Virtual Atom Smashers

Peter Skands (CERN Theoretical Physics Dept)



High School Teachers Program 2013 CERN

Background

Who am I?

Theorist (PhD 2004 from Lund U, Sweden) working on improving solutions to Quantum Chromodynamics (QCD)

- ... to write good "Monte Carlo event generators"
 - Used by experiments to give "theory predictions", to compare with data
 - Used to design and optimize detectors and analysis strategies
 - Used by theorists to explore new solutions, new ideas, new physics

Not a computer scientist. But the numerical calculations I (want to) do require a lot of power

→ distributed computing: farms / GRID / clouds In particular the LHC@Home 2.0 project "Test4Theory"



Background

Who am I?

Theorist (PhD 2004 from Lund U, Sweden) working on improving solutions to Quantum Chromodynamics (QCD)

... to write good "Monte Carlo event generators"

- Used by experiments to give "theory predictions", to compare with data
- Used to design and optimize detectors and analysis strategies
- Used by theorists to explore new solutions, new ideas, new physics

Not a computer scientist. But the numerical calculations I (want to) do require a lot of power

→ distributed computing: farms / GRID / clouds In particular the LHC@Home 2.0 project "Test4Theory"



Why did Mick invite me?

To explore possibilities for high school students/teachers in these areas? Some half-baked ideas and examples.

Nutshell





Theory

Experiment

Adjust this to agree with this

Nutshell





Theory

Experiment

Adjust this to agree with this

→ Science

In Practice





VINCIA

- "Virtual Colliders" = Simulation Codes
- Particle Physics Models, Algorithms, ...
- → Simulated Particle Collisions





Real Universe → Experiments & Data

Particle Accelerators, Detectors, and Statistical Analyses

→ Published Measurements



Tools for HEP Experiments



PYTHIA



Sjöstrand, Mrenna & Skands, Computer Physics Communications 178 (2008) 852

The most widely used event generator in HEP

Select incoming beams, energy, process type, ...

Generates collisions, particle decays, radiation, hadronization, ...

→ Simulated Events

What can you do with that?

Can count particles, identify their types, study their spectra Compare with measured experimental results Study effects of changing the model parameters \rightarrow physics

What does it require?

PYTHIA is written by \sim 7 people. It is \sim 6MB of standalone C++ code. Compiles easily on Macs and Linux (Windows not recommended). A bit of C++ experience and an ordinary laptop or desktop.

Help: many examples and a written tutorial (so far oriented towards scientists and university students). If useful for teaching, interesting to develop material targeted at younger people?

Theory ↔ Data Comparisons

Task: determine "best" parameters for theory models

→ Compare against thousands of measurements, taken under different conditions, by different experiments, at different colliders

+ do this for many simulators & versions, with different setups

LEP Tevatron SLC LHC ISR HERA SPS RHIC Quite technical Quite tedious

Ask someone else

Theory ↔ Data Comparisons

Task: determine "best" parameters for theory models

→ Compare against thousands of measurements, taken under different conditions, by different experiments, at different colliders

+ do this for many simulators & versions, with different setups

LEP Tevatron SLC LHC ISR HERA SPS RHIC Quite technical Quite tedious

Ask someone else everyone

Theory ↔ Data Comparisons

LEP

SLC

HERA

Task: determine "best" parameters for theory models

→ Compare against thousands of measurements, taken under different conditions, by different experiments, at different colliders

+ do this for many simulators & versions, with different setups

LHC@home 2.0





Tevatron

SPS

RHIC

Quite technical Quite tedious

Ask someone else everyone

> J. Blomer, P. Buncic, I. Charalimpidis, F. Grey, A. Haratyunyan, A. Karneyeu, D. Lombrana-Gonzalez, M. Marquina, B.Segal, P. Skands,

7000 Volunteers - 20000 Hosts Over 900 billion simulated collision events

LHC@Home 2.0 - Test4Theory

Idea: ship volunteers a virtual atom smasher (to help do high-energy theory simulations)

Runs when computer is idle. Sleeps when user is working.

Problem: Lots of different machines, architectures (tedious, technical)

Use Virtualization (CernVM) → provides standardized computing environment on any machine (in our case: Scientific Linux)

Replica of our normal working environment.

Separation of IT and Science

Virtualization: never previously done for a volunteer cloud

LHC@home 2.0 Test4Theory volunteers' machines seen during the past 24 hours (7011 machines overall)



LHC@home 2.0 Test4Theory volunteers' machines seen during the past 24 hours (7011 machines overall)



More than the sum of its parts

Test4Theory combines:

- Theoretical high-energy physics
- In particular the strong nuclear force

Random-number based Monte Carlo techniques

- & Markov-chain algorithms
- Programming
- Distributed computing
- In particular virtualization & cloud computing
- Citizen science
- Experimental particle physics Including statistical data analysis

Used to provide comprehensive tests of the theory models used at LHC.

Accurate theory → accurate measurements



Possible projects or study topics?

Maths

Basic statistics (and data analysis, can also be related to social studies, polling, etc) Random numbers Markov chains

Physics

The strong nuclear force (and explicit simulations of it) Setting up and running your own virtual atom smasher, and learning about the particles it produces

Computer Science

Virtual Machines (e.g., Linux-CernVM), with the Test4Theory project as a concrete example

Distributed computing, using the Test4Theory volunteer cloud as a real-world example that individuals can participate in

Example in Statistics: Fostering Healthy Skepticism

A friend pointed me to a story on FB

A Scottish study had found that the IQ of the eldest child was higher than that of the younger siblings

Disturbing news (I am the younger one)

A professor somewhere was commenting on what this meant about how we treat each child differently, etc.

The difference was 1%, on average

They had examined 10,000 cases

The statistical uncertainty for 10,000 counts is: 1%

With very little statistics knowledge, students could be given cases like this to debunk, and see it work in practice

Other more CERN-related examples could draw from data analysis and particle physics. How do you know you've made a discovery? Significance.

Charges Stopped or kicked

Charges Stopped or kicked

Associated field (fluctuations) continues

Charges Stopped or kicked

Radiation

Associated field (fluctuations) continues

Radiation

Charges Stopped or kicked

Radiation

a.k.a. Bremsstrahlung Synchrotron Radiation

Radiation

The harder they stop, the harder the fluctations that continue to become radiation

Jets = Fractals!

Bjorken scaling

To first approximation, gauge theories (like QCD) are SCALE INVARIANT

A quantum fluctuation inside a fluctuation inside a fluctuation ...

A gluon emits a gluon emits a gluon emits a gluon ...

If the coupling "constant" of the strong force was a constant, this would be absolutely true



Asymptotic Freedom

"What this year's Laureates discovered was something that, at first sight, seemed completely contradictory. The interpretation of their mathematical result was that the closer the quarks are to each other, the *weaker* is the 'colour charge'. When the quarks are really close to each other, the force is so weak that they behave almost as free particles. This phenomenon is called 'asymptotic freedom'. The converse is true when the quarks move apart: the force becomes stronger when the distance increases."



Nobelprize.org

2004

The Official Web Site of the Nobel Prize

The Nobel Prize in Physics 2004 David J. Gross, H. David Politzer, Frank Wilczek



David J. GrossH. David PolitzerFrank WilczekThe Nobel Prize in Physics 2004 was awarded jointly to David J. Gross, H. David Politzer and FrankWilczek "for the discovery of asymptotic freedom in the theory of the strong interaction".

Photos: Copyright © The Nobel Foundation

Asymptotic Freedom

"What this year's Laureates discovered was something that, at first sight, seemed completely contradictory. The interpretation of their mathematical result was that the closer the quarks are to each other, the *weaker* is the 'colour charge'. When the quarks are really close to *1 each other, the force is so weak that they behave almost as free particles. This phenomenon is called 'asymptotic freedom'. The converse is true when the quarks move apart: *2 the force becomes stronger when the distance increases."

Nobelprize.org

2004

The Official Web Site of the Nobel Prize

The Nobel Prize in Physics 2004 David J. Gross, H. David Politzer, Frank Wilczek



David J. GrossH. David PolitzerFrank WilczekThe Nobel Prize in Physics 2004 was awarded jointly to David J. Gross, H. David Politzer and FrankWilczek "for the discovery of asymptotic freedom in the theory of the strong interaction".

Photos: Copyright © The Nobel Foundation



^{*1} The force still goes to ∞ as $r \rightarrow 0$ (Coulomb potential), just less slowly

^{*2} The potential grows linearly as $r \rightarrow \infty$, so the force actually becomes constant (even this is only true in "quenched" QCD. In real QCD, the force eventually vanishes for r>>1fm)

The Strong Coupling "Constant"

The Strong Coupling "Constant" as function of energy scale, Q

From PDG Review on QCD. by Dissertori & Salam 0.5 July 2009 $\alpha_{s}(\mathbf{Q})$ △ ▲ Deep Inelastic Scattering $\circ \bullet e^+e^-$ Annihilation 0.4 □

■ Heavy Quarkonia 0.3 0.2 Freedom? Unification 0.1 \equiv QCD $\alpha_s(M_Z) = 0.1184 \pm 0.0007$ 1 10 100 **O** [GeV]

At low scales

Coupling $a_s(Q)$ actually runs rather fast with Q

Perturbative solution diverges at a scale Λ_{QCD} somewhere below

≈ 1 GeV

So, to specify the strength of the strong force, we usually give the value of a_s at a unique reference scale that everyone agrees on: M_Z = 91.2 GeV/c

The Strong Coupling "Constant"



At low scales

Coupling $a_s(Q)$ actually runs rather fast with Q

Perturbative solution diverges at a scale Λ_{QCD} somewhere below

 \approx 1 GeV

So, to specify the strength of the strong force, we usually give the value of a_s at a unique reference scale that everyone agrees on: M_Z = 91.2 GeV/c

Confinement

We don't see quarks and gluons ...

Mesons

Quark-Antiquark Bound States $\pi^0, \pi^{\pm}, K^0, K^{\pm}, \eta, \dots$





Quark-Quark-Quark Bound States

 $p^{\pm}, n^0, \Lambda^0, \ldots$

Potential between a quark and an antiquark as function of distance, R



Potential between a quark and an antiquark as function of distance, R



Potential between a quark and an antiquark as function of distance, R



Potential between a quark and an antiquark as function of distance, R

K(R) 0.9linear par 0.8 total 0.7 Short Distances ~ "Coulomb" 0.6 Coulomb part 0.5 0.4 $V(R) = V_{p} + K R - e/R + f/R^{2}$ Partons 0.3 12 16 24 8 20R

Long Distances ~ Linear Potential



Quarks (and gluons) confined inside hadrons

Potential between a quark and an antiquark as function of distance, R

Long Distances ~ Linear Potential



Quarks (and gluons) confined inside hadrons





(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:



(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:



Assume you know the area of <u>this</u> shape: πR^2 (an overestimate)

(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:

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



Assume you know the area of <u>this</u> shape: πR² (an overestimate)

(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:

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



Assume you know the area of <u>this</u> shape: πR² (an overestimate)

(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:

Now get a few friends, some balls, and throw random shots inside the circle (but 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 <u>this</u> shape: πR² (an overestimate)

(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:

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

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



(apologies, I did not have much time to adapt these slides)

You want: to know the area of this shape:

Now get a few friends, some balls, and throw random shots inside the circle (but 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 <u>this</u> shape: πR² (an overestimate)



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

 $A_{a} \approx N_{hit}/N_{miss} \times \Pi R^{2}$

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

Convergence:

 $\label{eq:alpha} \begin{array}{l} \textbf{Calculus:} \ \{A\} \ \text{converges to B} \\ \text{if an n exists for which} \\ |A_{i>n} - B| < \epsilon, \text{ for any } \epsilon > 0 \end{array}$

Monte Carlo: {A} converges to B if n exists for which the probability for |A_{i>n} - B| < ε, for any ε > 0, is > P, for any P[0<P<1]</p> A Monte Carlo technique: is any technique making use of random numbers to solve a problem

Convergence:

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

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

I will not tell you how to *write* a Random-number generator (interesting topic & history in its own right)

Instead, I <u>assume</u> that you can write a computer code and link to a random-number generator, from a library

E.g., ROOT includes one that you can use if you like.

PYTHIA also includes one

From the PYTHIA 8 HTML documentation, under <u>"Random Numbers"</u>:

Random numbers R uniformly distributed in 0 < R < 1 are obtained with

```
Pythia8::Rndm::flat();
```

+ Other methods for exp, x^*exp , 1D Gauss, 2D Gauss.

Example: Number of summer students who will get hit by a car during the next 3 weeks

Complicated Function:

Time-dependent

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

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

No two students are the same

Need to compute probability for each and sum

(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.



Approximate Student

by most accident-prone student / famous person / movie star / ...

(not making fun of handicapped people ...)

This extreme guess will be the equivalent of our simple overestimate from before:



Off we go...

Throw random accidents according to:



Off we go...

Throw random accidents according to:



Too

Off we go...

Throw random accidents according to:



Off we go...

Throw random accidents according to:



Simple

Verestimate

Off we go...

Throw random accidents according to:



 $\mathsf{R} = (t_e - t_0) \Delta x \quad \alpha_{\max} \; n_{\text{stud}} \; \rho_{c\max}$

Hit rate for most accident-prone student Rush-hour density of cars Simple Overestimate

Accept trial hit (i,x,t) with probability

Prob(accept) =
$$\frac{\alpha_i(x,t) \ \rho_i(x,t) \ \rho_c(x,t)}{\alpha_{\max} \ n_{\text{stud}} \ \rho_{c\max}}$$

Using the following:

 ρ_c : actual density of cars at location x at time t ρ_i : actual density of student i at location x at time t α_i : The actual "hit rate" (OK, not really known, but can make one up)

Accept trial hit (i,x,t) with probability

Prob(accept) =
$$\frac{\alpha_i(x,t) \ \rho_i(x,t) \ \rho_c(x,t)}{\alpha_{\max} \ n_{\text{stud}} \ \rho_{c\max}}$$

Using the following:

 ho_c : actual density of cars at location x at time t ho_i : actual density of student i at location x at time t ho_i : The actual "hit rate" (OK, not really known, but can make one up)

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



The "Jeppsson" Project April 2010

The Jeppsson Project

April 2009: FB message from friend of friend: can a 15-yr old be a oneweek intern at CERN?

We were developing a run-time display for our simulation anyway.

April 2010: simple text editor to edit input cards. Run-time display to compare output histograms to data.

The Jeppsson Project

April 2009: FB message from friend of friend: can a 15-yr old be a oneweek intern at CERN?

We were developing a run-time display for our simulation anyway.

April 2010: simple text editor to edit input cards. Run-time display to compare output histograms to data.

! * Strong-force Co Vincia:alphaSValue	upling = 0.138
! * Hadronic Energy Vincia:cutoffScale	y Scale = 0.45
! * String paramete	ers
StringZ:aLund	= 0.38
StringZ:bLund	= 0.62
StringPT:sigma	= 0.26

! * Quark flavor parameters StringFlav:probStoUD = 0.21 StringFlav:mesonUDvector = 0.35 StringFlav:mesonSvector = 0.55 StringFlav:probQQtoQ = 0.08 StringFlav:probQQtoQQ = 1.00 StringFlav:probQQ1toQQ0 = 0.03 StringFlav:decupletSup = 1.00 StringFlav:etaSup = 0.60 StringFlav:etaPrimeSup = 0.10

The Jeppsson Project

April 2009: FB message from friend of friend: can a 15-yr old be a oneweek intern at CERN?

We were developing a run-time display for our simulation anyway.

April 2010: simple text editor to edit input cards. Run-time display to compare output histograms to data.



! * Strong-force Co	upling
Vincia:alphaSValue	= 0.138
! * Hadronic Energy	Scale
Vincia:cutoffScale	= 0.45
! * String paramete	ers
StringZ:aLund	= 0.38
StringZ:bLund	= 0.62
StringPT:sigma	= 0.26

! * Quark flavor parameters StringFlav:probStoUD = 0.21 StringFlav:mesonUDvector = 0.35 StringFlav:mesonSvector = 0.55 StringFlav:probQQtoQ = 0.08 StringFlav:probSQtoQQ = 1.00 StringFlav:probQQ1toQQ0 = 0.03 StringFlav:decupletSup = 1.00 StringFlav:etaSup = 0.60 StringFlav:etaPrimeSup = 0.10

The Jeppss

April 2009: FB message from friend of friend: can a 15-yr old be a oneweek intern at CERN?

We were developing a run-time display for our simulation anyway.

April 2010: simple text editor to edit input cards. Run-time display to compare output histograms to data.

May 2010: Parameters released as new defaults.



The Jeppss

April 2009: FB message from friend of friend: can a 15-yr old be a oneweek intern at CERN?

We were developing a run-time display for our simulation anyway.

April 2010: simple text editor to edit input cards. Run-time display to compare output histograms to data.

May 2010: Parameters released as new defaults.

→ Citizen Cyberlab Pilot Project



The Citizen Cyberlab ICT Project

Standalone 3-yr Project funded by EU (2012-2015)

CERN Task: create citizen science pilot project in particle physics

The EU funds a 2-year "CERN fellowship", started in May: Ioannis Charalimpidis

We will

Develop an application that lets citizen scientists learn about, interact with, and optimize high-energy physics simulations, by comparing them to real data

 \rightarrow feedback to scientists

How?

Combine the framework and lessons from Test4Theory / LHC@home 2.0 with those from the Jeppsson project \rightarrow **Atom Smasher Application**

Provide content, explanations, visualizations (modifiable and open)

Organize one or more **citizen-science events** at CERN (e.g., hack fests, event for CERN open day in September), host **summer students** (IT or Physics masters student), ...

The Citizen Cyberlab ICT Project

Standalone 3-yr Project funded by EU (2012-2015)

CERN Task: create citizen science pilot project in particle physics

The EU funds a 2-year "CERN fellowship", started in May: Ioannis Charalimpidis

We will

Develop an application that lets citizen scientists learn about, interact with, and optimize high-energy physics simulations, by comparing them to real data

 \rightarrow feedback to scientists

We now have ~ 6 months to develop the first rough prototype

How?

Combine the framework and lessons from Test4Theory / LHC@home 2.0 with those from the Jeppsson project \rightarrow Atom Smasher Application

Provide content, explanations, visualizations (modifiable and open)

Organize one or more **citizen-science events** at CERN (e.g., hack fests, event for CERN open day in September), host **summer students** (IT or Physics masters student), ...

Possible study / project topics?

Maths

Basic statistics (and data analysis, can also be related to social studies, polling, etc) Random numbers Markov chains

Physics

The strong nuclear force (and explicit simulations of it) Setting up and running your own virtual atom smasher, and learning about the particles it produces

Computer Science

Virtual Machines (e.g., Linux-CernVM), with the Test4Theory project as a concrete example

Distributed computing, using the Test4Theory volunteer cloud as a real-world example inviting participation