# Technical Aspects of Varying Hadronic Model Parameters and Studying Sensitivity of Simulated Results

Julia Yarba

22nd Geant4 Collaboration Meeting

09/29/2016

Fermilab

# (Relatively) Recent Initiative in the HAD Group

- Challenge:
  - how sensitive are Geant4 predictions to the variations of model parameters, and what uncertainties are associated with a Geant4 physics model, or a group of models, involved in simulation and optimization of a detector design
- Current Efforts:
  - Expand model(s) configuration interface(s)
  - Develop set of tools to study the effects of varying model parameters on the simulated results
  - Refine validity of ranges of the model(s) parameters
  - Determine correlations among parameters

# People Involved

- Laura Fields
- Krzysztof Genser
- Robert Hatcher
- Michael Kelsey
- Gabriel Perdue
- Hans Wenzel
- Dennis Wright
- Julia Yarba

# Configurable Geant4 HAD Models (I)

- PreCompound (since 10.3)
  - 9 configurable parameters and 14 switches
- Bertini (since 10.1)
  - 11 configurable parameters and 5 switches
  - Documentations by Dennis W. (password protected): https://cdcvs.fnal.gov/redmine/attachments/download/32147/Bertini_params.pdf
  - Note: documentation covers only 5 parameters and 2 switches
- FTF(P) – work in progress, 1st configuration interface in 10.3.ref08
  - ”diffraction” part of FTF for now; baryon projectile as the 1st use case
    - “hadronization” part (string fragmentation) is left for later
  - 18 configurable parameters and 4 switches
  - Documentation by Vladimir U.: http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/ForToolkitDeveloper/html/ch03s05.html#sect.ExtdFuncHadPhys.ChangeFTFParam

# Configurable Geant4 HAD Models (II)

- Configuration interfaces:
  - PreCompound – classical C++ interface (get/set)
  - Bertini – G4UI-based interface, in transition to newly implemented G4HadronicDeveloperParameters
  - FTF – C++ interface based on G4HadronicDeveloperParameters
- G4HadronicDeveloperParameters - Tatsumi's presentation in 6A
- All changes of parameters must be made very early in the application code, and the state should be "PreInit"
- Unique configuration vs multiple configurations:
  - Currently, multiple configurations of PreCompound and Bertini are possible
  - However, G4HadronicDeveloperParameters limits it to "one round of change per job" which directly applies to FTF

# Software to Study Effect of Parameters Variations

- Simulation with multiple configurations of Geant4 model(s) in the multi-parameters space
  - Process-level or based on a physics list
- Collective analysis of multiple variants of the resulting physics observables and simultaneous benchmarking vs relevant experimental datasets (one or many, as applicable)
  - Programmatic (C++) access to the DoSSiER, to extract experimental data
- Based on Art Framework: http://art.fnal.gov
  - Flexible run-time configurable workflow
    - Easy to extend configuration interface
    - Multiple instances of the same component can run in the same job
  - Comprehensive bookkeeping (ROOT-based)
    - E.g. we can keep track of model and parameters, metadata (beam, energy/momentum, target,…) + physics observables of interest
  - Modular, easy to extend software design

# Using Professor Tuning Toolkit (I)

- "Scans" by a single parameter are a good starting point and give a reasonable feel what what parameters are "sensitive" and where
- However, understanding correlations of parameters in only possible in the multi-parameters space
- Hence, enters Professor: http://professor.hepforge.org
  - "Fundamentally, the idea of Professor is to reduce the exponentially expensive process of brute-force tuning to a scaling closer to a power law in the number of parameters, while allowing for massive parallelisation and systematically improving the scan results by use of a deterministic parameterisation of the generator's response to changes in the steering parameters." – from Professor's web site
  - A set of parameters $P_i=\{x_i , y_i , z_i ,\ldots\}$ is a "point" in the multi-parameter space
  - Randomly sample multi-parameter space
    - For each $P_i$ simulate data combinatorics: beam × energy × target …
    - Derived quantities are histograms
  - Each simulated (histogram) bin content is $f(P_i)$ - polynomial approximation
  - Fit experimental data with $f(P_i)$ to explore sensitivity and coupling of parameters

# Using Professor Tuning Toolkit (II)

## Number of parameters vs polynomial order vs number of "points"

3 dimensional parameter space:

| Polynomial order | Minimum samples |
|---|---|
| 0 | 1 |
| 1 | 4 |
| 2 | 10 |
| 3 | 20 |
| 4 | 35 |
| 5 | 56 |
| 6 | 84 |
| 7 | 120 |
| 8 | 165 |
| 9 | 220 |
| 10 | 286 |

20 dimensional parameter space:

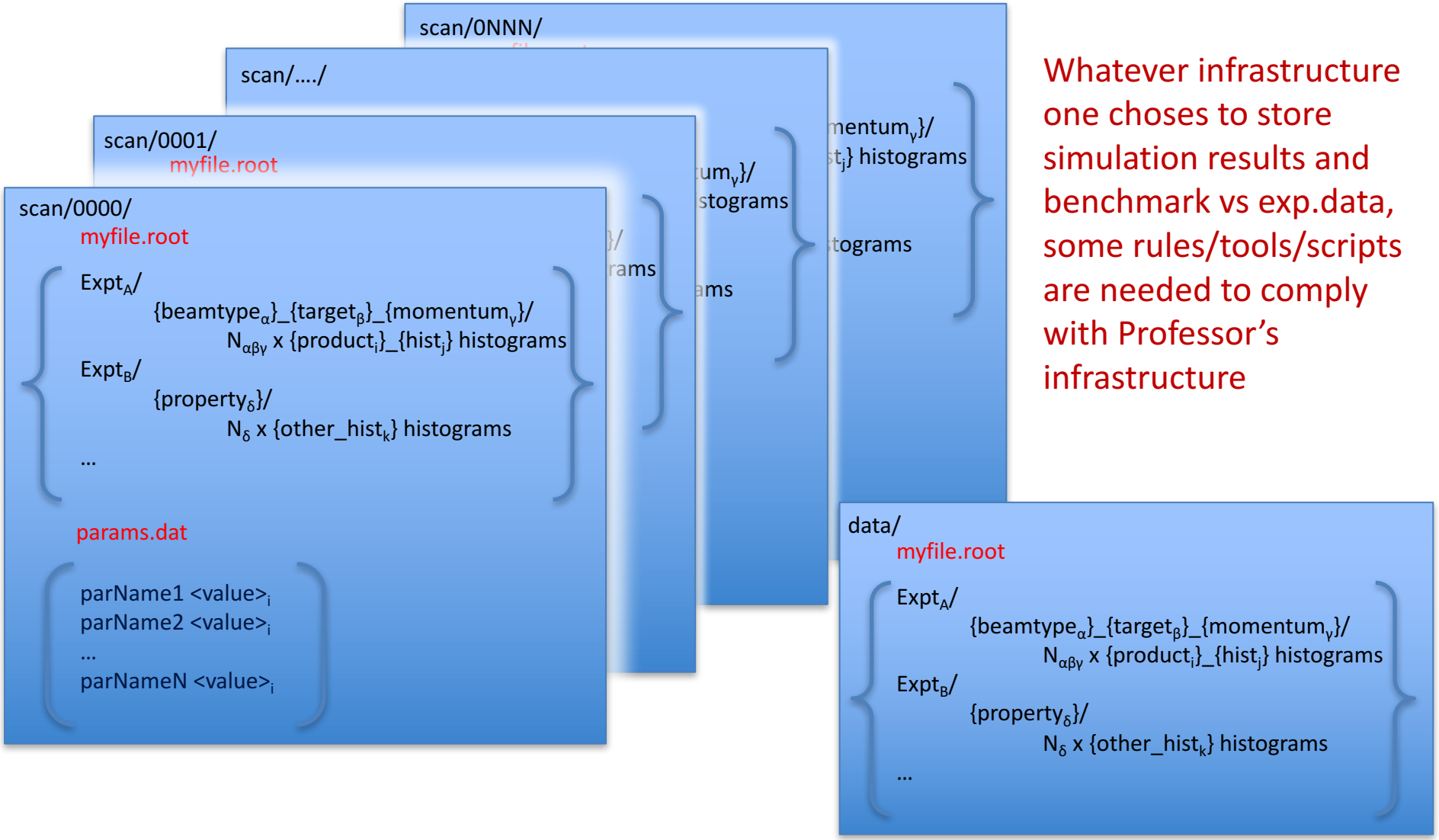
| Polynomial order | Minimum samples |
|---|---|
| 0 | 1 |
| 1 | 21 |
| 2 | 231 |
| 3 | 1771 |
| 4 | 10626 |
| 5 | 53130 |
| 6 | 230230 |
| 7 | 888030 |
| 8 | 3108105 |
| 9 | 10015005 |
| 10 | 30045015 |

50 dimensional parameter space:

| Polynomial order | Minimum samples |
|---|---|
| 0 | 1 |
| 1 | 51 |
| 2 | 1326 |
| 3 | 23426 |
| 4 | 316251 |
| 5 | 3478761 |
| 6 | 32468436 |
| 7 | 264385836 |

Hopefully in the range of suggested variations of the parameters the results would be relatively smooth so that we can use low order polynomials

# Using Professor Tuning Toolkit (III)




Whatever infrastructure one choses to store simulation results and benchmark vs exp.data, some rules/tools/scripts are needed to comply with Professor's infrastructure

# Latest News from Robert Hatcher

- Robert has ran through Professor analysis of pion production in 5GeV/c $\pi^-$+Cu simulated with Bertini and compared vs HARP data (only one dataset for now)
- … and he was able to obtain a PRELIMINARY correlation matris
- … which tells that e.g. XSecScale is not correlated to anything else
- It is too early to show any specific numbers into a presentation
- However, it is a encouraging proof of principle

# Documentation

- ”Internal” documentation exists although it is password-protected because we are not yet ready to share it with general users
- Obvously, anyone in Geant4 can be granted access
- Documentation on the art-based software:
  - https://cdcvs.fnal.gov/redmine/projects/g4mps/wiki/Phase2_App_01052016
- Notes on experience with Professor:
  - https://cdcvs.fnal.gov/redmine/projects/professor-based-parameter-variation/wiki

# Summary

- As the configuration interfaces to Geant4 HAD models get expanded, we are in the processes of exploring sensitivity of simulated physics observables to the variations of model parameters
- This requires certain infrastructure run multiple variants of a model (or a physics list) and to collect outputs for further comprehensive analysis
- We have implemented a set of tools for this purpose
- Understanding correlations of parameters is an important task that requires analysis in the multi-parameter space and may involve multiple datasets
- We have implemented a set of interface tools to the Professor toolkit and are currently exploring how far we can go with it