

Computer Physics

in Particle Physics

Peter Skands

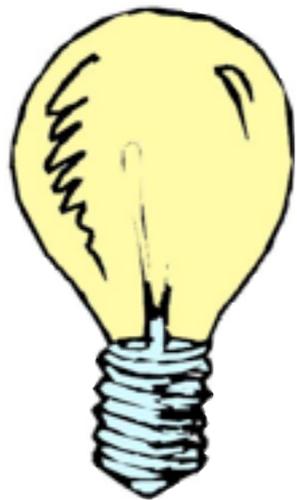
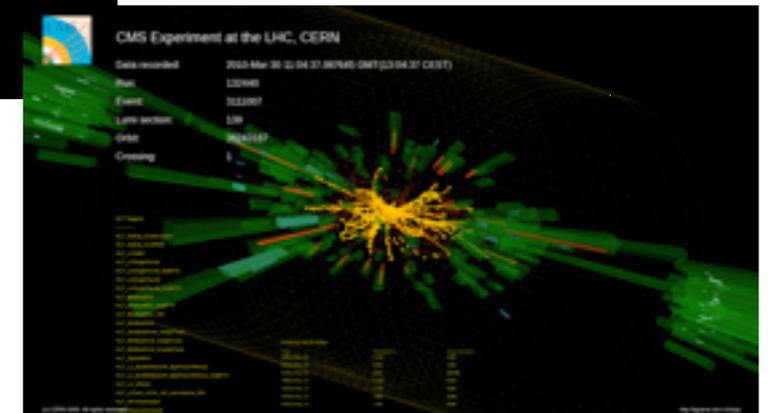
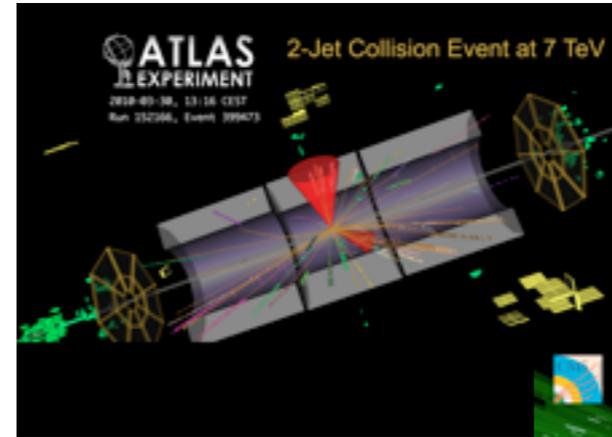
Theoretical Physics, CERN

"Nothing"

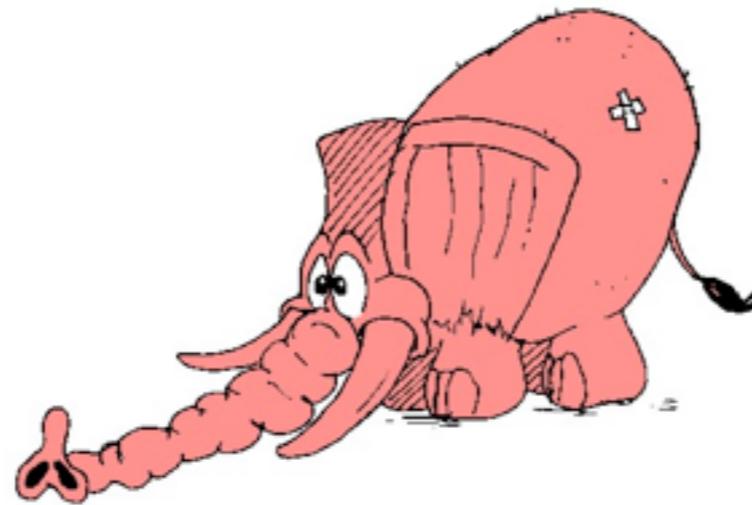
Glue action density: $2.4 \times 2.4 \times 3.6$ fm
QCD Lattice simulation from
D. B. Leinweber, hep-lat/0004025

Collider Physics

Comparisons
to Collider
observables

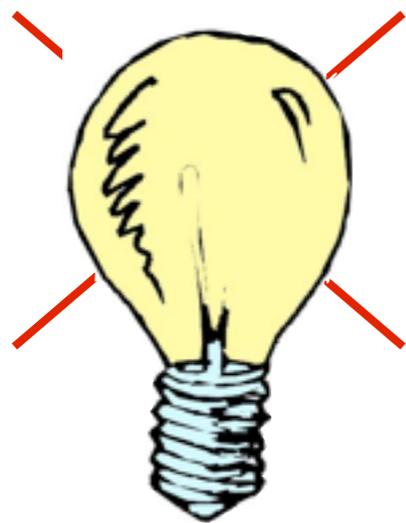
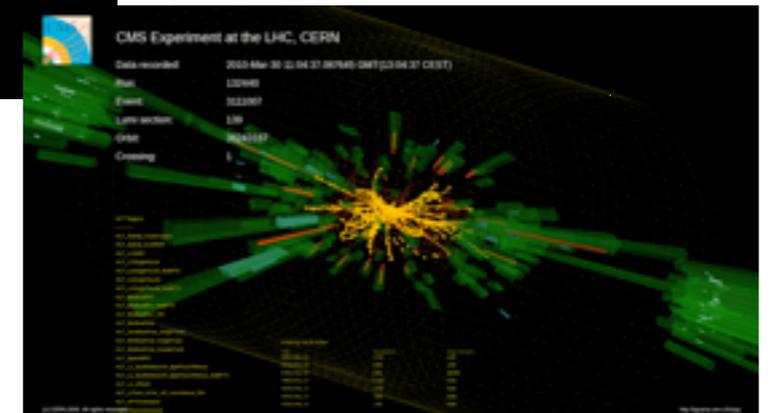
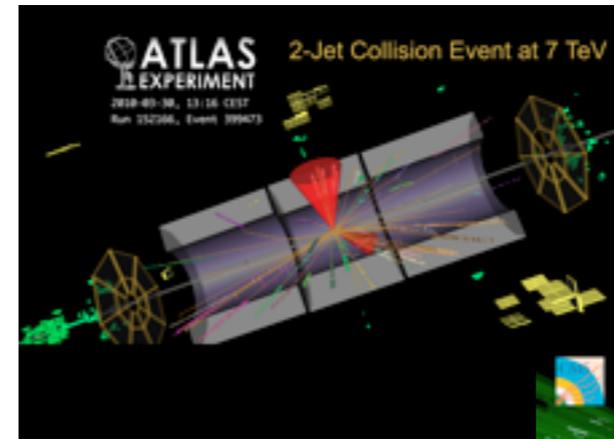


L=...

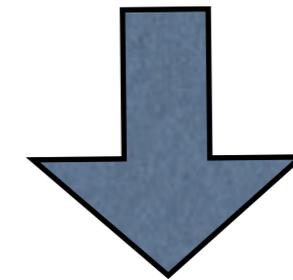
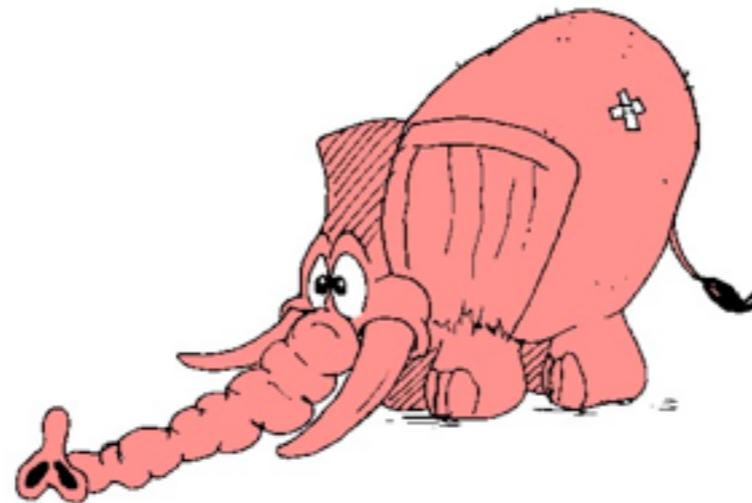


Collider Physics

Comparisons
to Collider
observables



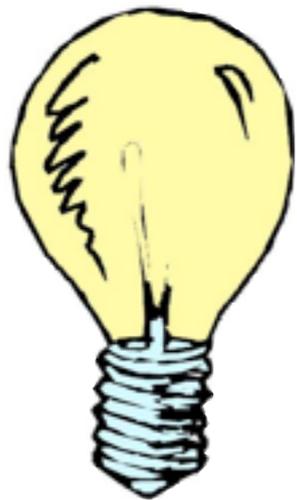
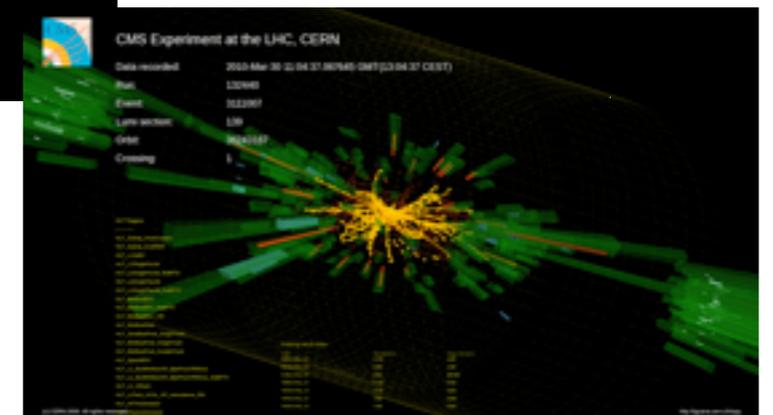
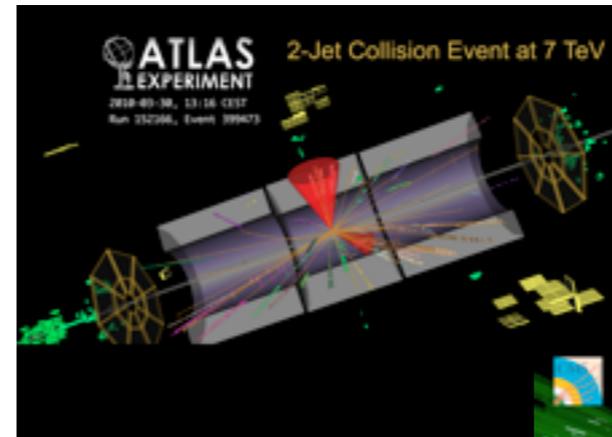
$L = \dots$



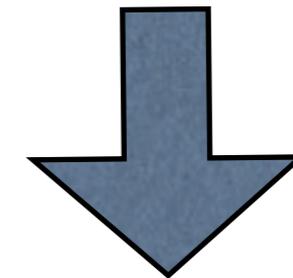
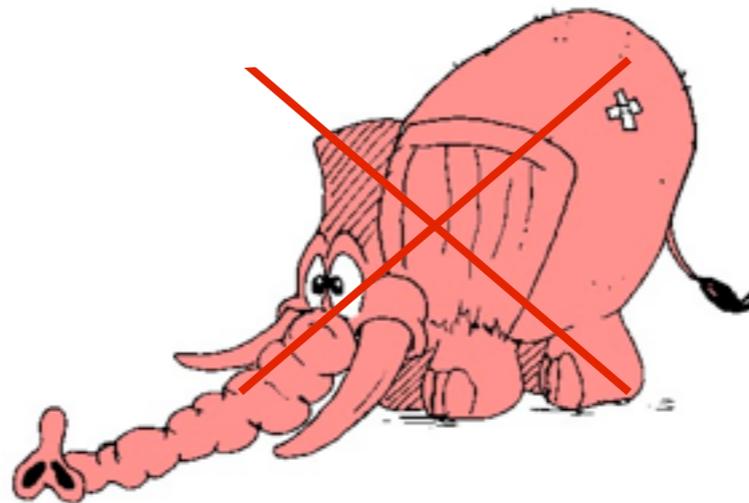
A) Theoretical Idea
is wrong

Collider Physics

Comparisons
to Collider
observables



L=...



A) Theoretical Idea

B) SM Physics Model
is wrong

Topics

Lecture 1:

Numerical Integration
Monte Carlo methods
Importance Sampling
The Veto Algorithm

+ on Friday

Practical Exercises:

PYTHIA 8 kickstart
(check the instructions)

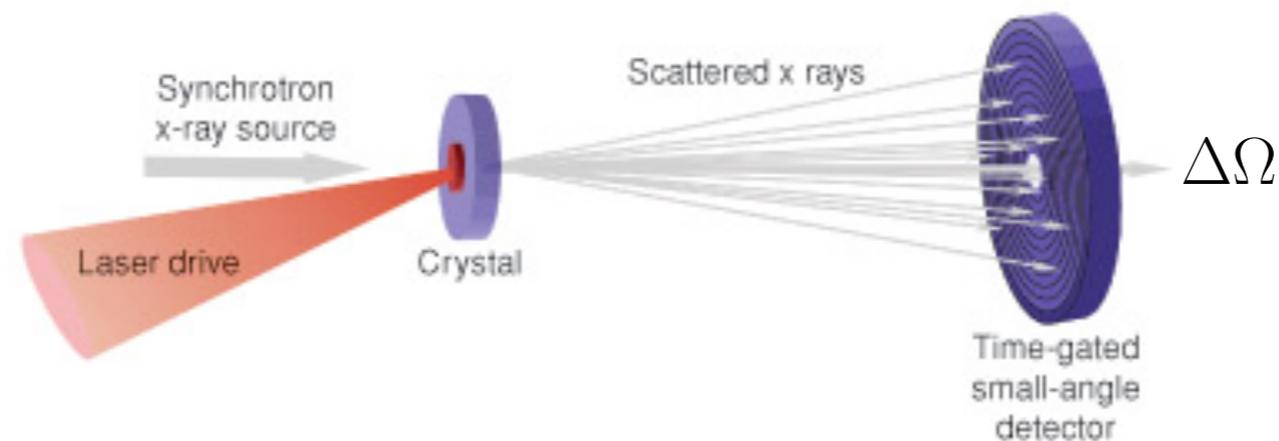
Lecture 2:

Application of these methods to simulations
of collider physics: Monte Carlo Event Generators

Why Integrals?

Think: scattering experiments

→ Integrate differential cross sections over specific phase space regions

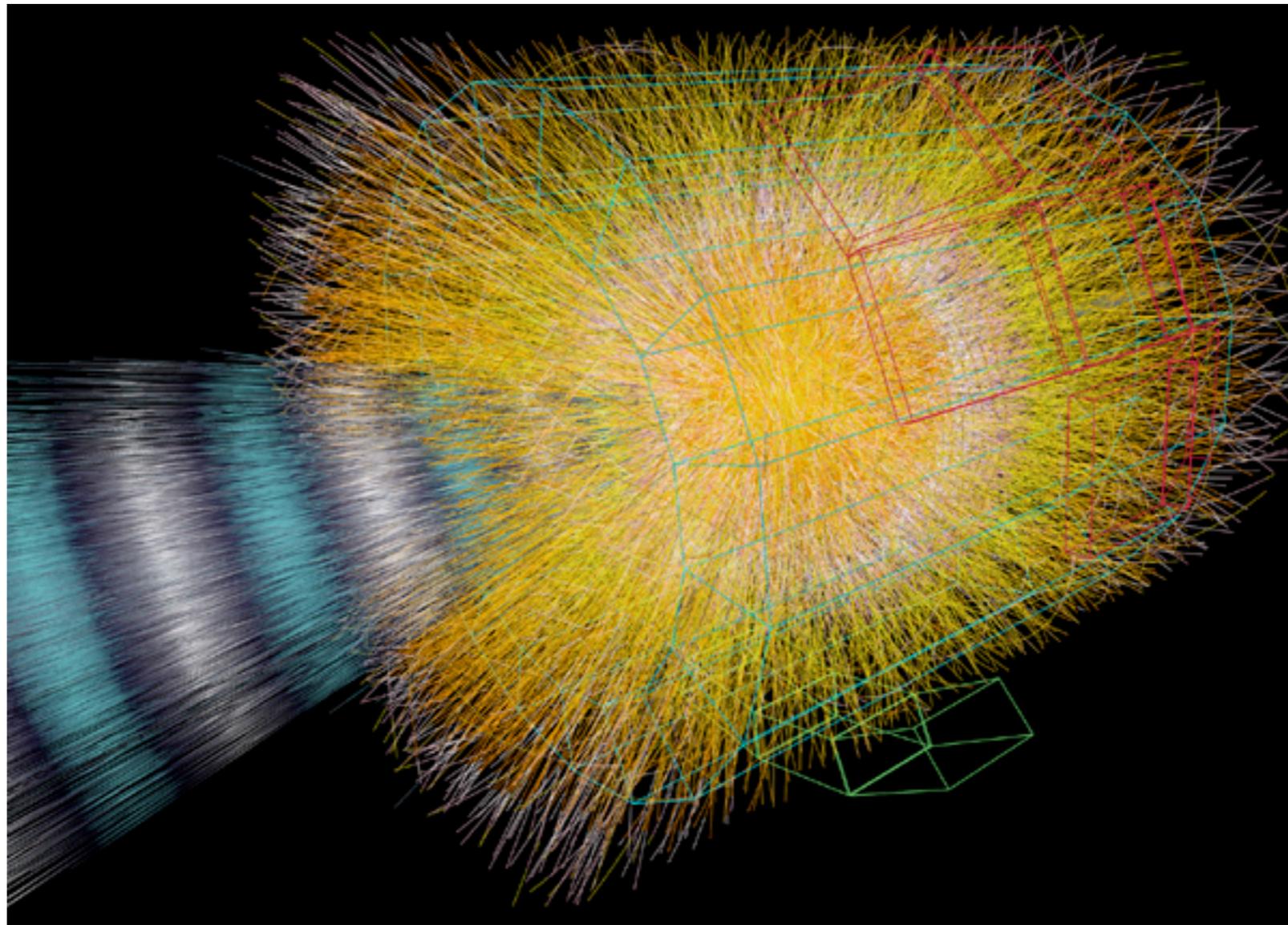


Predicted number of counts
= integral over solid angle

$$N_{\text{count}}(\Delta\Omega) \propto \int_{\Delta\Omega} d\Omega \frac{d\sigma}{d\Omega}$$

In particle physics:
sum (= integrate) over all quantum histories

ALICE Collision

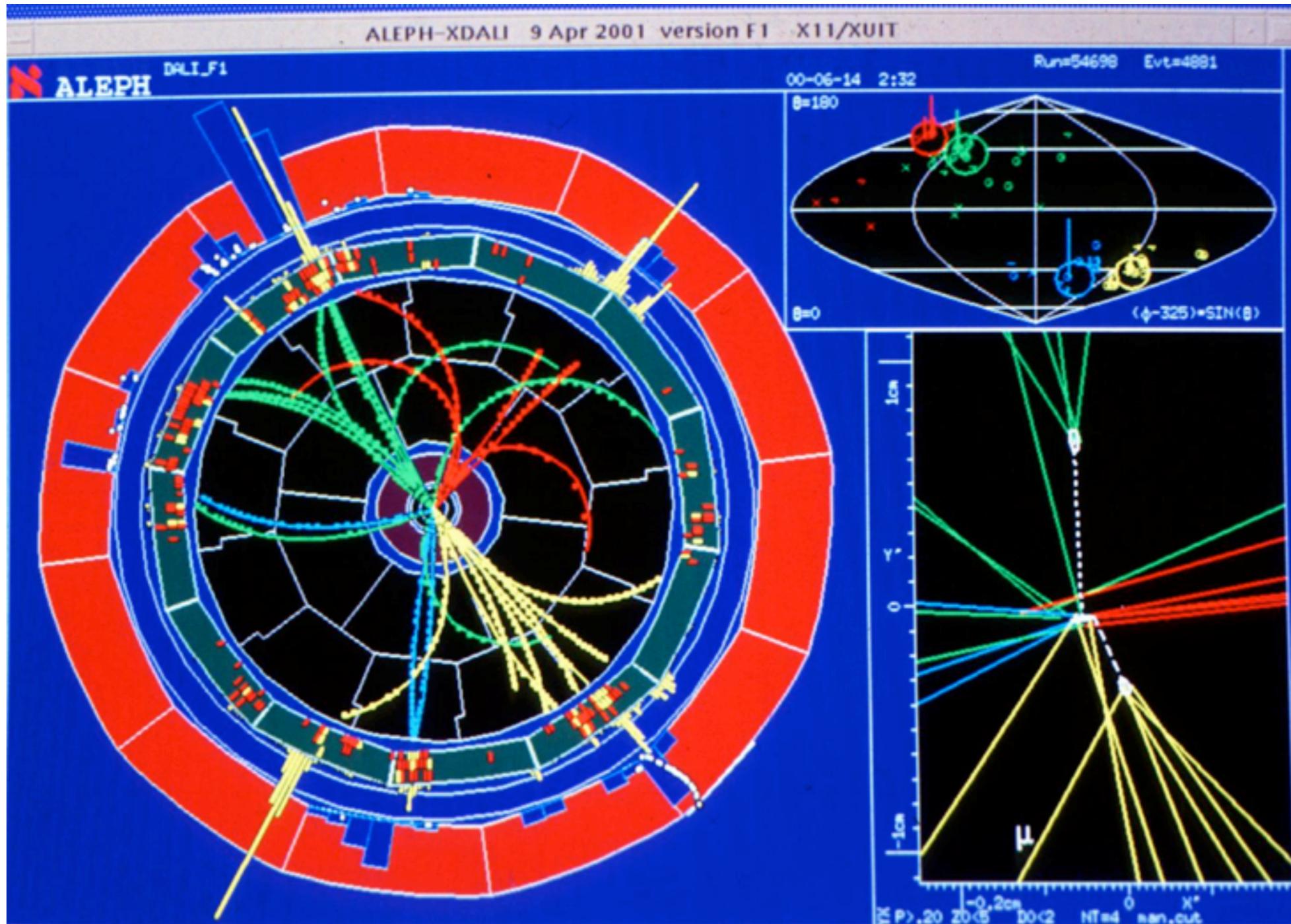


↪ More complicated integrals ...

Why Integrals?

Why Numerical?

Let's go back and look at something slightly simpler ...



4-jet event in ALEPH at LEP (a Higgs candidate)

Now compute the backgrounds ...

Why Numerical?

Part of $Z \rightarrow 4$ jets ...

5.3 Four-parton tree-level antenna functions

The tree-level four-parton quark-antiquark antenna contains three final states: quark-gluon-gluon-antiquark at leading and subleading colour, A_4^0 and \tilde{A}_4^0 and quark-antiquark-quark-antiquark for non-identical quark flavours B_4^0 as well as the identical-flavour-only contribution C_4^0 . The quark-antiquark-quark-antiquark final state with identical quark flavours is thus described by the sum of antennae for non-identical flavour and identical-flavour-only. The antennae for the $q\bar{q}g\bar{g}$ final state are:

$$A_4^0(1_q, 3_g, 4_g, 2_{\bar{q}}) = a_4^0(1, 3, 4, 2) + a_4^0(2, 4, 3, 1), \quad (5.27)$$

$$\tilde{A}_4^0(1_q, 3_g, 4_g, 2_{\bar{q}}) = \tilde{a}_4^0(1, 3, 4, 2) + \tilde{a}_4^0(2, 4, 3, 1) + \tilde{a}_4^0(1, 4, 3, 2) + \tilde{a}_4^0(2, 3, 4, 1), \quad (5.28)$$

$$\begin{aligned} a_4^0(1, 3, 4, 2) = & \frac{1}{s_{1234}} \left\{ \frac{1}{2s_{13}s_{24}s_{34}} [2s_{12}s_{14} + 2s_{12}s_{23} + 2s_{12}^2 + s_{14}^2 + s_{23}^2] \right. \\ & + \frac{1}{2s_{13}s_{24}s_{134}s_{234}} [3s_{12}s_{34}^2 - 4s_{12}^2s_{34} + 2s_{12}^3 - s_{34}^3] \\ & + \frac{1}{s_{13}s_{24}s_{134}} [3s_{12}s_{23} - 3s_{12}s_{34} + 4s_{12}^2 - s_{23}s_{34} + s_{23}^2 + s_{34}^2] \\ & + \frac{3}{2s_{13}s_{24}} [2s_{12} + s_{14} + s_{23}] + \frac{1}{s_{13}s_{34}} [4s_{12} + 3s_{23} + 2s_{24}] \\ & + \frac{1}{s_{13}s_{134}^2} [s_{12}s_{34} + s_{23}s_{34} + s_{24}s_{34}] \\ & + \frac{1}{s_{13}s_{134}s_{234}} [3s_{12}s_{24} + 6s_{12}s_{34} - 4s_{12}^2 - 3s_{24}s_{34} - s_{24}^2 - 3s_{34}^2] \\ & + \frac{1}{s_{13}s_{134}} [-6s_{12} - 3s_{23} - s_{24} + 2s_{34}] \\ & + \frac{1}{s_{24}s_{34}s_{134}} [2s_{12}s_{14} + 2s_{12}s_{23} + 2s_{12}^2 + 2s_{14}s_{23} + s_{14}^2 + s_{23}^2] \\ & + \frac{1}{s_{24}s_{134}} [-4s_{12} - s_{14} - s_{23} + s_{34}] + \frac{1}{s_{34}^2} [s_{12} + 2s_{13} - 2s_{14} - s_{34}] \\ & + \frac{1}{s_{34}^2s_{134}^2} [2s_{12}s_{14}^2 + 2s_{14}^2s_{23} + 2s_{14}^2s_{24}] - \frac{2s_{12}s_{14}s_{24}}{s_{34}^2s_{134}s_{234}} \\ & + \frac{1}{s_{34}^2s_{134}} [-2s_{12}s_{14} - 4s_{14}s_{24} + 2s_{14}^2] \\ & + \frac{1}{s_{34}s_{134}s_{234}} [-2s_{12}s_{14} - 4s_{12}^2 + 2s_{14}s_{24} - s_{14}^2 - s_{24}^2] \\ & + \frac{1}{s_{34}s_{134}} [-8s_{12} - 2s_{23} - 2s_{24}] + \frac{1}{s_{134}^2} [s_{12} + s_{23} + s_{24}] \\ & \left. + \frac{3}{2s_{134}s_{234}} [2s_{12} + s_{14} - s_{24} - s_{34}] + \frac{1}{2s_{134}} + \mathcal{O}(\epsilon) \right\}, \end{aligned} \quad (5.29)$$

$$\begin{aligned} \tilde{a}_4^0(1, 3, 4, 2) = & \frac{1}{s_{1234}} \left\{ \frac{1}{s_{13}s_{24}s_{134}s_{234}} \left[\frac{3}{2}s_{12}s_{34}^2 - 2s_{12}^2s_{34} + s_{12}^3 - \frac{1}{2}s_{34}^3 \right] \right. \\ & + \frac{1}{s_{13}s_{24}s_{134}} [3s_{12}s_{23} - 3s_{12}s_{34} + 4s_{12}^2 - s_{23}s_{34} + s_{23}^2 + s_{34}^2] \\ & + \frac{s_{12}^3}{s_{13}s_{24}(s_{13} + s_{23})(s_{14} + s_{24})} + \frac{1}{s_{13}s_{24}(s_{13} + s_{23})} \left[\frac{1}{2}s_{12}s_{14} + s_{12}^2 \right] \\ & + \frac{1}{s_{13}s_{24}(s_{14} + s_{24})} \left[\frac{1}{2}s_{12}s_{23} + s_{12}^2 \right] + \frac{1}{s_{13}s_{24}} \left[3s_{12} + \frac{3}{2}s_{14} + \frac{3}{2}s_{23} \right] \\ & + \frac{1}{s_{13}s_{134}^2} [s_{12}s_{34} + s_{23}s_{34} + s_{24}s_{34}] + \frac{2s_{12}^3}{s_{13}s_{134}s_{234}(s_{13} + s_{23})} \\ & + \frac{1}{s_{13}s_{134}s_{234}} [3s_{12}s_{34} - s_{24}s_{34} - 2s_{34}^2] \\ & + \frac{1}{s_{13}s_{134}(s_{13} + s_{23})} [s_{12}s_{24} + s_{12}s_{34} + 2s_{12}^2] \\ & + \frac{1}{s_{13}s_{134}} [-s_{23} - s_{24} + 2s_{34}] + \frac{1}{s_{13}s_{234}(s_{13} + s_{23})} [s_{12}s_{14} + s_{12}s_{34} + 2s_{12}^2] \\ & + \frac{1}{s_{13}s_{234}} [-2s_{12} - 2s_{14} + s_{24} + 2s_{34}] \\ & + \frac{2s_{12}^3}{s_{13}(s_{13} + s_{23})(s_{14} + s_{24})(s_{13} + s_{14})} \\ & + \frac{1}{s_{13}(s_{13} + s_{23})(s_{13} + s_{14})} [s_{12}s_{24} + 2s_{12}^2] \\ & + \frac{1}{s_{13}(s_{14} + s_{24})(s_{13} + s_{14})} [s_{12}s_{23} + 2s_{12}^2] \\ & + \frac{2s_{12}}{s_{13}(s_{13} + s_{14})} - \frac{2}{s_{13}} + \frac{1}{s_{134}^2} [s_{12} + s_{23} + s_{24}] \\ & \left. + \frac{1}{s_{134}s_{234}} [s_{12} - s_{34}] + \frac{1}{s_{134}} + \mathcal{O}(\epsilon) \right\}. \end{aligned} \quad (5.30)$$

First computed by K. Ellis, D. Ross, A. Terrano, Nucl.Phys.B178 (1981) 421
This version from Gehrmann-de-Ridder, Gehrmann, Glover, JHEP 0509(2005)056

Why Numerical?

The non-identical quark antenna is:

$$B_4^0(1_q, 3_{q'}, 4_{\bar{q}'}, 2_{\bar{q}}) = b_4^0(1, 3, 4, 2) + b_4^0(2, 3, 4, 1) + b_4^0(1, 4, 3, 2) + b_4^0(2, 4, 3, 1), \quad (5.37)$$

with a sub-antenna function given by

$$b_4^0(1, 3, 4, 2) = \frac{1}{s_{1234}} \left\{ \frac{1}{s_{34}^2 s_{134}^2} [s_{12}s_{13}s_{14} + s_{13}s_{14}s_{23} - s_{13}^2 s_{24}] \right. \\ + \frac{1}{s_{34}^2 s_{134} s_{234}} [-s_{12}s_{13}s_{24} + s_{13}s_{14}s_{23} - s_{13} s_{24}^2] + \frac{1}{s_{34} s_{134}^2} [s_{12}s_{13} + s_{13}s_{23}] \\ \left. + \frac{1}{2s_{34} s_{134} s_{234}} [2s_{12}s_{13} + s_{12}^2] + \frac{s_{12}}{2s_{134} s_{234}} + \mathcal{O}(\epsilon) \right\}. \quad (5.38)$$

The identical-flavour-only quark-antiquark-quark-antiquark antenna is:

$$C_4^0(1_q, 3_q, 4_{\bar{q}}, 2_{\bar{q}}) = c_4^0(1, 2, 3, 4) + c_4^0(1, 4, 3, 2), \quad (5.42)$$

$$c_4^0(1, 2, 3, 4) = \frac{1}{s_{1234}} \left\{ -\frac{s_{12}s_{13}s_{14}}{2s_{23}s_{34}s_{123}s_{134}} + \frac{1}{2s_{23}s_{34}s_{134}s_{234}} [-s_{12}s_{13}s_{24} + s_{13}s_{14}s_{24}] \right. \\ - \frac{s_{13}s_{24}^2}{2s_{23}s_{34}s_{234}^2} - \frac{s_{12}s_{13}}{s_{23}s_{123}s_{134}} \\ + \frac{1}{2s_{23}s_{123}s_{234}} [-s_{12}s_{14} - s_{12}s_{34} - s_{12}^2 + s_{13}s_{24}] \\ + \frac{1}{2s_{23}s_{134}s_{234}} [s_{12}s_{14} + s_{12}s_{34} + s_{12}^2 + s_{13}s_{24}] - \frac{s_{13}}{2s_{123}s_{134}} \\ \left. + \frac{1}{s_{23}s_{234}^2} [s_{12}s_{24} + s_{14}s_{24}] + \frac{1}{2s_{123}s_{234}} [-s_{12} + s_{14}] + \mathcal{O}(\epsilon) \right\}. \quad (5.43)$$

Integrate over 4-particle phase space ...

This is one of the simplest processes ... computed at lowest order in the theory.

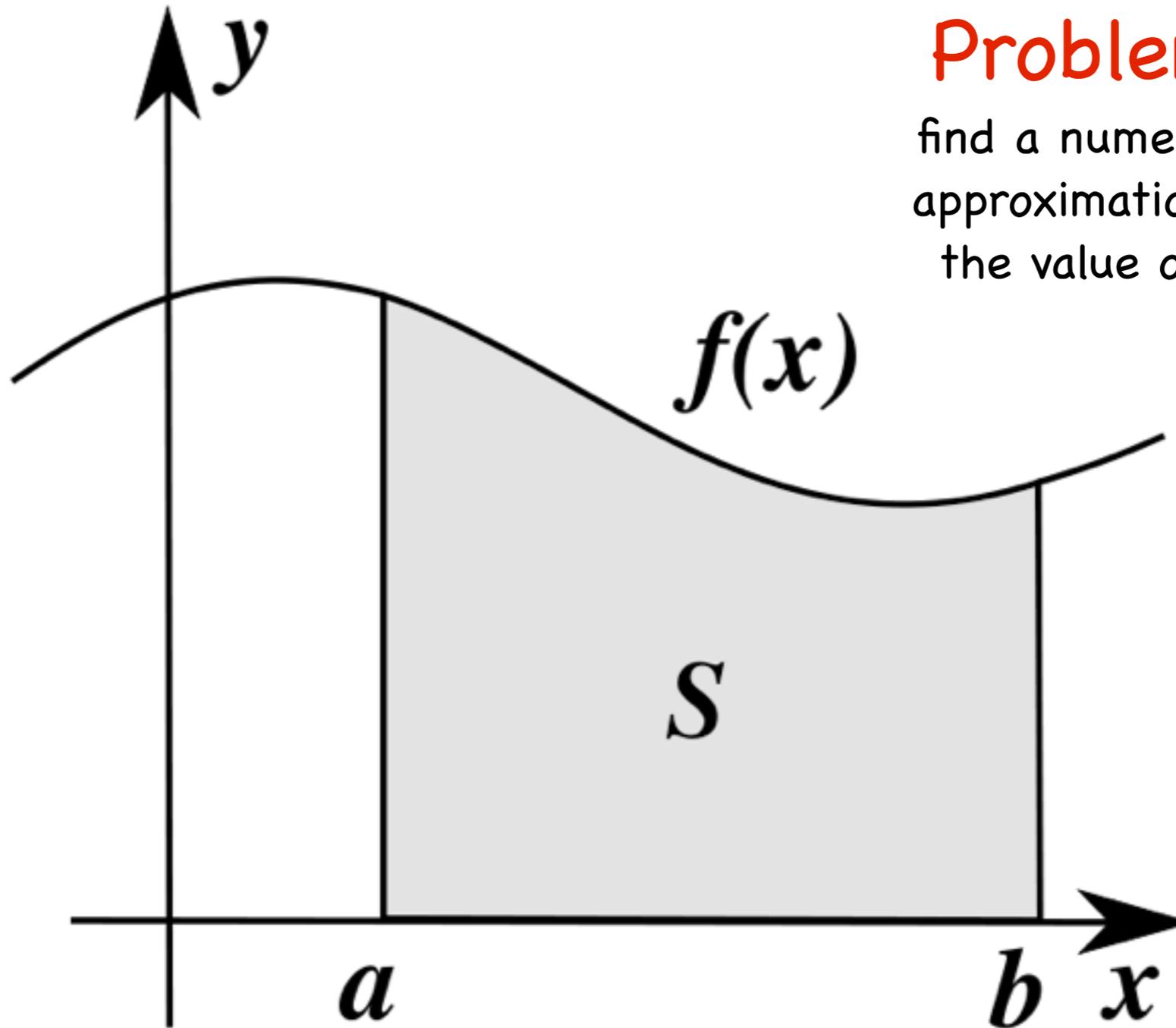
Now compute the quantum corrections: $Z \rightarrow 5, 6, \dots$

And higher orders of quantum fluctuations (quantum loops) ...

And hadronization, hadron decays, detector response, ...

Recap

Numerical Integration



Problem:

find a numerical approximation to the value of S

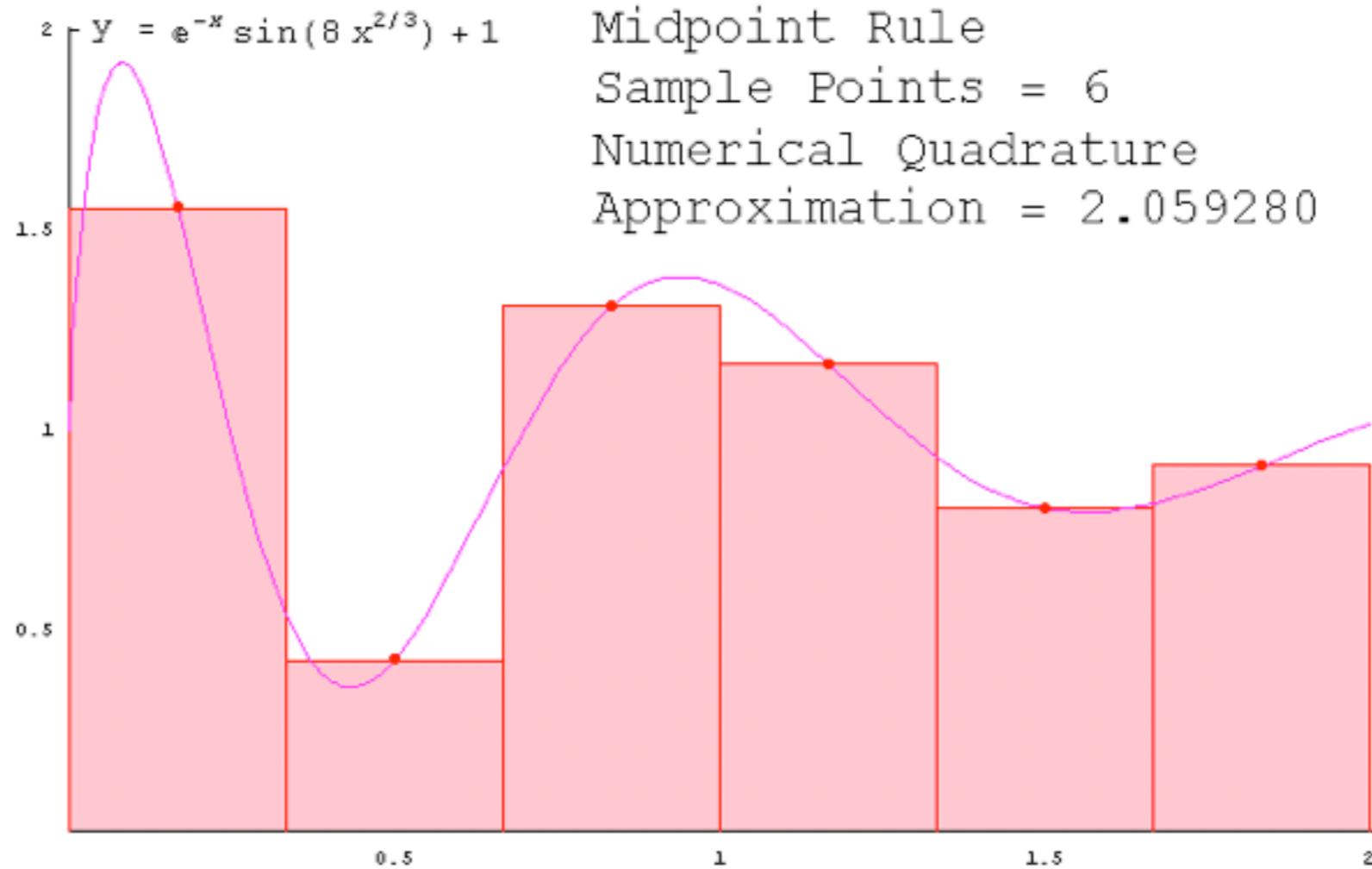
Recap

Riemann Sums

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(t_i)(x_{i+1} - x_i)$$



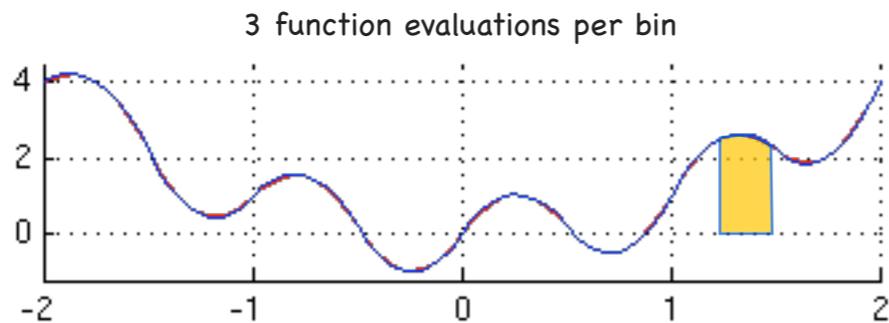
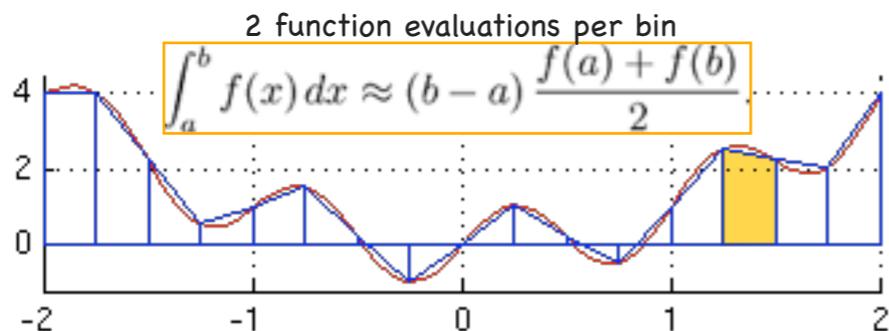
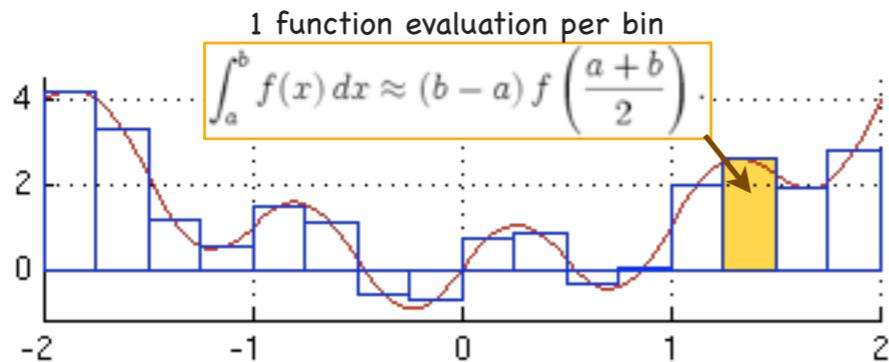
B. Riemann, (1826-1866)



Recap

Numerical Integration in 1D

Fixed-Grid n-point Quadrature Rules:



Midpoint (rectangular) Rule:

Divide into N "bins" of size Δ

Approximate $f(x) \approx$ constant in each bin

Sum over all **rectangles** inside your region

Trapezoidal Rule:

Approximate $f(x) \approx$ linear in each bin

Sum over all **trapeziums** inside your region

Simpson's Rule:

Approximate $f(x) \approx$ quadratic in each bin

Sum over all **simpsons** inside your region

etc ...

Convergence Rate

The most important question:

How long do I have to wait?

(How many points do I need for a given precision)?

Uncertainty as a function of number of points	$n_{\text{eval}} / \text{bin}$	Approx Conv. Rate (in 1D)
Trapezoidal Rule (2-point)	2	$1/n^2$
Simpson's Rule (3-point)	3	$1/n^4$
... m-point (Gauss quadrature)	m	$1/n^{2m-1}$

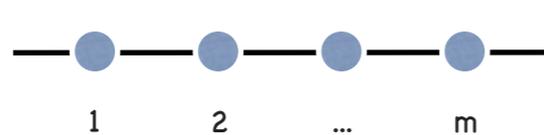
See, e.g., Numerical Recipes

See, e.g., F. James, "Monte Carlo Theory and Practice"

Higher Dimensions

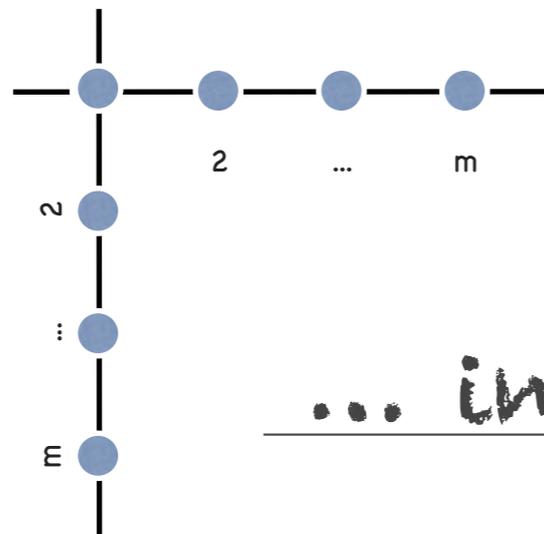
Fixed-Grid (Product) Rules scale exponentially with D

N-point rule in 1 dimension



→ m function evaluations per bin

... in 2 dimensions



→ m^2 evaluations per bin

... in D dimensions → N^D per bin

E.g., to evaluate a 12-point rule in 10 dimensions,
need 1000 billion evaluations per bin

Convergence Rate

+ Convergence is slower in higher Dimensions!

→ More points for less precision



Uncertainty as a function of number of points	$n_{\text{eval}} / \text{bin}$	Approx Conv. Rate (in D dim)
Trapezoidal Rule (2-point)	2^D	$1/n^{2/D}$
Simpson's Rule (3-point)	3^D	$1/n^{4/D}$
... m-point (Gauss rule)	m^D	$1/n^{(2m-1)/D}$

See, e.g., Numerical Recipes

See, e.g., F. James, Monte Carlo Theory and Practice

Monte Carlo

A Monte Carlo technique: is any technique making use of random numbers to solve a problem

Convergence:

Calculus: $\{A\}$ converges to B
if an n exists for which
 $|A_{i>n} - B| < \epsilon$, for any $\epsilon > 0$

Monte Carlo: $\{A\}$ converges to B
if n exists for which
the probability for
 $|A_{i>n} - B| < \epsilon$, for any $\epsilon > 0$,
is $> P$, for any $P[0 < P < 1]$

“This risk, that **convergence is only given with a certain probability**, is inherent in Monte Carlo calculations and is the reason why this technique was named after the world’s most famous gambling casino. Indeed, the name is doubly appropriate because the **style of gambling** in the Monte Carlo casino, not to be confused with the noisy and tasteless gambling houses of Las Vegas and Reno, is serious and sophisticated.”

*F. James, “Monte Carlo theory and practice”,
Rept. Prog. Phys. 43 (1980) 1145*

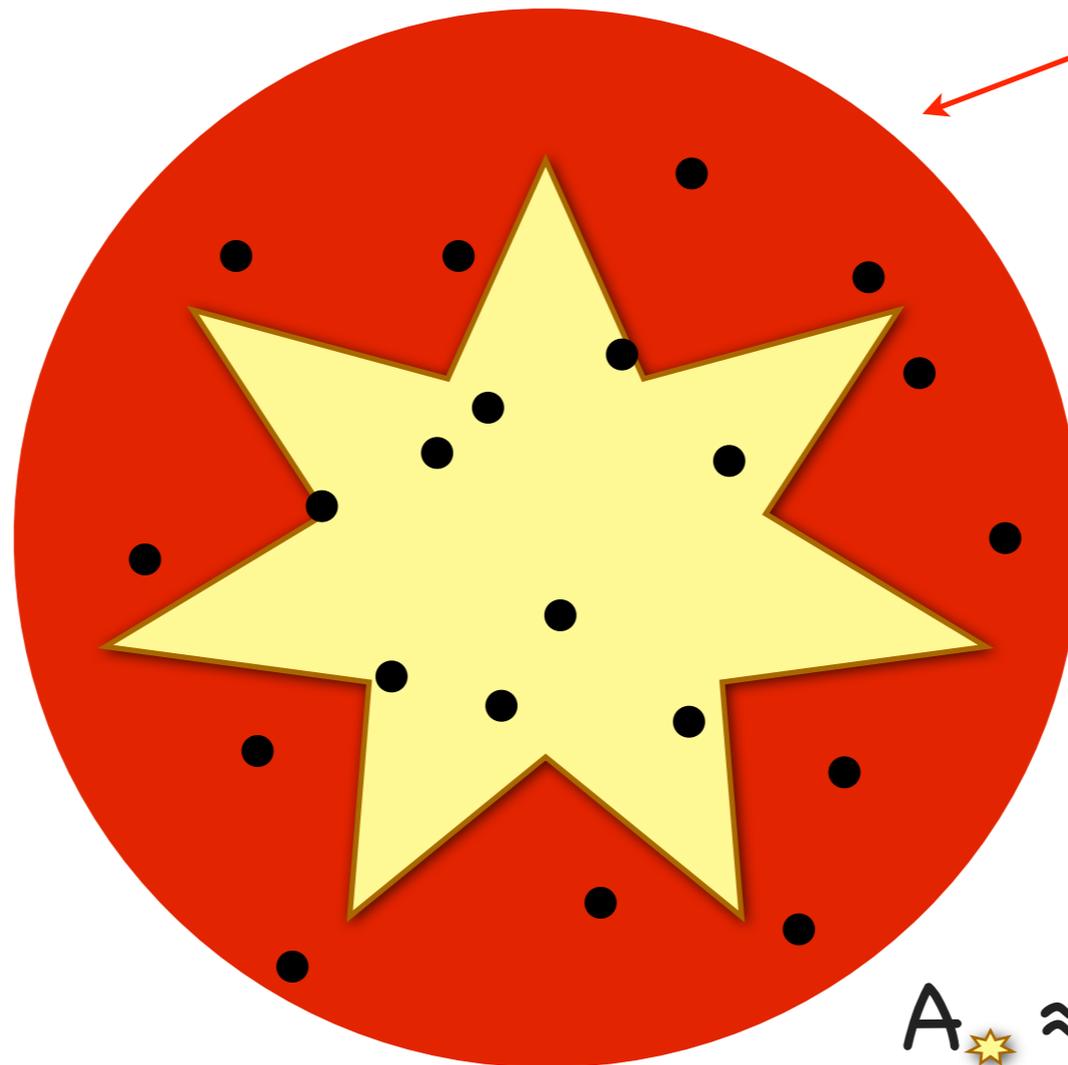
Random Numbers and Monte Carlo

Example 1: simple function (=constant); complicated boundary

Example: you want to know the area of this shape:

Now get a few friends, some balls, and throw random shots inside the circle
(PS: be careful to make your shots truly random)

Count how many shots hit the shape inside and how many miss



Assume you know the area of this shape:

πR^2
(an overestimate)



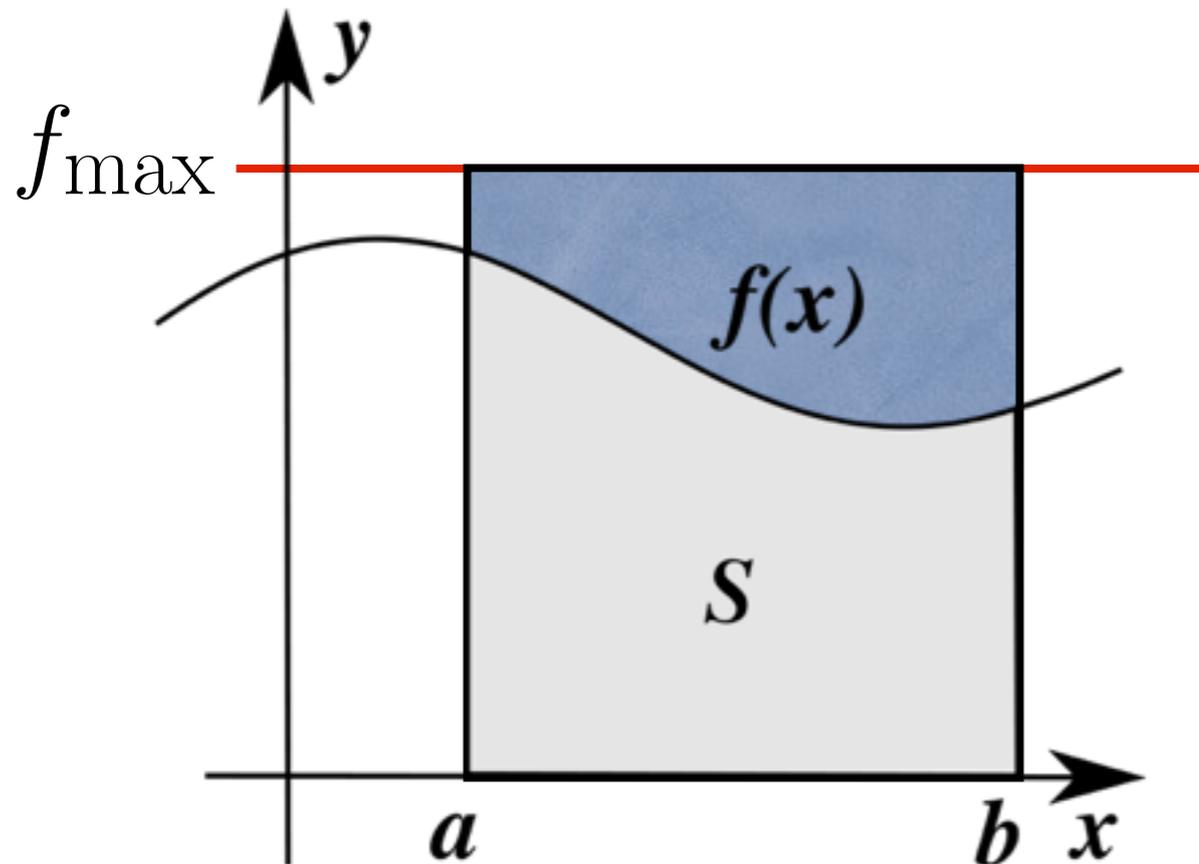
Earliest Example of MC calculation: Buffon's Needle (1777) to calculate π

G. Leclerc, Comte de Buffon (1707-1788)

$$A_{\star} \approx N_{\text{hit}} / N_{\text{miss}} \times \pi R^2$$

Random Numbers and Monte Carlo

Example 2: complicated function; simple boundary



Start from **overestimate**,

$$f_{\max}$$

Generate uniformly distributed random points between a and b

$$\frac{f(x_i)}{f_{\max}} = P_{\text{hit}}$$

The integral is then \approx

$$(b - a) f_{\max} \frac{1}{n} \sum_{i=1}^n \frac{f(x_i)}{f_{\max}}$$

area of rectangle

fraction that 'hit'

Justification

1. Law of large numbers

For a function, f , of random variables, x_i ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(x_i) = \frac{1}{b-a} \int_a^b f(x) dx$$

Monte Carlo Estimate

The Integral

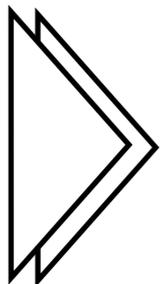


For infinite n:
Monte Carlo is a
consistent
estimator

2. Central Limit theorem

The sum of n independent random variables (of finite expectations and variances) is asymptotically Gaussian

(no matter how the individual random variables are distributed)



For finite n:

The Monte Carlo estimate is Gauss distributed around the true value

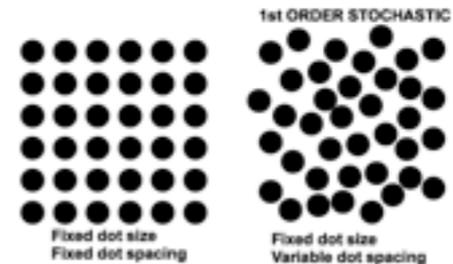
Convergence

MC = Monte Carlo

MC convergence is Stochastic!

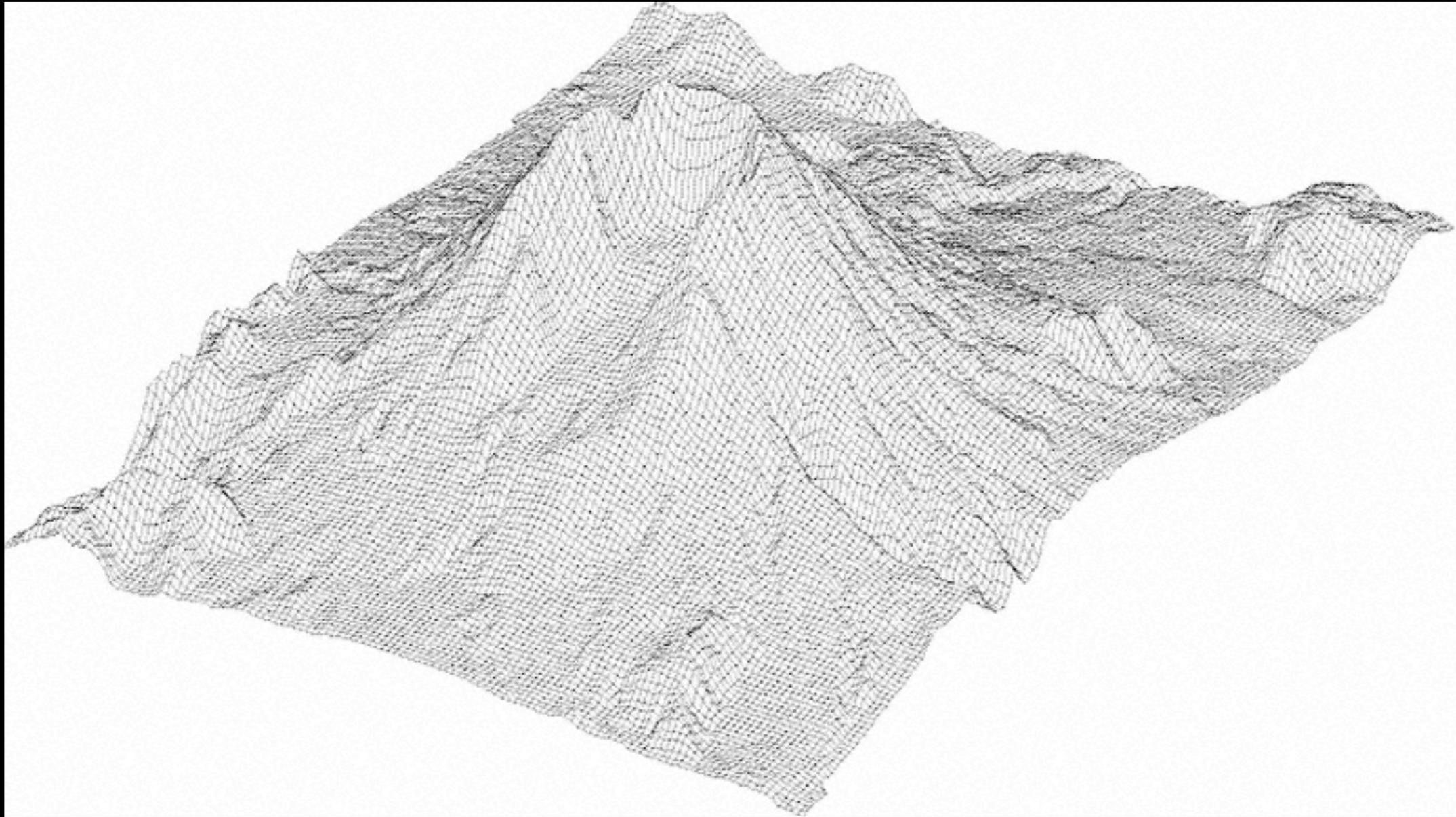
$\frac{1}{\sqrt{n}}$ in any dimension

+ can re-use previously generated points (\approx nesting)

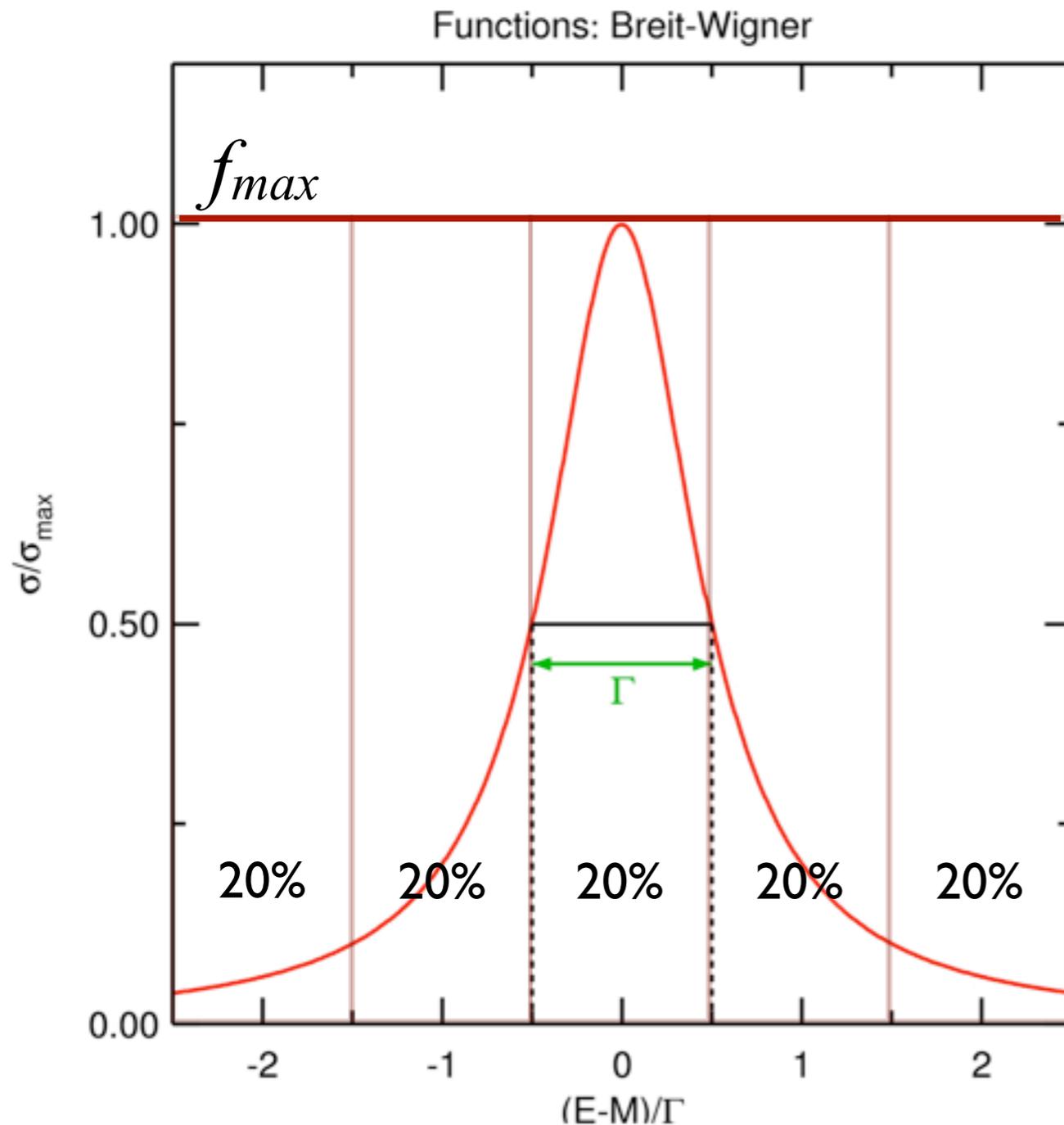


Uncertainty as a function of number of points	n_{eval} / bin	Approx Conv. Rate (in 1D)	Approx Conv. Rate (in D dim)
Trapezoidal Rule (2-point)	2^D	$1/n^2$	$1/n^{2/D}$
Simpson's Rule (3-point)	3^D	$1/n^4$	$1/n^{4/D}$
... m-point (Gauss rule)	m^D	$1/n^{2m-1}$	$1/n^{(2m-1)/D}$
Monte Carlo	1	$1/n^{1/2}$	$1/n^{1/2}$

Importance Sampling



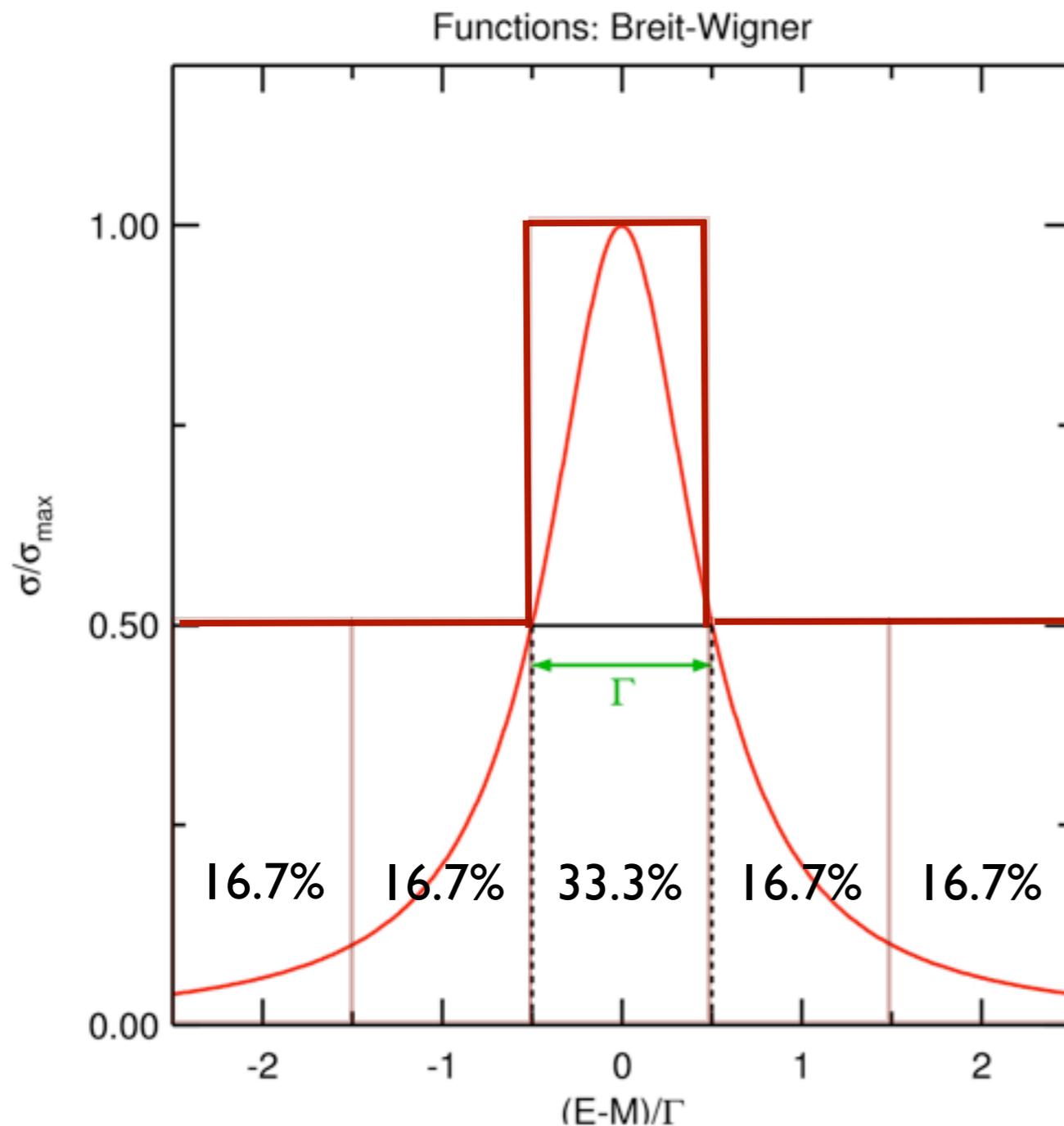
Peaked Functions



Precision on integral dominated by the points with $f \approx f_{max}$ (i.e., peak regions)

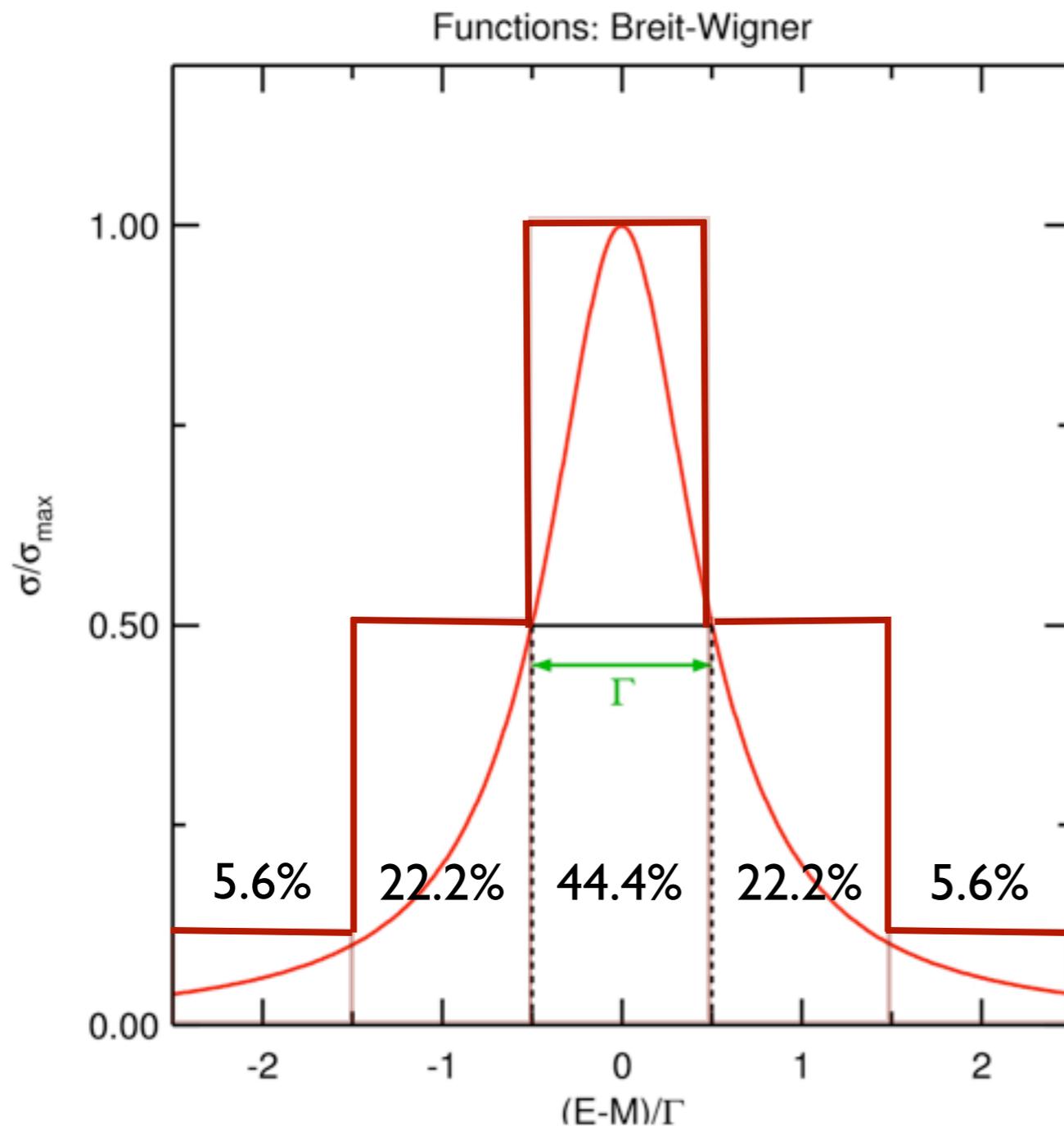
→ slow convergence if high, narrow peaks

Stratified Sampling



- make it twice as likely to throw points in the peak
- faster convergence for same number of function evaluations

Adaptive Sampling

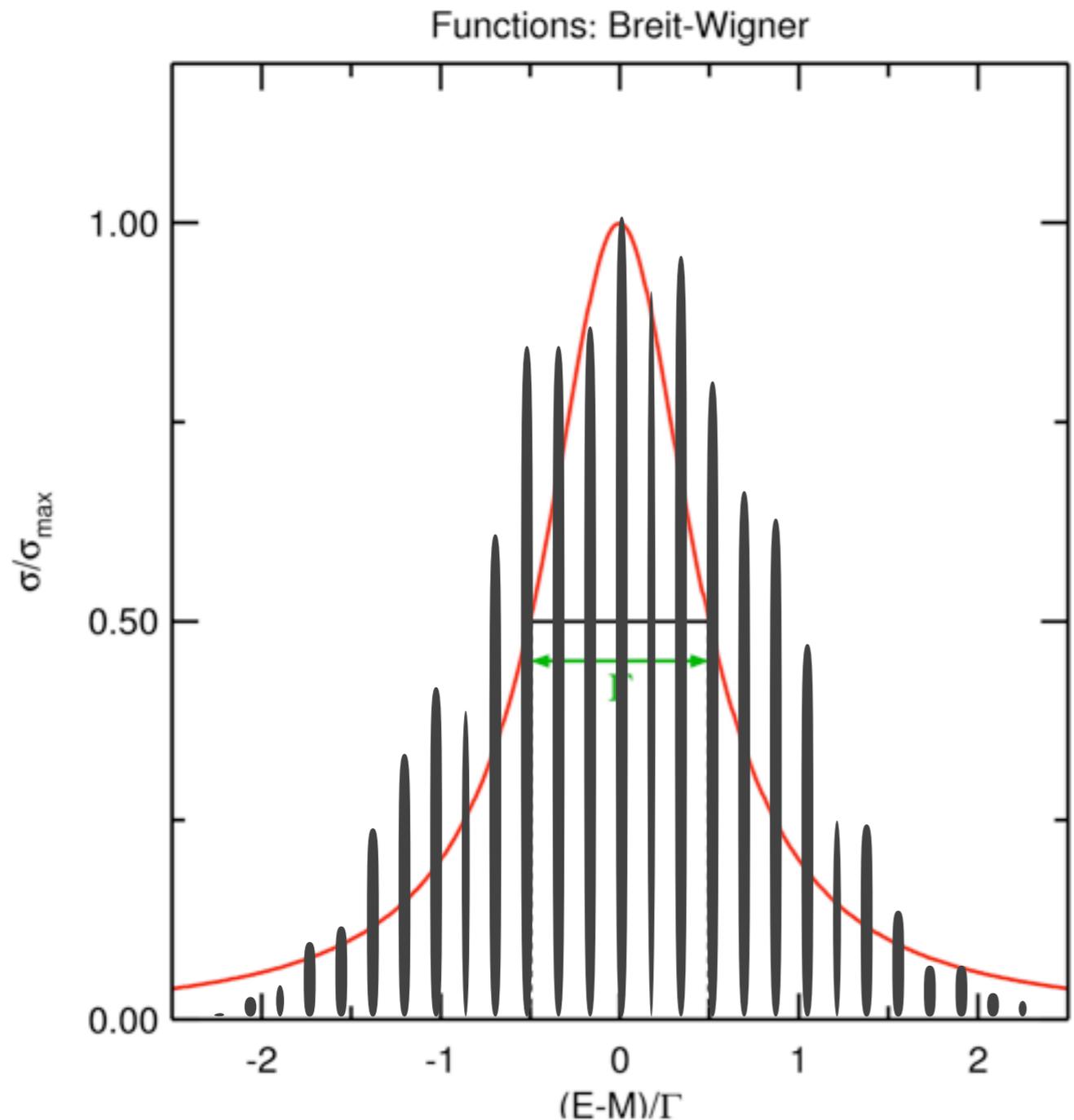


→ can even design algorithms that do this automatically as they run

→ Adaptive sampling

Importance Sampling

E.g., VEGAS algorithm, by G. Lepage



→ or throw points according to some smooth peaked function for which you have, or can construct, a random number generator (here: Gauss)

Why does this work?

- 1) You are inputting knowledge: obviously need to know where the peaks are to begin with ... (say you know, e.g., the location and width of a resonance)
- 2) Stratified sampling increases efficiency by combining n-point quadrature with the MC method, with further gains from adaptation
- 3) Importance sampling:

$$\int_a^b f(x)dx = \int_a^b \frac{f(x)}{g(x)}dG(x)$$

Effectively does flat MC with changed integration variables

Fast convergence if $f(x)/g(x) \approx 1$

The Veto Algorithm

Hit



Miss

How we do Monte Carlo

1. Take your system

- Set of radioactive nuclei
- Set of hard scattering processes
- Set of resonances that are going to decay
- Set of particles coming into your detector
- Set of cosmic photons traveling across the galaxy
- Set of molecules
- ...



How we do Monte Carlo

1. Take your system

2. Generate a "trial" (event/decay/interaction/...)

- Not easy to generate random numbers distributed according to exactly the right distribution?
- May have complicated dynamics, interactions ...
- → use a simpler "trial" distribution

- Flat with some stratification
- Or importance sample with simple overestimating function (for which you can generate random #s)

How we do Monte Carlo

Sounds deceptively simple, but ...
with it, you can integrate
arbitrarily complicated functions (in particular chains of nested functions),
over arbitrarily complicated regions,
in arbitrarily many dimensions ...

1. Take your system

2. Generate a "trial"

- **Accept trial** with probability $f(x)/g(x)$
 - $f(x)$ contains all the complicated dynamics
 - $g(x)$ is the simple trial function
- **If accept:** replace with new system state
- **If reject:** keep previous system state

no dependence on g in final result
- only affects convergence rate

And keep going: generate next trial ...



Example: Number of students who will get run over on their way here during the next 3 weeks

Complicated Function:

Time-dependent

Traffic density during day, week-days vs week-ends

(i.e., simulates non-trivial time evolution of system)

No two students are the same

Need to compute probability for each and sum

(i.e., simulates having several distinct types of "evolvers")

Multiple outcomes:

Hit → keep walking, or go to hospital?

Multiple hits = Product of single hits, or more complicated?

Monte Carlo Approach

Approximate Traffic

Simple overestimate:

highest recorded density
of most careless drivers,
driving at highest recorded speed

etc. (If this becomes too slow (computing time), try more clever "stratifications", adaptations, and/or importance sampling)



Approximate Student

by most accident-prone L- and R-hand traffic student
(overestimate)

Hit Generator

Off we go...

Throw random accidents according to:

$$R = \int_{t_0}^{t_e} dt \int dx \sum_{i=1}^{n_{\text{stud}}} \alpha_i(x, t) \rho_i(x, t) \rho_c(x, t)$$

Student-Car Coupling Density of Student i Density of Cars

Sum over students (possibly weighted by speed × drunkenness)

t_e : time of accident

Stratification

$$R = (t_e - t_0) \left(\alpha_{L, \max} N_L + \alpha_{R, \max} N_R \right) \rho_{c \max}$$

Coupling of most accident-prone left-hand-traffic student Coupling of most accident-prone right-hand-traffic student Rush-hour density of cars

Too Difficult

Simple Overestimate

Hit Generator

Trial Generator: (generate t_e)

$$R = (t_e - t_0) (\alpha_{L,\max} N_L + \alpha_{R,\max} N_R) \rho_{c\max}$$

t_e : time
of accident

Coupling of
most accident-prone
left-hand-traffic student

Coupling of
most accident-prone
right-hand-traffic student

Rush-hour density
of cars

Simple
Overestimate

(Also generate trial x_e , uniformly between Travelers lodge and Stias)

Accept with probability

$$P_{\text{accept}} = \frac{\alpha_i(x, t) \rho_i(x, t) \rho_c(x, t)}{(\alpha_{L,\max} N_L + \alpha_{R,\max} N_R) \rho_{c\max}}$$

→ True integral = number of accepted hits
(note: we didn't really trit multiple hits ... → Markov Chain)

Summary

Quantum Scattering Problems are common to many areas of physics:
To compute expectation value of observable: integrate over phase space

Complicated functions → Numerical Integration

High Dimensions → Monte Carlo (stochastic) convergence is fastest
Additional power by stratification and/or importance sampling



Additional Bonus → Veto algorithm → direct simulation of
arbitrarily complicated reaction chains → next lecture

Recommended Reading

F. James

Monte Carlo Theory and Practice

Rept.Prog.Phys.43 (1980) p.1145

S. Weinzierl

Topical lectures given at the Research School Subatomic physics, Amsterdam, June 2000

Introduction to Monte Carlo Methods

e-Print: hep-ph/0006269

S. Teukolsky, B. Flannery, W. Press, T. Vetterling

Numerical Recipes (in FORTRAN, C, ...)

<http://www.nr.com/>