Table of Contents

Proton Calibration Samples in Run I

Proton issues

In run 1 the proton coverage is not as good as desired by some analyses.

While an extremely high statistics, high purity sample of protons can be acquired through the decay of the Lamba0, these protons do not have the same kinematic coverage as those from beauty and charm decays. In general the calibration sample from Lambda0 suffers from a lack of protons with high transverse momentum.

In August 2014 a second calibration sample was included which used protons originating from an inclusive LambdaC decay. While LambdaC decays are also abundant in our data one key problem is that most of the trigger and/or stripping lines that collect these events place a PID requirement on the proton. This means that only a few routes are available and that these tend to be inefficient or prescaled. While the sample from LambdaC has improved coverage, it still suffers from low statistics. The overall calibration sample is two orders of magnitude lower than that of the Lamba0 calibration sample.

In June 2015 a third calibration sample is introduced. This is an additional sample of protons from LambdaC, but these LambdaC originate from Lb decays and are triggered through the muon daughter of the semi-leptonic decay of the Lb. This new sample is independent of the existing sample by virtue of the trigger paths. The new sample increases the statistics of the existing LambdaC sample by approx 60% however has a slightly softer PT spectrum. Hence while this solution is unlikely to be the magic fix the increase in statistics may alleviate some of the problems to a greater extent.

This is likely to be the last update to the proton samples prior to any restripping of the Run I data. Studies so far on other potential sources of kinematically selected protons have not identified other viable sources.

Due to the number of different proton samples, and the fact that they were produced at different times has lead to slightly complicated naming system. This is described in detail on this page.

Naming conventions

Sample	Source	Stripping samples	Particle type
Lam0	prompt L0 -> Ppi decays	S20, S20r1, S21, S21r1	P
IncLc	Lc ->pkpi - hadronic trigger	S20, S20r1, S21, S21r1	P_IncLc
LcfB	Lb ->Lc(pkpi) mu nu - Muon trigger	S21, S21r1	P_LcfB
TotLc	Merger of the two Lc samples	S21, S21r1	P_TotLc

Meaning of the Stripping Samples

S21 and S20 despite both being done under Reco-14 are not the same reconstruction due to on the fly changes in the CALO. Please use the one that corresponds to your sample. The differences for protons are very small. If you want the additional protons in different S21 samples for a S20 analysis it is suggested that you use S21 and apply some systematic for the difference in the performance (which can be estimated from the IncLc sample that is available in both).

The new Lc from Lb sample has been created using the S21 stripping line and hence has S21 in the filename. At the same time the other Lc sample was recreated, to allow for the merger of the two samples with relative ease. There will be small statistical differences between the S21 and S20 IncLc samples due to a slight change in selection strategy.

The calibration samples made present in Aug 2014 remain with the S20 naming scheme. Their presence is required so that mature analyses continue to have access to them and for reasons of data preservation.

New analyses should use TotLc + Lam0 where possible to access the largest calibration statistics.

The samples IncLc and LcfB are available separately in S21 incase some more detailed studies are required.

Accessing the various samples

The version of PIDPERFSCRIPTS must be v9r1 or higher

```
python MakePerfHistsRunRange.py 21 MagUp P_TotLc "[DLLK >0.0"]
```

The minimum arguments are " stripVersion " " polarity " " Particle type " " PID cut "

For protons the " Particle type " arguments are "P", "P_IncLc", "P_LcfB", "P_TotLc". If you use "P_IncLc" then make sure you specify the stripping version that you actually want. For all other cases there is only one option that will result in a file being accessed. If the wrong option is used the program will exit stating that the file was not found.

New analyses should only use "P" or "TotLc" to access either the sample from L0 or the combined Lc sample.

Which sample should i use? Which binning should i use?

The high statistics of the protons from Lam0 should not be ignored. They will provide much better coverage in regions of lower transverse momentum, or low momentum and high eta, than the Lc samples simply by virtue of statistics. They should be used as far as possible.

If you have an existing mature analysis you should continue to use a mix of protons from Lam0 and protons from Lc (P_IncLc S20)

If you are starting a new analysis or wish to improve your existing analysis then use a mix of protons from Lam0 and protons from both Lc samples (P_TotLc S21).

Since the coverage in your region of interest is likely to switch from the Lam0 sample to the Lc sample defining a good binning scheme is paramount.

A good binning scheme takes into account the region where your signal lies - (there is no point in computing efficiencies for tracks that don't lie in this region) and take into account the distribution of calibration statistics. There is a tradeoff between fine bins and having sufficient statistics within a bin to accurately determine the efficiency. Large bins can lead to larger systematic uncertainties if the efficiency across the bin has a strong variation. Any region that does not have sufficient coverage by the calibration samples should be removed from you signal sample. What counts as sufficient is, naturally, analysis dependent. The addition of the extra Lc calibration sample may allow for a finer binning or it may allow the total phasespace covered to be extended (or both).

It is therefore clear that analyses dependent on the proton samples should be defining their own binning schemes.

-- SnehaMalde - 2015-06-29

This topic: LHCb > LHCbPIDCalibProts4Run1
Topic revision: r2 - 2016-03-18 - SnehaMalde