Sprout

A software tool for your analysis

Malin Bohman

UPPSALA UNIVERSITET

FAIRNESS

Overview

Introduction

How Sprout came about

The Use Case of Sprout

How Sprout can help you in your analysis

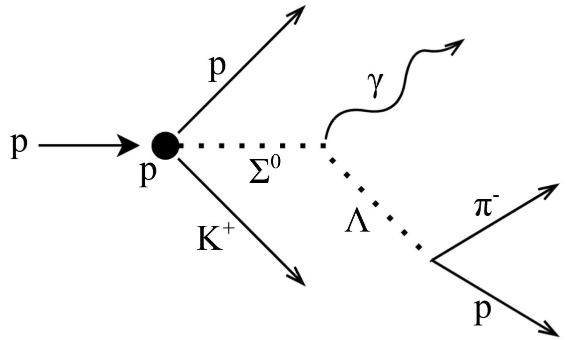
The Sprout Family

Overview of Sprout's features

UPPSALA UNIVERSITET

Introduction

- Sprout🌱: a little C++ project that grew from ROOT
- My background: Hyperon data analysis for Panda@HADES
- Data from proton-proton beam-time at HADES collected in 2022
- **Goal**: Measure *cross sections* for various Sigma0 production channels
- **Current focus**: Careful studies of Σ^0 *reconstruction efficiencies* for different channels

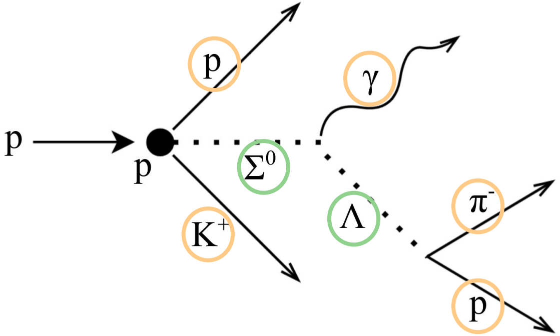


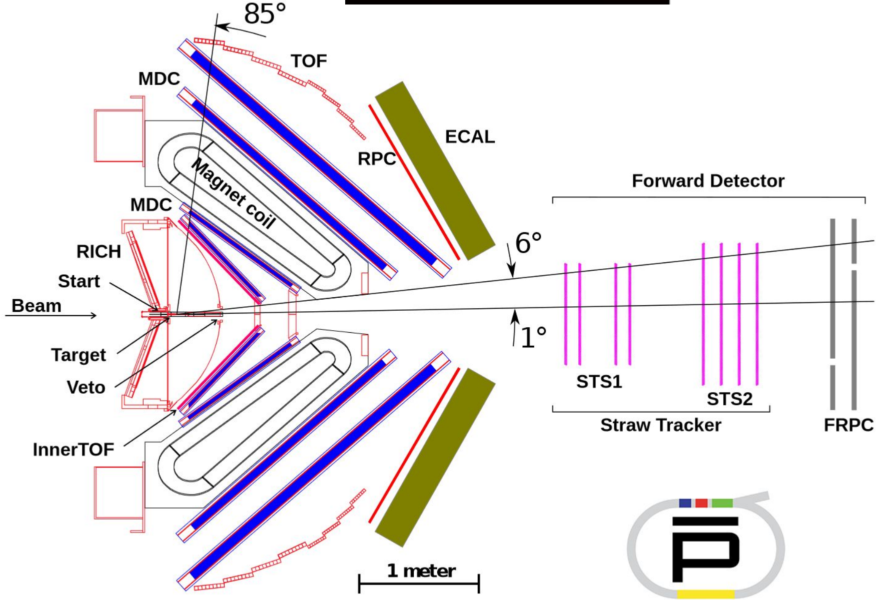
UPPSALA UNIVERSITET

Introduction

The data we're looking for:

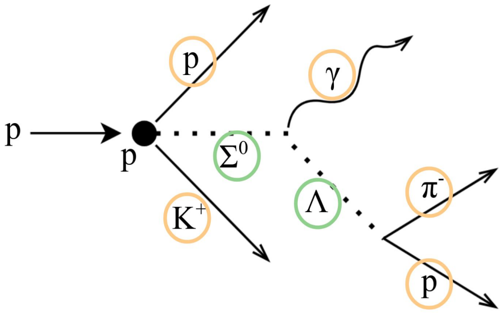
Reconstruct final-state particle four-vectors based on detector response to learn something about short-lived particles





UPPSALA UNIVERSITET

The Analysis Workflow



Event Reconstruction

Try to verify the presence of the particle of interest

In my case: assess the likelihood that the final state particles actually originated from a Σ^0 event

3 **Extract Measurement**

Perform a measurement on yield to extract physics of interest, e.g. cross section

1 **Preselection**

Sort through data to select events that are relevant to your analysis

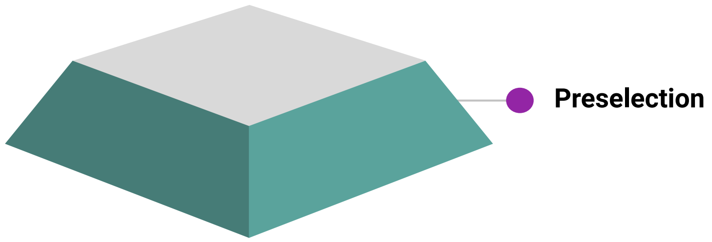
In my case: look for presence of final-state particles that could originate from a Σ^0 event

UPPSALA UNIVERSITET

Preselection

Look for presence of final-state particles that could originate from a Σ^0 event

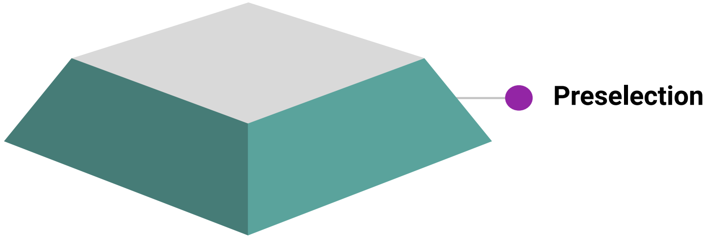
- Assess the data quality:
  - Various qualifiers for detector track quality
  - Use of different triggers
- Determine particle type from detector information
- Check for presence of relevant final-state particles



Preselection

Difficulties:

- Many different parameters and combinations to test and keep track of
- Often changing code to alter parameters
- Easy to lose track and prone to human error



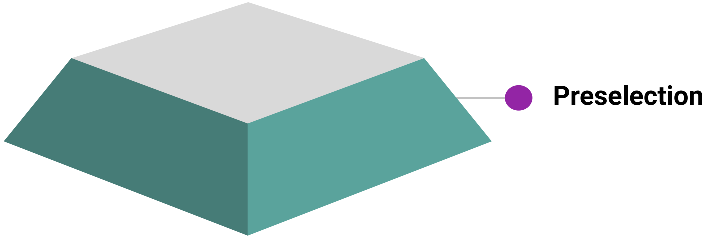
UPPSALA UNIVERSITET

Preselection

Difficulties:

- Many different parameters and combinations to test and keep track of
- Often changing code to alter parameters
- Easy to lose track and prone to human error

SproutParam
Offers a simple way to store and access parameters throughout your analysis

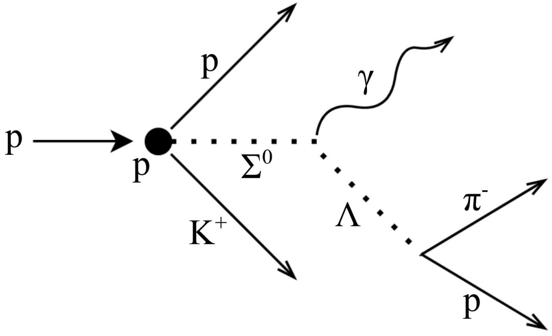


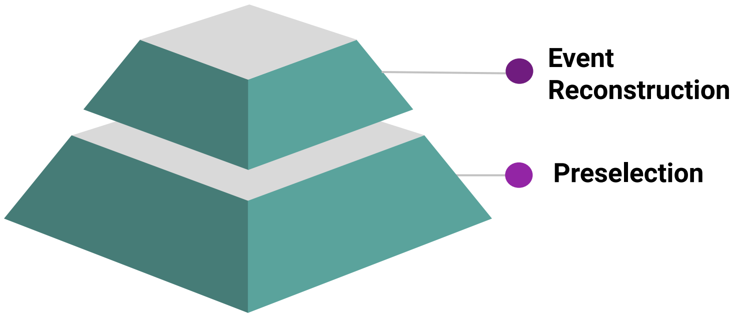
UPPSALA UNIVERSITET

Event Reconstruction

Assess the likelihood that the final state particles actually originated from a Σ^0 event

- Need to filter true Σ^0 events from *background*
  - Falsely identified particles
  - Particles from secondary reactions
  - Combinatorics
- Use *cuts* to decrease background



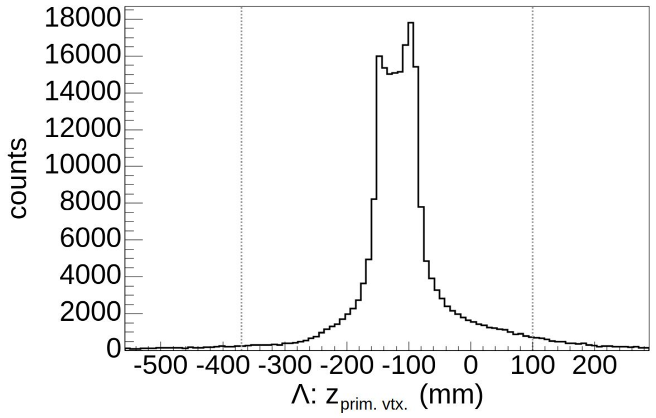


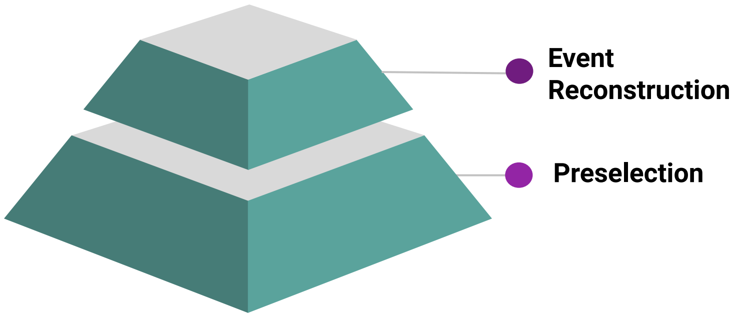
UPPSALA UNIVERSITET

Event Reconstruction

Assess the likelihood that the final state particles actually originated from a Σ^0 event

- Identify suitable cut variables to separate signal from background
- Use simulations to define a suitable cut value, e.g. by optimizing the signal yield
- Must ensure that simulations accurately describe the data in each cut variable



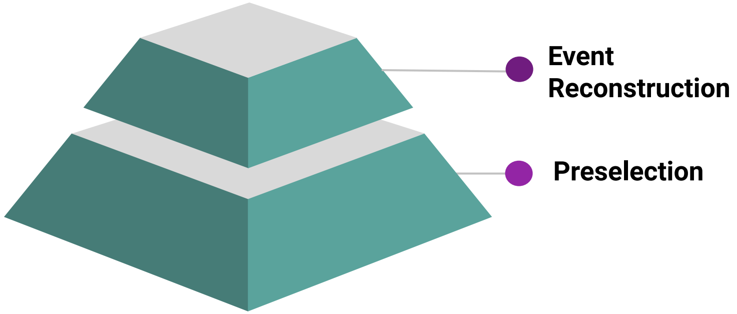


UPPSALA UNIVERSITET

Event Reconstruction

Difficulties:

- An analysis might contain *many* different cuts
  - Many variables to check for agreement between data and simulations
  - Each cut value tuned by trying out ~ 100 positions to optimize yield
  - A lot of quality assurance plots to generate and compare



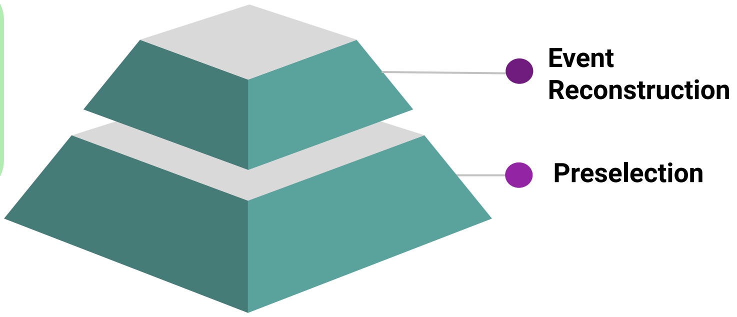
UPPSALA UNIVERSITET

Event Reconstruction

Difficulties:

- An analysis might contain *many* different cuts
  - Many variables to check for agreement between data and simulations
  - Each cut value tuned by trying out ~ 100 positions to optimize yield
  - A lot of quality assurance plots to generate and compare

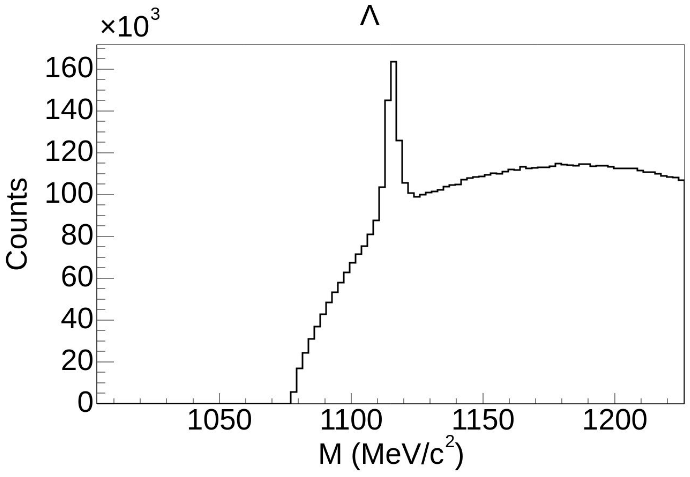
SproutCut
Automatically generates plots for cut quality assurance
Feature to come: perform bulk of the effort of cut tuning

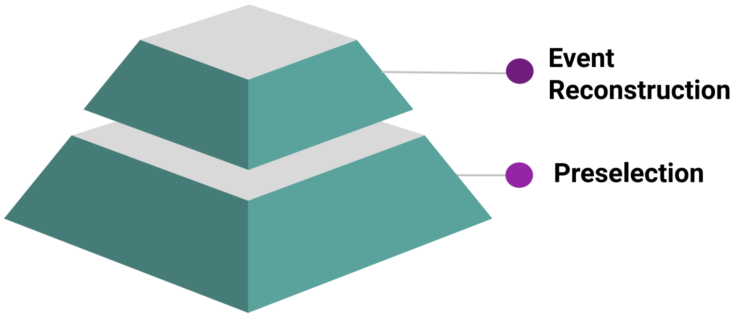


UPPSALA UNIVERSITET

Event Reconstruction

- Even after cuts, some background usually remains
- Easily seen from invariant mass histograms
- Commonly use fits to subtract background contributions from the signal yield



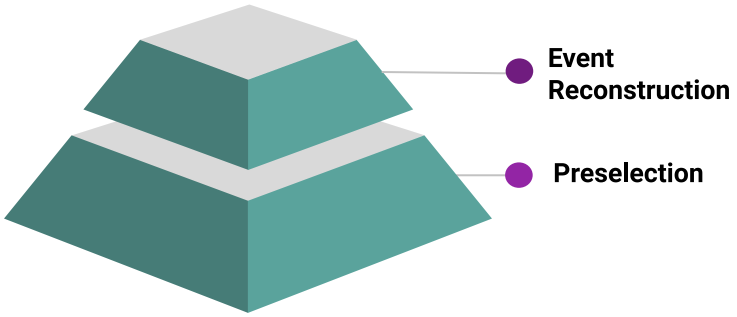


UPPSALA UNIVERSITET

Event Reconstruction

Difficulties:

- Each fit needs visual quality assurance. Does it fit the data accurately?
- Especially background distributions may differ over different kinematic regions, no one-model fits all
- Each fit may require different start parameters to not fail
- Some analyses may require ~ 100 fits



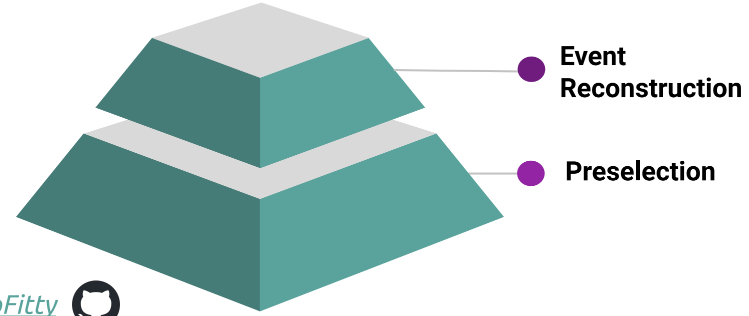
UPPSALA UNIVERSITET

Event Reconstruction

Difficulties:

- Each fit needs visual quality assurance. Does it fit the data accurately?
- Especially background distributions may differ over different kinematic regions, no one-model fits all
- Each fit may require different start parameters to not fail
- Some analyses may require ~ 100 fits

SproutFit
Minimizes your effort by storing fit parameters in a separate file, automatically finds start parameters and generates plots for quality assurance



UPPSALA UNIVERSITET

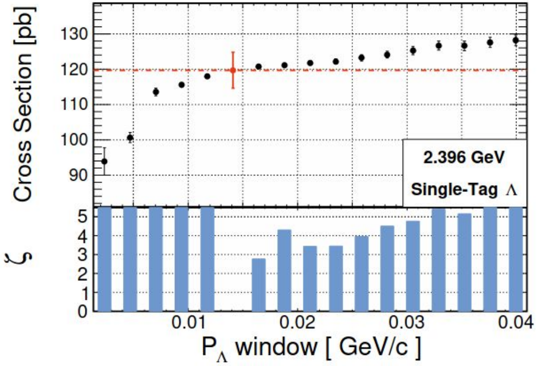
Inspired by Rafał Lalik's (Jagiellonian University) software package HelloFitty

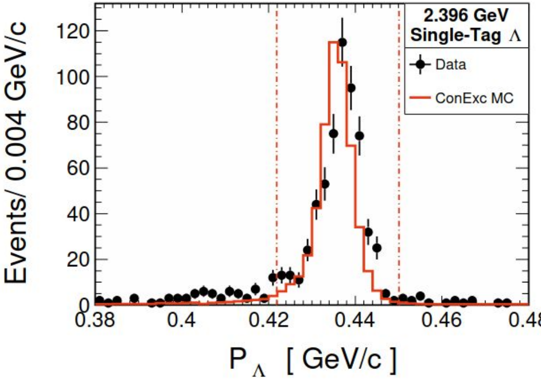
Extract Measurement

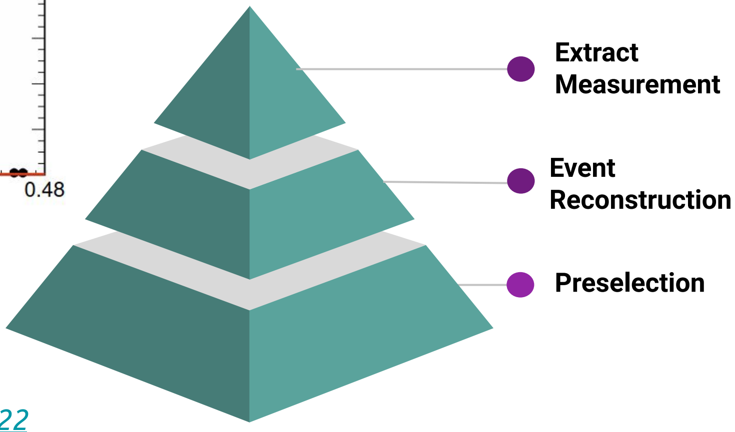
Perform a measurement on yield to extract physics of interest

Caveat: selection cuts may impact the measurement.

Potential source of systematic error!





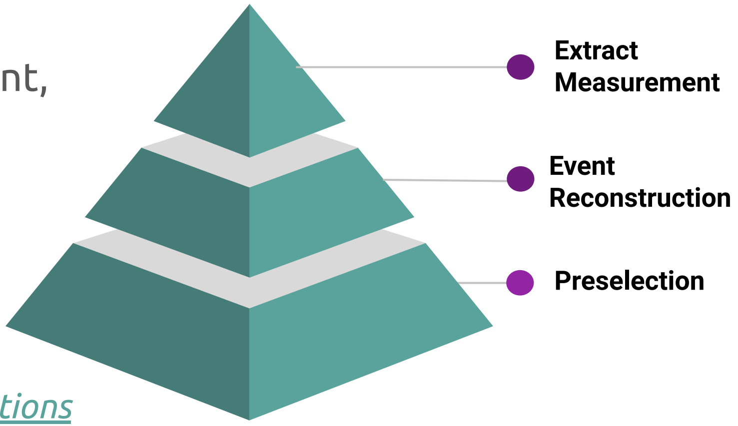


UPPSALA UNIVERSITET

Figures from: [Viktor Thorén, Hadron Physics in a Polarized World, 2022]

Extract Measurement

- Must accurately determine, minimize and evaluate sources of systematic errors
- Common (bad) practice: alter cut values and quote systematic uncertainty based on cumulation of observed variations in the relevant measurement
- May result in:
  - Penalised diligence:
    many checks = larger systematics
  - Larger systematics = overestimated agreement,
    fewer citations



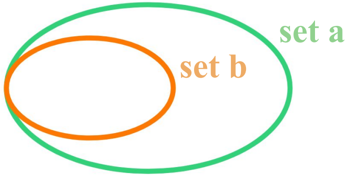
UPPSALA UNIVERSITET

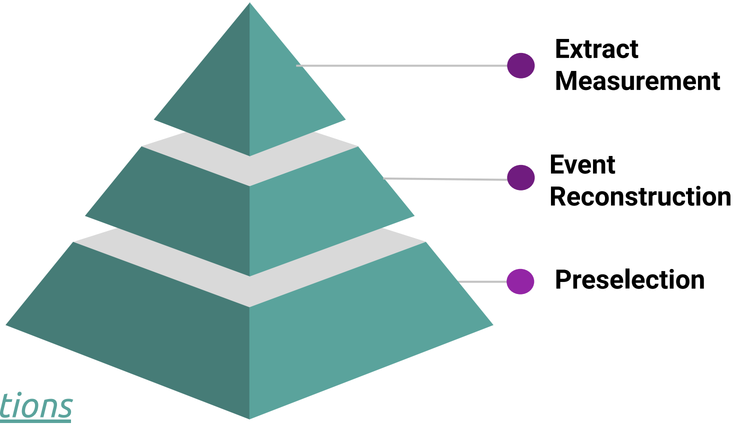
See *Roger Barlow, 2002,* *Systematic Errors: Facts and Fictions*

Extract Measurement

Better way: Determine *significant* sources of systematic errors with a *consistency check*

- Check passed ✓: do not quote as systematic uncertainty
- Check failed ✕: remove source of the error. If impossible, evaluate and quote systematic uncertainty
- R. Barlow suggests a check based on the uncorrelated uncertainty between data sets (e.g. two different cut values)
- Data sets containing the same events are highly statistically correlated





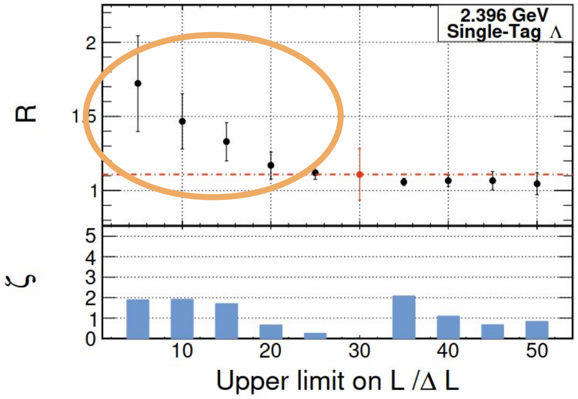
UPPSALA UNIVERSITET

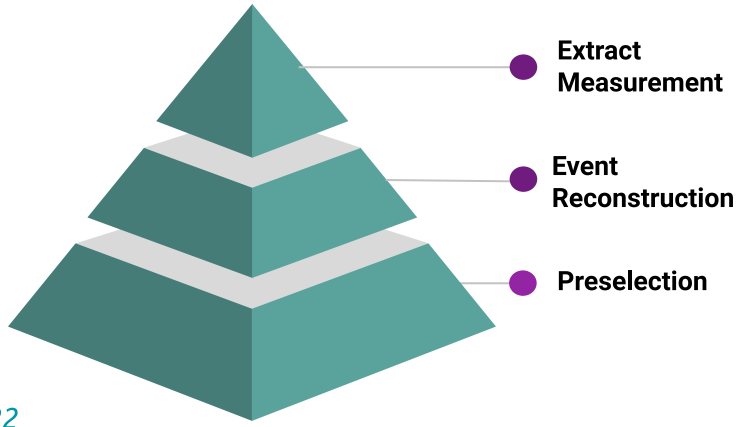
See *Roger Barlow, 2002,* Systematic Errors: Facts and Fictions

Extract Measurement

Difficulties:

- Many cuts throughout analysis for which a consistency check must be made
- Important to visually inspect how a measurement is impacted by each cut value
    - Look for trends - a critical region might be present even if check is passed





UPPSALA UNIVERSITET

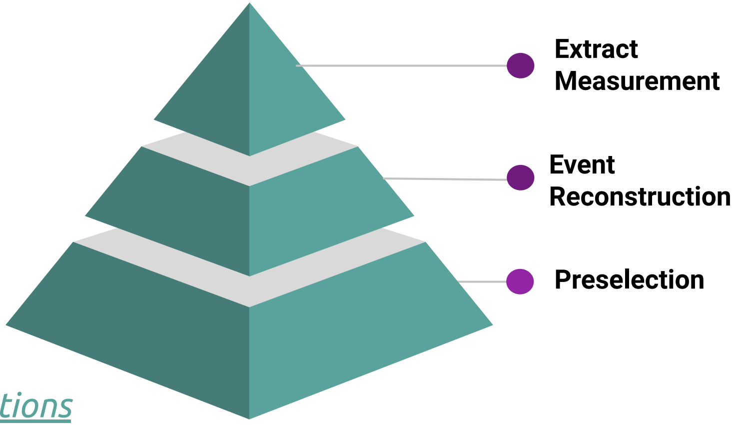
Figure from: Viktor Thorén, Hadron Physics in a Polarized World, 2022

Extract Measurement

Difficulties:

- Many cuts throughout analysis for which a consistency check must be made
- Important to visually inspect how a measurement is impacted by each cut value
  - Look for trends and critical regions

SproutCut - *features to come:*
Help you perform consistency checks and generate quality assurance plots to detect systematic errors



UPPSALA UNIVERSITET

See *Roger Barlow, 2002,* *Systematic Errors: Facts and Fictions*

The Sprout🌱 Family

- Data analysis relies heavily on ROOT
- ROOT is great, but:
  - Takes up many lines of code
  - Find myself writing a lot of similar code many times over
- General aims of Sprout🌱:
  - Automate the boilerplate
  - Make it easier to keep manageable and modular code

SproutParam

SproutPlot

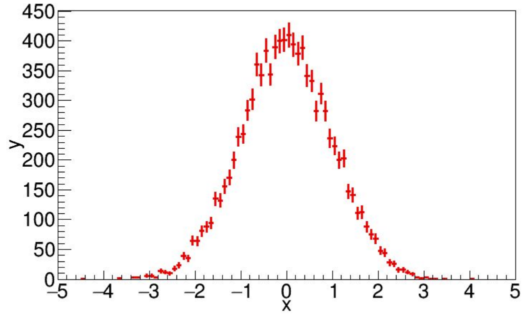
SproutFit

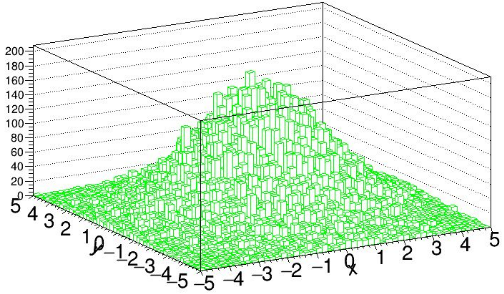
SproutCut

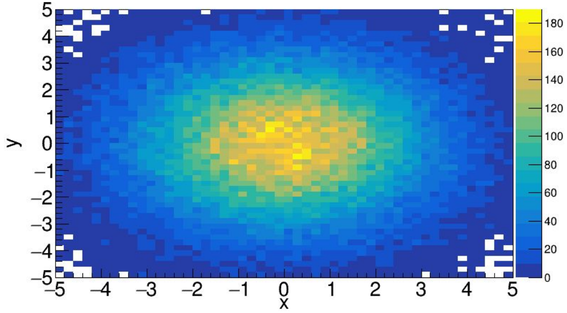
SproutPlot

An analysis involves hundreds of histograms

- SproutPlot keeps a collection of all of them
- Automatically style, draw and write all histograms at once…
- … but keep the ability to modify individual histograms as usual when needed
- Contains many more features you might find useful!







UPPSALA UNIVERSITET

SproutPlot

```
SproutPlot myPlots;

// create and fill histograms

TFile myFile("myFile.root", "recreate");
myFile.cd()
myPlots.writeHist();
```

```
// create and fill histograms

TFile myFile("myFile.root", "recreate");
myFile.cd();
h1.Write();
h2.Write();
h3.Write();
h4.Write();
h5.Write();
h6.Write();
h7.Write();
h8.Write();
h9.Write();
h10.Write();
```

SproutFit

- Fits user-specified signal+background distributions to histograms contained in a SproutPlot.
- Signal and background models specified by the user in a .txt file
- Automatically finds suitable start parameters
- Easily modify parameters of bad fits

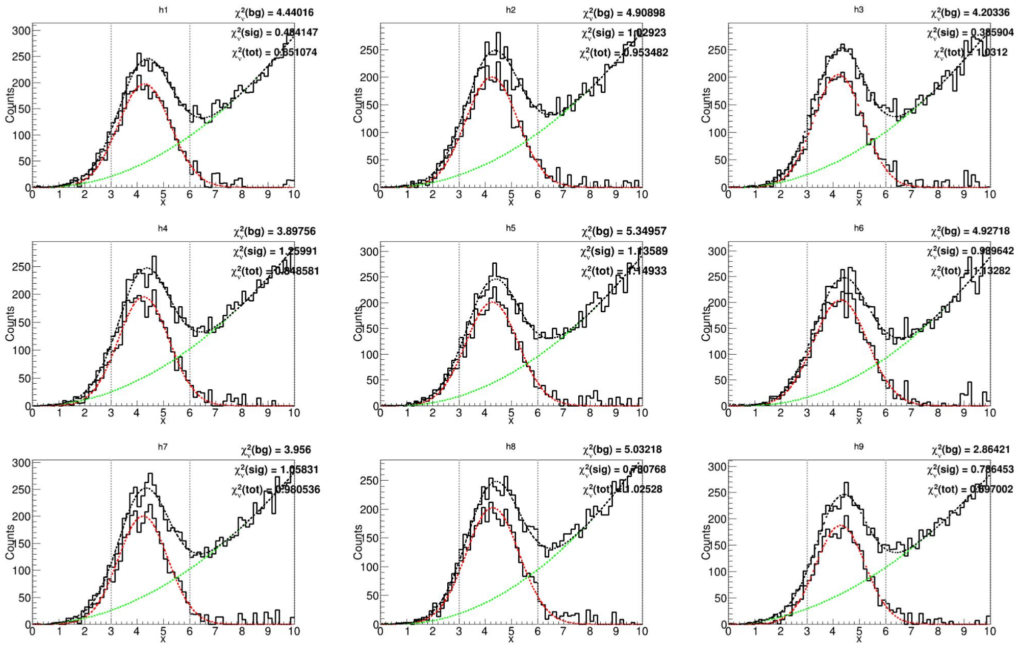
```
pol2 gaus -3.11928 4.99871 -2 2 28.7394  7.78128
pol2 gaus -3.11928 4.99871 -2 2 41.237  9.95509
pol2 gaus -3.11928 4.99871 -2 2 19.3014  7.24632
pol2 gaus -3.11928 4.99871 -2 2 29.7916  8.59777
pol2 gaus -3.11928 4.99871 -2 2 36.1238  8.90835
pol2 gaus -3.11928 4.99871 -2 2 72.5726  14.8853
pol2 gaus -3.11928 4.99871 -2 2 37.4449  9.37459
pol2 gaus -3.11928 4.99871 -2 2 34.4013  9.57495
pol2 gaus -3.11928 4.99871 -2 2 75.3767  14.8835
pol2 gaus -3.11928 4.99871 -2 2 22.0282  7.48923
pol2 gaus -3.11928 4.99871 -2 2 23.6093  7.68782
pol2 gaus -3.11928 4.99871 -2 2 24.4318  8.17412
pol2 gaus -3.11928 4.99871 -2 2 10.4908  6.32788
pol2 gaus -3.11928 4.99871 -2 2 29.1846  7.97549
pol2 gaus -3.11928 4.99871 -2 2 37.3783  9.18131
pol2 gaus -3.11928 4.99871 -2 2 22.4391  7.93826
pol2 gaus -3.11928 4.99871 -2 2 78.2404  14.4209
pol2 gaus -3.11928 4.99871 -2 2 24.6451  7.74814
pol2 gaus -3.11928 4.99871 -2 2 36.5364  10.5195
pol2 gaus -3.11928 4.99871 -2 2 34.7326  8.65592
pol2 gaus -3.11928 4.99871 -2 2 33.2054  9.41413
pol2 gaus -3.11928 4.99871 -2 2 32.8913  8.84999
pol2 gaus -3.11928 4.99871 -2 2 28.615  8.67445
pol2 gaus -3.11928 4.99871 -2 2 11.157  6.63691
pol2 gaus -3.11928 4.99871 -2 2 22.894  8.08729
```


Inspired by Rafał Lalik's (Jagiellonian University) software package HelloFitty

SproutFit

Perform many fits and produce quality assurance plots with ease:

```
SproutPlot myPlots;
// Add and fill histograms to myPlots
SproutFit fitter;
fitter.fit(myPlots,"myFits.png");
```



UPPSALA UNIVERSITET

Inspired by Rafał Lalik's (Jagiellonian University) software package HelloFitty

Access Sprout

Who can use Sprout?

- You use C++/ROOT as part of your analysis
- Sprout can be used from both executables and ROOT-macros

Download the package and access guides for set-up and usage at:

https://github.com/malle-b/Sprout

UPPSALA UNIVERSITET

Contact: malin.bohman@physics.uu.se

Thank you🌱

Backup

SproutParam

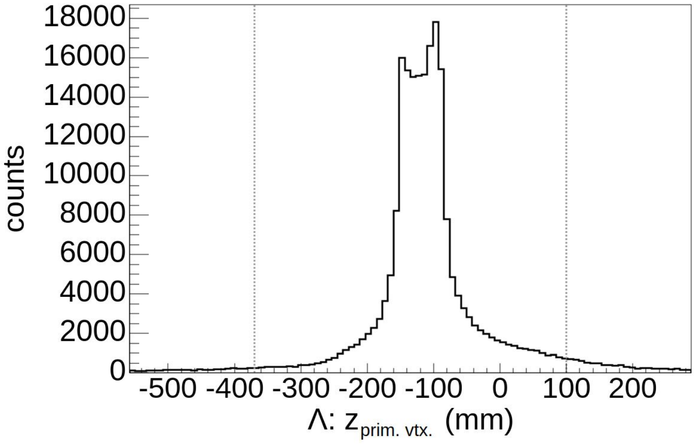
- Store parameters by name in .txt file and easily retrieve them from your analysis code
- Retrieval algorithm is O(1) on average
- Stored parameters can be changed easily without needing to recompile
- Easily define and run different analysis cases with different sets of parameters for comparison

```
# paths and file names
#-----------------------------
dst_list                /path/to/my_dst_list.txt
out_dir                 /path/to/output/folder
dst_loop_output         lambdaCands.root

# event selection
#-----------------------------
trigger_pt1             0
trigger_pt2             0
trigger_pt3             1
proton_pion_sel         1

# track selection
#-----------------------------
use_flag_Hadron         1
use_flag_kIsUsed        0
use_cut_pid             1
p_cut_file              Proton_Cut_file.root
pi_cut_file             Pion_Cut_file.root
use_hades_protons       1
use_fwd_protons         1

# corrections
#-----------------------------
beam_correction         1
energy_correction       1

# background reduction
#-----------------------------
invM_cut                1
```

UPPSALA
UNIVERSITET

SproutCut

- Prepare cuts by specifying name (for later access) and cut value
- SproutCut keeps a collection of all created cuts
- Upon applying a cut, automatically generates histogram for quality assurance
- Easily write all generated histograms at once
- See full documentation on GitHub for more and future features



UPPSALA UNIVERSITET