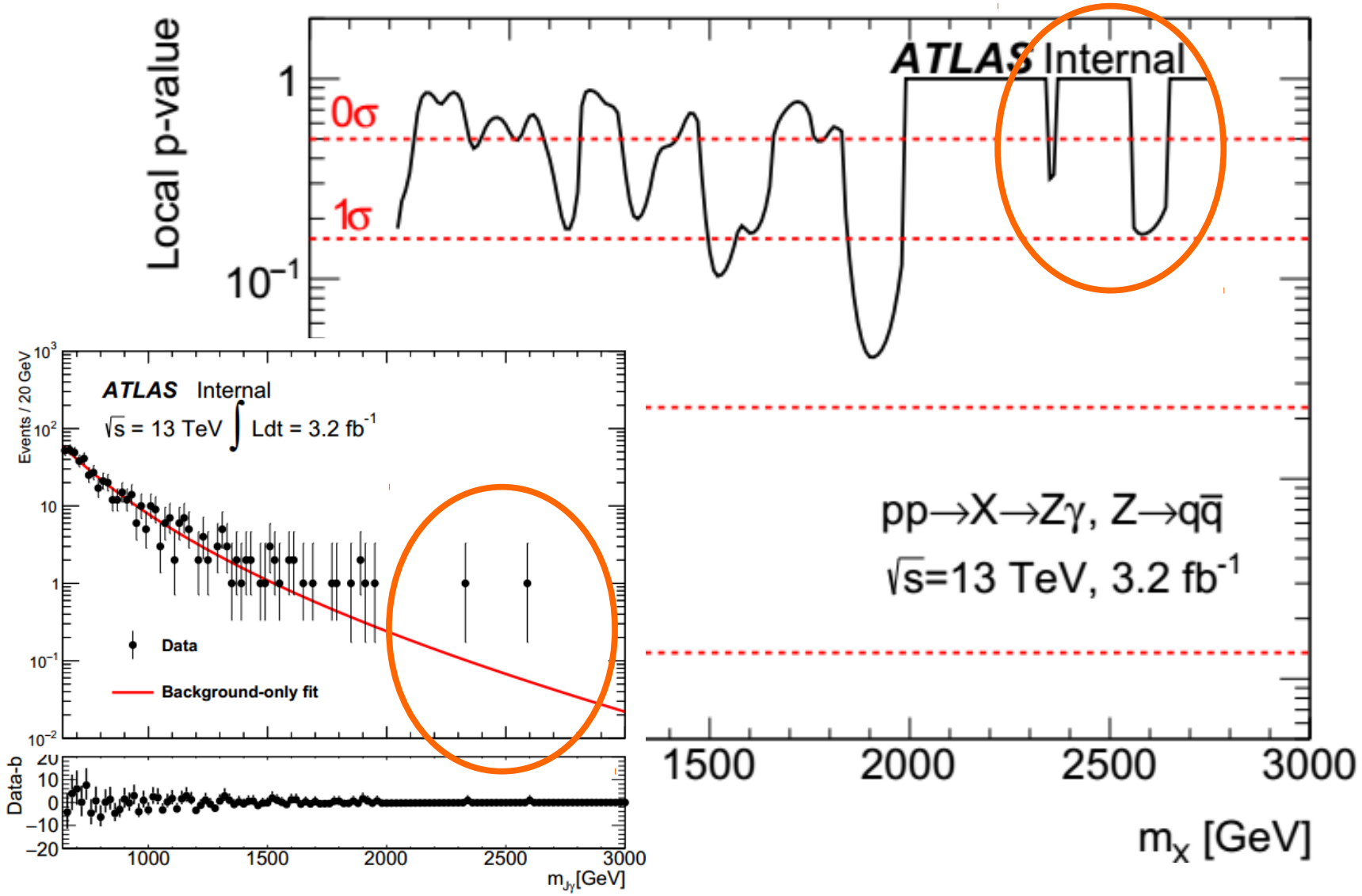
Unbinned fit
in Zy boosted analysis

Xiaohu SUN
IHEP
2016-06-17

p0 issue @ 2.6TeV

Bill asked:

Fig 13 b shows p-values where the event at 2.35 TeV makes a delta-function spike, while that at 2.6 TeV is much broader.



La cause et la solution

When fitting to the points with no data observed, the signal strength mu can be pushed to a very negative number to balance the expected background

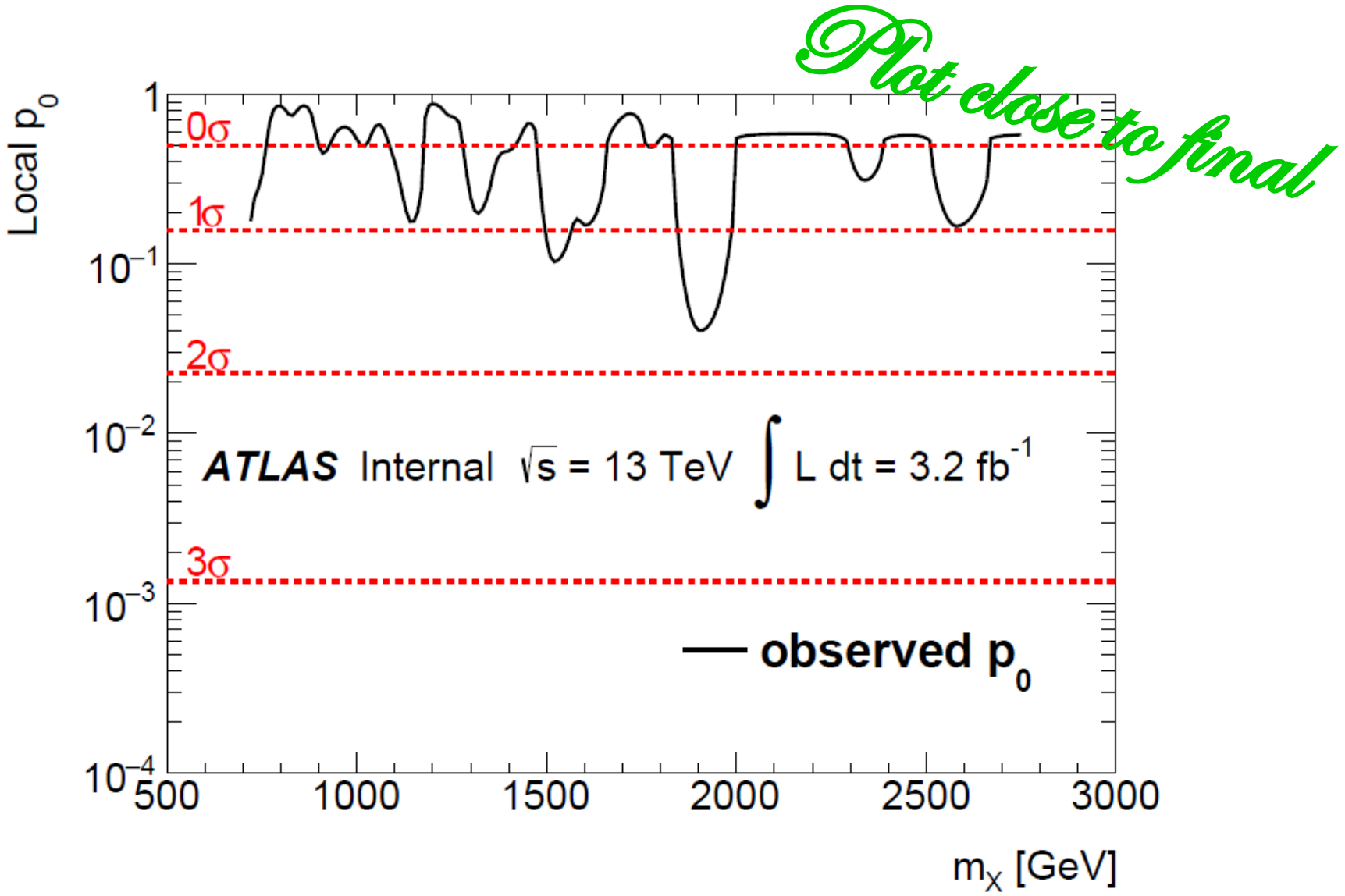
In high mass where observed data is few, one fits at a certain mass point, i.e. mu*S+B=0, which is totally OK; BUT mu*S+B goes to negative in higher mass points, where the expected background is small (at the tail of power-law)

Negative pdf value causes fit failures and prevents proper calculation of TS, resulting in wrong p0 as you see in high mass

We implemented Enrique's suggestion on limiting the negative boundary of mu. This effectively protected pdf values.

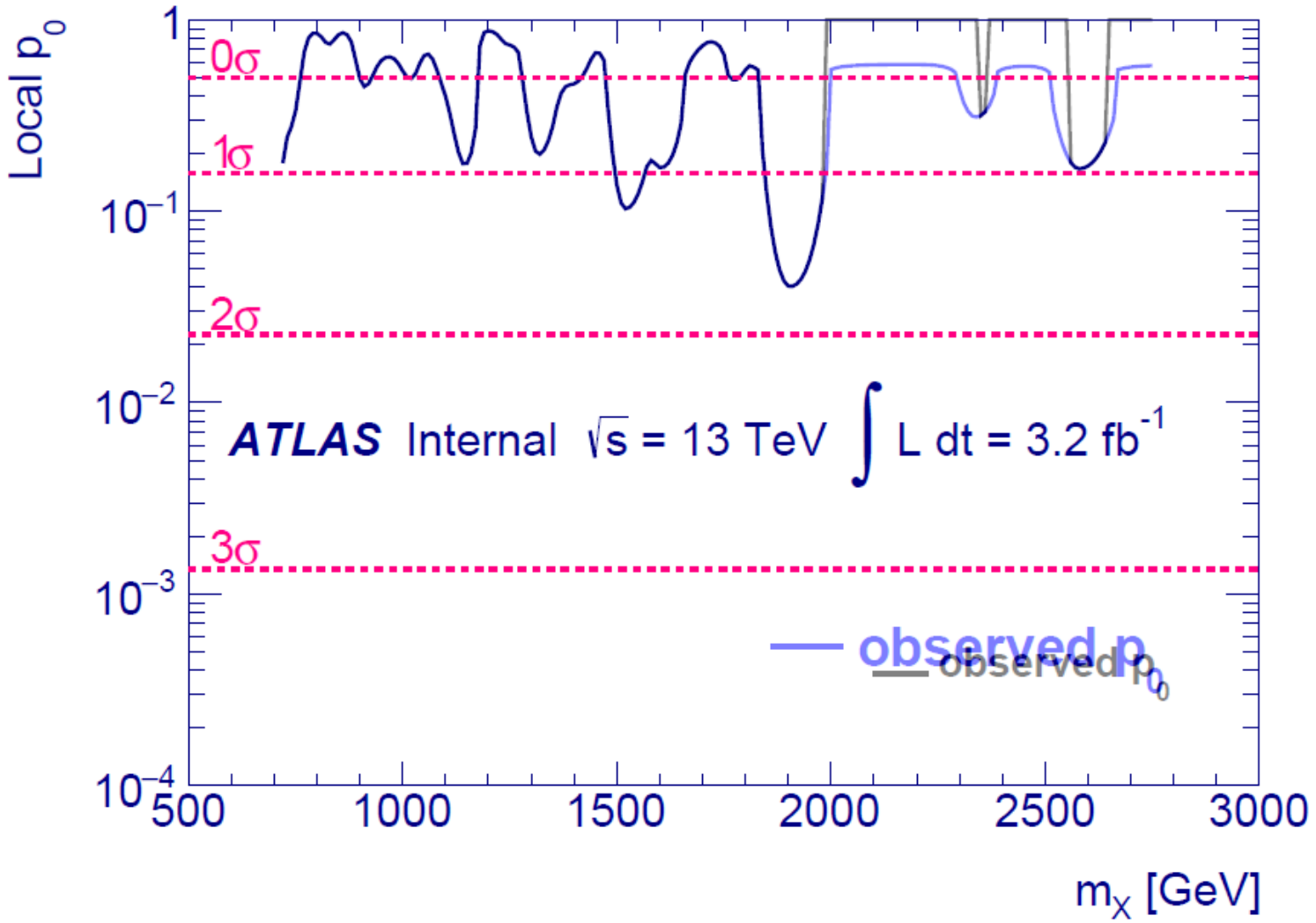
Therefore, mu is negatively limited in high mass region

Voila



There is no failure in Minuit now :)
Fit is healthy and p0 curve looks correct

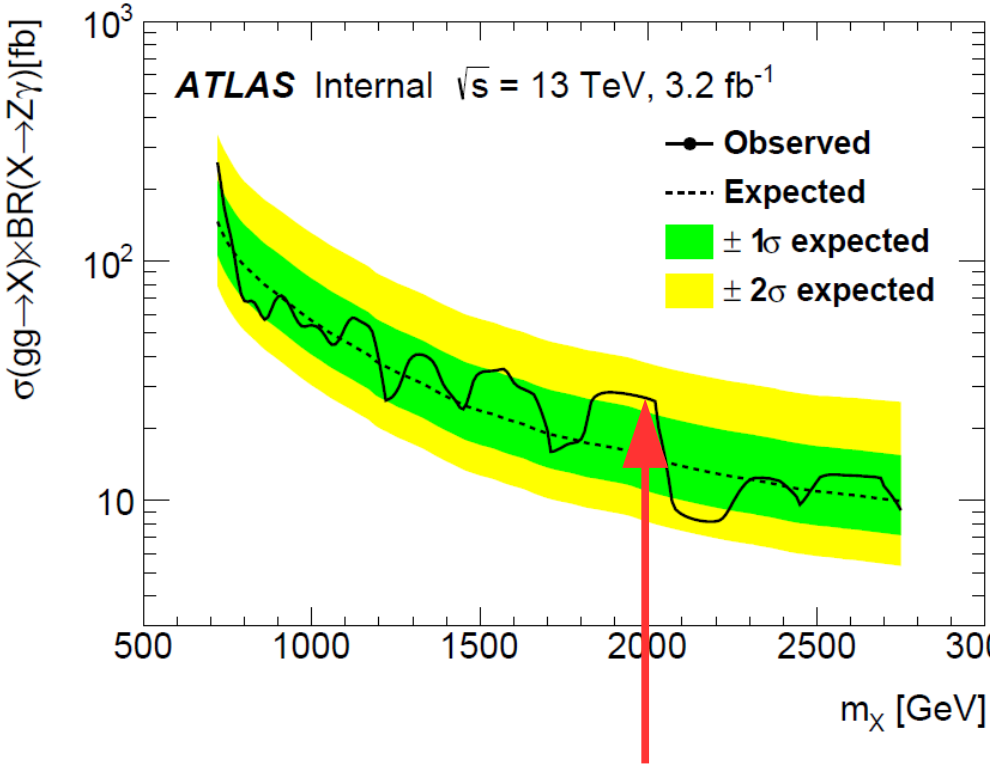
Check with the last picture

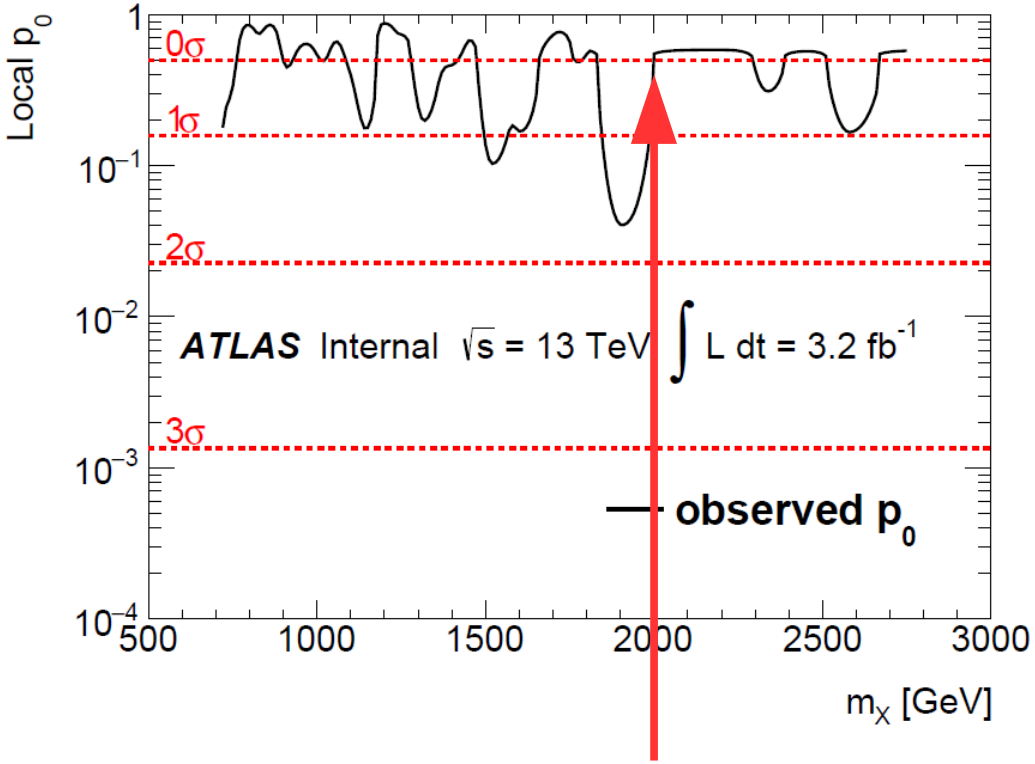


The last p0 curve is black on bottom
The new p0 curve is bleu on top

Another question from Bill

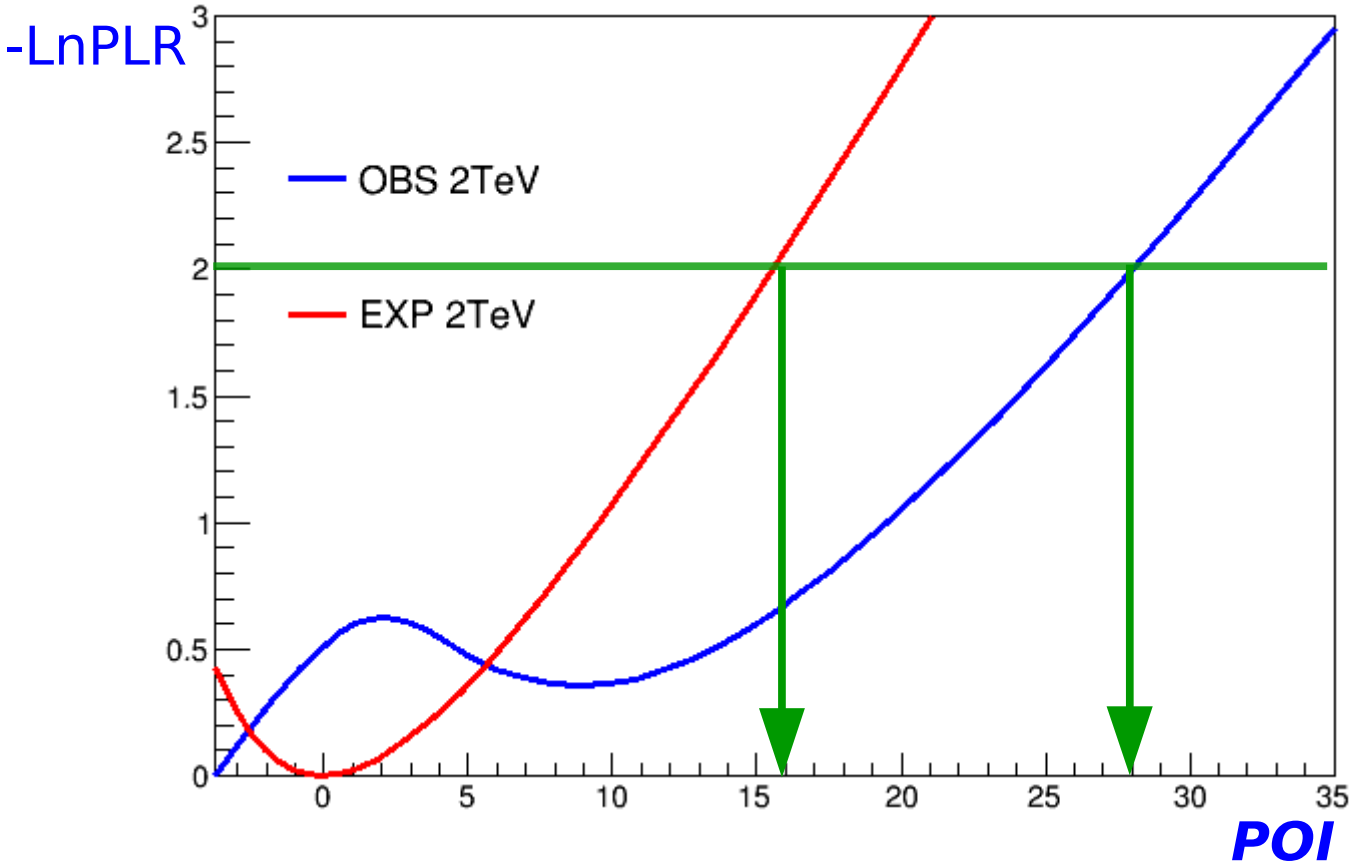
Also, at 2 TeV the p-value of the 1.9 TeV bump falls back to 1 - no sign of it at all. While in the limit plot, fig 4, there is a remant nearly to 2.1 TeV - but a flat 2.2 sigma from 1.9 to 2 TeV, which looks very different from the p-value.





Likelihood

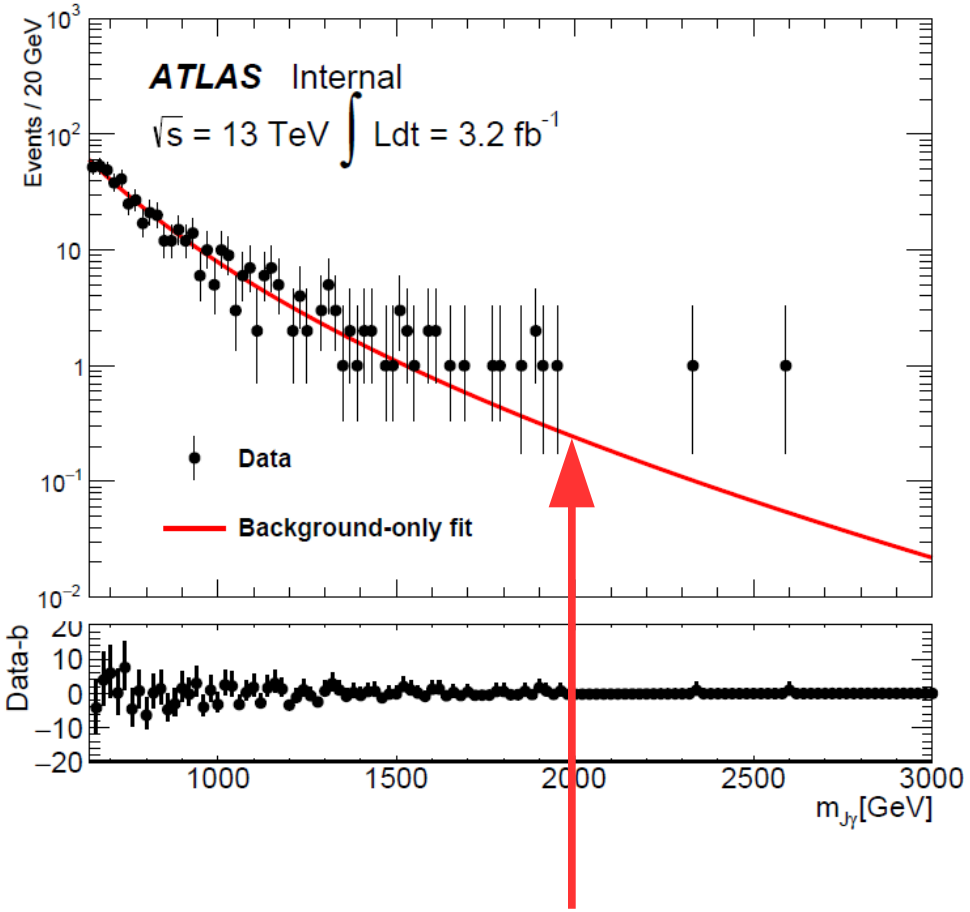
Best fitted POI is smaller in obs than in exp due to data deficit
But likelihood curve is broader in obs than in exp
Thus, limit is higher in obs than in exp

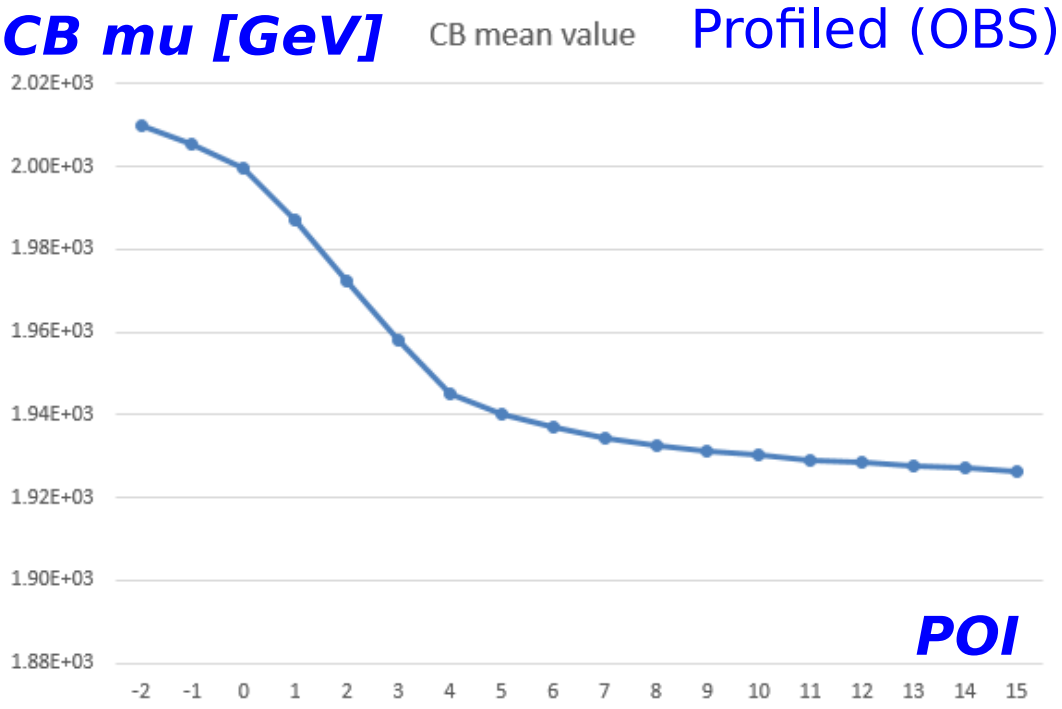


So far Bill's question is answered
But there is a local minimum in obs → explained in next page

Local minimum is due to data

There is NO data; best fit is a negative mu
This is the global minimum in likelihood curve

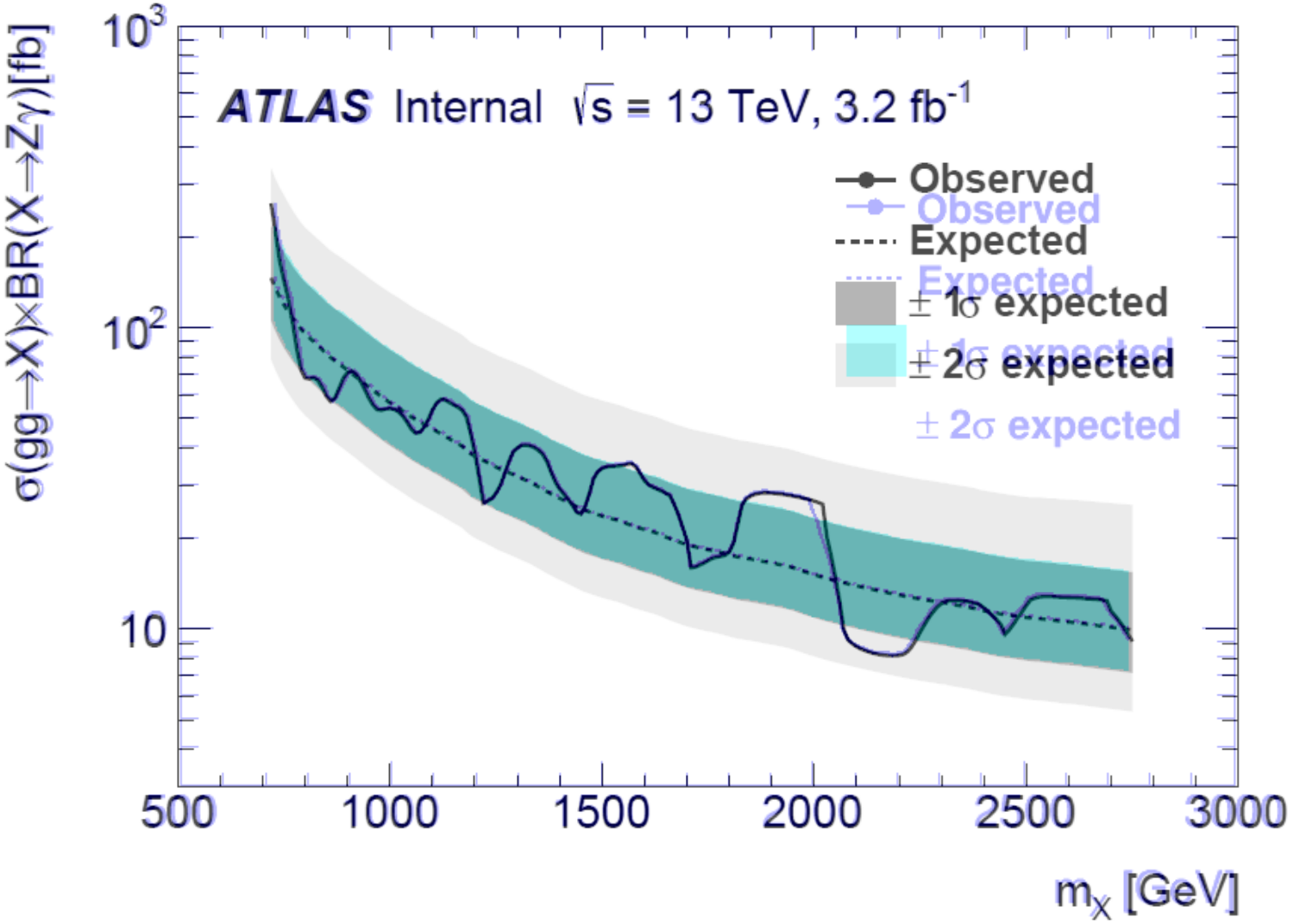




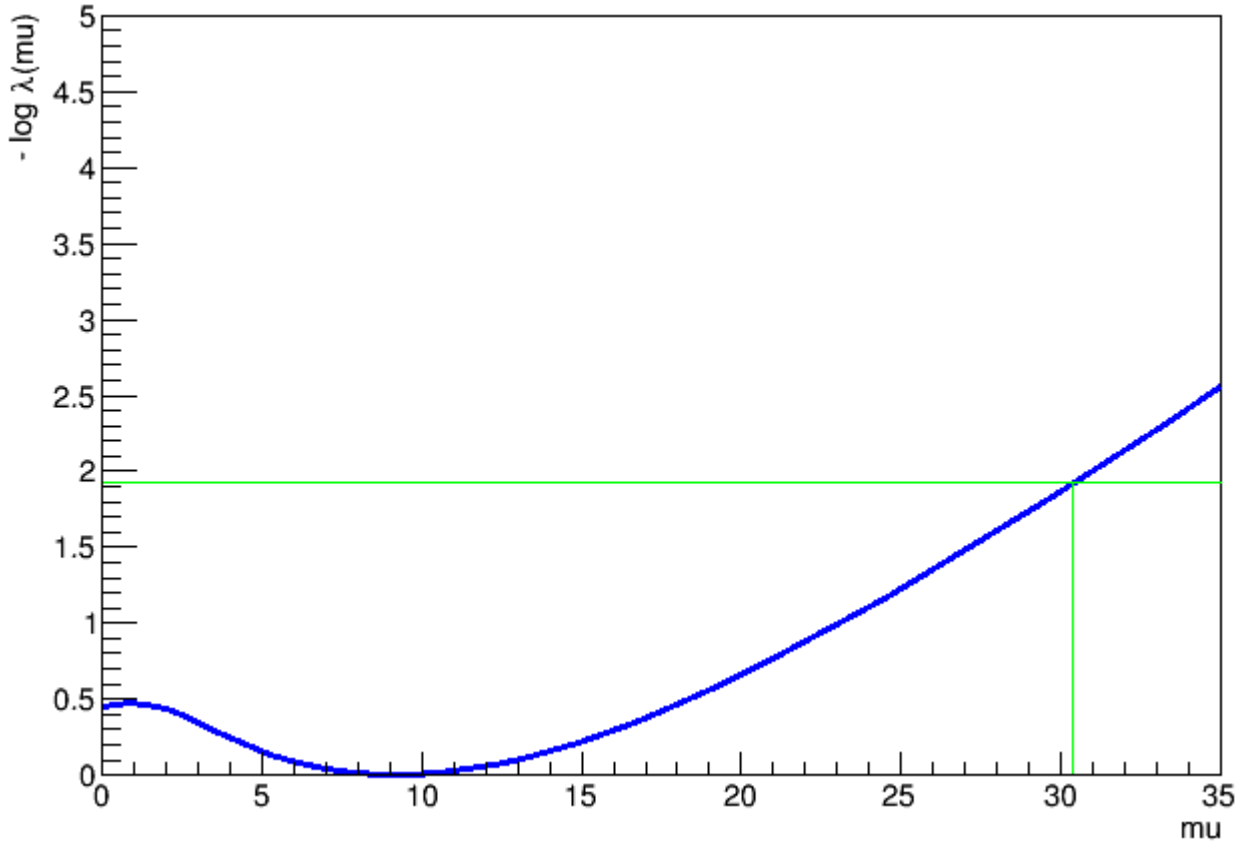
At 2TeV, minimize NLL as a func of POI
CB mean value is pulled to data excess at 1.93TeV
There we have a local minimum
And this makes PLR curve broader

Linear interpolation

1990-2070, points in between are linearly interpolated (obs limits)

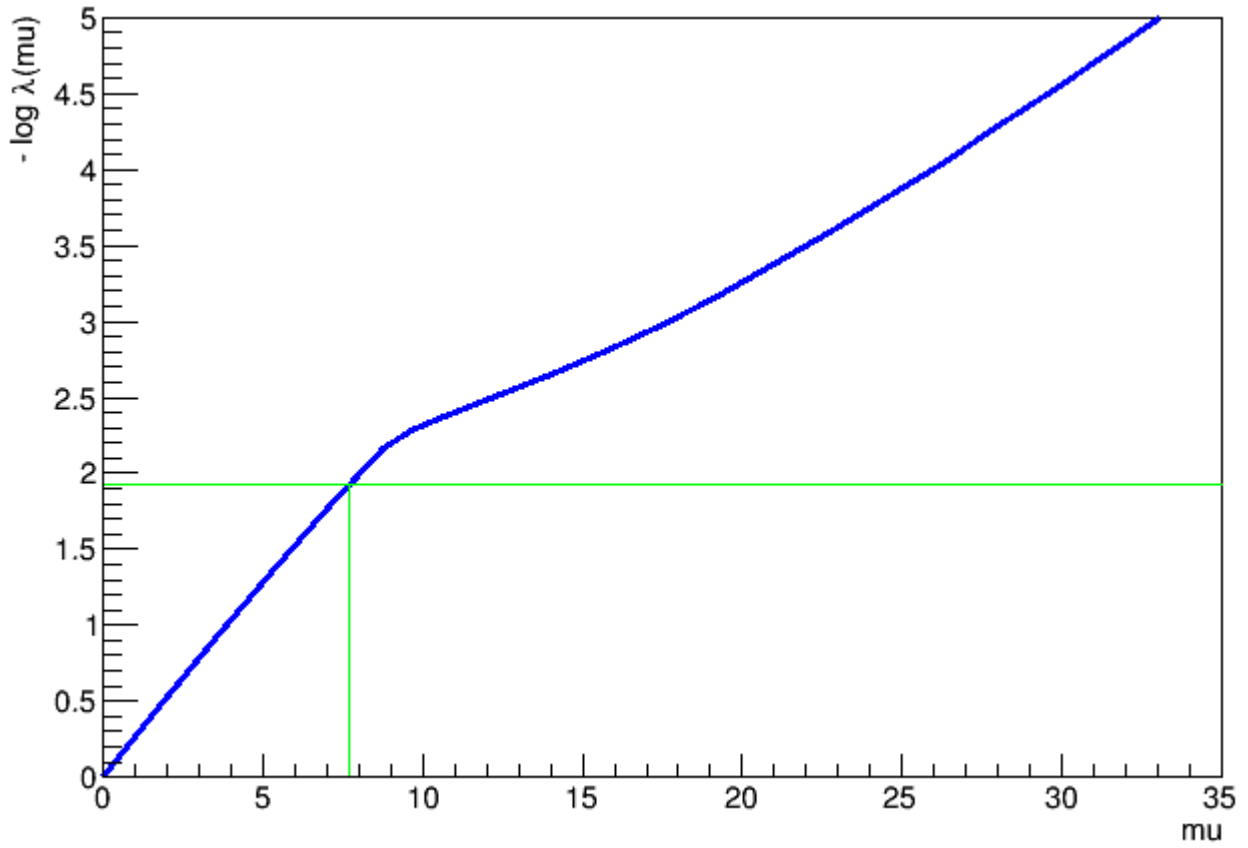


1990



Interpolation starts from 1990

2070



The local min moves out of -LnLambda=2.0 since 2070

Interpolation ends at 2070

Backup

Try capped p0

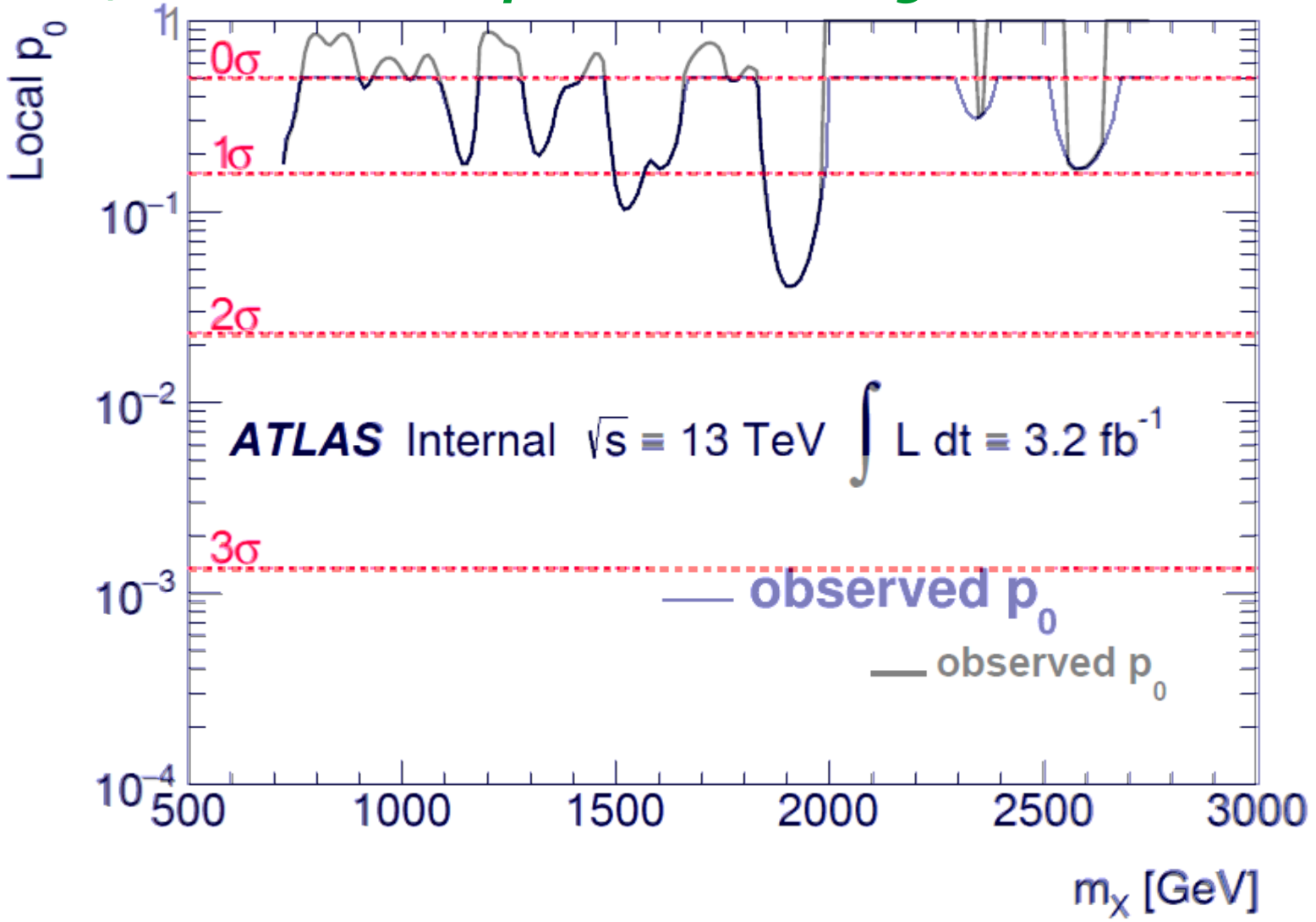
The uncapped p0 (our default) is on bottom in black

The capped p0 is on top in blue (AsymptoticCalculator is used)

Capped is much more stable in low stats region

But we are interested in uncapped p0

However, it tells how important the range of mu is



S