Evaluation of Measurement Uncertainty using Adaptive Monte Carlo Methods

Gerd Wübbeler[1], Peter M. Harris[2], Maurice G. Cox[2], Clemens Elster[1]

[1)] Physikalisch-Technische Bundesanstalt (PTB)
[2)] National Physical Laboratory (NPL)

Emerging Topics in Mathematics for Metrology – From Measurement Uncertainty to Metrology of Complex Systems

Physikalisch-Technische Bundesanstalt (PTB)

21-22 June 2010, Berlin, Germany

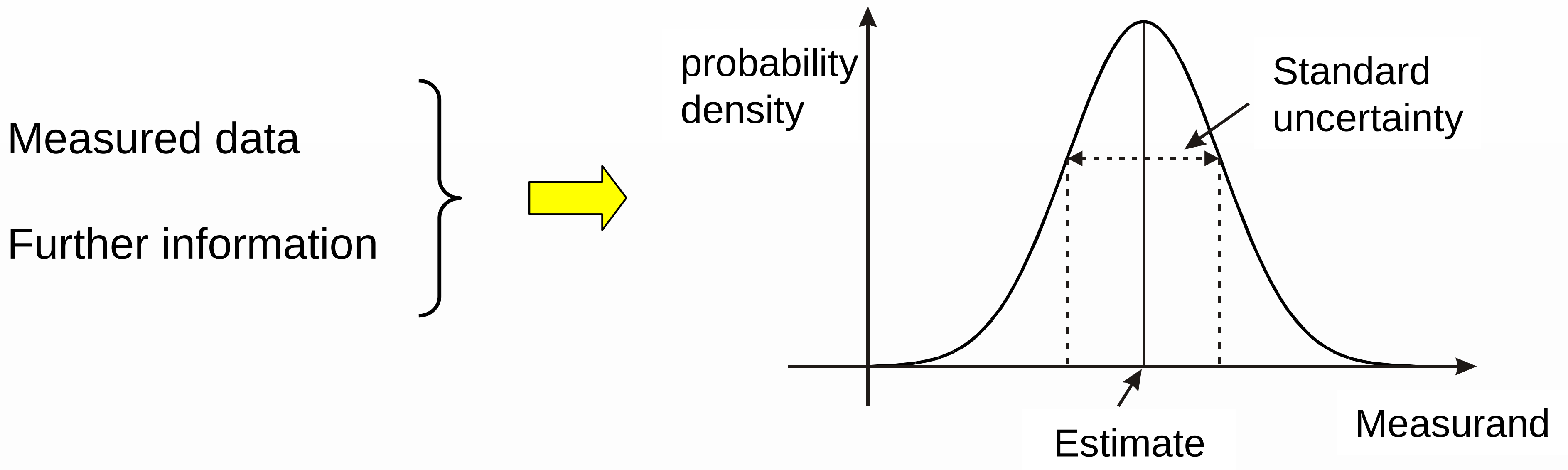
Content

- Evaluation of measurement uncertainty according to GUM S1
- GUM S1 adaptive Monte Carlo scheme
- Alternative approach: Stein's Two-stage scheme

GUM Supplement 1 (GUM S1)

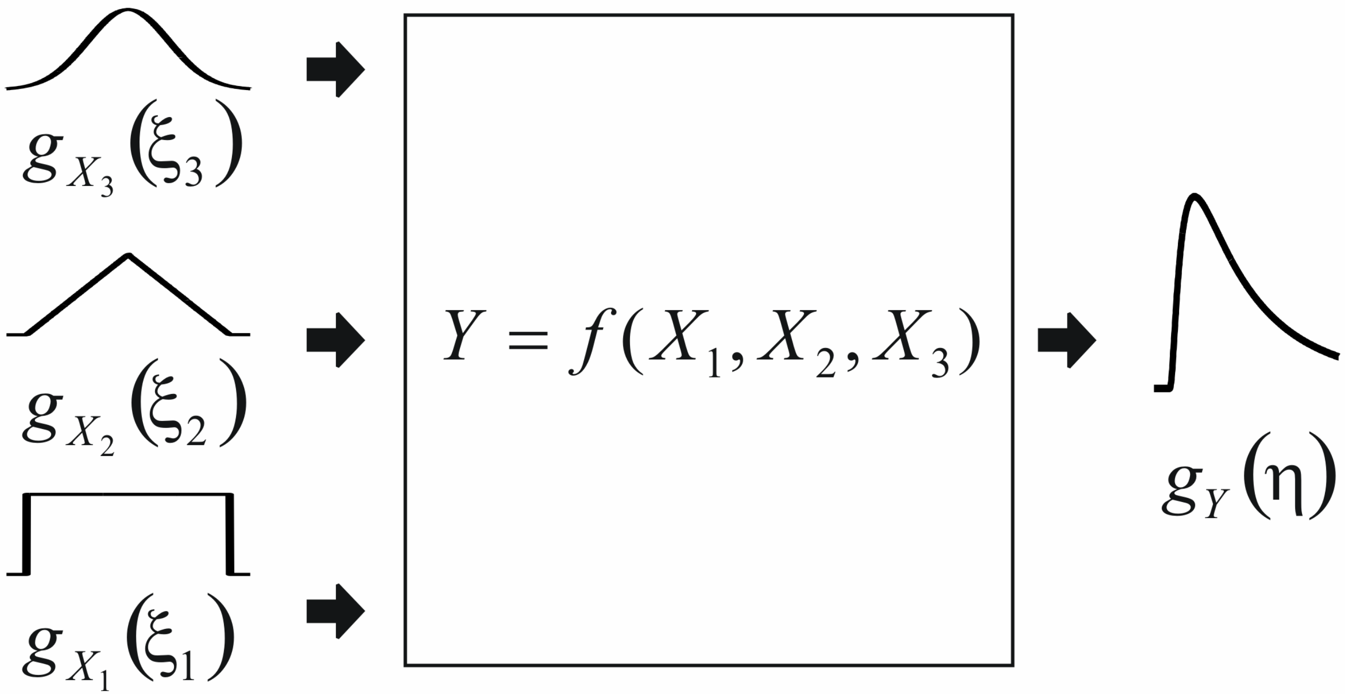
- **PDF** based method
- Numerical evaluation by a **Monte Carlo Method (MCM)**

Probability density function (PDF)



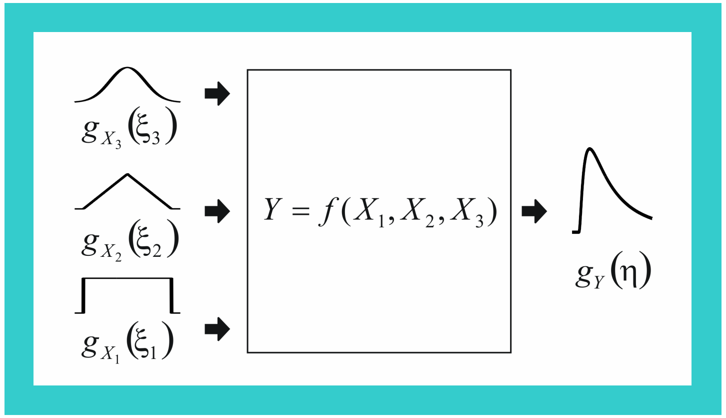
Propagation of distributions

PDFs for input quantities ⇨ Measurement model ⇨ PDF for measurand



Change-of-variables

GUM S1 Monte Carlo Method (MCM)



$$g_Y(\eta) = \int g_{\mathbf{X}}(\boldsymbol{\xi})\,\delta[\eta - f(\boldsymbol{\xi})]\,\mathrm{d}\boldsymbol{\xi}$$

Model

$Y = f(X_1, \ldots, X_N)$

PDF of input quantities

$g_{X_1,\ldots,X_N}(\xi_1, \ldots, \xi_N)$

$(\xi_1, \ldots, \xi_N)$ **random draw from** $g_{X_1,\ldots,X_N}(\xi_1, \ldots, \xi_N)$

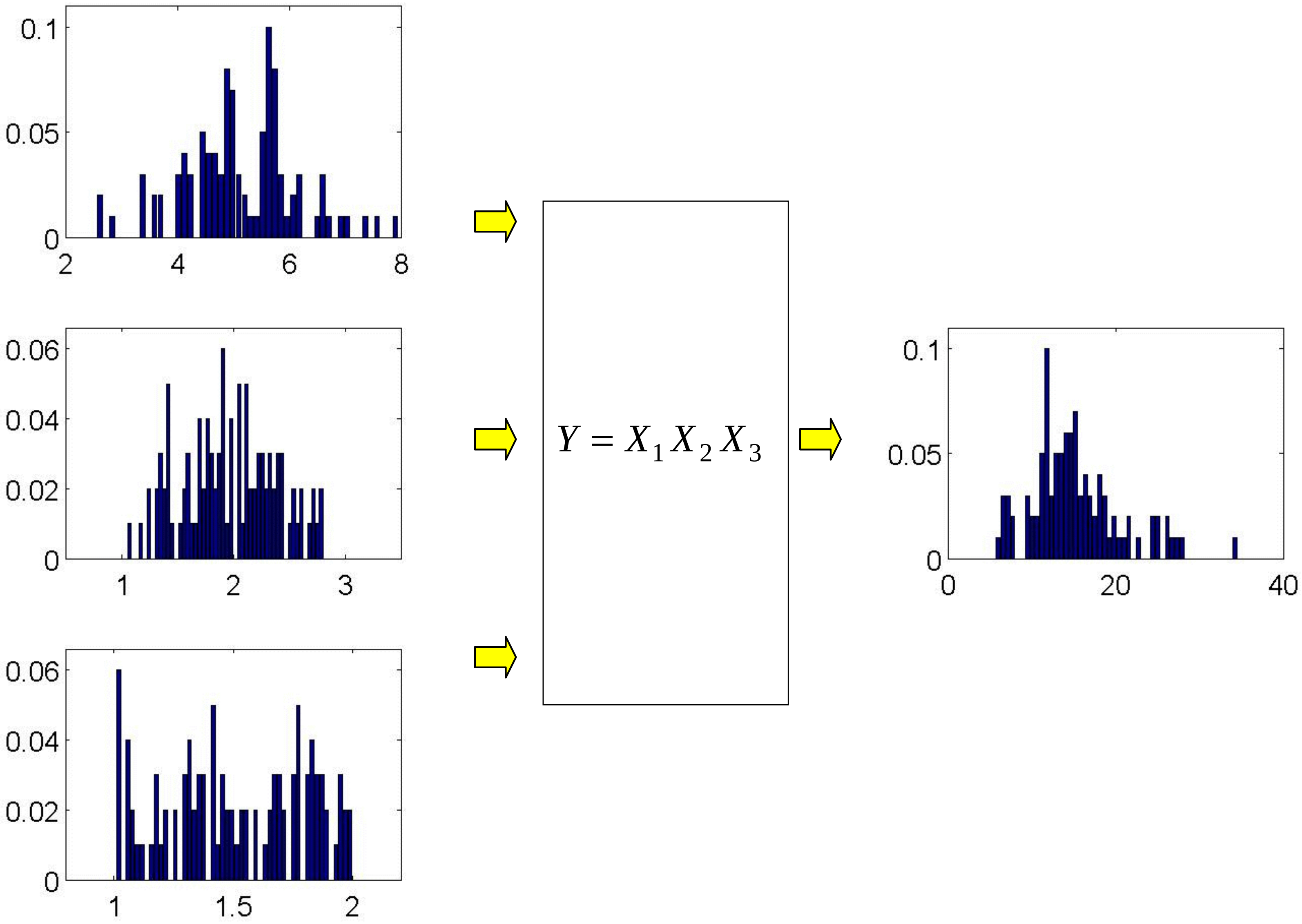
$\eta = f(\xi_1, \ldots, \xi_N)$ **evaluation of measurement model**

η **random sample from** $g_Y(\eta)$

many repetitions ➔ PDF $g_Y(\eta)$

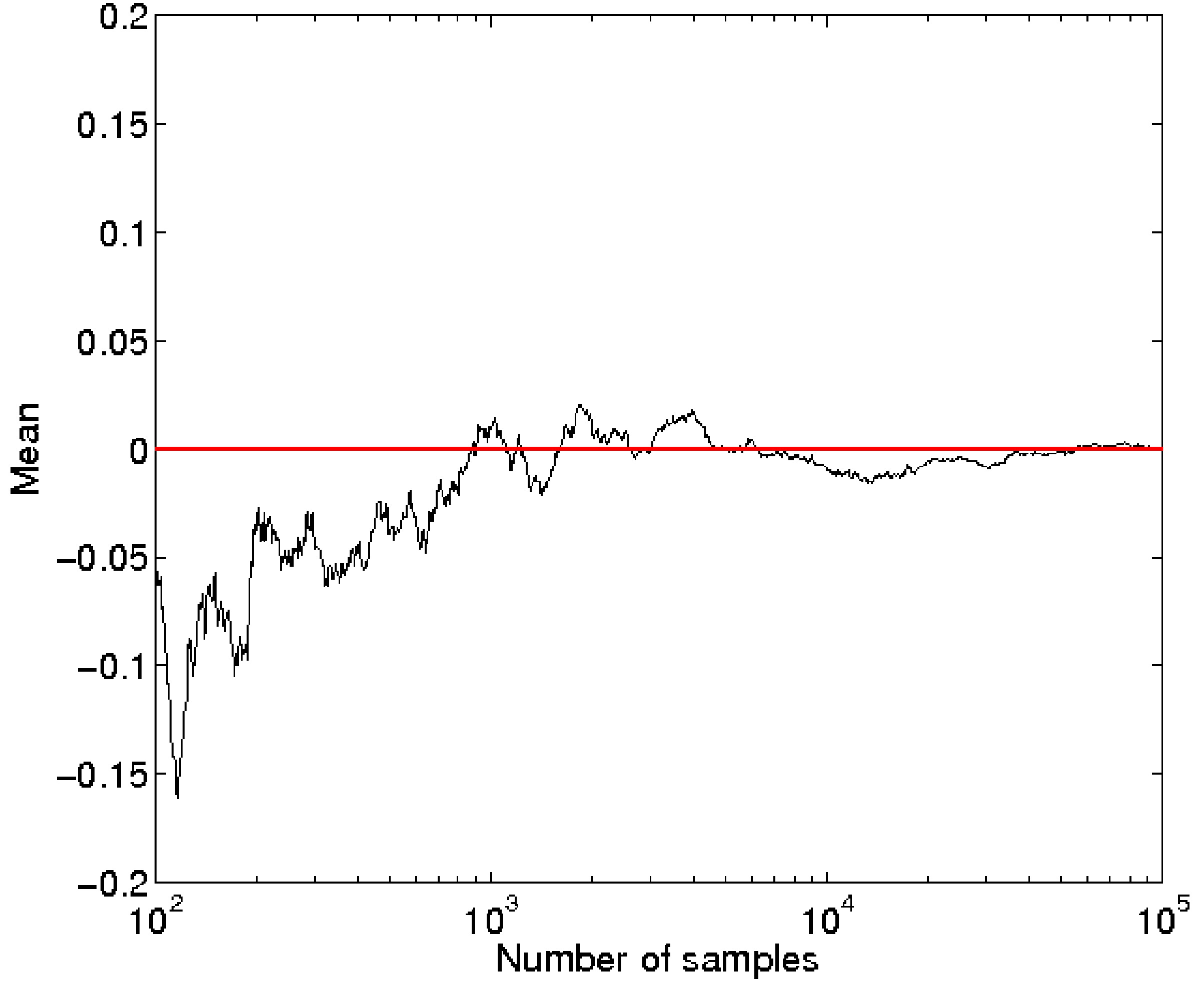
Illustration

10^2 **trials**



Convergence

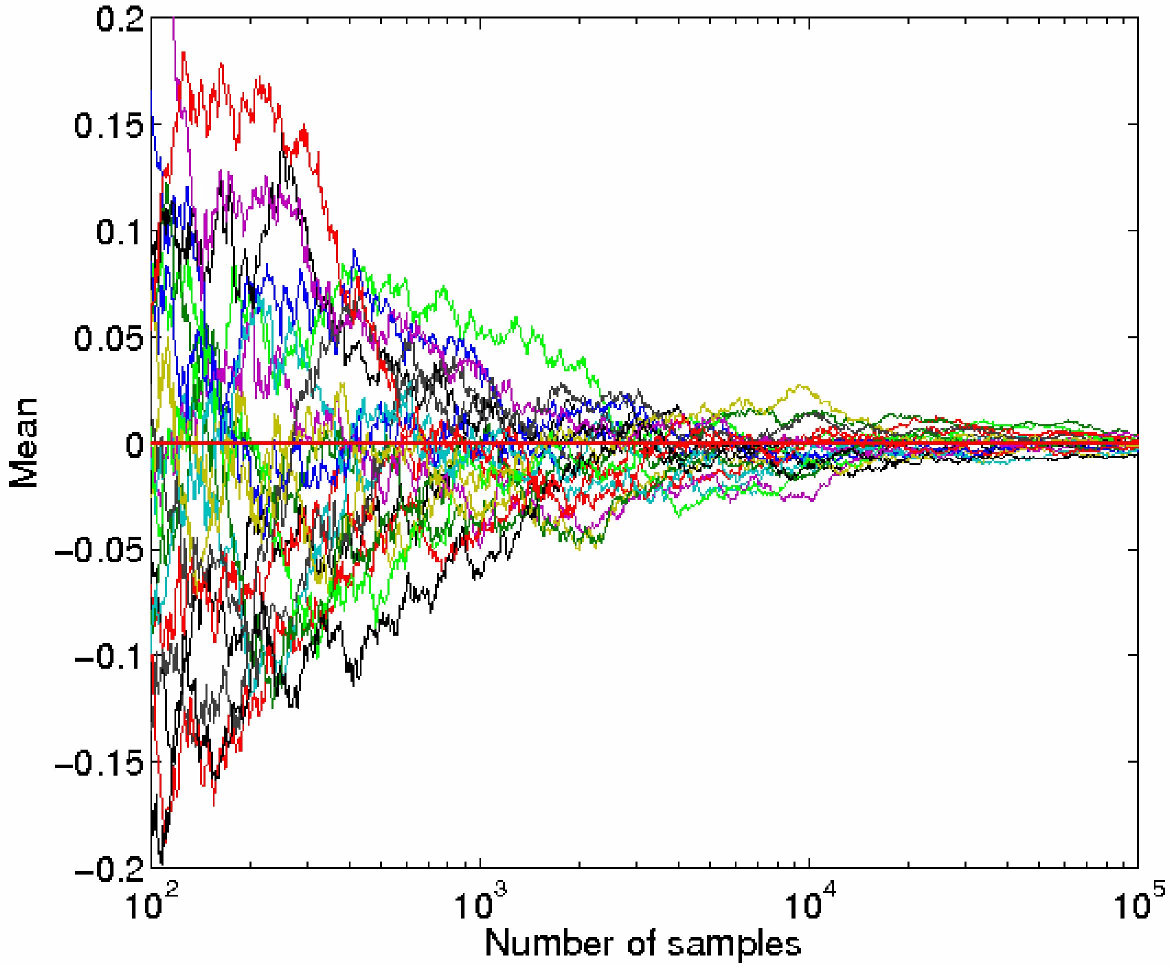
$$X_i \sim N(0,1)$$



Law of large numbers

Repetition of the MCM calculation

$$X_i \sim N(0,1)$$



MCM results exhibit random variations

GUM S1 Adaptive Monte Carlo scheme (7.9)

Goal Estimation of the expectation y with **accuracy** δ with a **coverage probability** of about 95 %.

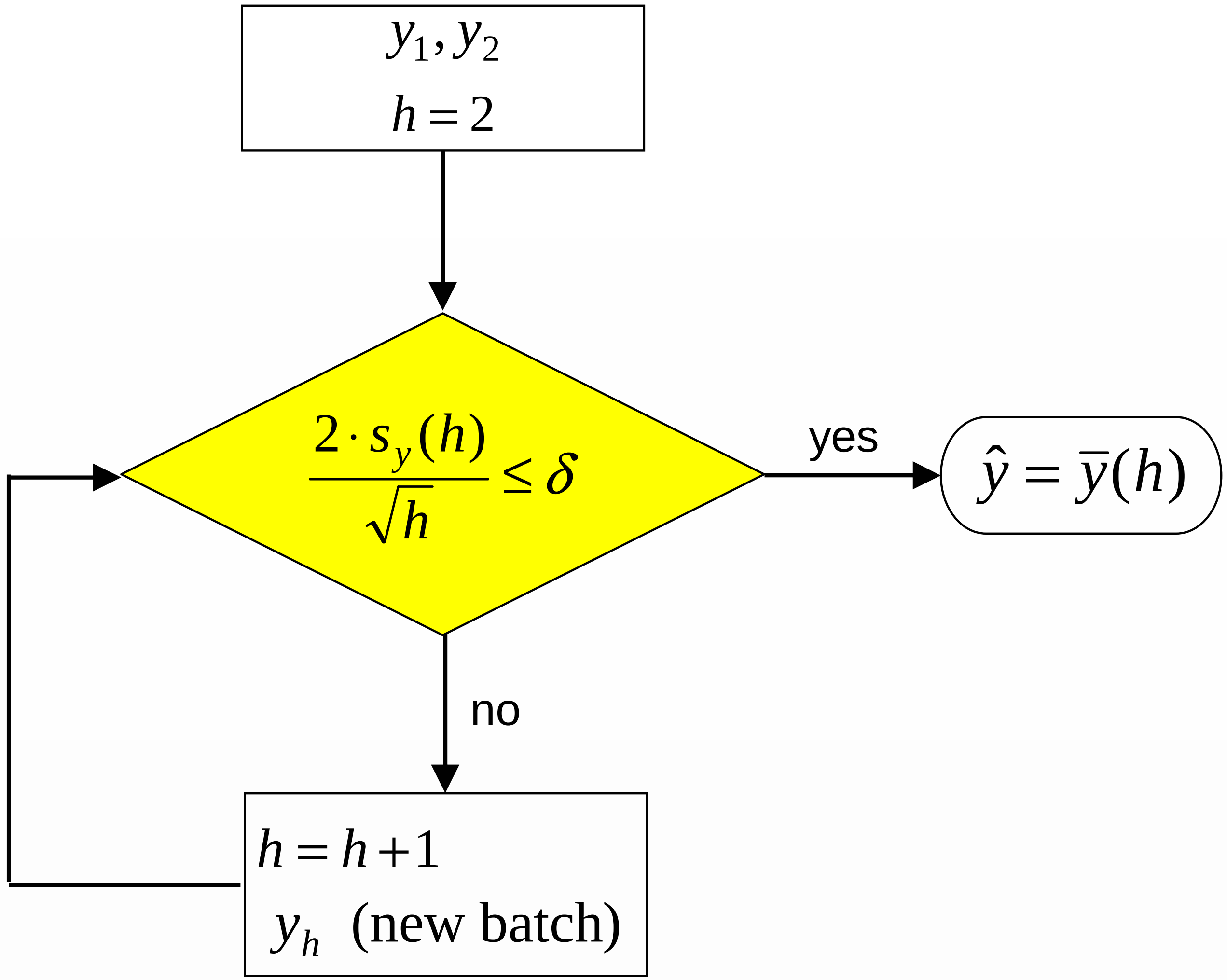
- **Sequential batch-processing mode** (e.g. 10 000 trials per batch)
- y_i **mean** of the trials within batch i
- y_i for sufficiently large batch size **Gaussian** distributed (central limit theorem)

$$(y_1, \ldots, y_h) \Rightarrow \begin{cases} \bar{y}(h) = \dfrac{1}{h} \displaystyle\sum_{i=1}^{h} y_i \\ s_y^2(h) = \dfrac{1}{h-1} \displaystyle\sum_{i=1}^{h} (y_i - \bar{y}(h))^2 \end{cases}$$

GUM S1 Adaptive Monte Carlo scheme (7.9)

Start: Batch 1 and 2

Stopping-rule



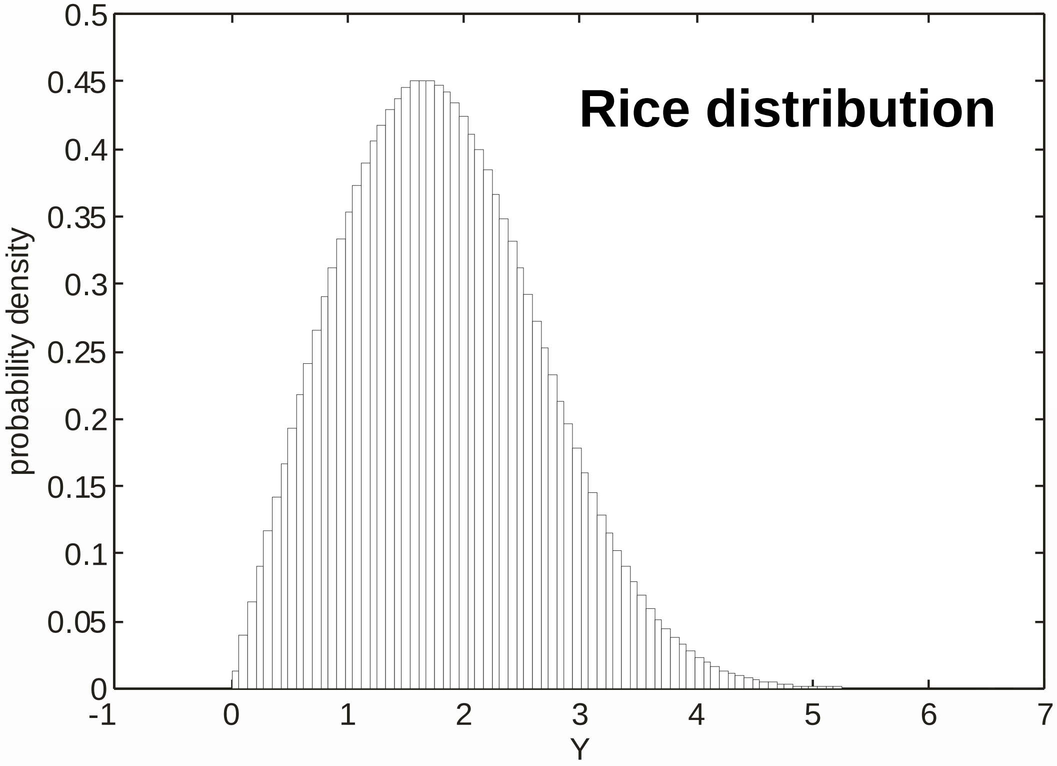
Assessment of the adaptive schemes

Model $Y = \sqrt{X_1^2 + X_2^2}$

Estimates $x_1 = x_2 = 1$

Uncertainties $u(x_1) = u(x_2) = 1$

⇨ Gaussian distributions (uncorrelated)



y	= 1.812 9
$u(y)$	= 0.844 6

Assessment of the adaptive schemes

Goal Determination of the expectation value of the Rice distribution with an accuracy of $\delta = 0.005$ (95 %)

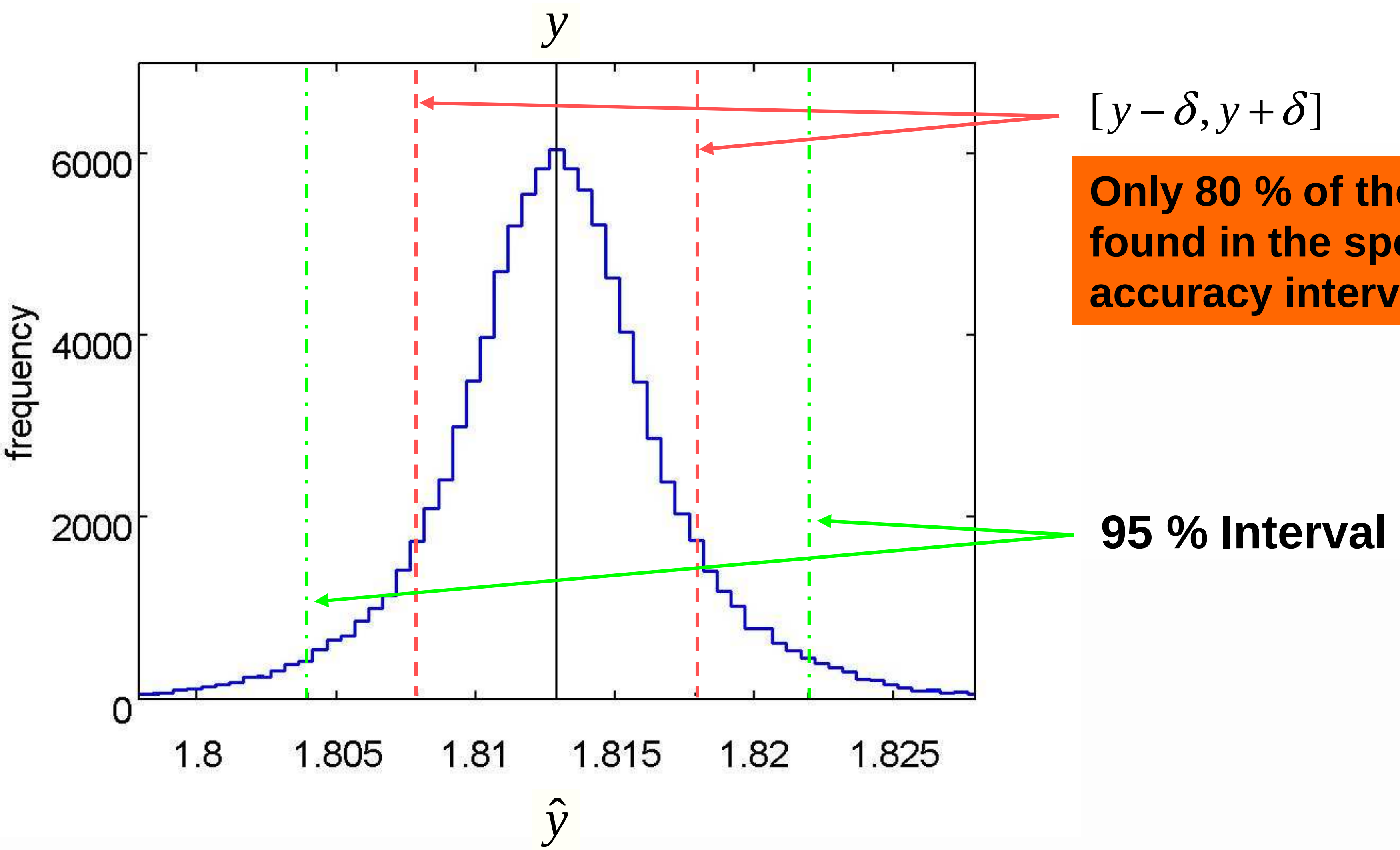
Assessment Adaptive scheme is repeatedly executed **100 000 times**

Result per run

- Estimate of the expectation value
- Number of required batches

Assessment of the GUM S1 adaptive scheme

Distribution of 10^5 estimates

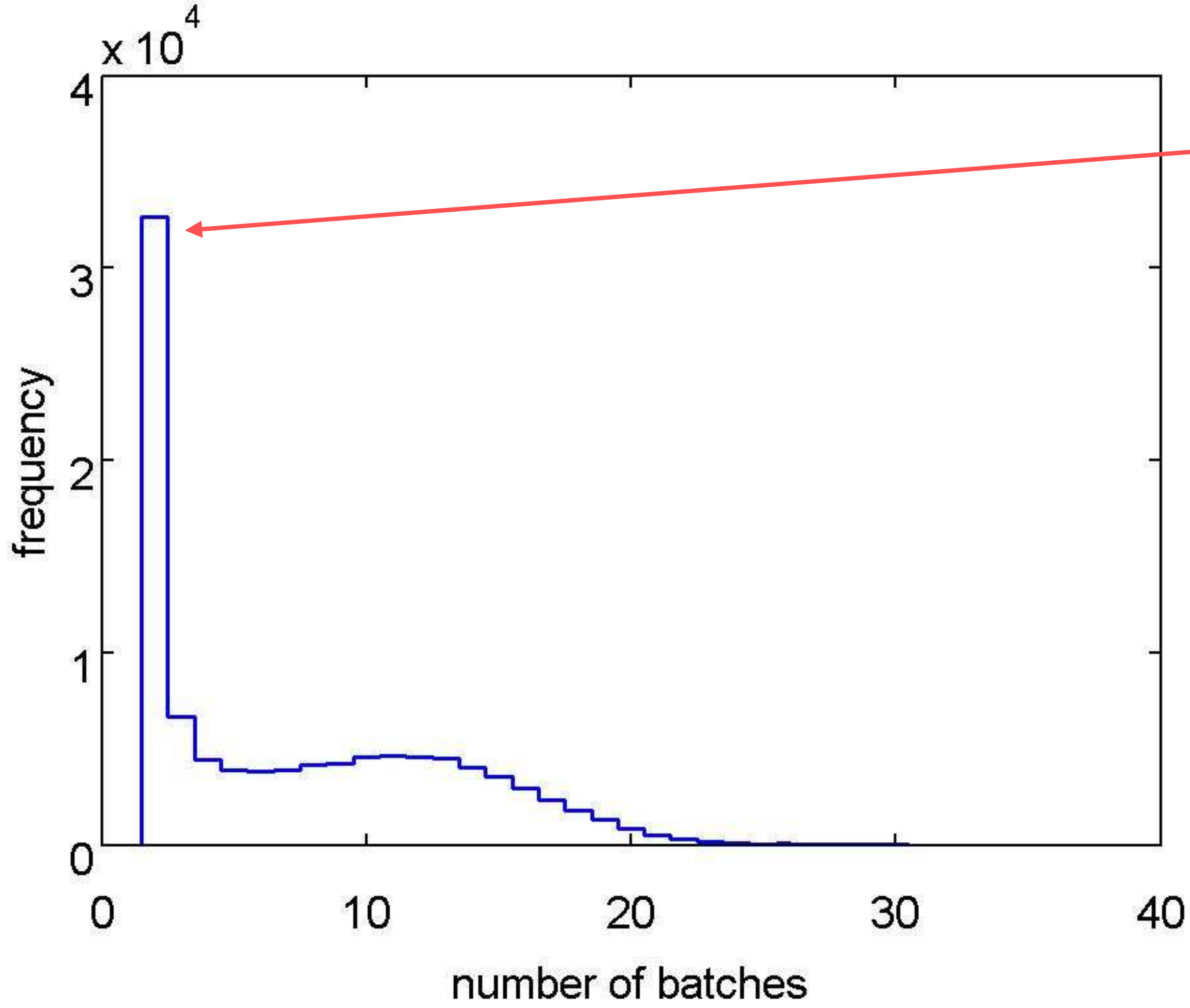


Only 80 % of the results found in the specified accuracy interval

(Batch size 10^4, δ=0.005)

Assessment of the GUM S1 adaptive scheme

Distribution of the number of batches



Large number of early terminations at h = 2 batches (about 32 %)

(Batch size 10^4, δ=0.005)

Reasons for the behavior of GUM S1 adaptive scheme

Sequential estimation scheme

- Random sample size (random number of batches)
- Multiple (dependent) testing for termination of sampling

- Confidence level of GUM S1 stopping rule adequate for fixed sample size
- Multiple testing and random sample size not taken into account

↓

Resulting confidence level does not necessarily meet the confidence level applied in individual test

Alternative Adaptive Monte-Carlo scheme

Goal	**Carry out the MCM until a prescribed accuracy is achieved at a specified confidence level (with lowest possible numerical effort)**

Stein‘s Two-stage scheme

Given $y_1, y_2, \ldots$ i.i.d. from $N(\mu, \sigma^2)$, μ, σ^2 unknown

Goal h and $\bar{y}(h) = \frac{1}{h}\sum_{i=1}^{h} y_i$

so that $[\bar{y}(h) - \delta, \bar{y}(h) + \delta]$ is a confidence interval for μ at confidence level 1-α

Step 1

Make $h_1 > 1$ random draws $y_1, \ldots, y_{h_1}$

➔Variance $s_y^2(h_1) = \frac{1}{h_1 - 1}\sum_{i=1}^{h_1}(y_i - \bar{y}(h_1))^2$

Two-stage scheme

Step 2

Number h_2 of additionally required drawings

$$h_2 = \max\left(\left\lfloor \frac{s_y^2(h_1)\cdot\left(t_{h_1-1,1-\alpha/2}\right)^2}{\delta^2}\right\rfloor - h_1 + 1,\ 0\right)$$

$h_2 > 0$ ➔ make h_2 further random draws

$h_2 = 0$ ➔ no additional random draws are made

} $\bar{y}(h_1 + h_2)$

$$\Pr(\bar{y}(h_1 + h_2) - \delta \le \mu \le \bar{y}(h_1 + h_2) + \delta) \ge 1 - \alpha$$

Proof C Stein, 1945, *Ann. Math. Statist.* **16** 243-58

Application of the two-stage scheme within GUM S1

Sequential batch-processing mode

y_i	**Mean of the trials in batch *i***	unbiased
$u_i^2(y)$	**Variance of the trials in batch *i***	

sufficiently large batch size (CLT) ➔ y_i , $u_i^2(y)$ approximately **Gaussian distributed**

Two-stage scheme applicable for

- **Estimate** y
- **Squared uncertainty** $u^2(y)$

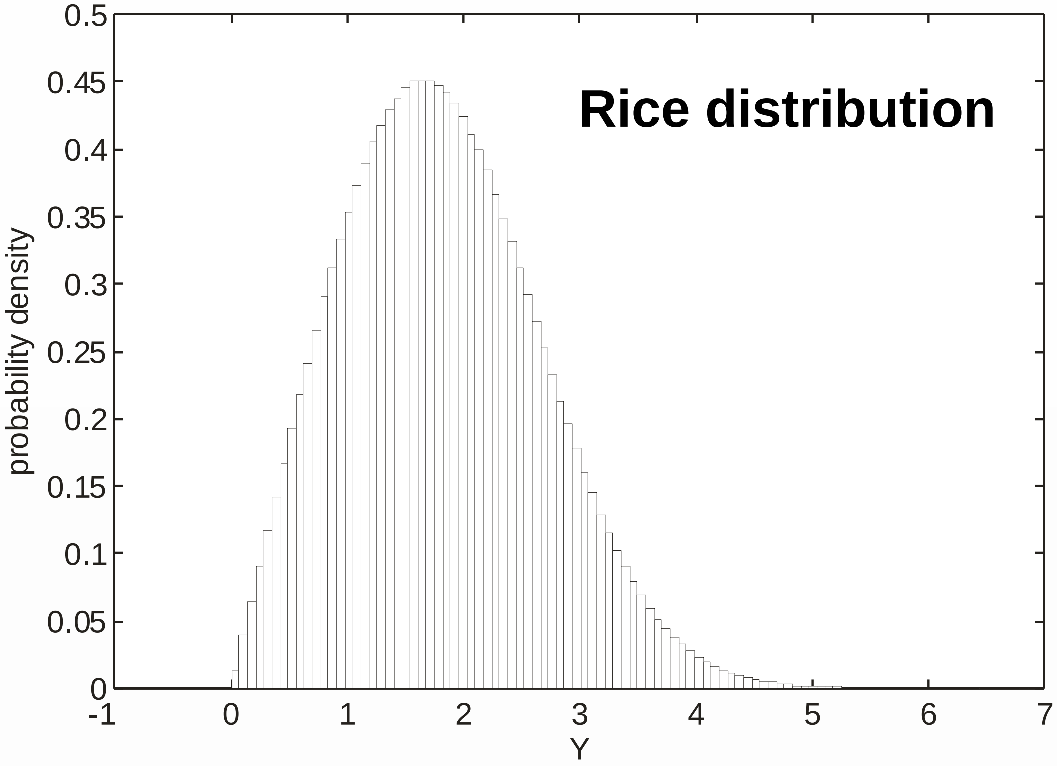
Assessment of the adaptive schemes

Model $Y = \sqrt{X_1^2 + X_2^2}$

Estimates $x_1 = x_2 = 1$

Uncertainties $u(x_1) = u(x_2) = 1$

⇨ Gaussian distributions (uncorrelated)

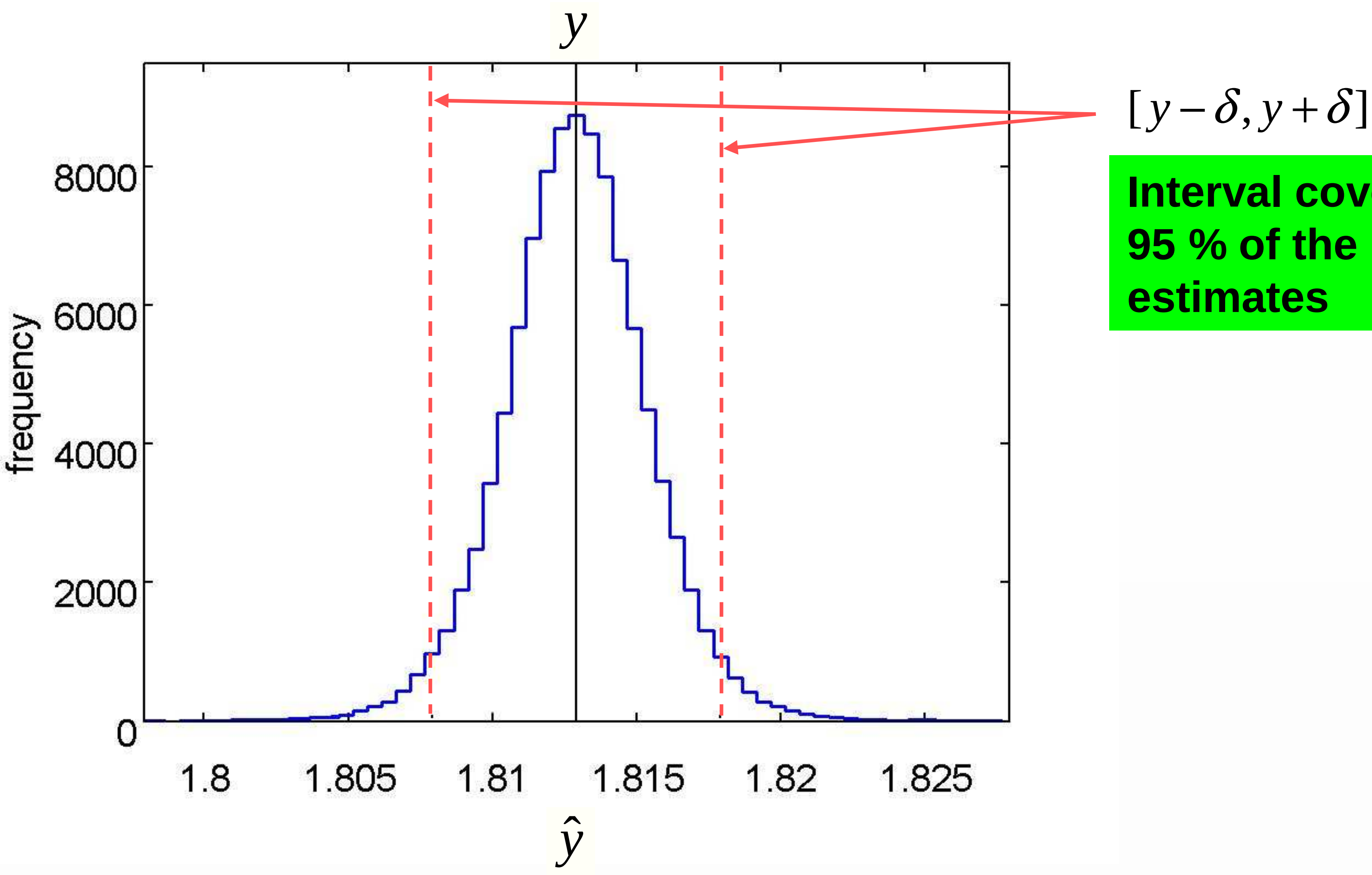


y = 1.812 9

$u(y)$ = 0.844 6

Assessment of the two-stage scheme

Distribution of 10^5 estimates

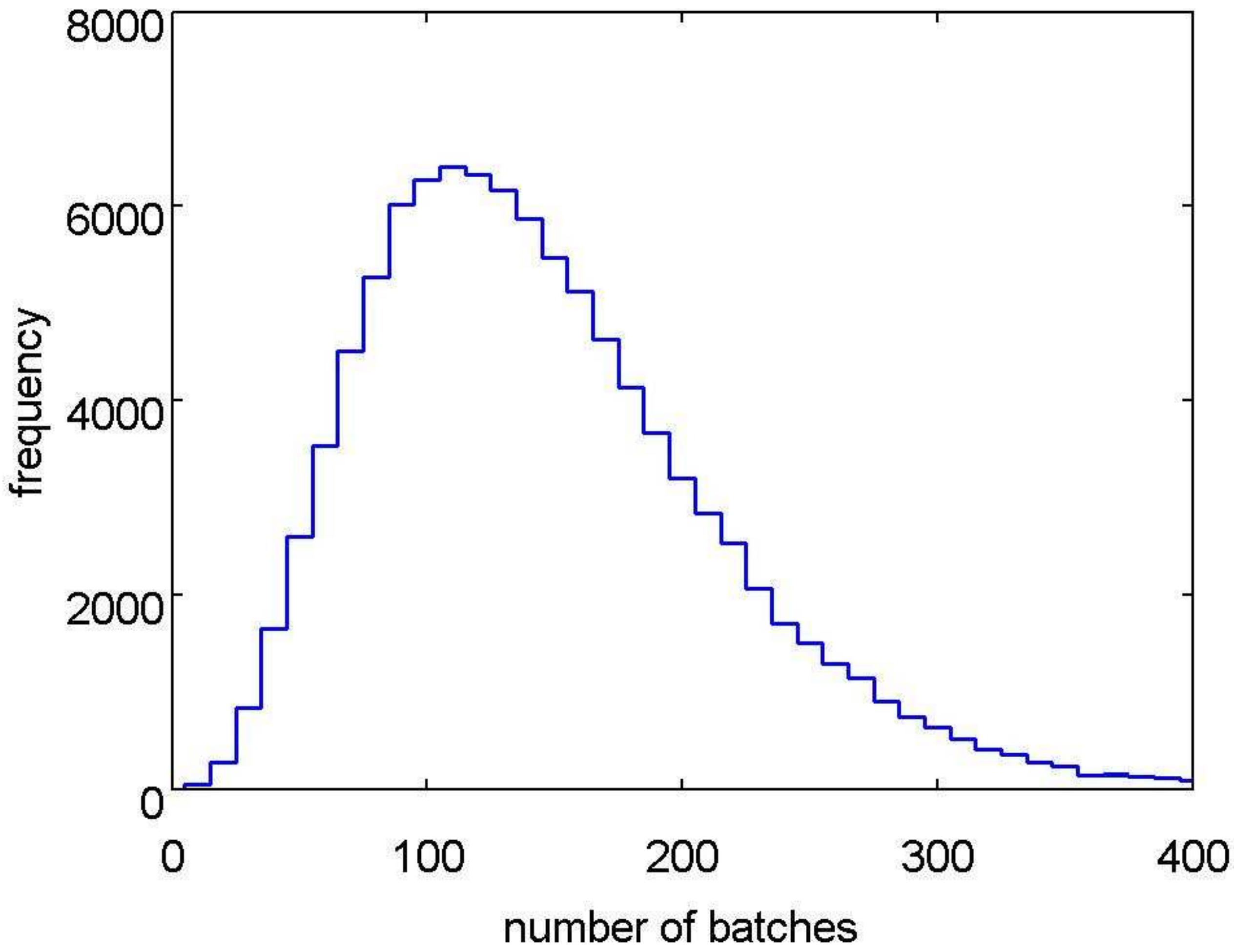


Interval covers 95 % of the estimates ✓

(h_1=10, Batch size 10^3, δ=0.005, α=0.05)

Assessment of the two-stage scheme

Distribution of the number of batches



No early terminations

(h_1=10, Batch size 10^3, δ=0.005, α=0.05)

Assessment of the two-stage scheme

Determination of the estimate of the measurand: dependence on batch size

Batch size	Success rate	ANT*
1	0.952(0.001) ✓	145 802(211)
10	0.951(0.001) ✓	146 176(218)
100	0.952(0.001) ✓	146 423(218)
1 000	0.951(0.001) ✓	146 383(218)
10 000	0.969(0.001) ✓	157 271(196)

$(h_1 = 10, \delta = 0.005, \alpha = 0.05)$

Result for a requested confidence level of 99.9 %

$(\alpha = 0.001, \quad \text{Batch size } 1000)$

➔ **Success rate** **0.999 08 (0.000 1)** ✓

ANT **654 760 (972)**

*) **ANT: Average Number of Trials**

Summary: Adaptive schemes for GUM S1

- GUM S1 adaptive scheme does not (intend to) meet a 95 % confidence level

- Alternative approach: Two-stage scheme

  - ✓ Attains specified confidence level for a Gaussian distribution (Proof by C. Stein, important for metrological applications)

  - ✓ Applicable for the estimate and the squared uncertainty

  - ✓ Allows prediction of computation time (number of trials)

Wübbeler, Harris, Cox, Elster **Metrologia** 47 (2010) 317–324