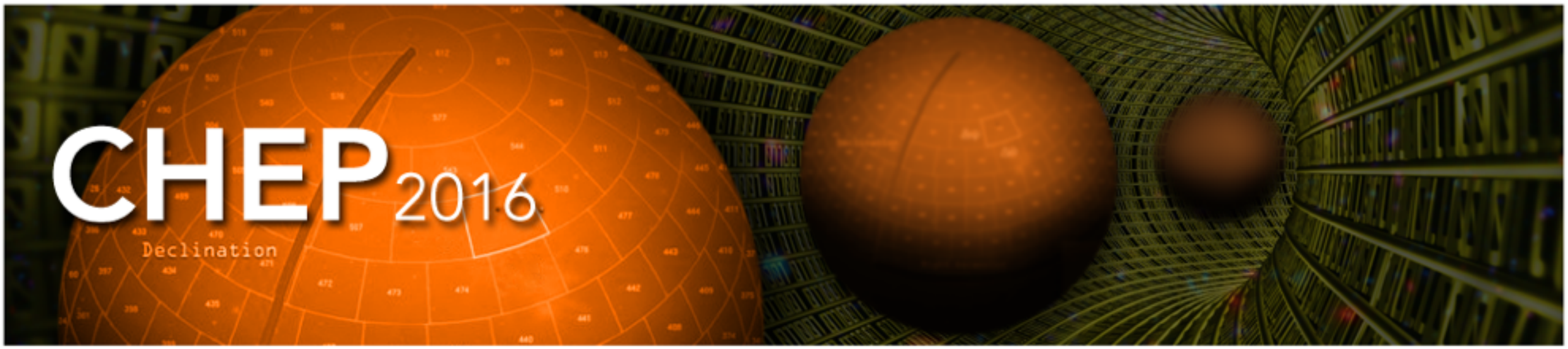
Validation of Electromagnetic Physics Models for Parallel Computing Architectures in the GeantV project

Marilena Bandieramonte
(marilena.bandieramonte@cern.ch)
for the GeantV team

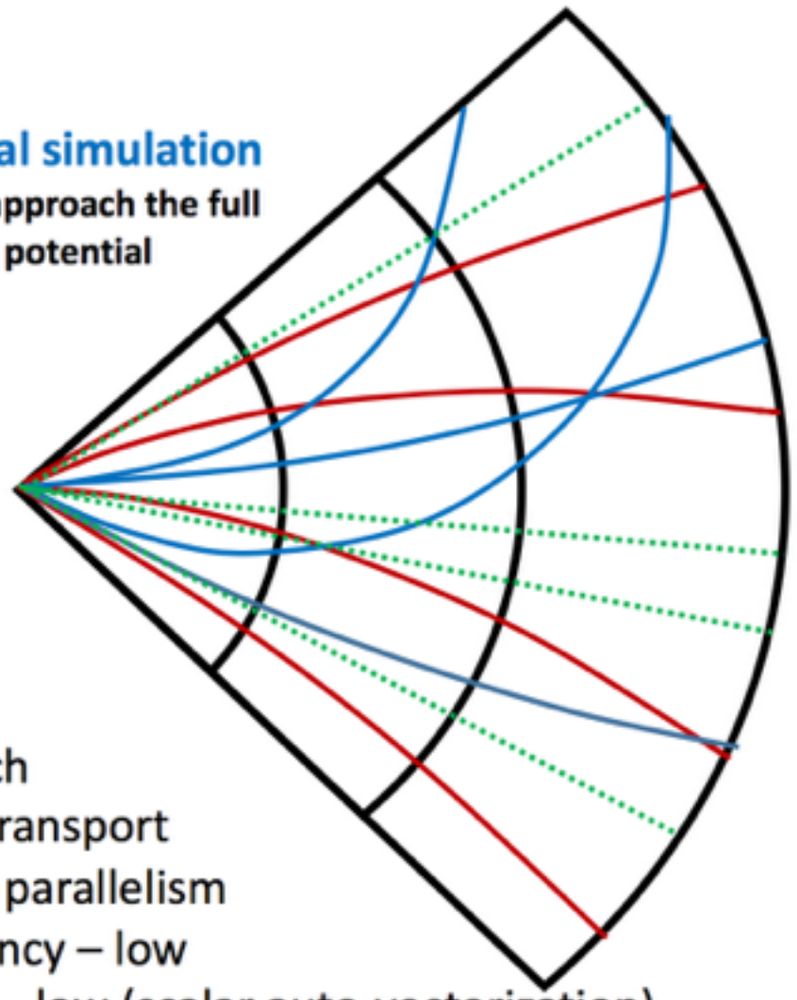
13th Oct 2016, San Francisco



22nd International Conference on Computing in High Energy and Nuclear Physics, Hosted by SLAC and LBNL, Fall 2016

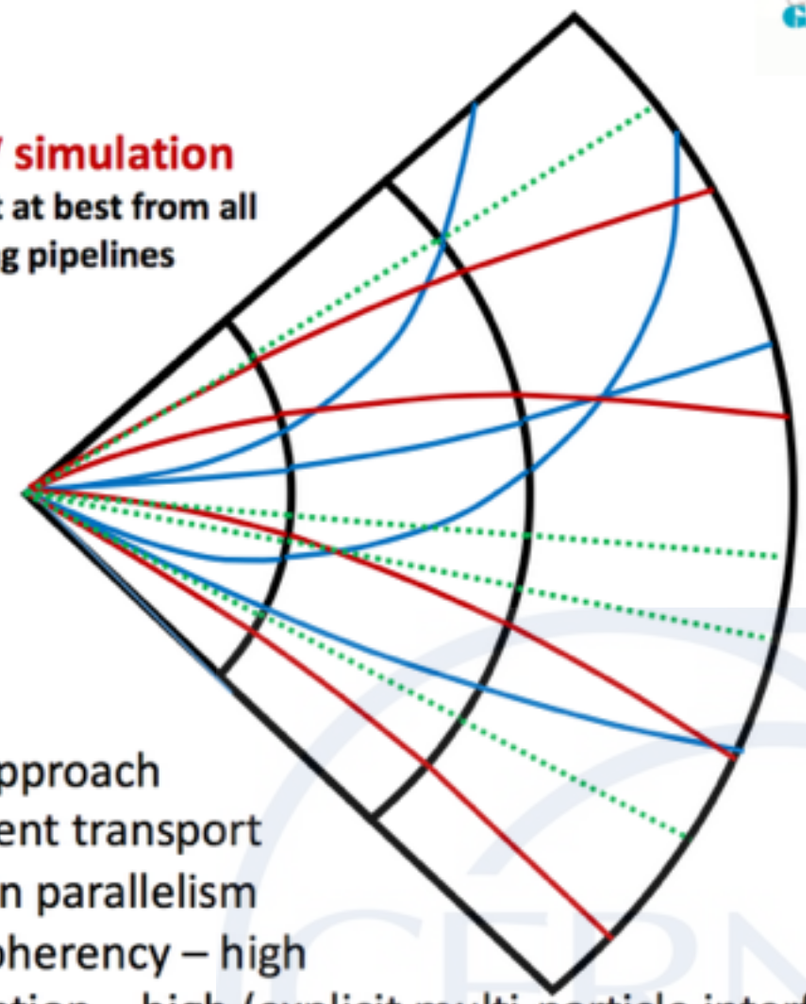
GeantV – Adapting simulation to modern hardware

Classical simulation
hard to approach the full machine potential



- Stack approach
- Single event transport
- Embarrassing parallelism
- Cache coherency – low
- Vectorization – low (scalar auto-vectorization)

GeantV simulation
can profit at best from all processing pipelines



- Basket approach
- Multi event transport
- Fine grain parallelism
- Cache coherency – high
- Vectorization – high (explicit multi-particle interfaces)

Vectorizing physics models: Goals

- ***80% of CPU*** time is spent on electrons and photons in Geant4 for typical collider experiments. About 30-40% is on Electromagnetic (EM) processes

  - write EM physics models dealing with multiple tracks
    - accurate
    - fast
    - portable

  - exploit both ***SIMD*** (vector pipeline) and ***SIMT*** (accelerators) execution models

    - ***common source code*** between scalar, vector and accelerator (GPU, Xeon Phi)

- Two approaches for the physics: bottom-up, top-down

- **Top-down Approach:** Focus on sampling techniques

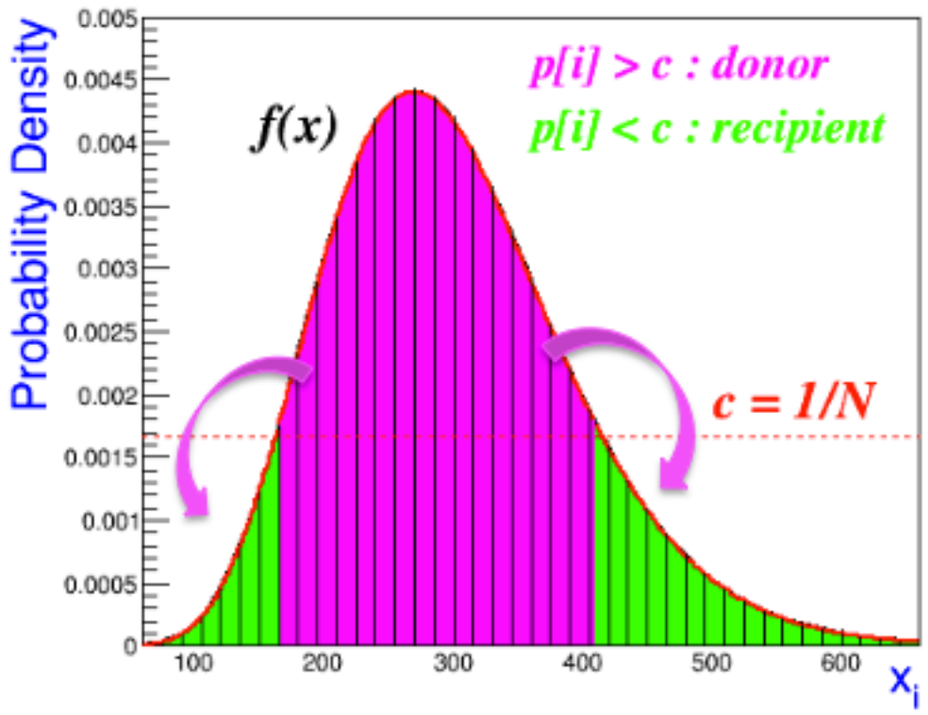
  - Geant4 typically uses composition and rejection sampling techniques

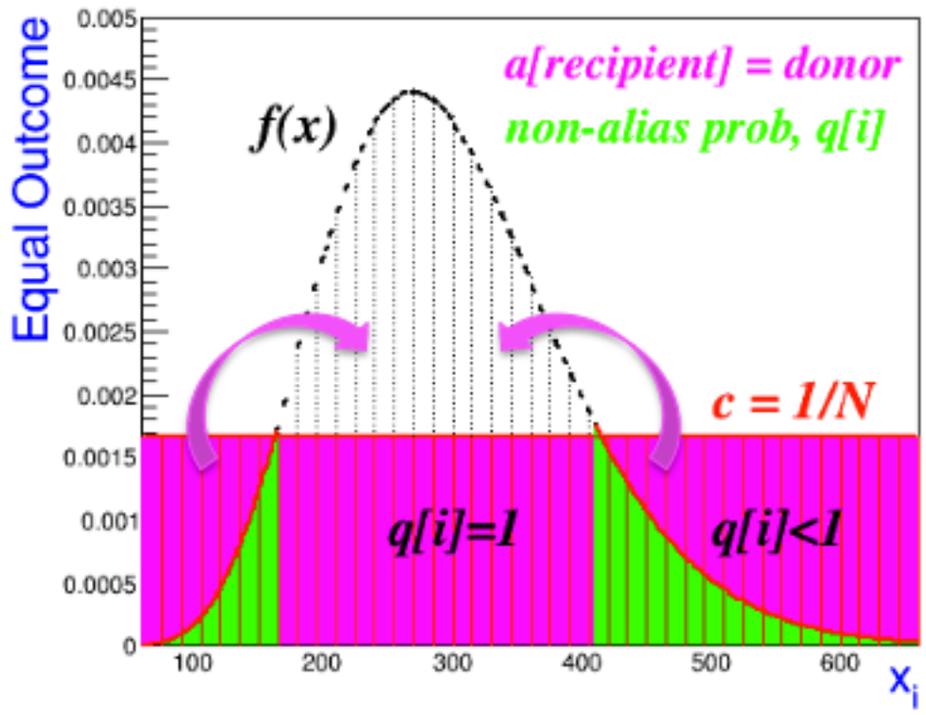
  - Explore alternative **sampling techniques**

  - **Statistical Validation** of vectorized physics models

Alias Sampling Method for N Discrete Outcomes

- Recast a N-discrete p.d.f to N equal probable events, each with likelihood 1/N = c *(A. J. Walker, 1974)* – effectively vectorizable
- Reproduce the original distribution by one trial sampling:





Alias index: a[recipient] = donor
Non-alias probability: q[i]
For any random (x_i), accept if rand(0,1) < q[i] or take the alias

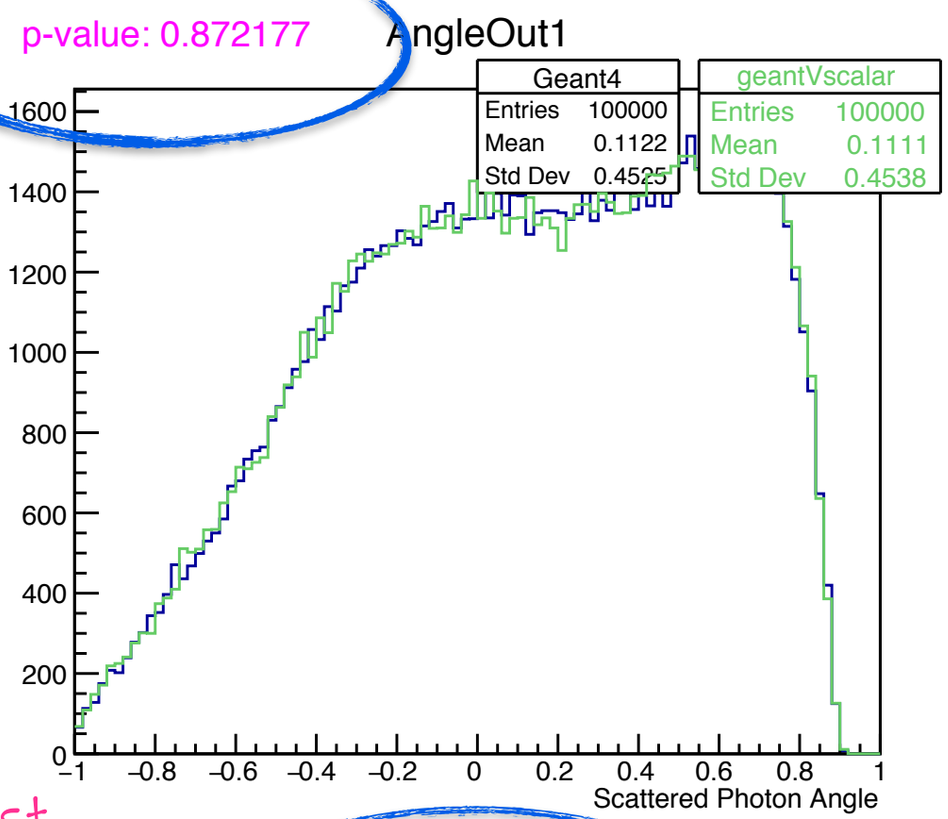
Pearson χ^2-test for comparing weighted and unweighted histograms

- Comparison of two histograms expect hypotheses that two histograms represent identical distributions. To make a decision **p-value** should be calculated.

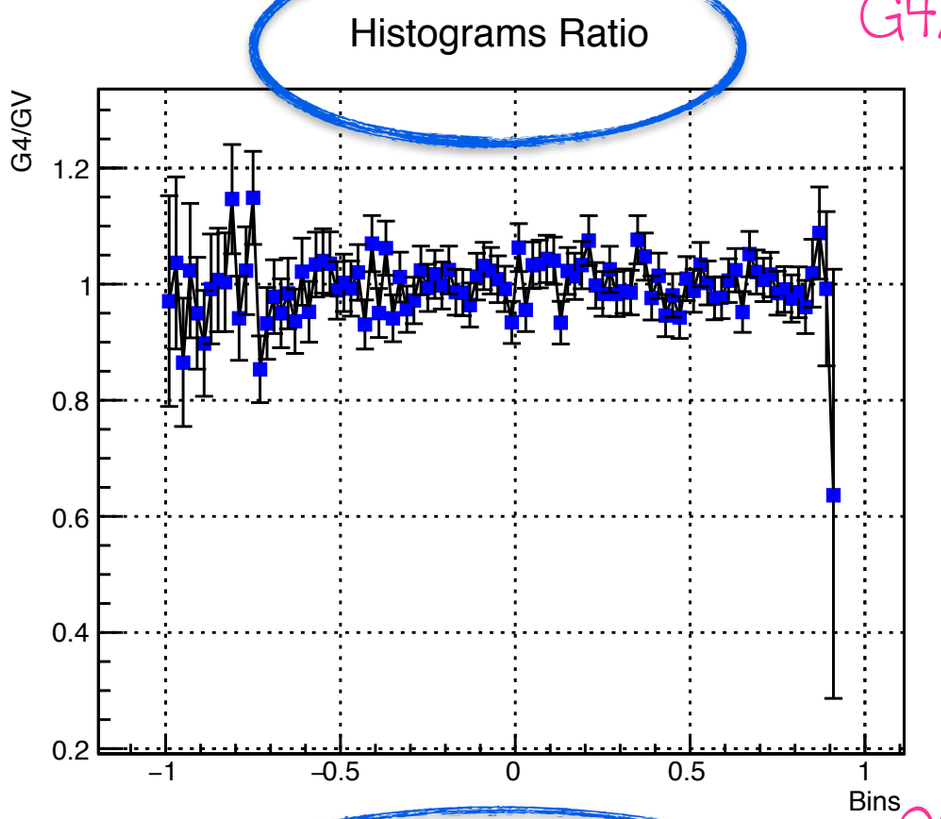
- The hypotheses of identity is rejected if the p-value is lower than some significance level.
    - Traditionally significance levels 0.1, 0.05 and 0.01 are used.
    - **Chosen threshold 0.05 -> If p-value < 0.05 hypothesis of identity is rejected!**

- The comparison procedure should include an **analysis of the residuals** which is often helpful in identifying the bins of histograms responsible for a significant overall #chi^{2} value.

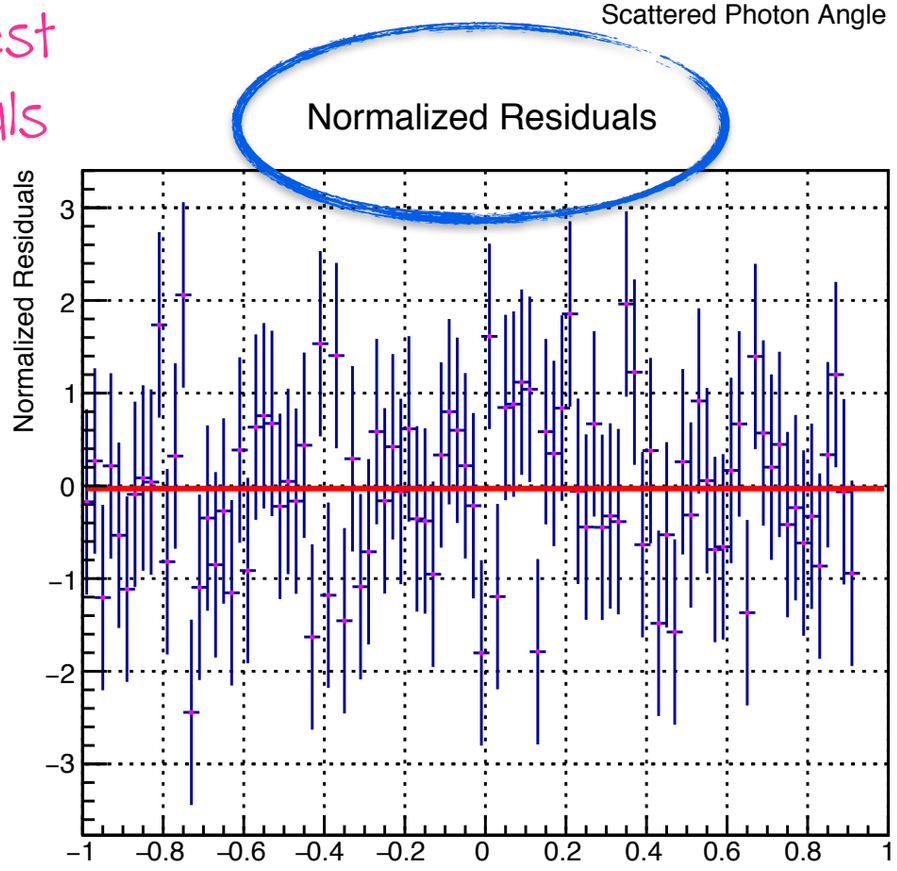
p-value: 0.872177



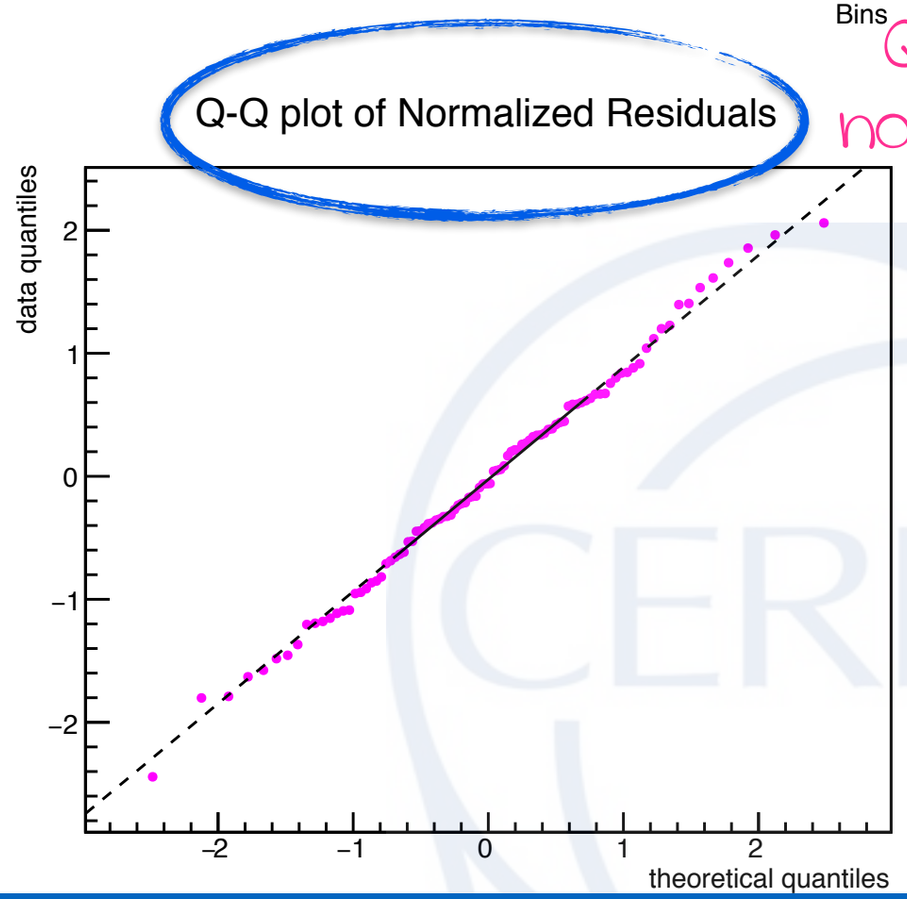
G4/GV ratio



chi^2 test residuals



QQ-Plot for normalized res



EM physics processes under development

Primary	Process	Model	Secondaries	Survivor
γ	Compton Scattering	Klein-Nishina	e^-	γ
	Pair-Production	Bethe-Heitler	e^- e^+	-
	Photo-Electric Effect	Sauter-Gavrila	e^-	-
e^-	Ionization	Moller-Bhabha	e^-	e^-
	Bremsstrahlung	Seltzer-Berger	γ	e^-
	Multiple Scattering	Goudsmit-Saunderson	-	e^-

Gamma processes

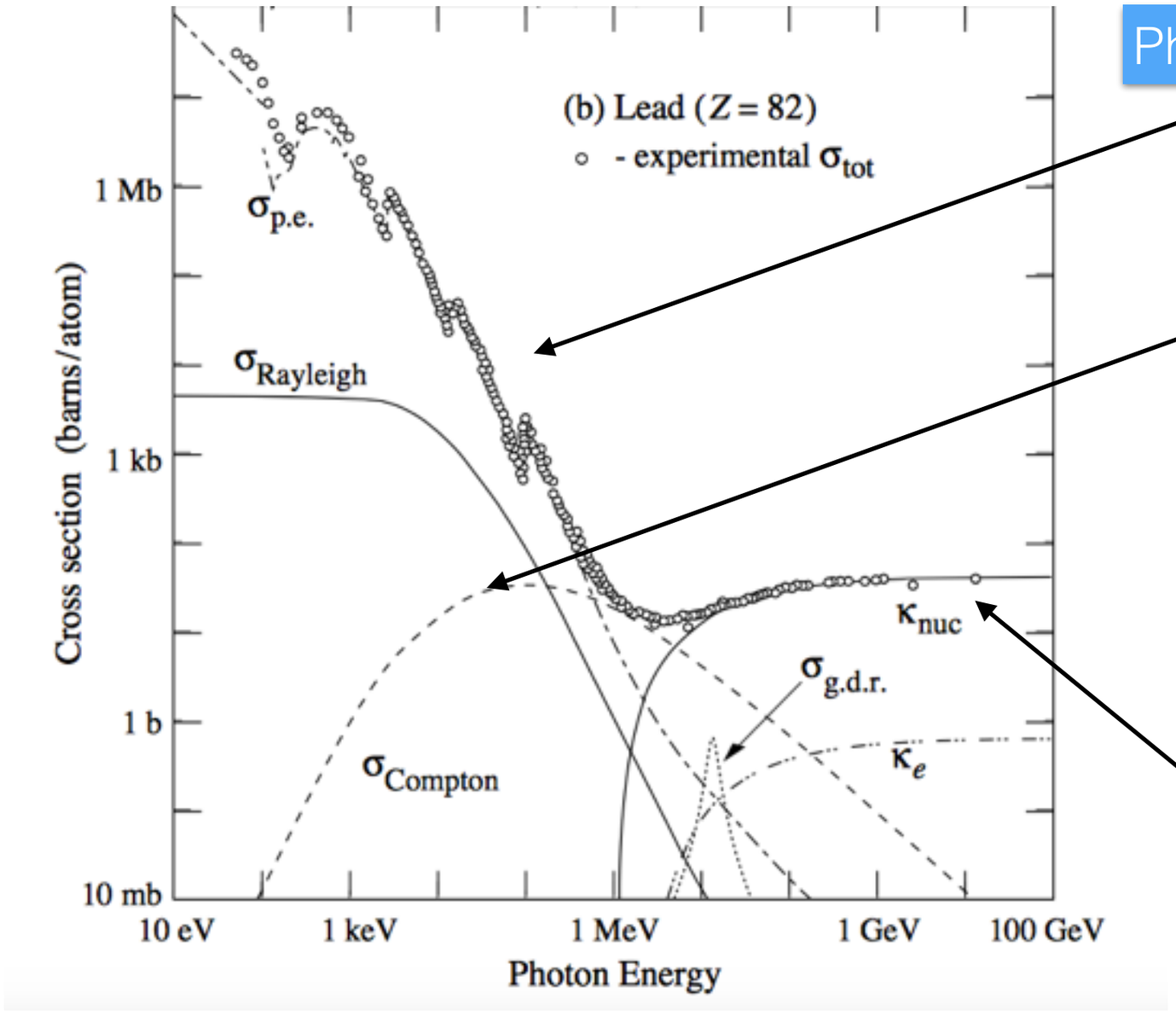


Photo-electric effect*

SauterGavrila

Compton**

KleinNishina

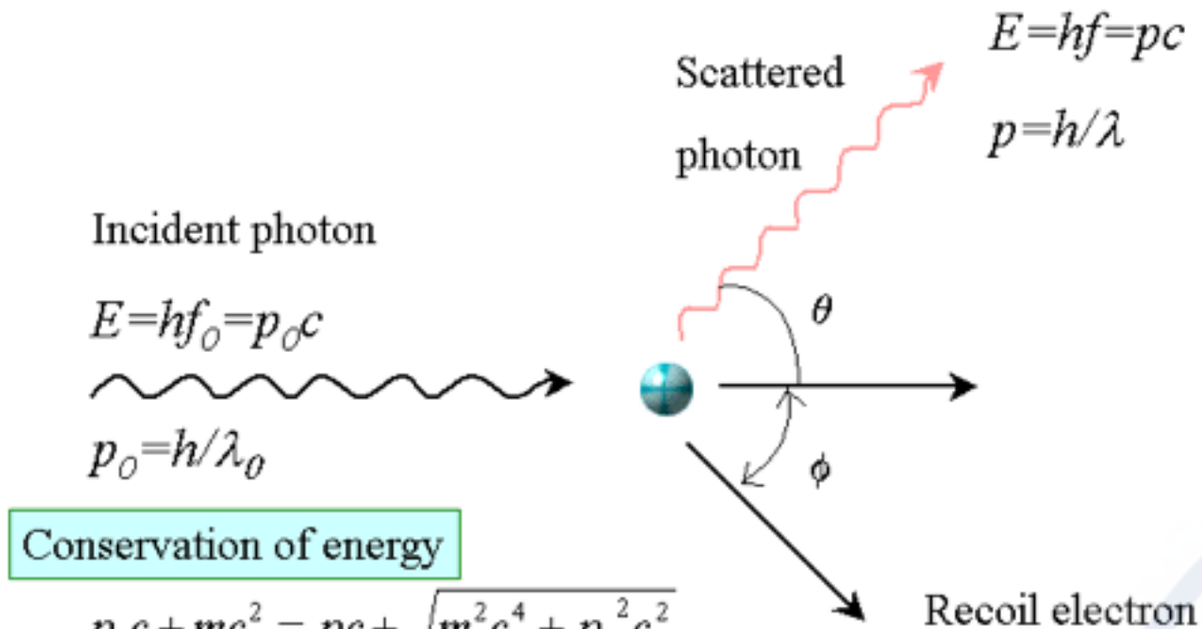
Pair production***

Bethe-Heitler

* Photo-electric should not be validated with Geant4 in the energy range, (E/m_electron) < 50 ~ 25 MeV (above this threshold, Geant4 returns cos(theta) =1 as a good approximation)
** Compton should be validated for E > 10KeV (to be consistent with Geant4)
*** Pair production: should be validated in the energy range, E > 2*m_electron ~ 1 MeV

Klein-Nishina Validation

Compton Scattering



Conservation of energy

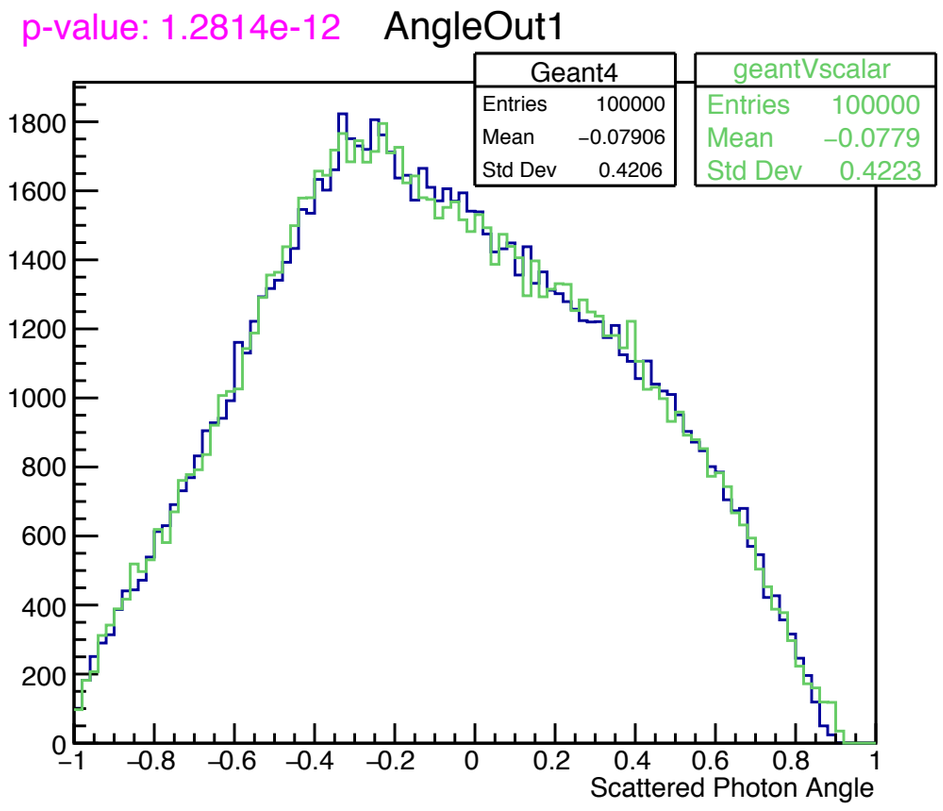
$$p_0c + mc^2 = pc + \sqrt{m^2c^4 + p_e{}^2c^2}$$

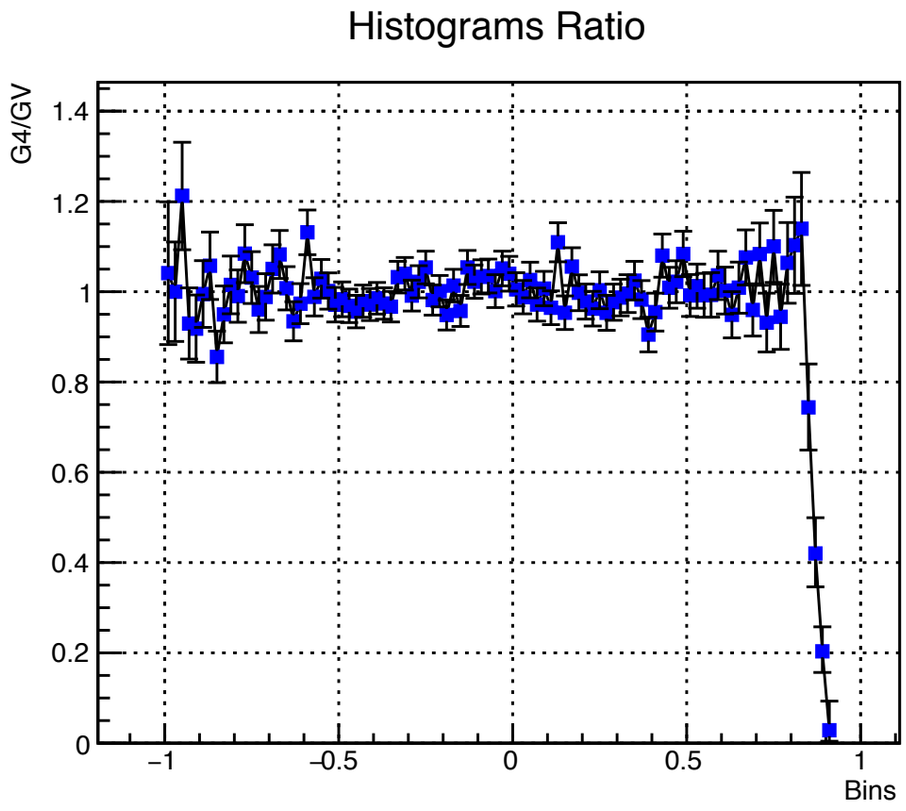
Conservation of momentum

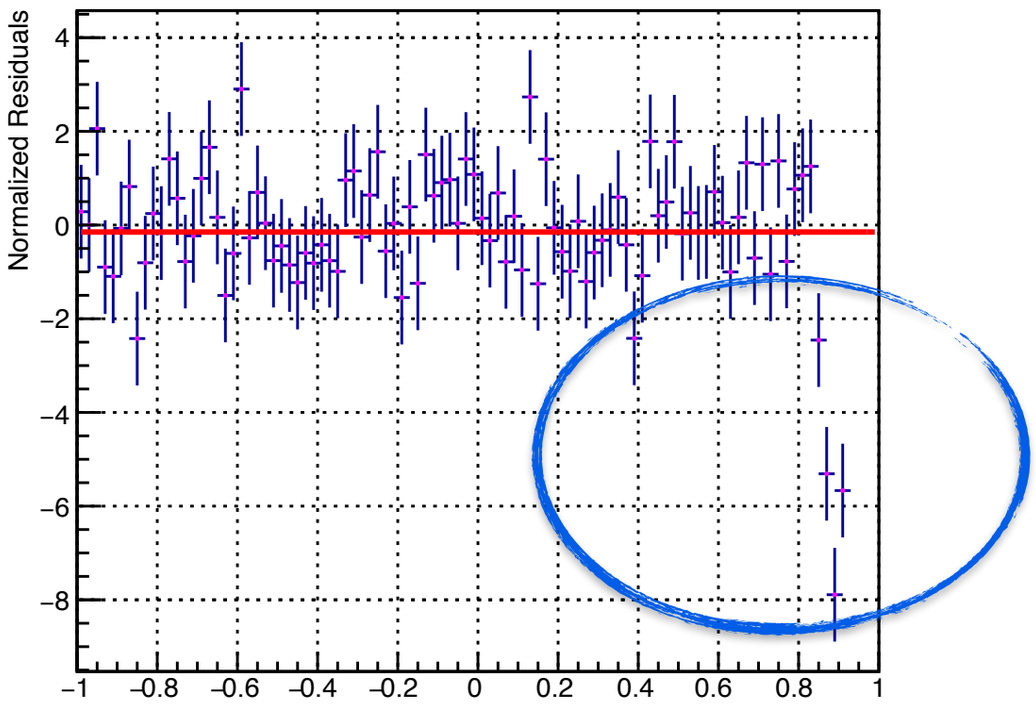
$$p_0^2 + p^2 - 2p_0p\cos(\theta) = p_e^2$$

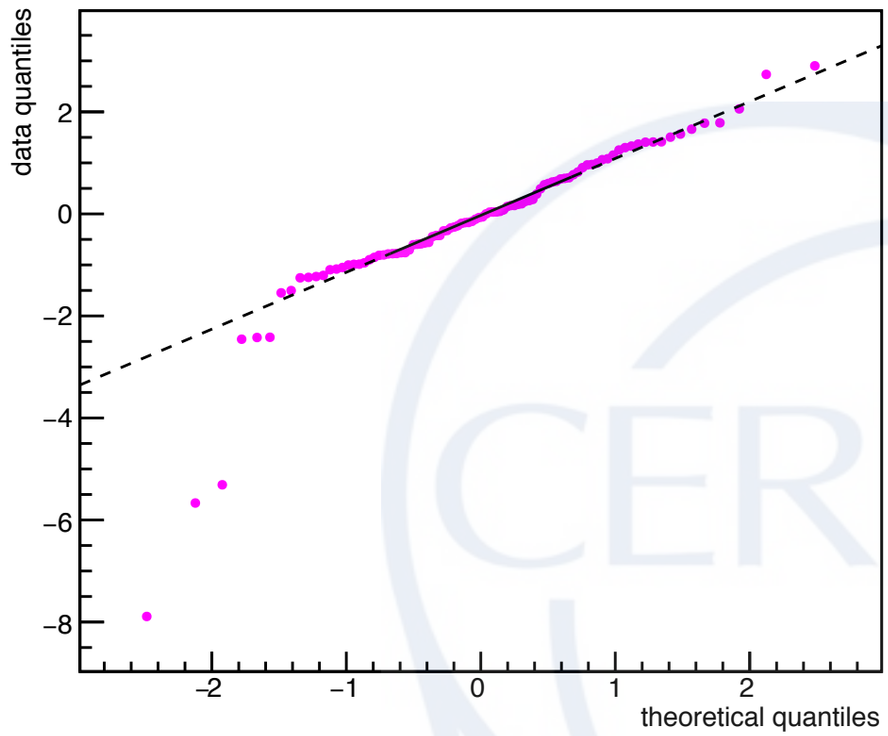
EnergyIn	*ValidationQ*	*Chi-2 test p-value*	*A-D test p-value*	*K-S test p-value*
0.01 MeV	EnergyOut1	0.650629	1	1
	EnergyOut2	0.664124	1	1
	AngleOut1	1.2814e-12	0.999999	0.999998
	AngleOut2	1.59571e-171	1	1
0.1 MeV	EnergyOut1	0.342675	1	1
	EnergyOut2	0.209657	1	1
	AngleOut1	0.17035	0.997342	0.999633
	AngleOut2	1.75073e-68	0.999999	0.999998
1 MeV	EnergyOut1	0.241563	0.999999	0.999633
	EnergyOut2	0.172393	0.999989	0.993765
	AngleOut1	0.379543	0.986458	0.906206
	AngleOut2	0.104473	1	0.999633
10 MeV	EnergyOut1	7.19275e-32	0.98934	0.906206
	EnergyOut2	7.90652e-21	0.998548	0.993765
	AngleOut1	0.720494	0.999999	0.999998
	AngleOut2	6.33771e-15	1	0.999998
100 MeV	EnergyOut1	0.522992	0.999778	0.993765
	EnergyOut2	0.0483391	1	0.999633
	AngleOut1	0.0463506	0.999999	0.999633
	AngleOut2	0.0137113	0.999992	0.906206
1000 MeV	EnergyOut1	5.03277e-06	0.940721	0.699374
	EnergyOut2	2.58997e-07	0.93819	0.906206
	AngleOut1	0	0.157212	0.28093
	AngleOut2	0	0.569165	0.0783231
10000 MeV	EnergyOut1	5.20263e-308	0.00121127	7.14257e-05
	EnergyOut2	0	0.00050695	1.90803e-05
	AngleOut1	0	0.000419885	1.90803e-05
	AngleOut2	0	0.0139016	1.90803e-05

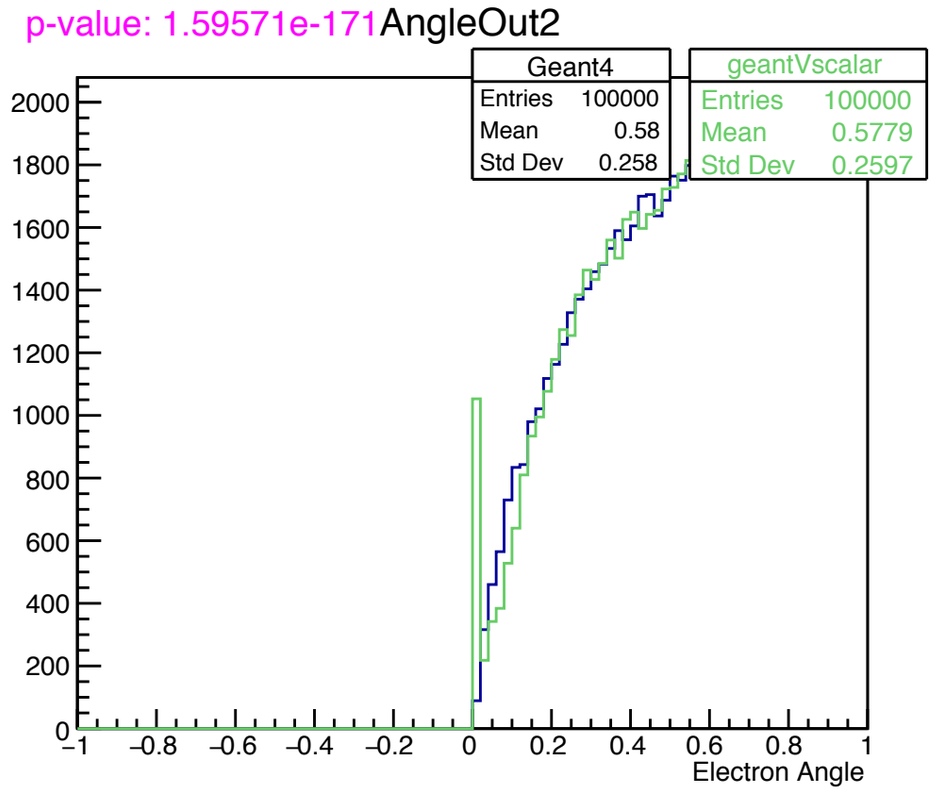
p-value: 1.2814e-12

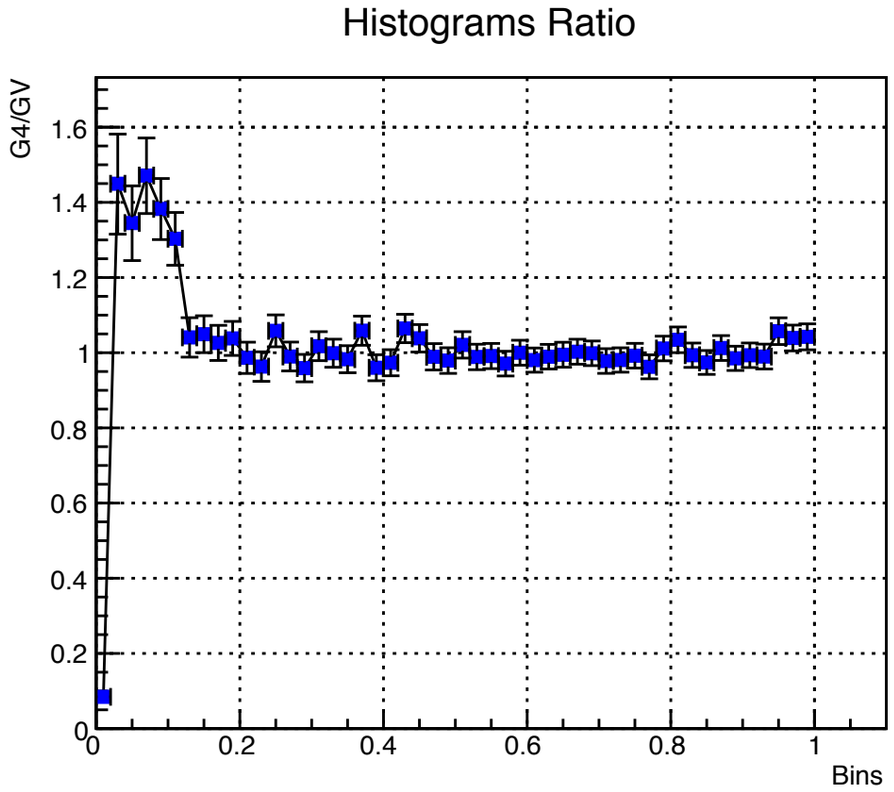


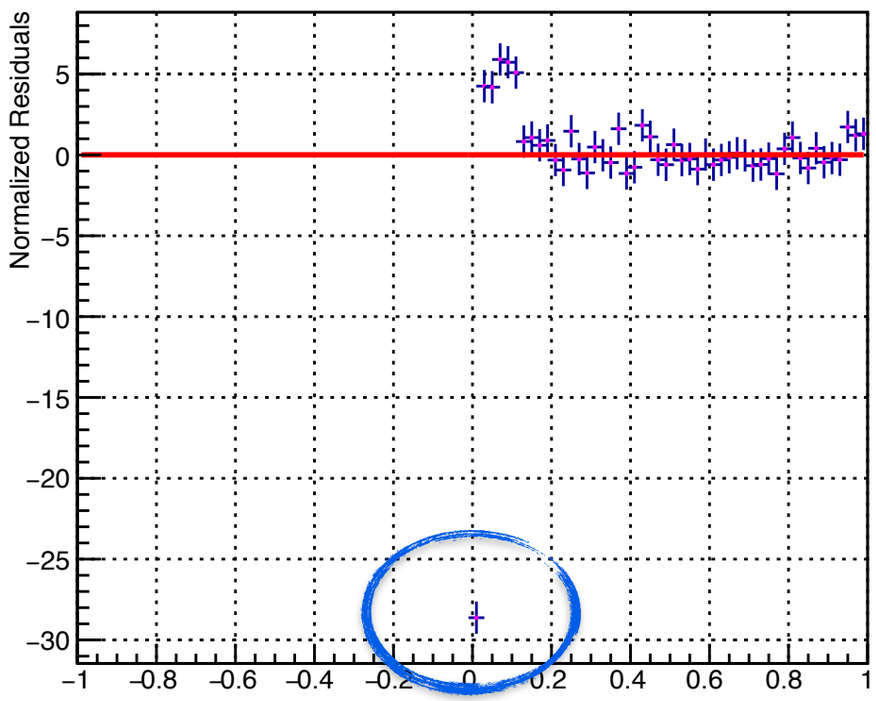


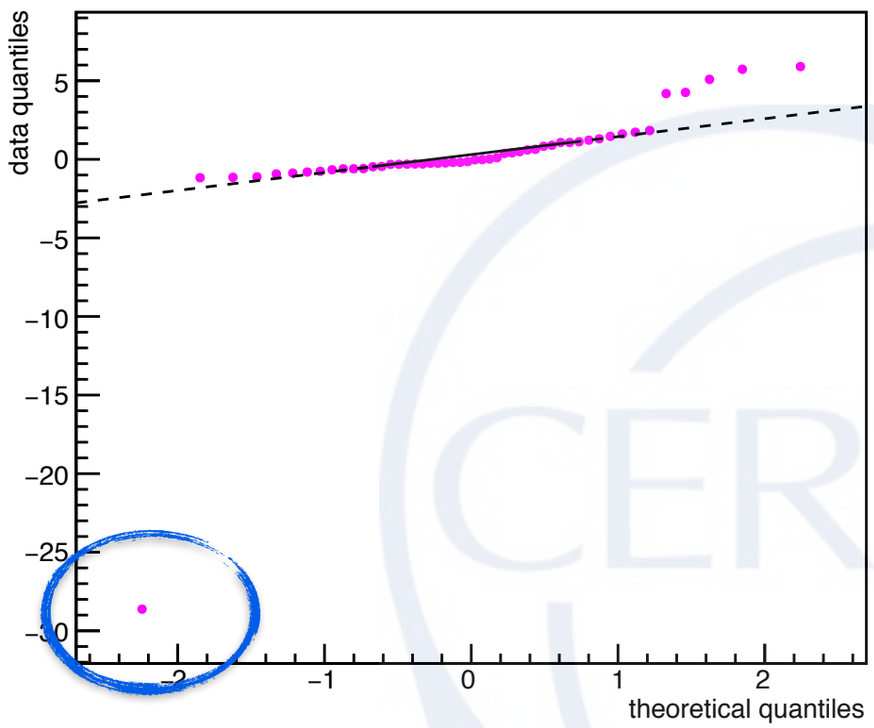





p-value: 1.59571e-171
AngleOut2
Geant4
Entries 100000
Mean 0.58
Std Dev 0.258
geantVscalar
Entries 100000
Mean 0.5779
Std Dev 0.2597
Electron Angle


Histograms Ratio
G4/GV
Bins


Normalized Residuals
Normalized Residuals

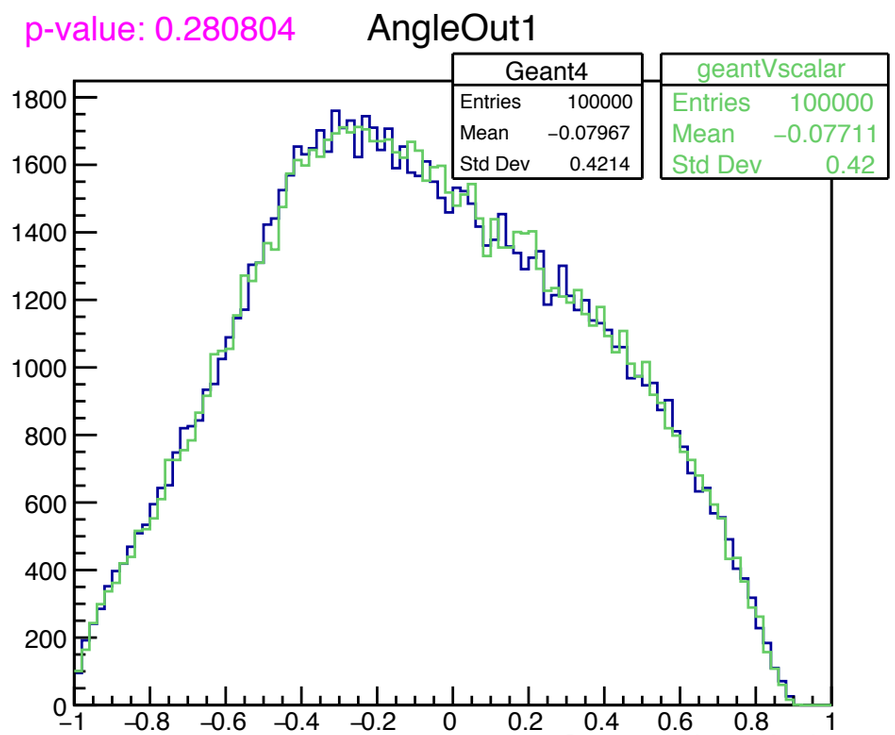

Q-Q plot of Normalized Residuals
data quantiles
theoretical quantiles

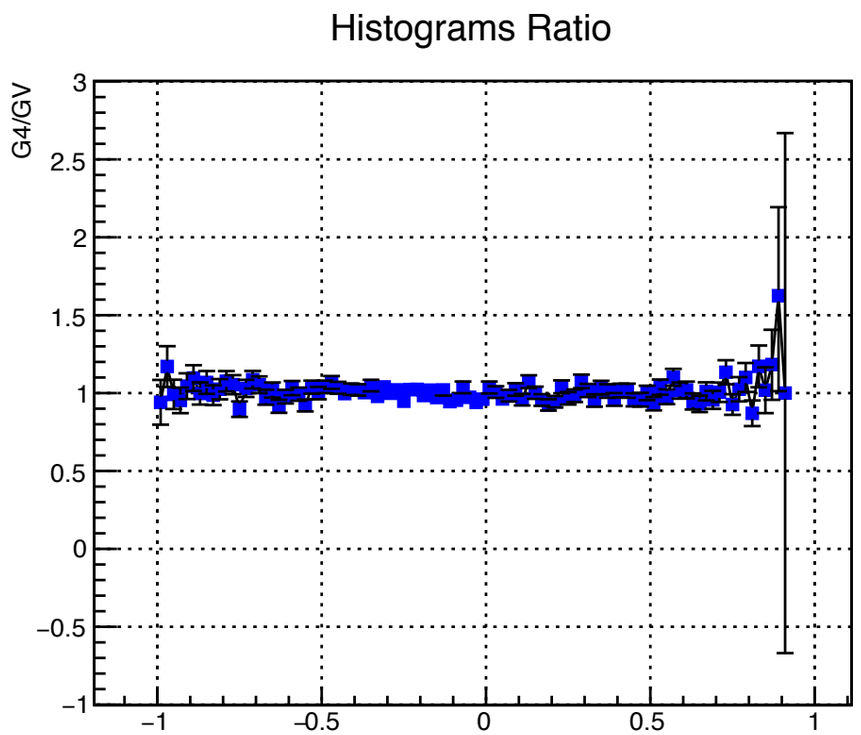
EnergyIn	*ValidationQ*	*Chi-2 test p-value*	*A-D test p-value*	*K-S test p-value*
0.01 MeV	EnergyOut1	0.31306	1	1
	EnergyOut2	0.795741	1	1
	AngleOut1	0.280804	0.999835	0.999998
	AngleOut2	0.141958	0.999999	0.999998
0.1 MeV	EnergyOut1	0.857909	1	1
	EnergyOut2	0.858931	1	1
	AngleOut1	0.560302	1	1
	AngleOut2	1.5663e-15	0.999554	0.967068
1 MeV	EnergyOut1	0.0103478	0.997621	0.906206
	EnergyOut2	0.0475544	0.999389	0.967068
	AngleOut1	0.0167979	0.997876	0.993765
	AngleOut2	5.14327e-56	0.999348	0.999633
10 MeV	EnergyOut1	0.817173	1	0.999633
	EnergyOut2	0.875291	0.999997	0.993765
	AngleOut1	0.872177	0.999999	0.999998
	AngleOut2	0.995689	1	0.999998
100 MeV	EnergyOut1	0.011512	0.876119	0.467558
	EnergyOut2	0.124762	0.981562	0.699374
	AngleOut1	0.0264691	0.999996	0.999633
	AngleOut2	0.387263	1	1
1000 MeV	EnergyOut1	5.03277e-06	0.940721	0.699374
	EnergyOut2	2.58997e-07	0.93819	0.906206
	AngleOut1	0	0.157212	0.28093
	AngleOut2	0	0.569165	0.0783231
10000 MeV	EnergyOut1	5.20263e-308	0.00121127	7.14257e-05
	EnergyOut2	0	0.00050695	1.90803e-05
	AngleOut1	0	0.000419885	1.90803e-05
	AngleOut2	0	0.0139016	1.90803e-05

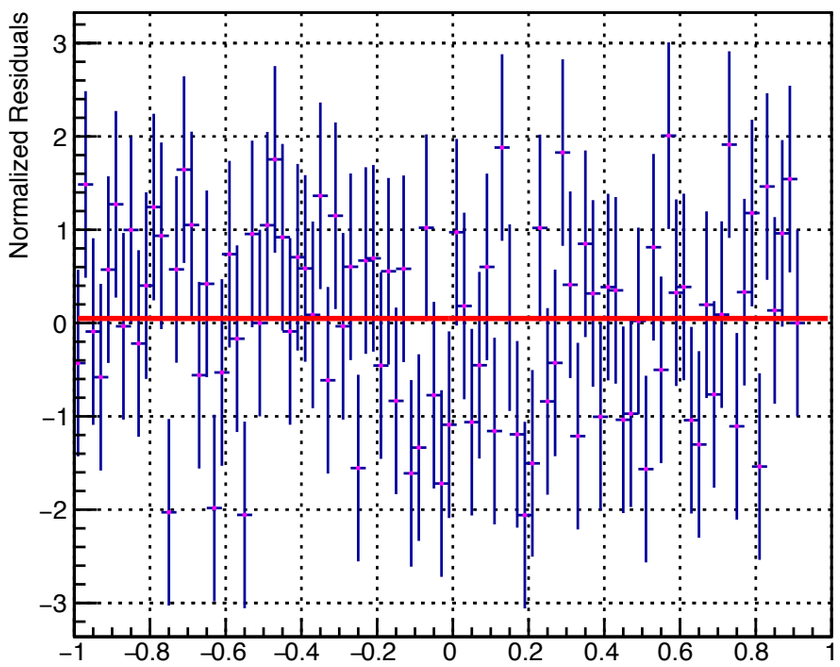
E_in=0.01MeV

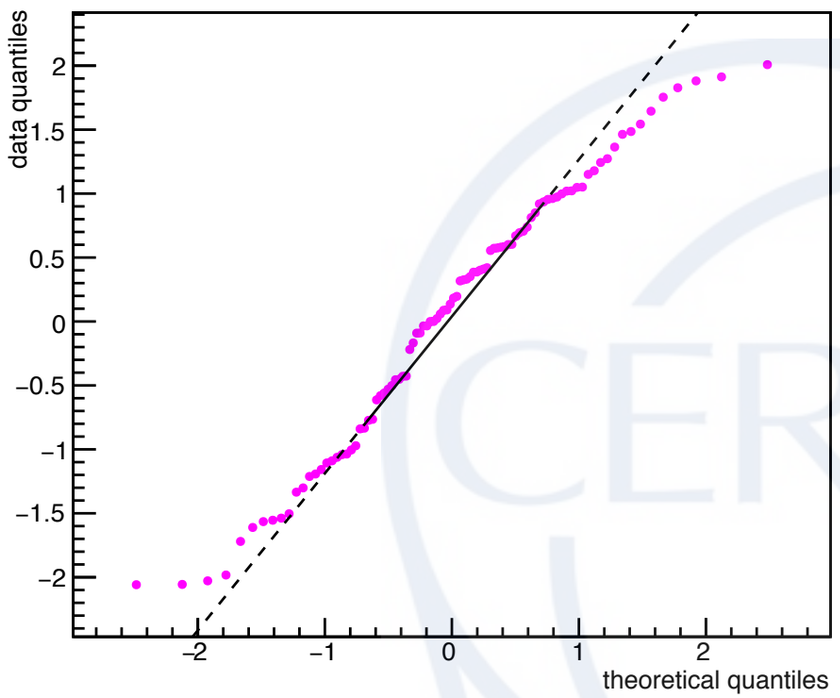
Fixed

p-value: 0.280804



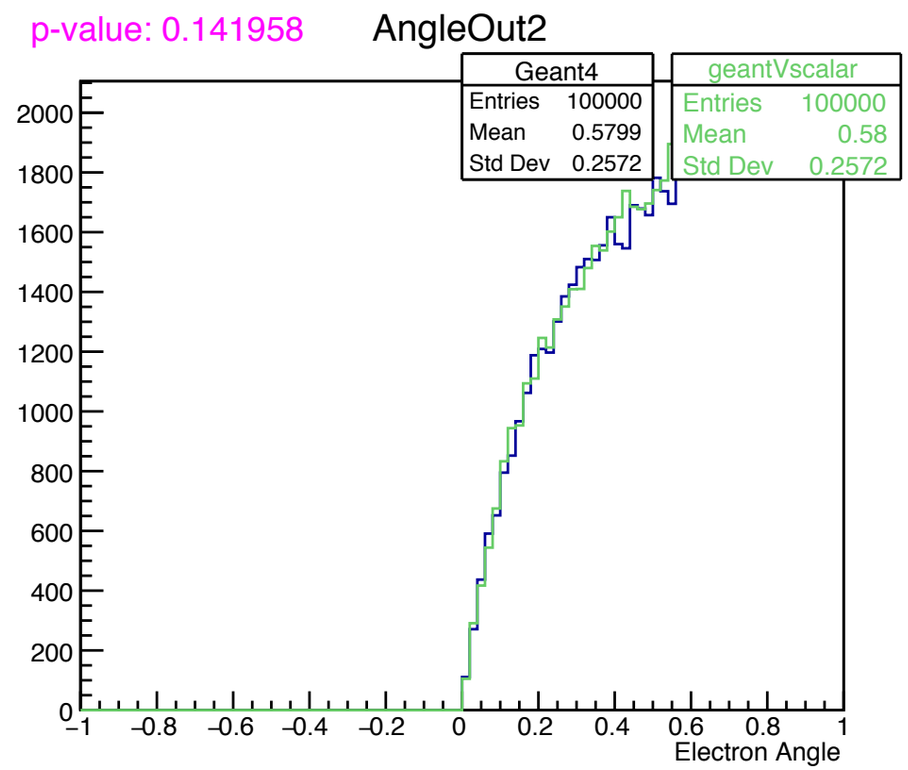


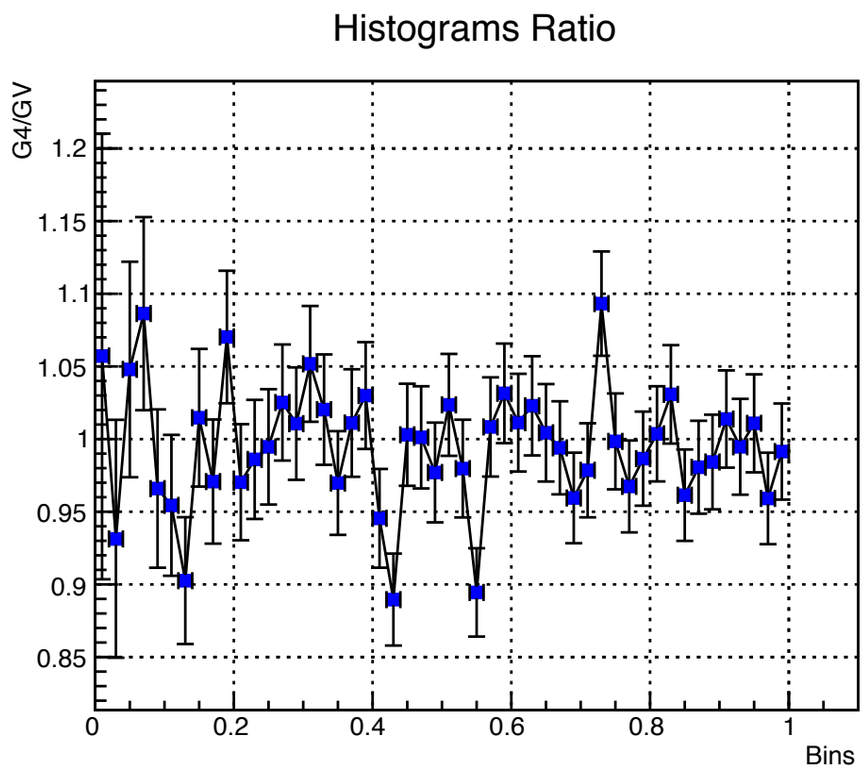


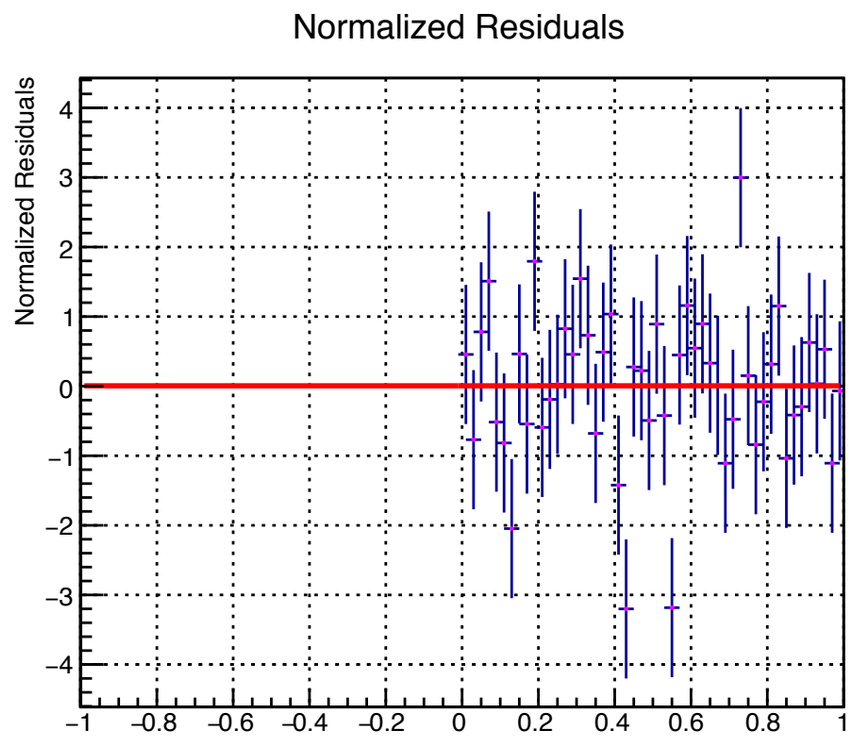


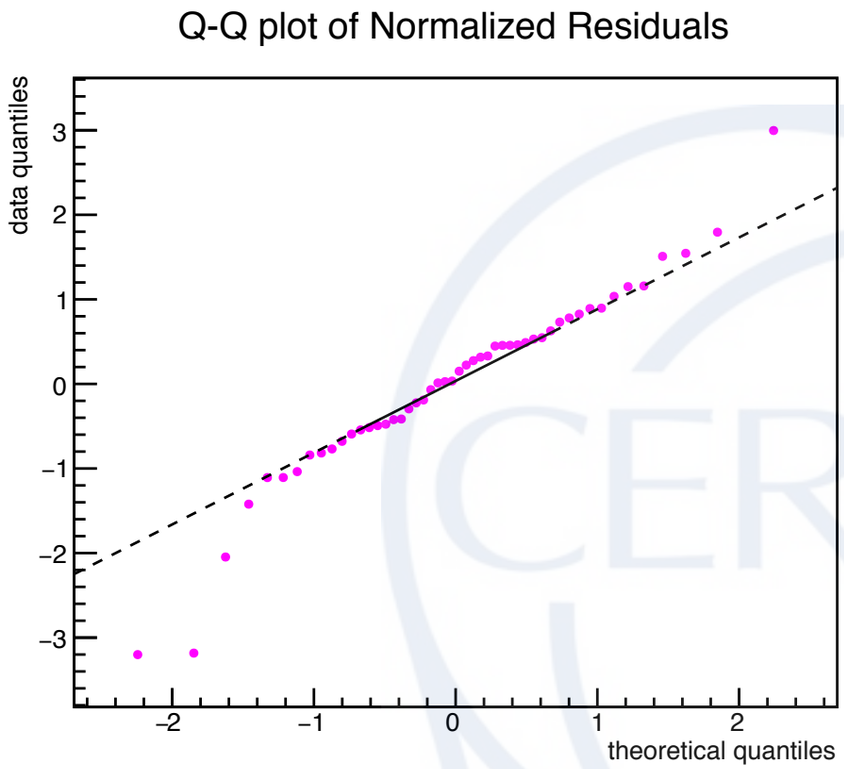
E_in=0.01MeV

Fixed





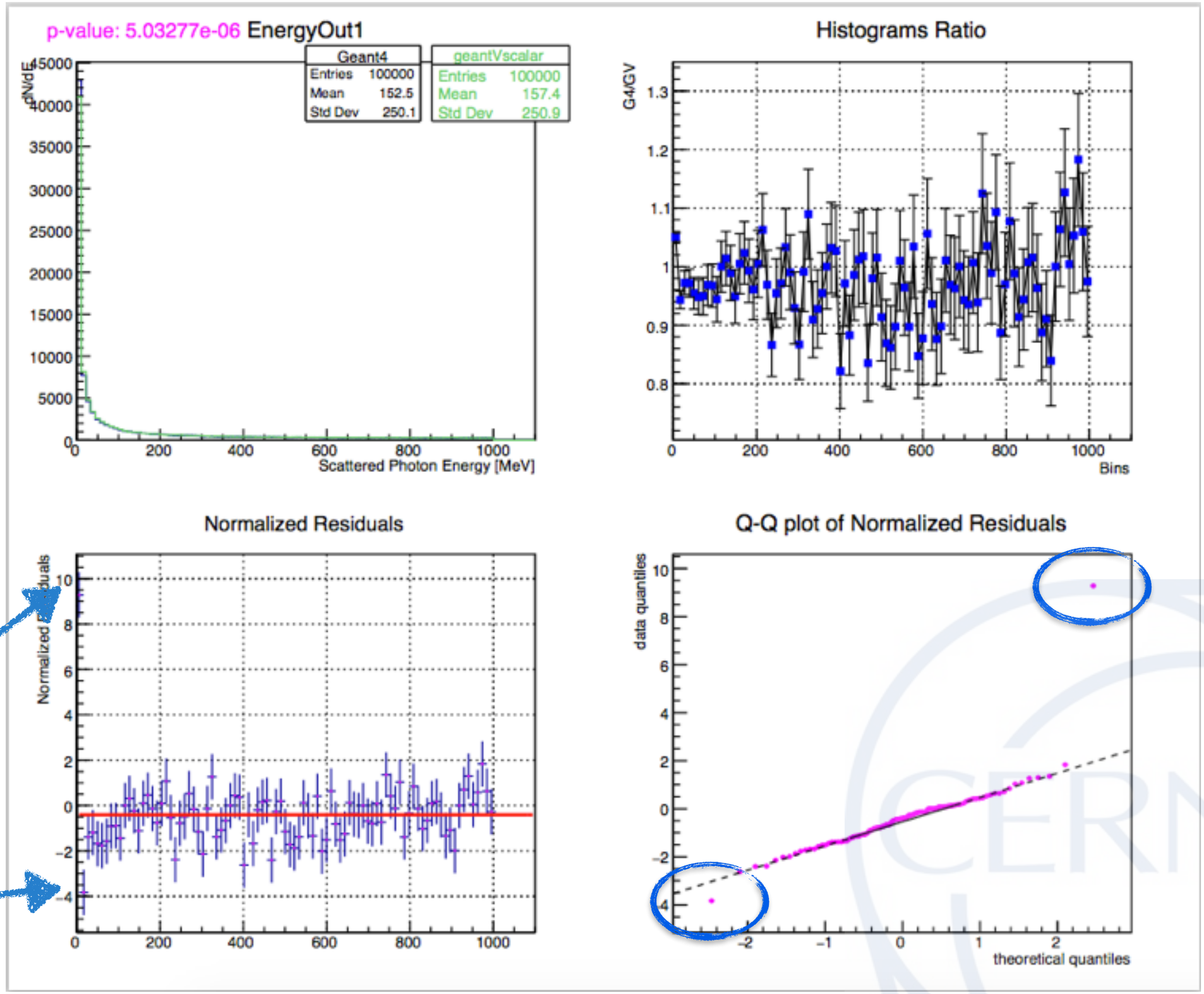




P-value Table

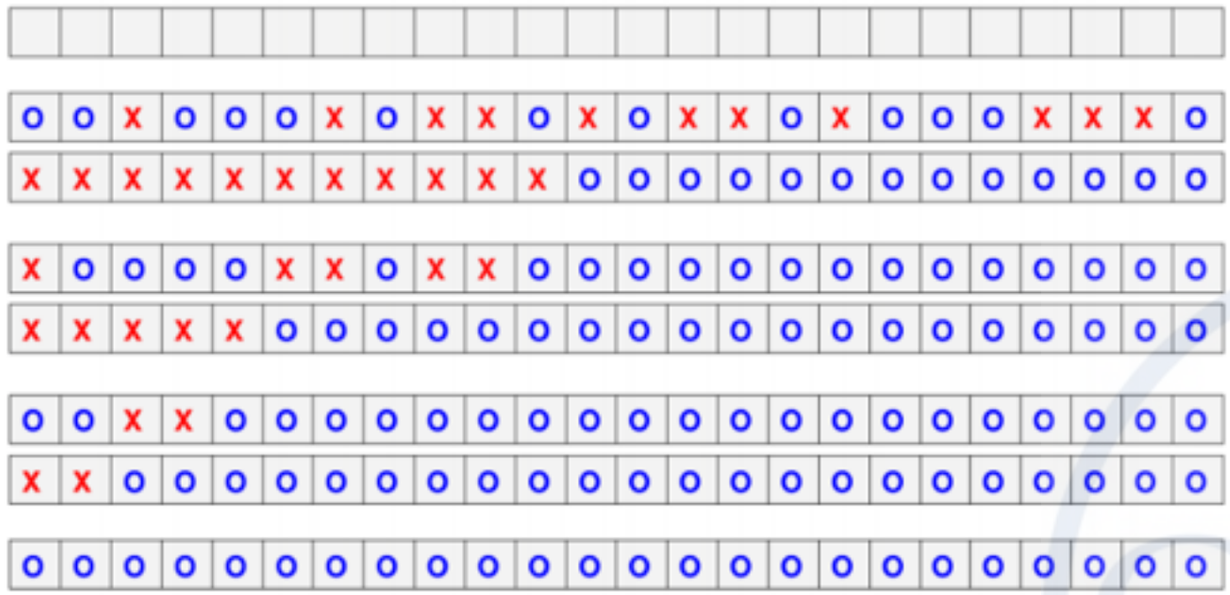
EnergyIn	*ValidationQ*	*Chi-2 test p-value*	*A-D test p-value*	*K-S test p-value*
0.01 MeV	EnergyOut1	0.131912	1	1
	EnergyOut2	0.492766	1	1
	AngleOut1	0.878847	1	0.999998
	AngleOut2	0.590701	1	0.999998
0.1 MeV	EnergyOut1	0.07002	1	1
	EnergyOut2	0.0333841	1	1
	AngleOut1	0.120768	0.999238	0.993765
	AngleOut2	0.897628	1	0.999998
1 MeV	EnergyOut1	0.228669	0.978313	0.812748
	EnergyOut2	0.220132	0.996374	0.906206
	AngleOut1	0.0470413	0.999927	0.999633
	AngleOut2	4.68076e-46	0.995252	0.993765
10 MeV	EnergyOut1	0.492207	0.983997	0.967068
	EnergyOut2	0.29053	0.977717	0.967068
	AngleOut1	0.289294	0.999259	0.999633
	AngleOut2	0.139839	1	1
100 MeV	EnergyOut1	0.564205	0.999778	0.967068
	EnergyOut2	0.817869	0.999553	0.967068
	AngleOut1	0.357838	1	0.993765
	AngleOut2	0.572742	1	1
1000 MeV	EnergyOut1	0.379965	0.998104	0.967068
	EnergyOut2	0.219433	0.990911	0.967068
	AngleOut1	1.18777e-163	0.568697	0.812748
	AngleOut2	1.52104e-193	0.896832	0.467558

E_in=1000 MeV


p-value: 5.03277e-06 EnergyOut1
Geant4
Entries 100000
Mean 152.5
Std Dev 250.1
geantVscalar
Entries 100000
Mean 157.4
Std Dev 250.9
dN/dE
Scattered Photon Energy [MeV]
Histograms Ratio
G4/GV
Bins
Normalized Residuals
Normalized Residuals
Q-Q plot of Normalized Residuals
data quantiles
theoretical quantiles

Alternative Sampling Methods

- Limitation of the discrete alias sampling method
  - the alias method with a finite bin size is subject to have biased outcomes if pdf is neither near constant nor linear within a bin
  - Ex.: $d\sigma/dE$ [$\gamma(E,0) \rightarrow \gamma(E',\sin\theta)$]: for large E and small E' -> E'/E^3
- Alternative techniques using the composition and rejection
  - Parallel (vector) + Sequential (scalar) loop over the vector width
  - Shuffling (try and unpack, overhead for reorganizing data)



  - hybrid (mixture of different methods in the parameter space)

See Soon's talk for performance

P-value Table

EnergyIn	*ValidationQ*	*Chi-2 test p-value*	*A-D test p-value*	*K-S test p-value*
1 MeV	EnergyOut1	0.580762	1	0.999998
	EnergyOut2	0.187933	1	0.999633
	AngleOut1	0.355483	1	0.999998
	AngleOut2	0.823083	1	1
10 MeV	EnergyOut1	0.00633212	0.999966	0.967068
	EnergyOut2	0.00199365	0.999996	0.999633
	AngleOut1	0.0351077	1	1
	AngleOut2	0.100245	1	1
100 MeV	EnergyOut1	0.834322	0.999999	0.993765
	EnergyOut2	0.903547	0.999992	0.999633
	AngleOut1	0.795684	1	1
	AngleOut2	0.809488	1	1
250 MeV	EnergyOut1	0.930848	1	0.999998
	EnergyOut2	0.856883	1	1
	AngleOut1	0.0656924	1	1
	AngleOut2	0.721054	1	1
500 MeV	EnergyOut1	0.937476	1	0.999998
	EnergyOut2	0.841945	1	0.999998
	AngleOut1	0.0737359	1	1
	AngleOut2	0.114897	1	1
1000 MeV	EnergyOut1	0.523284	1	0.999998
	EnergyOut2	0.777374	1	0.999998
	AngleOut1	0.638076	1	1
	AngleOut2	0.656735	1	1
10000 MeV	EnergyOut1	0.73755	0.999999	0.993765
	EnergyOut2	0.615561	1	0.999633
	AngleOut1	0.247631	1	1
	AngleOut2	0.360034	1	1
100000 MeV	EnergyOut1	0.121317	0.999999	0.993765
	EnergyOut2	0.19909	1	0.999633
	AngleOut1	0.881816	1	1
	AngleOut2	0.86112	1	1

Adaptive binning and transformation for the Alias method

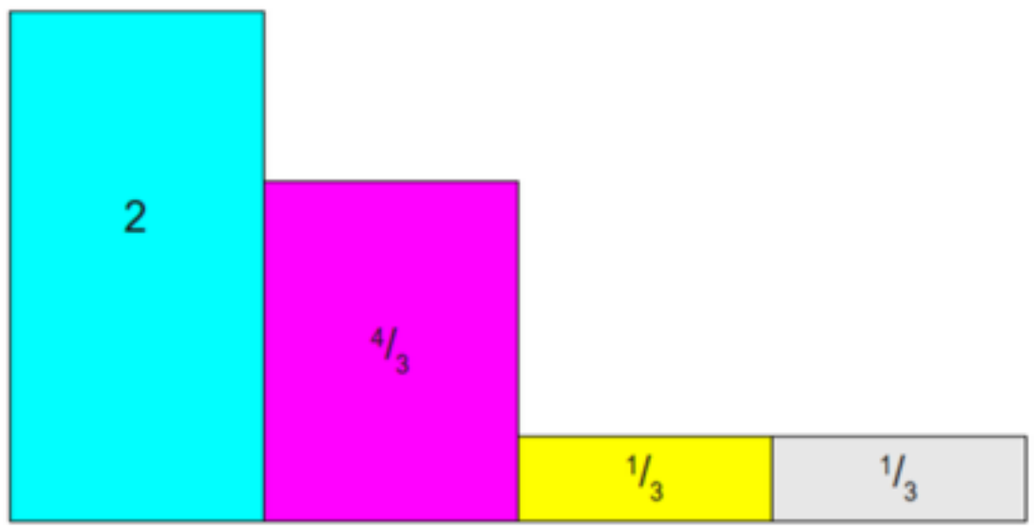
Considerations and further developments

- The actual implementation of the Alias sampling introduces discretization errors, especially when the sampled distributions are very steep
  - Found and fixed problems for Alias-Compton up to 100MeV
  - Investigating the threshold energy (boundary between Alias and Shuffling sampling methods 100 MeV at the moment)
- *Tools automation* is advanced, easing future validations
  - P-value Table automatically generated
  - Statistical analysis and graphs automatically generated
- Integration of Compton VecPhys Process with the GeantV Scheduler
- **Ongoing developments:**
  - Investigation of *different sampling techniques* and *improvement of* the Alias sampling (non-linear binning, transformation)
  - Extend validation for intermediate energies (around 100MeV and middle of the bins)

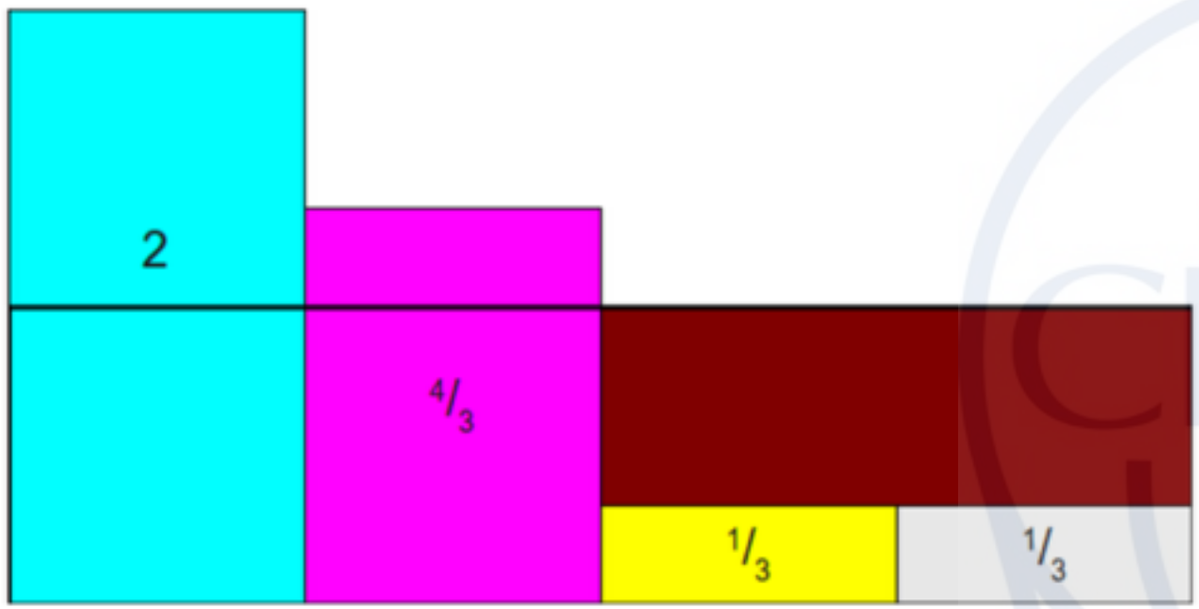
Backup

Alias method

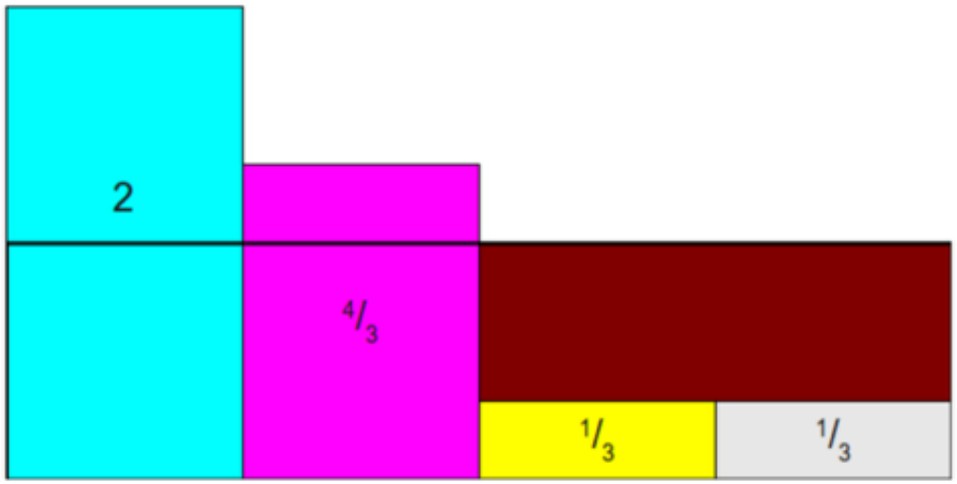
Initial pdf (equal likelihood=1/4)

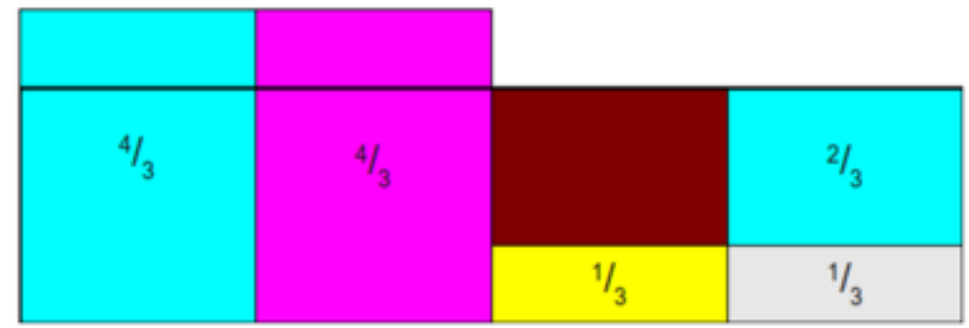


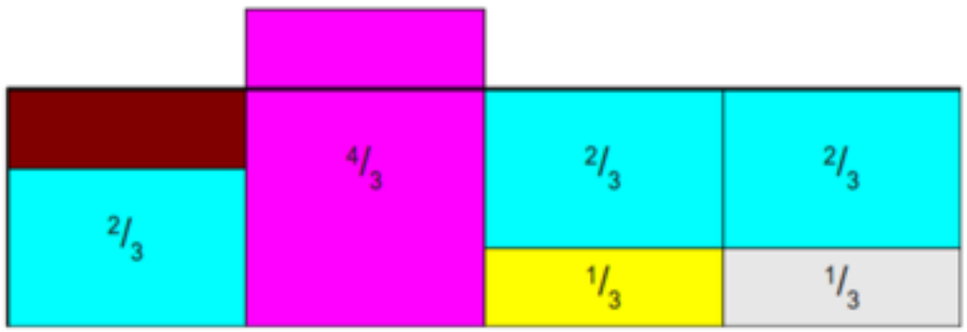
Scaled probabilities so that a prob of 1/4 would weight 1



Alias method







Algorithm: Alias Method

- **Initialization**:
  1. Create arrays $Alias$ and $Prob$, each of size n.
  2. Create a balanced binary search tree T.
  3. Insert $n \cdot p_i$ into T for each probability i.
  4. For $j = 1$ to $n - 1$:
     1. Find and remove the smallest value in T; call it p_l.
     2. Find and remove the largest value in T; call it p_g.
     3. Set $Prob[l] = p_l$.
     4. Set $Alias[l] = g$.
     5. Set $p_g := p_g - (1 - p_l)$.
     6. Add p_g to T.
  5. Let i be the last probability remaining, which must have weight 1.
  6. Set $Prob[i] = 1$.
- **Generation**:
  1. Generate a fair die roll from an n-sided die; call the side i.
  2. Flip a biased coin that comes up heads with probability $Prob[i]$.
  3. If the coin comes up "heads," return i.
  4. Otherwise, return $Alias[i]$.

Improvements of the algorithm - Vose's algorithm

This algorithm was originally described in the paper "A Linear Algorithm For Generating Random Numbers With a Given Distribution" by Michael Vose

The idea behind Vose's algorithm is to maintain two worklists, one containing the elements whose height is less than 1 and one containing the elements whose height is at least 1, and to repeatedly pair the first elements of each worklist.

On each iteration, we consume the element from the "small" worklist, and potentially move the remainder of the element from the "large" worklist into the "small" worklist.

The algorithm maintains several invariants:
- The elements of the "small" worklist are all less than 1.
- The elements of the "large" worklist are all at least 1.

Algorithm	Initialization Time		Generation Time		Memory Usage	
	Best	Worst	Best	Worst	Best	Worst
Alias Method	$O(n \log n)$		$\Theta(1)$		$\Theta(n)$	
Vose's Alias Method	$\Theta(n)$		$\Theta(1)$		$\Theta(n)$	

Ref: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=992137

Algorithm: (Unstable) Vose's Alias Method

- **Initialization**:
  1. Create arrays $Alias$ and $Prob$, each of size n.
  2. Create two worklists, $Small$ and $Large$.
  3. Multiply each probability by n.
  4. For each scaled probability p_i:
     1. If $p_i < 1$, add i to $Small$.
     2. Otherwise ($p_i \geq 1$), add i to $Large$.
  5. While $Small$ is not empty:
     1. Remove the first element from $Small$; call it l.
     2. Remove the first element from $Large$; call it g.
     3. Set $Prob[l] = p_l$.
     4. Set $Alias[l] = g$.
     5. Set $p_g := p_g - (1 - p_l)$.
     6. If $p_g < 1$, add g to $Small$.
     7. Otherwise ($p_g \geq 1$), add g to $Large$.
  6. While $Large$ is not empty:
     1. Remove the first element from $Large$; call it g.
     2. Set $Prob[g] = 1$.
- **Generation**:
  1. Generate a fair die roll from an n-sided die; call the side i.
  2. Flip a biased coin that comes up heads with probability $Prob[i]$.
  3. If the coin comes up "heads," return i.
  4. Otherwise, return $Alias[i]$.

Is it a stable version?

Unfortunately, the above algorithm, as written, is not numerically stable. Two sources of inaccuracy:

- The computation to determine whether or not a probability belongs to the Small or Large group may be inaccurate. Specifically, it may be possible that scaling up the probabilities by a factor of n may cause probabilities equal to 1/n to end up being slightly less than 1 (ending up in the Small list rather than in the Large one)

- The computation that subtracts the appropriate probability mass from a larger probability is not numerically stable and may introduce significant rounding errors.

A Vose's Alias method stable implementation

We will update the inner loop of the algorithm so that it terminates whenever either of the two worklists are empty, so we don't accidentally end up looking at nonexistent elements from the Large worklist.

Second, when one worklist is empty, we'll set the remaining probabilities of the elements in the other worklist to all be 1 since, mathematically, this should only occur if all of the remaining probabilites are precisely equal to 1.

Finally, we'll replace the computation that updates the large probabilities with a slightly more stable computation.

Algorithm: Vose's Alias Method

- **Initialization**:
  1. Create arrays $Alias$ and $Prob$, each of size n.
  2. Create two worklists, $Small$ and $Large$.
  3. Multiply each probability by n.
  4. For each scaled probability p_i:
     1. If $p_i < 1$, add i to $Small$.
     2. Otherwise ($p_i \geq 1$), add i to $Large$.
  5. While $Small$ and $Large$ are not empty: *($Large$ might be emptied first)*
     1. Remove the first element from $Small$; call it l.
     2. Remove the first element from $Large$; call it g.
     3. Set $Prob[l] = p_l$.
     4. Set $Alias[l] = g$.
     5. Set $p_g := (p_g + p_l) - 1$. *(This is a more numerically stable option.)*
     6. If $p_g < 1$, add g to $Small$.
     7. Otherwise ($p_g \geq 1$), add g to $Large$.
  6. While $Large$ is not empty:
     1. Remove the first element from $Large$; call it g.
     2. Set $Prob[g] = 1$.
  7. While $Small$ is not empty: *This is only possible due to numerical instability.*
     1. Remove the first element from $Small$; call it l.
     2. Set $Prob[l] = 1$.
- **Generation**:
  1. Generate a fair die roll from an n-sided die; call the side i.
  2. Flip a biased coin that comes up heads with probability $Prob[i]$.
  3. If the coin comes up "heads," return i.
  4. Otherwise, return $Alias[i]$.