Fast integration using quasi-random numbers

J.Bossert, M.Feindt, U.Kerzel
University of Karlsruhe

ACAT 05

bmb+f - Förderschwerpunkt
Elementarteilchenphysik
Großgeräte der physikalischen Grundlagenforschung

Outline

- Numerical integration
- Discrepancy
  - The Koksma-Hlawka inequality
- Pseudo- and quasi-random numbers
  - Generators for quasi-random numbers
  - Optimising the discrepancy
  - Quasi-random numbers in *n* dimensions
  - Comparison
- Examples

Numerical integration

- Integral evaluation using MC technique
  - interprete integral as expectation value

    $$If = \int f(x)dx = vol(B)\,\langle f(x)\rangle$$

  - estimate $\langle f(x)\rangle$ by averaging over *N* samples:

    $$\langle f(x)\rangle \approx Q_n f = 1/N \sum_i f(x_i)$$

    $\rightarrow$ use uniformly distributed pseudo-random numbers to sample $f(x_i)$

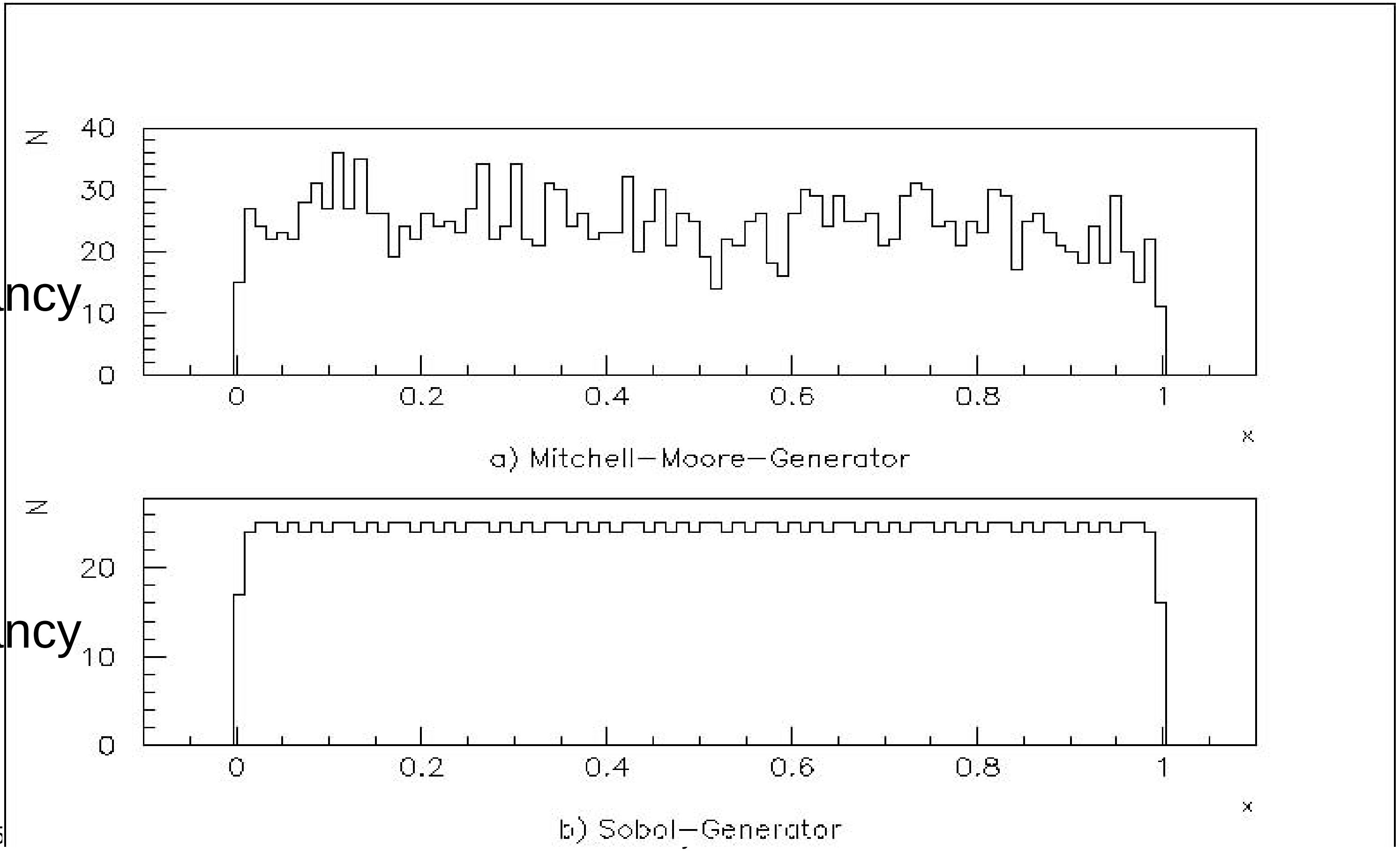
  - define error: $\epsilon = |If - Qf|$

Discrepancy

- measure of roughness
  (deviation from desired flat distribution)

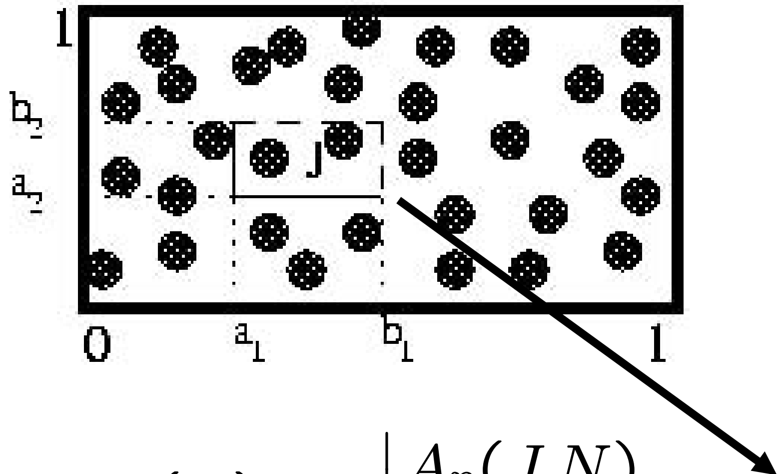
large discrepancy

small discrepancy



Discrepancy cont...

- in 2 dimensions:



- local discrepancy $R_N(J) = \left| \frac{A_p(J,N)}{N} - vol(J) \right|$

$A_p(J,N)$ # points in J *N* # points in unit square

(number of points in J proportional to ratio of area of J to unit square)

- generalised to *s* dimensions

Discrepancy cont...

- Discrepancy of a ensemble P with norm:

  - $L_\infty : \quad D^*_\infty(P_N) = \sup_{J(x)\in I^s} |R_N(J)|$
  - $L_2 : \quad D^*_2(P_N) = \left(\int_{I^s} (R_N(J))^2\right)^{\frac{1}{2}}$

$$D^*_\infty(P_N) \geq D^*_2(P_N)$$

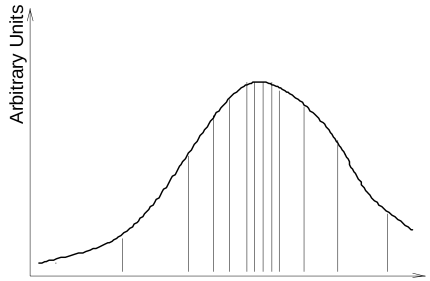
*: one corner of J in (0,0) of unit square

Koksma-Hlawka inequality

- relates error on numerical integration with discrepancy:

$$V(f)D^*_\infty(P_N) \geq \epsilon = |If - Qf|$$

- two handles to minimise error ε:
  - *minimise variation* $V(f)$ of function f
    - variable transformation
    - importance sampling



  - sample with numbers with *low discrepancy* $D^*_\infty(P_N)$

Pseudo-random numbers

- follow deterministic pattern
  - created by e.g. linear congruence generator

$$y_{i+1} = (ay_i + c)\, mod\ m \rightarrow x_i = y_i/m$$

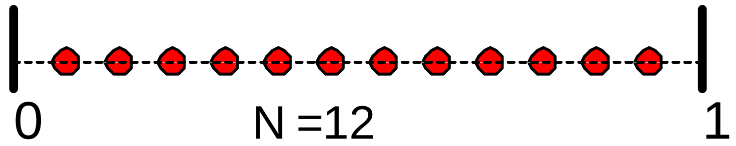
- statistically independent from each other
  - simulated "real" random numbers
  - Beware of period of generator (e.g. Ranlux: 10^{165})

- Discrepancy:

$$D^*_\infty(P_N) = \mathcal{O}\left(\sqrt{\frac{\log\log N}{N}}\right) \longrightarrow \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \qquad \boxed{\epsilon \propto \frac{1}{\sqrt{N}}}$$

Lattice

- 1d: equidistant points have minimal discrepancy



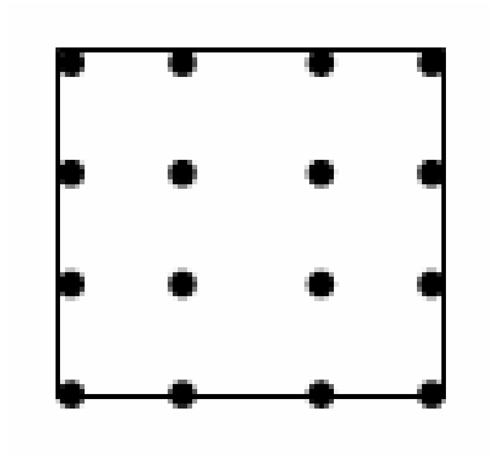
- first idea: extend to *s* dimensions

  $\rightarrow$ lattice

- need N=n^d points

  (otherwise no lattice)

e.g. n=4, d=2
$\rightarrow$ N=16



Quasi-random numbers

- constructed to be evenly distributed
- *not* independent from each other
- need to know total number N from beginning
  - good for integration, *not* simulation
- low discrepancy series:

$$D^*_\infty = \mathcal{O}\left(\frac{(\log N)^{s-1}}{N}\right)$$

- $\Rightarrow$ faster convergence, smaller error

$$D^*_\infty(P_N) \longrightarrow \mathcal{O}\left(\frac{1}{N}\right) \qquad \boxed{\epsilon \propto \frac{1}{N}}$$

Quasi-random series

- van der Corput series:

$$n = \sum_j d_j b^j \; ; \; n = d_N \dots d_2 d_1 d_0$$

radically inverse function:

$$\varphi_b : \mathcal{N}_0 \to [0,1) \; \varphi_b : d_N \dots d_2 d_1 d_0 \mapsto 0.d_0 d_1 d_2 \dots d_N$$

$$\boxed{x_n := \varphi_b(n-1)}$$

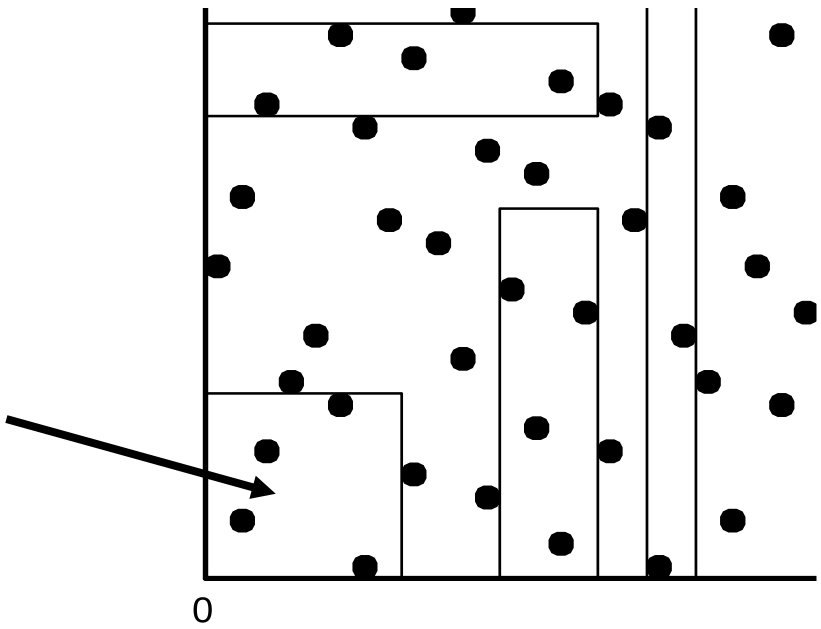
- other series:
  - Halton: extending Corput series to several dimensions
  - Hammersly: replace 1st dim. of Halton series by lattice
    - $\Rightarrow$ lower discrepancy

(T,M,S) nets and (T,S) series

- <u>(t,s) series:</u> class of quasi-random numbers using radically inverse functions with low discrepancy
- <u>(t,m,s) net:</u> each elementary interval E (Vol(E) = b^{t-m}) contains b^t points of series with b^m total points

basis

elementary interval E



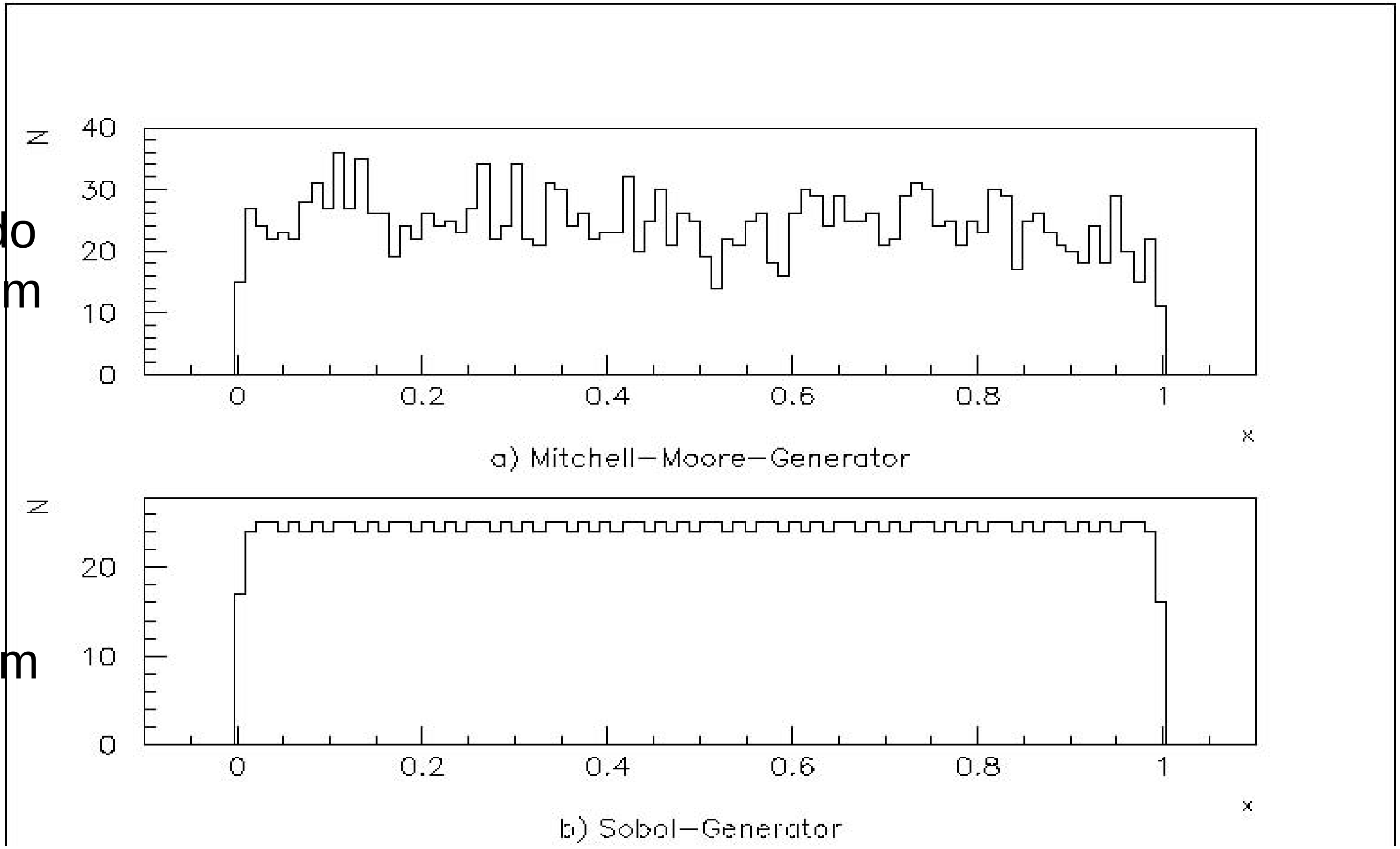
(2,6,2) net:
2^6 = 64 total points
2^2 = 4 points in E

Pseudo- vs. Quasi-random

- generate 2048 numbers

pseudo
random

quasi
random

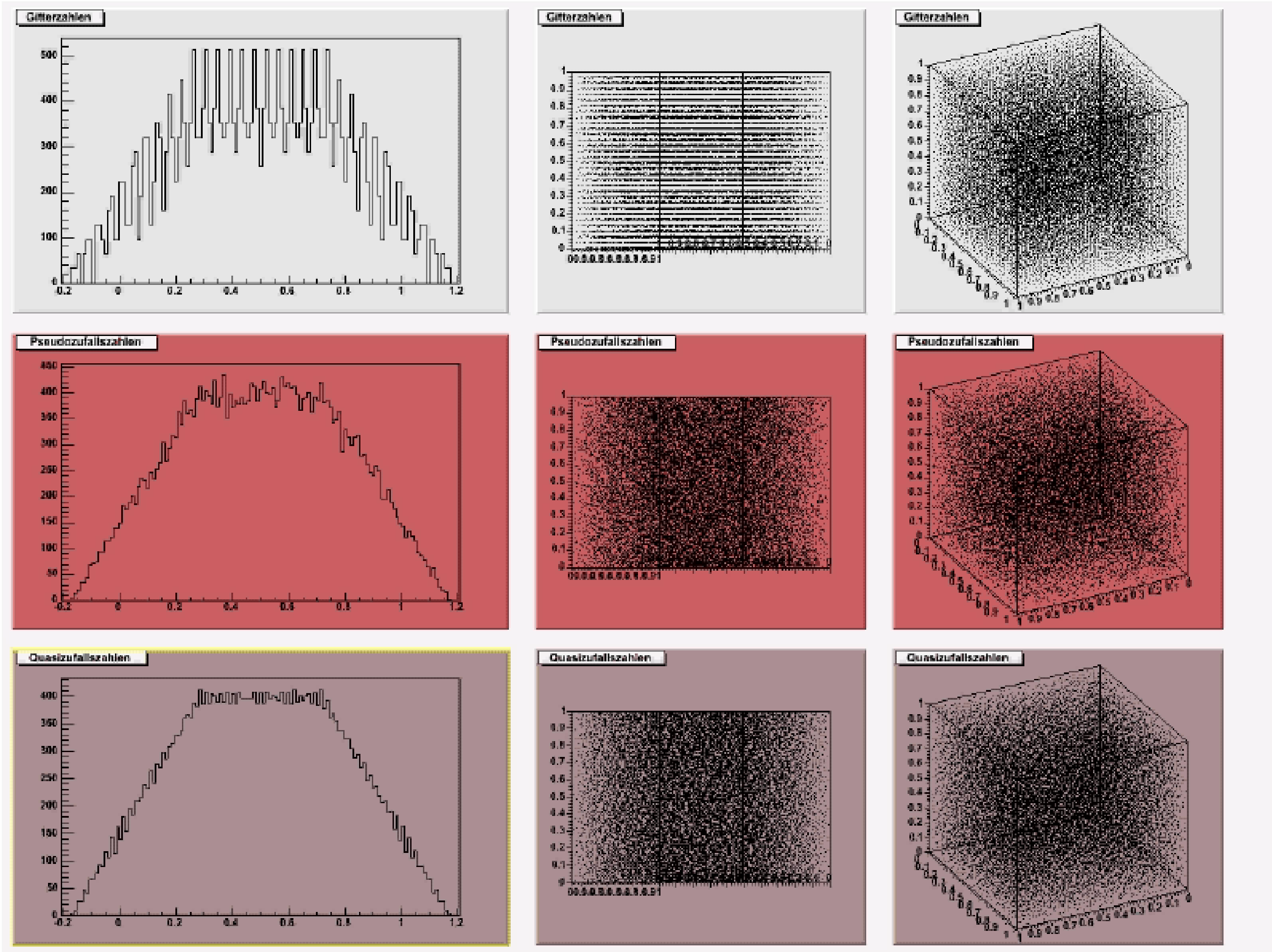


Comparison

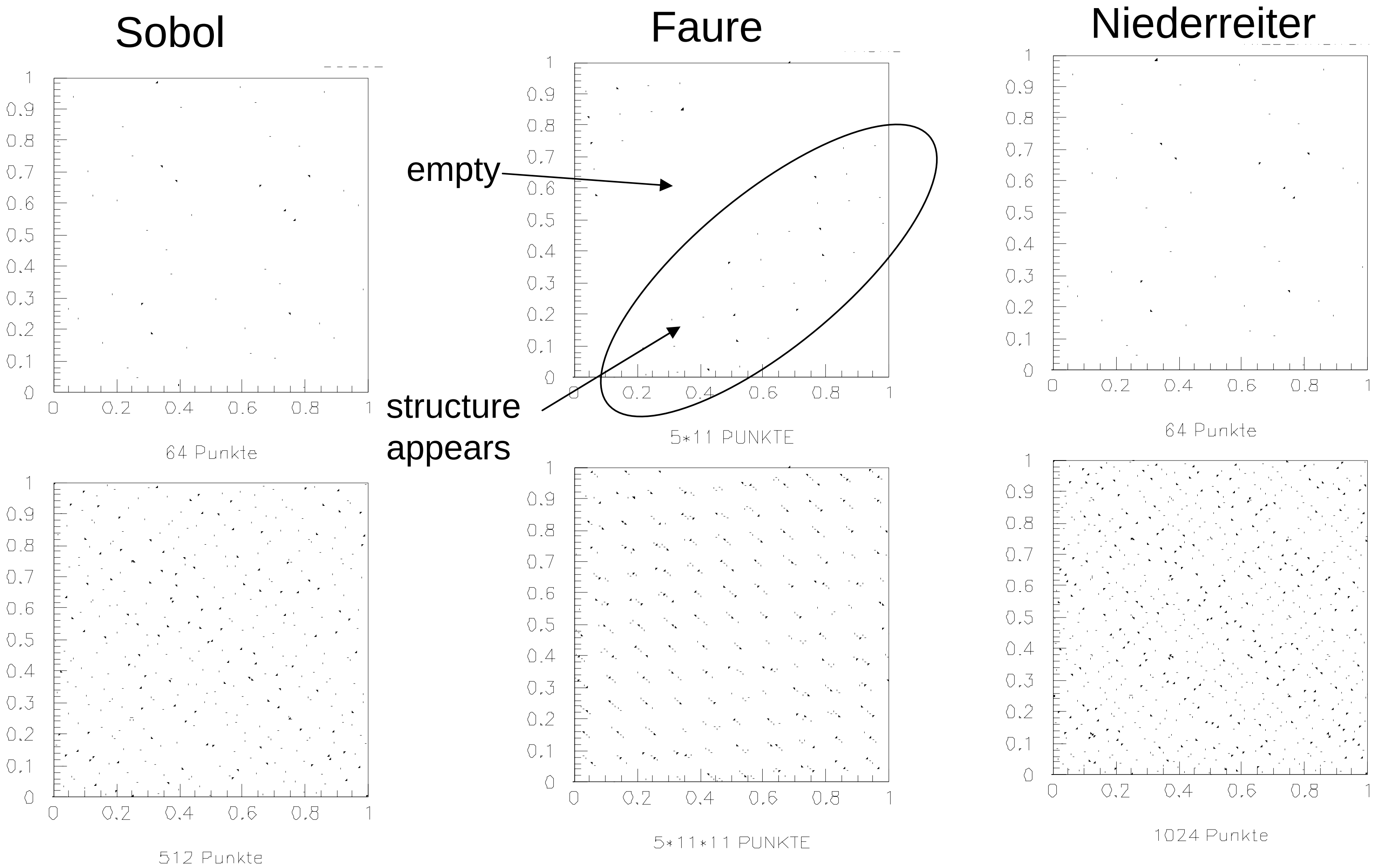
lattice

pseudo-random

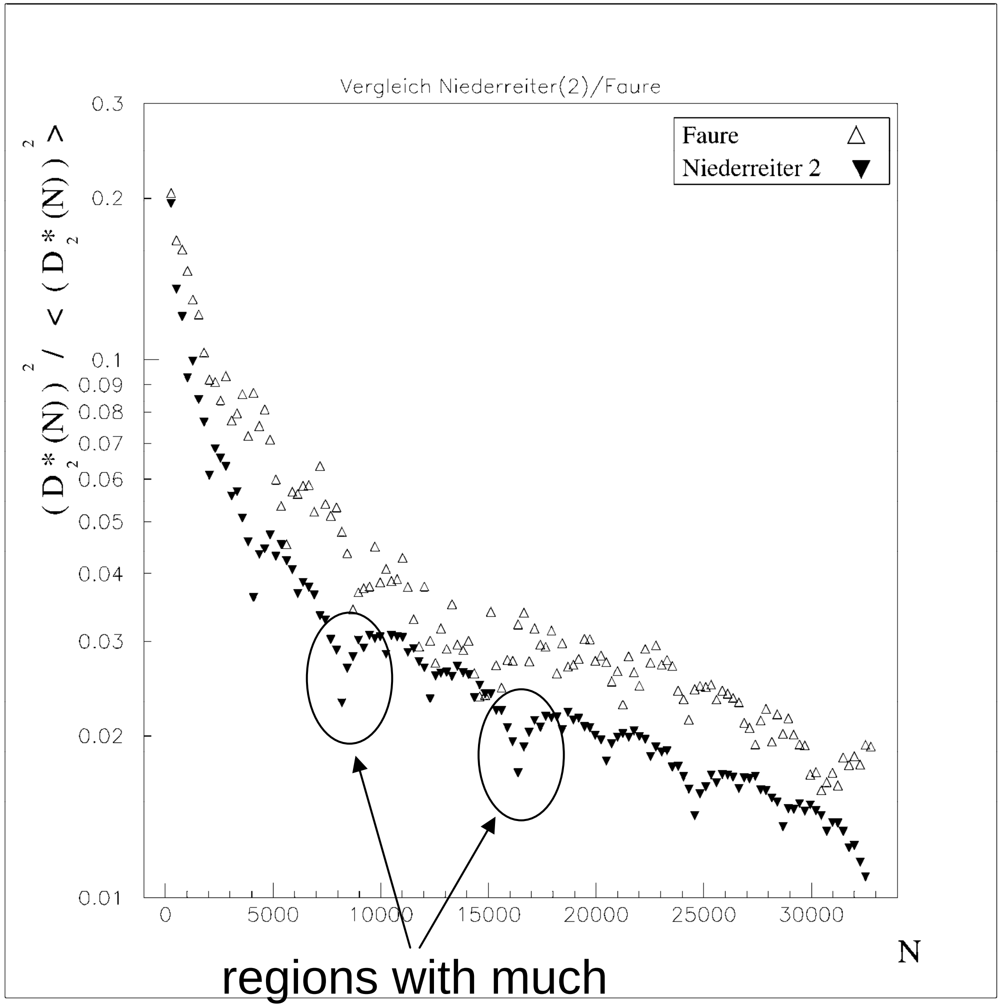
quasi-random



Generator examples



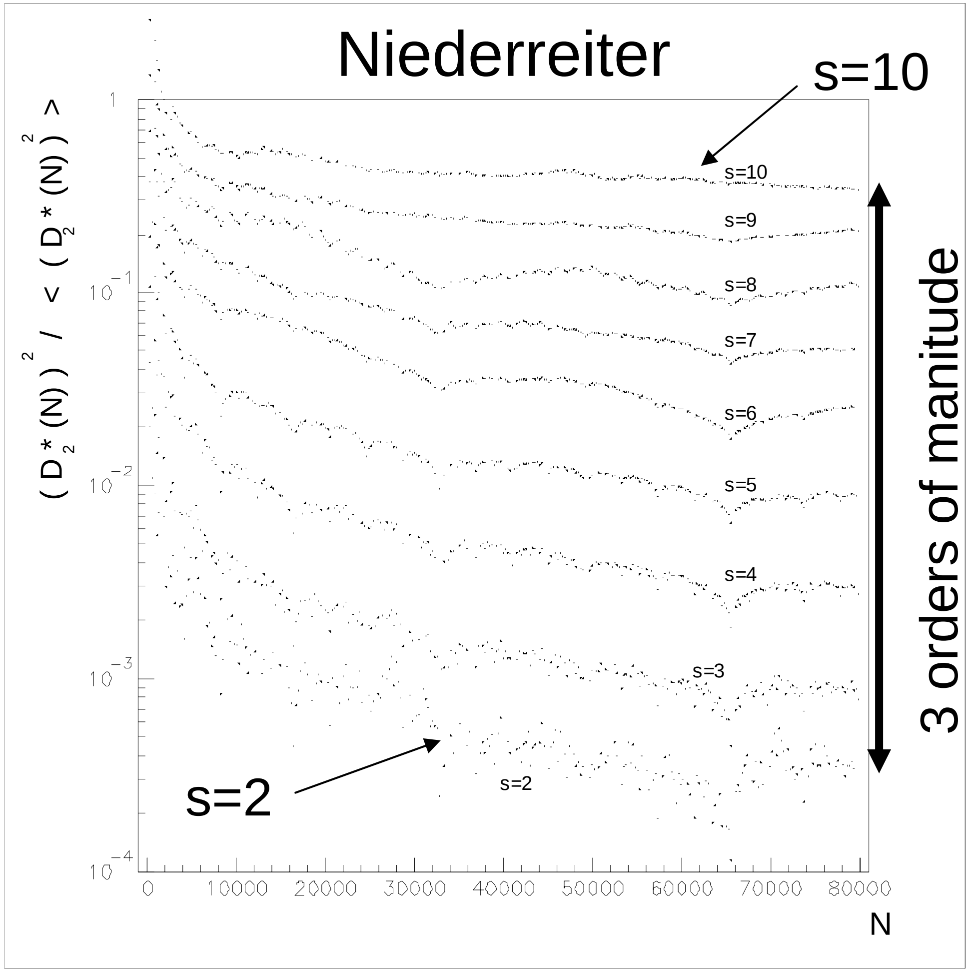
Optimising discrepancy



regions with much lower discrepancy

- Minima for $N = 2^k$
- "tune" discrepancy by carefully choosing needed N

several dimensions



compare discrepancy in several dimensions

→ quasi-random numbers good in few dimensions:

$$s \approx 2 \ldots 10$$

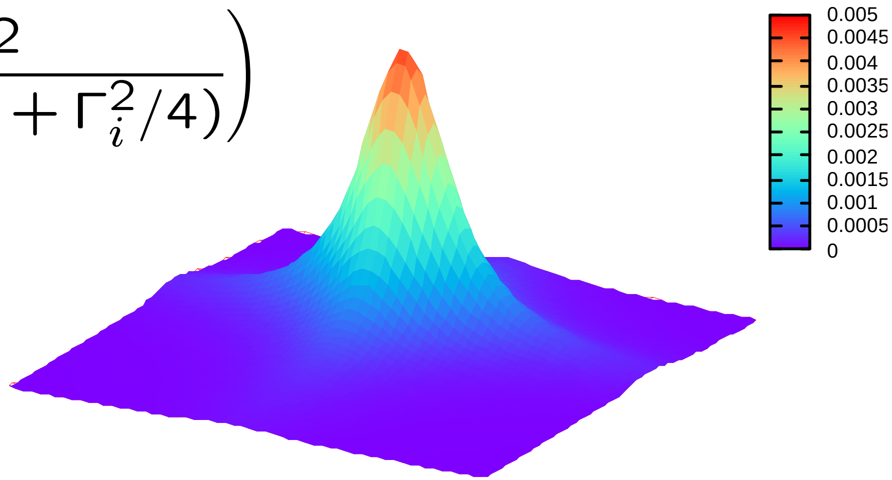
Example: Breit-Wigner

numerical integration of multi-dim. BW

$$\int \prod_{i=1}^{s} \left(\frac{\Gamma_i/2}{\pi((x-x_0)^2 + \Gamma_i^2/4)} \right)$$

e.g. 2 dimensions:

x_{01}	x_{02}	Γ_1	Γ_2
44.05	57.49	10.84	8.12

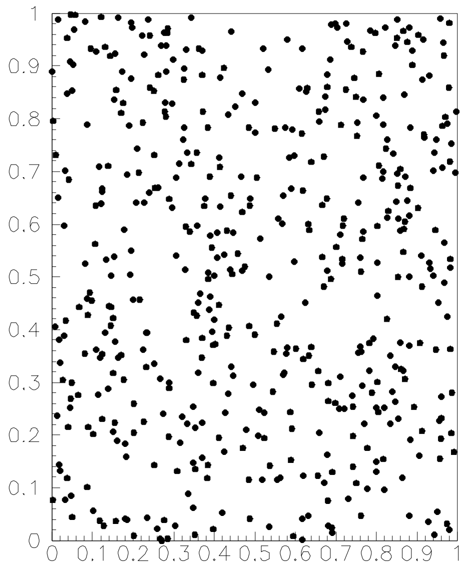


deviation from real integral value: ($8.8116863 \cdot 10^{-1}$)

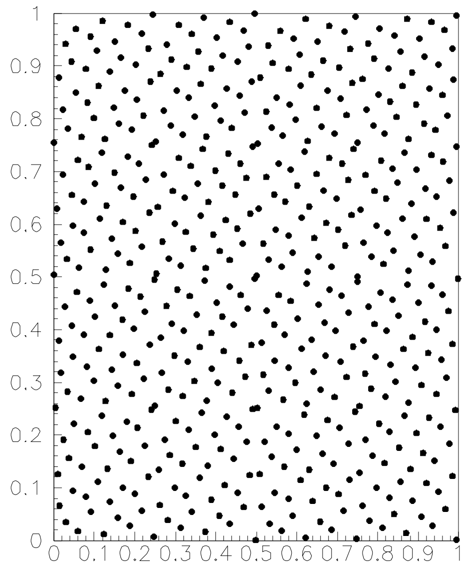
N	Pseudo	Halton	Faure	Sobol	Niederreiter	Hammersly	lattice
1024	$1.03 \cdot 10^{-1}$	$1.72 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$1.03 \cdot 10^{-2}$	$2.36 \cdot 10^{-3}$	$7.37 \cdot 10^{-4}$	$4.61 \cdot 10^{-4}$
8192	$3.60 \cdot 10^{-2}$	$1.98 \cdot 10^{-3}$	$1.94 \cdot 10^{-3}$	$6.20 \cdot 10^{-4}$	$1.73 \cdot 10^{-4}$	$1.24 \cdot 10^{-5}$	-----------
16384	$2.53 \cdot 10^{-2}$	$9.08 \cdot 10^{-4}$	$9.56 \cdot 10^{-4}$	$2.96 \cdot 10^{-4}$	$7.11 \cdot 10^{-5}$	$5.04 \cdot 10^{-6}$	$9.66 \cdot 10^{-5}$

Optimising numerical integration

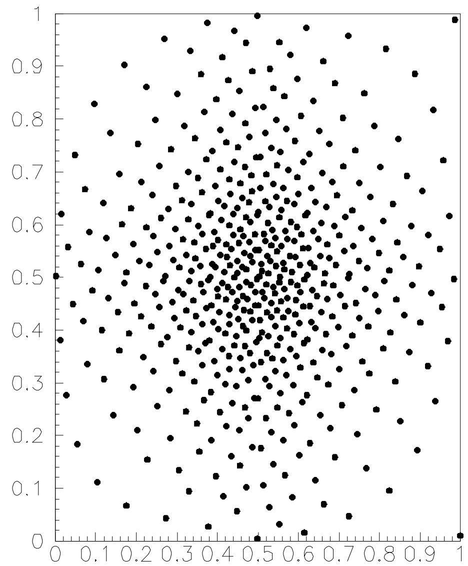
- <u>Example:</u> convolution of BW with resolution <u>per event</u> (B^0 mixing analysis)
  - use quasi-random numbers ⇒ fewer numbers N needed for evaluation
  - transform for optimal function sampling ⇒ importance-sampling



Pseudo random (RANLUX)



Quasi random (Niedereiter)



Quasi random transformed

Q-VEGAS

- VEGAS: package for numerical integration in several dimensions:
  - start with uniform intervals
  - evaluate function values in these intervals
  - iteratively adopt interval structure to shape of function

- Q-VEGAS: use quasi-random numbers instead of pseudo-random numbers:
  - faster and more accurate evaluation

Summary

- Low discrepancy (=evenly distributed) numbers important in numerical integration.
- Quasi-random numbers superior in few dimensions:
  - faster convergence
  - higher accuracy

  → $\epsilon \propto \frac{1}{N}$ instead of $\frac{1}{\sqrt{N}}$

- Wide range of applications, e.g. Q-Vegas