Introduction to Monte Carlo

Topics

- Alyssa: Search for SUSY
- Alice: Discovery of CP Violation
- George: Discovery of the Neutron
- Dasom: Neutrino Oscillations
- Kevin: W and Z
- Kelsey: Observation of top quark
- Kripa : history of the proton
- Bertie: Dark Matter
- Josh: Search for the Higgs at the LHC
- DJ: Discovery of the muon
- Joseph: Discovery of the Tau

Schedule : 10-15 min each

- February 23rd
  - Kripa
  - George
  - DJ
- February 28th
  - Alice
  - Joseph
  - Kevin
  - Bertie
- February 29th
  - Kelsey
  - Dasom
  - Josh
  - Alyssa

What to prepare

- Read paper and prepare a 10-15 presentation explaining
  - What is the underlying physics they accomplished?
  - What was the experimental setup ?
  - How did they analyze the data?
- Make short presentation on the order of ~7-10 slides explaining those 3 things above
- Come and ask me questions if you don’t understand something. As you will see some papers are incredibly brief !!

Last Two Weeks

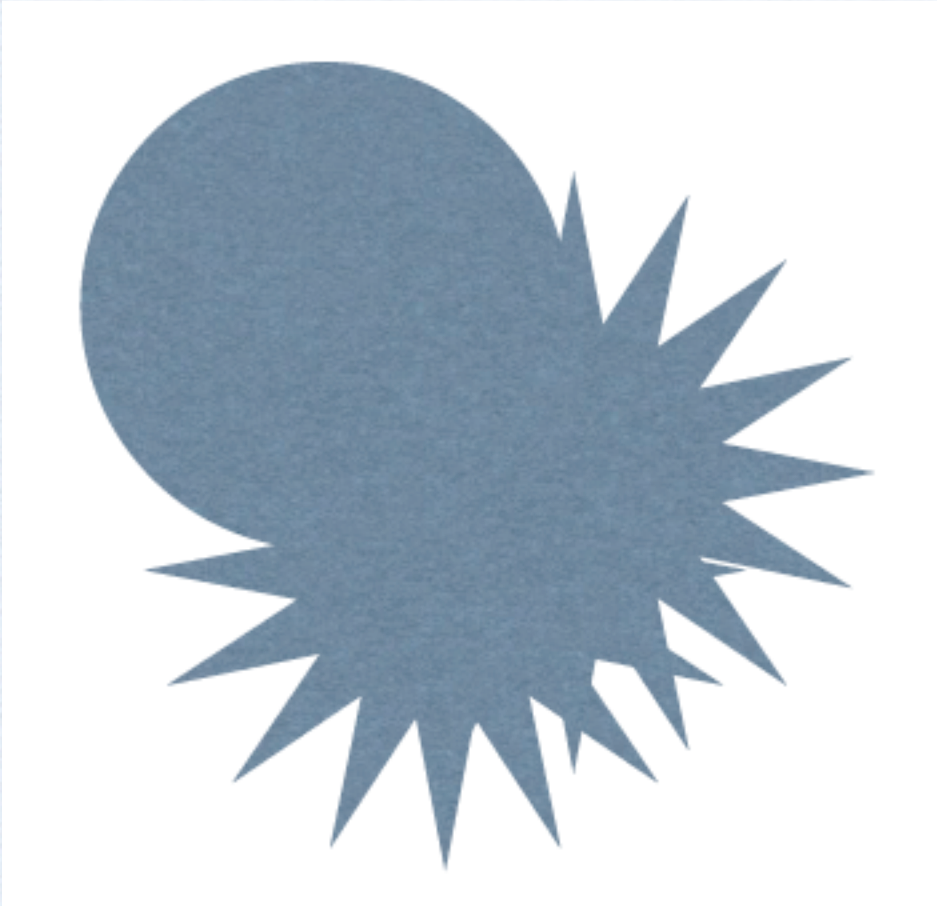
- According to the schedule I have - the last two weeks you have a conflict on Thursday
- If that is true I would propose to move that class to Wednesday

What is Monte Carlo

- A general computational method to approximate very complicated calculations which sometimes do not have analytic solutions
- Perform a large number of experiments using random number generators to see what happens

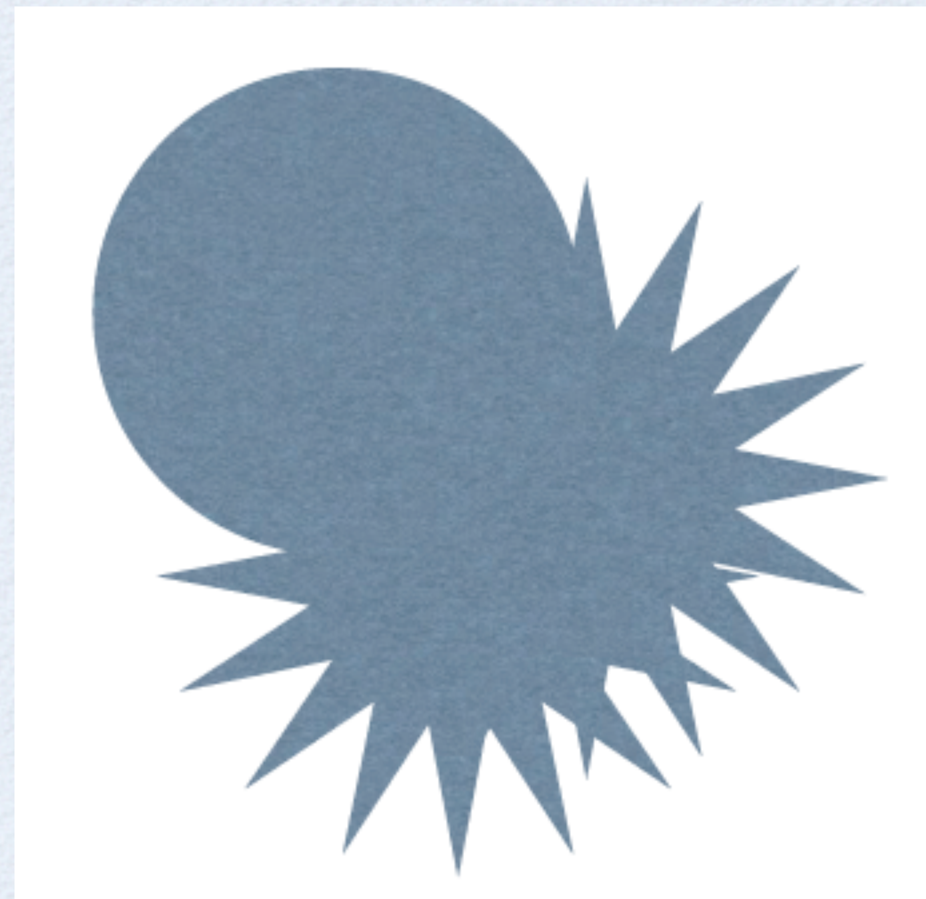
Simple Example

- Say we had some complicated shape that we would like to calculate the area of

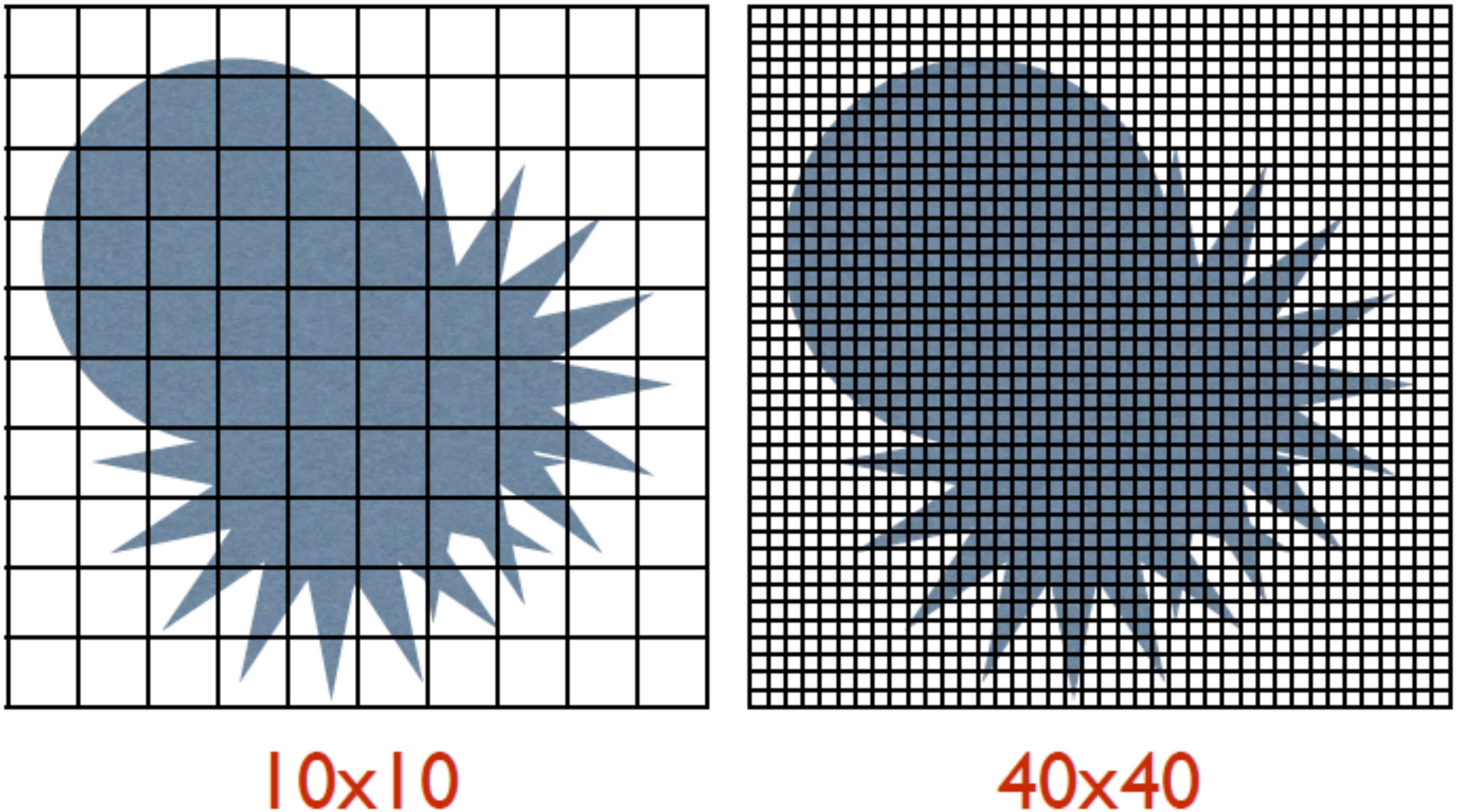


Any ideas

- Difficult to do analytically
- easier to do numerically



Basic Idea



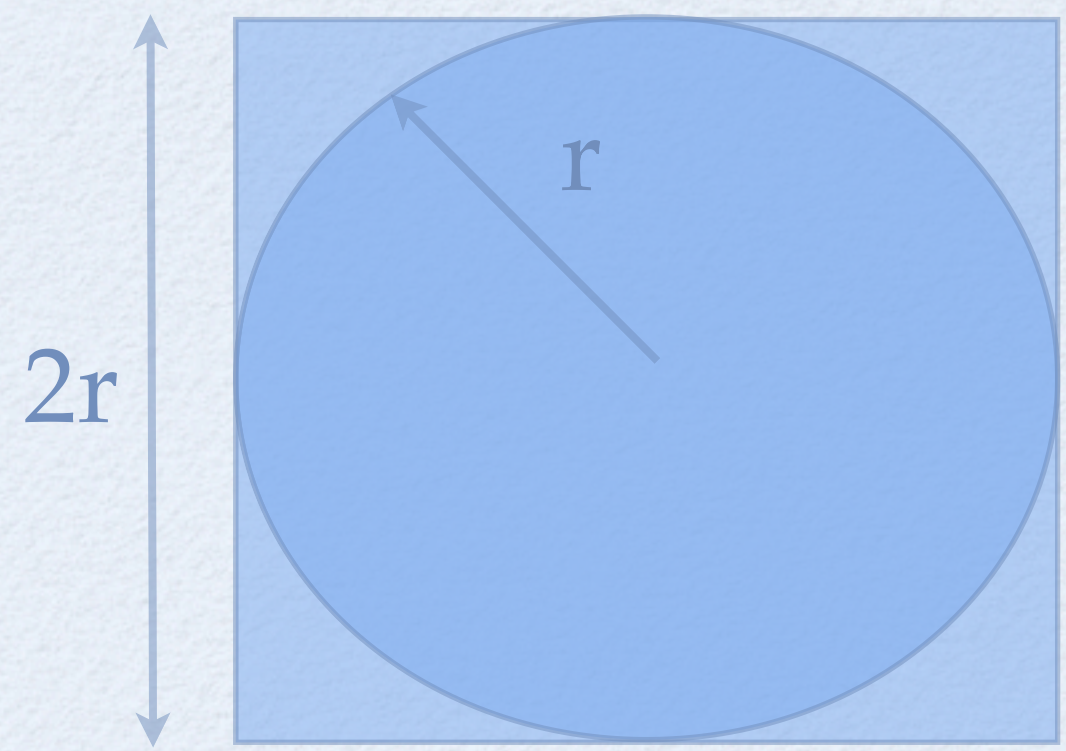
Area = (# Hits)/(# Total) x total area

Estimation of Pi

- How could we use this method to estimate pi

EXAMPLE

- Imagine a square of side of length 2r
- inside that square is a circle of radius of r



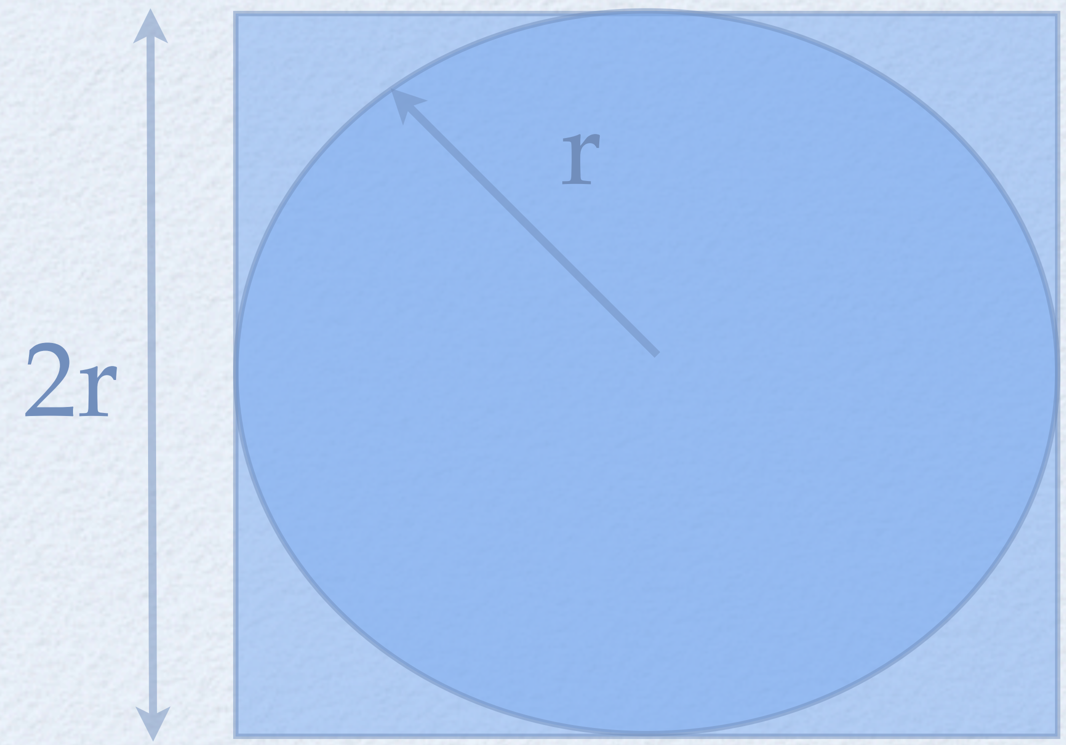
Example

The areas can be written

$$A_{square} = (2r) \times (2r) = 4r^2$$

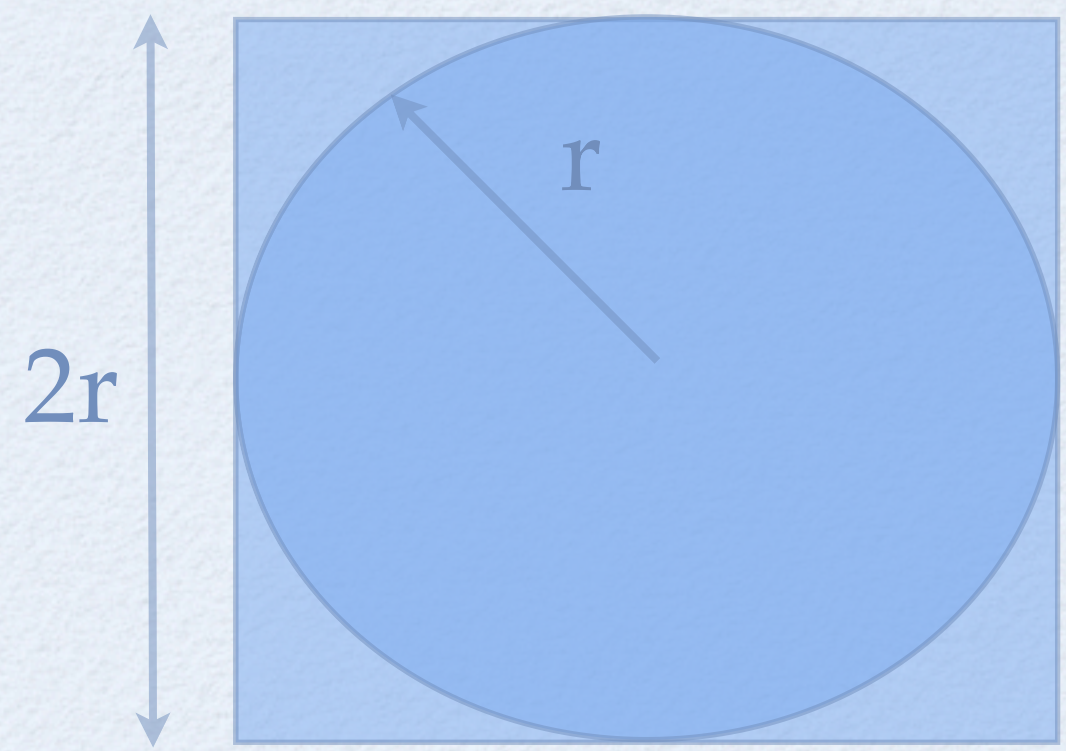
$$A_{circle} = \pi r^2$$

$$\frac{A_{circle}}{A_{square}} = \frac{\pi}{4}$$



Example

Now imagine we throw darts at a square wall with a circle drawn in like this



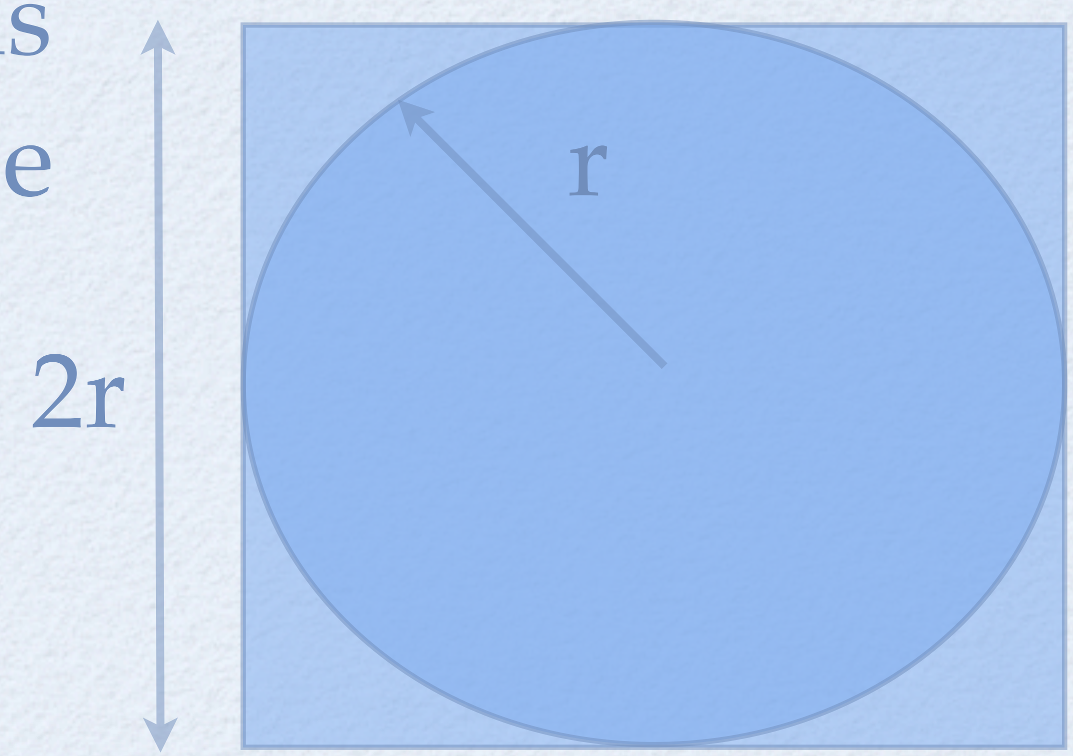
$$\frac{A_{circle}}{A_{square}} = \frac{\pi}{4}$$

EXAMPLE

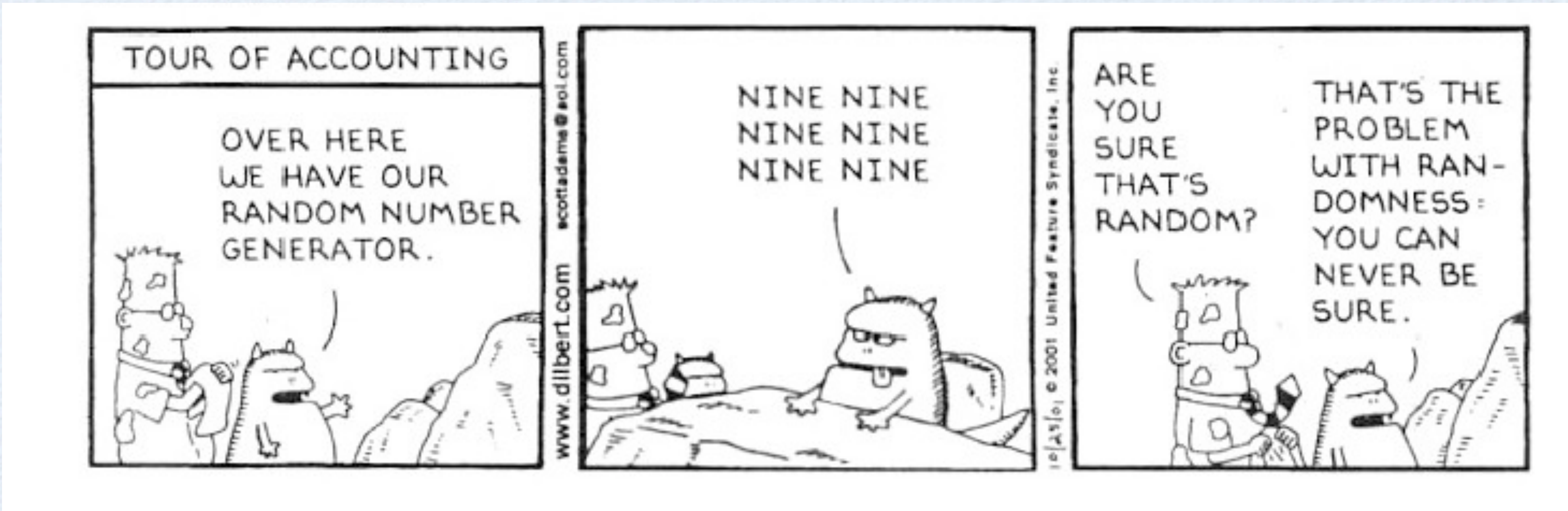
If thrown randomly it will land with equal probability anywhere. Hence, the ratio that is inside the circle to the ratio inside that inside the square is simply

$$\frac{\pi}{4}$$

$$\frac{A_{circle}}{A_{square}} = \frac{\pi}{4}$$



Aside : Random #s



- How do we generate them?

We would like

- In a uniform distribution of random numbers in [0,1] every number has the same chance of showing up

- Note that 0.000000001 is just as likely as 0.5

To obtain random numbers:

- Use some chaotic system like roulette, lotto, 6-49, ...
- Use a process, inherently random, like radioactive decay
- Tables of a few million truly random numbers exist .....

(.....until ~ 30 years ago.....)

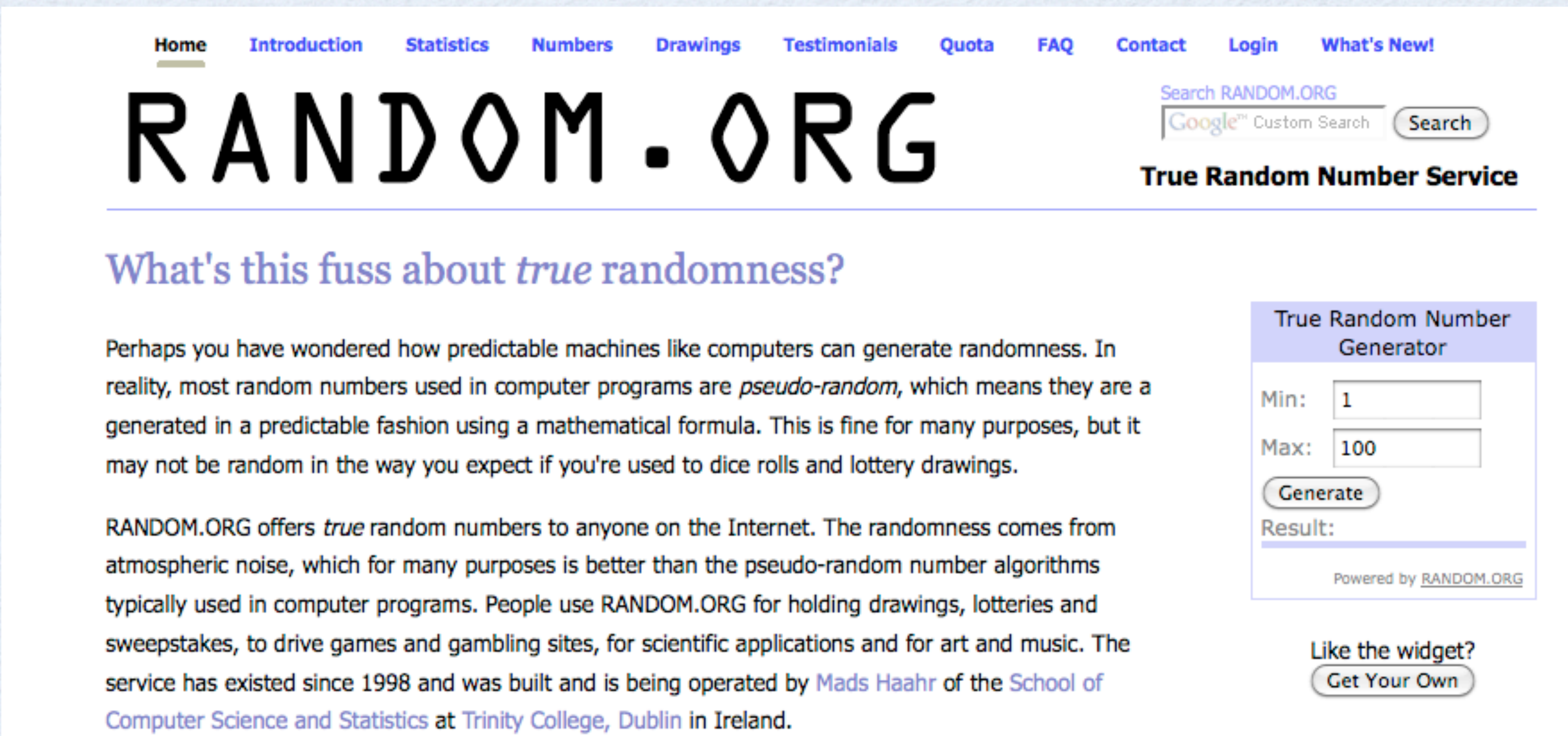
BUT not enough for most applications

→ .... we have true random number generators ...

TRUE RANDOM #S

- atmospheric noise, which is quite easy to pick up with a normal radio: used by RANDOM.ORG

- much more can be found on the web ....

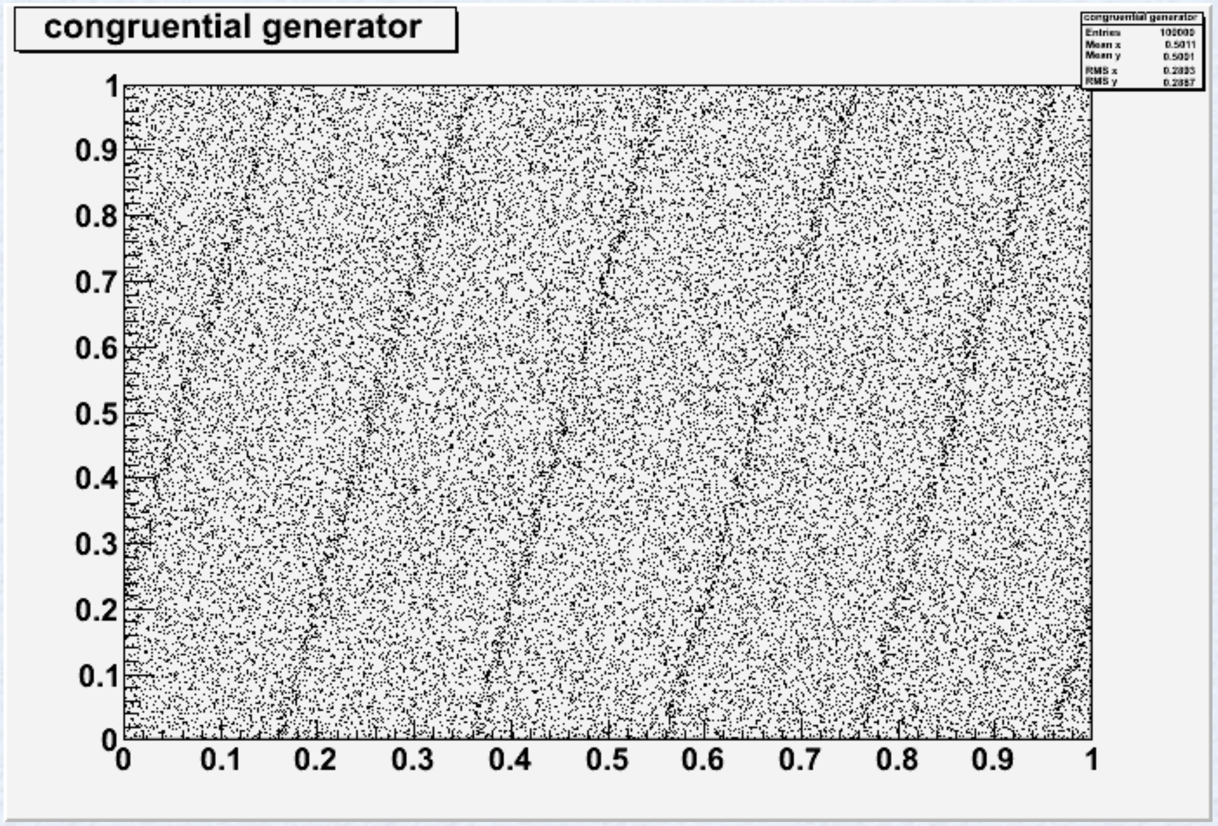


Psuedo Random #s

- Are a sequence of numbers generated by a computer algorithm, usually uniform in the range [0,1]
- more precisely: algo's generate integers between 0 and M, and then rn=In/M
- A very early example: Middles Square (John van Neumann, 1946) generate a sequence, start with a number of 10 digits, square it, then take the middle 10 digits from the answer, as the next number etc.:
- 57721566492 = 33317792380594909291
- Hmmmm, sequence is not random, since each number is determined from the previous, but it appears to be random
- this algorithm has problems ..... BUT a more complex algo does not necessarily lead to better random sequences .... Better use an algo that is well understood

Periodicity

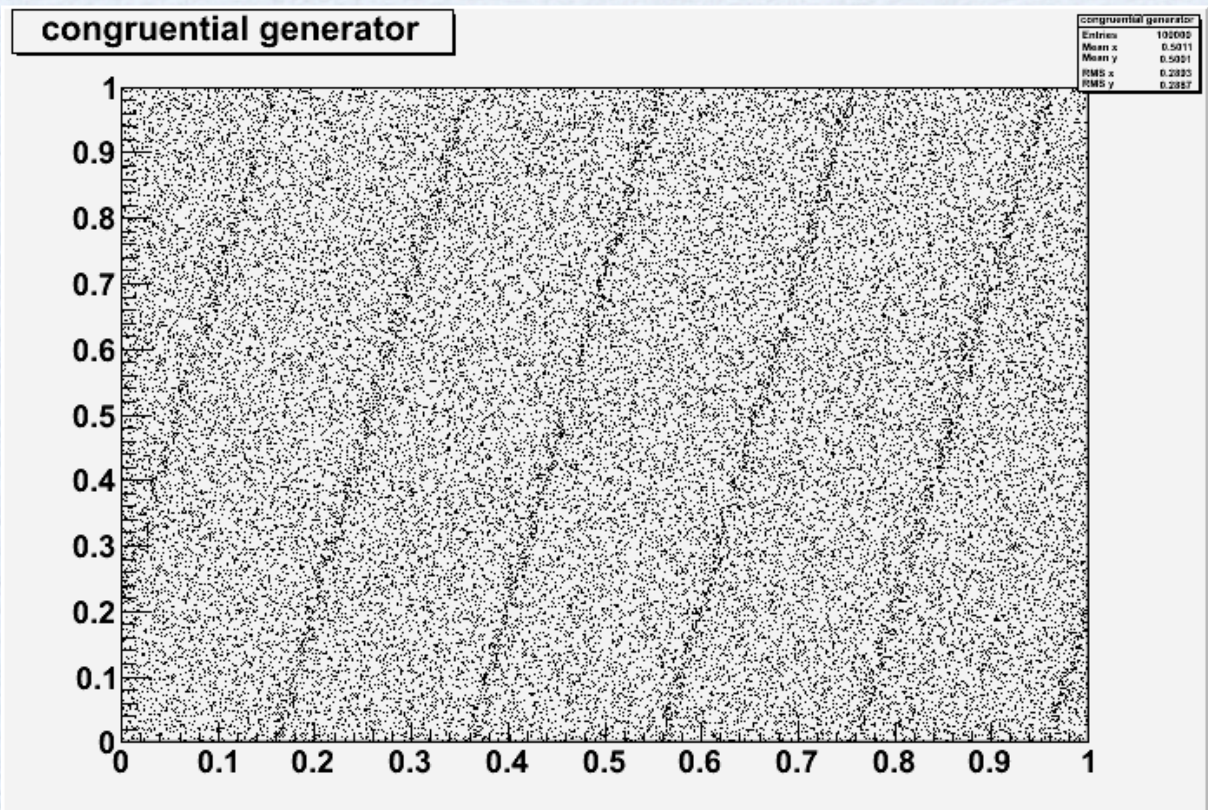
Simple Generator



Pretty good but you can see periodicity !

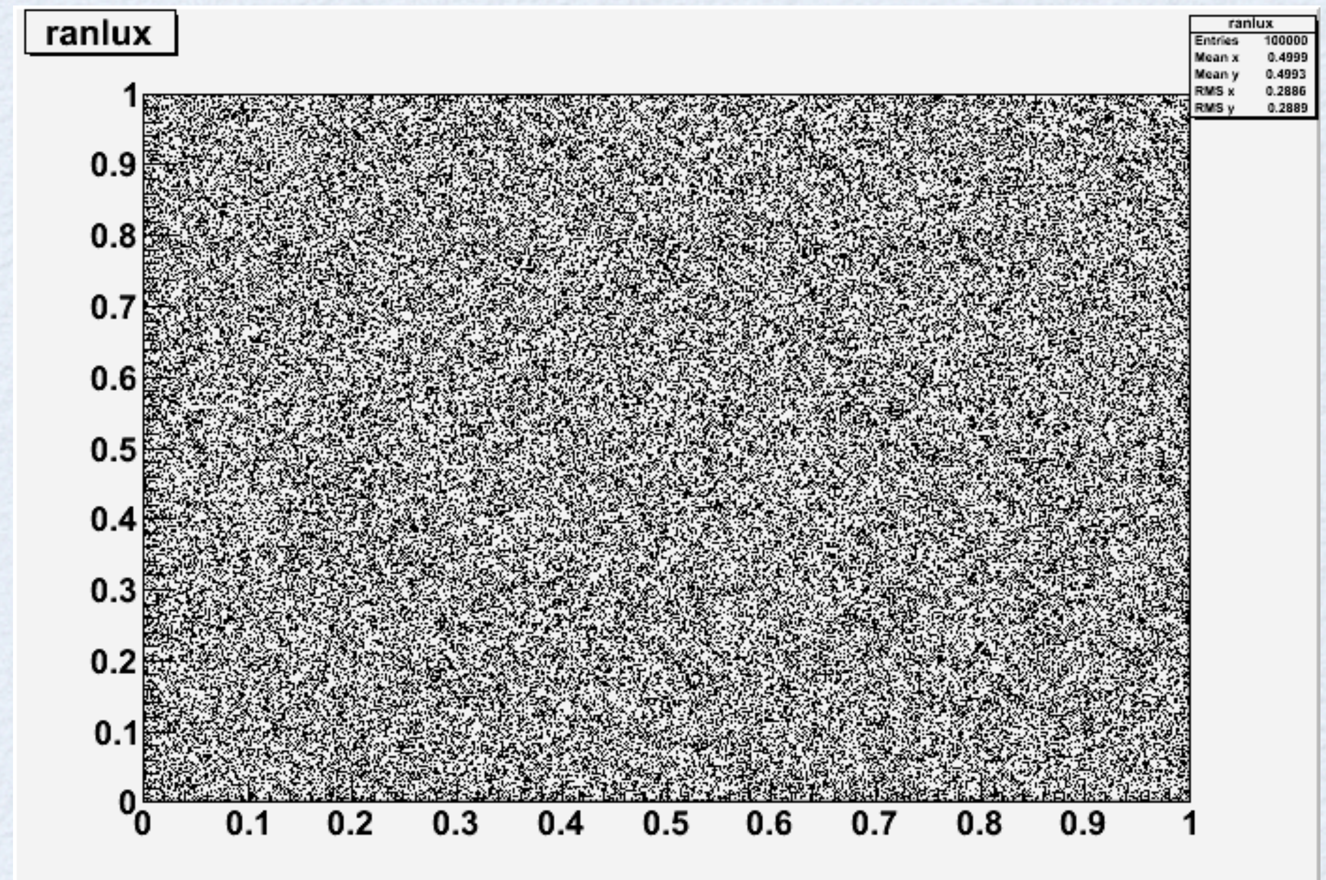
PERIODICITY

Simple Generator



Pretty good but you can see periodicity !

Sophisticated Generator: RANLUX



OK BUT WHY THIS AND NOT

Use in Science and Math

$$\int_{x_1}^{x_2} f(x)dx = (x_2 - x_1)\langle f(x)\rangle$$

$$\langle f(x)\rangle \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

- Integration:
    - We are trying to calculate the error under a complicated function that we don't have a simple analytic expression for
    - can use the 'throw dart method'
- Turns out it is a popular method because the uncertainty on the integral scales well with dimension of problem

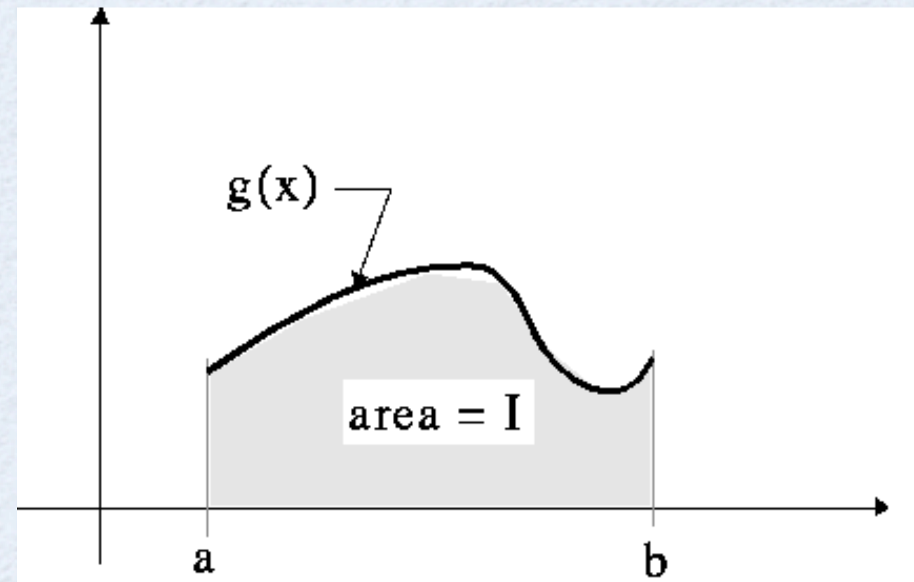


Figure 9. Monte Carlo Integration

Multi Dimensions

- Turns out that because the method uses random or psuedo-random numbers the uncertainty in the limit of a large number trials is independent of the dimensionality of the integral
- Not true of other algorithms
  - Trapazoid rule
  - Simposon's rule

- Monte Carlo error remains $\propto 1/\sqrt{N}$
- Trapezium rule $\propto 1/N^{2/d}$
- Simpson's rule $\propto 1/N^{4/d}$

In Particle Physics

- Phase space is 3 dimensions per particles
  - ~250 hadrons at an LHC event
- Accuracy remains the same for monte carlo method but is greatly reduced for other classic methods that work well in one or low dimensions
- Uncertainty estimate is much easier than in other methods

History and Name

- Method formally developed by John Neumann during the World War II, but already known before. It was used to study radiation shielding and distance that neutrons would likely travel through material.

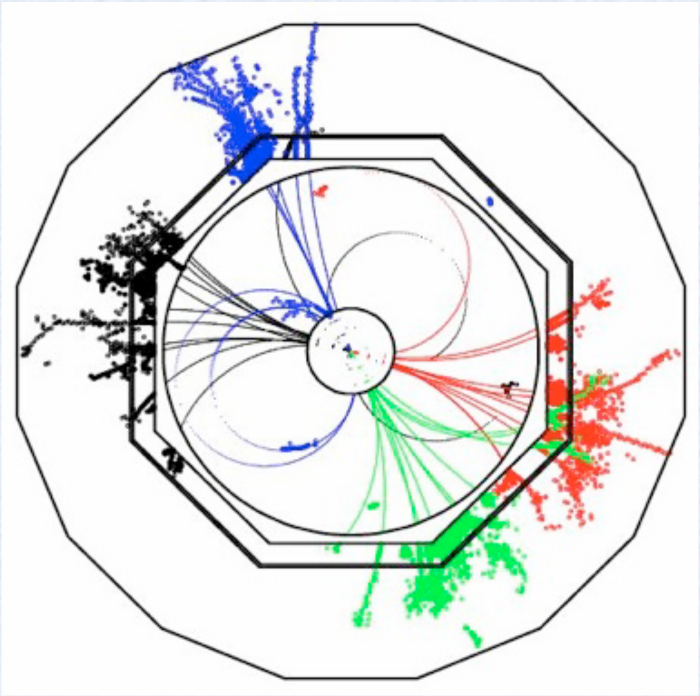
- Von Neumann chose the codename "Monte Carlo". The name is a reference to the Monte Carlo Casino in Monaco where Ulam's uncle would borrow money to gamble

Why Monte Carlo

- Monte Carlo assumes the system is described by probability density functions (PDF) which can be modeled. It does not need to write down and solve equation analytically / numerically. PDF comes from
  - Data driven
  - Theory driven
  - Data + Theory fitting

Where do we use it

- Detector design and optimization
    - Complicate and huge detector
    - Very expensive
- Simulation of particle interactions with detector's material
- Physics analysis
    - New predicted physics
    - Event selection
    - Background estimation
    - Efficiencies of detector/algorithm/...



Use in HEP

Choose model, constraints, parameters, decay chain of interest

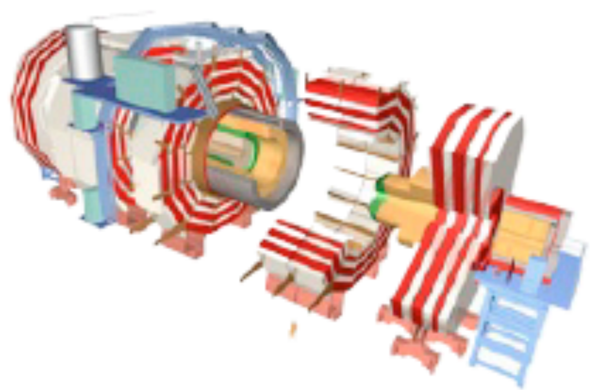
Proposed Theory

Kinematics, information from a known (detectable) particles

Generator

Detector Simulation
- Hardware
- Software

Simulation Digitization Triggering



Experiment Triggering

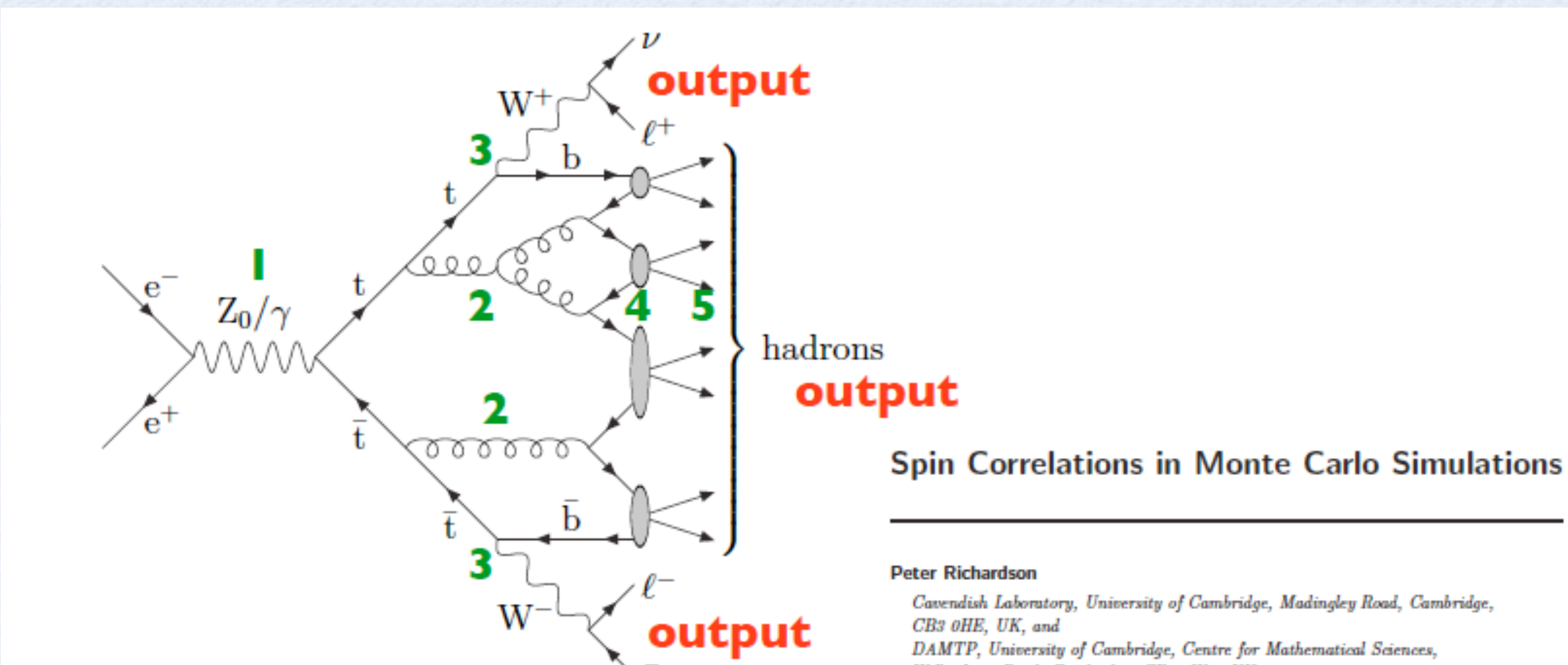
Offline software
- Event selection

Reconstruction Analysis

Results
Improvement

Event Generator

- Say we would like to know what a top event at the international linear collider
- We would like to have an idea of what this looks like before we build the detector!



First Step

- We generate events according to the quantum mechanical probability for it to occur
- Up until now we have considered Monte Carlo as a numeric trick for integration
- If the function to be integrated is a probability density function we can turn it into an "event generator"

$$\sigma = \int_0^1 \frac{d\sigma}{dx}\, dx$$

What is an event?

- A real event is a particular interaction - say a proton proton collision at the LHC, a neutrino interaction , etc
- We are trying to simulate that with a computer program and come up with a list of 4-vectors for events whose properties are drawn from the same probability distribution function as that in nature

Monte Carlo Events

$$\sigma = \int_0^1 \frac{d\sigma}{dx}\, dx$$

- Say σ is cross-section for a certain process
- d σ/dx is called the differential cross-section and is related to the probability of an event to have property x
- Weighted Events: Can generate events with weights given by dσ/dx
- Unweighted Events: No weights but only keep a fraction of the events and decide to keep them with probability :

$$(d\sigma/dx)/(d\sigma/dx)_{\mathrm{max}}$$

WHAT YOU GET

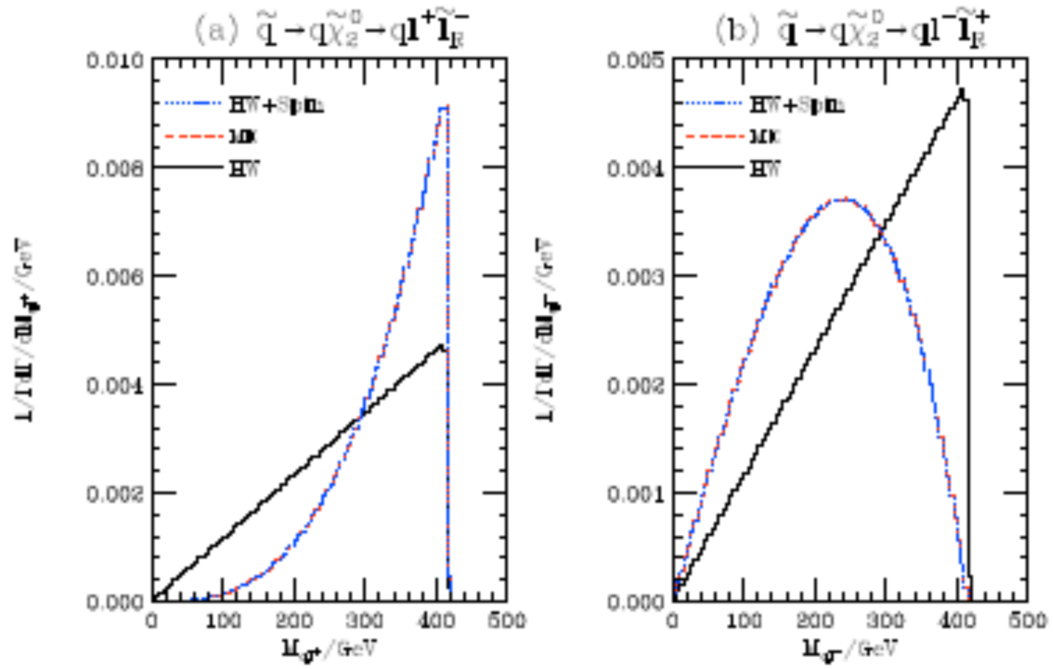


Figure 3: Distribution of the mass of quark and lepton produced in the decay $\tilde{q}_L \to q\tilde{\chi}^0_2 \to q\ell^{\pm}\tilde{\ell}^{\mp}_R$. The solid line gives the result of phase space, the dashed line gives the full result and the dotted line the result of the spin correlation algorithm.

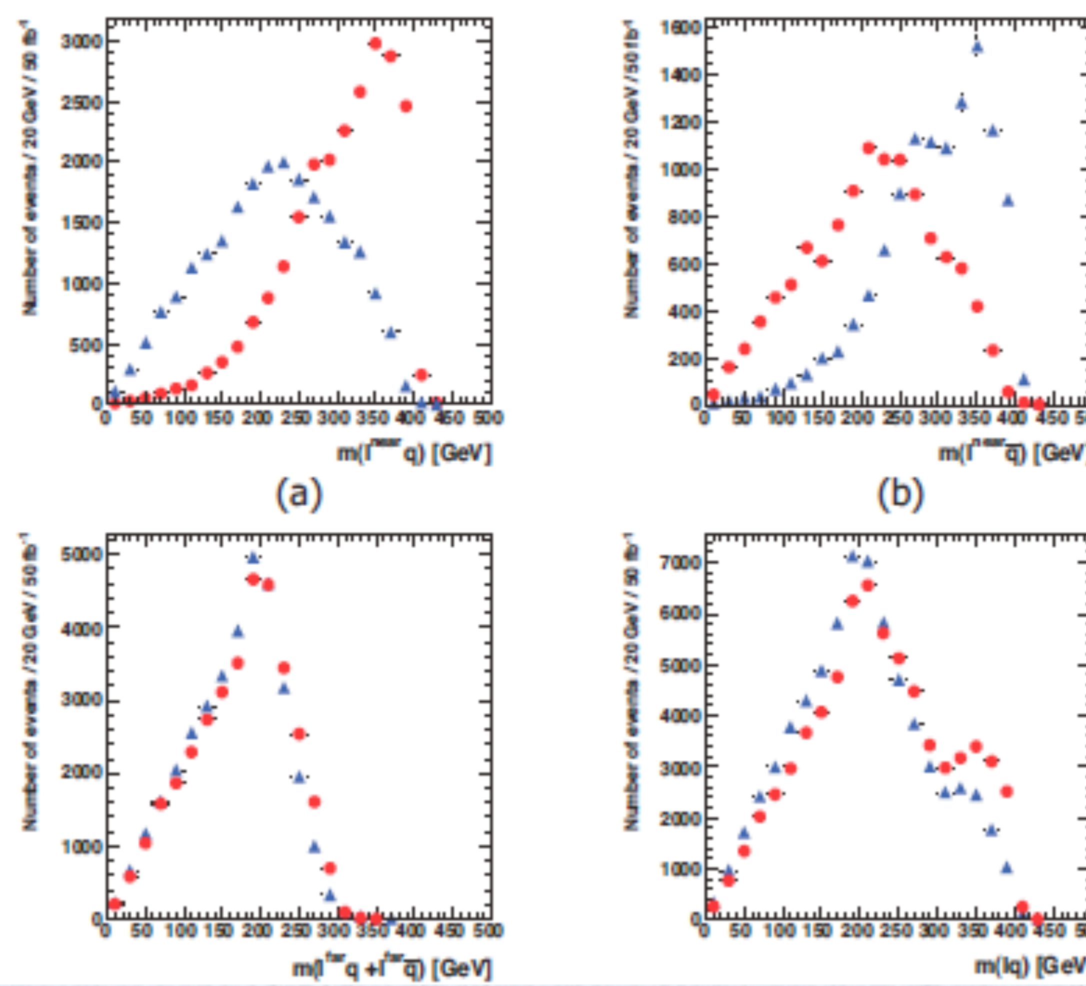
Proposed by generator

Spin Correlations in Monte Carlo Simulations

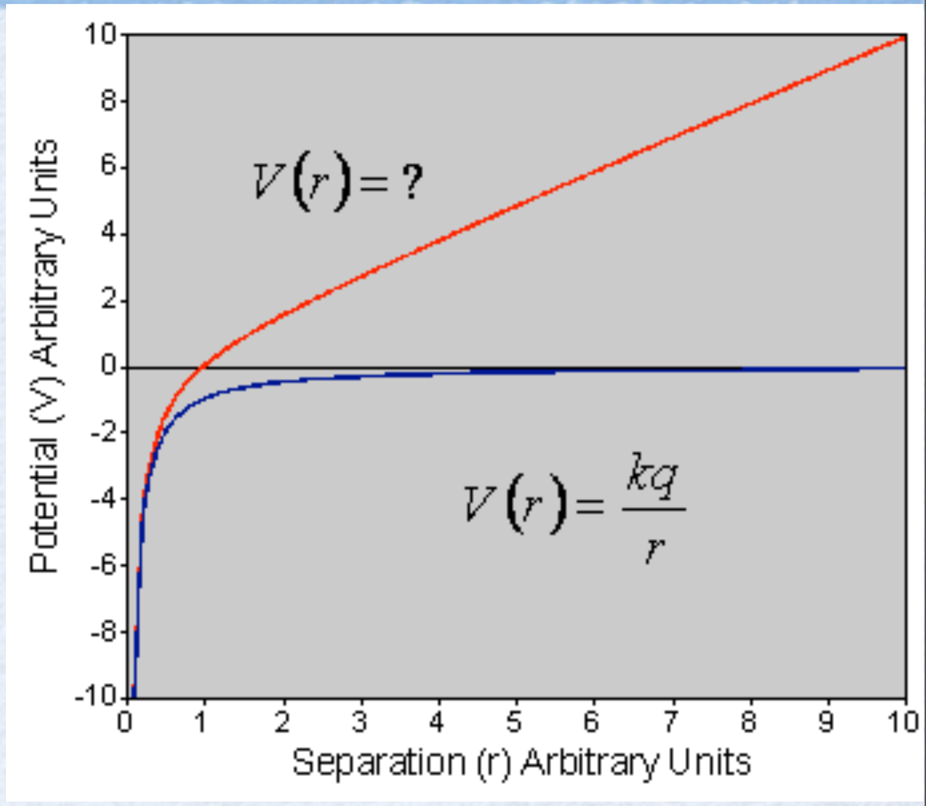
Peter Richardson

Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, CB3 0HE, UK, and
DAMTP, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, UK.

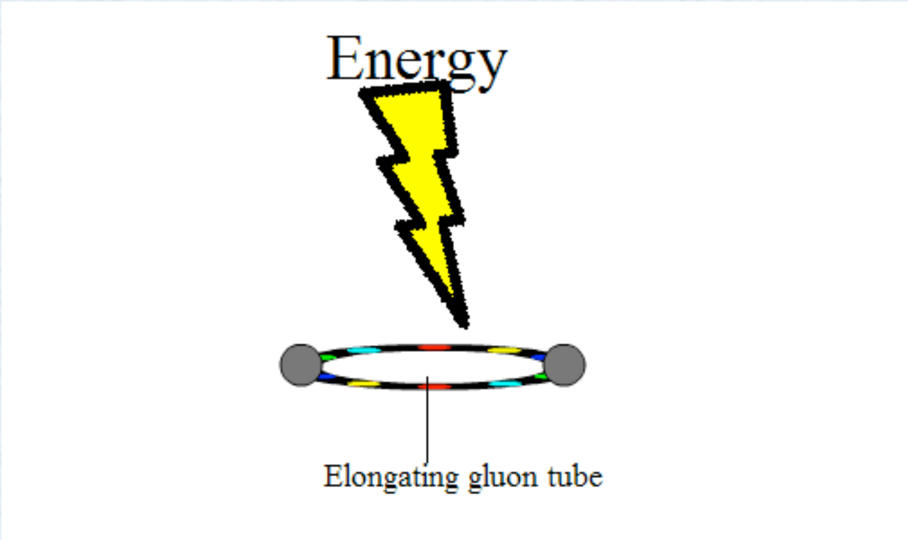
What you got from generator



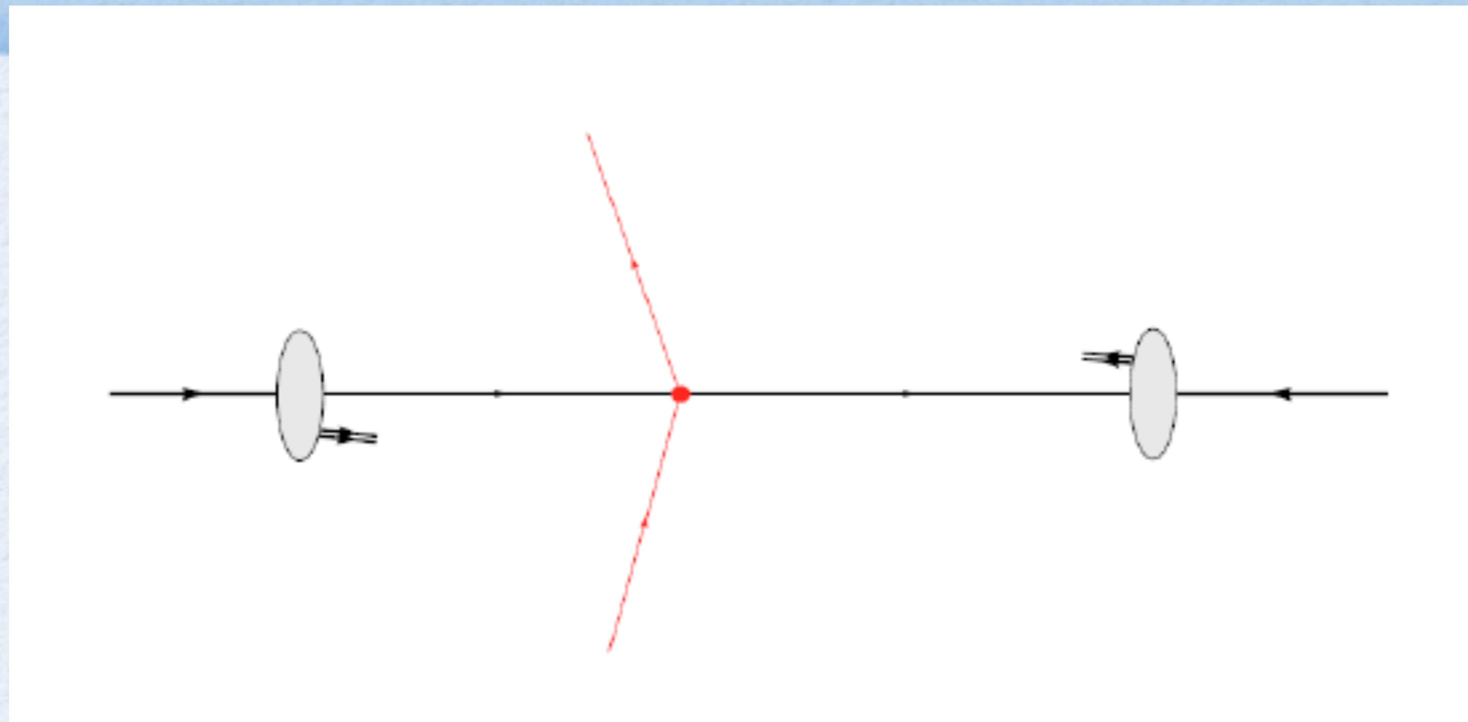
Aside: Strong Interactions



- Turns out that particles that feel the strong force have a quantum charge which we call color
- Unlike electrical or gravitational interactions the potential energy gets larger with distance !
- Enough energy to produce a pair of virtual particles which then form stable bound states

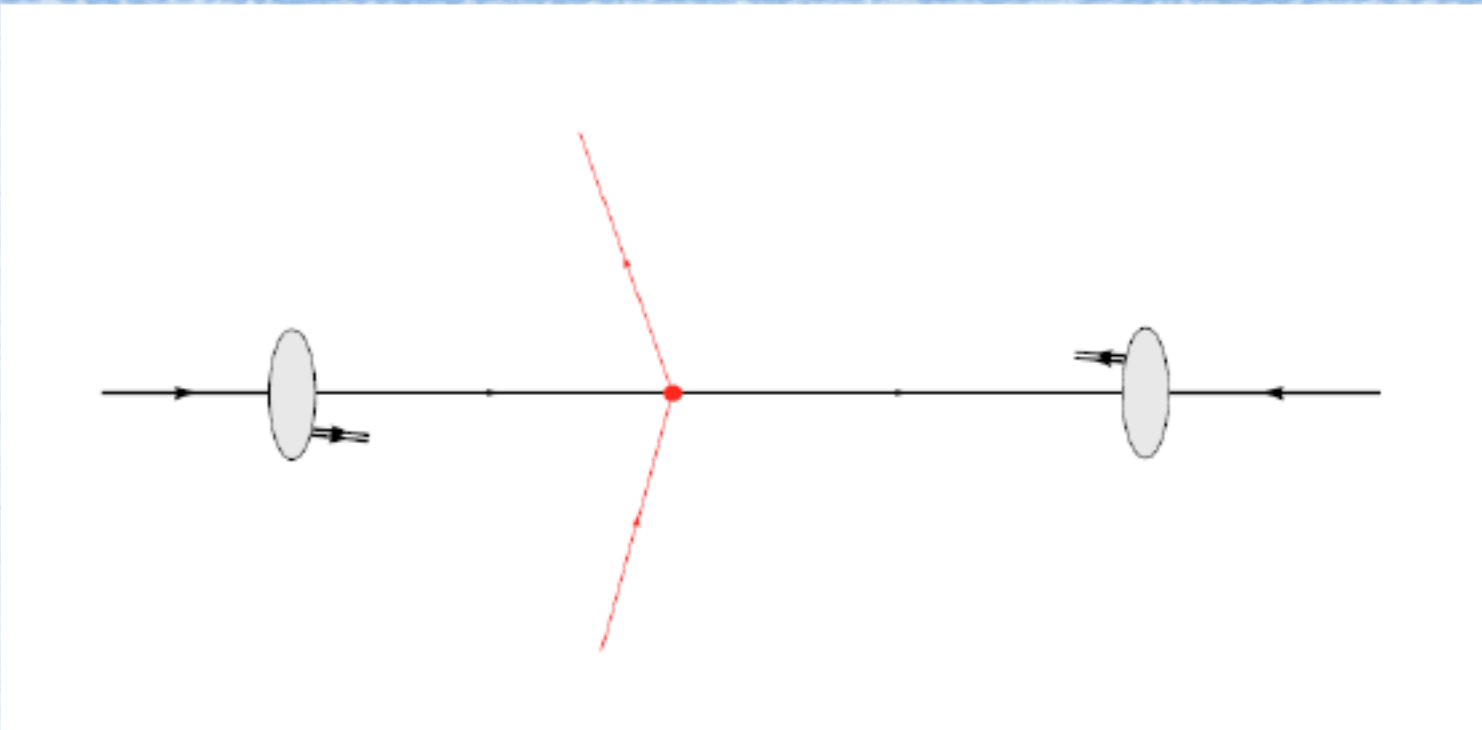


COMPLICATION



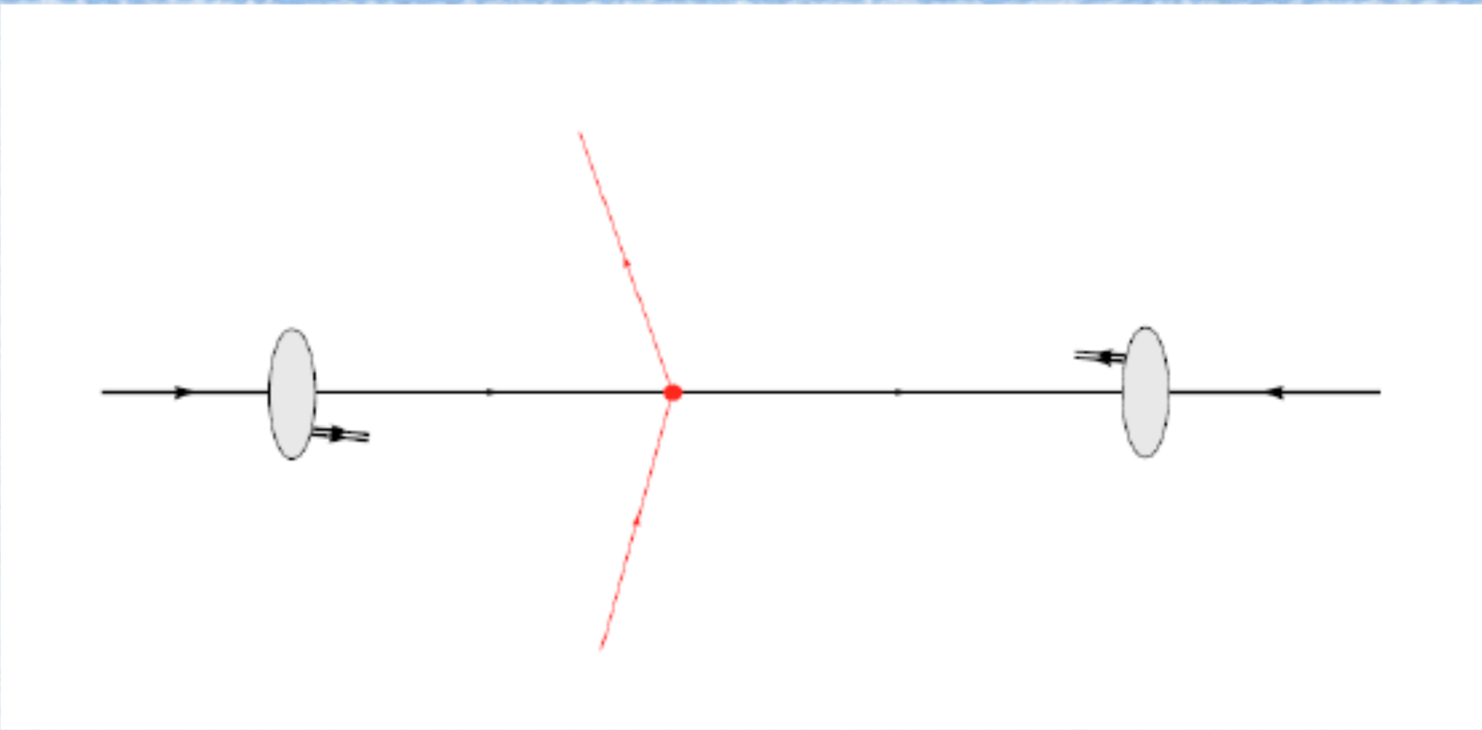
- Proton collisions are in detail fantastically complicated

What is a proton

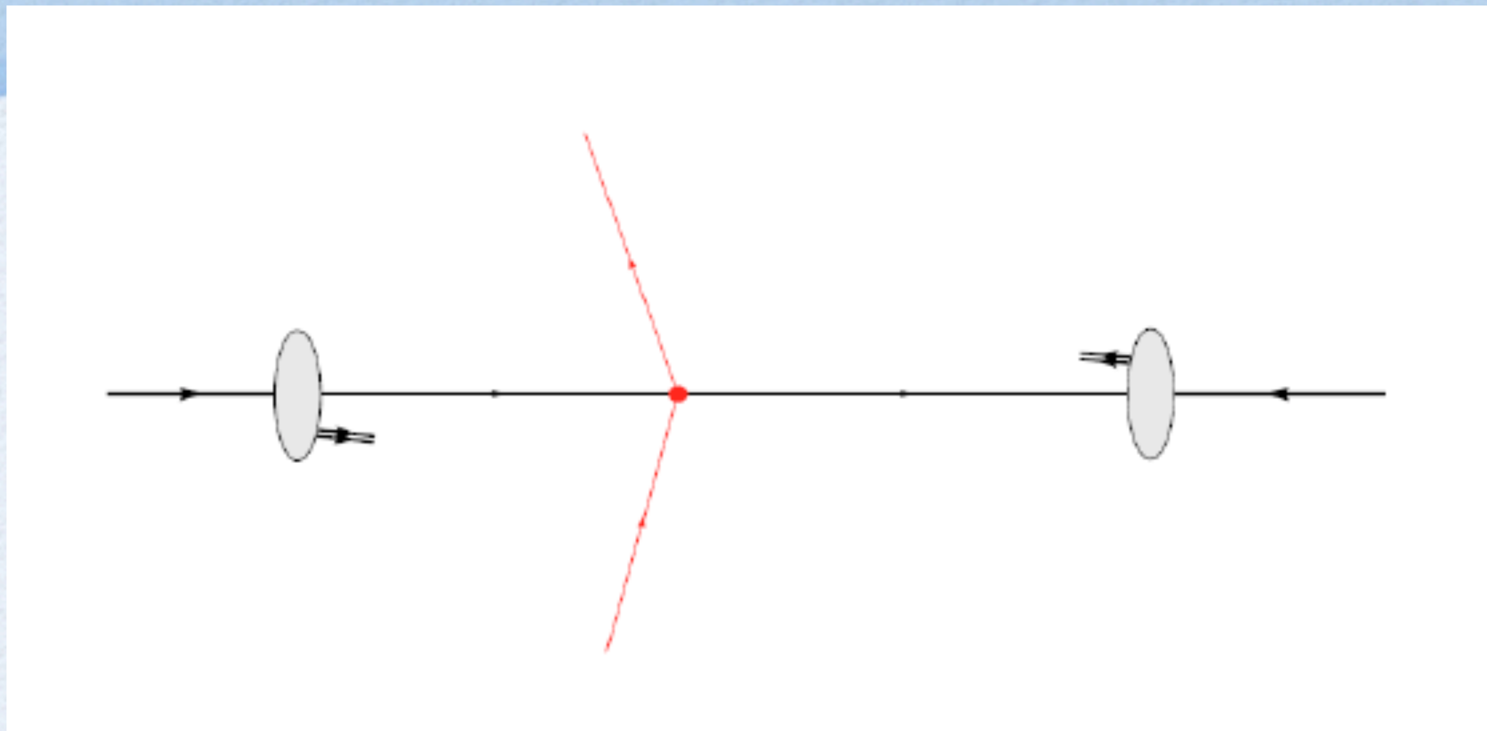


- Each proton contains 3 valence quarks (uud)

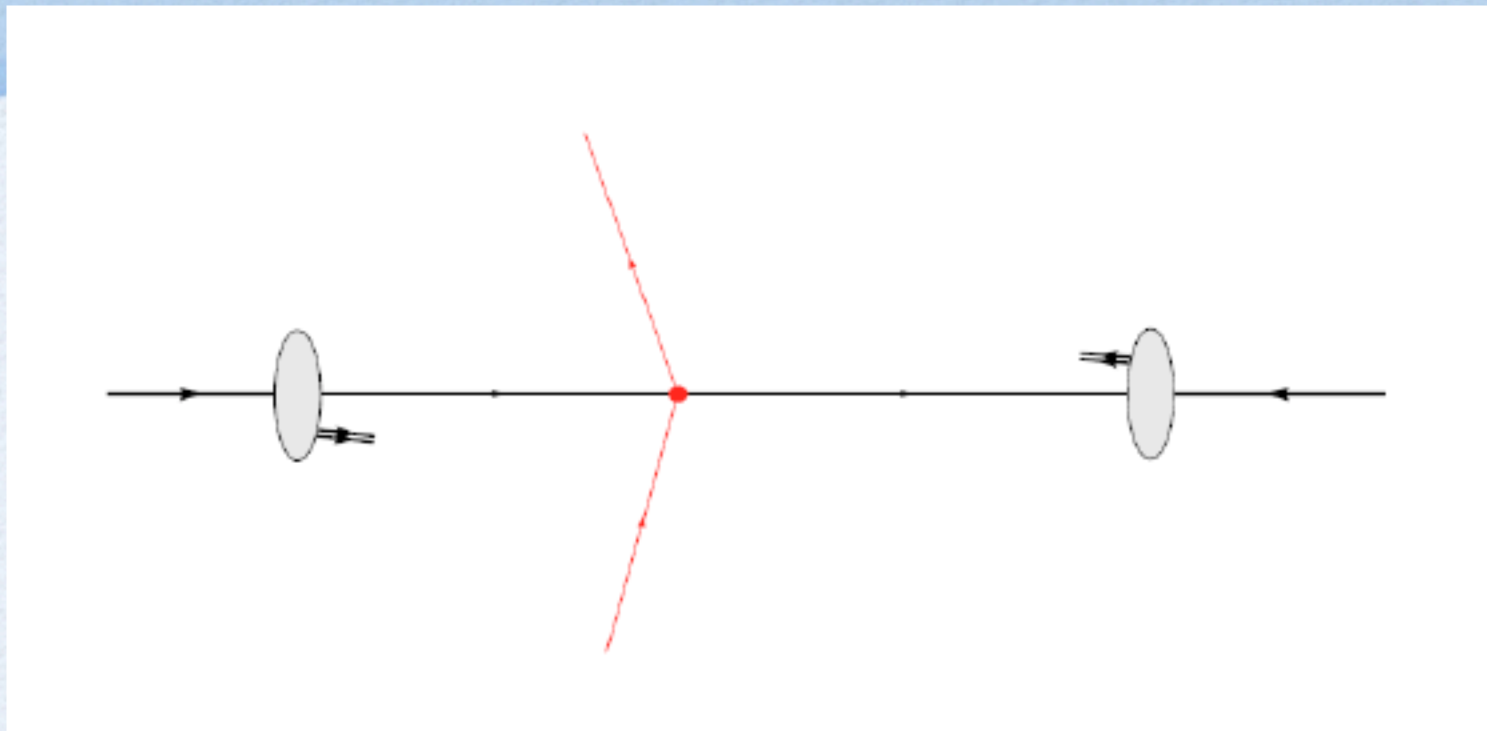
What is a proton



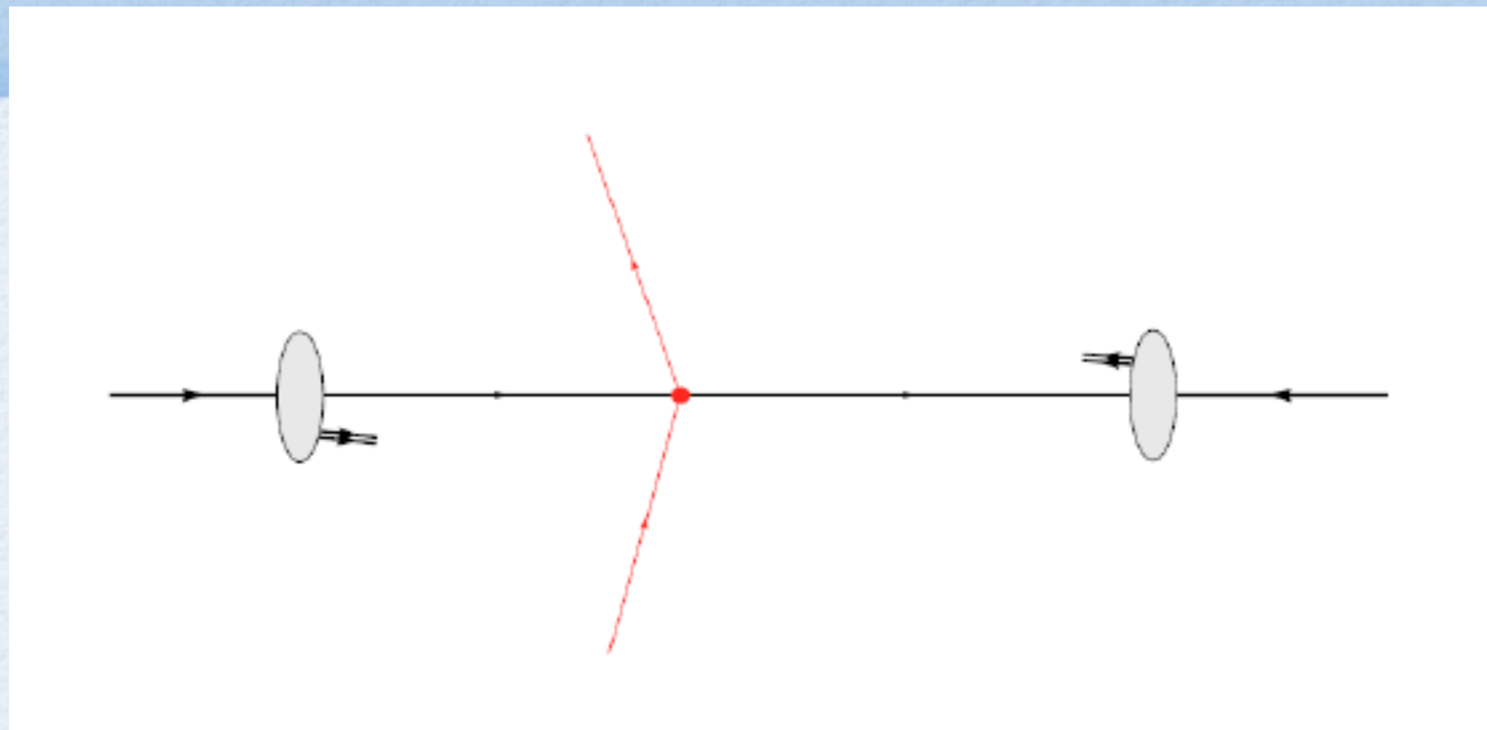
- Each proton contains 3 valence quarks (uud)
- and gluons



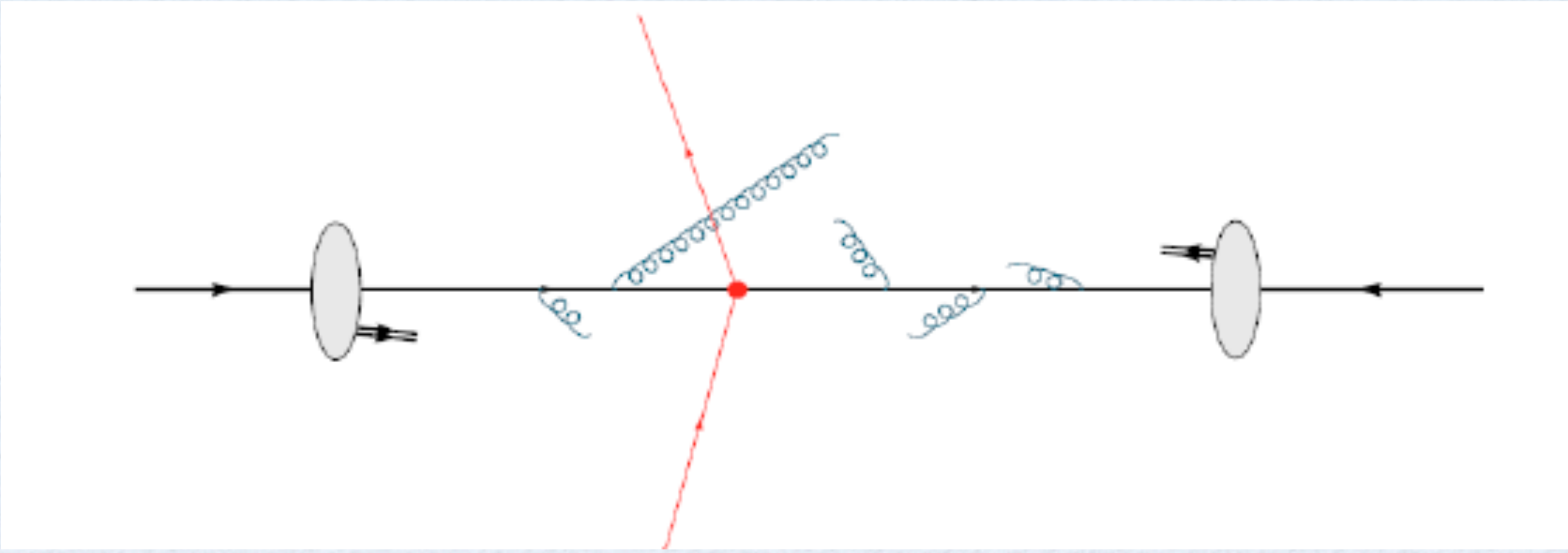
- Each proton contains 3 valence quarks (uud)
- and gluons
- and - quantum mechanical fluctuations of virtual particles known collectively as "sea quarks"



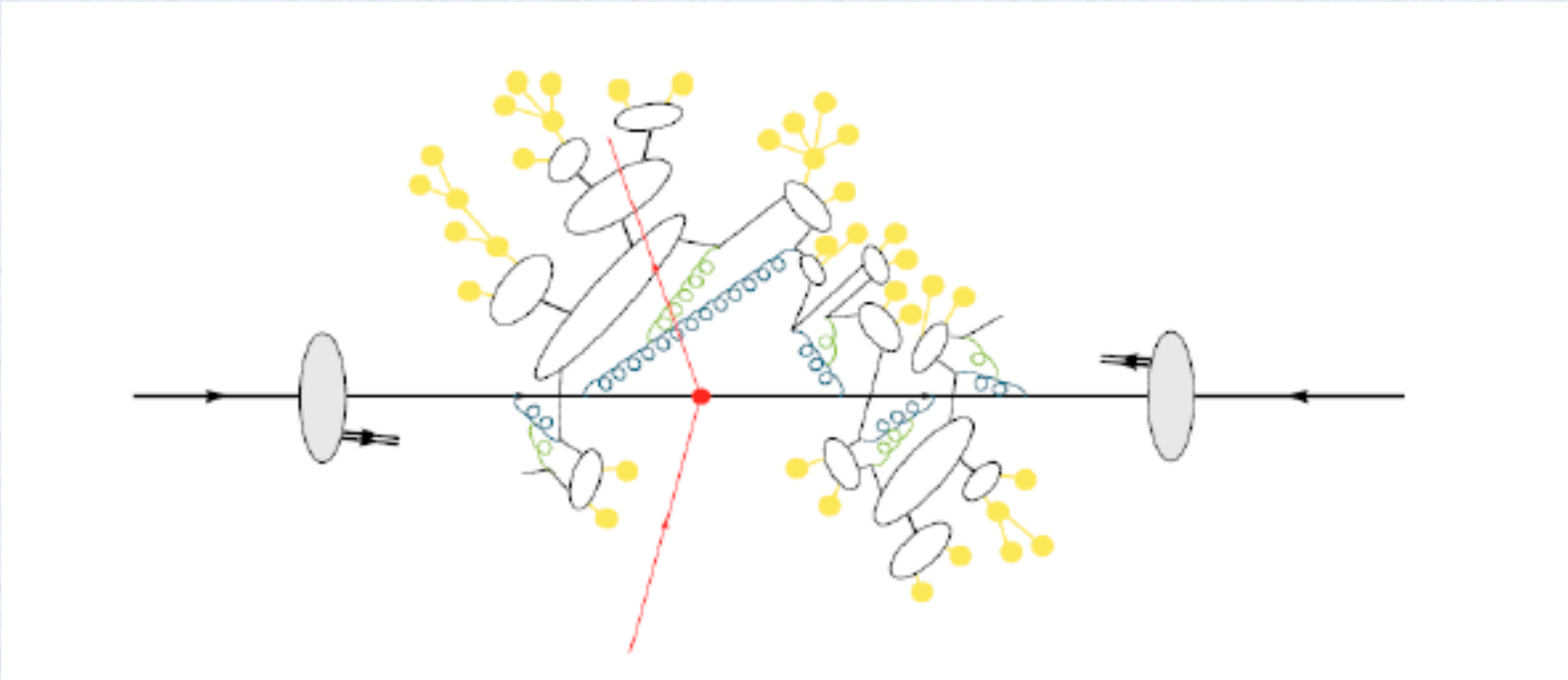
- Each ‘hard’ collision is between one of the constituent particles
- The red line is the outgoing hard particles in a “2 to 2” process
  - two particles come in - two particles leave



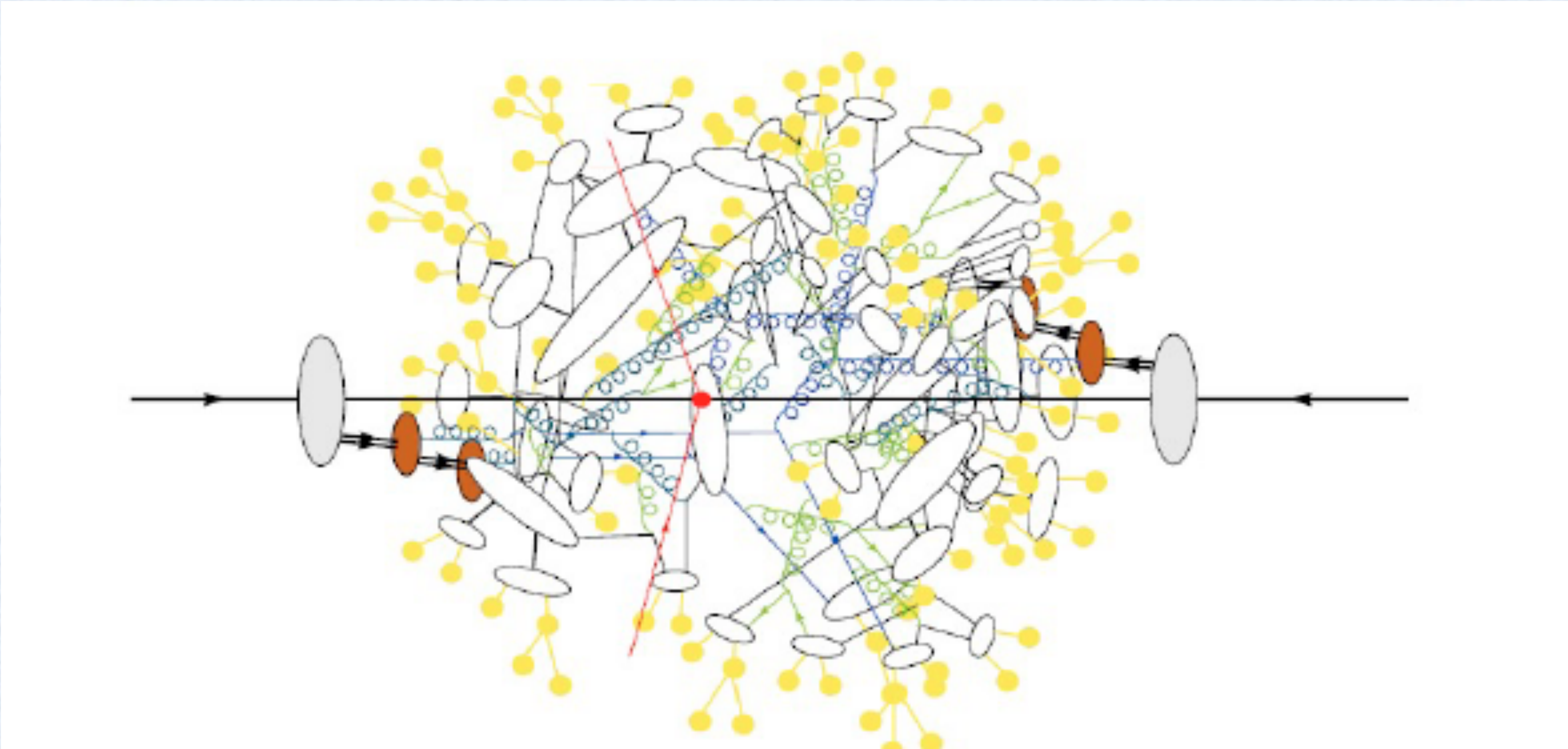
- On top of this the initial and final (if also strongly interacting) produce more strong particles



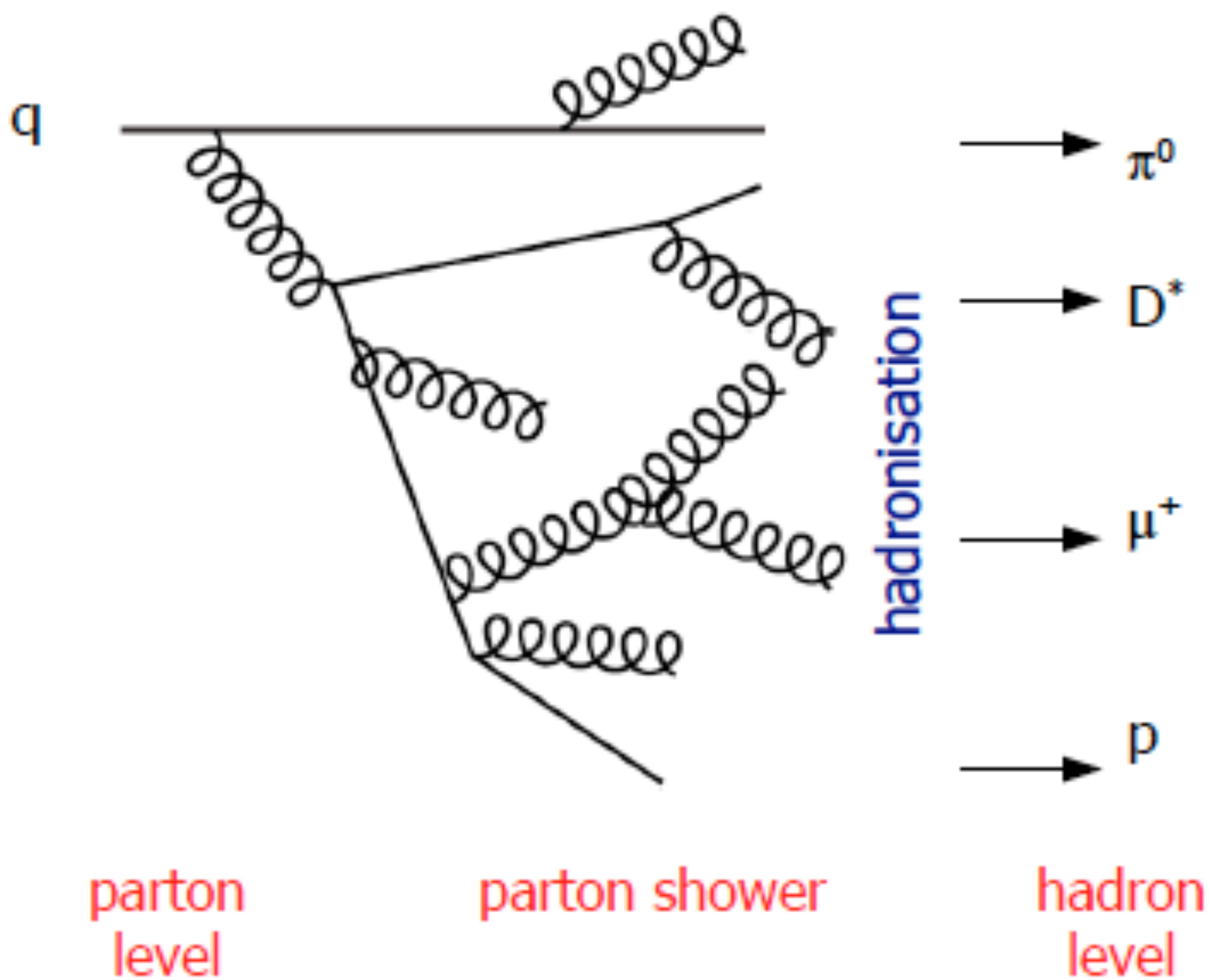
- but then those gluons 'hadronize' or turn into quasi-stable particles (pions, protons, etc)



- But then those hadrons decay..



- and then you have to pay attention to what happens to the rest of the protons...



- Generator stops with set of "stable" final state particles
- Complete 4-vector info is known about every particle
- All parent/daughter relations are kept track off
- High energy parton state known as parton level
- Stable particle state known as hadron level

IN DETAIL

```
                 EVENT       1: 7000.00 GEV/C P      ON  7000.00 GEV/C P       PROCESS:  3000
                 SEEDS:      41410 &       50928  STATUS:  100 ERROR:   0  WEIGHT:  3.2279E-02


                                                        ---INITIAL STATE---

 IHEP   ID       IDPDG IST MO1 MO2 DA1 DA2   P-X     P-Y     P-Z  ENERGY    MASS     V-X       V-Y       V-Z       V-C*T
    1 P           2212 101   0   0   0   0   0.00    0.00 7000.0 7000.0    0.94 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    2 P           2212 102   0   0   0   0   0.00    0.00-7000.0 7000.0    0.94 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    3 CMF            0 103   1   2   0   0   0.00    0.00    0.014000.014000.00 0.000E+00 0.000E+00 0.000E+00 0.000E+00

                                                        ---HARD SUBPROCESS---

 IHEP   ID       IDPDG IST MO1 MO2 DA1 DA2   P-X     P-Y     P-Z  ENERGY    MASS     V-X       V-Y       V-Z       V-C*T
    4 GLUON         21 121   6   7   9   8   0.00    0.00  370.3  370.3    0.75 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    5 UQRK           2 122   6   8  23   7   0.00    0.00-1825.6 1825.6    0.32 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    6 HARD           0 120   4   5   7   8  -0.93  -37.21-1455.3 2196.1 1644.30 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    7 GLUINO   1000021 123   6   5  33   4 109.61   -6.30-1522.1 1641.6  604.81 0.000E+00 0.000E+00 0.000E+00 0.000E+00
    8 SSUR     2000002 124   6   4  35   5 -109.61   6.30   66.8  554.3  539.15 0.000E+00 0.000E+00 0.000E+00 0.000E+00

                                                        ---PARTON SHOWERS---

 IHEP   ID       IDPDG IST MO1 MO2 DA1 DA2   P-X     P-Y     P-Z  ENERGY    MASS     V-X       V-Y       V-Z       V-C*T
    9 GLUON         94 141   4   6  11  22  -4.81  -25.73  387.2  354.0 -158.95 0.000E+00 0.000E+00 0.000E+00 0.000E+00
   10 CONE           0 100   4   7   0   0   1.00   -0.06   -1.3    1.6    0.00 0.000E+00 0.000E+00 0.000E+00 0.000E+00
   11 GLUON         21   2   9  12 348 349  -1.71   -0.67   -1.5    2.5    0.75-1.884E-13 1.924E-14-3.534E-13 1.456E-13
   12 GLUON         21   2   9  13 350 351  -0.88    0.78   -0.9    1.7    0.75-1.884E-13 1.924E-14-3.534E-13 1.456E-13
   13 GLUON         21   2   9  14 352 353  -2.10    1.29   -0.6    2.6    0.75 7.628E-16 1.278E-14-1.909E-13-1.791E-13
   14 GLUON         21   2   9  15 354 355  -1.94    1.53    3.4    4.3    0.75 4.095E-15 2.568E-14-7.810E-13-7.659E-13
   15 GLUON         21   2   9  47 356 357  -2.13   -0.76   95.8   95.8    0.75-5.939E-15 3.248E-14-4.639E-12-4.624E-12
   16 UD          2101   3   9   0  46   0   0.00    0.00 4705.8 4705.8    0.64 3.135E-13 9.477E-13-5.015E-10-5.014E-10
   17 UQRK           2   2   9  18 358 301   0.37    1.05 1723.7 1723.7    0.32 3.135E-13 9.477E-13-5.015E-10-5.014E-10
   18 GLUON         21   2   9  19 359 360   2.18    2.40   27.5   27.7    0.75 2.027E-13 2.212E-13 1.855E-12 1.884E-12
   19 GLUON         21   2   9  20 361 362   3.43    3.08   57.2   57.4    0.75 2.027E-13 2.212E-13 1.855E-12 1.884E-12
```

MC

- Remember: You need PDF to build MC program
  - Learn to describe particle production/decay by matrix element (amplitude) of hat process (Explain in QFT)
  - Learn QCD (to deal with jet fragmentation, parton)
  - Computational + Mathematical skills are needed
- Real life
- There are few theoretical groups who provide us the MC event generator:
  - Lund University, Sweden (PYTHIA)
  - INP, Krakow, Poland (TAUOLA, PHOTOS)
  - -Other groups
- - HERWIG
- - ISAJET

Experimentalists

- HEP experimentalists do not write MC generator ourselves (^_^)!!
- We use/modify MC generators proposed by theoretical groups.
- Your best MC generator is a generator which serve you results that you believe/study.
- Read manual (VERY IMPORTANT)

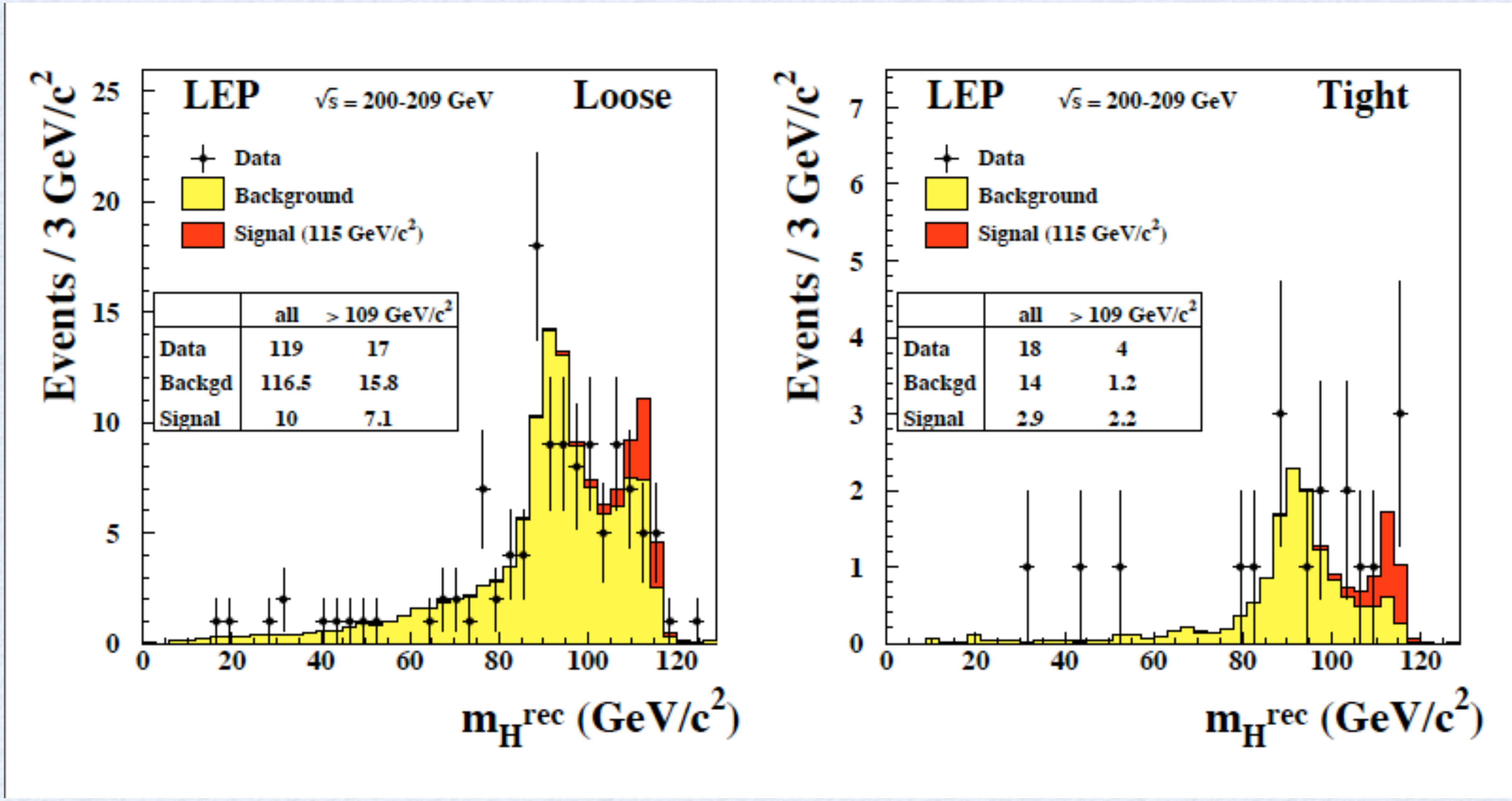
But...that is not enough

- Why?

Detector Simulation

- We need to know how our detector will see the productions from collisions
- Detector simulation tracks the particles through detector material (simulating their interactions with material)
  - Input: group of (quasi)-stable of particles from MC generator group of particles from particle-gun
  - Output: Depend on your analysis
- Popular detector simulation
  - GEANT4 (http://geant4.web.cern.ch)
  - FLUKA (http://www.fluka.org/fluka.php)

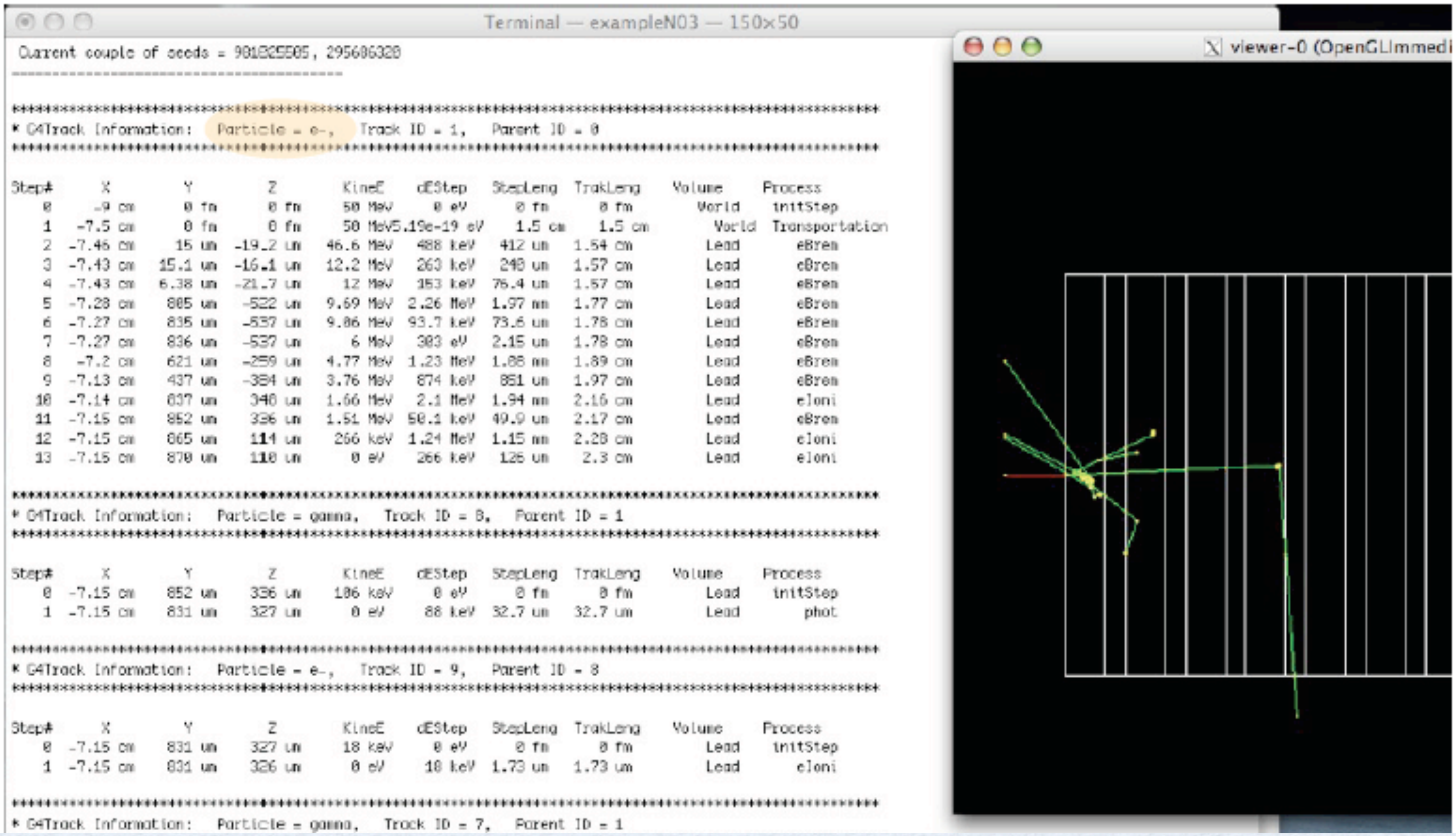
EXAMPLE



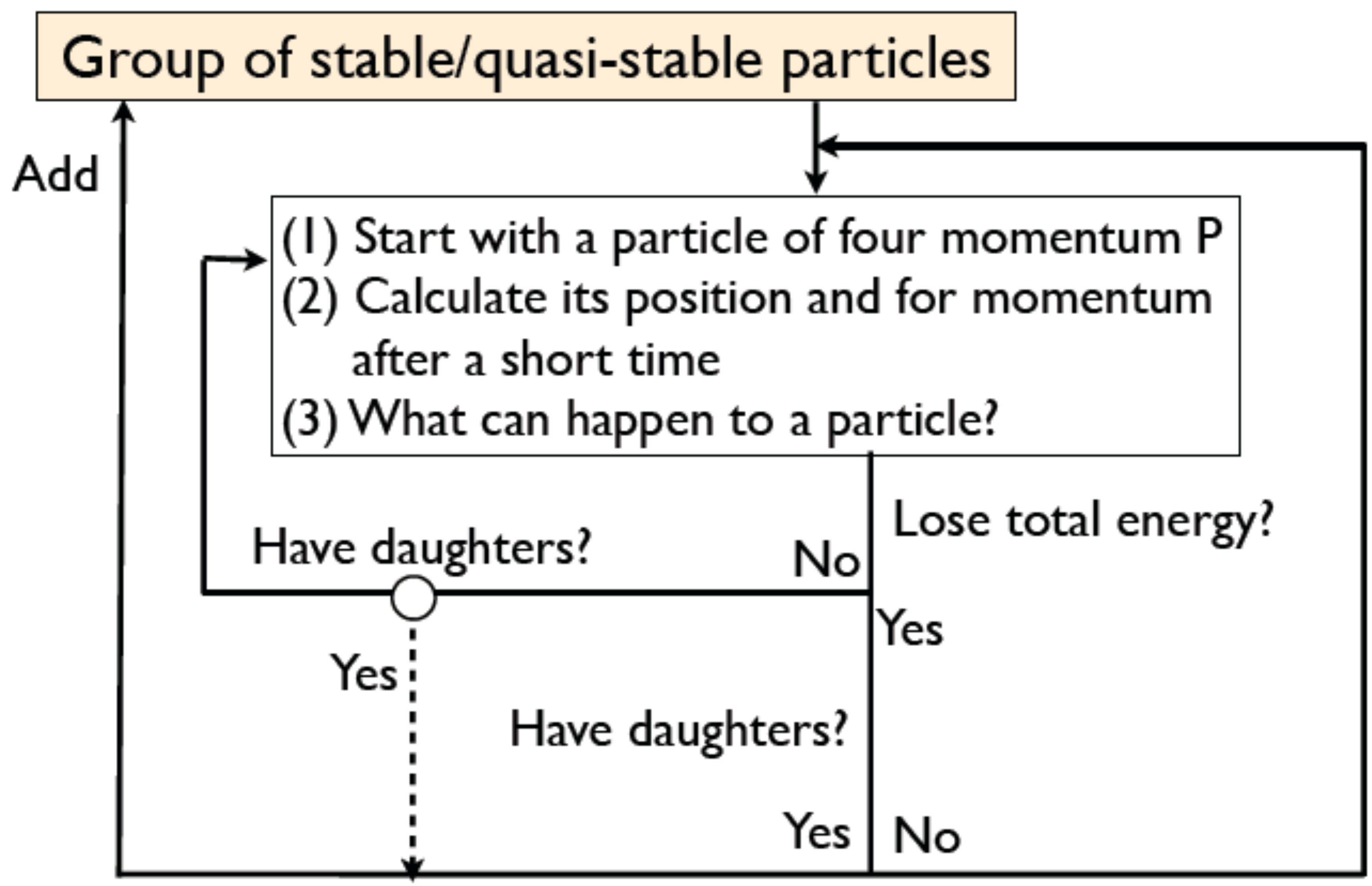
When you see plots like this. Most often the background and signal come at least in part from Monte Carlo simulations

Example: ATLAS calorimeter

Simulation of the calorimeter with Pb layers and liquid Ar detection gaps



What happens inside



Propagating particles

In electron case:

☑ With probability **p1**: ionize the gas, loose some momentum, produce N secondary electrons with momentum p_{e1N1},...

☑ Do nothing with probability **1-p1**

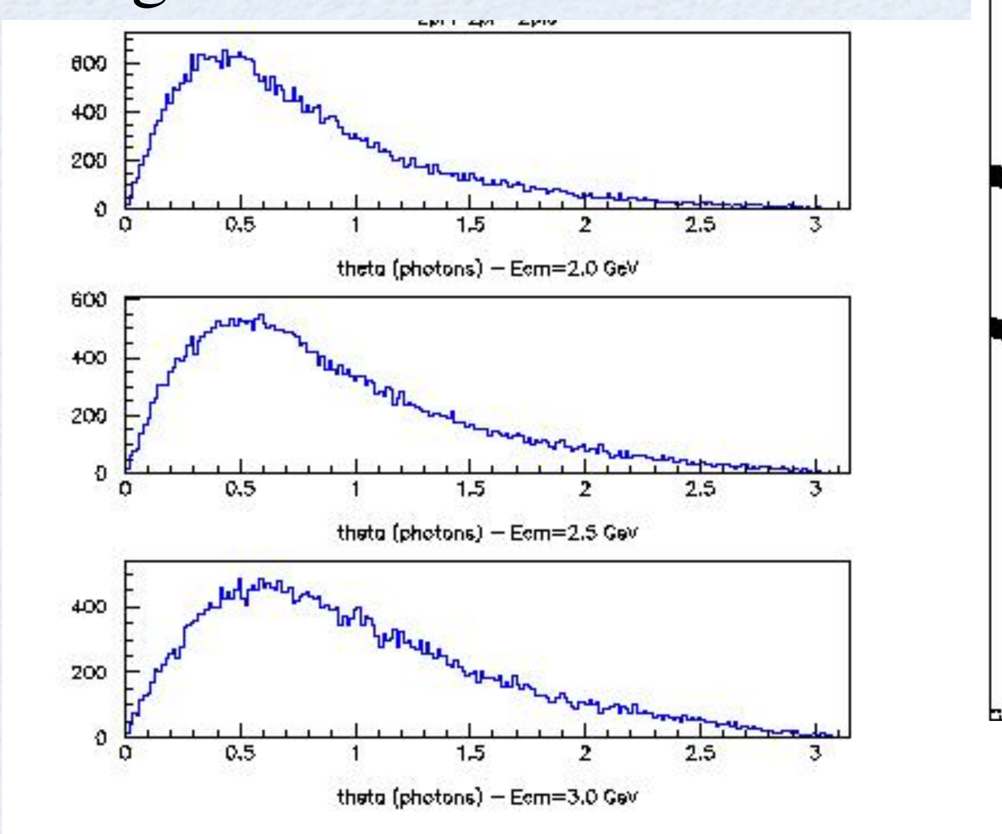
☑ Generate random number **r** in the range [0,1]

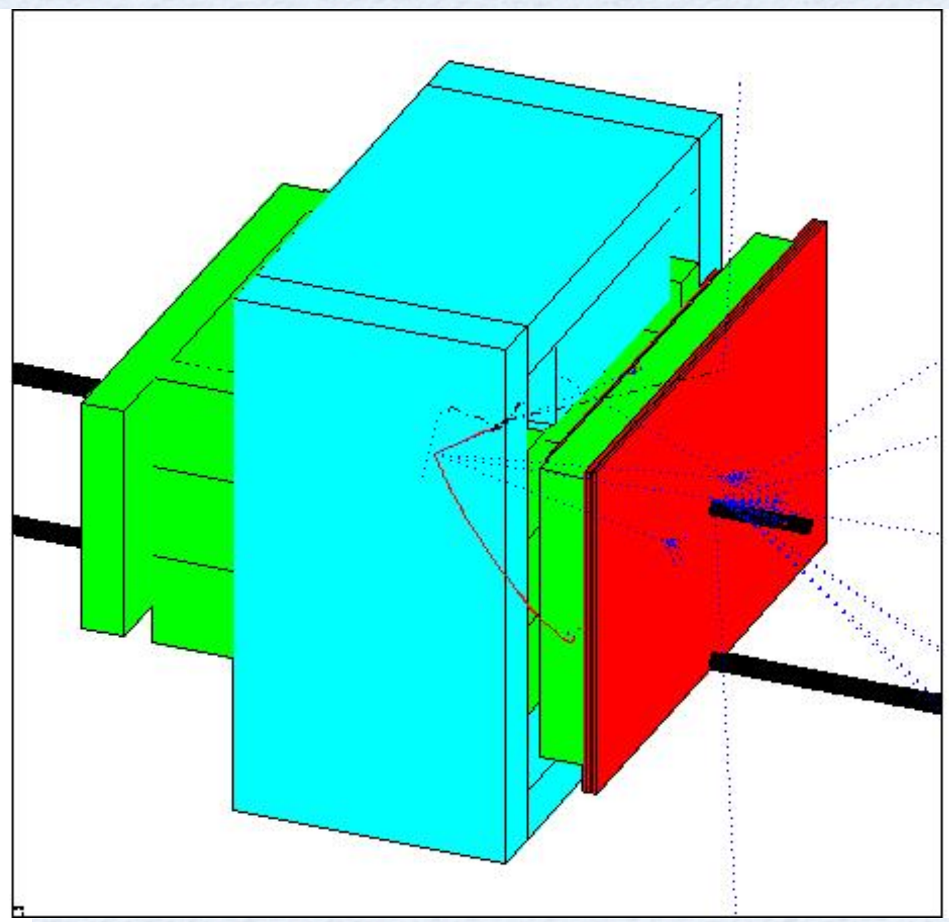
☑ If **r<p1**, generate momenta of secondary electrons, add them to particles list, reduce the momentum of initial electron.

Detector Acceptance

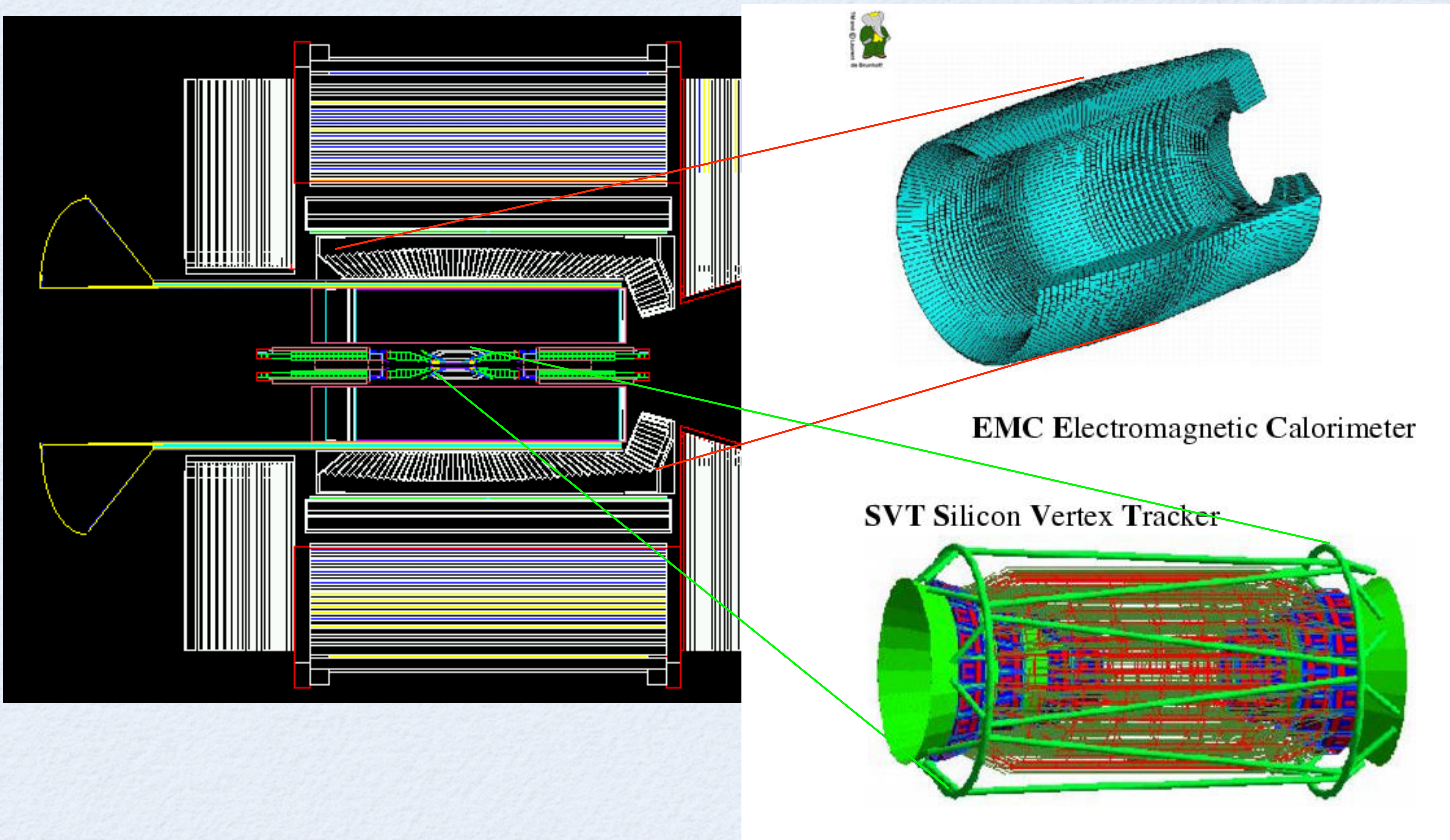
Study of the process:
$e^+ e^- \rightarrow \pi^+ \pi^- \pi^0$

Angular distribution of π^0 decay products for 3 different energies.



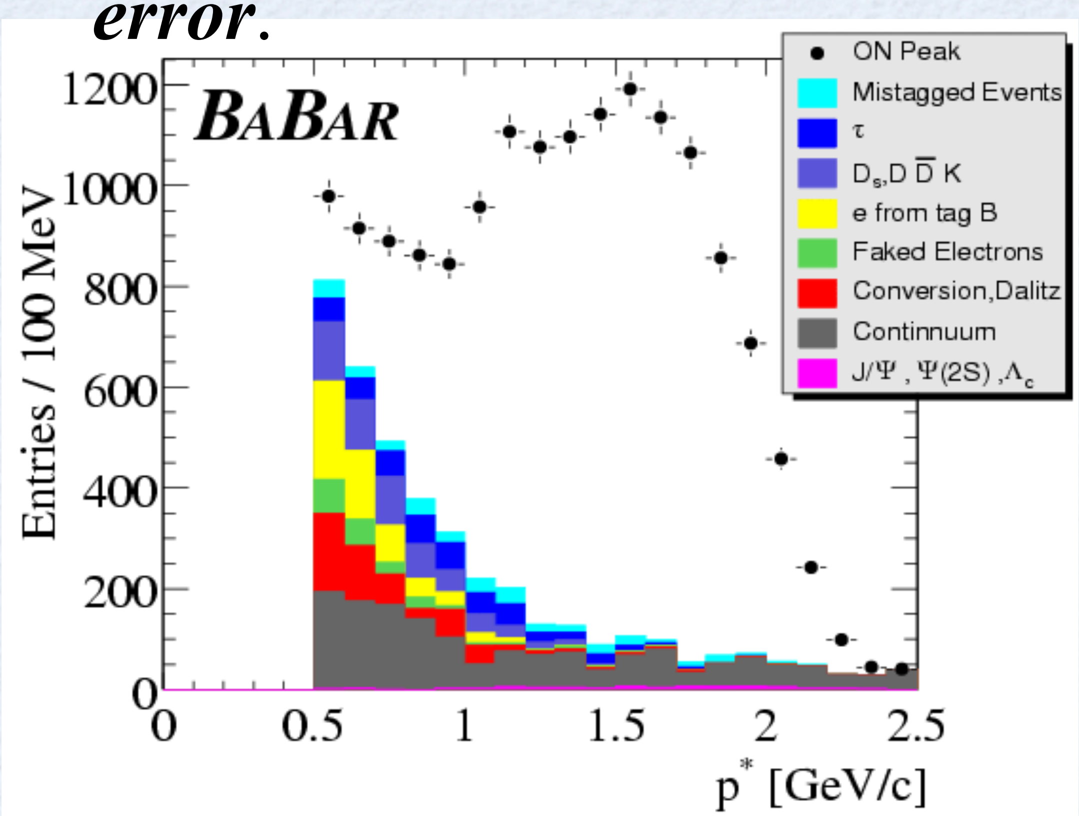


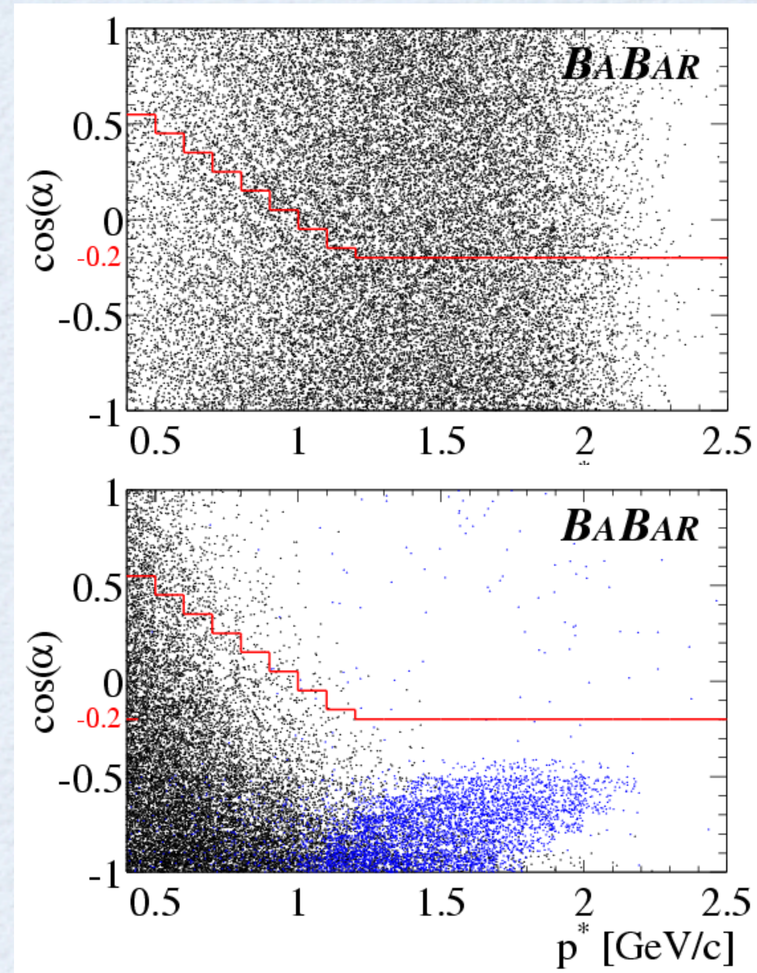
Detector Design



ANALYSIS

- Use MC simulation to compute the ***signal efficiency*** and *background contamination*.
- Optimize the selection criteria to get the ***smallest error***.



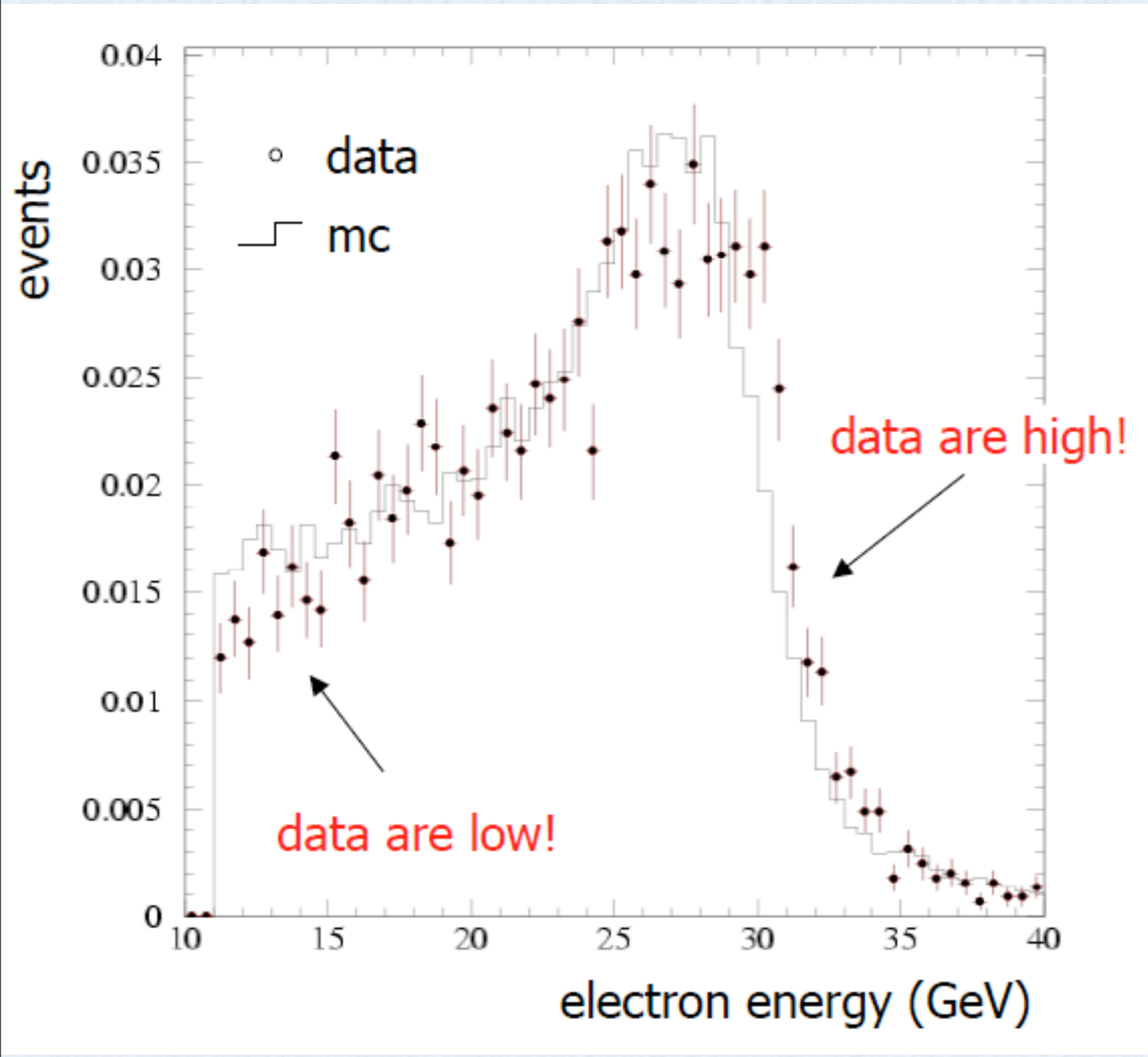


Signal MC

Background MC

Analysis

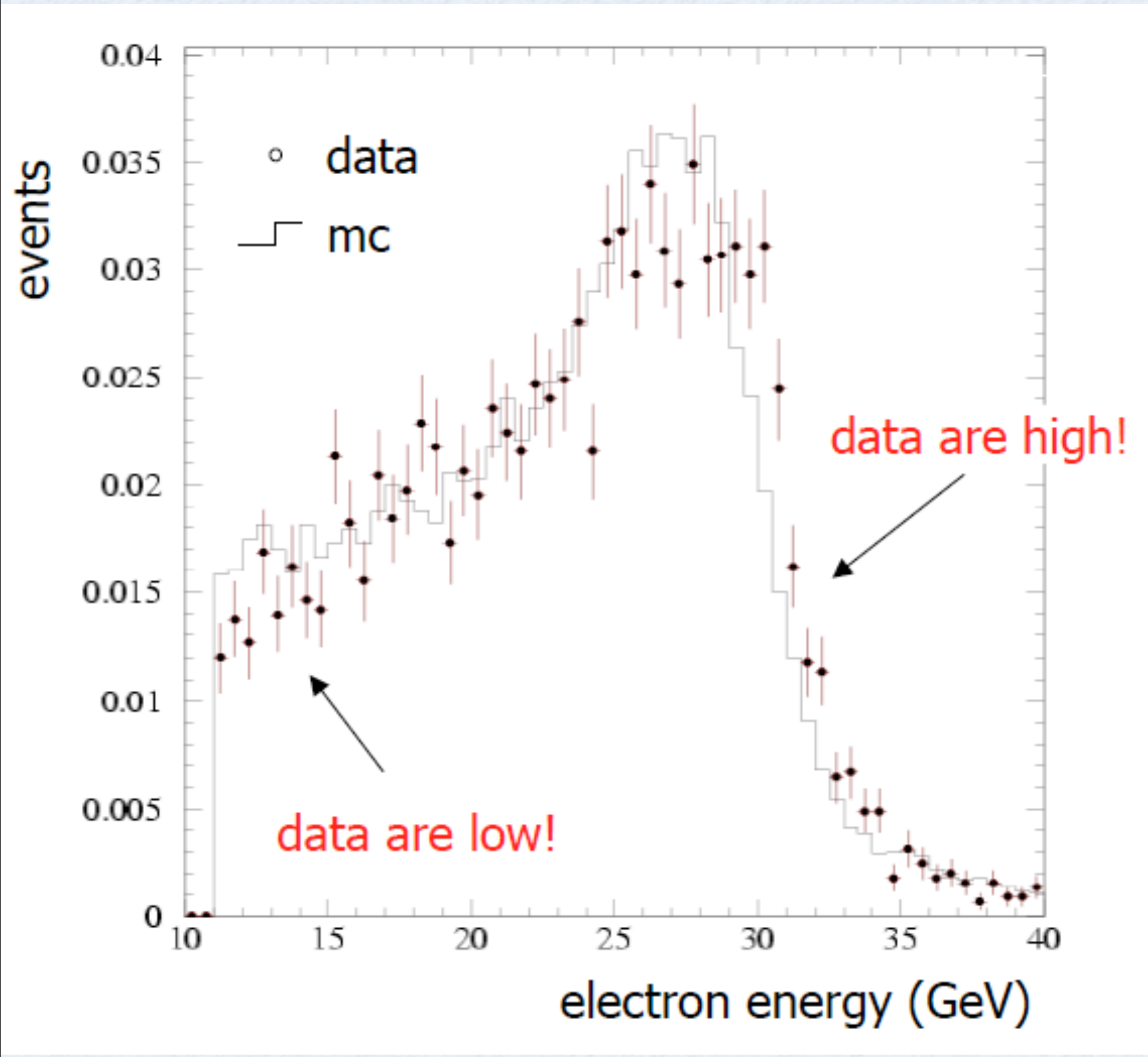
Data from deep inelastic scattering at DESY



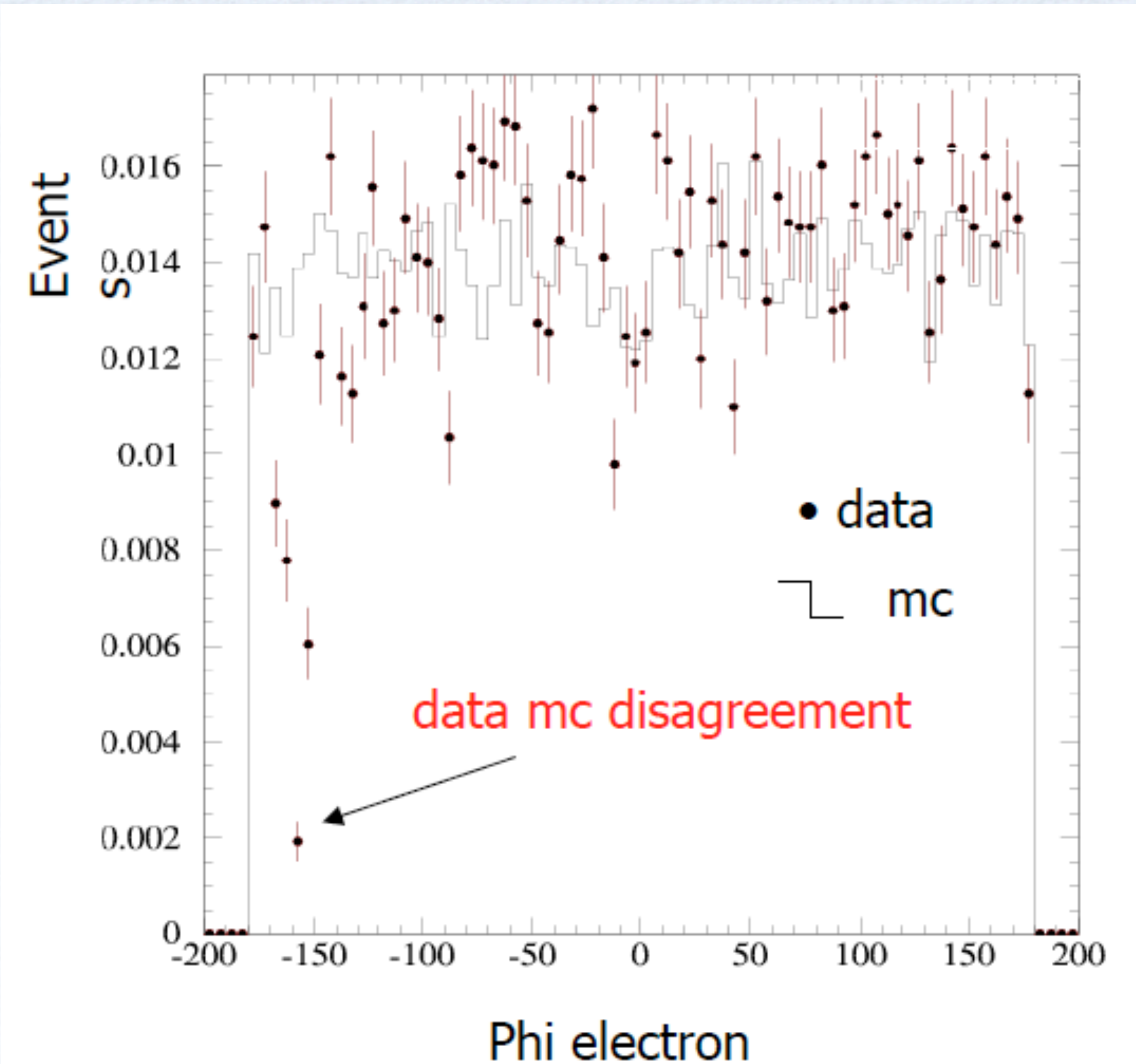
- Often detector resolution and acceptance it is difficult to make a direct comparison with theory
- Compare a simulation with all known physics effects to the measured data

Analysis

Data from deep inelastic scattering at DESY



- Monte Carlo over predicts due to problem with model of underlying physics
- Monte Carlo under predicts at high energy due to simulation having too good resolution



- Normalized Distribution of azimuthal angle - no physics in phi
- Any non uniformity comes from detector acceptance or detector issue
- Taken when a sector of the tracker was turned off in data (still present in simulation)

Next...

- More practice with C++ and root