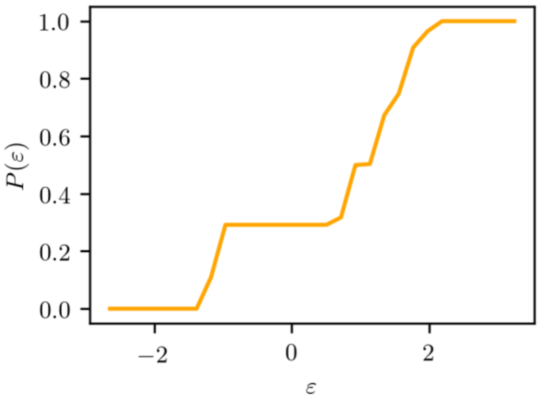
- Prediction of variance is better
- But: is the assumption correct?



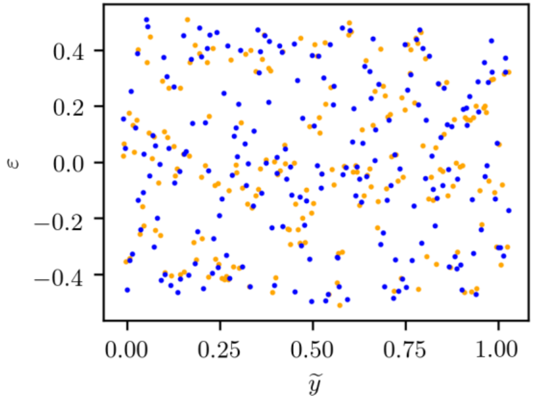


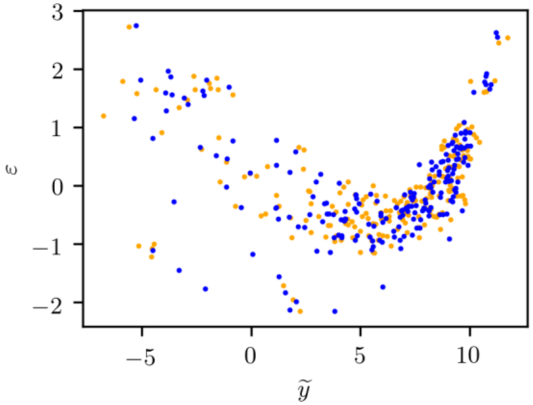
- Idea: local cumulative density functions

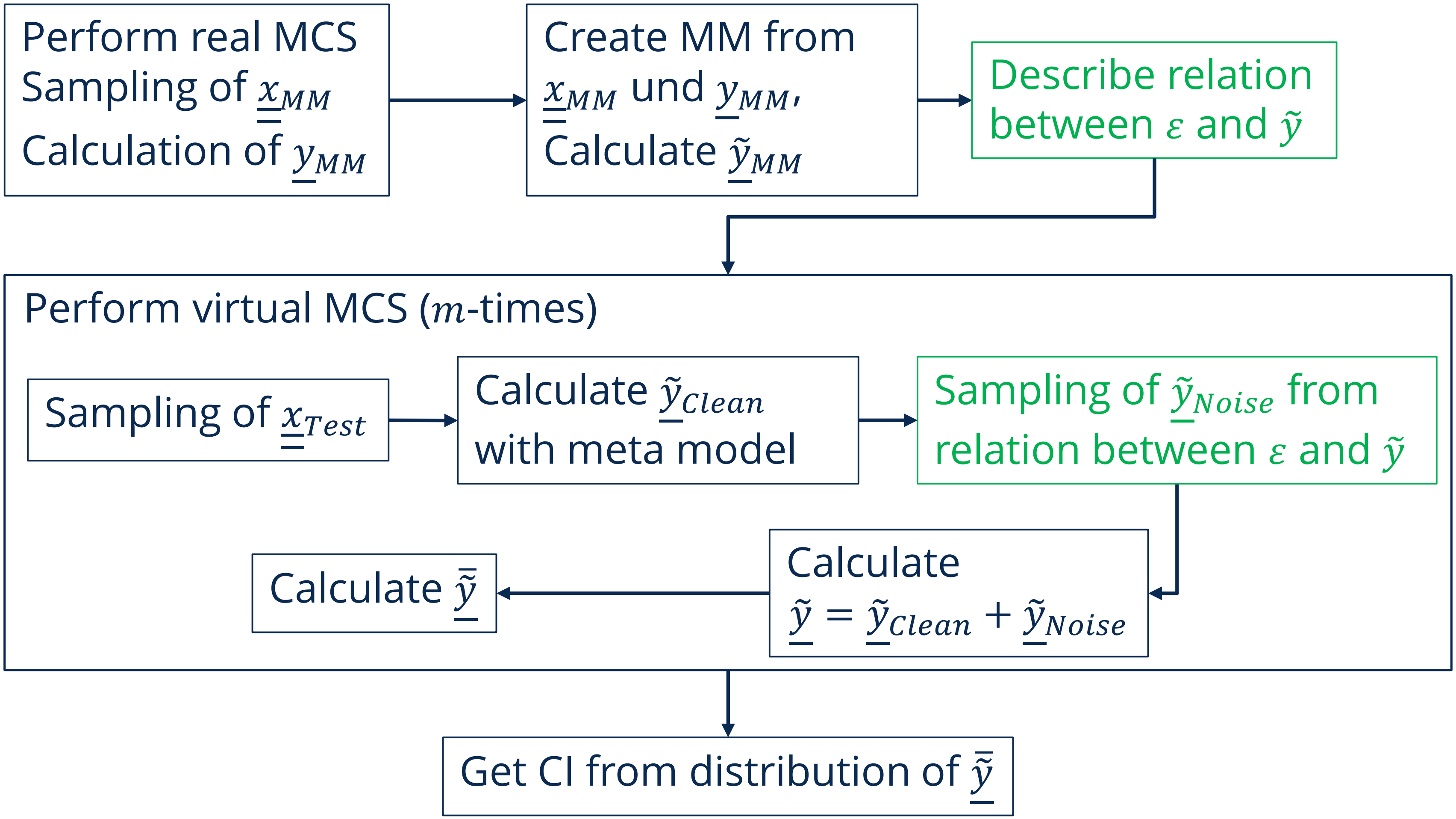




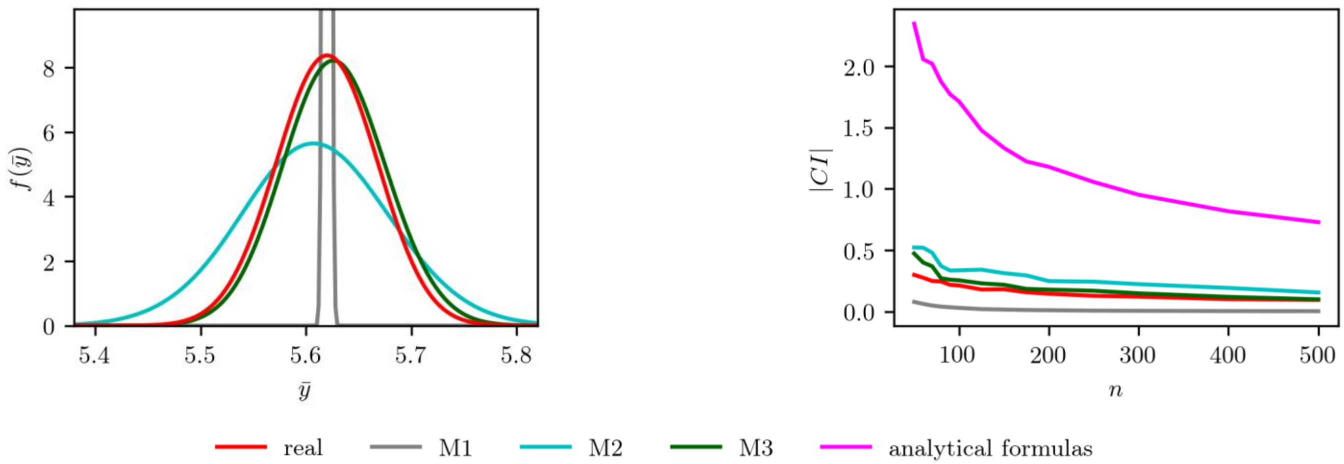
- Comparison:






Perform real MCS
Sampling of $\underline{x}_{MM}$
Calculation of $\underline{y}_{MM}$
Create MM from $\underline{x}_{MM}$ und $\underline{y}_{MM}$,
Calculate $\underline{\tilde{y}}_{MM}$
Describe relation between ε and $\tilde{y}$
Perform virtual MCS (m-times)
Sampling of $\underline{x}_{Test}$
Calculate $\underline{\tilde{y}}_{Clean}$ with meta model
Sampling of $\underline{\tilde{y}}_{Noise}$ from relation between ε and $\tilde{y}$
Calculate $\underline{\tilde{y}} = \underline{\tilde{y}}_{Clean} + \underline{\tilde{y}}_{Noise}$
Calculate $\underline{\bar{\tilde{y}}}$
Get CI from distribution of $\underline{\bar{\tilde{y}}}$

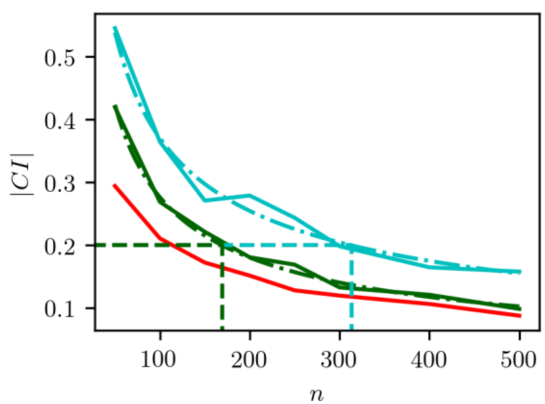
- LHS:

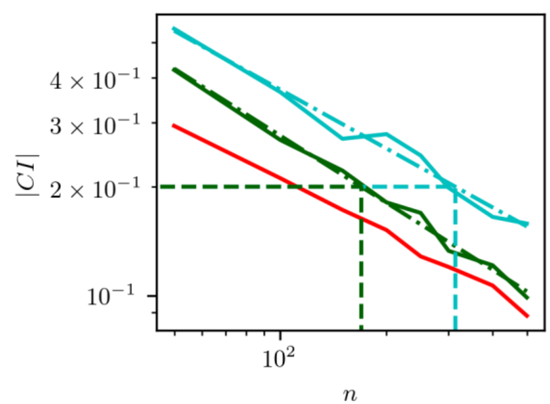

$f(\bar{y})$
0
2
4
6
8
5.4
5.5
5.6
5.7
5.8
$\bar{y}$
$|CI|$
0.0
0.5
1.0
1.5
2.0
100
200
300
400
500
n
real
M1
M2
M3
analytical formulas

- Certain size of CI must be achieved
- Size of CI known after MCS

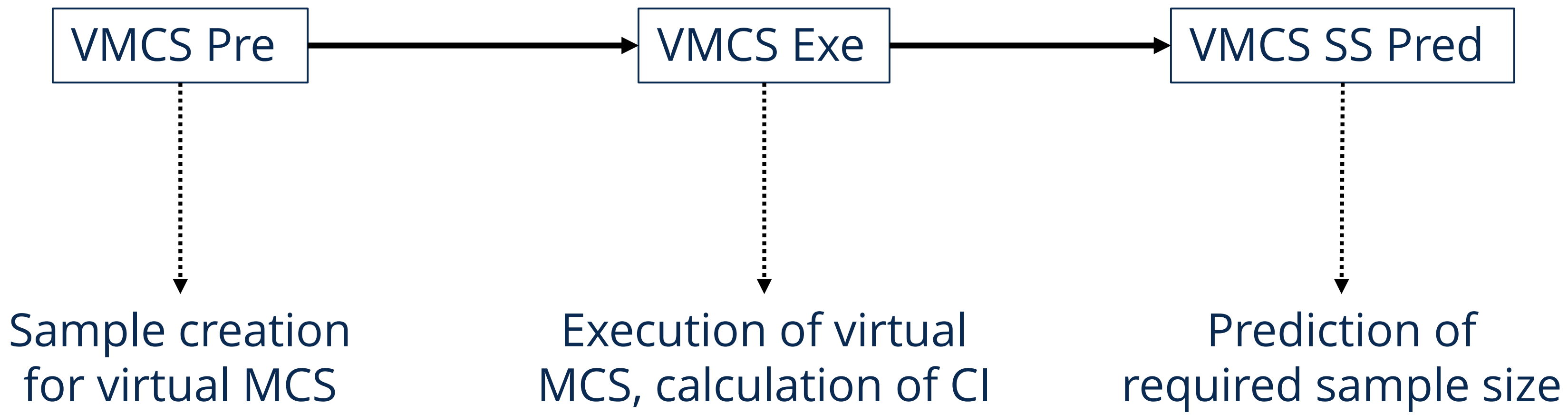
- Methods can be used to predict required sample size

- Steps:
  - Create MM, perform virtual MCS at different sample sizes
  - Determine size of CI at different sample sizes
  - Approximate evolution of CI size
  - Calculate required sample size

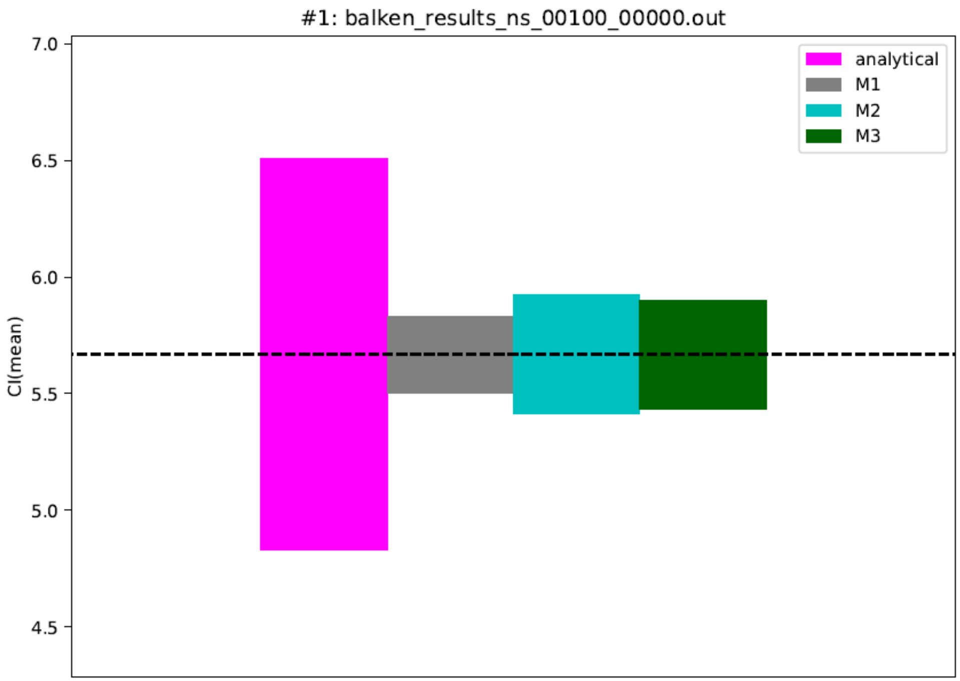




- Framework to calculate CI from given MCS
- Easy to handle, flexible usage
- Basis: MCS from ProSi



- Output: xml-Format or visualization



5. Summary

- Goal: Predict confidence interval of MCS with LHS more precisely
- Idea:
  - Approximate system behavior with meta model
  - Simulate MCS
- 3 Methods:
  - M1: Meta model
  - M2: Meta model + ε from PDF / CDF
  - M3: Meta model + ε from relation between $\tilde{y}$ and ε
- **SRS:** M1, M2, M3, **LHS:** M2, M3
- CI is predicted more precisely
- Framework for application was developed

- Outlook:
  - Application of methods on turbomachinery example
  - Use of other meta model types
  - Performing of further tests

TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Mechanical Science and Engineering| Institute of Fluid Mechanics | Chair of Turbomachinery and Flight Propulsion

Thank you for your attention!

Andriy Prots, Lars Högner, Matthias Voigt, Ronald Mailach

Andriy.Prots@tu-dresden.de

DRESDEN concept Exzellenz aus Wissenschaft und Kultur

Dresden, October 10, 2019