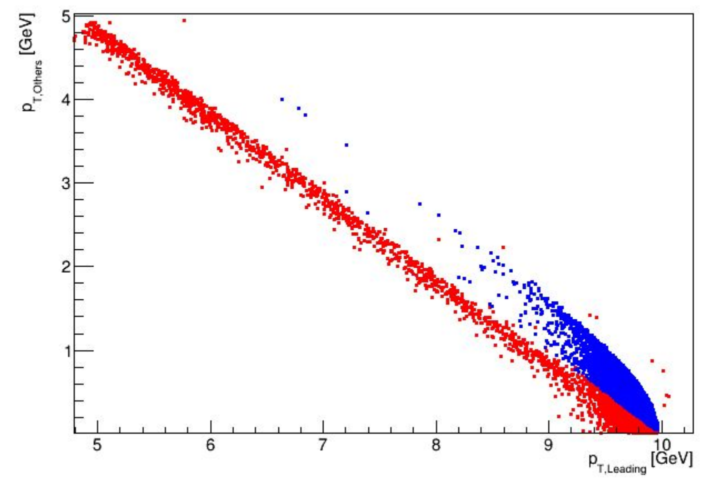
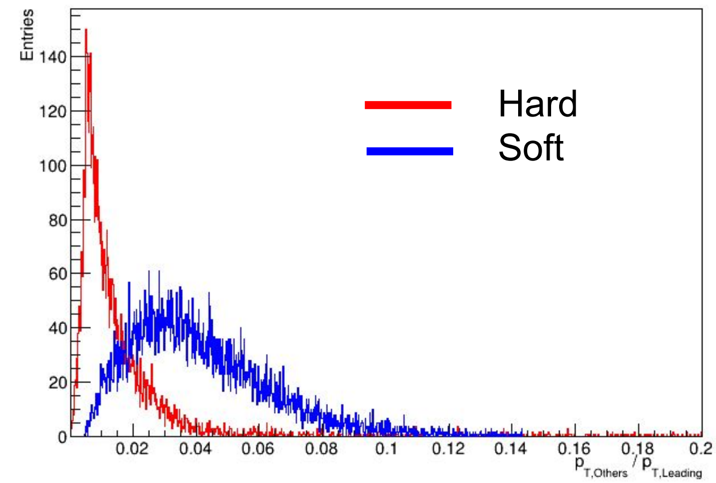
Method for parametrisation of nuclear interaction

Parametrised nuclear interaction

- Dataset production
  - Geant4 used as reference truth data
  - Particle gun used for <u>different, fixed</u> initial momenta
  - Record of final state (FS) particles + properties
- Filter events
  - Only interested in common hadrons
  - Remove leptons, photons, nuclei etc. from FS
    - These are usually small / rare contributions
  - Remove events with decays
    - Significant different kinematics
    - Handled by different module
  - Remove FS particles below 50 MeV
    - i.e. particles captured in detector material
- Event labeling / categorisation
  - Events from different categories have different FSs
  - Labels:
    - Nuclear interaction in event?
    - FS Multiplicity
    - Soft / hard nuclear interaction

2 particles FS from 10 GeV π^+



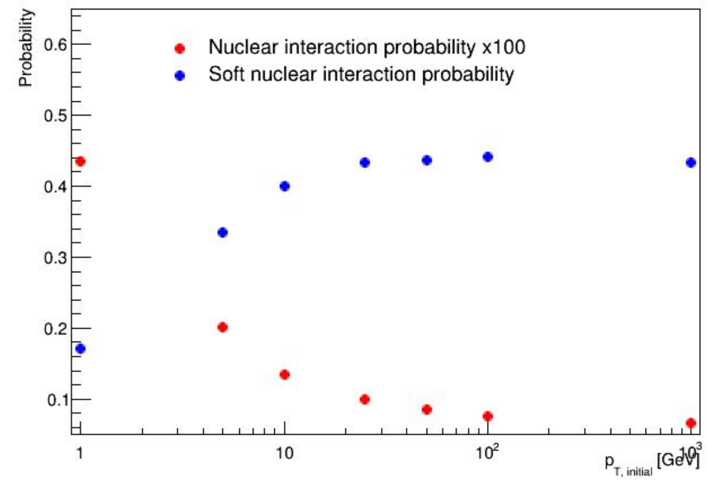


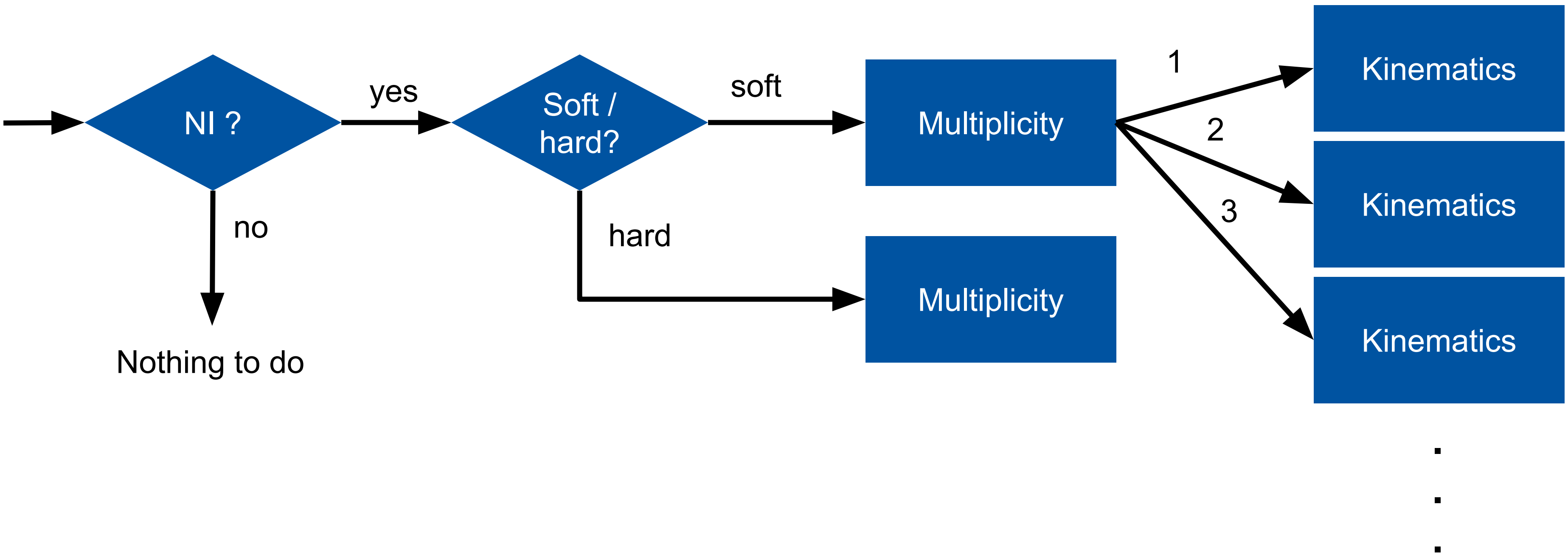
Labeled soft if all fulfilled:

- Did the initial particle survive? (same PDG ID & most of the initial momentum)
- Kinematic separation by $\Sigma_i p_{T,i} \geq p_{T,initial}$
- Is there a FS that looks like that?

Workflow

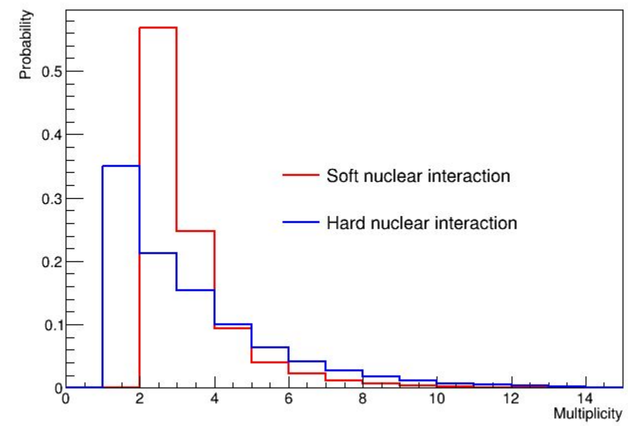
- Process divided into several steps
  - Did the nuclear interaction (NI) occur?
    - $P(d|\lambda) = exp(-d/\lambda)$
    - Treat momentum dependence
  - Soft / hard interaction
  - Multiplicity
  - FS particle properties





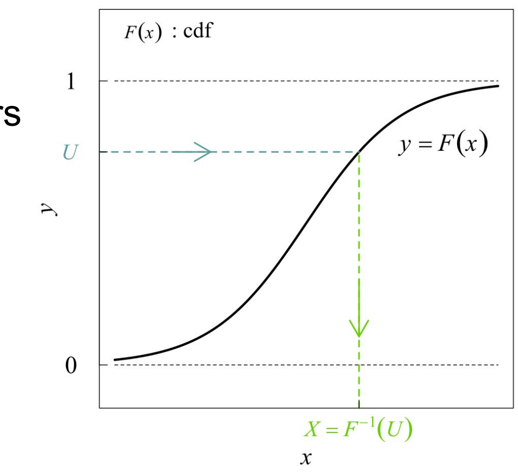
Multiplicity

- Multiplicity distribution derived from data
- Sampling according to distribution required



- Inverse transform method
  - Consider the CDF
  - $y = F(x) \propto U(0,1) \Leftrightarrow x \propto F^{-1}(U)$
  - Sample from uniform distribution rather than from F(x)
  - **Bin lookup** instead of searching for F → No fit functions!

- At that point the interaction type and the multiplicity is known
- Particle kinematics need to be calculated next
- Starting with the momentum:
  - Each generated particle is constrained by the momentum of others
  - Ordering required for structuring
    - FS particles are ordered by momentum (= generations)
  - Individual FS momentum distribution derived for each generation
  - Inverse transform method applicable (theoretically…)



Correlated sampling

- Goal: Find $\vec{p} = \{p_1, p_2, \ldots, p_n\}$ for multiplicity n and absolute values p_i
- Momenta constrained by correlation
  - Distributions transformed to $erf^{-1}(2F(x)-1) \propto G(0,1)$ with the gaussian distribution G
  - This is equivalent to a mapping from Uniform to Gauss distribution
  - Considering $F(x) \Rightarrow \vec{F}(\vec{p})$ which leads to

$$G(\vec{p}'|\vec{0},\vec{1}) \propto exp(-\frac{1}{2}\vec{p}'^T \Sigma^{-1} \vec{p}')$$

Explicit correlation!

- Correlated sampling
  - De-correlate the distribution = diagonalise Σ
  - Transformation of G into eigenspace of Σ with $\Sigma_{i,j} = \begin{cases} e_i & i = j \\ 0 & i \neq j \end{cases}$ with the eigenvalues $\{e_i\}$

  $\longrightarrow$ Sample distribution becomes:

$$G(\vec{p}''|\vec{0},\overrightarrow{\sqrt{e}}) \propto exp(-\frac{1}{2}\Sigma_i \frac{p_i''^2}{e_i}) = \Pi_i \underbrace{G(p_i''|0,\sqrt{e_i})}$$

Independent distributions!

- Produced samples $\vec{p}''$ are transformed back to recieve $\vec{p}$
- Momentum samples are related relatively to each other
  - No constrained for their combination, yet
  - Combined constraint given by adding the distribution of the **sum of momenta** → n+1 distributions
  - Allows re-scaling afterwards
- Convenience scaling: Parametrisation of $\vec{p} = \left\{\frac{p_1}{p_{initial}}, \frac{p_2}{p_{initial}}, \ldots, \frac{p_n}{p_{initial}}, \frac{\Sigma_i p_i}{p_{initial}}\right\}$

Angular distributions

- Sampled momenta allow calculation of polar opening angles θ_{ij} via invariant mass between particle i and j

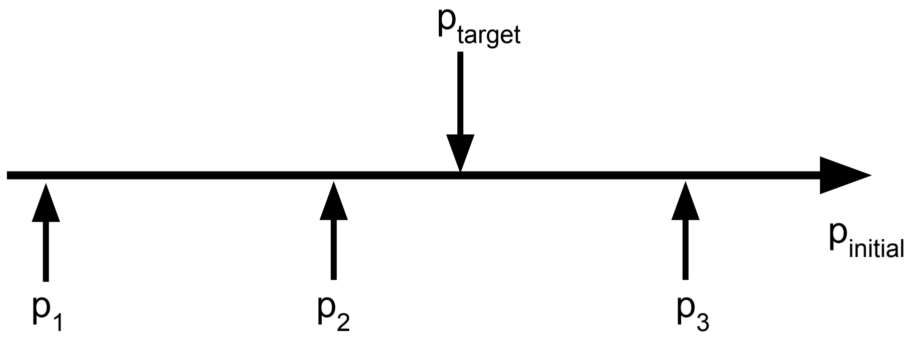
$$M_{ij} = \sqrt{2p_i p_j(1 - cos(\theta_{ij}))} \Leftrightarrow \theta_{ij} = cos^{-1}\left(1 - \frac{M_{ij}^2}{2p_i p_j}\right)$$

- First approach used j = i + 1
  - Dangerous since each value would affect the next one
  - An unfortunate sampling would produce strange results
  - More stable choice: Consider angle wrt initial particle (i = 0)
- Same sampling idea used as for the momenta
  - But without re-scaling → n distributions
  - Produces set of invariant masses
- Matching
  - Domain of cos^{-1} is [-1,1] but p_j and M_{0j} are sampled independently
  - A match is achieved iff $\forall j : \frac{M_{0,j}^2}{2p_0 p_j} \in [0, 2]$
  - If this is not the case:
    - As these are 2 correlated samples, no value could be modified independently
    - Need to re-sample a set → It was chosen to be the momenta
    - This scenario occurs in low density regions in the M-p-phase space
- Calculated polar opening & (assumed) azimuthal angle $\propto U(0, 2\pi)$ can be translated to particle angles

PDG ID and interpolation idea

- FS particles are recorded and re-ordered
- PDG ID association by creation a dedicated table
- Model:
  - i-th particle only creates the (i+1)th particle
  - Initial particle produces the 1st particle
- Look-up table derived from data for which PDG ID produces which PDG IDs with which probability

- Until now: Parametrisation using a sample with fixed initial momentum
- Next step: Interpolation using the parametrisations
- Strategy:
  - Repeat the parametrisation for initial momenta $p_{initial}$ → Parametrisations for $p_1,...,p_n$ obtained
  - Construct FS for p_{target} by using only the two neighbours p_2 and p_3



Interpolation

- Spacing between the samples
  - If the distributions of p_2 look similar to the ones from p_3 then the ones from p_{target} should be similar, too
    - Refers to the general shape (except a scaling factor)
  - If $p_{initial}$ becomes bigger the shapes of the distributions change slower
    - Spacing increases as $p_{initial}$ is increased
- How to obtain the target distribution from p_2 and p_3?
- Two ideas were tested:
  - (Considering only the kinematics)
  - Distributions are assumed to be similar → Spectra of eigenvalues are also similar (for M&p)
    - Define a **weighted set of eigenvalues** from p_2 and p_3 that looks like a set from p_{target}
    - Sample from this distribution (in eigenspace)
    - Perform the back-transformations for p_2 and p_3 independently
    - Receive a **weighted** combination from both sets
    - It was observed that the **matching rate dropped** significantly
  - Use the weights as probabilities to choose between p_2 and p_3
    - Perform all samples & back-transformations using the chosen one
    - Scale results to p_{target}
    - Provided a **high matching rate** but allows only **one weighting scheme**
- Latter more promising
- Quality of interpolation depends (only) on the weighting scheme

Summary

- Parametrisation for nuclear interaction derived
- Process is separable into individual sub-steps
- Allows sampling from histograms without requiring fit functions
- Out-of-the-box applicable for arbitrary initial momentum
- Parametrisation can be extended for interpolation
  - Only requires weighting of neighbouring distributions
- Spacing between parametrisation adaptable
  - New samples extend the set of parametrisations
    - No re-parametrisation of prior samples required!
  - Interpolation quality can be improved by additional samples
    - Similarity can be improved
    - Importance of weighting scheme can be suppressed