Geant 4

Special Purpose Statistical Analysis Tools
- open to discussion -

Julia Yarba, FNAL

21st Geant4 Collaboration Workshop

Sept. 12-16, 2016

Outline

- Motivation: Possible Tasks and Required Features
- RooMUHistos Package
- Suggestions (instead of Summary)

Motivation

- As we move into exploring effects on the Geant4 predictions by the model parameter variations, we need to analyze:
  - Sensitivity of Geant4 predictions
    - Group predictions
    - Combine selected groups of predictions
    - Perform statistical calculations including all groups of variants or selected selected groups
  - Collectively benchmark predictions (including variants) vs experimental data
- Proper bookkeeping becomes a task on itself
- Statistical calculations are more complex than “TH1-to-TH1”
- We need certain infrastructure to organize and streamline the analysis part

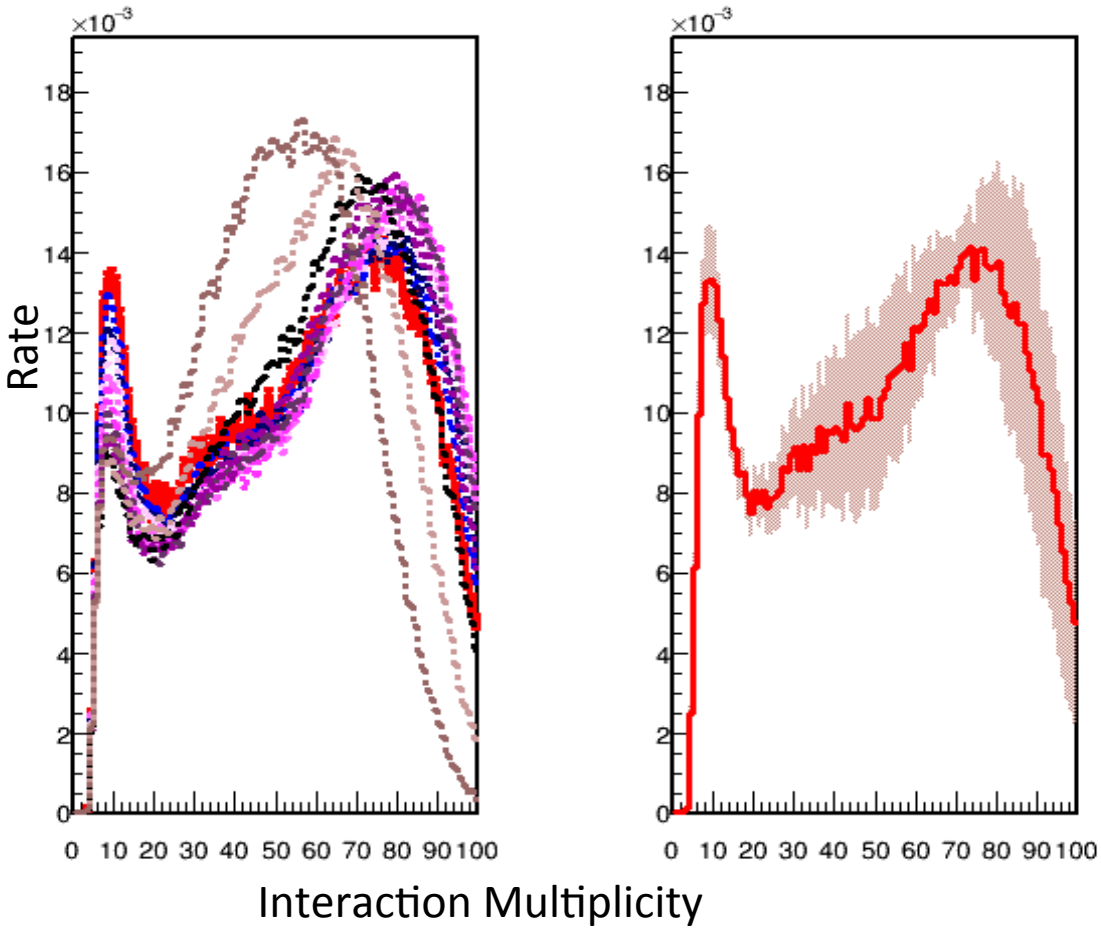
RooMUHistos – General Information

- Based only on ROOT – can be considered as an “extension”
  - well, also a bit on boost but this can go away if migrated to C++11
- Open-source - available via Git
  - https://github.com/ManyUniverseAna/RooMUHistos
- Easy to build
- Easily fits with other packages, including modern frameworks
- Usage examples provided
- Has attractive features that I could not find in native ROOT
  - Correct me if I am wrong…
- Does not contradict use of “bare” Geant4 and/or other analysis tools
- Obviously, has some room for improvement – what packages doesn’t ?
- Note: MU==ManyUniverses, where a Universe represents results obtained with one or another set of parameters

RooMUHistos – Selected Features

- Build-in functions for working with the case where a default value and a single and/or multiple variations around that value are considered
- Combine in one object all variants, together with the default predictions; organizes “ingredients” to compute uncertainties
  - Concept of “band”: a collection of “likeminded” histograms
  - There can be multiple bands (for different sources of uncertainty)
- Allows to present experimental data or MC results with statistical and systematic errors
  - Errors can be presented in a form of matrix
- Functions to proper calculate error matrices over MC variants
- Functions to take MC and/or data error matrices into account when calculate chi2
- Calculation (and plotting, if desired) of fractional uncertainties of all systematics contained in all error bands
- Systematic error sources can be grouped

Example use of RooMUHistos to demonstrate the effect of Radius Scale variations on the multiplicity of a single hadronic interaction



Default Bertini is shown **in red**

Left plot:

- Default Bertini + 10 variants of Radius scale from 95% to 50% of “standard”, step=5%

Right plot:

- Default Bertini predictions outfitted with **error band** calculated (as interquartile spread) from 10 variants of Radius Scale

Interaction: 5GeV/c proton on Pb

Statistics: 200K events/variant

RooMUHistos – Ownership and Support

- Originally developed in MINERvA collaboration
- Forked with permission
- “Lost” Python bindings for PyROOT (time/manpower constraints during the fork) but it can restored if required
- The package as it stands is "owned" by FNAL-SCD-PDS group
- We are open to accommodating modifications to suit if needed

RooMUHistos – Suggestions (instead of Summary)

- We are currently “trying out” RooMUHistos in the Geant4 model parameter study project
- It appears to be a convenient tool, in particular to implement various “end-of-job” and/or “summary” applications, especially in the case where multiple variants of an observable are involved
- If there is another similar tool, we should consider how to consolidate efforts; we definitely do not want to duplicate
- Otherwise we will suggest considering to adopt RooMUHistos as part of Geant4 statistical analysis domain
- RooMUHistos is also planned for use in GENIE validation domain

  Note: GENIE is a package to model neutrino interactions