Improved tuning method for Monte Carlo generators

Fabian Klimpel

38. IMPRS workshop 2017

Munich, 10.07.2017

Technische Universität München

$\Delta p \cdot \Delta q \geq \frac{1}{2}\hbar$

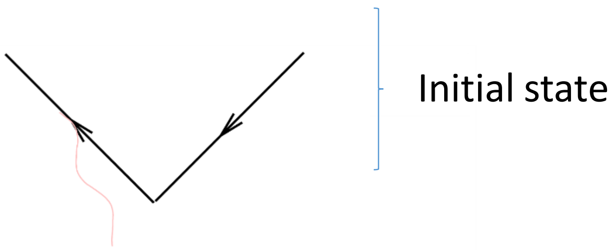
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Overview

1. Introduction to parameter tuning
2. Reproduction of a previous tuning
3. Introducing new approaches for parameter tuning
4. Summary and outlook

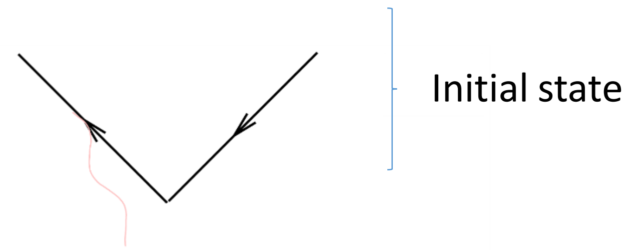
What is parameter tuning?

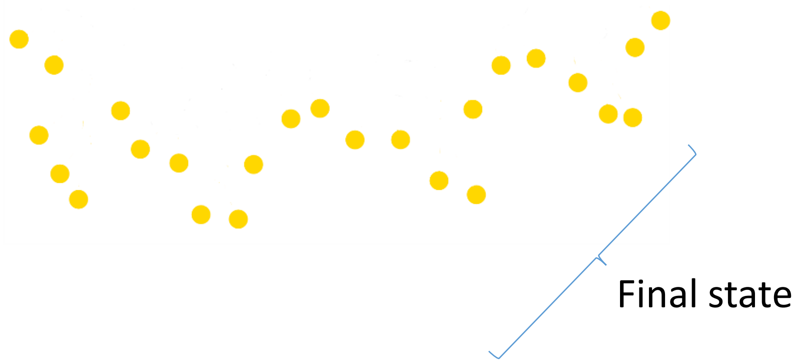
- <u>Experimental setup (e^+e^-):</u>
- Before the particle collision:
    - Initial state is (approximately) known



What is parameter tuning?

- Experimental setup (e^+e^-):
- Before the particle collision:
  - Initial state is (approximately) known
- After the collision:
  - Final state is measured (particle level)

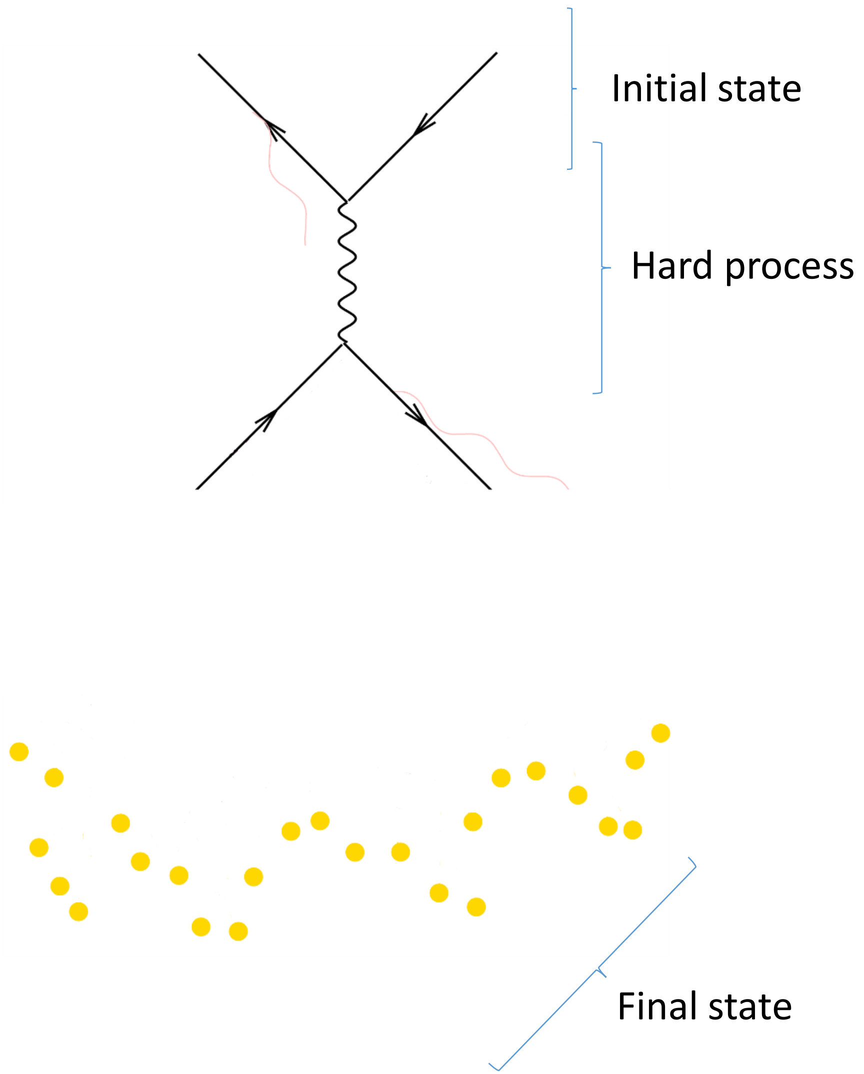




What is parameter tuning?

- <u>Experimental setup (e^+e^-):</u>
- Before the particle collision:
  - Initial state is (approximately) known
- After the collision:
  - Final state is measured (particle level)

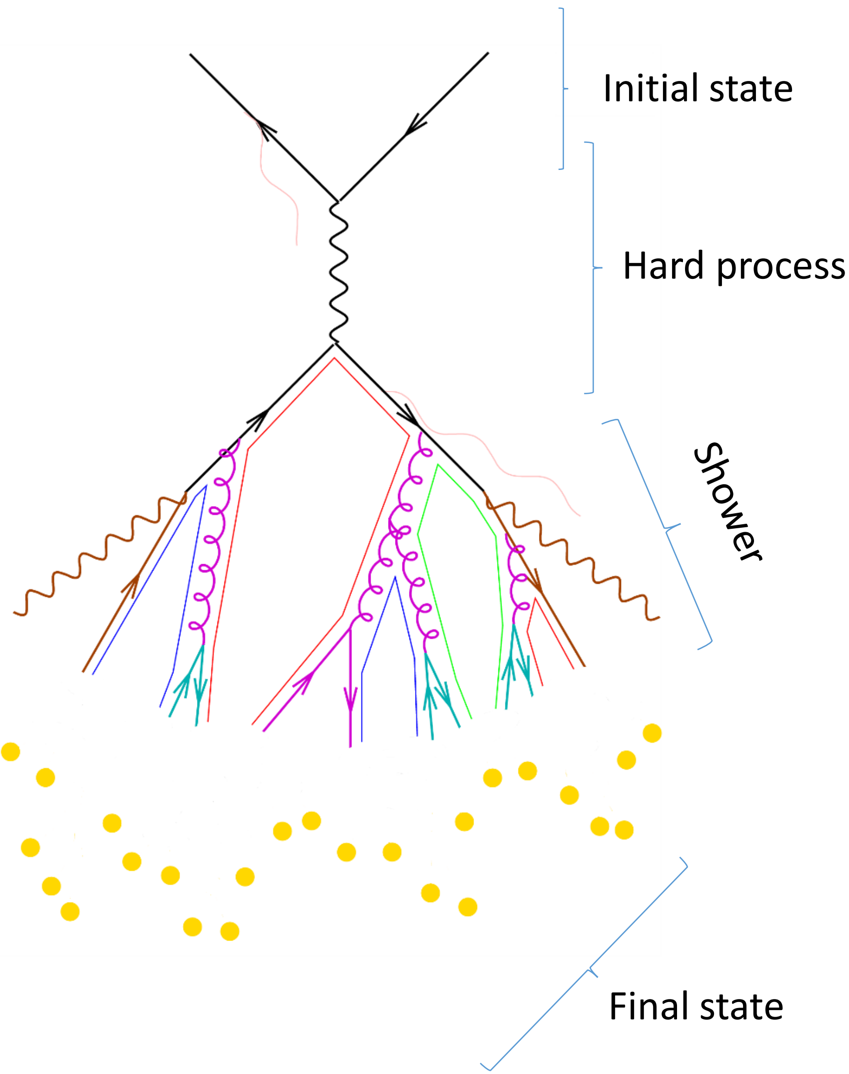
- <u>Monte Carlo (MC) generators:</u>
- Hard process: calculable



What is parameter tuning?

- Experimental setup (e^+e^-):
- Before the particle collision:
  - Initial state is (approximately) known
- After the collision:
  - Final state is measured (particle level)

- Monte Carlo (MC) generators:
- Hard process: calculable
- Parton shower: „modelled QCD“



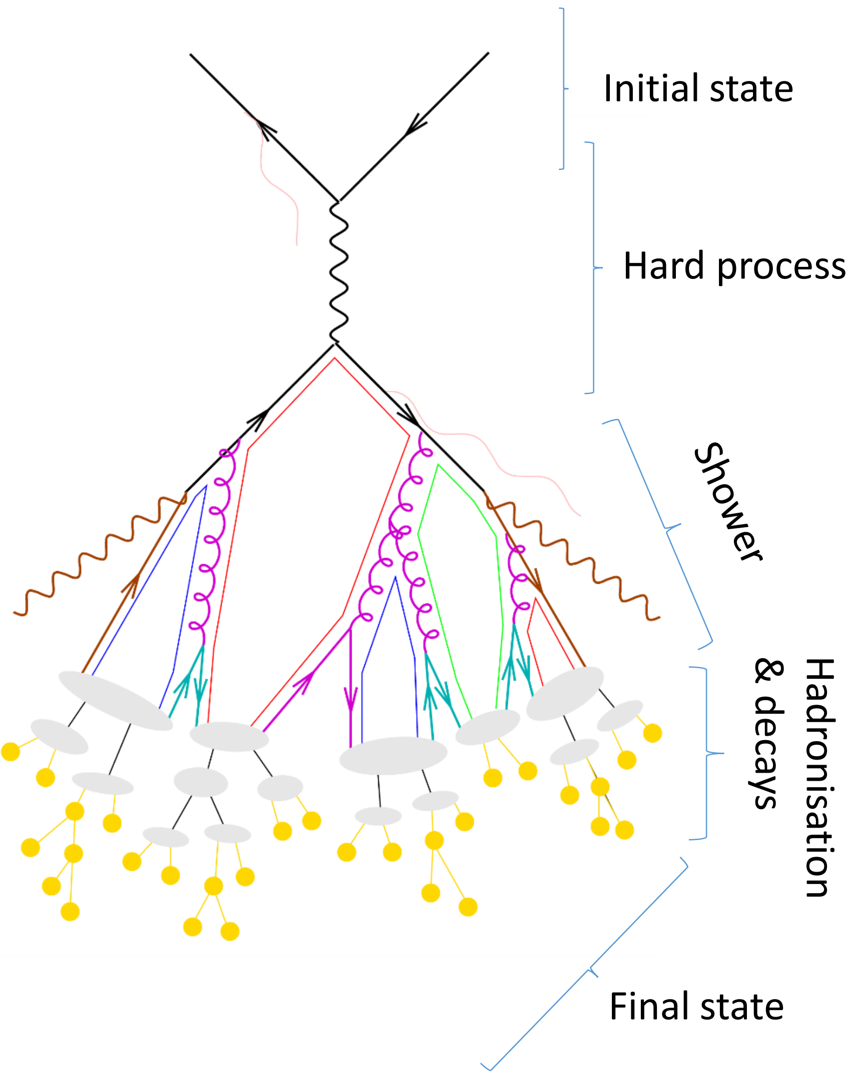
What is parameter tuning?

- <u>Experimental setup (e^+e^-):</u>
- Before the particle collision:
  - Initial state is (approximately) known
- After the collision:
  - Final state is measured (particle level)

- <u>Monte Carlo (MC) generators:</u>
- Hard process: calculable
- Parton shower: „modelled QCD“
- Hadronisation: modelled

➢ Adjust model parameters to optimize reproduction of measurement

- **But**: How to get their values?

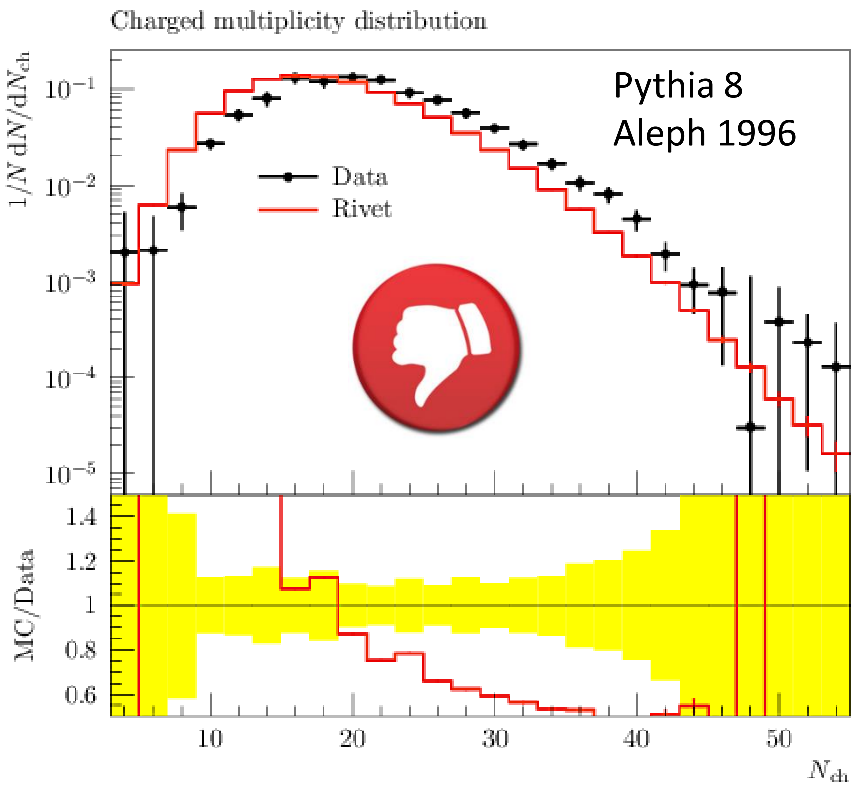


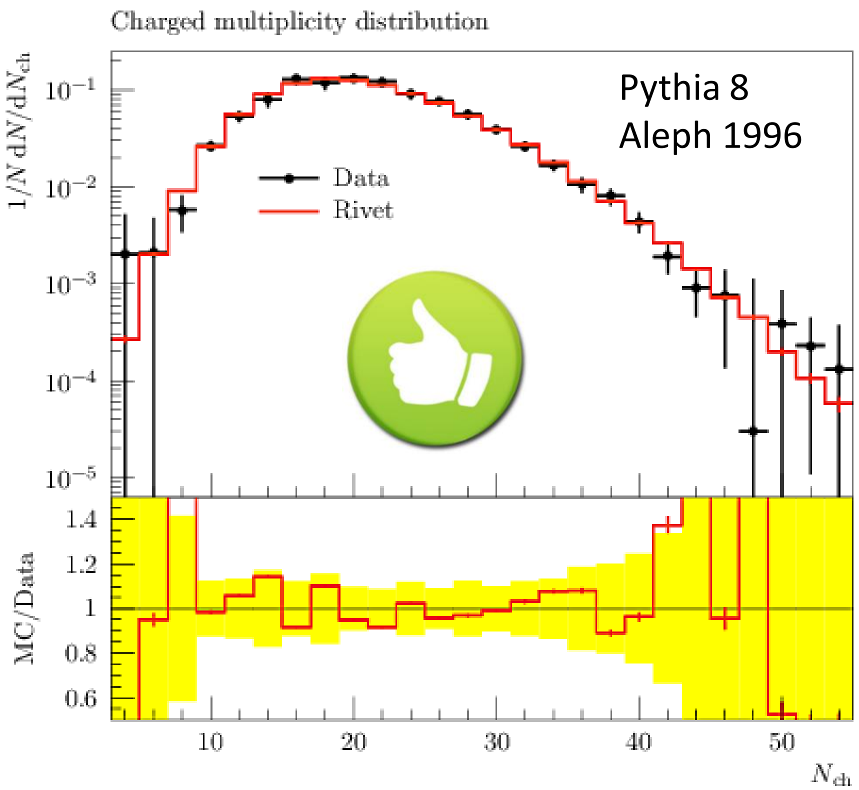
Parameter tuning - general

- Model parameters can be varied
  - ➢ parameters need values that describe the data best = tuning

Parameter tuning - general

- Model parameters can be varied
  - ➢ parameters need values that describe the data best = tuning
- Example: e^+e^- collisions at $\sqrt{s} = 91.2$ GeV with two different parameter sets



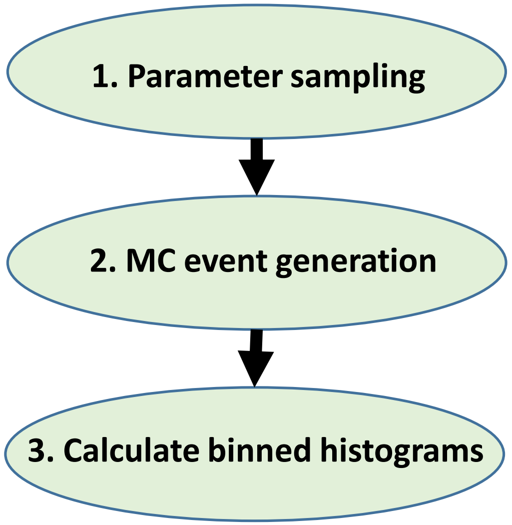


How to tune? – general

1. Sample random model parameter vectors $\boldsymbol{p}_{\mathrm{i}}$ in predefined ranges

2. Use every $\boldsymbol{p}_{\mathrm{i}}$ (= „run“) as input for the MC generator
   - ➢ High CPU consumption

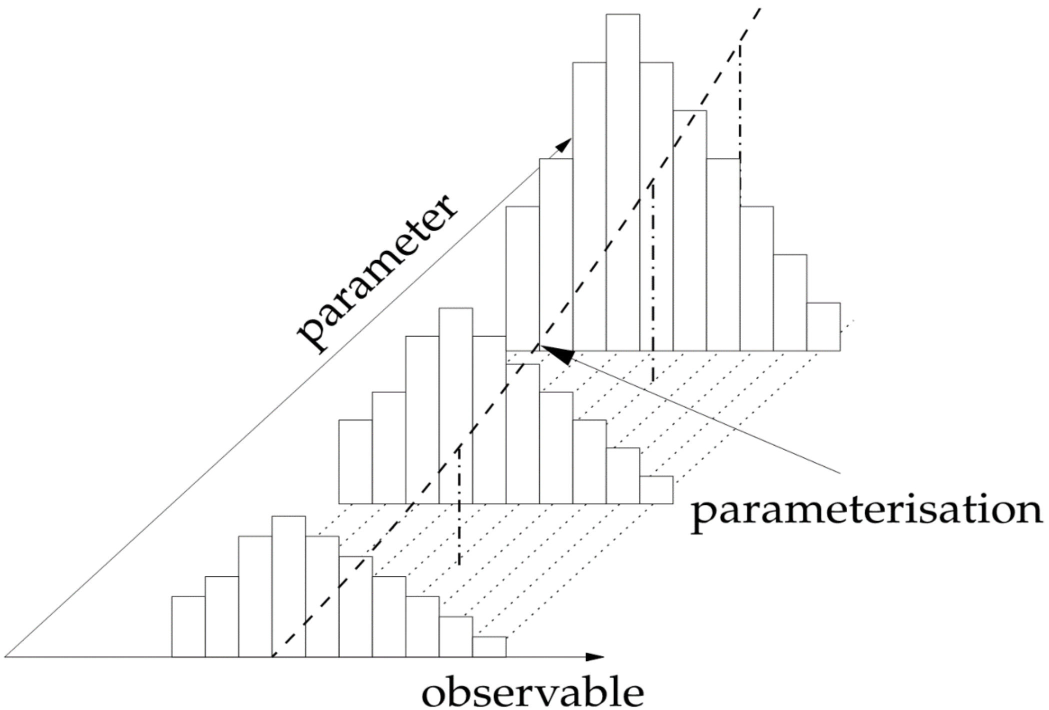
3. Extract observables from each run
   - ➢ binned histograms for each observable



How to tune? – Interpolation

- Assumption: sufficiently smooth change in a bin *b* while changing the value of a parameter

➢ Each bin can be parametrised by a polynomial function $f^{(b)}(\boldsymbol{p})$



Source: H. Schulz, Systematic Event Generator Tuning with Professor

How to tune? – general

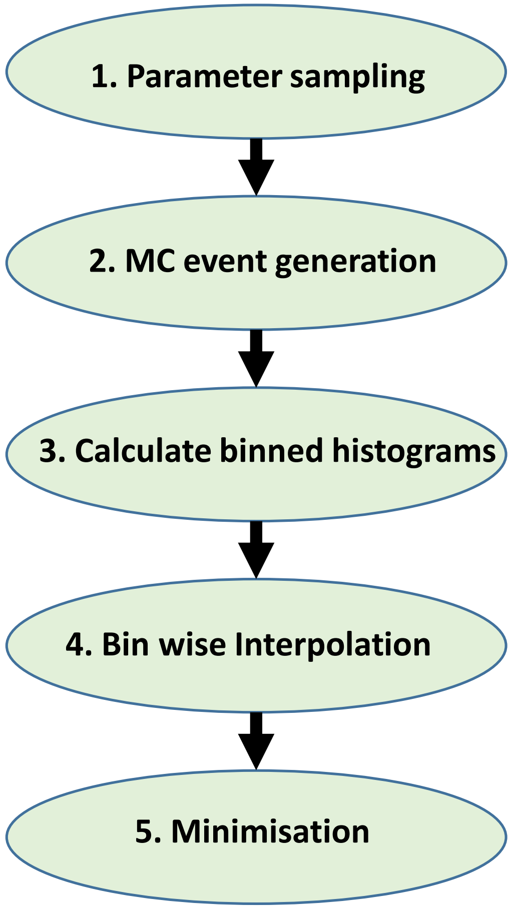
1. Sample random model parameter vectors $\boldsymbol{p}_{\mathrm{i}}$ in predefined ranges

2. Use every $\boldsymbol{p}_{\mathrm{i}}$ (= „run“) as input for the MC generator
   - ➢ High CPU consumption

3. Extract observables from each run
   - ➢ binned histograms for each observable

4. Calculate $f^{(b)}(\boldsymbol{p})$ as function of model parameters $\boldsymbol{p}$

5. Minimise $\chi^2(\boldsymbol{p})$



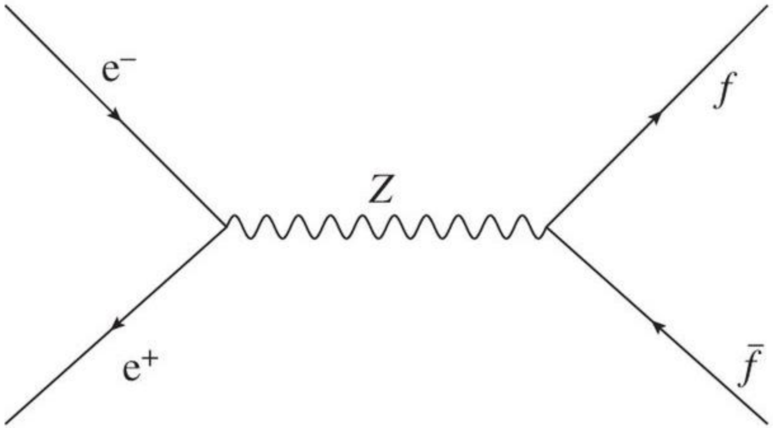
Overview

1. Introduction to parameter tuning
2. Reproduction of a previous tuning
3. Introducing new approaches for parameter tuning
4. Summary and outlook

Tuning application

Source: N. Fischer, Angular Correlation and Soft Jets as Probes of Parton Showers

- Reproduced standard Pythia tune
- Hard process:

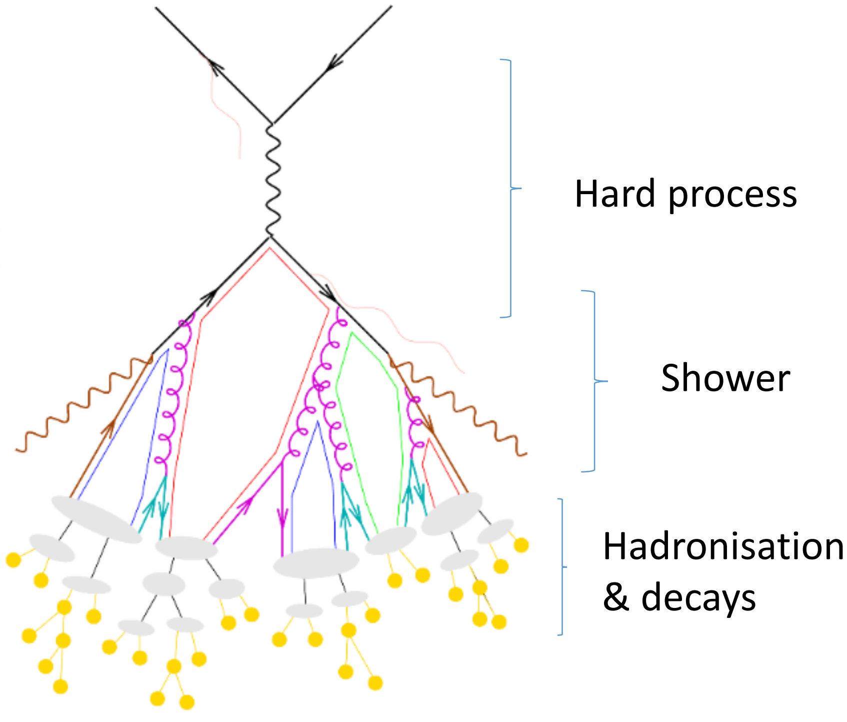


f = u,d,s,c,b quarks

- Events simulated with Pythia 8 at $\sqrt{s} = m(Z^0) = 91.2$ GeV, LO + LL
- Settings were extracted from standard tuning
- Data from LEP (Aleph, Opal, Delphi), PETRA (Jade) & PDG combinations
  - Phys. Rept., 294:1 (1998); PLB 512:30 (2001); ZP C73:11 (1996); EPJ C17:19 (2000); PLB 667:1 (2008)

Which parameters to tune?

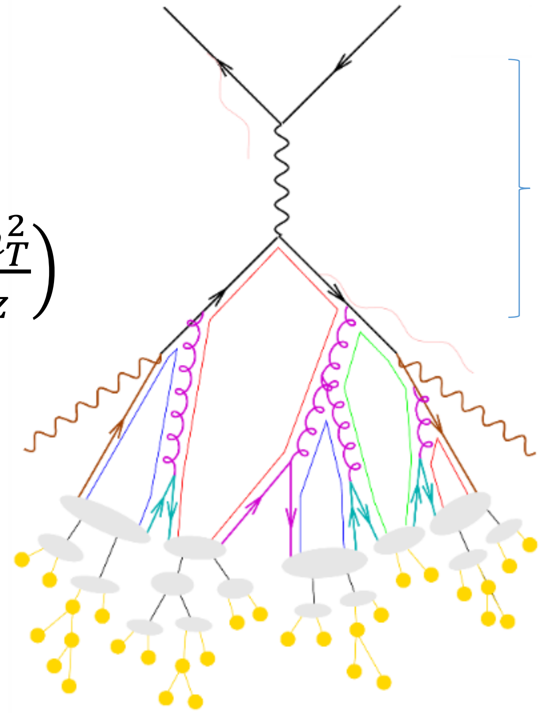
1. Parton shower model based on QCD:
   - High energy partons radiate gluons with probability ~α_s → shower
   - Shower cutoff at $p_{T,min}$
     - ➢ phase transition, confinement



Which parameters to tune?

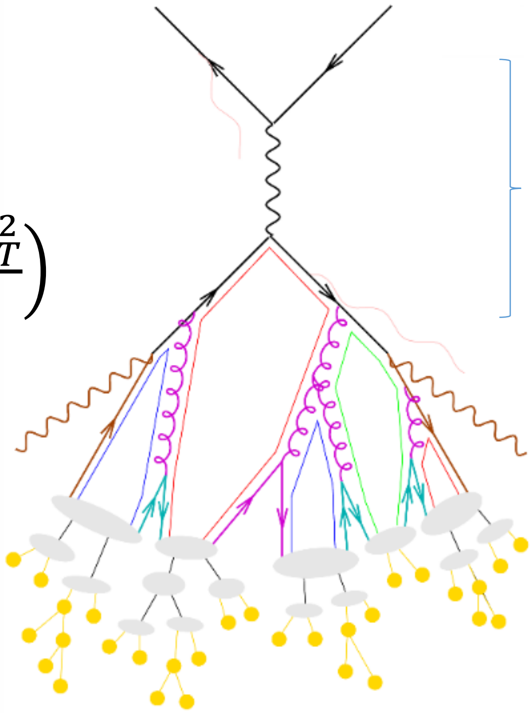
1. Parton shower model based on QCD:
   - High energy partons radiate gluons with probability $\sim\boldsymbol{\alpha_s} \rightarrow$ shower
   - Shower cutoff at $\mathbf{p_{T,min}}$
     - phase transition, confinement

2. Hadronisation (Lund-string model):
   - $f(z) \propto \frac{1}{z}(1-z)^{a_{\mathbf{Lund}}} \exp\left(-\boldsymbol{b}_{\mathbf{Lund}} \frac{m_T^2}{z}\right)$
   - Baryons: $a = a_{Lund} + \boldsymbol{a}_{\mathbf{ExtraDiquark}}$
   - $P(p_T) \propto \exp(-\frac{1}{2}\frac{p_T^2}{\boldsymbol{\sigma}^2})$



Which parameters to tune?

1. Parton shower model based on QCD:
   - High energy partons radiate gluons with probability $\sim\boldsymbol{\alpha_s}$ → shower
   - Shower cutoff at $\boldsymbol{p_{T,min}}$
     - phase transition, confinement



2. Hadronisation (Lund-string model):
   - $f(z) \propto \frac{1}{z}(1-z)^{\boldsymbol{a_{\mathrm{Lund}}}} \exp\left(-\boldsymbol{b_{\mathrm{Lund}}} \frac{m_T^2}{z}\right)$
   - Baryons: $a = a_{Lund} + \boldsymbol{a_{\mathrm{ExtraDiquark}}}$
   - $P(p_T) \propto \exp(-\frac{1}{2}\frac{p_T^2}{\boldsymbol{\sigma}^2})$

3. Hadron decays from PDG tables

➢ overall 6 parameters to tune using the Professor framework

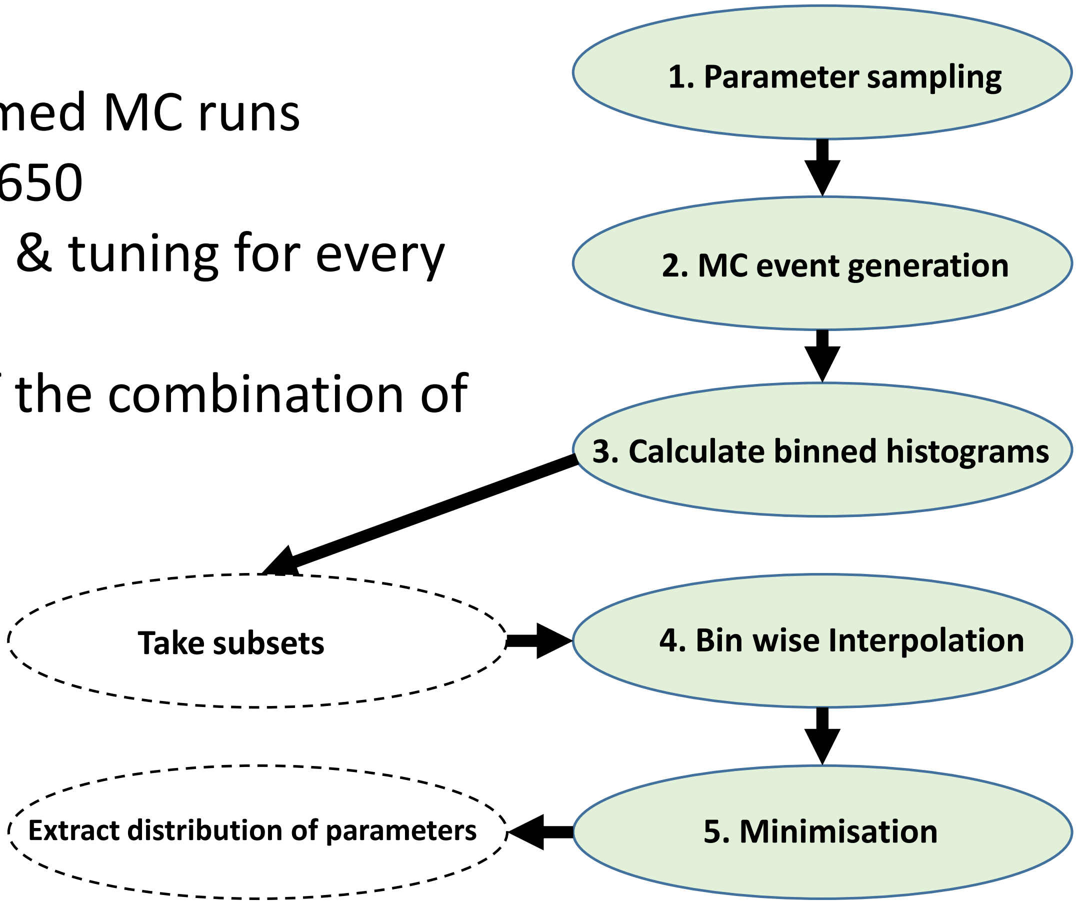
How to tune? – bootstrapping

Known from previous studies:

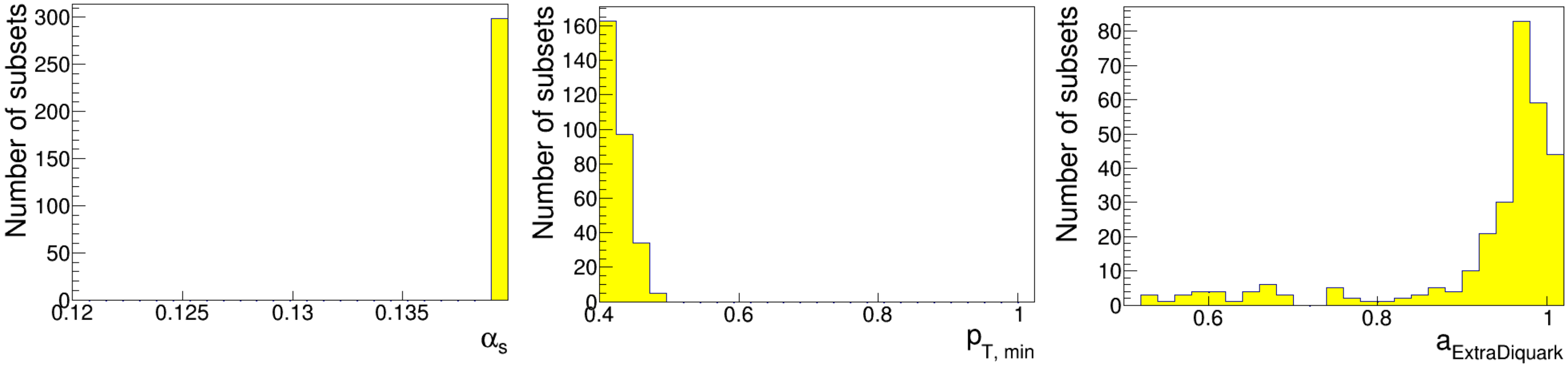
a) Take subsets of all performed MC runs
   - 300 times 500 out of 650

b) Perform the interpolation & tuning for every subset independently

c) Extract the distribution of the combination of tuned parameter values

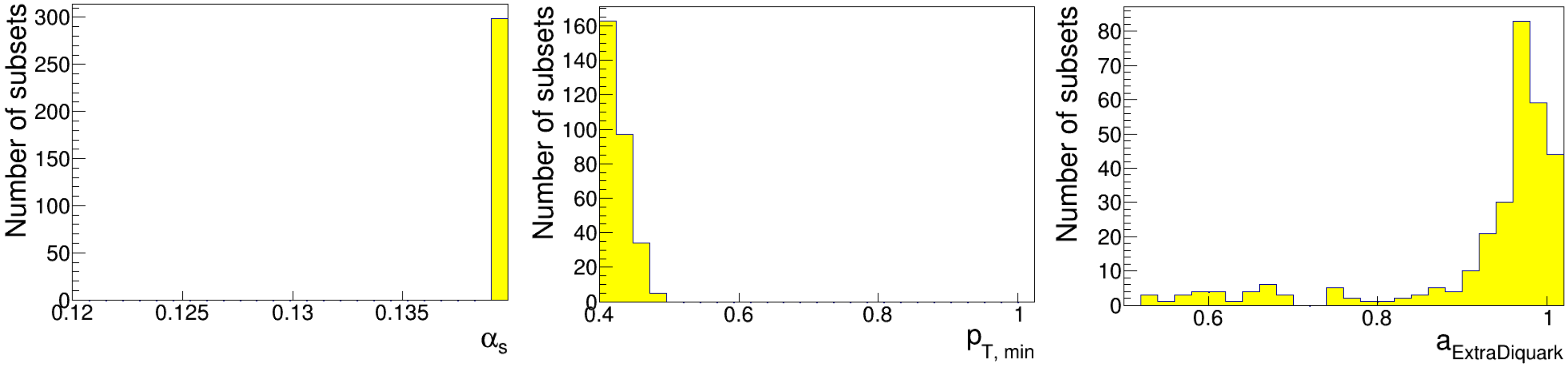


Distribution of tuned parameters with fixed ranges



- Parameter values at fit-range limit
- A better tune could lay outside the parameter ranges

Distribution of tuned parameters with fixed ranges

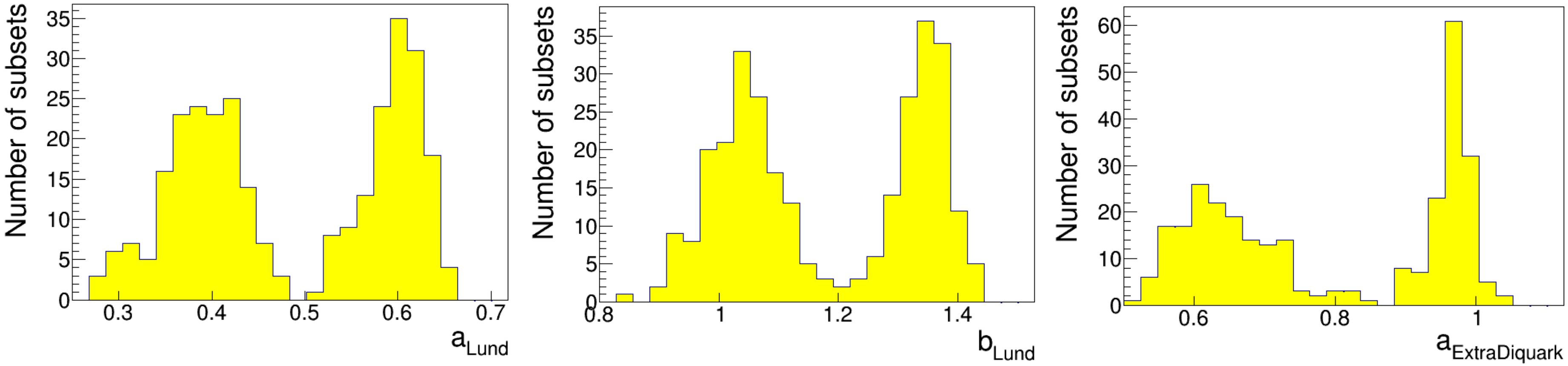


- Parameter values at fit-range limit
- A better tune could lay outside the parameter ranges

➢ Need to extend parameter range

➢ Keep interpolation functions unchanged

Distribution of tuned parameters without fixed ranges



- Extend parameter ranges, but no new MC samples
  - extrapolation in some parameters occurs
- Double peaks should not exist
  - result is sensitive to input set
- Is there a problem while minimising?
  - re-minimise by using another method

Overview

1. Introduction to parameter tuning
2. Reproduction of a previous tuning
3. Introducing new approaches for parameter tuning
4. Summary and outlook

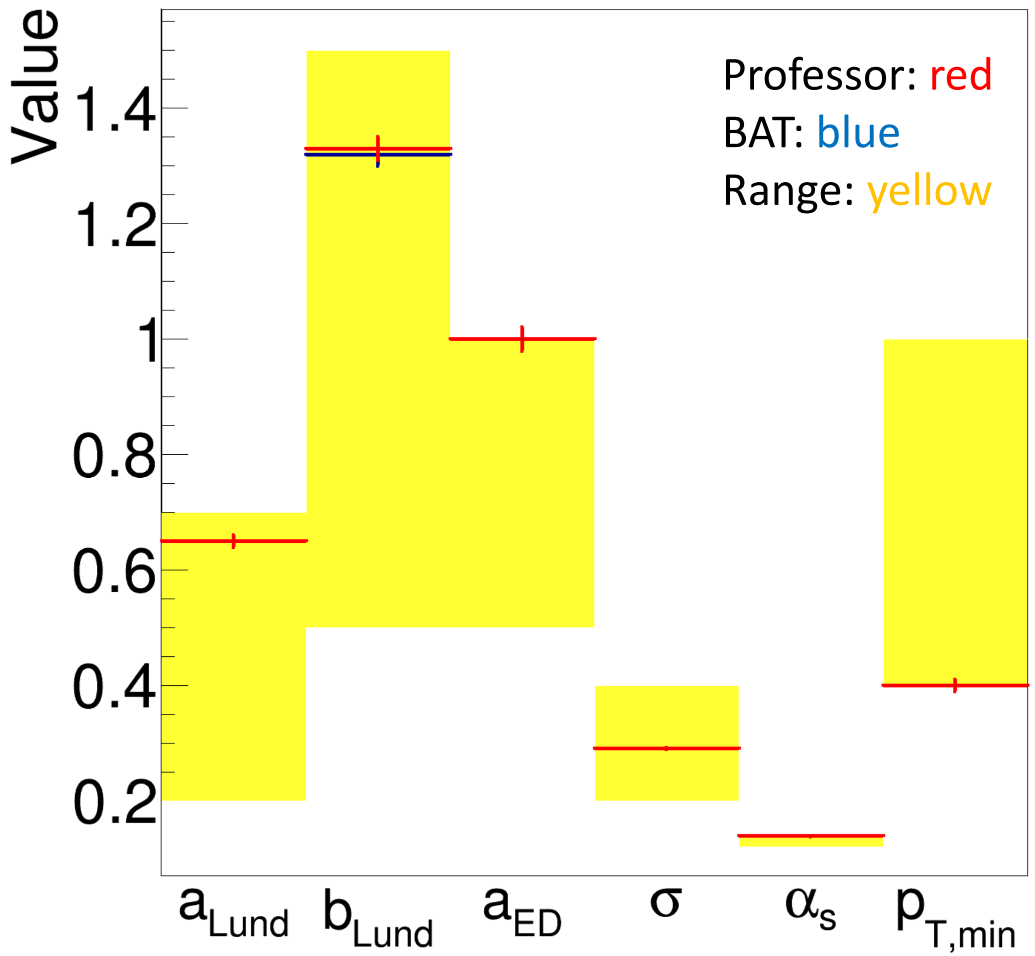
Bayesian Analysis Toolkit (BAT)

- Idea: Use BAT as control tune for Professor

- Working principle:
  - Based on self adapting Markov Chains
  - Steps determined by Metropolis-Hastings algorithm
  - Multiple Markov Chains should converge to same result

- Benefits of the algorithm:
  - Metropolis-Hastings algorithm reproduces a function
  - BAT collects information about the posterior likelihood
  - The algorithm can find the maximum posterior likelihood and thus optimal parameters

BAT - results



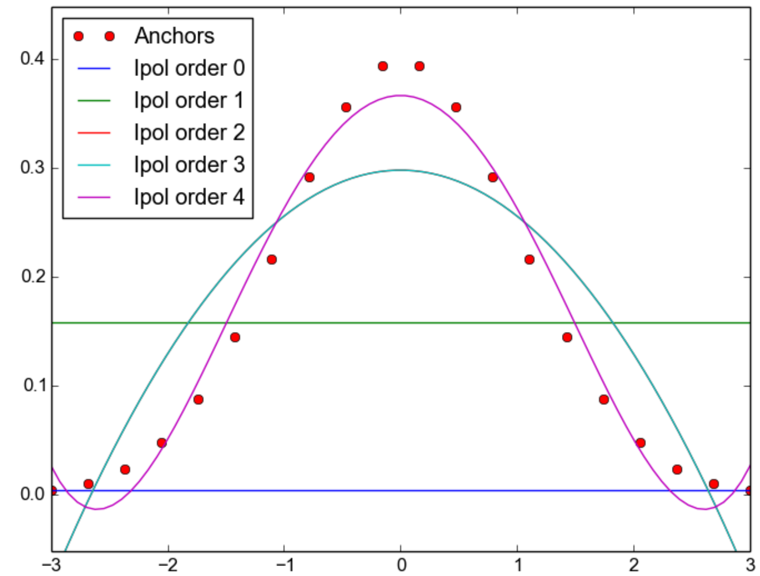
➢ The problem is not the minimisation!

Interpolation

- The problem could be produced while interpolating

a) Professor uses a fixed order polynom function for interpolation
   - possible over-/underfitting?

Simplified example of underfitting:



Underlying function: Gauss distribution

Interpolation

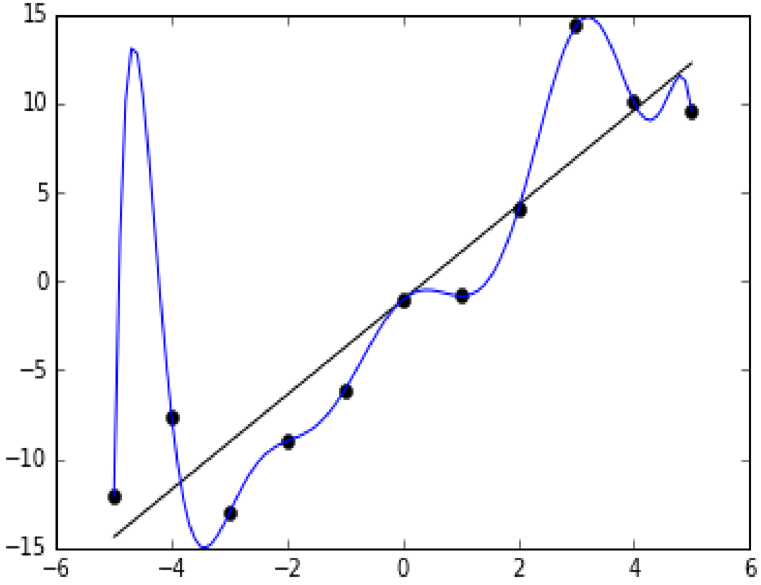
- The problem could be produced while interpolating

a) Professor uses a fixed order polynomial function for interpolation
   - possible over-/underfitting?

b) The quality of the fit should not be judged upon χ^2 only
   - Small $\chi^2 \nrightarrow$ good fit

➢ another approach is needed to avoid this

Simplified example of overfitting:



Source: https://en.wikipedia.org/wiki/Overfitting#/media/File:Overfitted_Data.png

New Interpolation

a) Avoid under-/overfitting

- Iteratively increase the number of terms in polynomial
- until interpolation „is good“

Construction based on:
IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, 561 (2010)

New Interpolation

a) Avoid under-/overfitting

- Iteratively increase the number of terms in polynomial
- until interpolation „is good“

Construction based on:
IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, 561 (2010)

b) Improved quality criterion:

- Introduce shape dependent parameter:

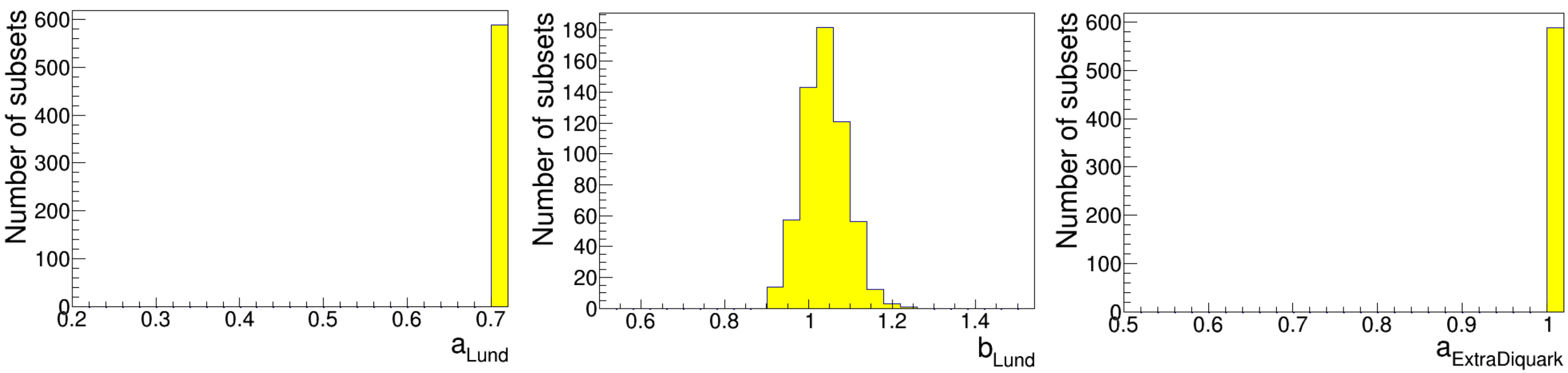
$$D_{Smooth} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{n}_{\text{Simulation data,i}} \cdot \boldsymbol{n}_{\text{Interpolation,i}}$$

Normalised gradient of the simulated data — Normalised gradient of the interpolation

➢ Minimise new criterion: $f = \chi^2 \frac{(1 - D_{\text{Smooth}})}{(1 + D_{\text{Smooth}})}$

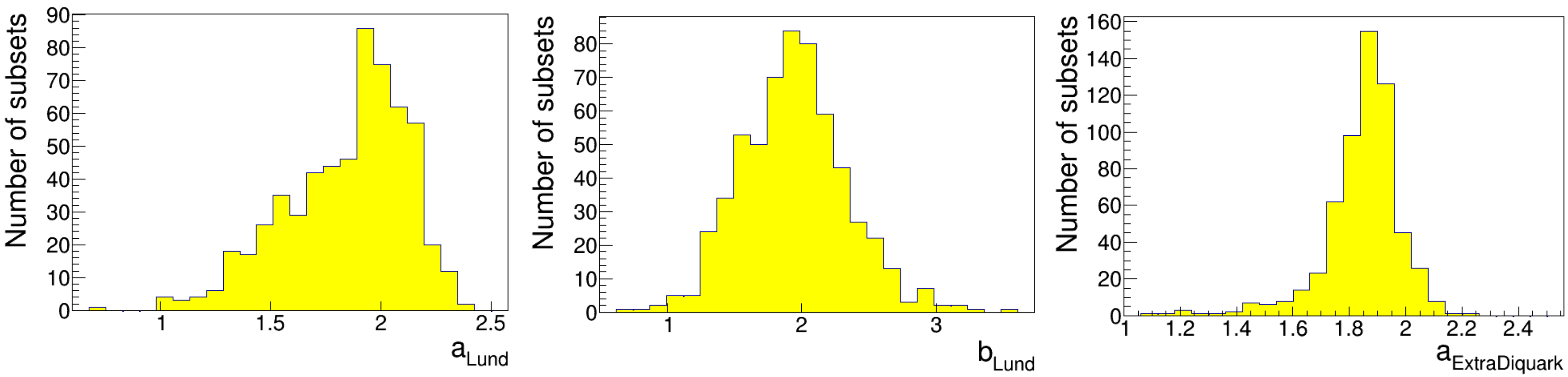
Distribution of tuned parameters

- Fixed ranges:
  - Parameter values at the limit
  - ➢ Allowing extrapolation



- Fixed ranges:
  - Parameter values at the limit
  - ➢ Allowing extrapolation

- Without fixed ranges:
  - Parameter values distributed around a single centre
  - ➢ Final values can be extracted



Error propagation of interpolation

- Until now: $\begin{pmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{pmatrix}$

➢ Consider the correlation between the fit parameters

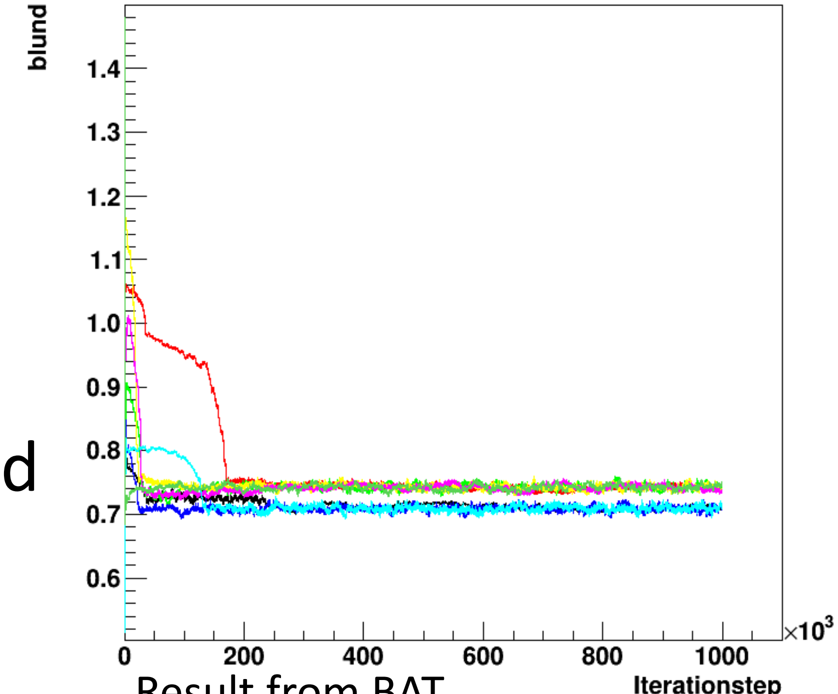
Error propagation of interpolation

- Until now: $\begin{pmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{pmatrix}$

➢Consider the correlation between the fit parameters

- But: Complicated likelihood & multiple attractors appear

➢further investigations are needed



Result from BAT

Overview

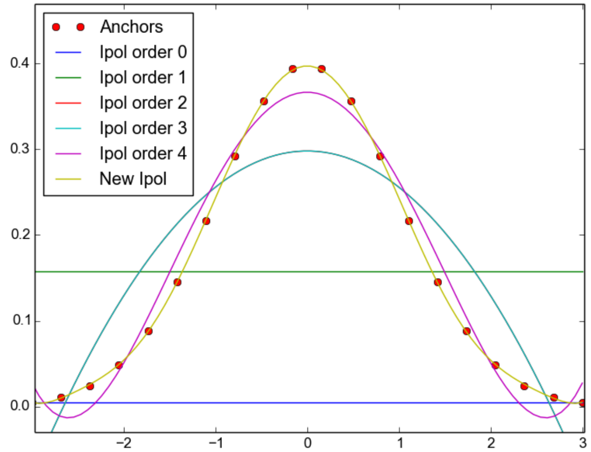
1. Introduction to parameter tuning
2. Reproduction of a previous tuning
3. Introducing new approaches for parameter tuning
4. Summary and outlook

Summary

- Professor tuning maybe unstable
- Problem caused by the interpolation algorithm

Summary

- Professor tuning maybe unstable
- Problem caused by the interpolation algorithm
- Using new approach to increase stability of interpolation with
  1. An iterative construction that increases the number of terms
  2. A new quality criterion f = $\chi^2 \frac{(1-D_{\text{Smooth}})}{(1+D_{\text{Smooth}})}$



- Using covariances of fit parameters is statistically consistent but causes further problems

Outlook

- Work on uncertainty calculation:
  - Investigation of the source the problems
    (e.g. parameter ranges, selected observables or weights)

Outlook

- Work on uncertainty calculation:
  - Investigation of the source the problems
    (e.g. parameter ranges, selected observables or weights)

- Outlier rejection if:
  - no parameter set describes the data well
  - precision of MC prediction is limited by underlying theory

Outlook

- Work on uncertainty calculation:
  - Investigation of the source the problems
    (e.g. parameter ranges, selected observables or weights)
- Outlier rejection if:
  - no parameter set describes the data well
  - precision of MC prediction is limited by underlying theory
- Redefine parameter ranges to avoid running into limit and repeat tune (work in progress)

Outlook

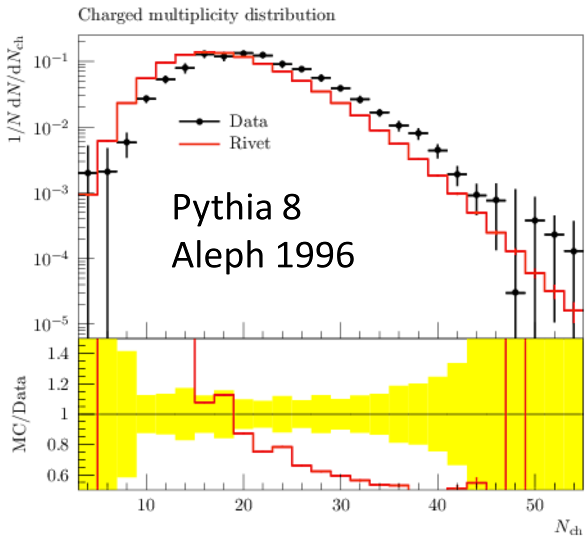
- Work on uncertainty calculation:
  - Investigation of the source the problems (e.g. parameter ranges, selected observables or weights)
- Outlier rejection if:
  - no parameter set describes the data well
  - precision of MC prediction is limited by underlying theory
- Redefine parameter ranges to avoid running into limit and repeat tune (work in progress)
- Implementation of the new interpolation algorithm into the Professor framework

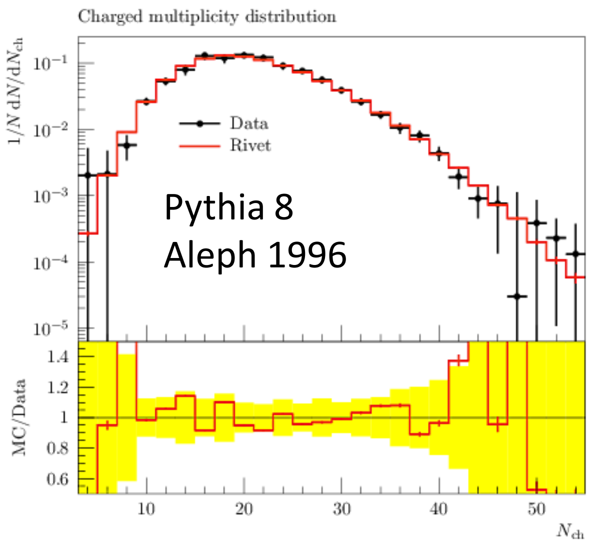
Thank you for your attention

Backup

Backup: Parameter sets of example distributions

Parameter	Left Set	Right Set
a_{Lund}	0.42	0.56
b_{Lund} [GeV^{-2}]	0.82	1.00
$a_{\mathrm{ExtraDiquark}}$	0.99	0.93
σ [GeV]	0.259	0.399
α_S	0.120	0.126
$p_{\mathrm{T,min}}$ [GeV]	0.42	0.95





Backup: gof

Interpolation function

Sum over every observable and bin

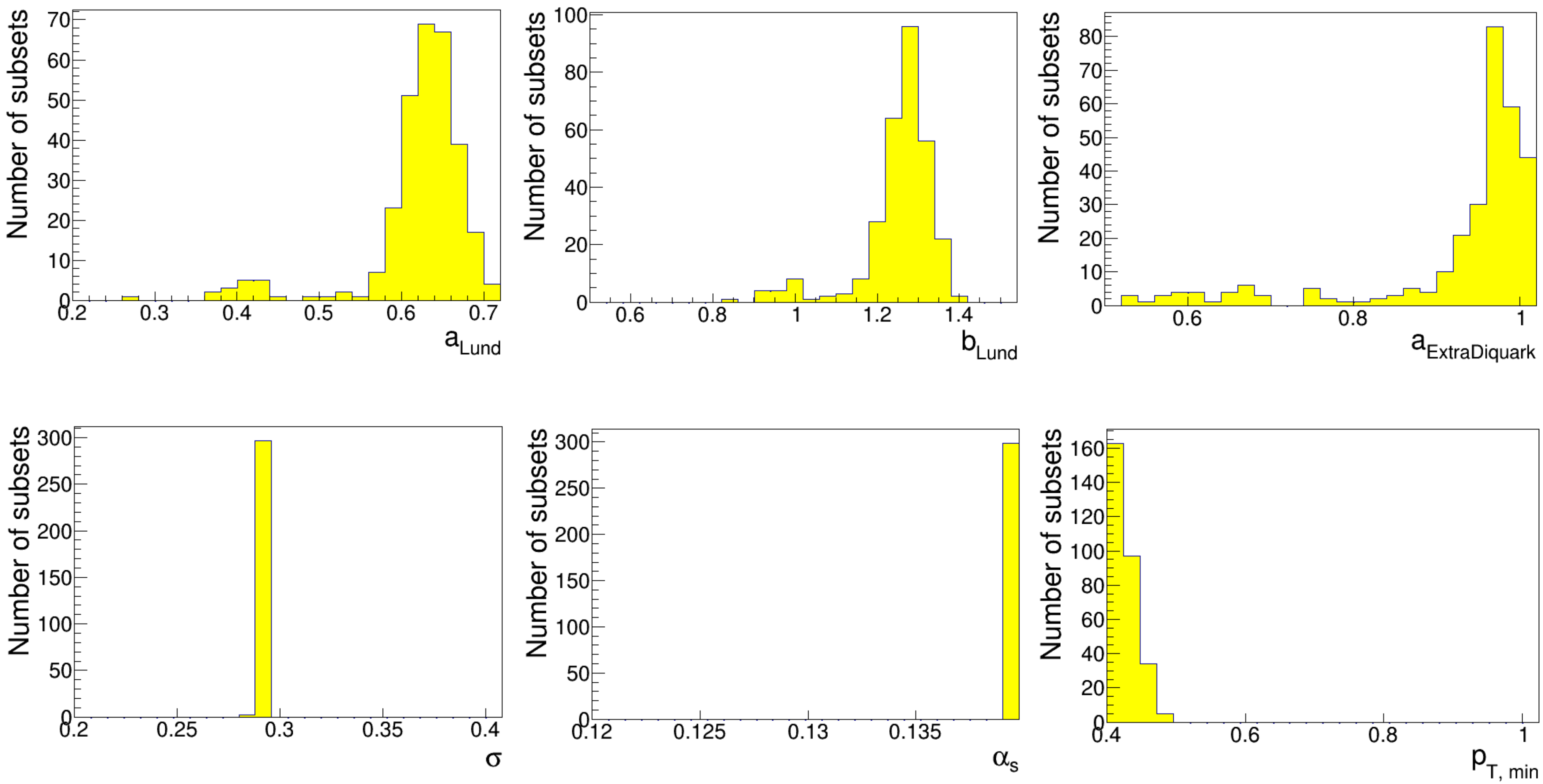
Measurement

$$\chi^2(\boldsymbol{p}) = \sum_{O} \sum_{b \in O} w_b \frac{\left(f^{(b)}(\boldsymbol{p}) - R_b\right)^2}{\Delta_b^2}$$

Weights

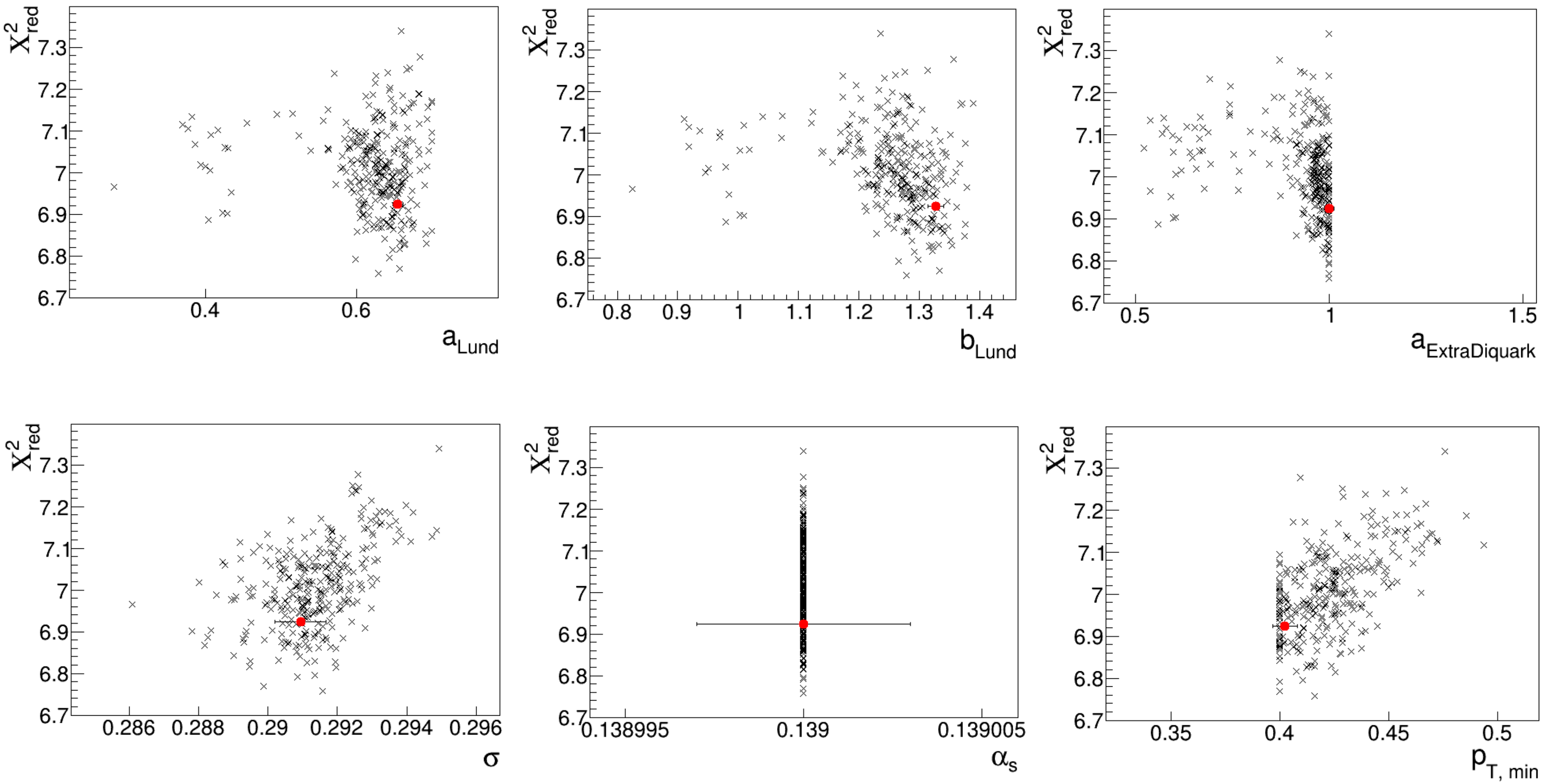
Sum of variances

Backup: Professor with parameter limits I



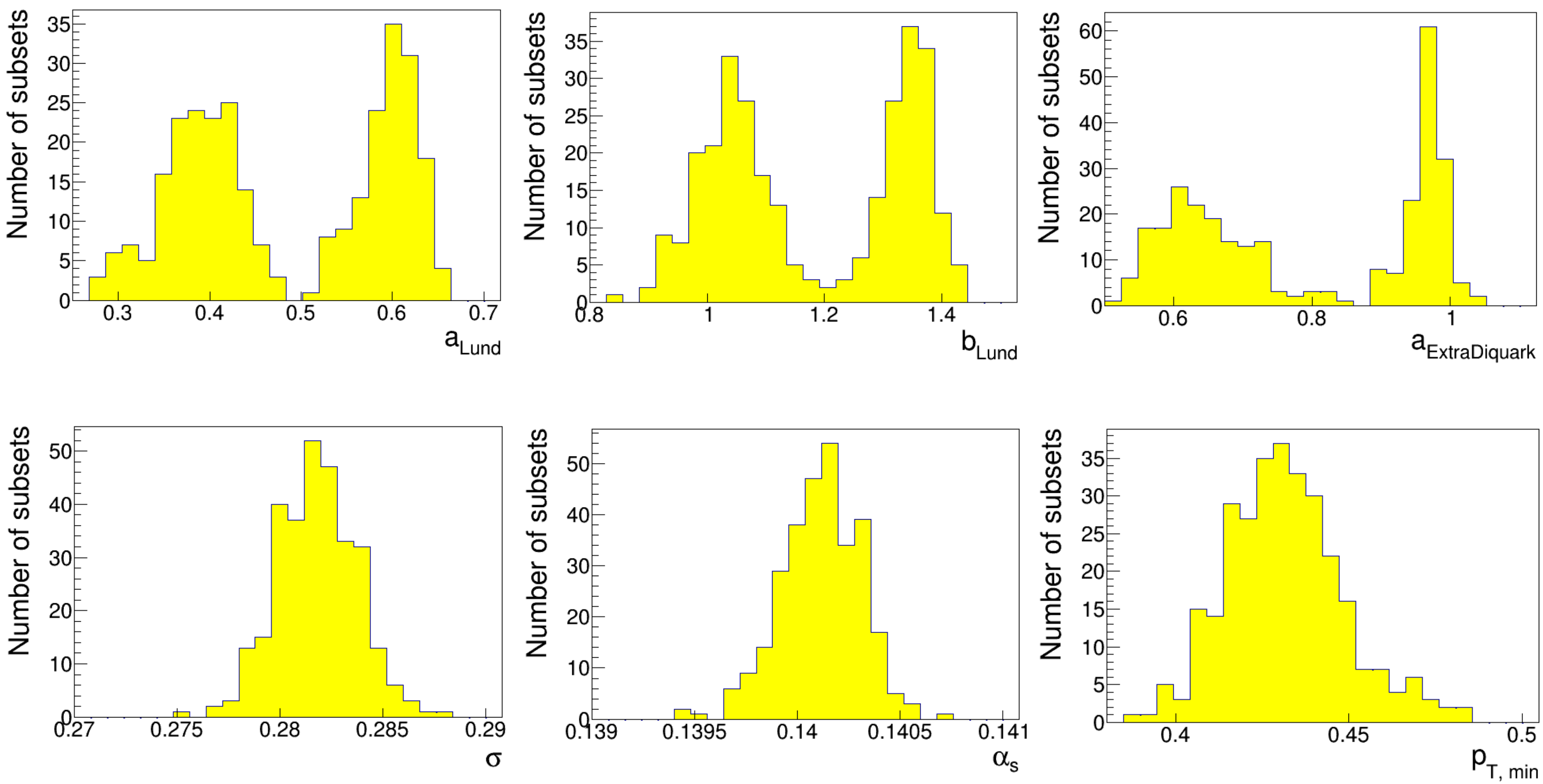
650 samples

Backup: Professor with parameter limits II



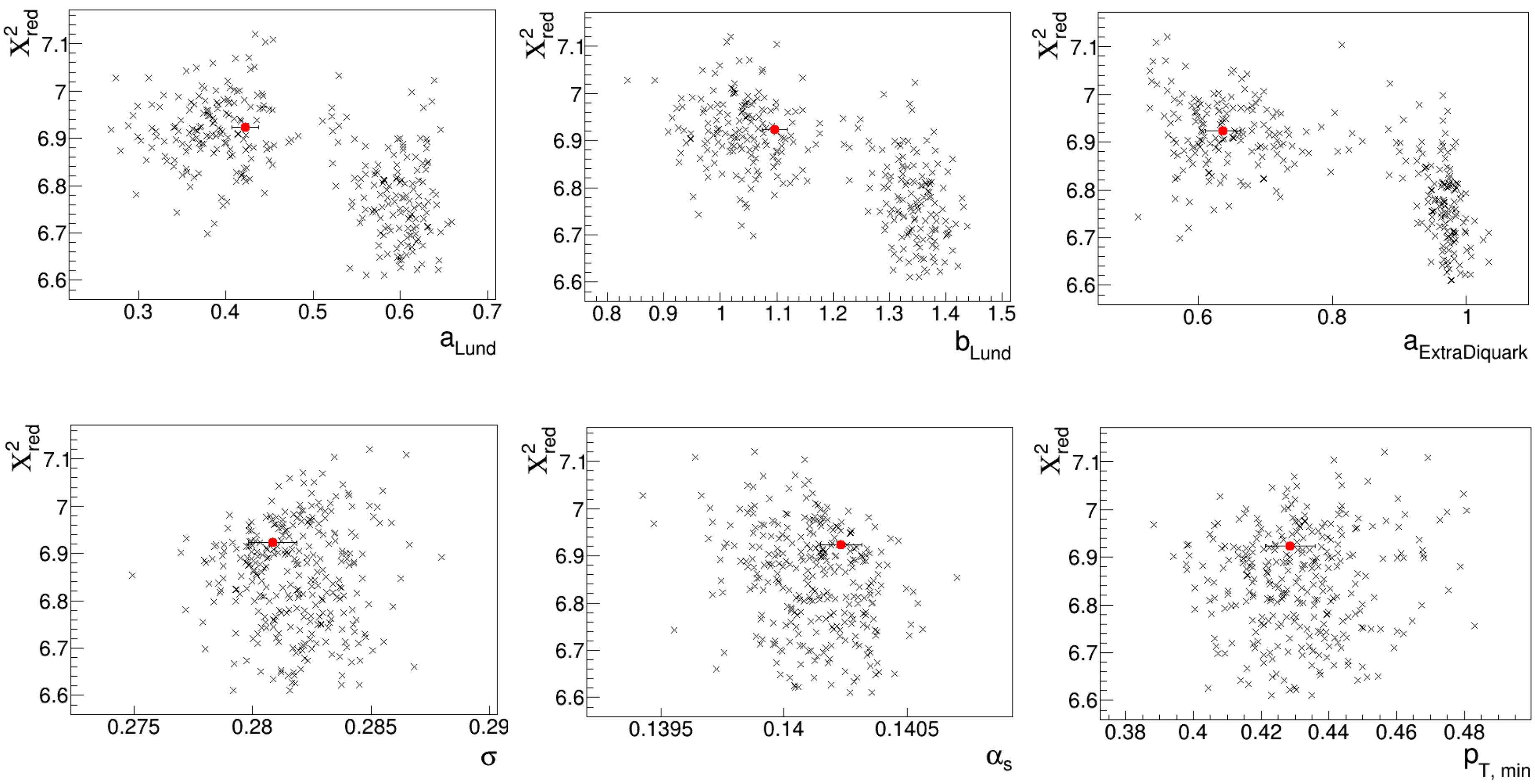
650 samples

Backup: Professor without parameter limits I



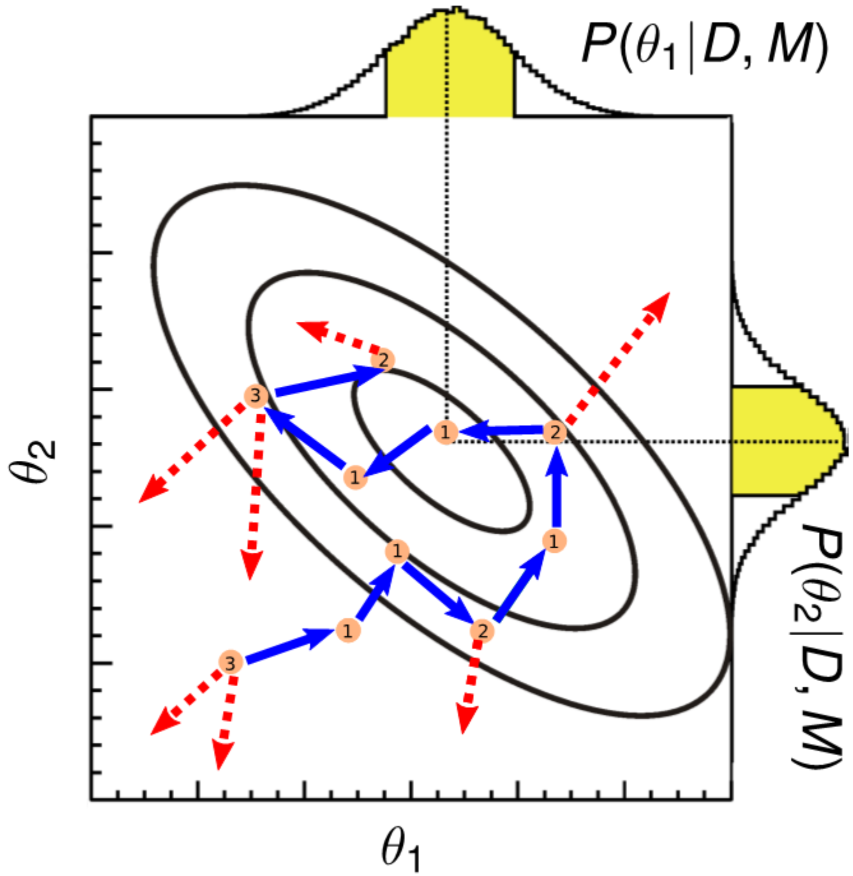
650 samples

Backup: Professor without parameter limits II

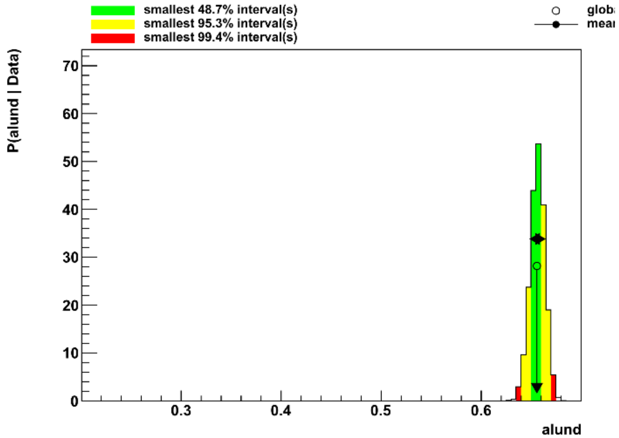


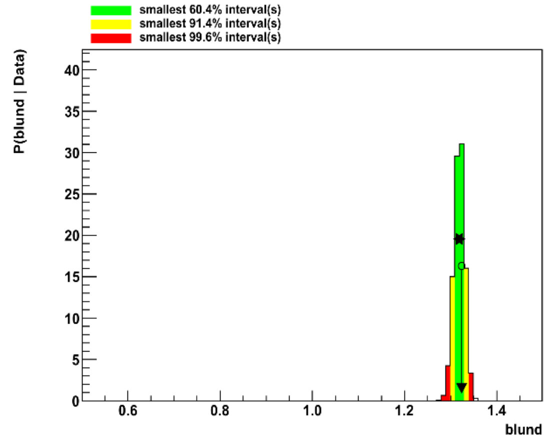
650 samples

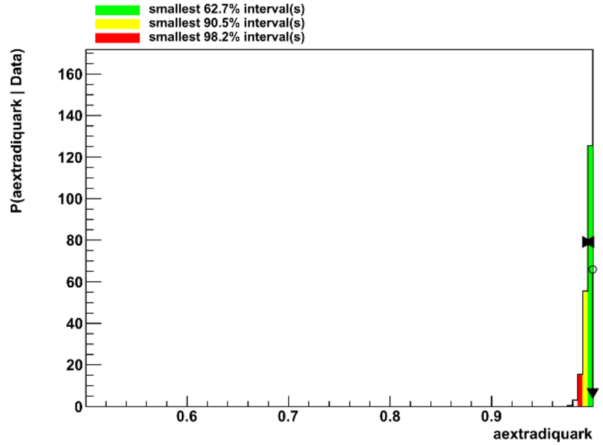
Backup: Markov Chains

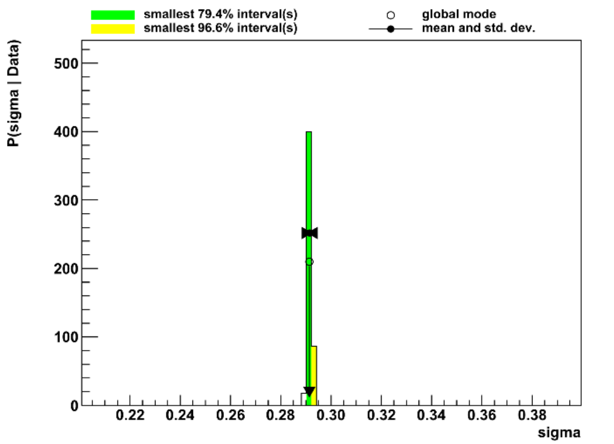


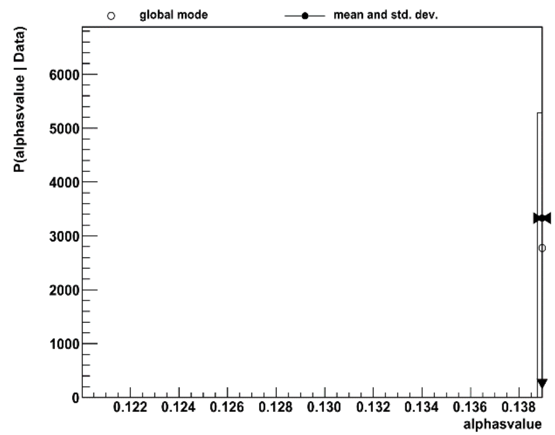
Backup: BAT results for Professor Ipol

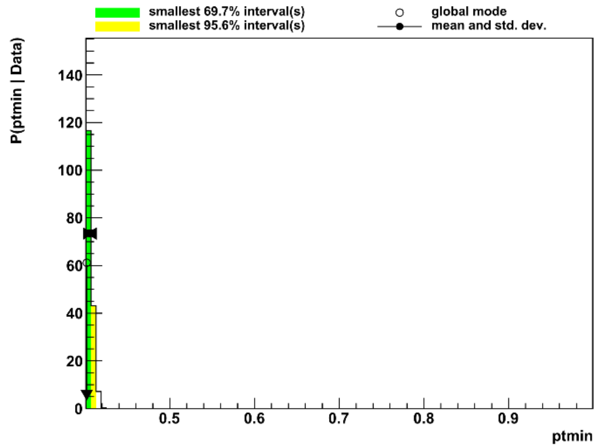












Backup: Comparison BAT <-> runcombs

Prof

Parameter	Value	Error
a_{Lund}	0.622	0.064
$b_{\mathrm{Lund}}[\mathrm{GeV}^{-2}]$	1.253	0.090
$a_{\mathrm{ExtraDiquark}}$	0.928	0.113
σ [GeV]	0.292	0.001
α_s	0.139	0.001
$p_{\mathrm{T,min}}$ [GeV]	0.426	0.019

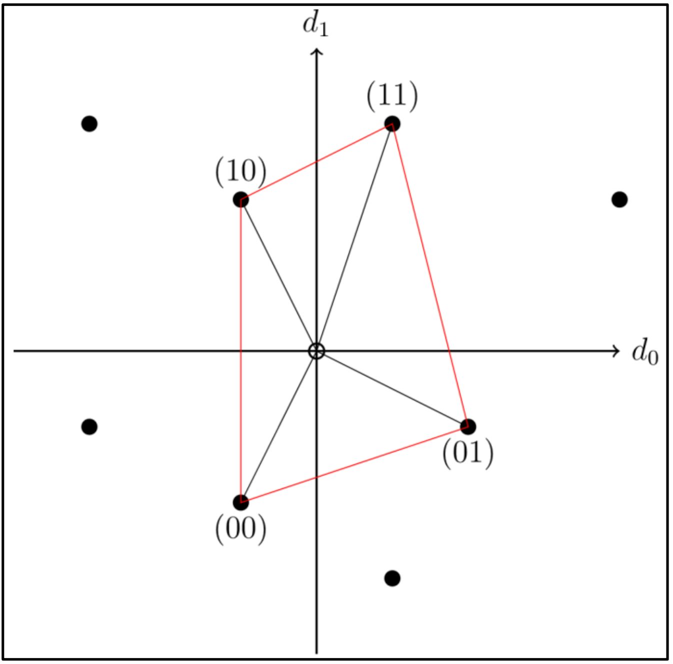
BAT

Parameter	Value	Error
a_{Lund}	0.655	0.008
$b_{\mathrm{Lund}}[\mathrm{GeV}^{-2}]$	1.324	0.014
$a_{\mathrm{ExtraDiquark}}$	1.000	0.008
σ [GeV]	0.291	0.001
α_s	0.139	0.001
$p_{\mathrm{T,min}}$ [GeV]	0.402	0.009

Uncertainty calculation unreliable because parameters at edges

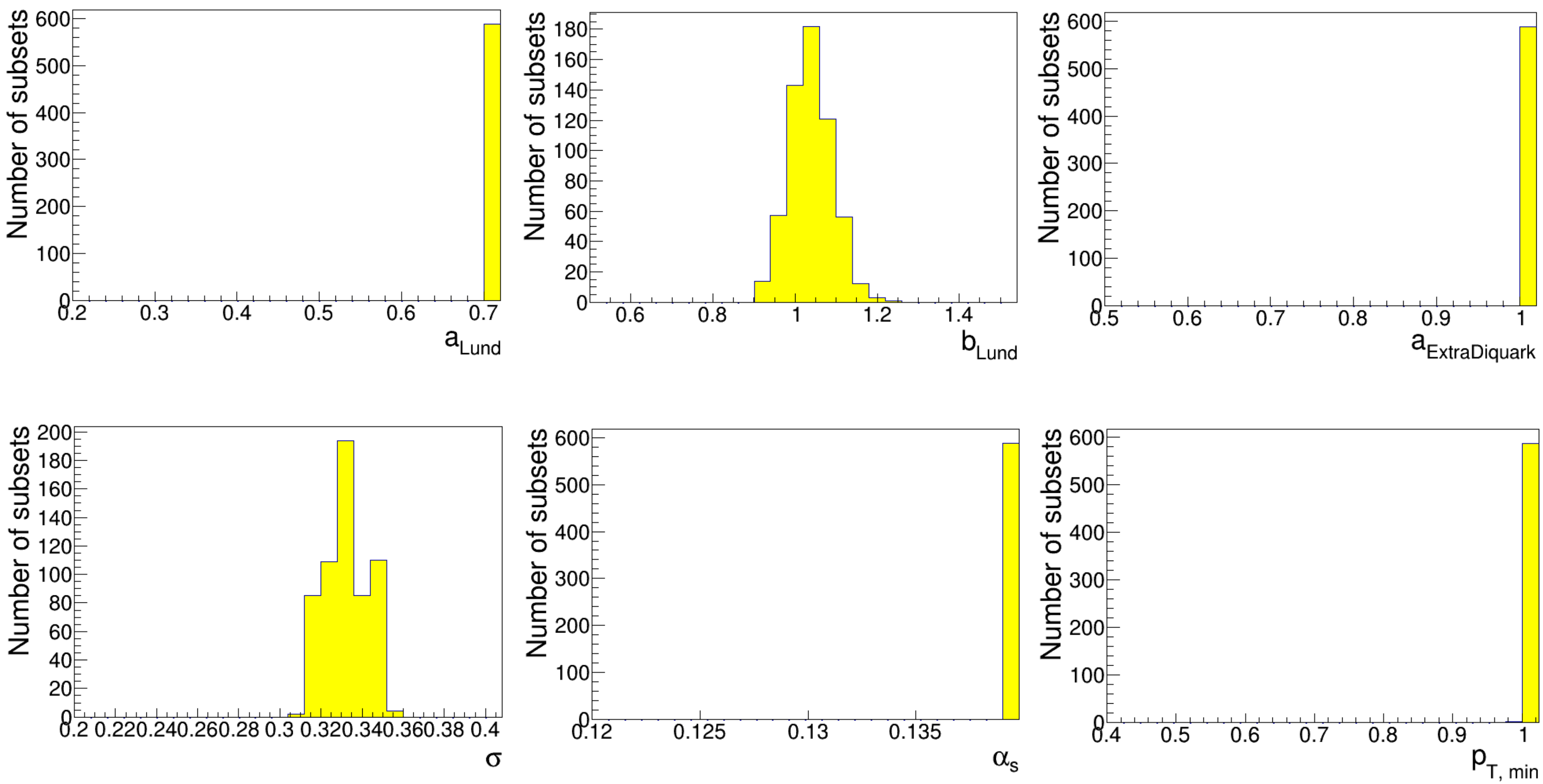
Backup: New Interpolation: Gradients

- $\boldsymbol{n}_{\text{Interpolation,i}} = \frac{1}{\|\nabla f_{\text{Interpolation}}\|} \nabla f_{\text{Interpolation}} \Big|_{\boldsymbol{x}_{\text{i}}}$
  - Polynomial function → analytical gradient
- $\boldsymbol{n}_{\text{Simulation data,i}} = \frac{1}{\|\nabla f_{\text{Simulation data}}\|} \nabla f_{\text{Simulation data}} \Big|_{\boldsymbol{x}_{\text{i}}}$



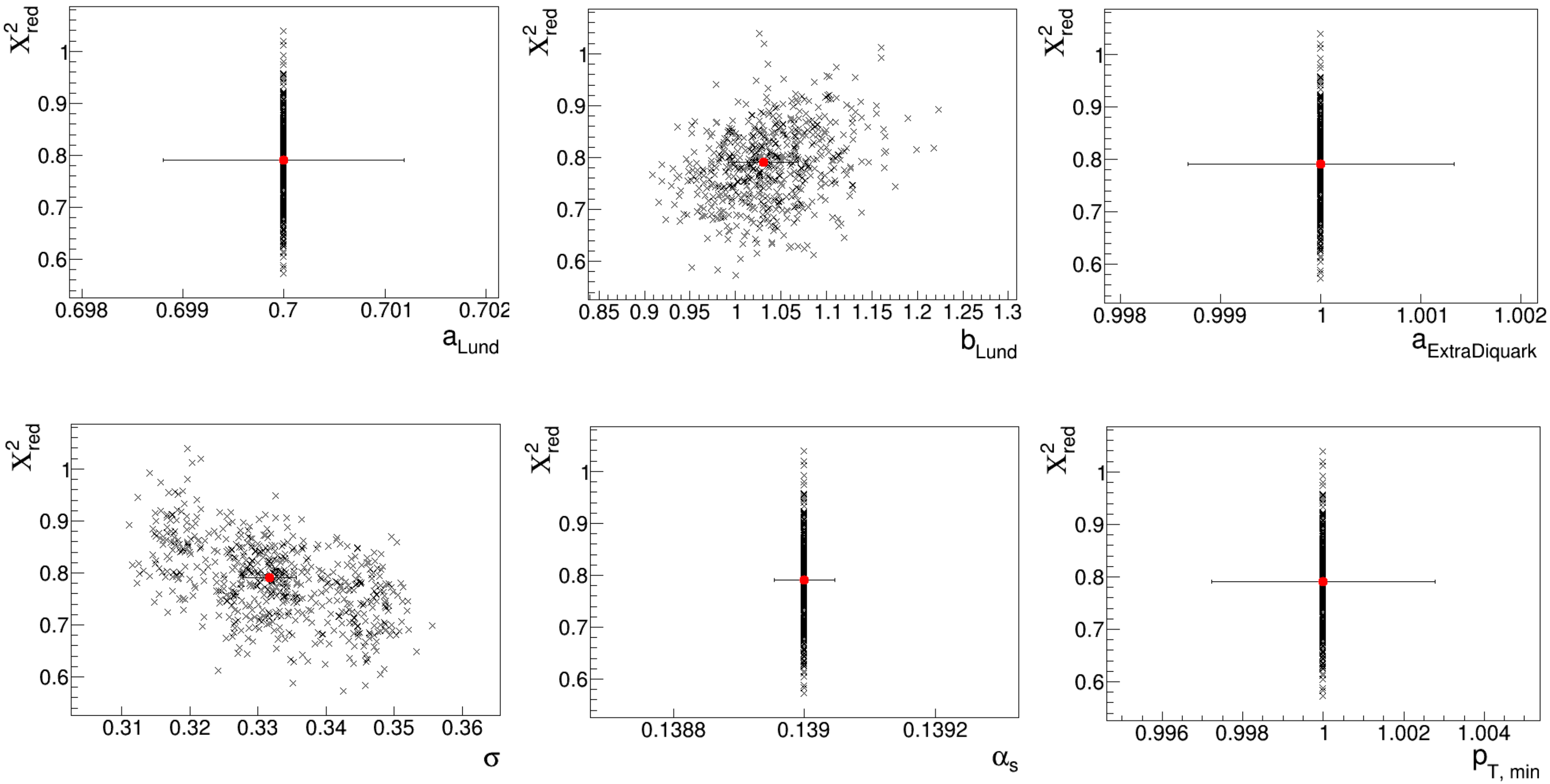
- <u>Approach:</u>
- Select $2^{\text{dimension}}$ closest points
- Re-weight using the distance to center
- Interpolate with first order polynomial
- Calculate gradient analytically

Backup: New Ipol with parameter limits I



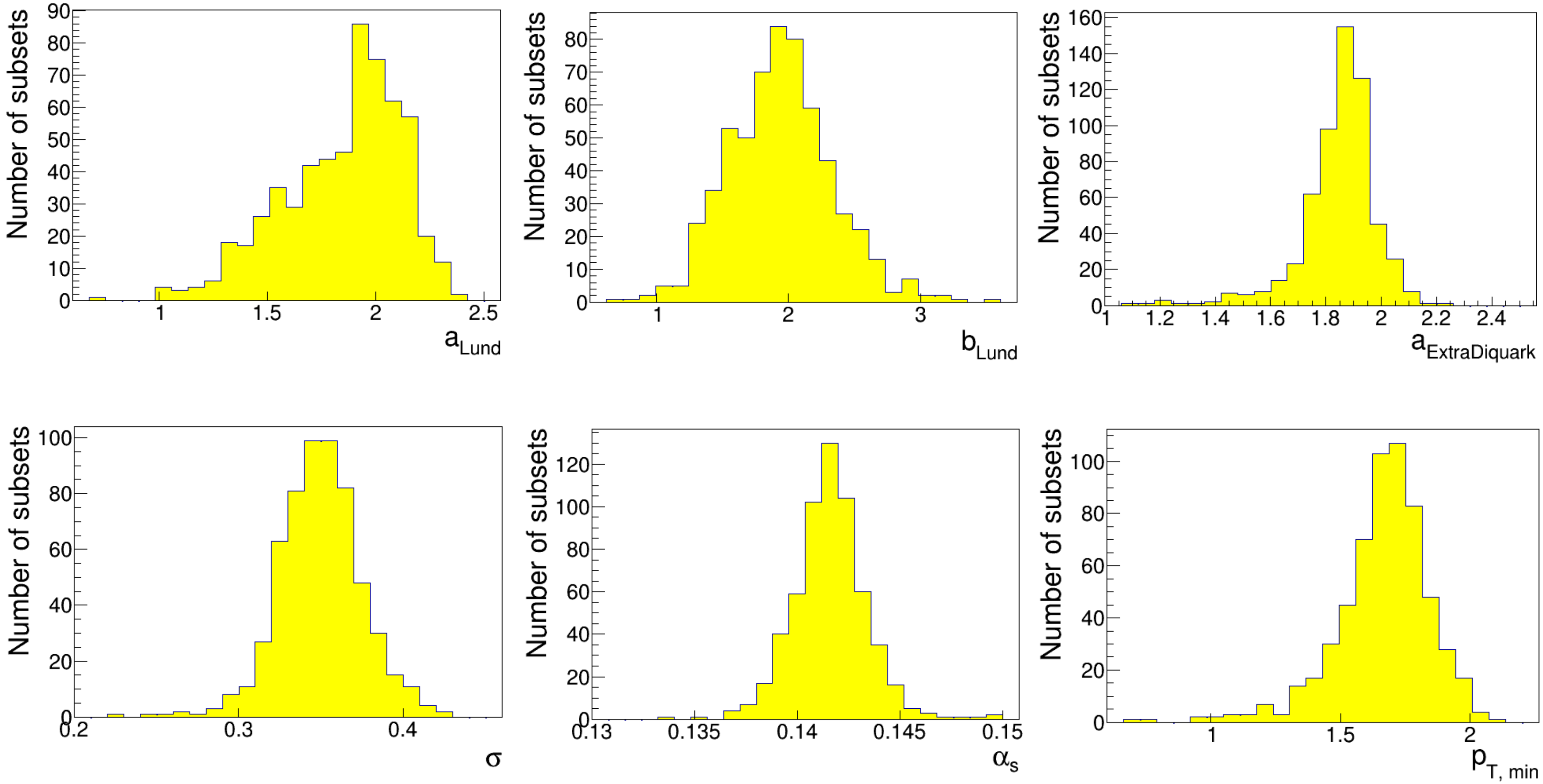
1500 samples

Backup: New Ipol with parameter limits II



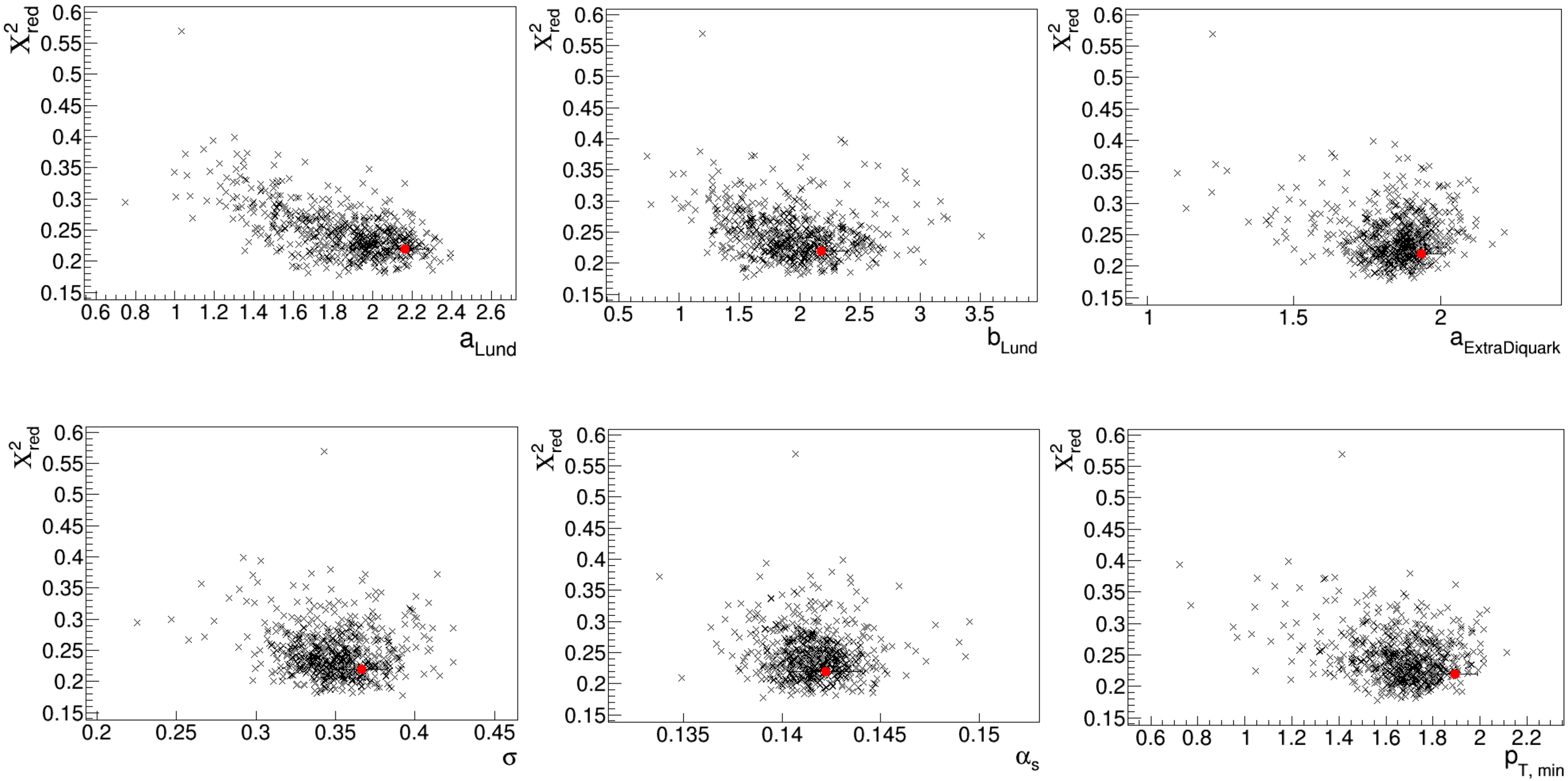
1500 samples

Backup: New Ipol without parameter limits I



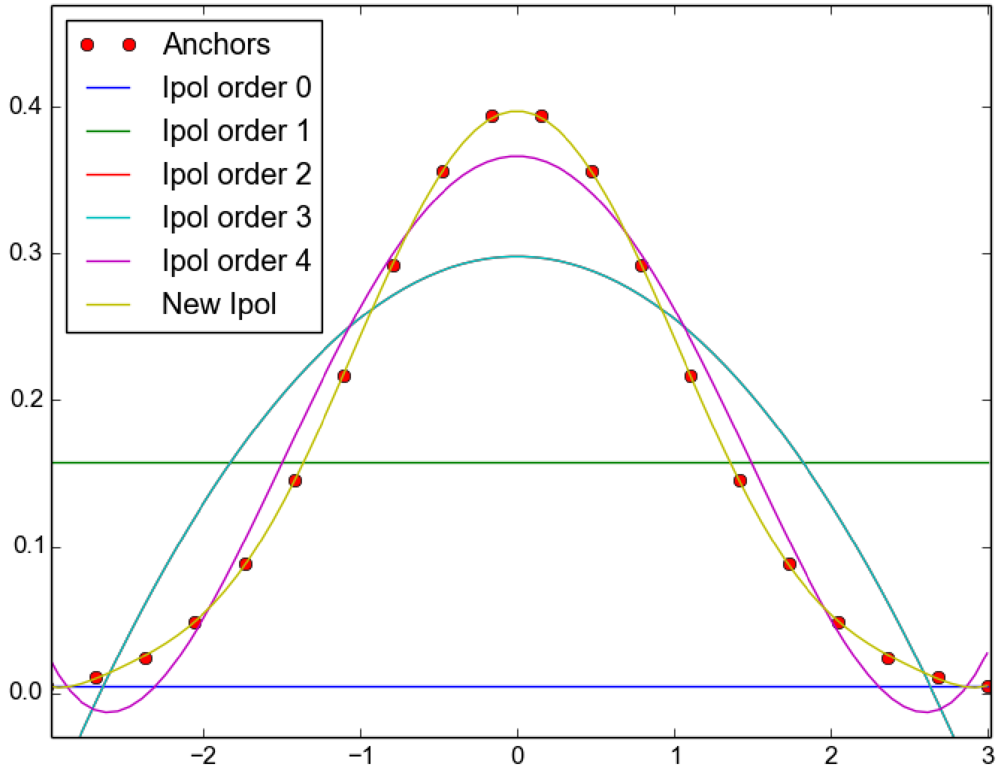
1500 samples

Backup: New Ipol without parameter limits II



1500 samples

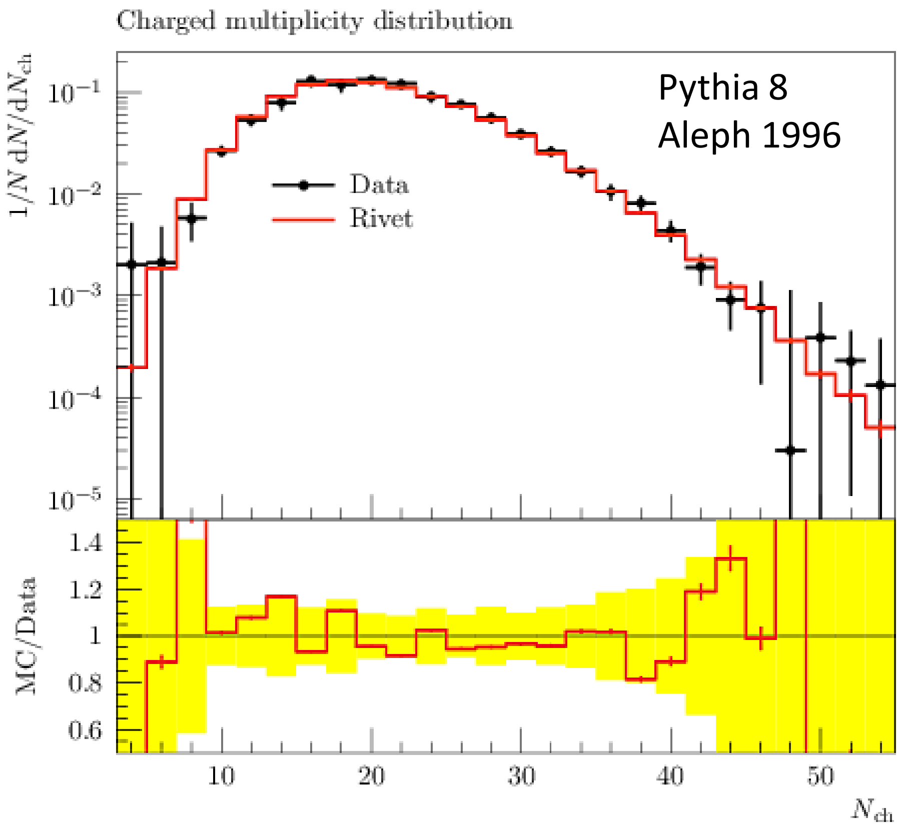
Backup: Fitting the normal distribution



Backup: Comparison by re-simulating

- Simulation with tuned parameter set with limits

Professor:



New interpolation: