Simulation Basics in CMS

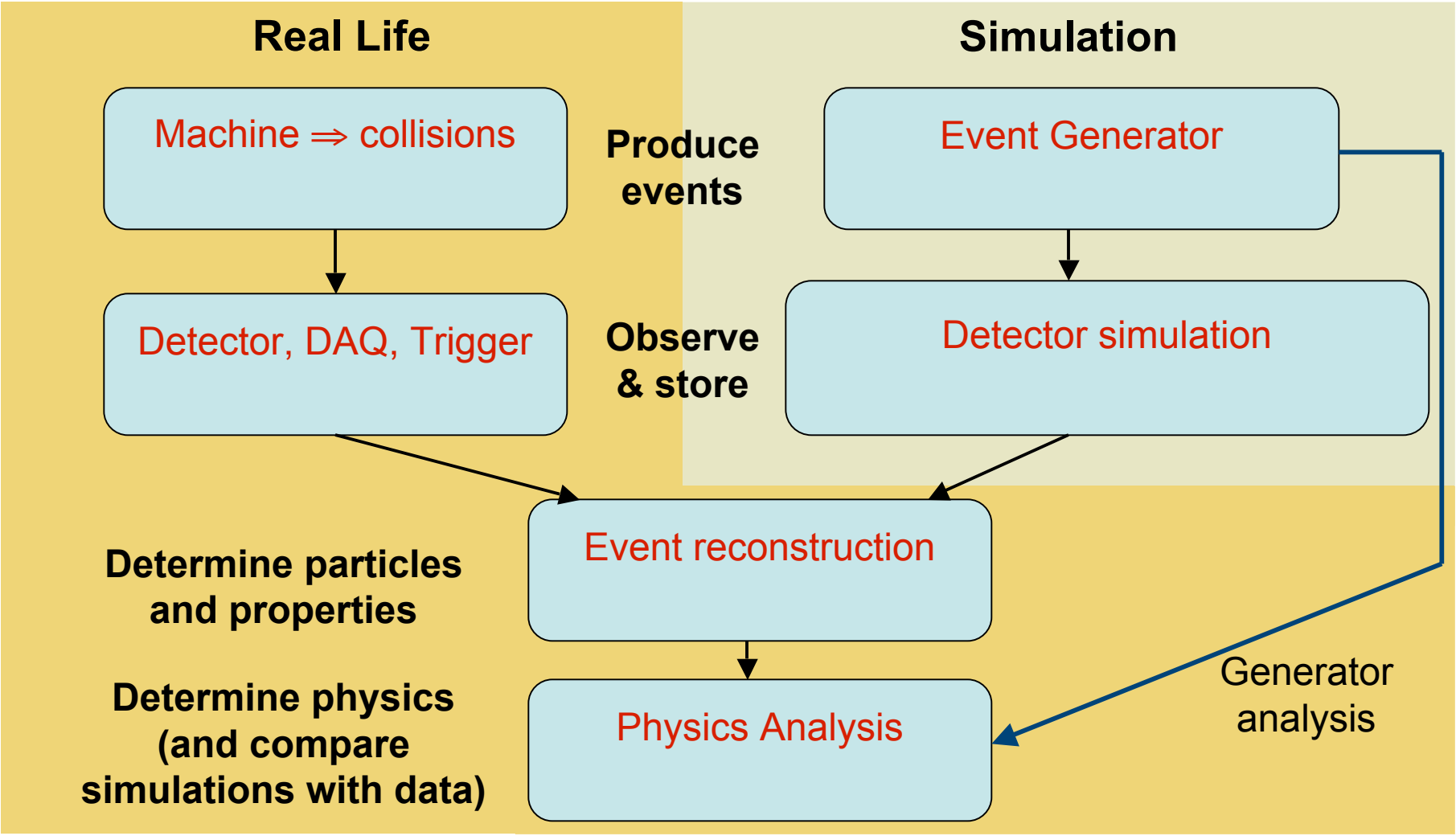
- Introduction to simulations
  - The Simulation process
  - Beyond the tutorial (asides)
- An interesting quote

Harry Cheung
Fermilab
All USCMS Meeting, at LPC
5/26/2006

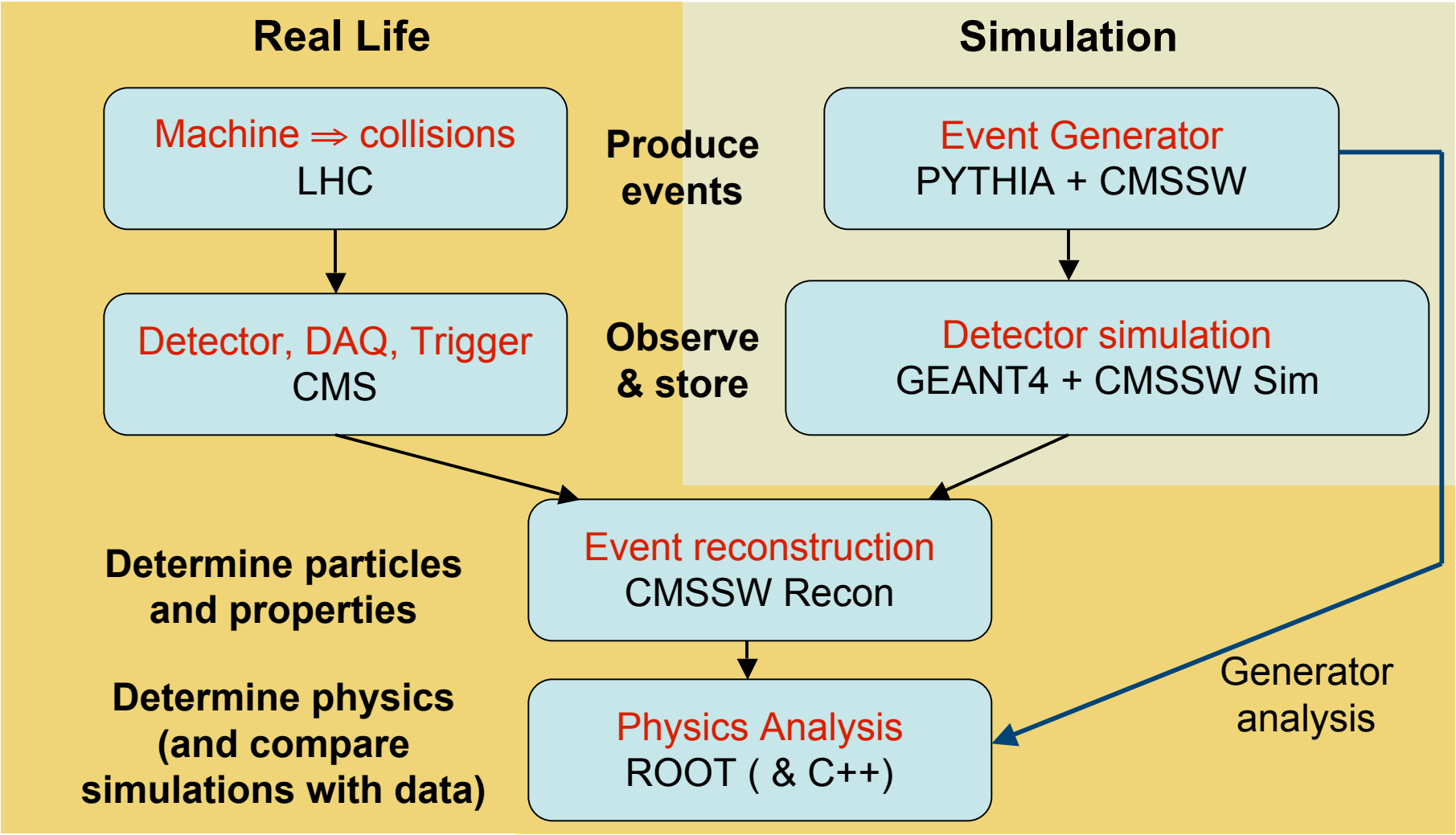
Introduction: Why Simulations?

- Simulations are used throughout a physics experiment:
  - Design of the experiment
  - Design and optimization of individual detectors
  - Predicting backgrounds
  - Understanding the data (simulations = numeric calculations)
- …including during physics data analysis:
  - Correcting for detector effects on data (acceptance, resolution)
  - For simulation based corrections to the data (variations of response, due to linearity, cracks)
    - Correcting the data *vs.* simulation of the effect
  - Extracting signals from data (backgrounds, shapes)
- So simulations must be validated:
  - "One time" validation: of the geometry, physics processes
  - Regular Incremental validation: for different code versions
  - **What is good enough?**

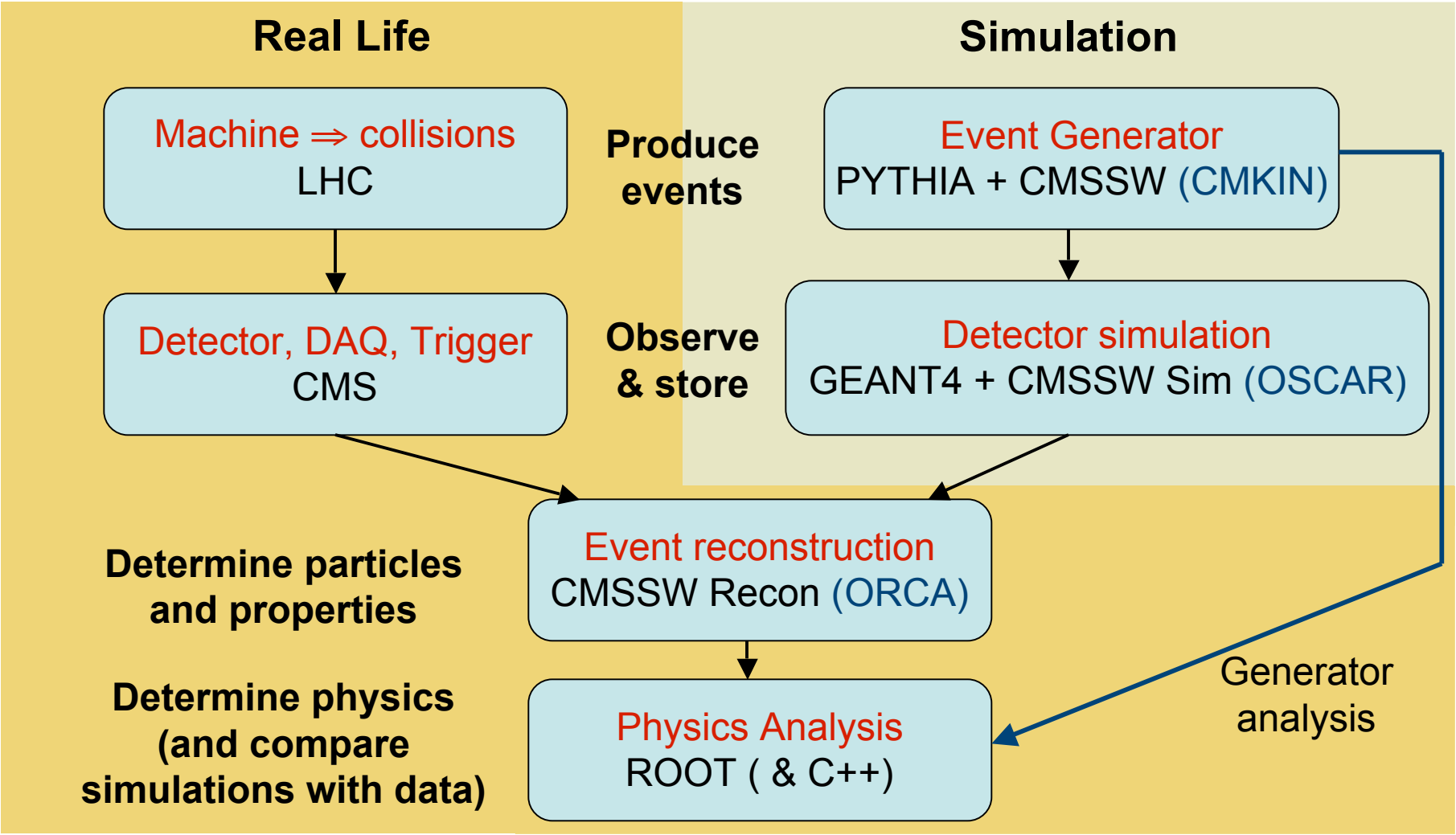
Intro: The Simulation Process



Intro: The Simulation Process



Intro: The Simulation Process



Intro: CMS Simulation Tutorial

- Already some excellent simulation tutorial material:
  - From May 2006 CPT week (written by Fabio Cossutti, based on original tutorial from Julia Yarba) https://twiki.cern.ch/twiki/bin/view/CMS/May06CPTweekSimuChain
  - Tutorial being incorporated into the CMS Workbook (by Anne Heavey and Fabio Cossutti) https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookGeneSimDigi
  - Some simulation output samples are available (thanks to Dave Evans and Oliver Gutsche), see https://twiki.cern.ch/twiki/bin/view/CMS/TutorialSamples and the LPC Physics Working group will generate more samples for physics (see Boaz Klima), or see Hans Wenzel if you want help with running in batch to generate your own.
- Other tutorials available from May 2006 CPT week
  - Validation, reconstruction, analysis tools, and “batch job” running https://twiki.cern.ch/twiki/bin/view/CMS/May06CPTweekTutorials

Intro: Simulation Tutorial Output

- Looking at the ROOT output file from the tutorials
  - Use bare ROOT to browse the events
  - Use ROOT macros (see tutorial)
  - Use CMSSW and write an EDAnalyzer (see next slide)
- Some of the objects you will find:
  - edmHepMCProduct_*label* gives info on the particles produced by the generator. *label* (name in cfg) gives the source of this object
    - PythiaSource or flatRandomEGunSource
    - VtxSmeared (same but with smeared vertex, if used)
    - Format is HepMC::GenEvent from CLHEP project, see http://cmsdoc.cern.ch/swdev/lxr/CMSSW/source/clhep/CLHEP/HepMC/GenEvent.h?v=0.6.0 or http://proj-clhep.web.cern.ch/proj-clhep/
  - EmbdSimTracks_*SimG4Object*, EmbdSimVertexs_*SimG4Object* produced by OscarProducer contain info for generator particles selected for Geant4 tracing, and some secondaries produced while tracing through CMS detector.

Intro: Looking More at the Output

- Some of the objects you will find (continued):
  - PCaloHits_SimG4Object_*type* and PSimHits_SimG4Object_*type* produced by OscarProducer contain the simulated hits before digitization (detector response)
  - EventAux and Provenance contain the event id information and information on the products respectively and come as part of the CMSSW framework.
  - *_*detector*digi produced by the digitization steps
- Looking at the data with the CMSSW framework:
  - For examples see the validation packages (also see https://twiki.cern.ch/twiki/bin/view/CMS/May06CPTweekValidationDemo)
    - Using EDAnalyzer: Validation/EcalDigis/test/EcalDigisAnalysis.cfg, and Validation/EcalDigis/src/EcalDigisValidation.cc
    - For EDProducer: Validation/GlobalHits/test/GlobalValProducer.cfg and Validation/GlobalHits/src/GlobalValProducer.cc
  - Also talk to people working on the different LPC working groups

Intro: The Generator

- A single particle gun gives events with one (or more) particle(s) from same vertex for tests
- A generator is needed to get events of interest, a general purpose generator is needed to do many things
  - The hard scattering process, parton showers, resonance decays, underlying events, hadronization, and ordinary decays
  - Available generator in CMSSW is Pythia 6.227 (see http://www.thep.lu.se/~torbjorn/Pythia.html) and the link there for the Pythia manual (you need this!)
  - The tutorial gives an example of minimum bias events
  - More examples are coming, for now see for example RecoTracker/RoadSearchCloudMaker/test/Sim_pythiaHZZmumumumu_To_Tracks.cfg (thanks to Oliver Gutsche) to see how to change the Pythia parameters to get $H \rightarrow ZZ \rightarrow \mu\mu\mu\mu$. See also JetMet example ("simulation of jets").
- An excellent tutorial series on PYTHIA given by Torbjörn Sjöstrand http://agenda.cern.ch/fullAgenda.php?ida=a042790

Aside: More Generators

- There are specialized generators for various physics:
  - There are many! E.g. ALPGEN, GR@PPA, EvtGen, AcerMC,… see Sjostrand's lectures (link on previous slide)
  - Still need a general purpose generator, PYTHIA, HERWIG, ISAJET, SHERPA, each have many parameters.
  - Only PYTHIA at the moment in CMSSW, and we have resident experts on PYTHIA, Steve Mrenna (WH7E), Peter Skands (WH3W). They gave talks at CDF together with T. Sjöstrand http://www-cdf.fnal.gov/physics/lectures/pythia_Dec2004.html and Bryan Webber gave talks on HERWIG http://www-cdf.fnal.gov/physics/lectures/herwig_Oct2004.html
- CMS has a dedicated generator group
  - See http://cmsdoc.cern.ch/cms/PRS/gentools/
  - Contact Filip Moortgat if you want to get involved in generators
  - If you have files generated in with CMKIN, you may be able to use them, you can check with Julia Yarba (see Workbook)

Intro: Other Physics Processes

- Besides the initial collision there are other interactions:
  - Interaction of generated particles with CMS detector and other material, secondaries generated will also interact and/or decay
  - This is handled in GEANT4 as part of the detector simulation, includes electromagnetic and hadronic interactions, and decays
  - Some processes are well known or calculable but still a large choice of processes, e.g. for electromagnetic we have
    - Photon processes (compton scattering, gamma conversion, photoelectric effect, muon pair production)
    - Electron/positron processes (ionization and delta ray production, bremsstrahlung, positron annihilation, synchrotron radiation)
    - Muon processes (Ionization and delta ray production, Bremsstrahlung, e+e- pair creation)
    - Hadron/ion processes (ionization)
    - Multiple scattering
  - Need to choose thresholds for secondary particle production
  - Cannot do each calculation, need cross section tables and approximations

Intro: Geant4 and Physics Tables

- More complicated for poorly known processes like hadron interactions
  - Implementation of physics processes using data-driven or theory-driven models, or parameterized cross-sections (in Geant4 the user is meant to select these!)
- Geant4 Physics Tables to make things simpler
  - Gives for each particle type which physics processes to use and how the cross sections are computed for each
    - LHEP: GHEISHA ported from Geant3 (parameterized inelastic)
    - **QGSP:** GHEISHA (E<25 GeV), q-g string model (E>25 GeV)
    - QGSC: as QGSP but chiral invariant PS decay fragmentation
    - FTFP: as QGSP but with fragmentation ala FRITJOF
  - Tutorial uses QGSP, for more information see links below
    - Geant4 Users Guide: physics processes, physics reference manual, physics table description, and example physics tables.
    - SLAC Geant4 group physics tables.

Aside: What About Decays?

- Decays can be in the generator and in the Geant4 Physics Table (I'm still figuring out which does what…)
  - PYTHIA can decay the particles it generates, there are many parameters in PYTHIA you can set to control decays, but heavy quark (b and c) decays are not handled in the best way.
  - Geant4 also has decay tables that can be used and constructed as part of making the Physics Table - usually for the simpler particle decays.
  - To see what is going on:
    - Look at the PYTHIA parameter settings for decays (and you should probably talk to a PYTHIA expert)
    - For Geant4 in CMS, the decay tables are made when particle types are "constructed" in making the Physics Table. Decays appear to be hardwired in code, see for example on cmsuaf at the file /uscmst1/prod/sw/cms/lcg/external/geant4-share/7.1.clhep1922-LCaptureFix/shared/source/particles/hadrons/mesons/src/G4Eta.cc for more information.

Aside: Geant4 *vs* Geant3

- Geant4 is in C++ and is set up to easily be tailored for more types of physics (astroparticle, nuclear physics,…)
- Geant4 geometry is done differently and there are new volume types, and a nice visualization tool. (In CMSSW, the geometry is split into a part in XML but also can be in algorithms in code.)
- Many other differences, see the Geant4 introduction.
- What particle id numbers do we use inside Geant?
  - CMSSW loads a particle ID table (in PDG id scheme) from $HEPPDT_PARAM_PATH/data/PDG_mass_width_2004.mc, see also http://pdg.lbl.gov/2005/mcdata/mc_particle_id_contents.html.
- Does Geant4 have the Geant3-style energy cutoffs for tracing?
  - Not by default. One can give instead a range (in length) for particular particles, materials or volumes. This is translated internally into energy thresholds for production of secondaries. All particles produced will be traced to zero energy. (Production thresholds are actually more complicated, see this part in the Geant4 users manual.)
- Multiple scattering is done differently than Geant3 , must look at the Geant4 Physics reference manual

Intro: The Digitization Step

- Detector response must be simulated
  - Geant4 is used to trace particles through the CMS detector and simulate interactions, production of secondaries, decays, multiple scattering
  - Response of each detector is done in the digi step. Do not want to use something like Geant4 to simulate the detector response, instead we parameterize the response from quantities given by the Geant4 simulation (track entry/exit points, energy deposition)
  - Some detectors need more parameterization than others, e.g. for the pixel sensors, each pixel is not defined in Geant4. "Simhits" must be stored and used at the digi step to turn on the right pixels and give the right charge deposition in each pixel that should be hit (complicated by presence of B-field).
  - Some detectors are complicated because signals are sampled every 25ns - previous and future crossings can affect the current one
- Detector response best done by the PRS groups as this is complicated and intimate knowledge of the detector and readout is required. Also linked with physics validation.

Intro: Simulations Validation

- The value of a simulation depends on how well the output mimics real data.
  - Validation of the physics processes (e.g. cross sections)
    - Physics Table, and Range cutoffs or approximations used
  - Validation of the geometry, including all materials
  - Validation of the detector responses (test beam or real data)
  - Misc. - e.g. environment, like beam backgrounds (can't take out)
    - Beam related, pileup, cosmic…
  - Validation of the actual computer code and updates
  - Deciding on what is good enough!
- Physics Validation is hard!
  - E.g. see the talk I gave at an LPC Physics working group meeting on Jan 26
- CMS has a Physics Validation group within the CMS simulations group (See Daniel Elvira)

Intro: Multiple Interactions

- The tutorial PYTHIA source will give only one pp collision
  - The pp crossing rate is 40 MHz for 25ns crossing time. At 10^{34} $cm^{-2}s^{-1}$ and with $\sigma_{tot}\approx$100mb, we will have 1GHz of interactions, or ≈25 interactions per crossing. This needs to be simulated.
- This pileup is done with the mixing module
  - Normally assume independent, so number of events is taken from a Poisson distribution with mean ≈25 or whatever. Also need to specify how many bunches before and after to add.
  - Min-bias events read from pre-generated file(s) in a ring buffer
  - More information (I'm still figuring this out myself - its not trivial) at https://twiki.cern.ch/twiki/bin/view/CMS/MixingModule
- Must pay attention for physics of rare events, an event in the pileup min-bias file with the "right" fluctuation can appear multiple times and give you "peak" ("delta function").

An Interesting Quote

- A quote from J.D.Bjorken usually given by T.Sjöstrand as a final word of warning:
  - "The Monte Carlo simulation has become the major means of visualization of not only detector performance but also of physics phenomena. So far so good. But it often happens that the physics simulations provided by the Monte Carlo generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data."
  - "I am prepared to believe that the computer-literate generation (of which I am a little too old to be a member) is in principle no less competent and in fact benefits relative to us in the older generation by having these marvelous tools. They do allow one to look at, indeed visualize, the problems in new ways. But I also fear a kind of "terminal illness", perhaps traceable to the influence of television at an early age. There the way one learns is simply to passively stare into a screen and wait for the truth to be delivered. A number of physicists nowadays seem to do just this."
- From a talk given by J.D.Bjorken at the 75th anniversary celebration of the Max-Plank Institute of Physics, Munich, germany, Dec. 10, 1992 (See SLAC Beam Line, 1992 vol.22, No.4 page 8.

So Contribute to the CMS Simulation!

- The simulation is ready for some serious testing, run it!
- The simulation is complicated and should not just be taken without question. It will need a lot of effort before we can use it for serious physics. You can contribute!
  - Running the simulation and giving feedback
  - Finding problems and investigating or testing solutions
  - Finding possible extensions or improvements
    - E.g. Adding new event generators or decay tables
    - E.g. GFlash electromagnetic and hadronic shower parameterization
    - E.g. FAMOS (fast simulation for CMS - check with P.Janot)
  - Help with physics validation in PRS groups (testbeam work)
- Come talk to Daniel Elvira or myself if you want ideas of what to work on
- Thanks to Julia Yarba for answering my many questions!