LHC physics

Maurizio Pierini
CERN

EPISODE I

SETTING UP A SEARCH AT THE LHC

MAURIZIO PIERINI

CERN

European Research Council

THE PLAN

- Lecture 1: **SETTING UP A SEARCH AT THE LHC**
  - Searching for SUSY in practice: strategy, trigger, reconstruction
  - Designing a search: Simplified Models
  - Building a search: signal region, control regions, statistics tools

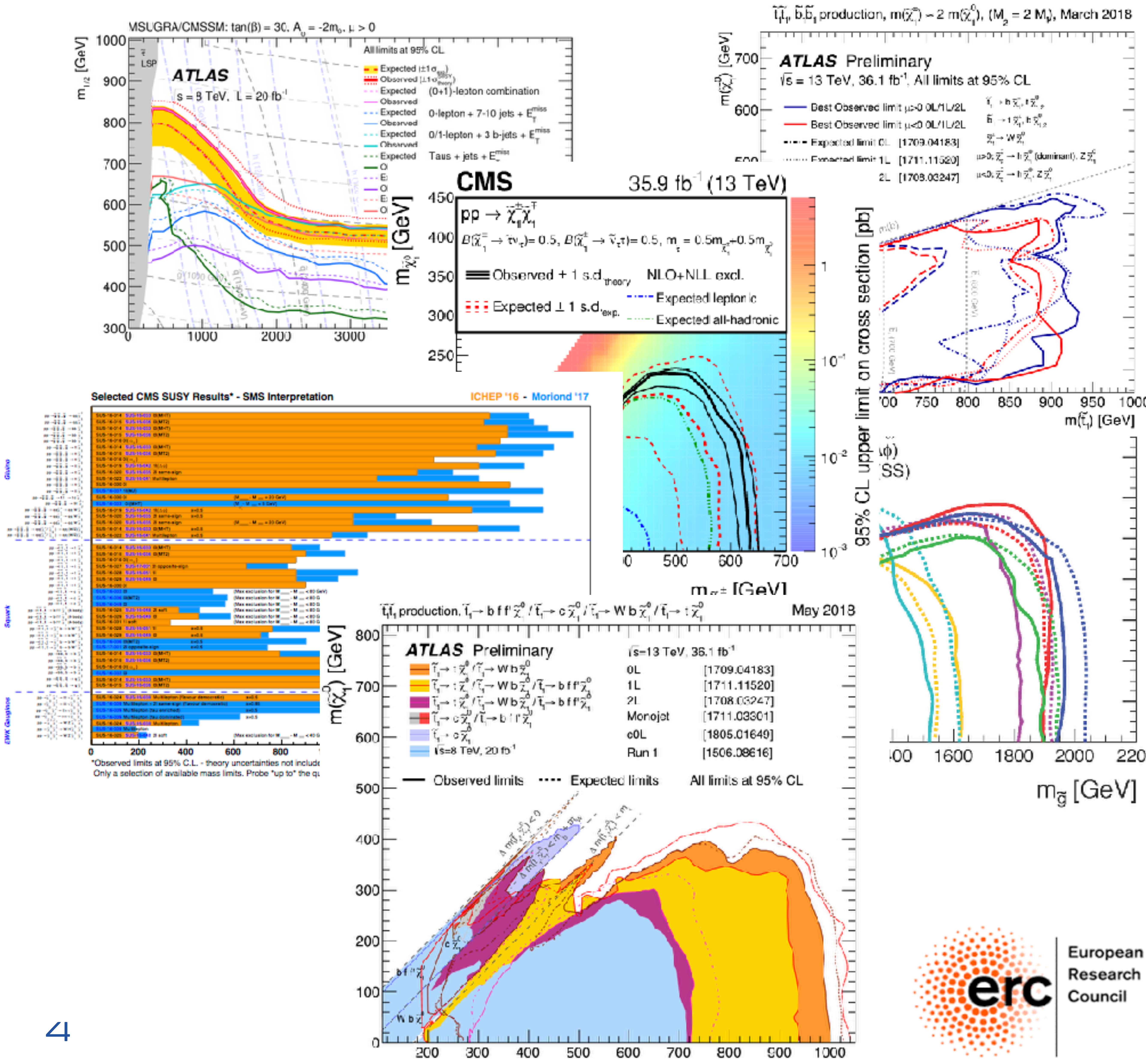
- Lecture 2: **R-PARITY CONSERVING SUSY**
  - DM direct production
  - DM cascade production

- Lecture 3: **BEYOND MET-BASED SEARCHES**
  - RPV SUSY
  - Displaced particles

European Research Council

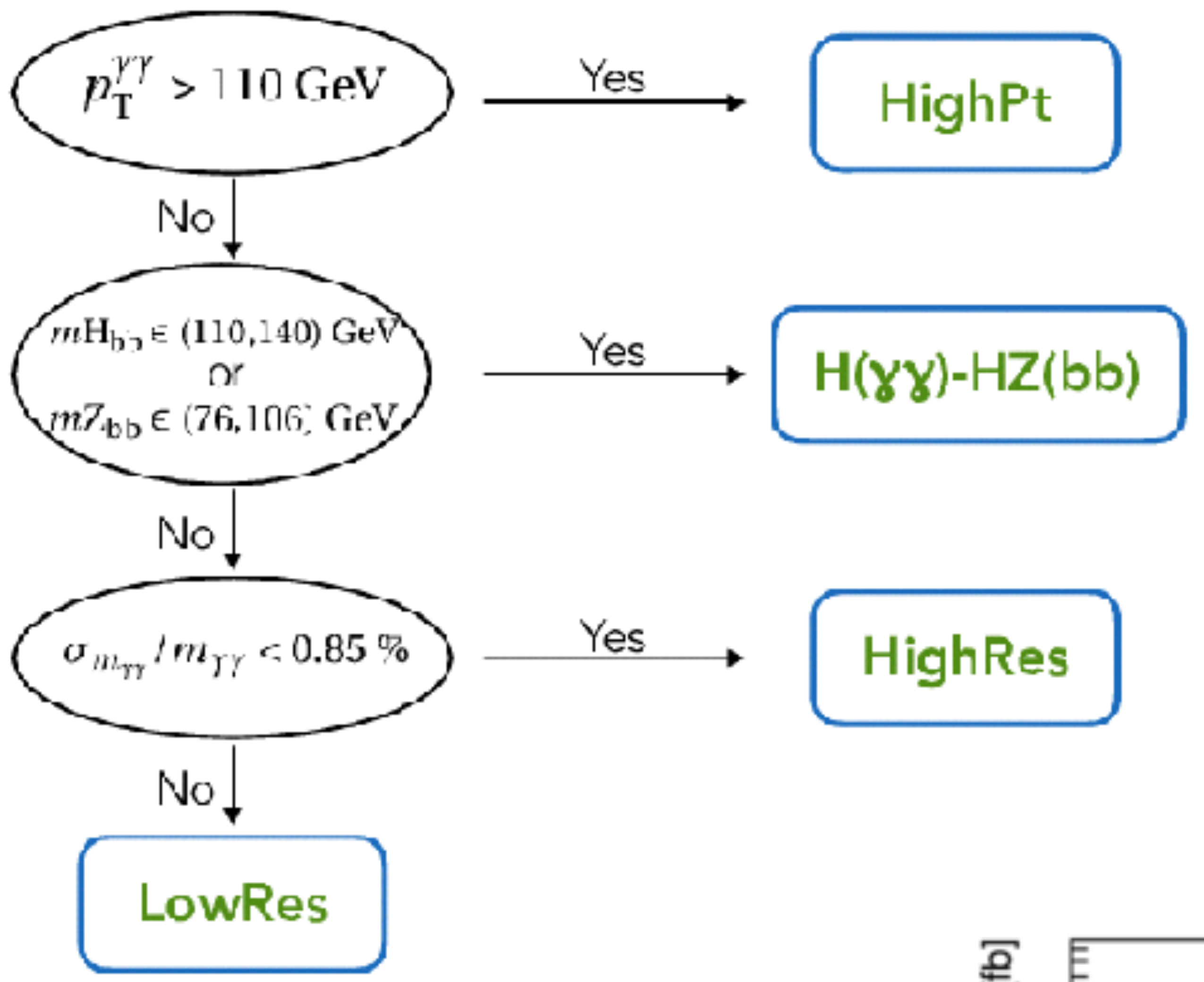
WHAT THESE LECTURES ARE NOT

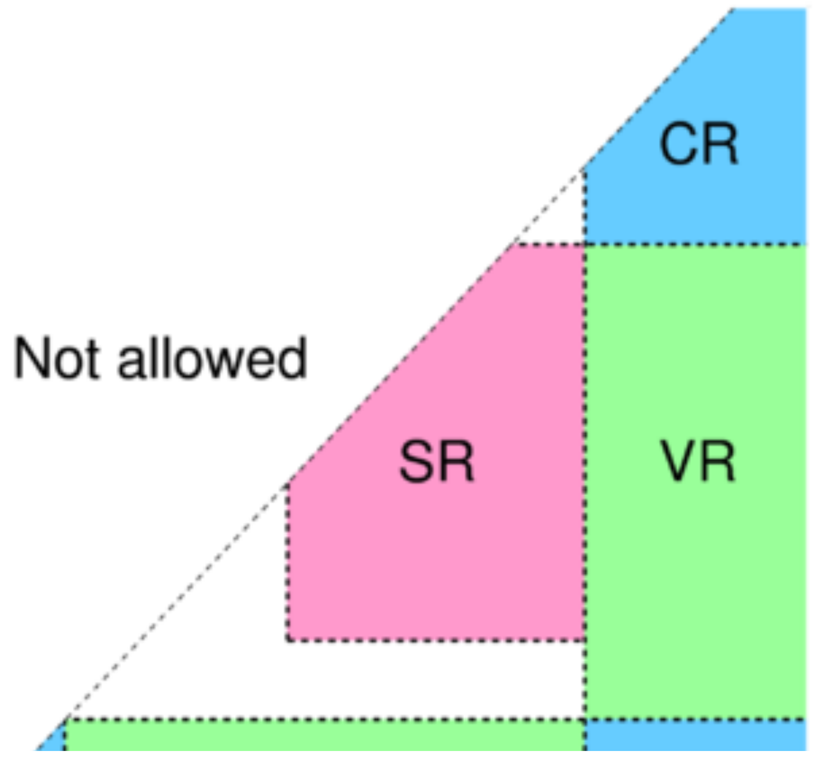
- *I don't intend to give you the status-of-the-art picture of searches @LHC*
  - *what is excluded in which final state up to which mass*
- *You have a full conference coming soon for that, hopefully with new results*

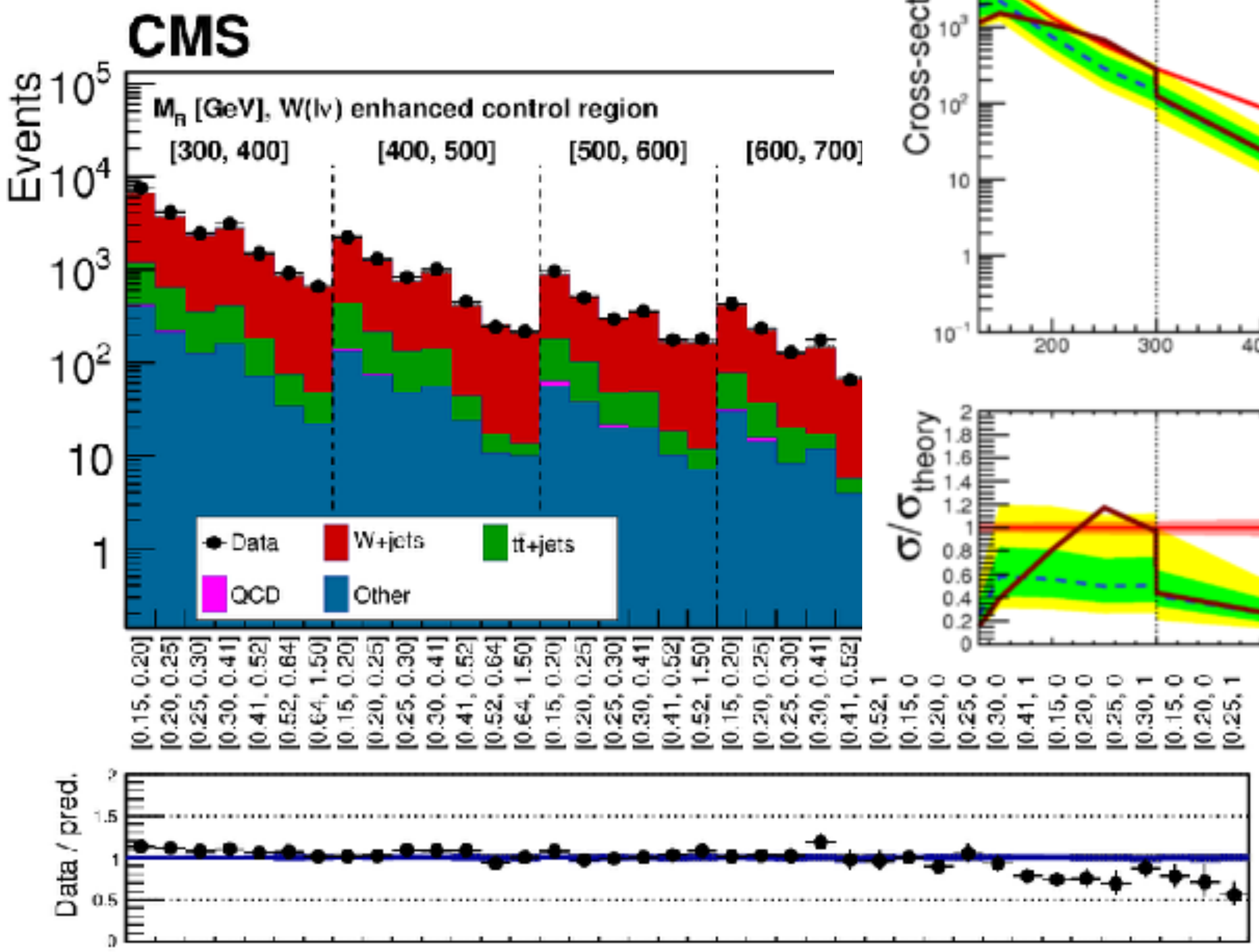


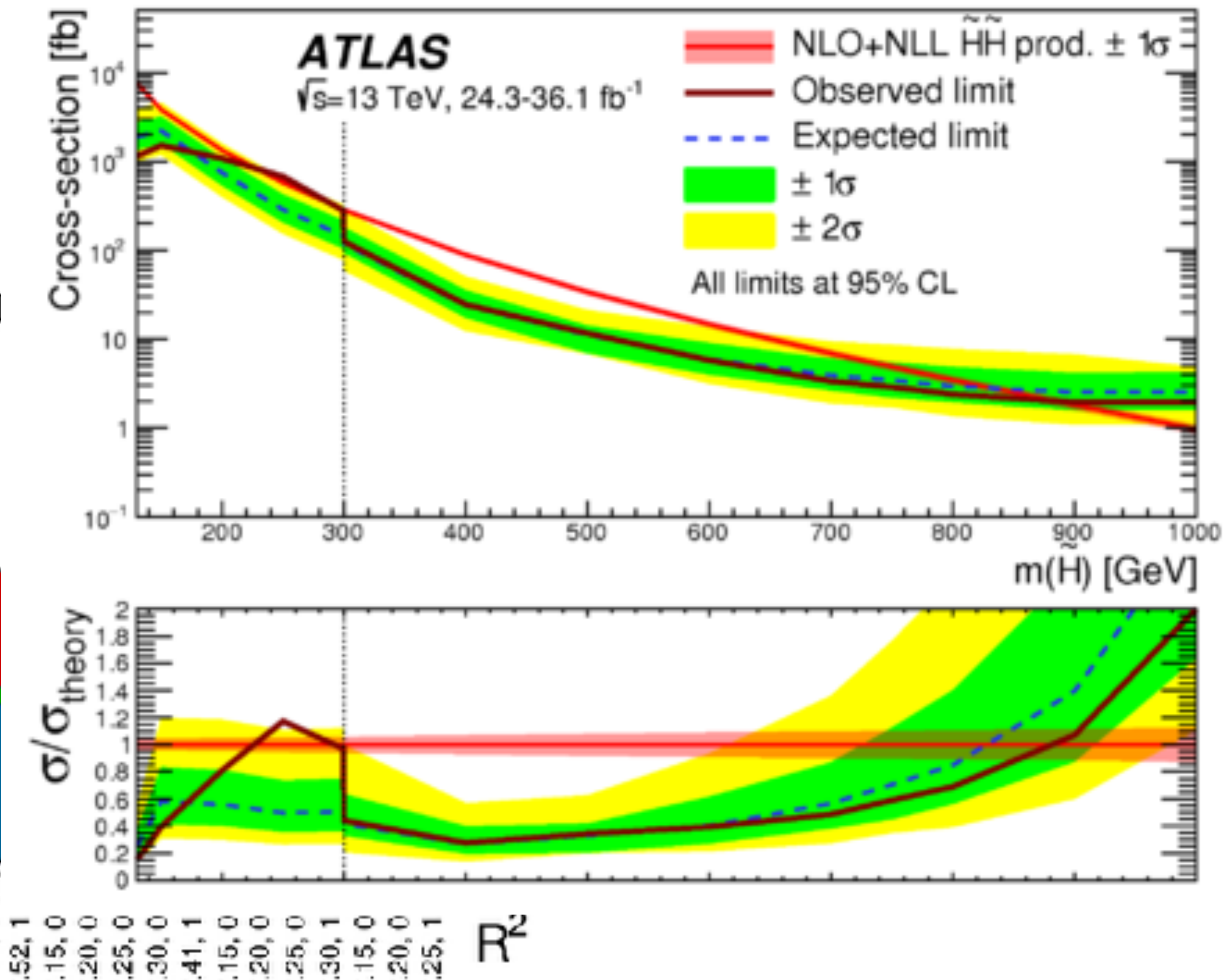
What These Lectures Are

- I would like to give you an idea of how a search is built
- I will use SUSY for practical reasons, but I will touch other searches too
- What I will say applies in general, not just to SUSY searches
- For experimentalists, it might be useful to design the strategy towards your next discovery
- For theorists, it might be useful to understand experimental results (i.e., where the assumptions come in, to which level results generalise, etc)



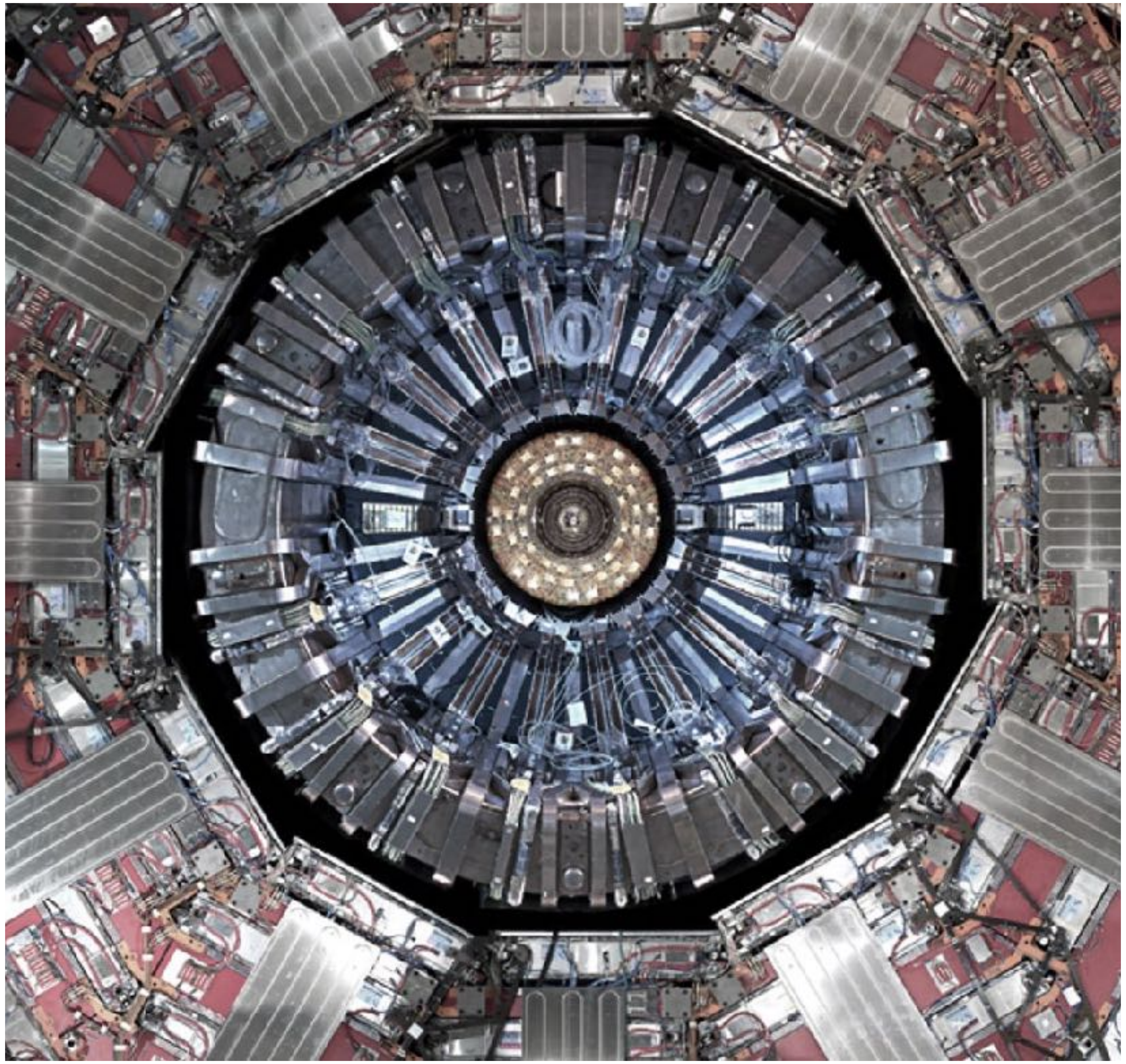






WARNING

- *The slides are CMS-biased, for obvious practical reasons*
  - *I am in CMS, I know CMS more, and it costs me less time to prepare lectures this way*
- *The large part of what I will say applies also to ATLAS*
- *If not, I will make it clear to you*



For decades physicists have been working on a beautiful theory that has promised to lead to a deeper understanding of the quantum world. **Now they stand at a crossroads:** prove it right in the next year or confront an epochal paradigm shift

By Joseph Lykken and Maria Spiropulu

PARTICLE PHYSICS

Supersymmetry and the Crisis in Physics

IN BRIEF

Supersymmetry postulates that every known particle has a hidden superpartner. Physicists love supersymmetry because it solves a number of problems that crop up when they try to extend our understanding of quantum mechanics. It would also potentially solve the mystery of the universe's missing dark matter.

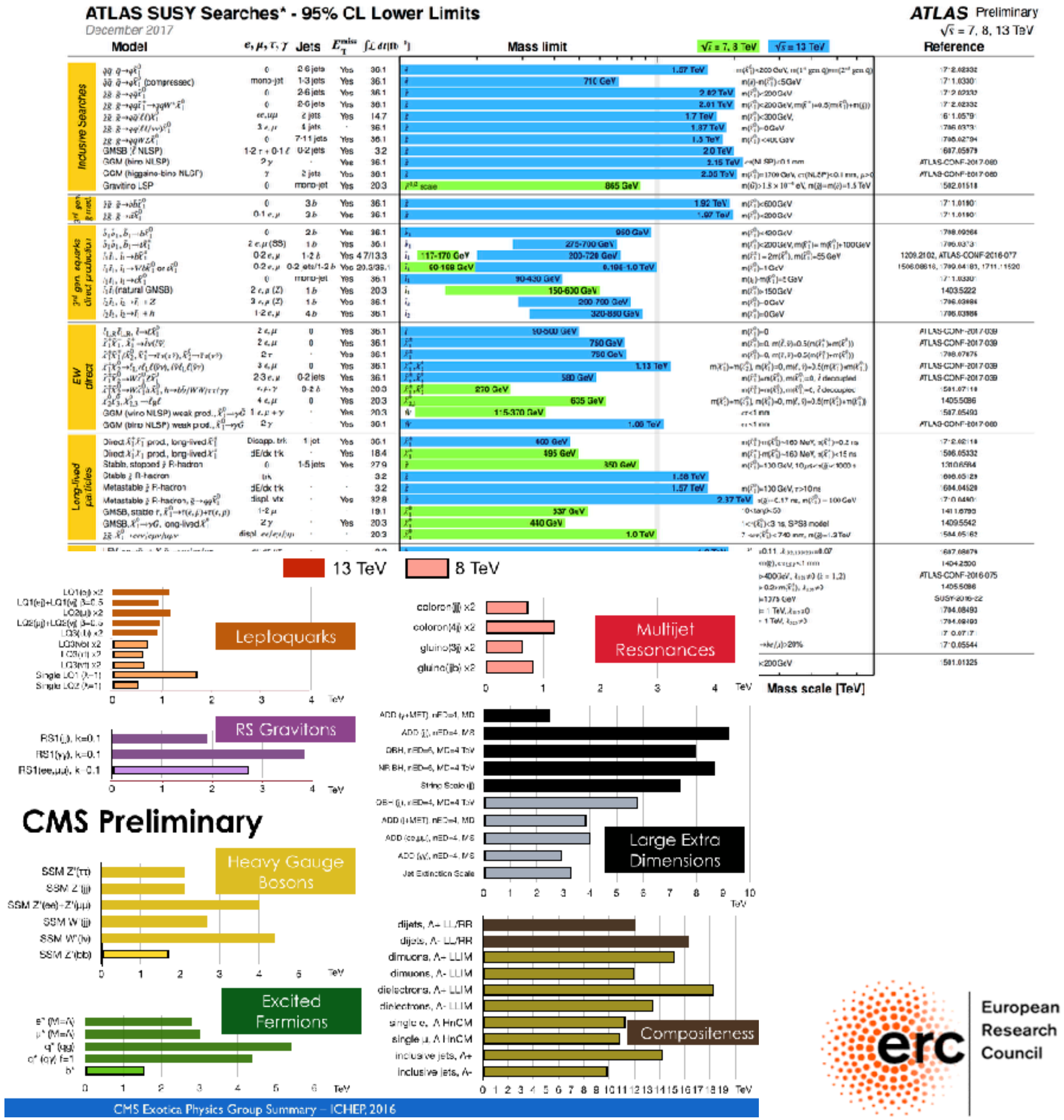
Physicists hoped to find evidence of supersymmetry in experiments at the Large Hadron Collider (LHC). To date, they have not. If no evidence arises in the next run of the LHC, supersymmetry will be in trouble. **The failure to find superpartners** is brewing a crisis in physics, forcing researchers to question assumptions from which they have been working for decades.

THE BIG PICTURE

European Research Council

EIGHT YEARS OF SEARCHES…

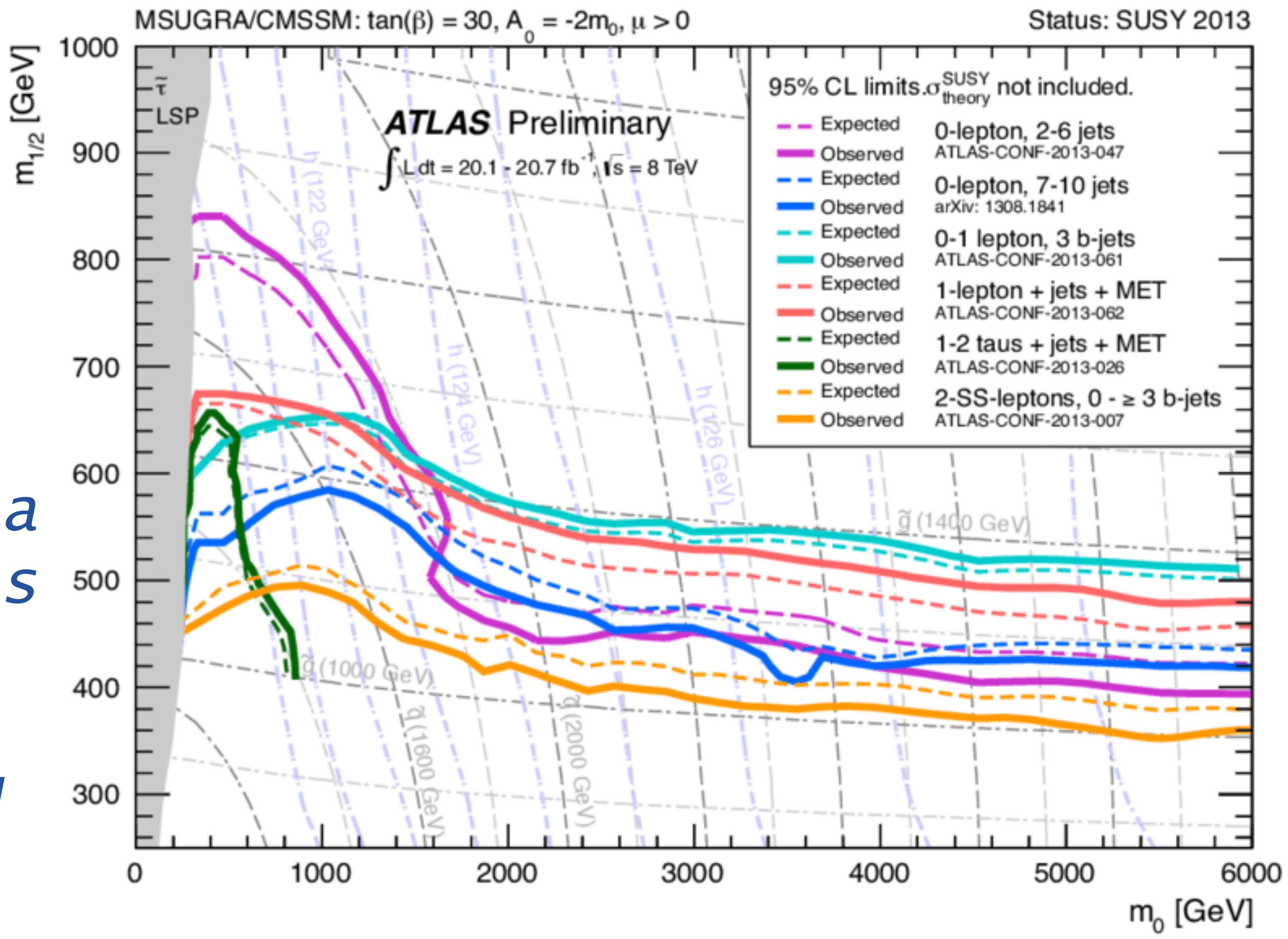
- *Extensively searched in all possible directions*
- *Done more than expected, with new ideas and original approaches to data taking*
- *Our new-physics target evolved towards more complicated scenarios*
  - *EXAMPLE: the SUSY we search for today is very different than what is in the ATLAS/CMS TDRs*



European Research Council

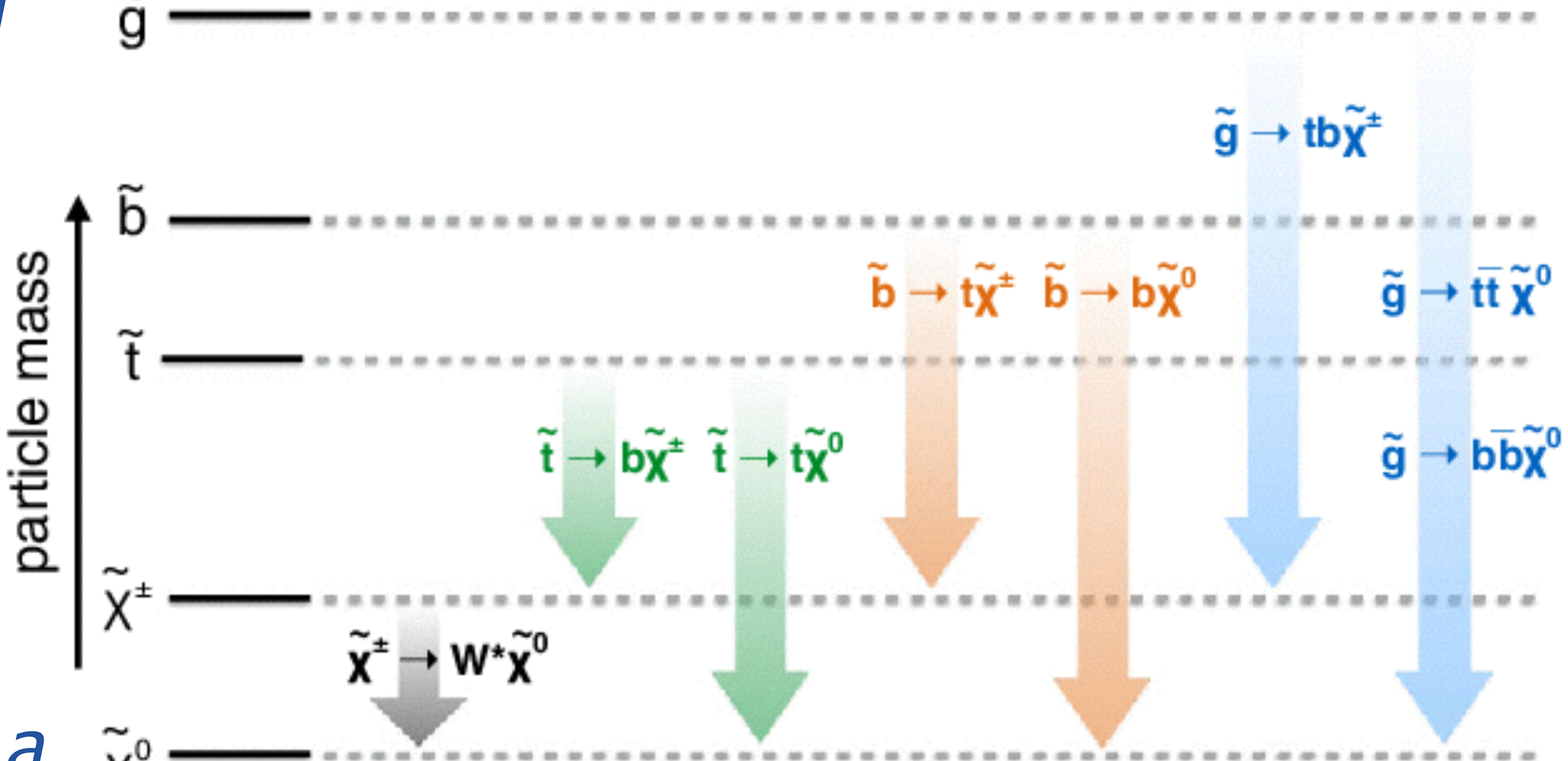
SUSY: A MOVING TARGET

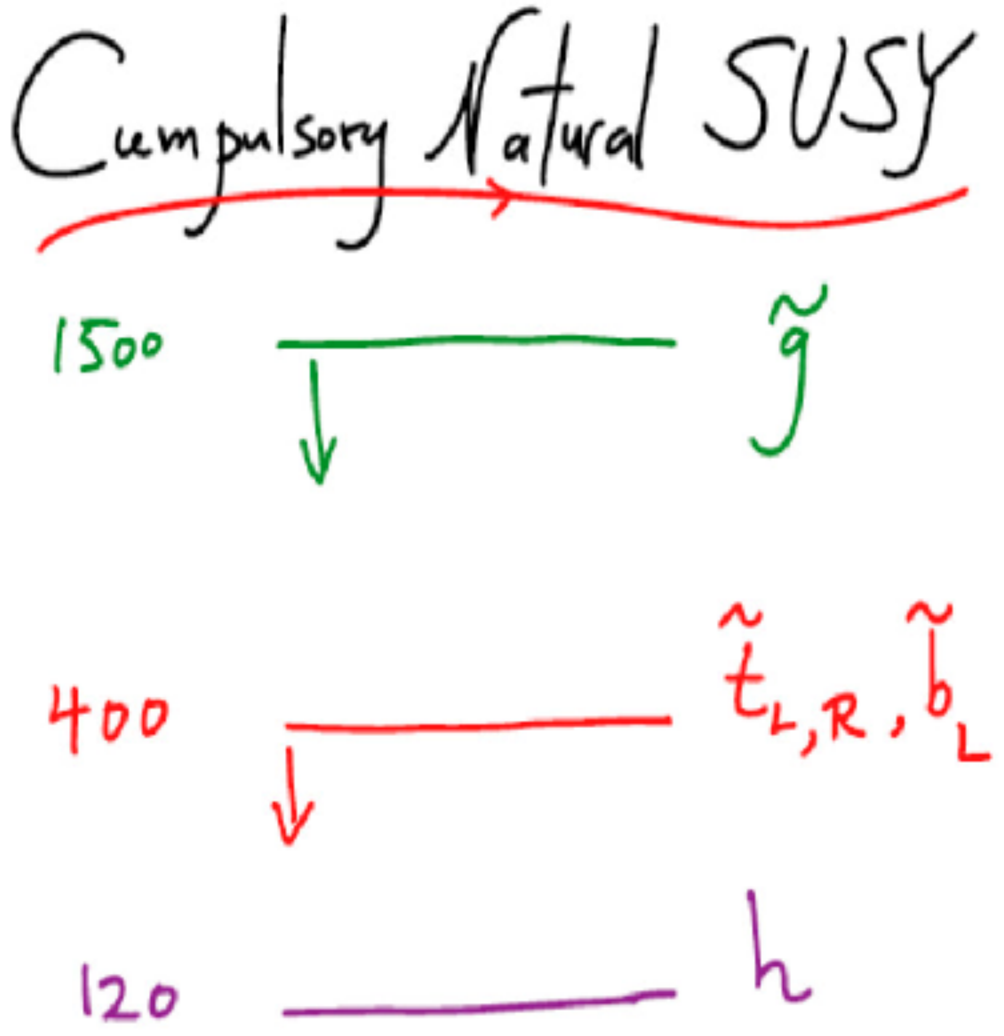
- *We started looking for mSugra-inspired models. Thanks to large gluino and squark cross sections, exclusions became soon very strong*
- *We then moved to Natural-SUSY scenarios, with focus on t and b squarks*
- *We moved to simplified models as a generalization of search strategies (with 100% BR assumptions)*
- *We recently generalized simplified models to BR-independent results*
- *And we extended model interpretation to large-dimensional scans (pMSSM)*



SUSY: A MOVING TARGET

- *We started looking for mSugra-inspired models. Thanks to large gluino and squark cross sections, exclusions became soon very strong*
- ***We then moved to Natural-SUSY scenarios, with focus on t and b squarks***
- *We moved to simplified models as a generalization of search strategies (with 100% BR assumptions)*
- *We recently generalized simplified models to BR-independent results*
- *And we extended model interpretation to large-dimensional scans (pMSSM)*





European Research Council

SUSY: A MOVING TARGET

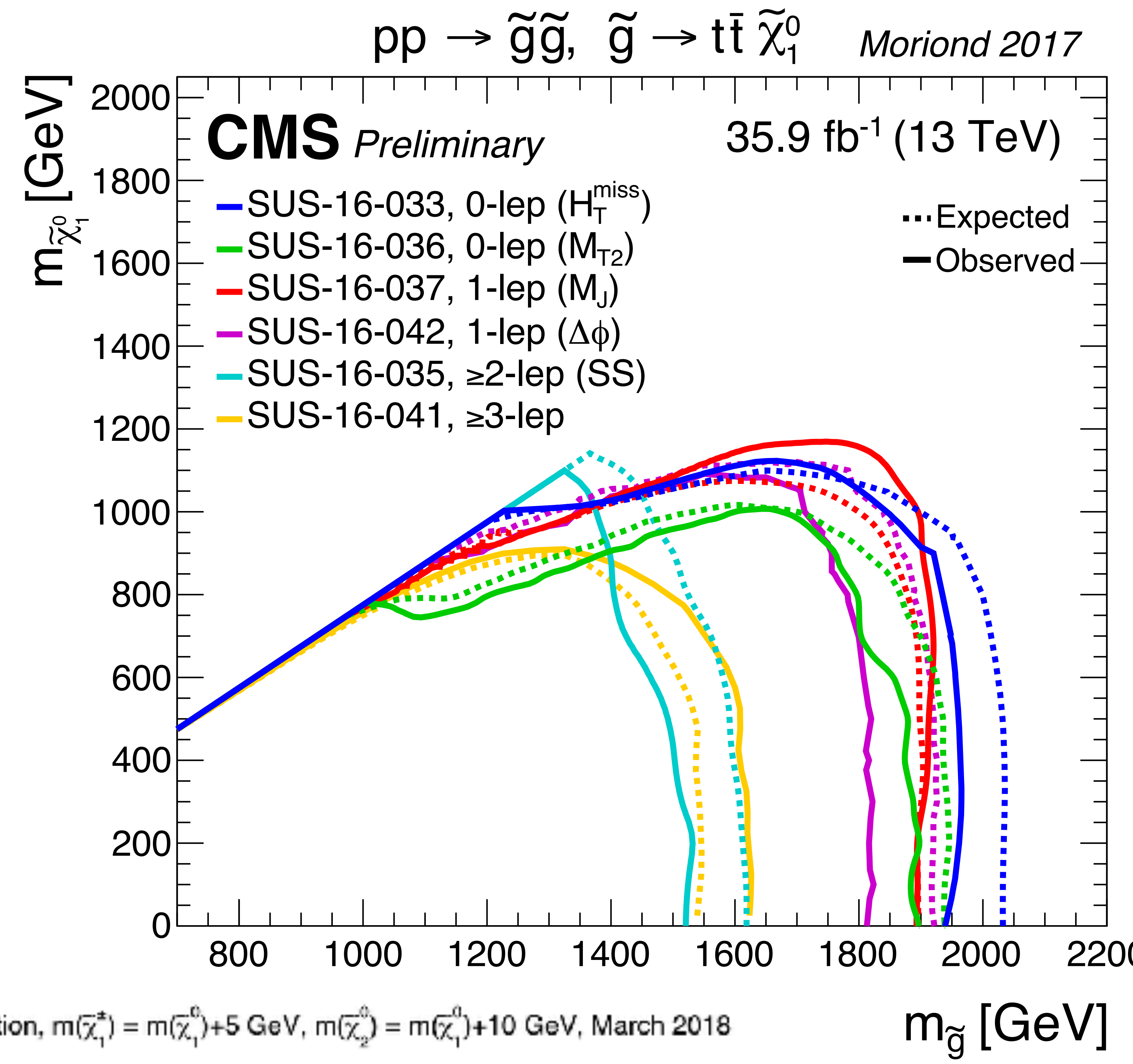
- *We started looking for mSugra-inspired models. Thanks to large gluino and squark cross sections, exclusions became soon very strong*

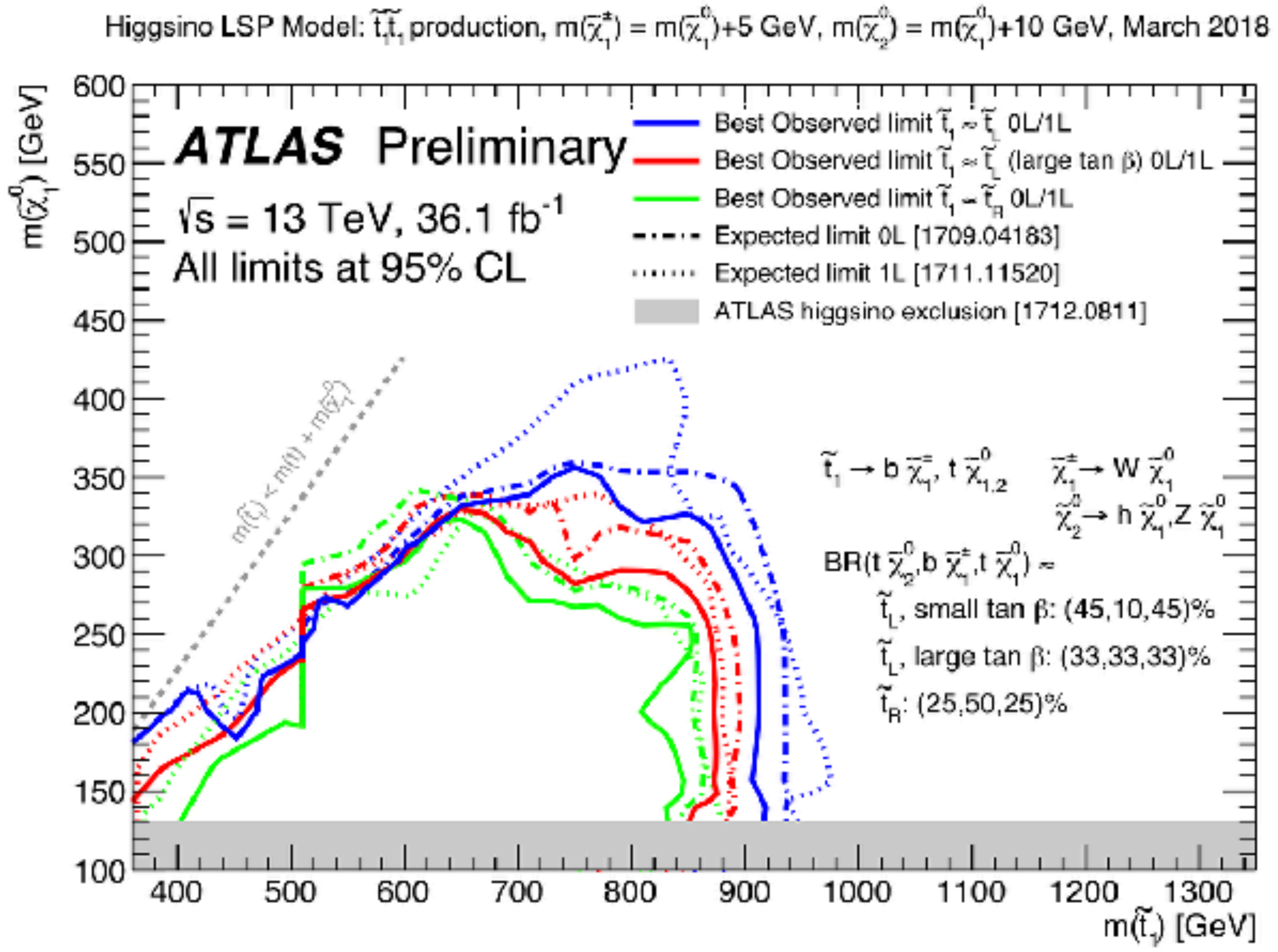
- *We then moved to Natural-SUSY scenarios, with focus on t and b squarks*

- ***We moved to simplified models as a generalization of search strategies (with 100% BR assumptions)***

- *We recently generalized simplified models to BR-independent results*

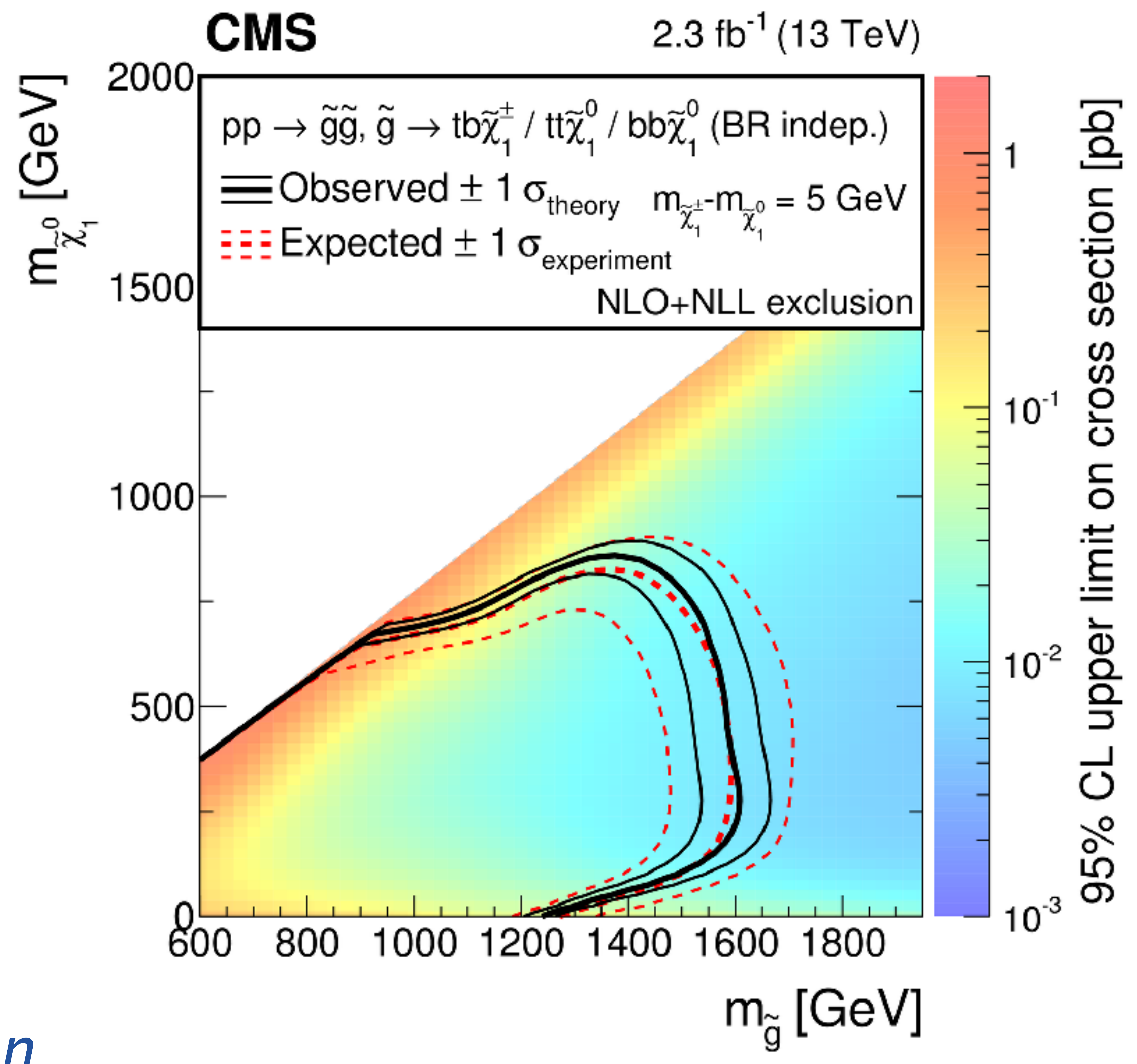
- *And we extended model interpretation to large-dimensional scans (pMSSM)*





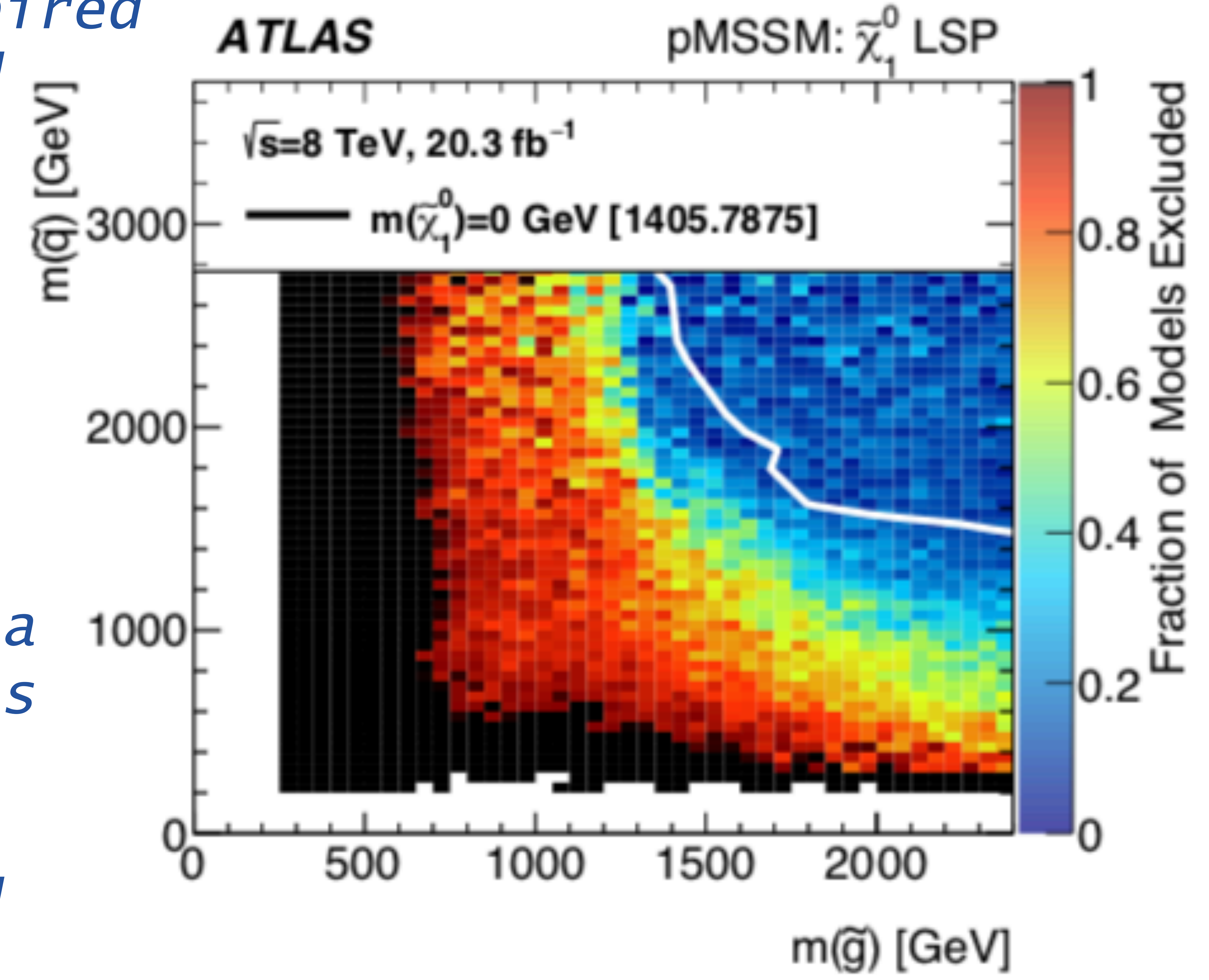
SUSY: A MOVING TARGET

- *We started looking for mSugra-inspired models. Thanks to large gluino and squark cross sections, exclusions became soon very strong*
- *We then moved to Natural-SUSY scenarios, with focus on t and b squarks*
- *We moved to simplified models as a generalization of search strategies (with 100% BR assumptions)*
- ***We recently generalized simplified models to BR-independent results***
- *And we extended model interpretation to large-dimensional scans (pMSSM)*



SUSY: A MOVING TARGET

- *We started looking for mSugra-inspired models. Thanks to large gluino and squark cross sections, exclusions became soon very strong*
- *We then moved to Natural-SUSY scenarios, with focus on t and b squarks*
- *We moved to simplified models as a generalization of search strategies (with 100% BR assumptions)*
- *We recently generalized simplified models to BR-independent results*
- ***And we extended model interpretation to large-dimensional scans (pMSSM)***

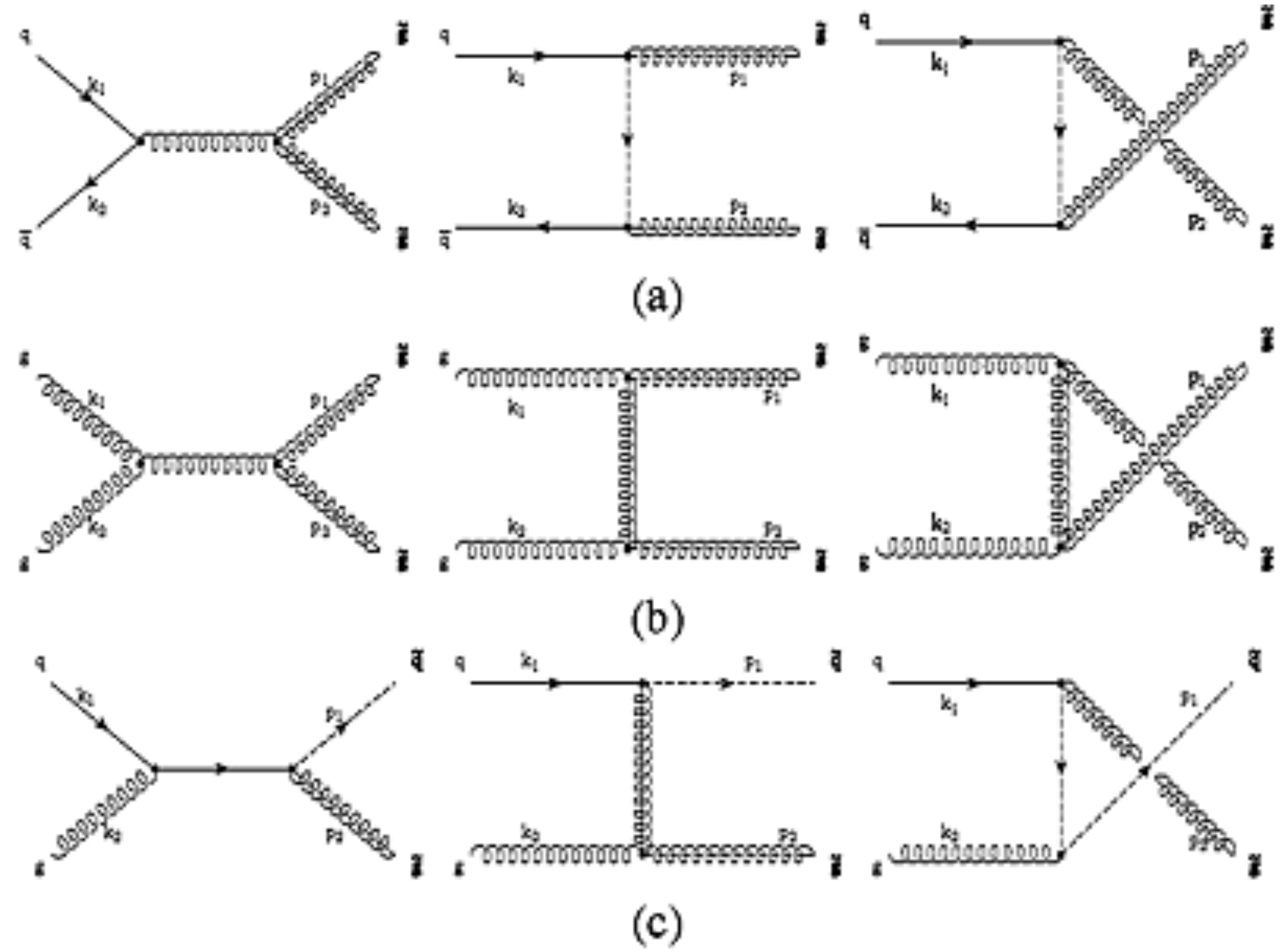


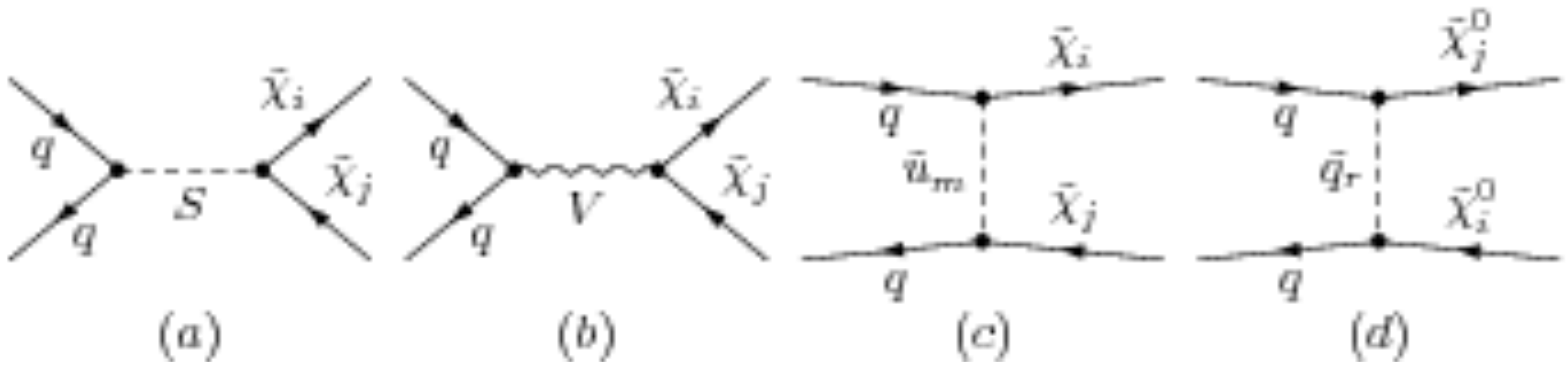
(a) All LSP types


CHECK LIST FOR
MY SUSY ANALYSIS
✓ SIGNATURE

SIGNAL: CHOOSE YOUR TARGET

- *LHC is a proton collider*
  - *whatever you want to produce needs to couple to quarks*
- *This is basically true for any MSSM particle*
- *Different production mechanisms contribute to determine the production cross section*

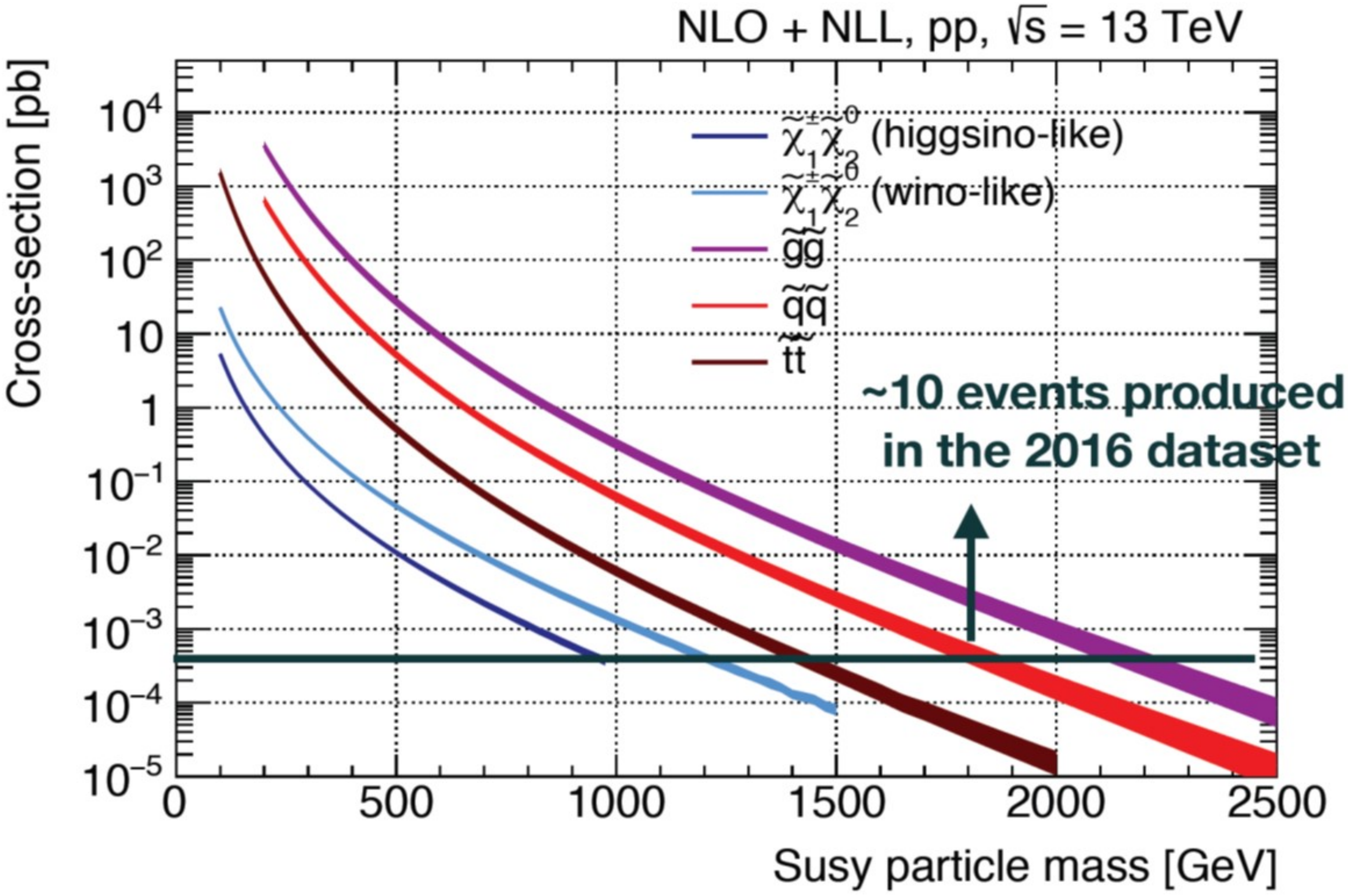




European Research Council

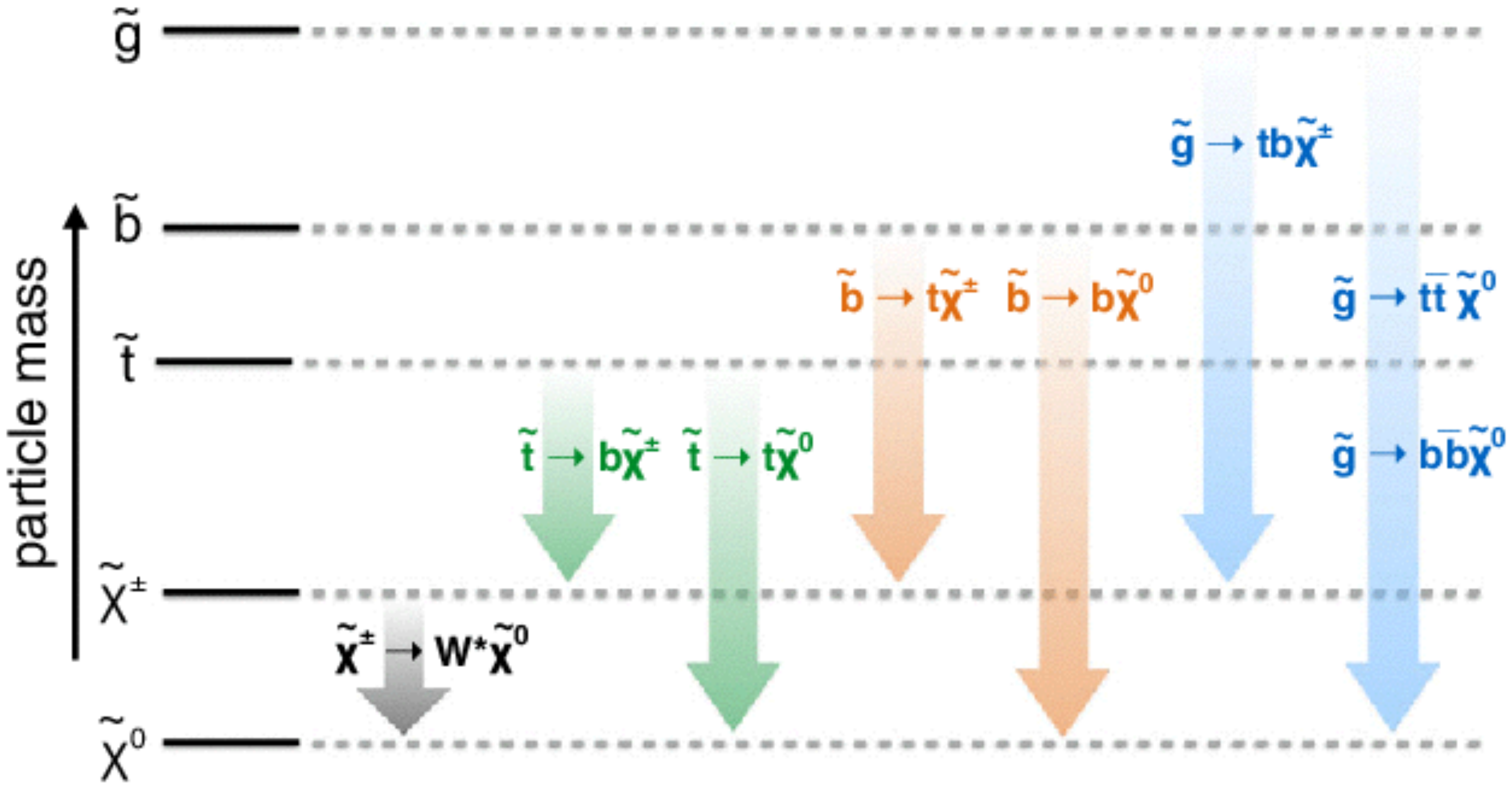
PRODUCTION CROSS SECTION

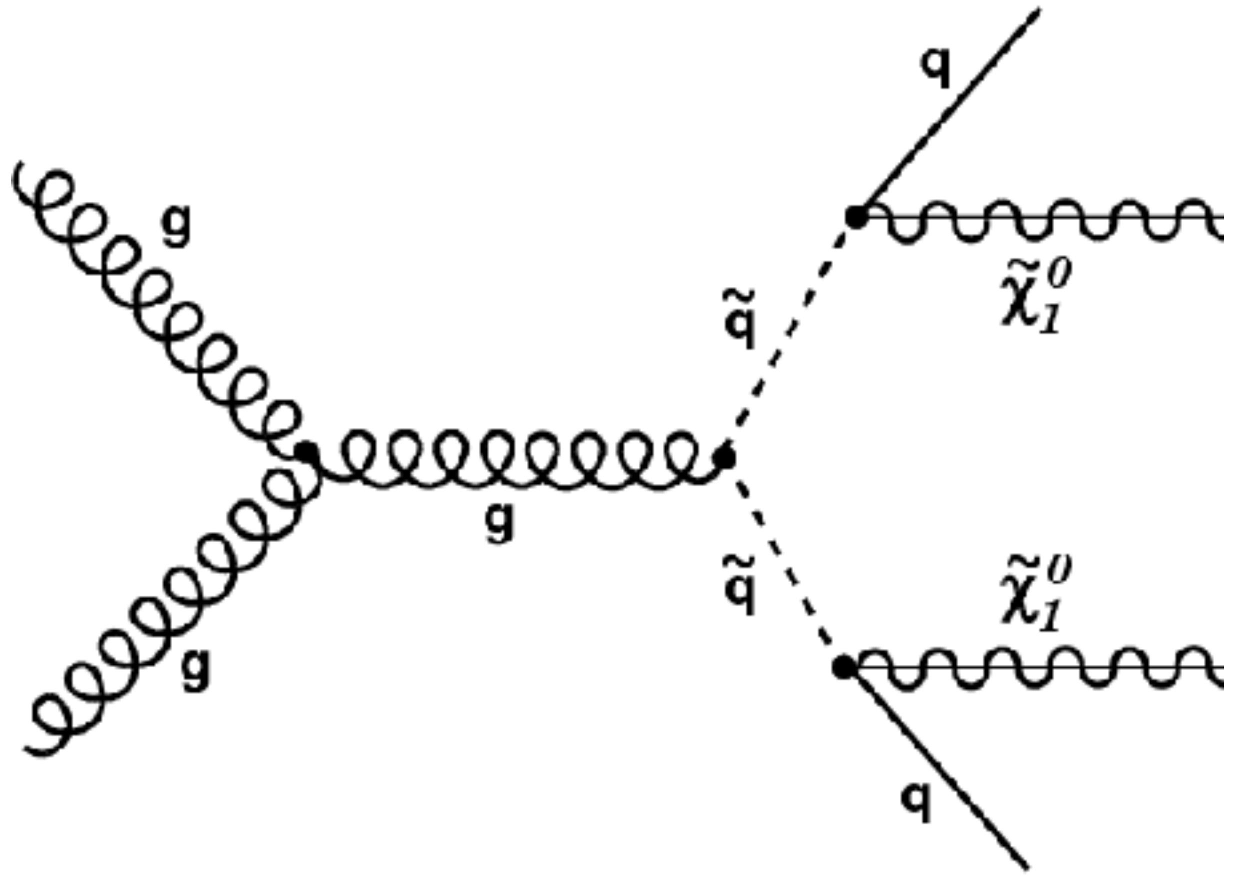
- *These mechanisms translate into a hierarchy of production cross sections*
  - *colored particles have larger cross sections*
  - *gluinos more than quarks, because of color enhancement*
  - *ewkinos have the smallest cross sections (but with a lot of data we are getting there)*



FINAL STATES

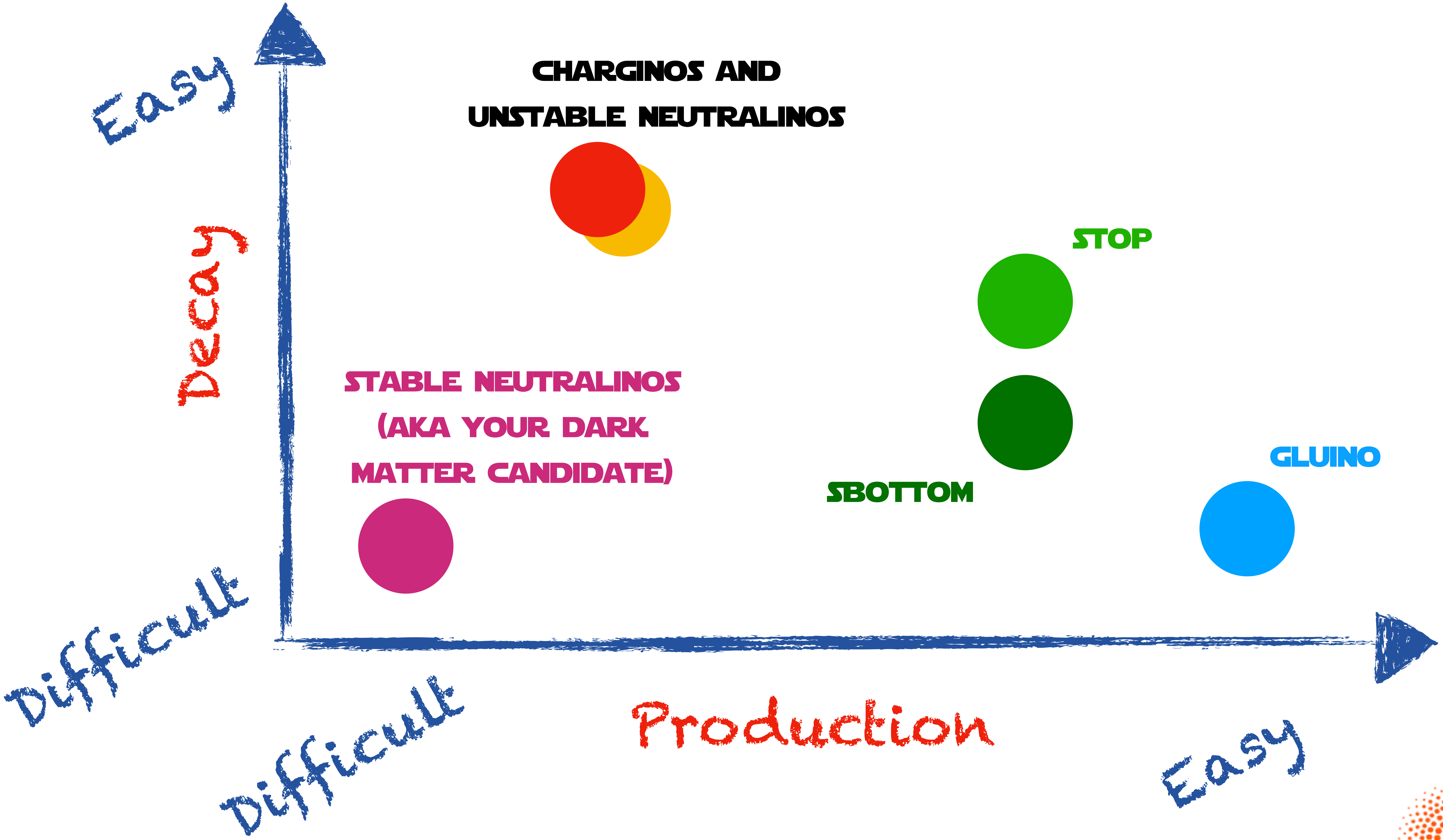
- Making SUSY particles is only part of the problem
- The next is the decay mode, starting with designing a trigger
  - ewkinos decay to leptons (via W/Z/H) or high-pT γ, which are rare in typical LHC collisions
  - squarks and gluons make jets, to which there is a large background
  - Among the squarks, stop and sbottom are better to handle, because they come with b-jets and sometimes with W bosons





European Research Council

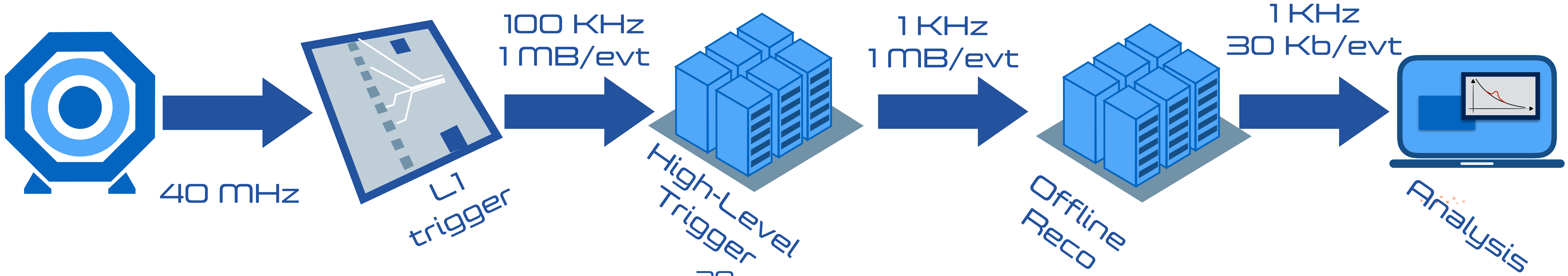
PRODUCTION VS DECAY



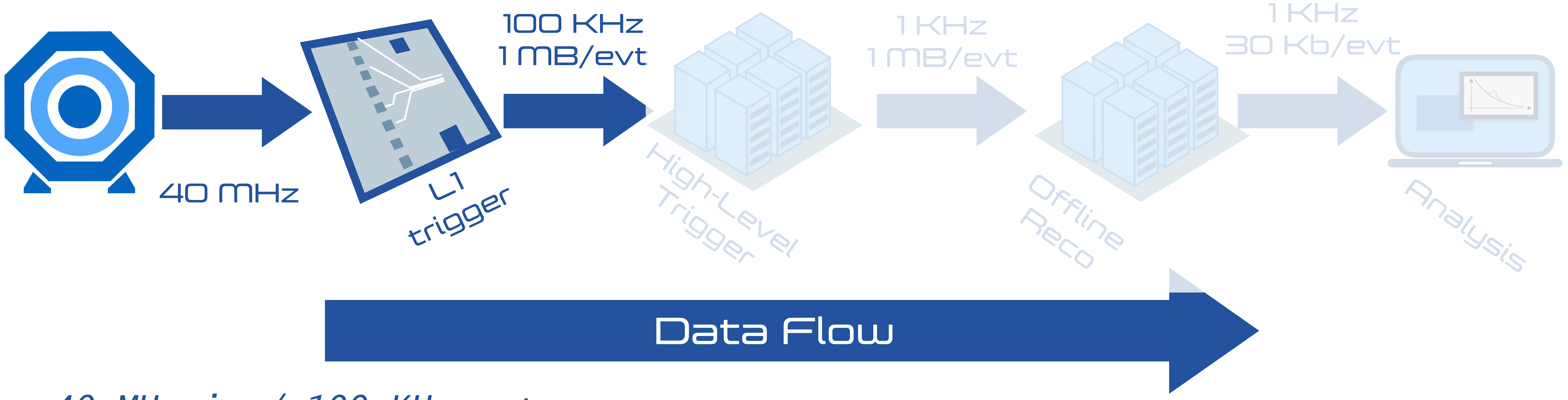

CHECK LIST FOR
MY SUSY ANALYSIS
✓ SIGNATURE
✓ TRIGGER

THE LHC DATA FLOW

- *Too many data, too large data -> need to filter online*
- *Selection done in two stages of reconstruction*
- *Accuracy increases with rate reduction and consequent latency increase*
- *Three main domains*
    - *Online selection (trigger)*
    - *Offline central reconstruction*
    - *Offline selection + data analysis*

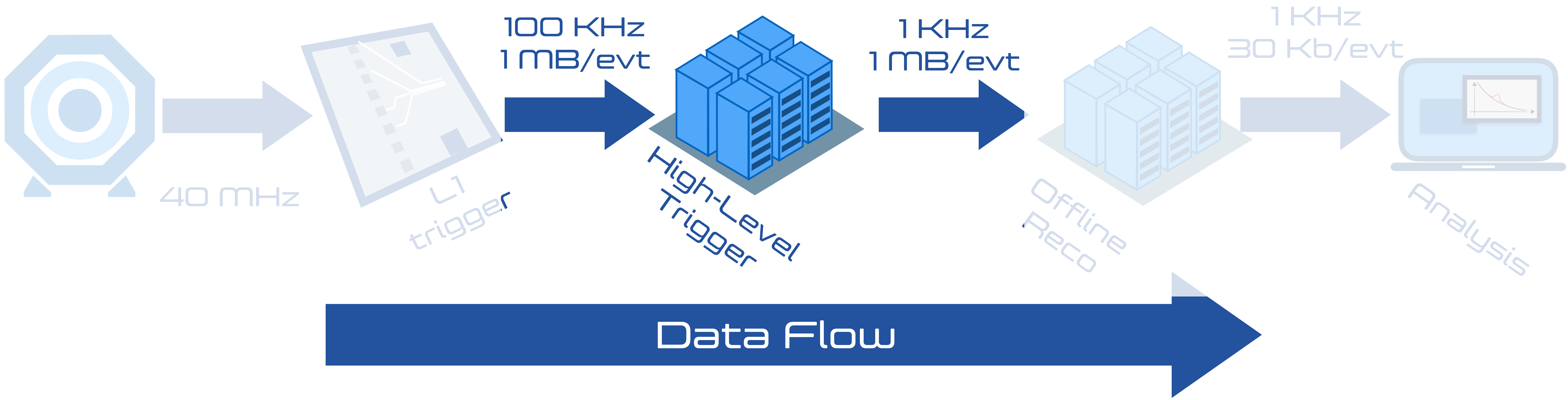


THE L1 TRIGGER



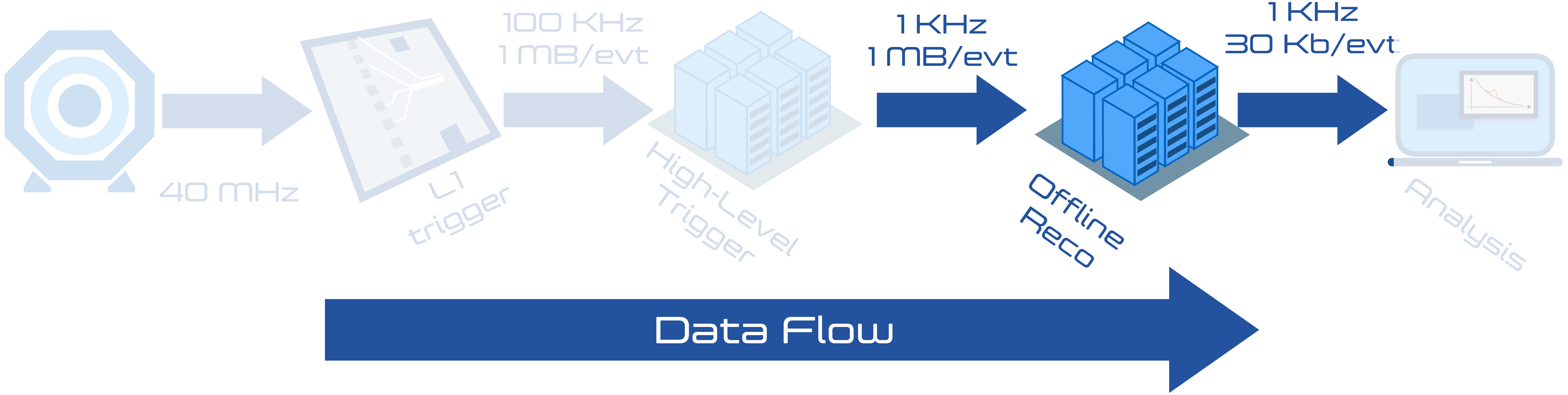
- 40 MHz in / 100 KHz out
- ~ 500 KB / event
- Processing time: ~10 μs
- Based on coarse local reconstructions
- Not all detectors available (e.g., no tracking in Run I/II)
- FPGAs / Hardware implemented

THE HLT TRIGGER



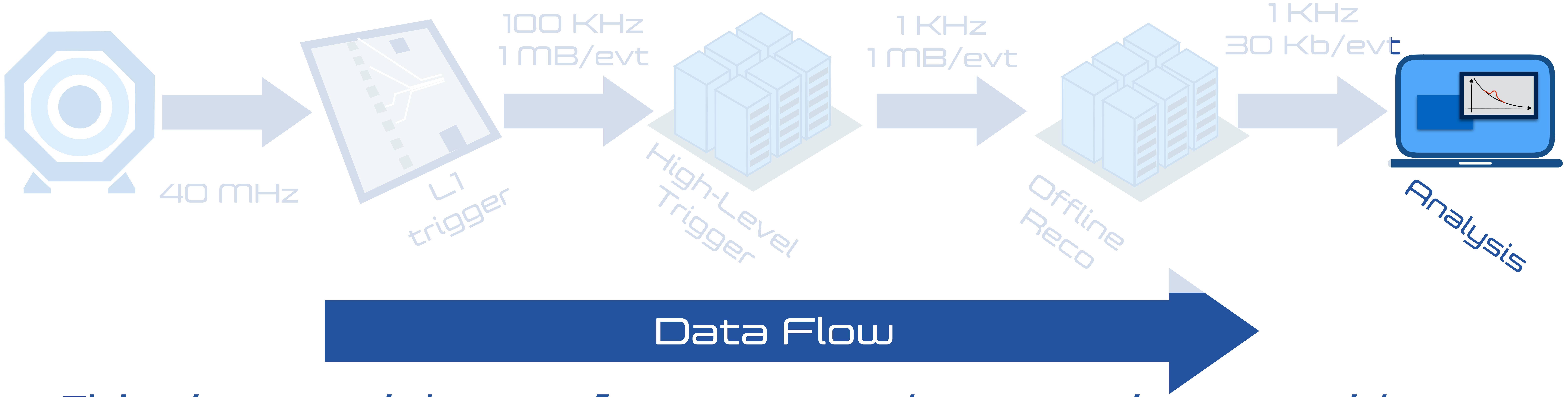
- *100 KHz in / 1 KHz out*
- *~ 500 KB / event*
- *Processing time: ~30 ms*
- *Based on simplified global reconstructions*
- *Software implemented on CPUs*

CENTRAL OFFLINE RECONSTRUCTION



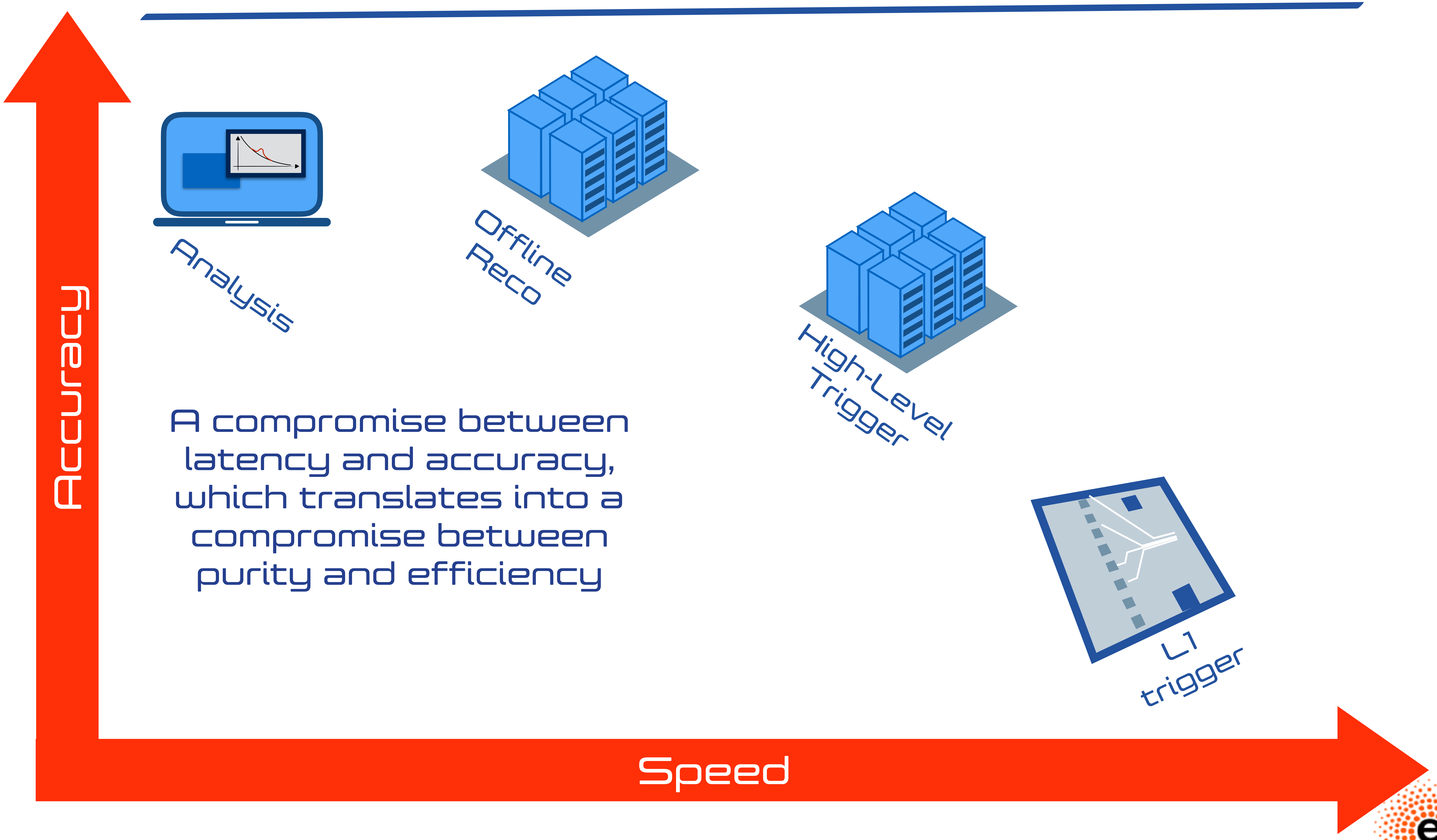
- *1 KHz in / 1.2 KHz out*
- *~ 1 MB reduced to 300/30/3 KB per event (AOD, miniAOD, nanoAOD)*
- *Processing time: tens of seconds to one minute*
- *Based on accurate global reconstructions*
- *Software implemented on CPUs*

OFFLINE ANALYSIS



- This is you with your laptop, an ssh connection, a grid certificate and/or some local computer far
- At this stage, anything is possible, with all the time you want (< postdoc contract/PhD duration)
- But you need to be sure that your interesting data made it up to here

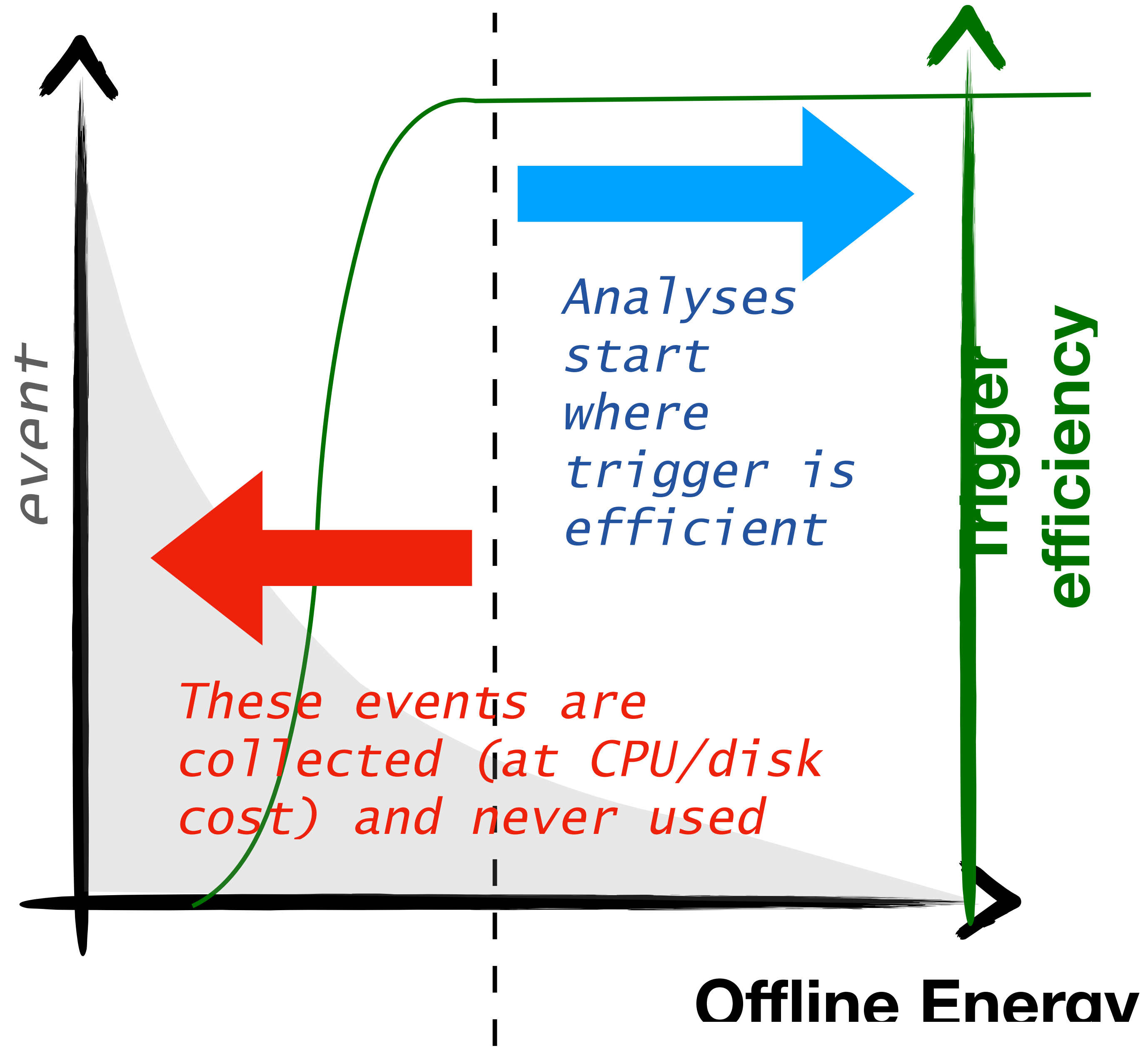
SPEED vs ACCURACY

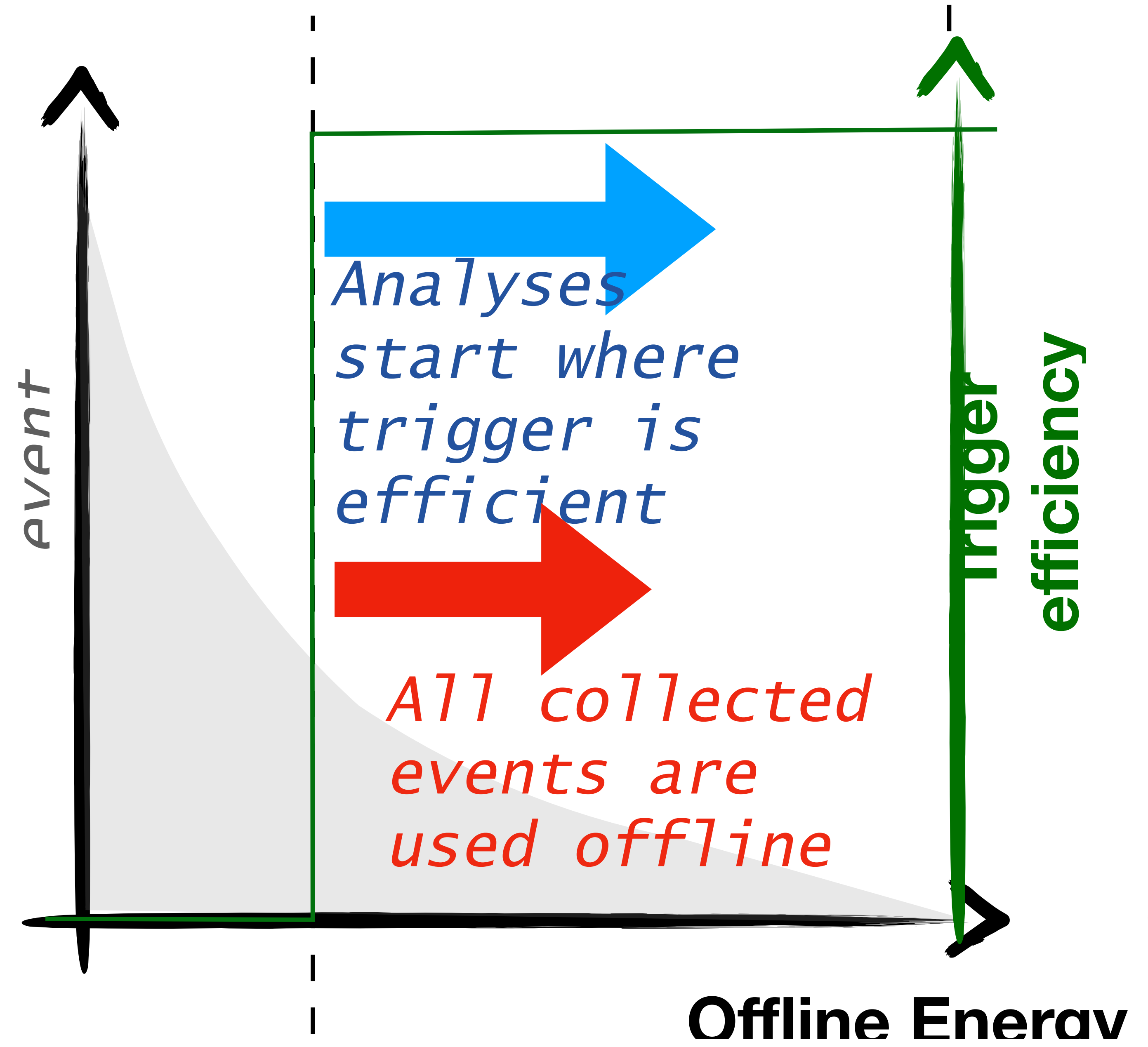


A compromise between latency and accuracy, which translates into a compromise between purity and efficiency

THE COST OF TRIGGERING

- For offline analysis, the online selection is some efficiency response function (aka turn-on curve)
- Usually, analyses work in the the regime where this function is flat
- Offline reco = online reco → sharp function: whatever you save is what you use
- Offline reco ≠online reco → slow raising function: many events are written for nothing (i.e., resources wasted)

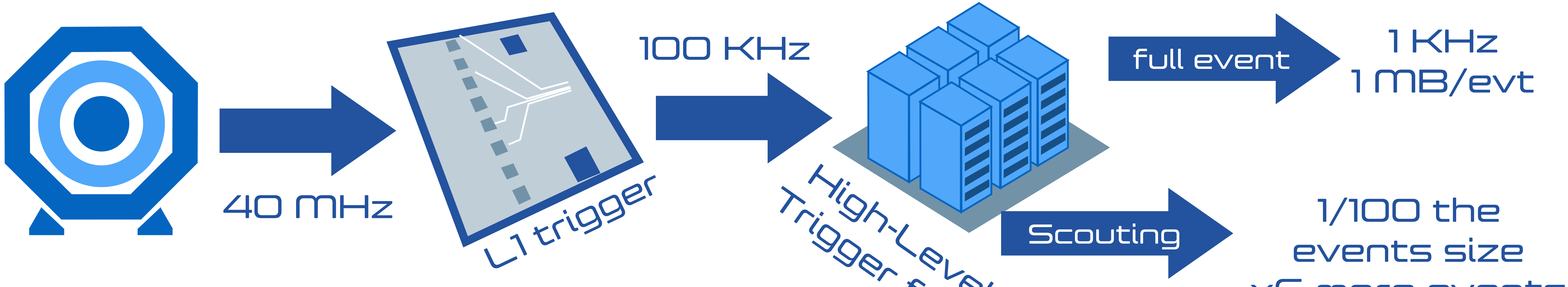




WHEN THE TRIGGER IS A PROBLEM

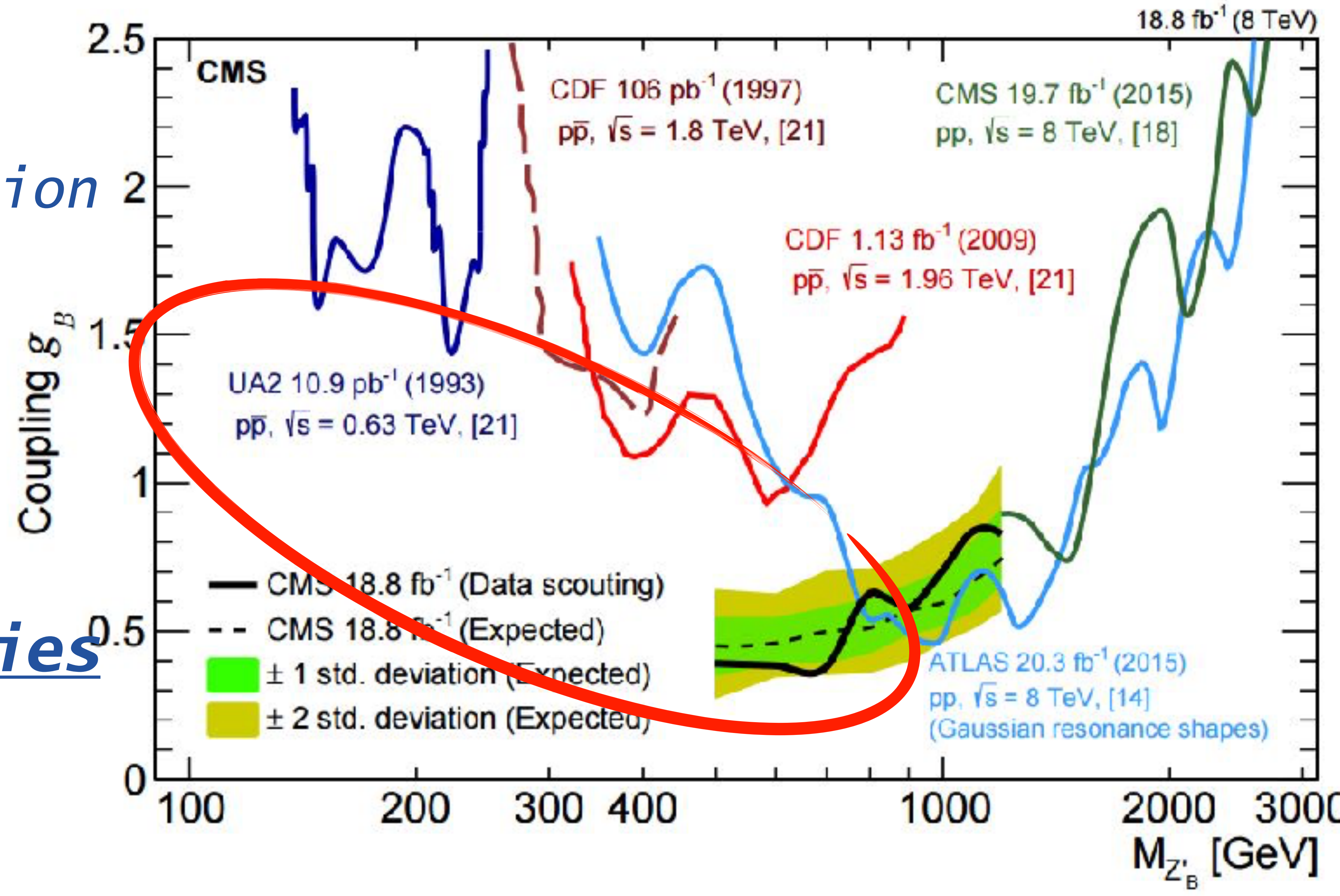
- *Whenever you don't have rare handles to trigger on (isolated & high-pT leptons or photons, large lepton multiplicity, a lot of missing E_T)*
- *When all you have is jets*
- *When your signature consists of moderate-pT objects*
- *When your signature consists of track-related features (displacement, stopping tracks, large dE/dx, etc)*
- *Trigger doesn't make things impossible. But can make things very inefficient*
  - *in which case, you have to wait for A LOT of data to probe certain scenarios*

THE WAY OUT: SCOUTING



Real-time new physics search with large datasets

- *Run reconstruction in the trigger farm*
- *Avoid resource limitations: write less information (a few floats) for more events*
- *Probes unexplored territory, previously left behind*

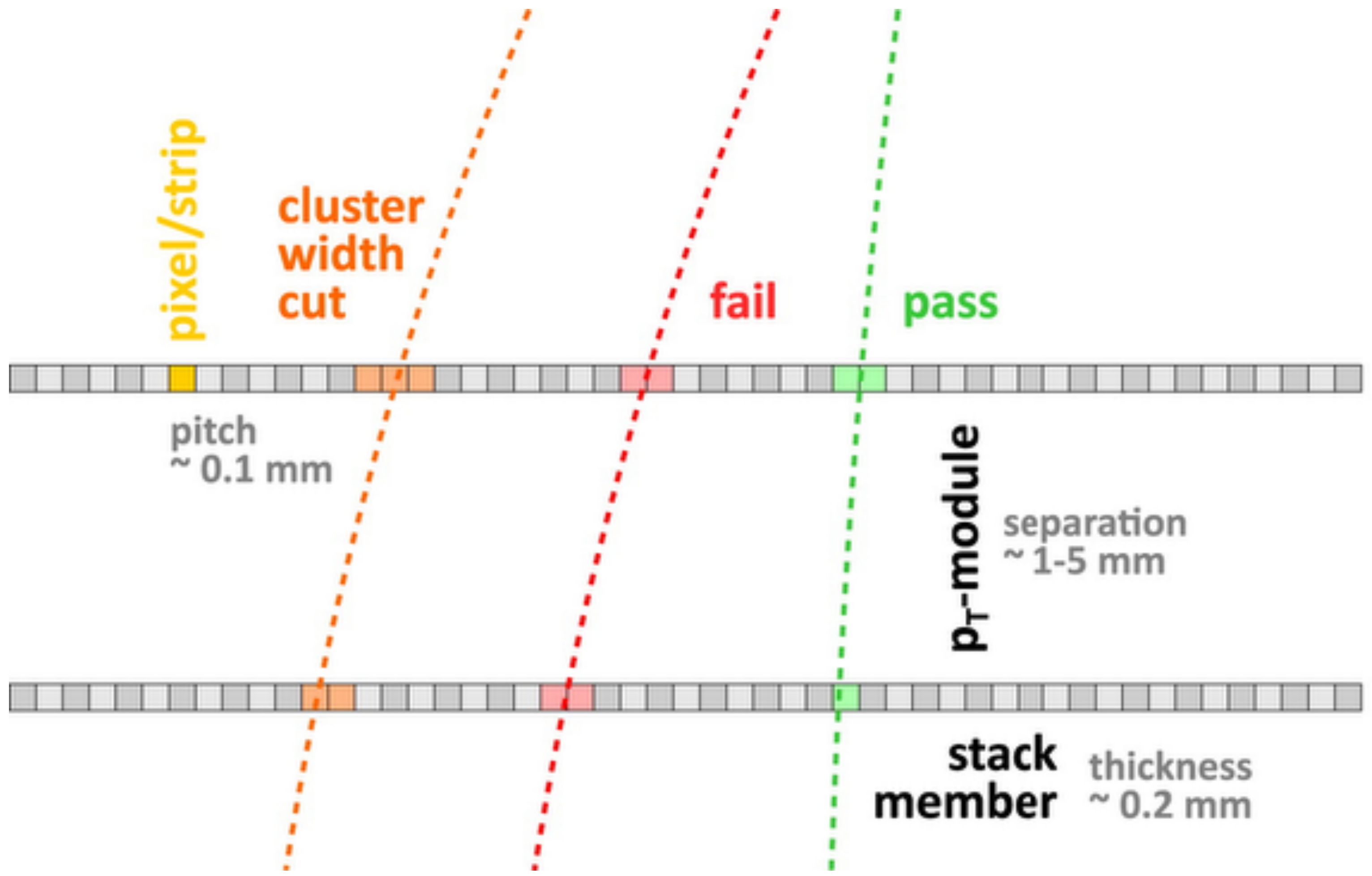


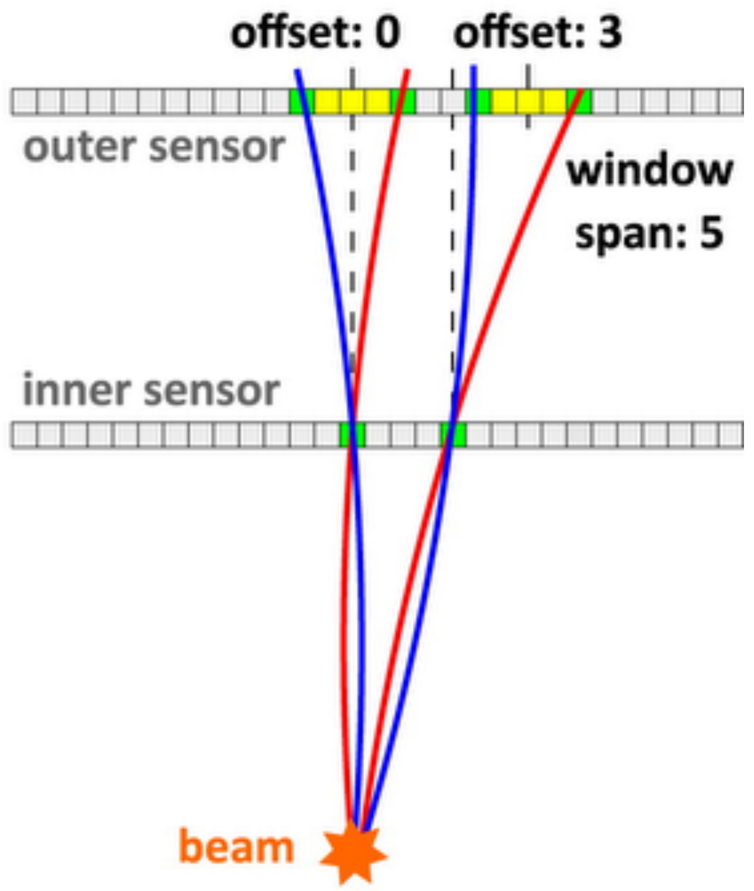
Problem: practical only for specific topologies

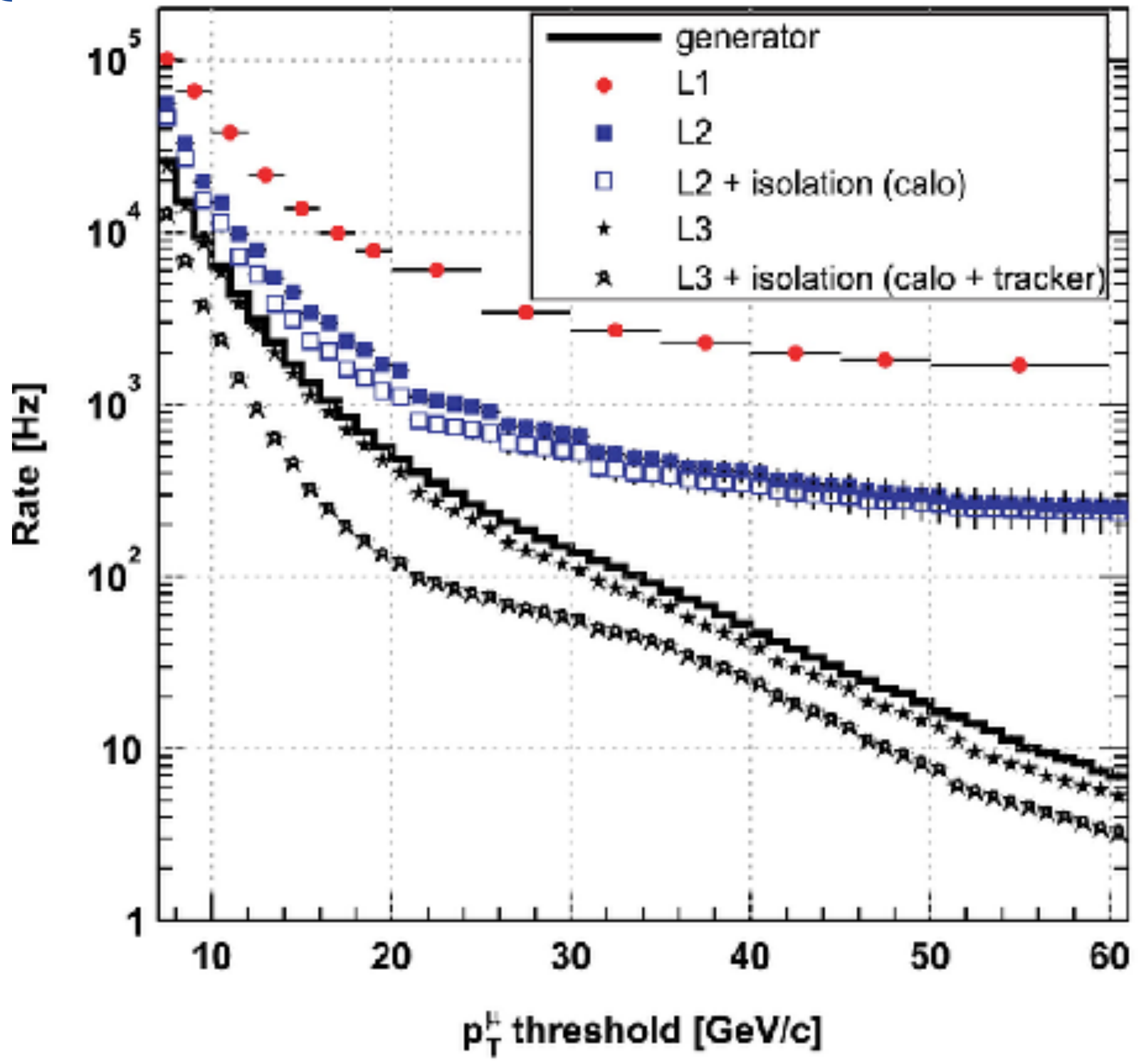
No application to SUSY (yet)

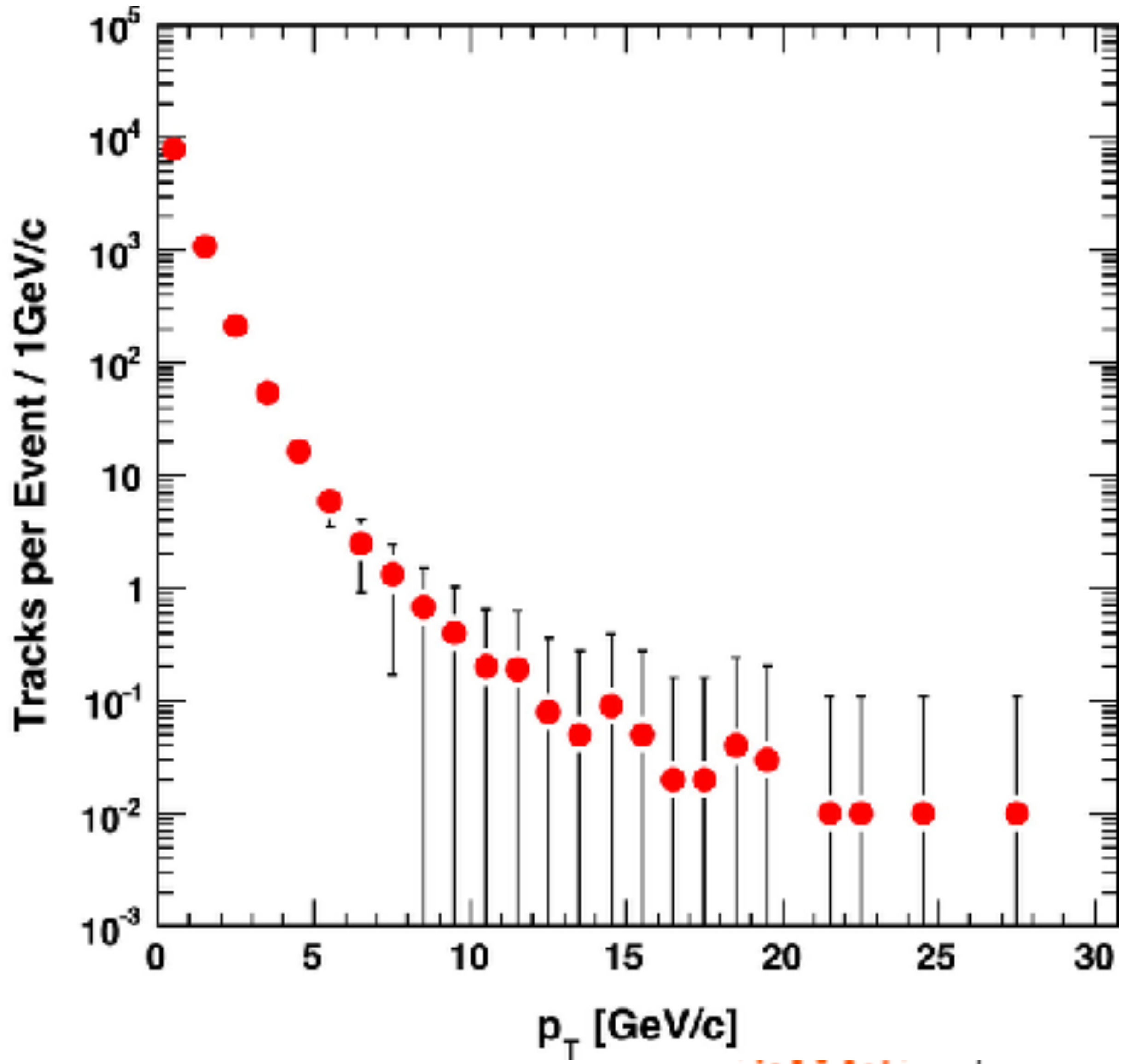
LIFE WILL BE EASIER

- *Major trigger upgrades are on the way*
  - *ATLAS introduced tracking capabilities at "L1.5"*
  - *Both experiments planning for tracking at 40 MHz for HL-LHC*
  - *Both experiments extending scouting usage*
  - *In RUN III, LHCb & ALICE moving to a real-time reconstruction system that will replace the HLT (i.e., scouting as a default)*





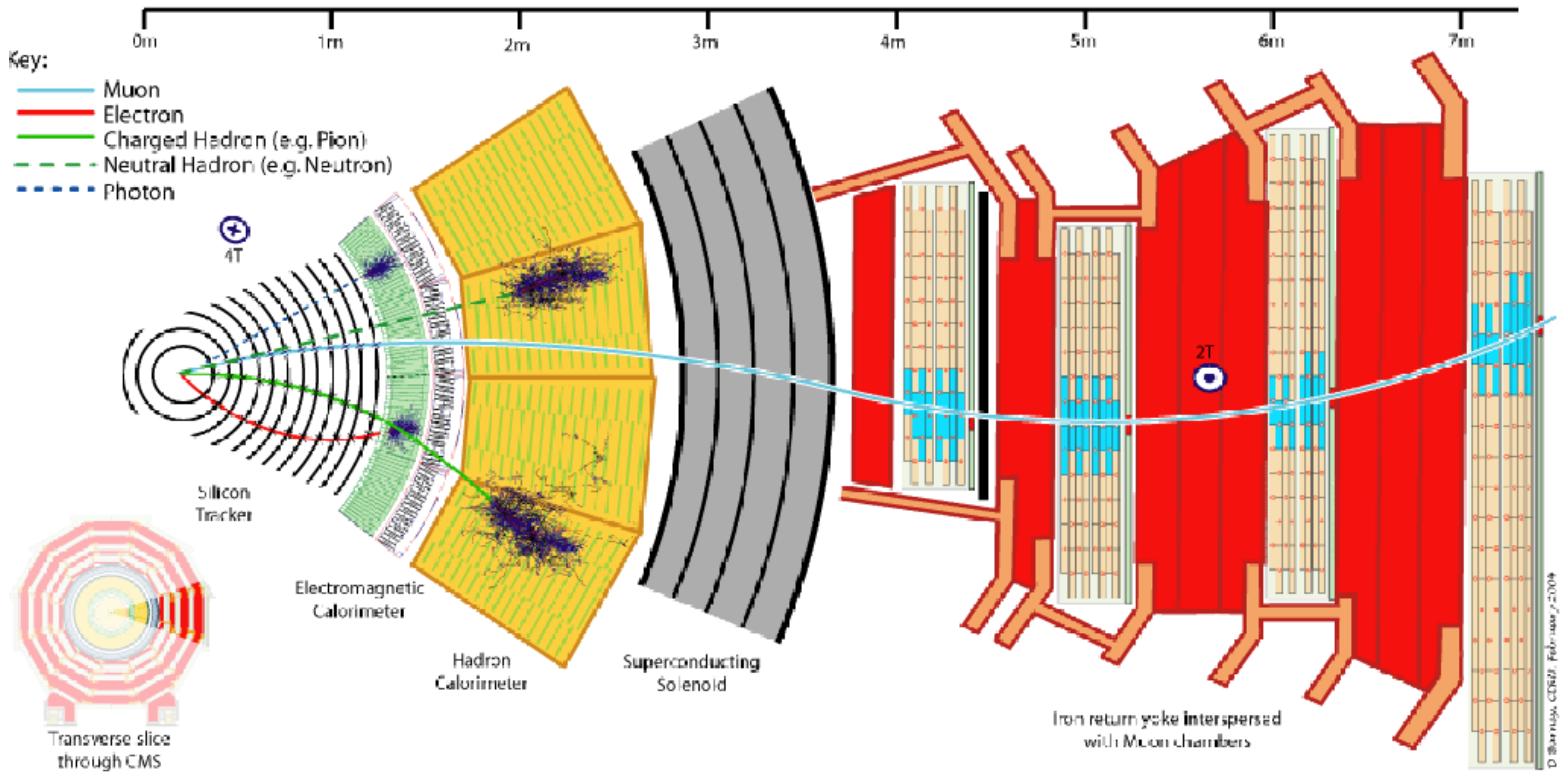


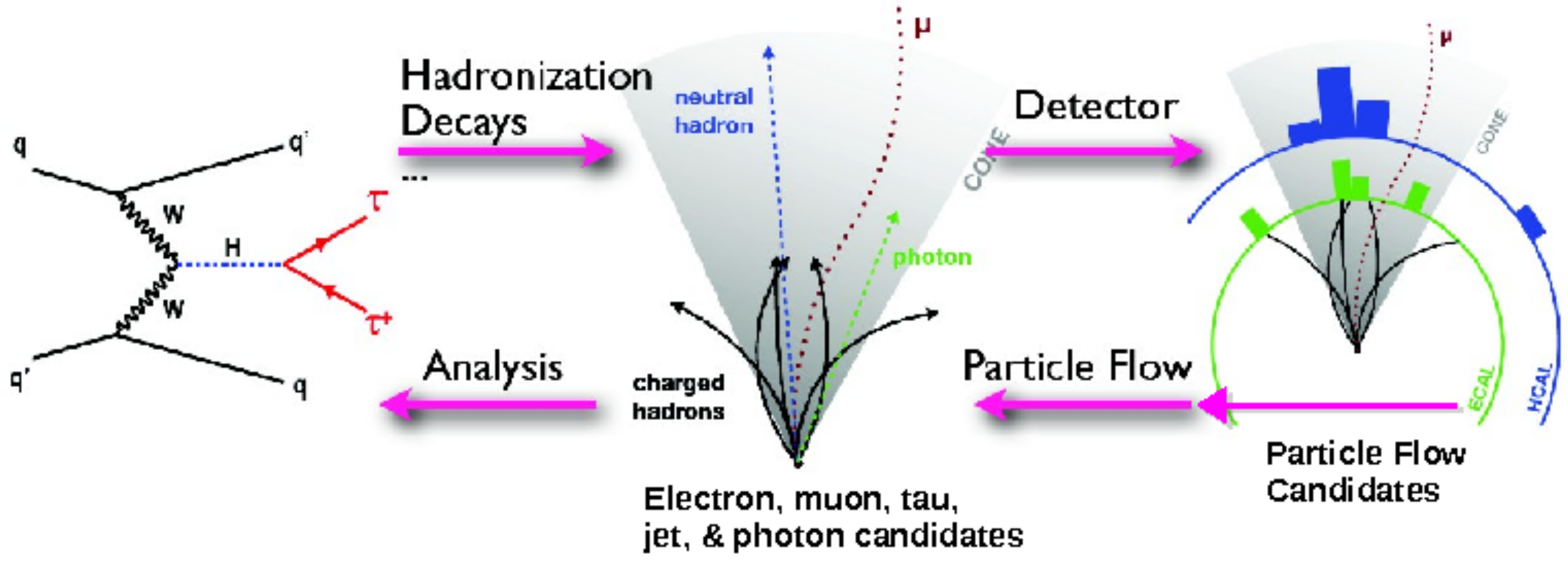



CHECK LIST FOR
MY SUSY ANALYSIS
✓ SIGNATURE
✓ TRIGGER
✓ RECONSTRUCTION

WHAT DO WE SEE IN THE DETECTORS

- Detectors are designed so that most of the particle is detected by at least one detector
- For jets, you care about the collective objects and not the individual constituents
  - you can use the energy deposits in calorimeters (standard reco)
  - you can first reconstruct the individual particles and then cluster them (particle flow)

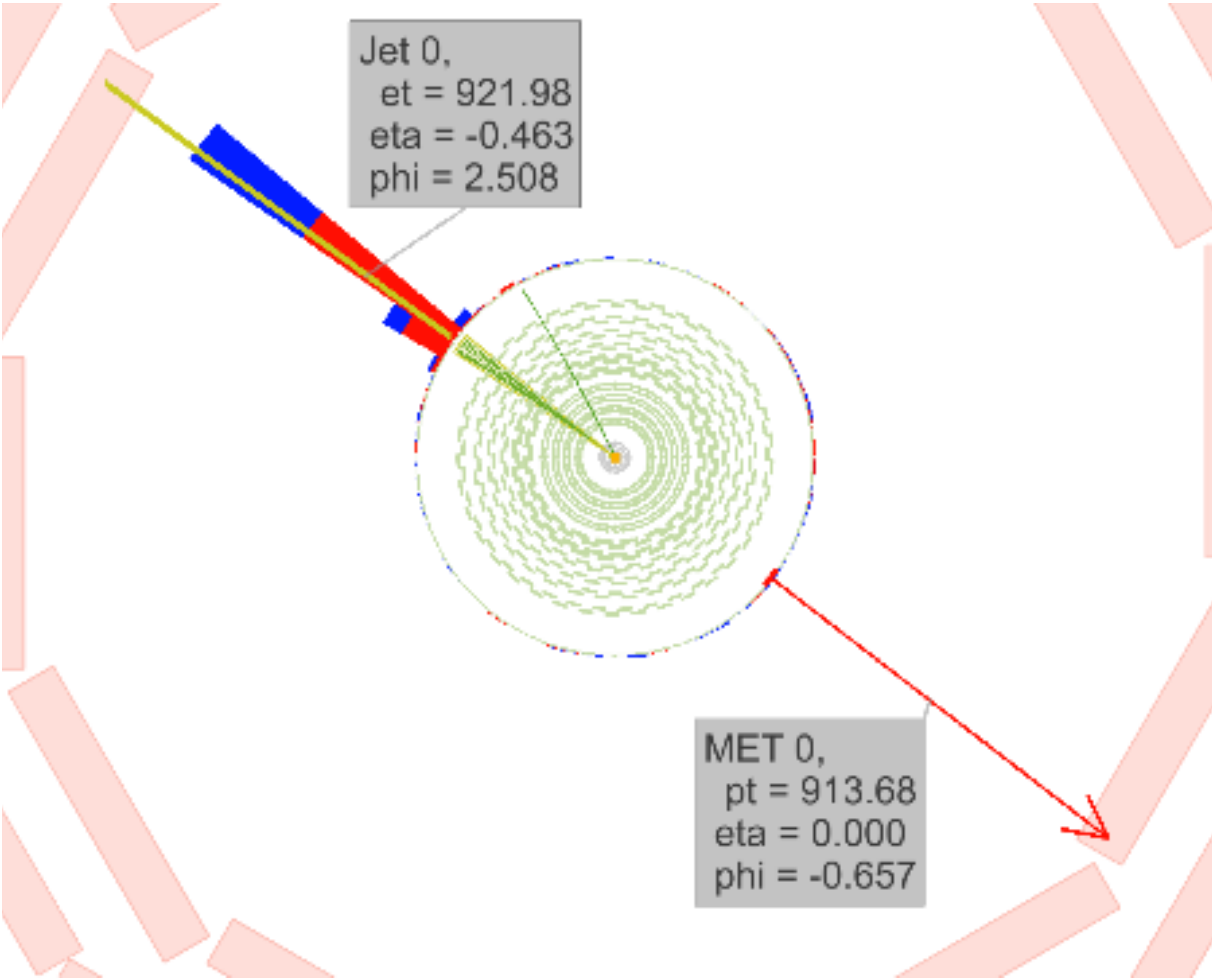




DETECTING INVISIBLES

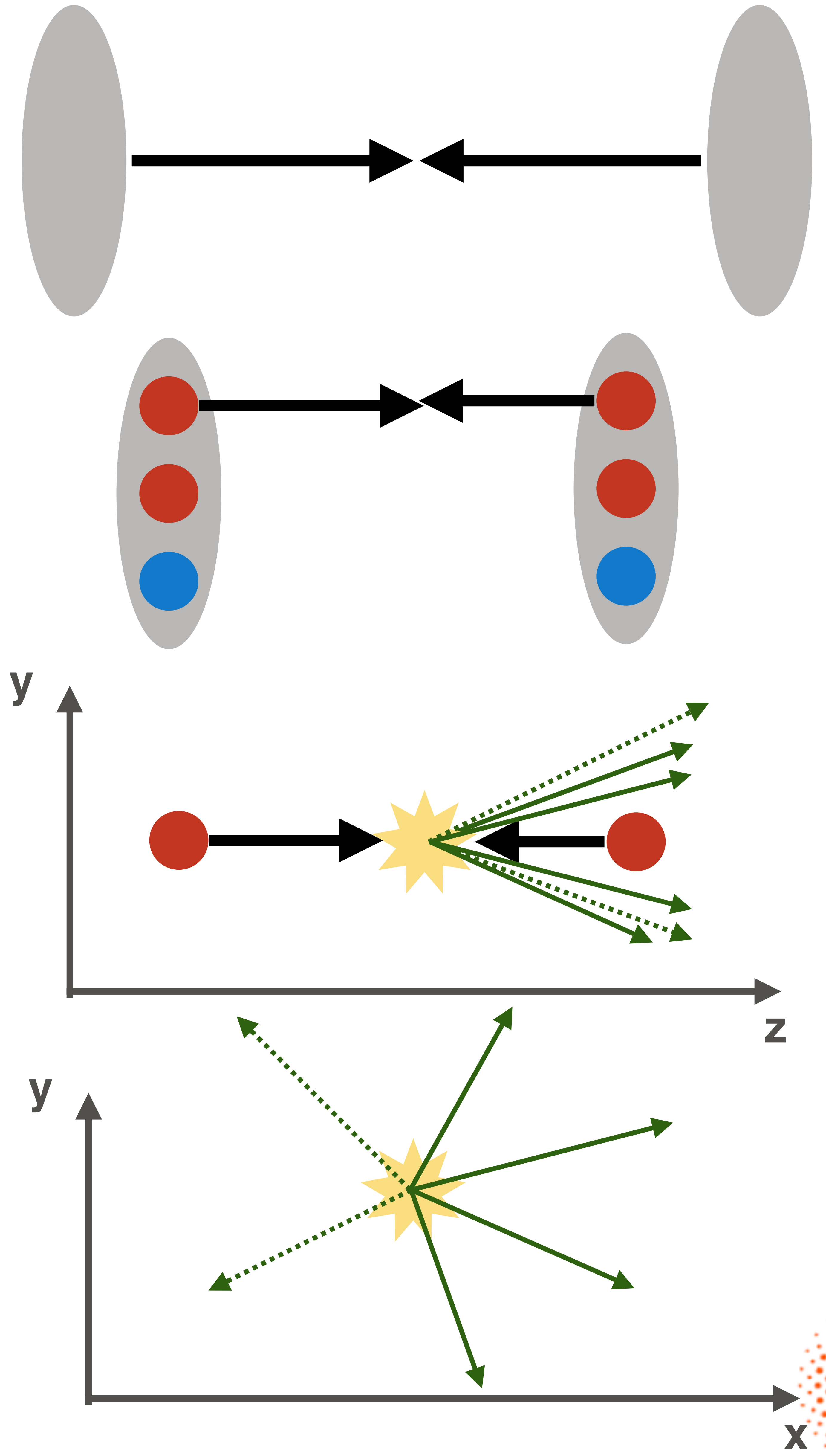
- Neutrinos other neutral stable particles (dark matter etc) don't interact in the detector
- Their collective presence can be detected measuring the missing transverse energy

$$E_T^{miss} = | - \Sigma_i \vec{p}_t^{\,i} |$$



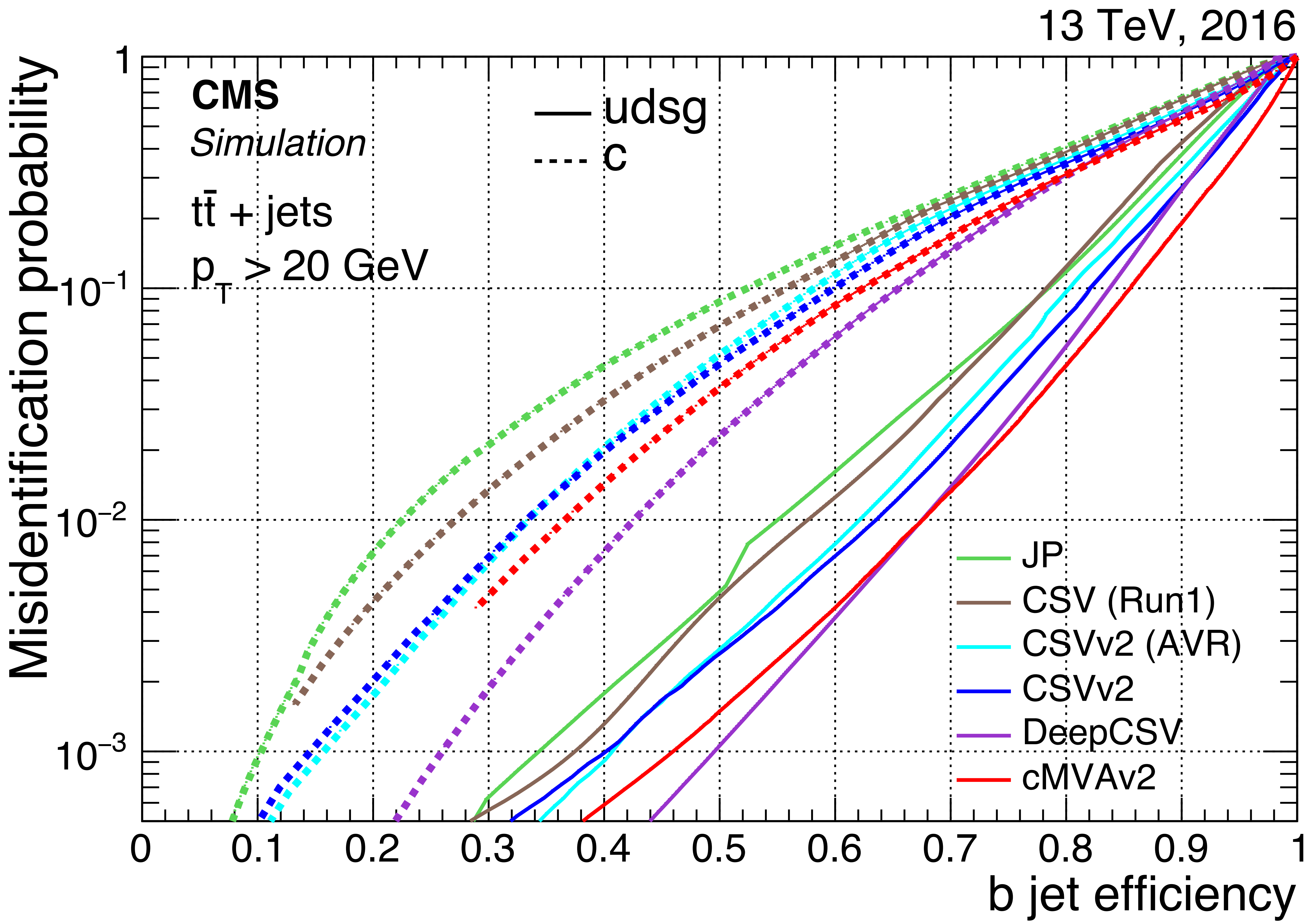
MISSING TRANSVERSE ENERGY

- *Two protons with same energy collide*
- *Actually, the collision is between quarks/gluons in the proton. They carry different fractions of the proton momentum*
- *As a result, there is a momentum imbalance ~ along the beam axes, but not in the transverse plane*
- *Transverse momenta should then balance. If some particle escaped undetected, the balance will be broken*



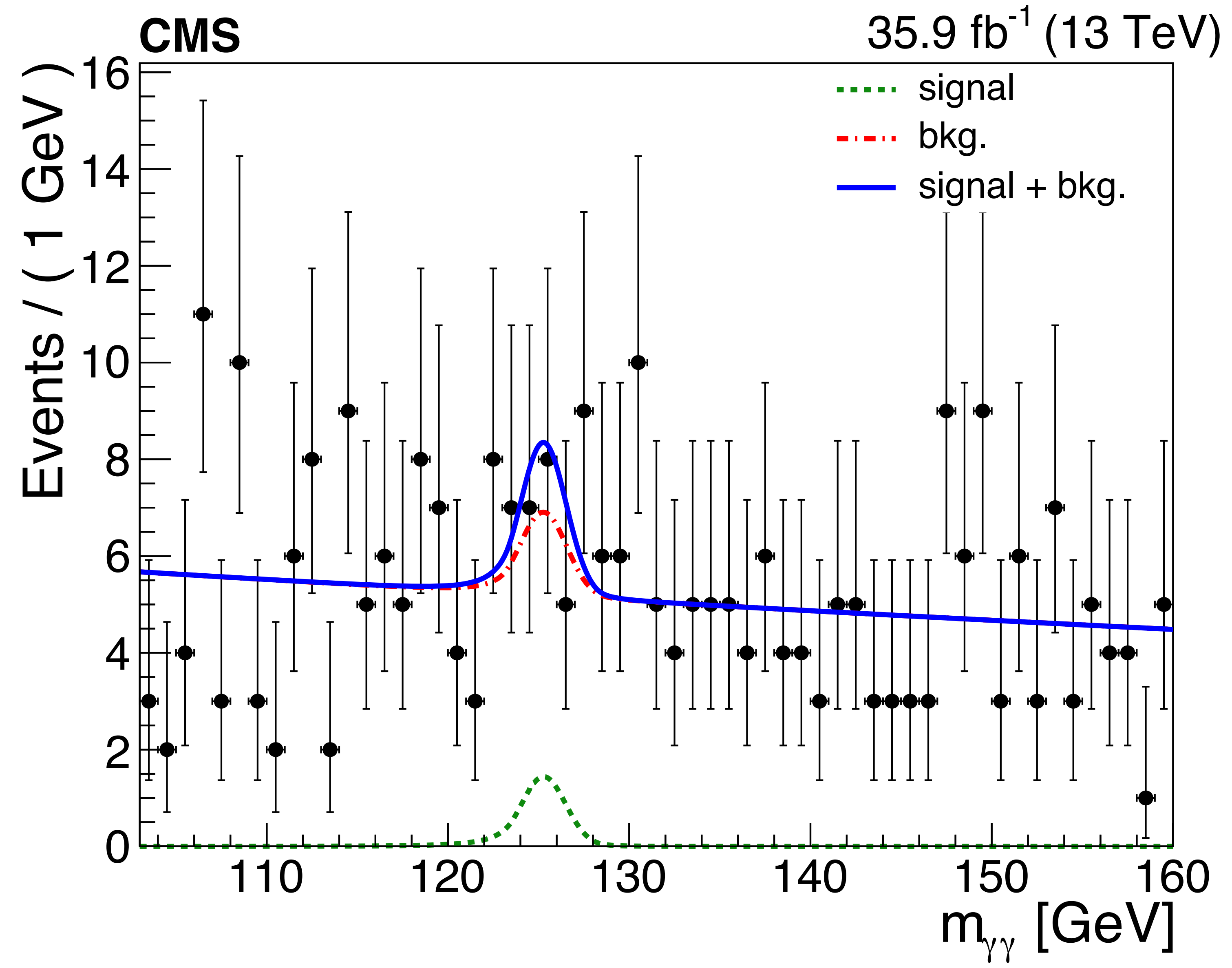
NEW SEARCH TOOLS

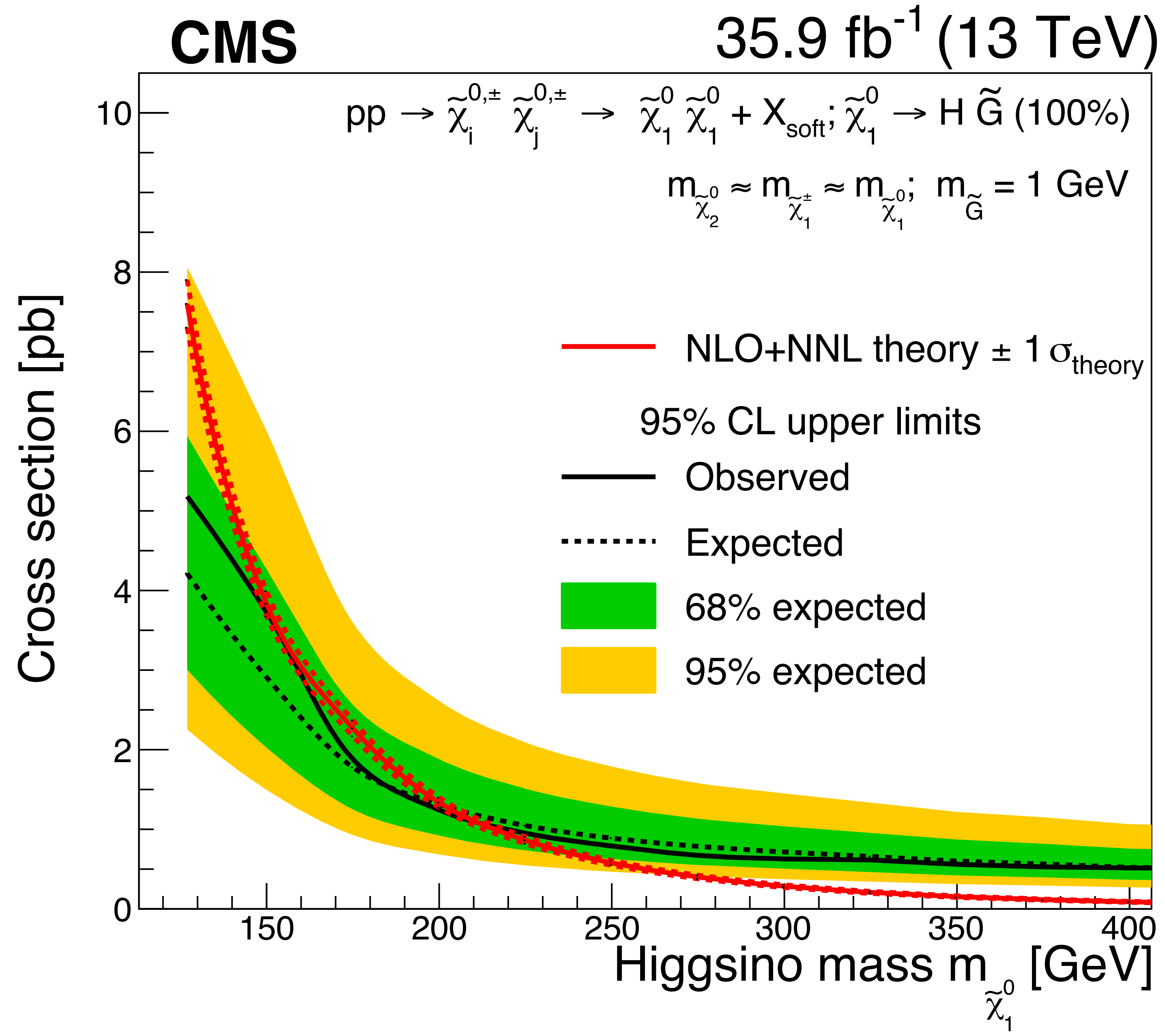
- *Today's BSM search today expanded in many new directions*
  - ***Better identification of complex objects (e.g., tau leptons, b-jets)***
  - *New Standard Model candles (e.g., the Higgs boson)*
  - *New reconstruction strategies (e.g., boosted jets)*
  - *Better understanding of the detector → better sensitivity to soft particles*
  - *More and more exotic signatures: displaced vertices, disappearing tracks, heavy stable charged particles, etc*



NEW SEARCH TOOLS

- *Today's BSM search today expanded in many new directions*
  - *Better identification of complex objects (e.g., tau leptons, b-jets)*
  - ***New Standard Model candles (e.g., the Higgs boson)***
  - *New reconstruction strategies (e.g., boosted jets)*
  - *Better understanding of the detector → better sensitivity to soft particles*
  - *More and more exotic signatures: displaced vertices, disappearing tracks, heavy stable charged particles, etc*

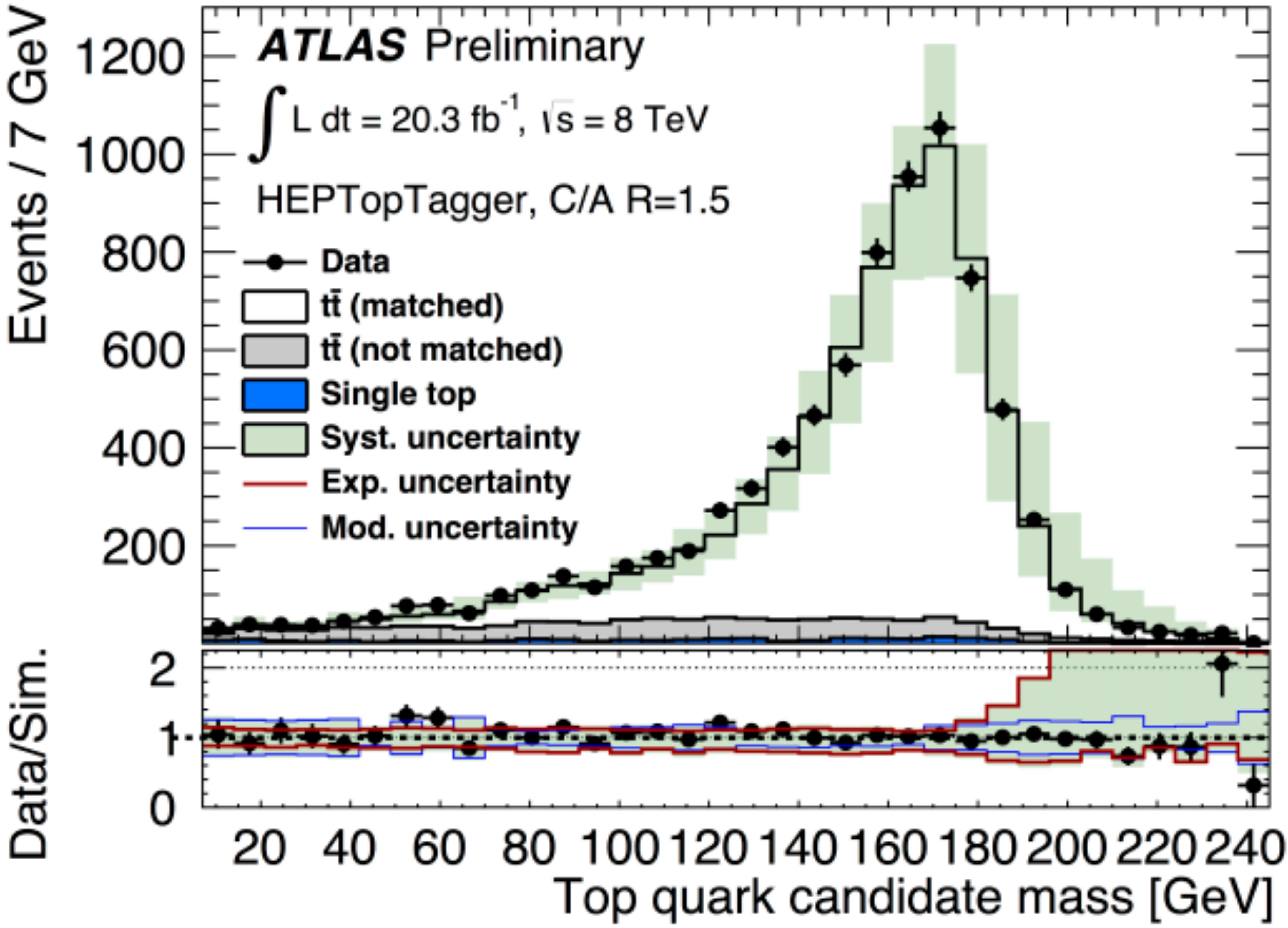




CMS-SUS-16-045

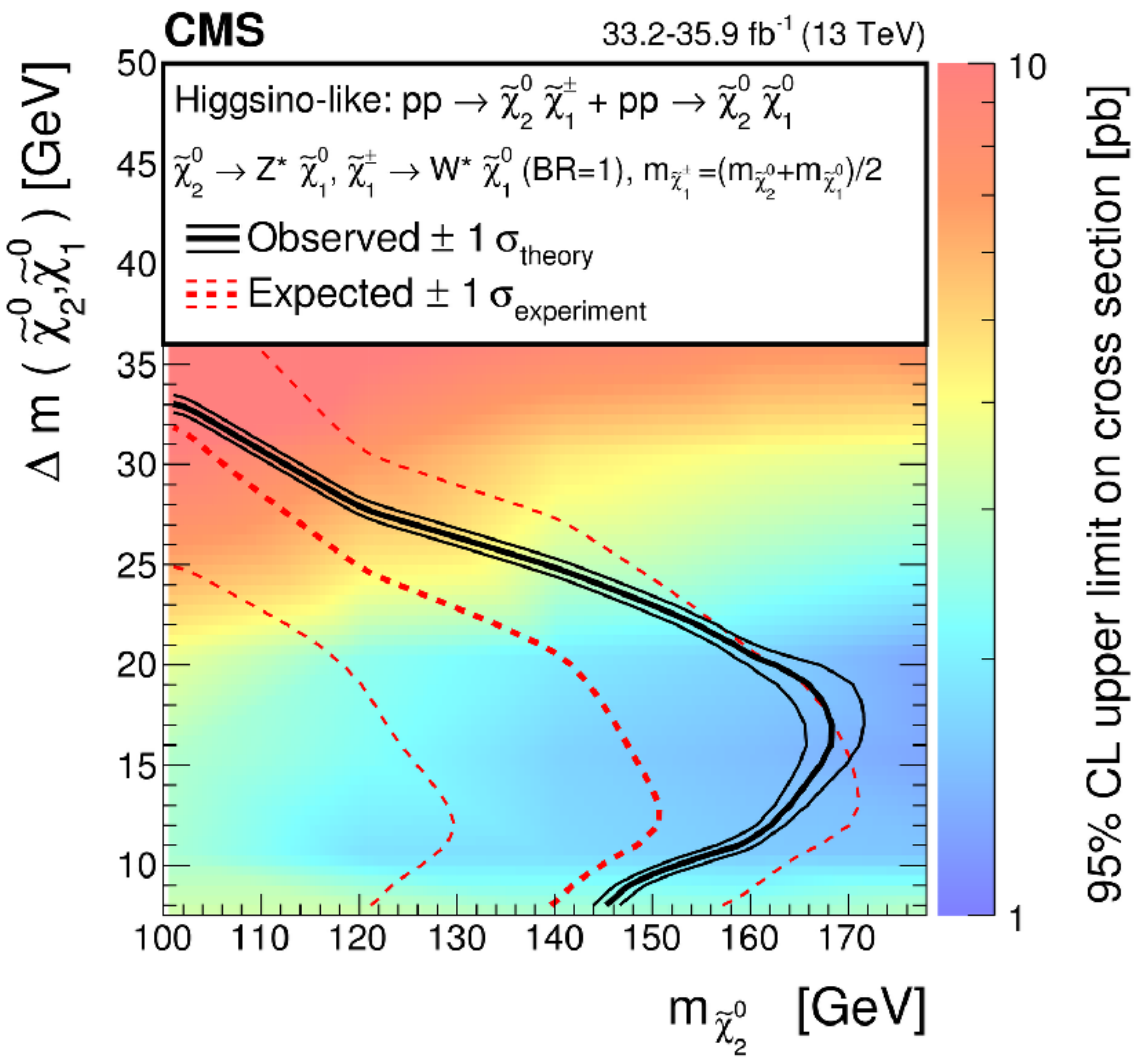
NEW SEARCH TOOLS

- *Today's BSM search today expanded in many new directions*
  - *Better identification of complex objects (e.g., tau leptons, b-jets)*
  - *New Standard Model candles (e.g., the Higgs boson)*
  - ***New reconstruction strategies (e.g., boosted jets)***
  - *Better understanding of the detector → better sensitivity to soft particles*
  - *More and more exotic signatures: displaced vertices, disappearing tracks, heavy stable charged particles, etc*



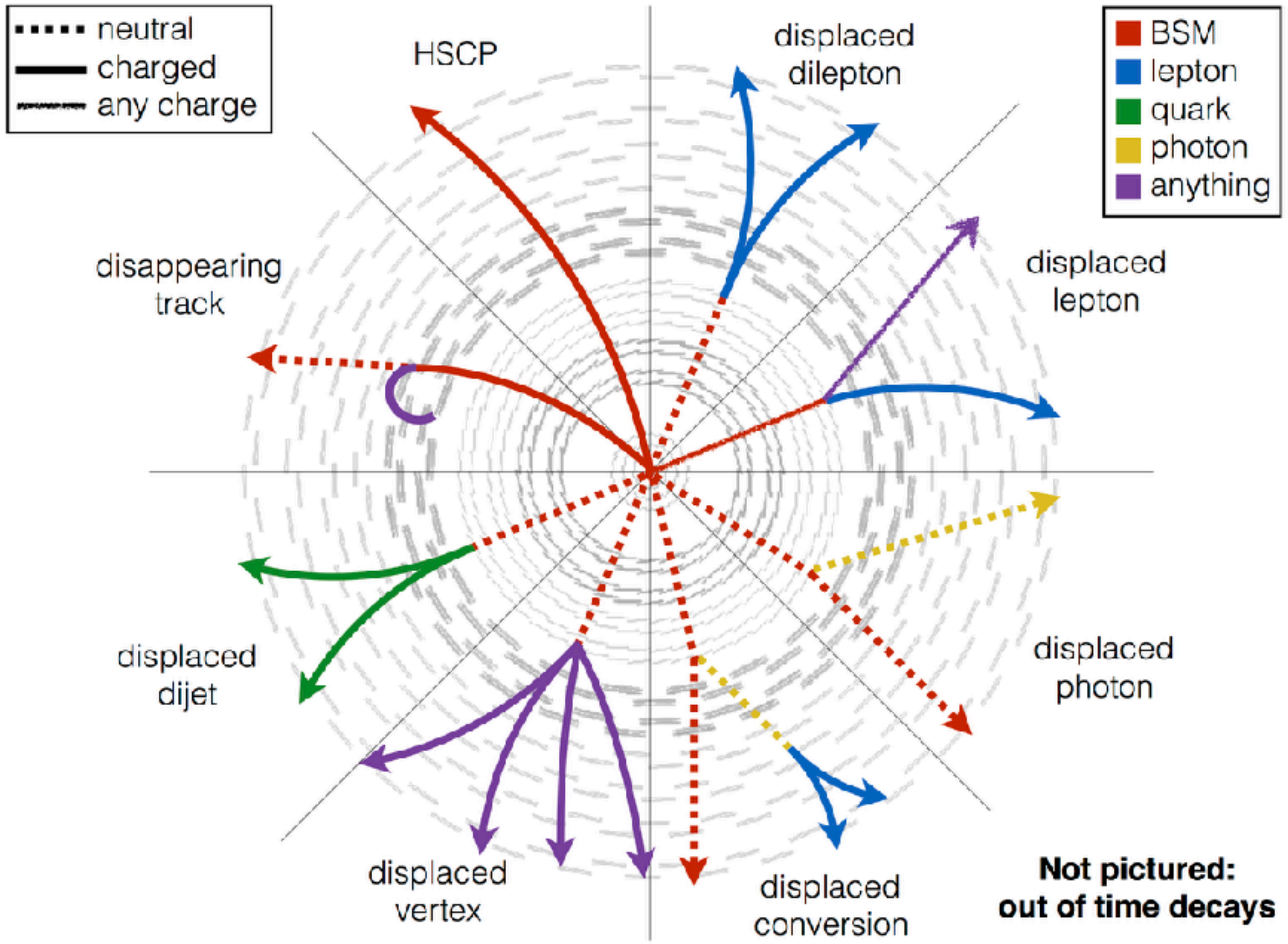
NEW SEARCH TOOLS

- *Today's BSM search today expanded in many new directions*
  - *Better identification of complex objects (e.g., tau leptons, b-jets)*
  - *New Standard Model candles (e.g., the Higgs boson)*
  - *New reconstruction strategies (e.g., boosted jets)*
  - ***Better understanding of the detector → better sensitivity to soft particles***
  - *More and more exotic signatures: displaced vertices, disappearing tracks, heavy stable charged particles, etc*



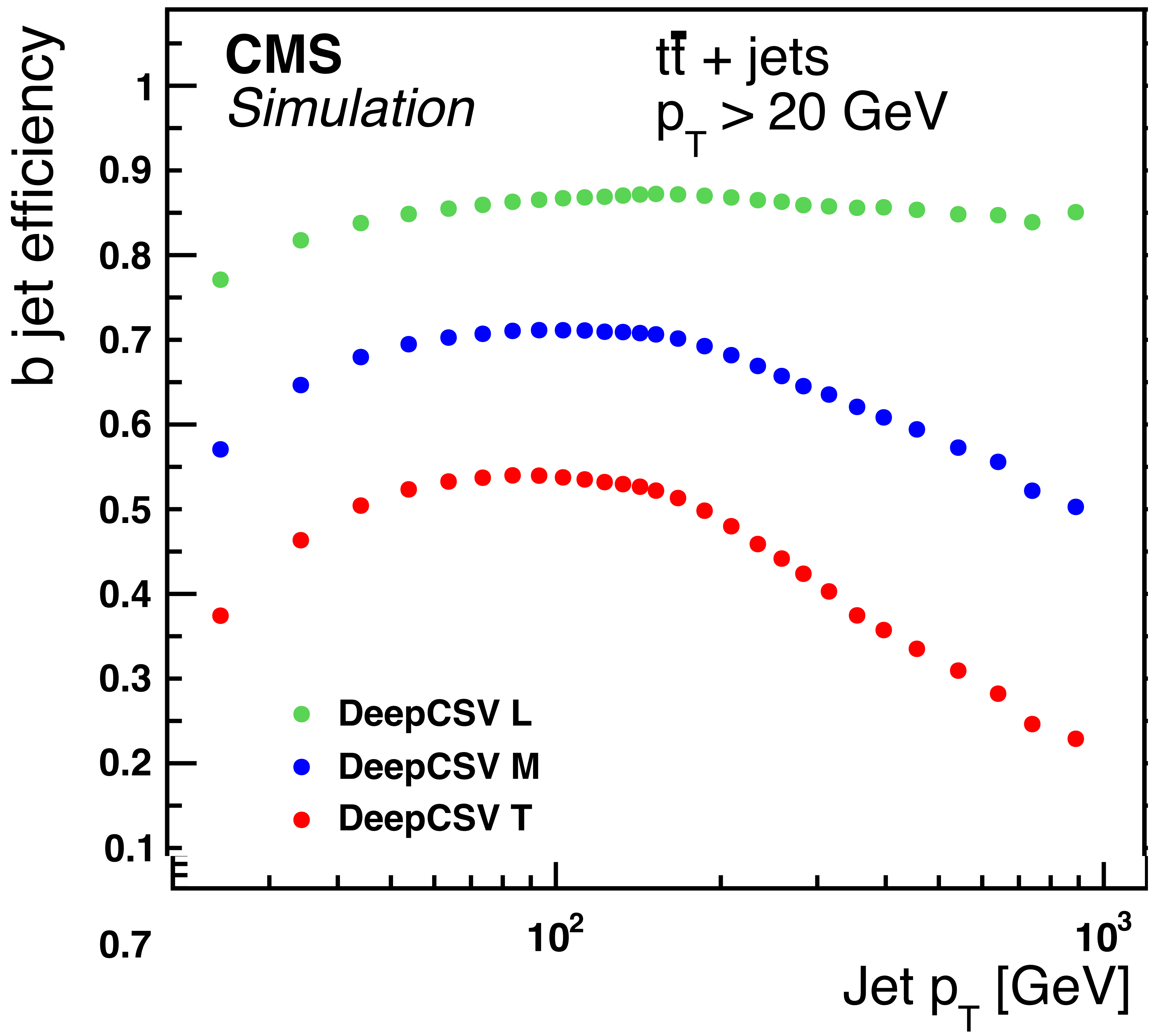
NEW SEARCH TOOLS

- *Today's BSM search today expanded in many new directions*
  - *Better identification of complex objects (e.g., tau leptons, b-jets)*
  - *New Standard Model candles (e.g., the Higgs boson)*
  - *New reconstruction strategies (e.g., boosted jets)*
  - *Better understanding of the detector → better sensitivity to soft particles*
  - ***More and more exotic signatures: displaced vertices, disappearing tracks, heavy stable charged particles, etc***



RECO IN SUMMARY

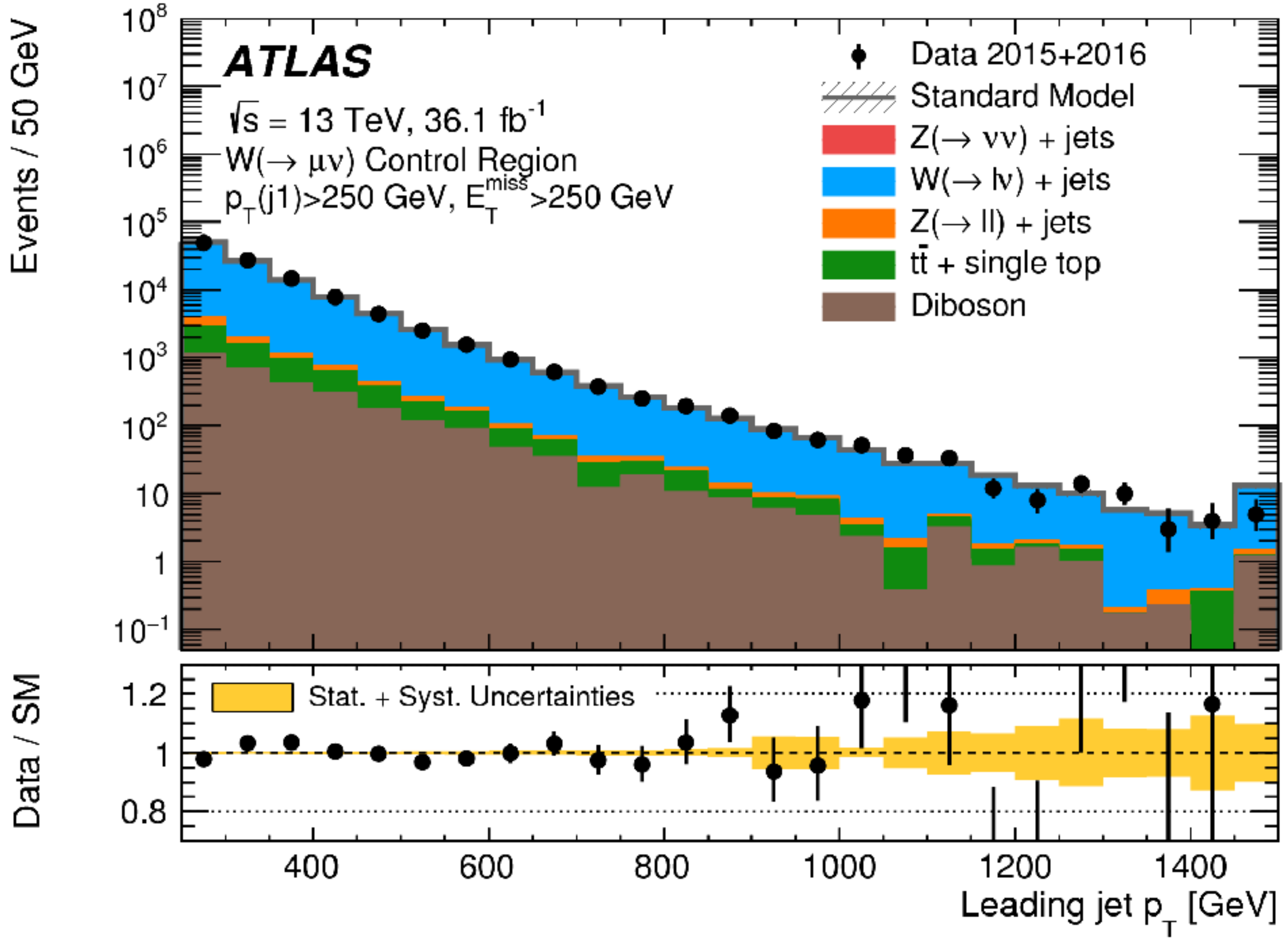
- *As long as you use standard objects (e.g., the kind of jets, leptons, etc used in SM, Higgs, Top analyses) you are ready to go*
- ***Complications may arise with soft or very-hard objects***
- *Reconstruction more complicated when you look into exotic signatures*
  - *displaced jets/tracks/leptons*
  - *stopping tracks*
  - …
- *The solution there is mainly case dependent, and we will talk about it in the last lecture*




CHECK LIST FOR
MY SUSY ANALYSIS
✓ SIGNATURE
✓ TRIGGER
✓ RECONSTRUCTION
✓ ANALYSIS STRATEGY

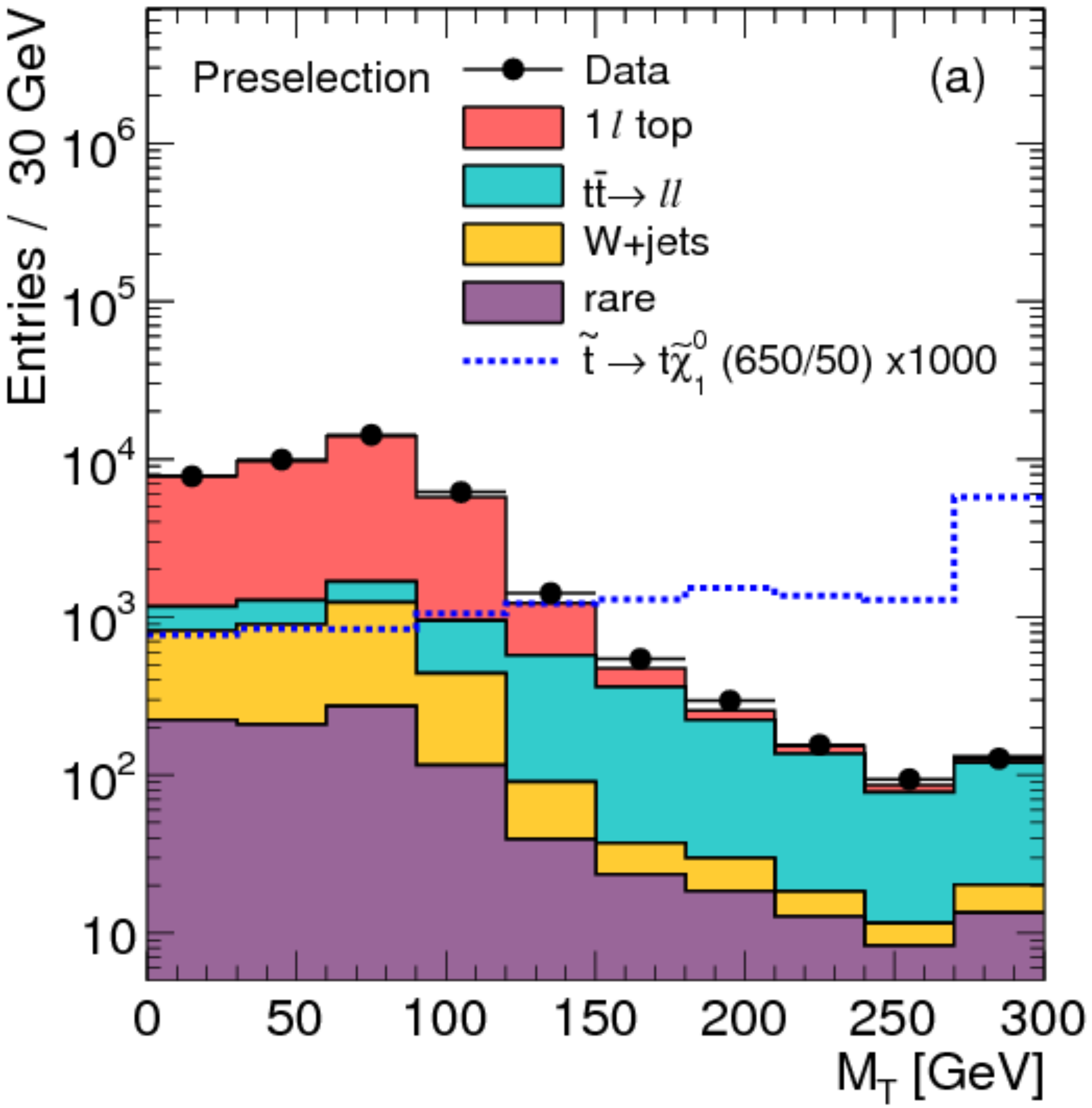
DESIGN YOUR SEARCH

- Based on your signature, you should have by now some final state in mind (e.g., jets + 1leptons + missing E_T)
- MC simulation would tell you the list of background processes from the Standard model, among which
  - Z+jets with Z to neutrinos
  - W+jets with W→ℓν
  - QCD with one lepton from meson decays
  - QCD with one jet faking a lepton (not for muons)
  - tt with at least one W→ℓν



ANALYSIS SELECTION

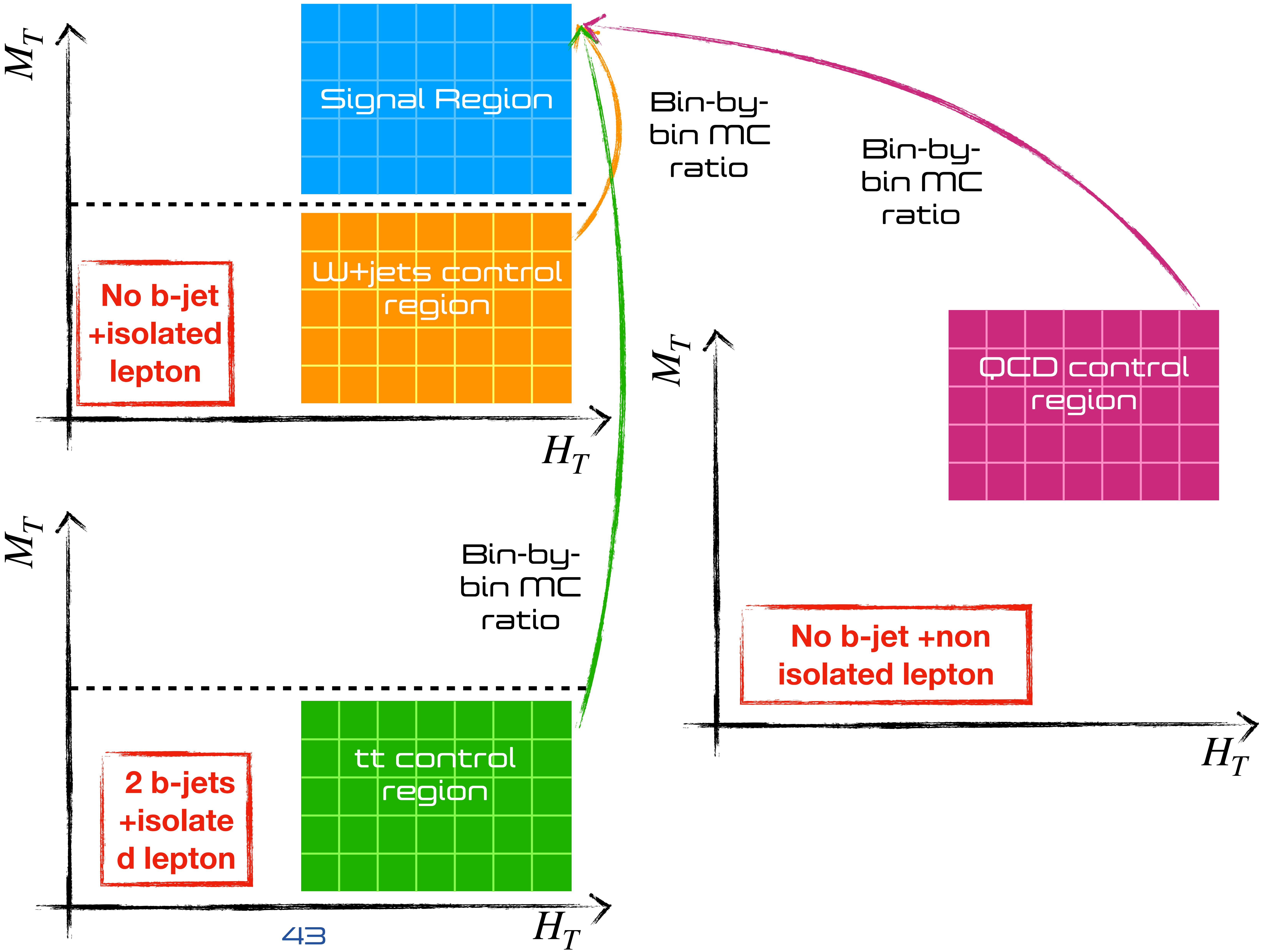
- *You need to have some physics motivated quantity that looks different for a signal and a background sample*
  - *Some kinematic quantity (see Lecture 2)*
  - *The presence of some special object (see Lecture 3)*
  - *Multiplicity of objects of some kind, e.g., leptons, b-jets, etc (see Lecture 3)*
- *Based on these quantities, one can focus the search on a subset of the events for which a signal enhancement is expected*



SIGNAL AND CONTROL REGIONS

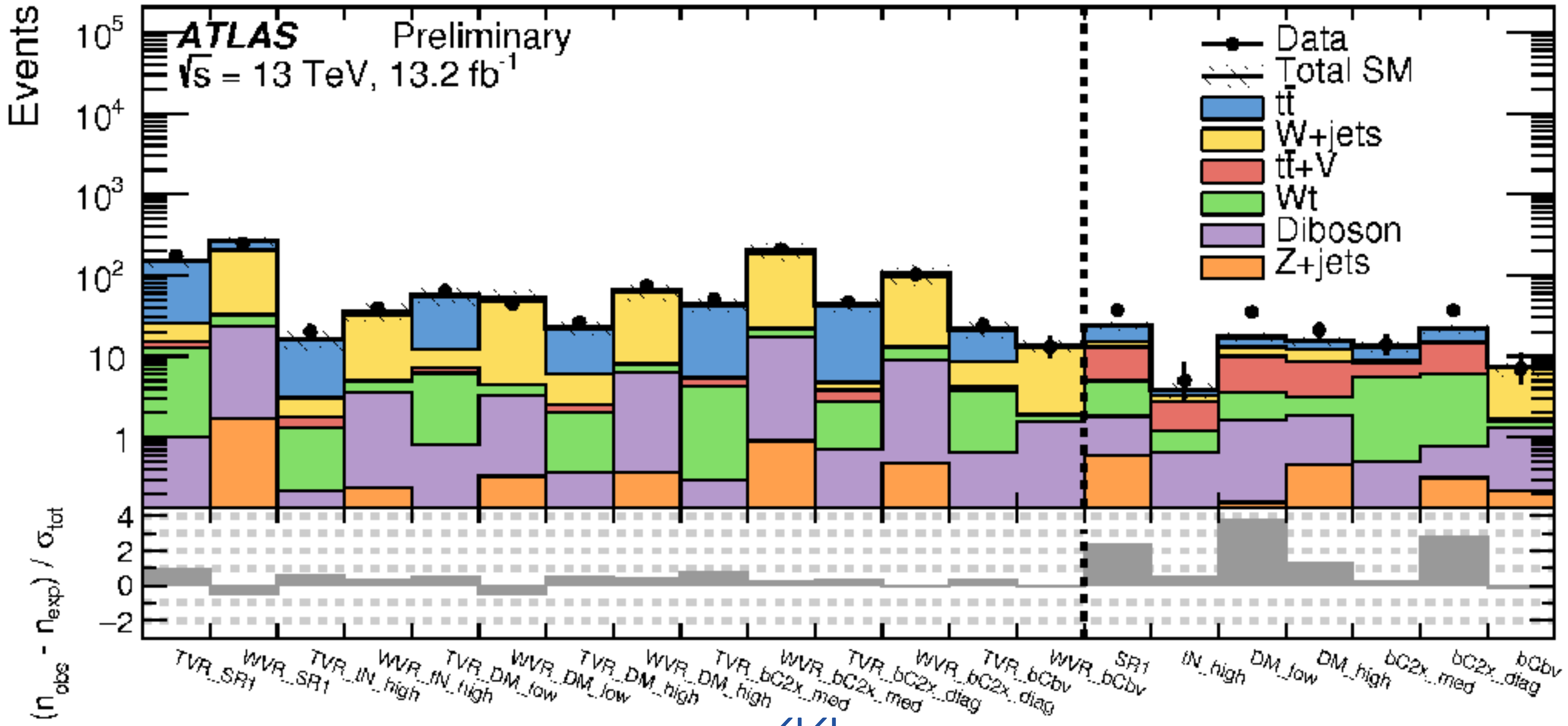
- *Together with the signal region, a set of control regions are designed, enriched of one kind of background*

- *Backgrounds can be measured in these regions and scaled according to transfer factors, predicted with MC simulations*



SIGNAL AND CONTROL REGIONS

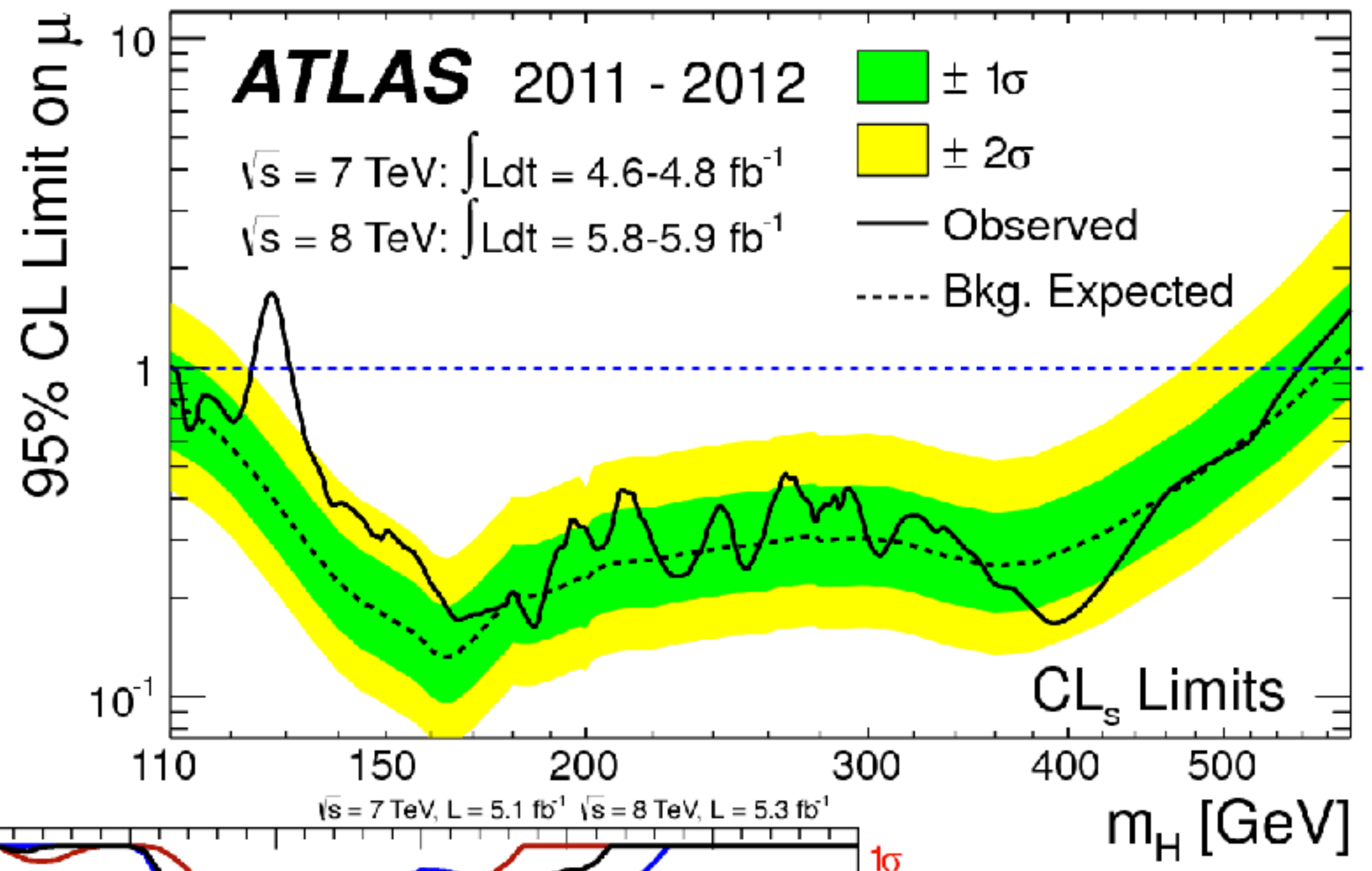
- *With this MC-assisted data-driven background prediction, you are more robust vs unexpected issues*
  - *Still, keep in mind that our MC simulation is more reliable than this, when far from the tails (template fits are often used for SM measurements)*
- *At this stage, your analysis translates into a multi-bin counting exercise, where the signal is searched as excess on prediction*

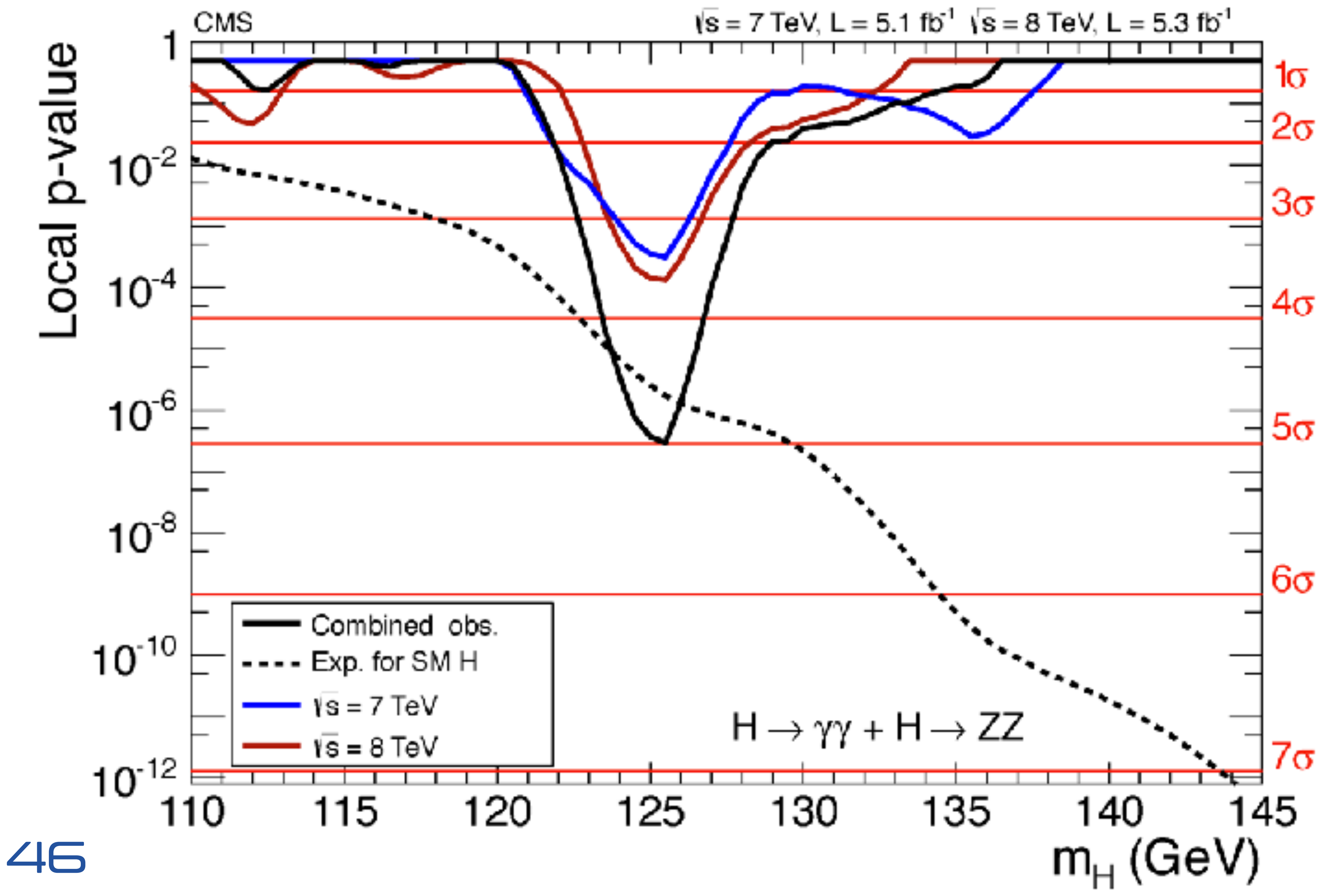



Check List for
My SUSY Analysis
✓ Signature
✓ Trigger
✓ Reconstruction
✓ Analysis strategy
✓ Statistical analysis

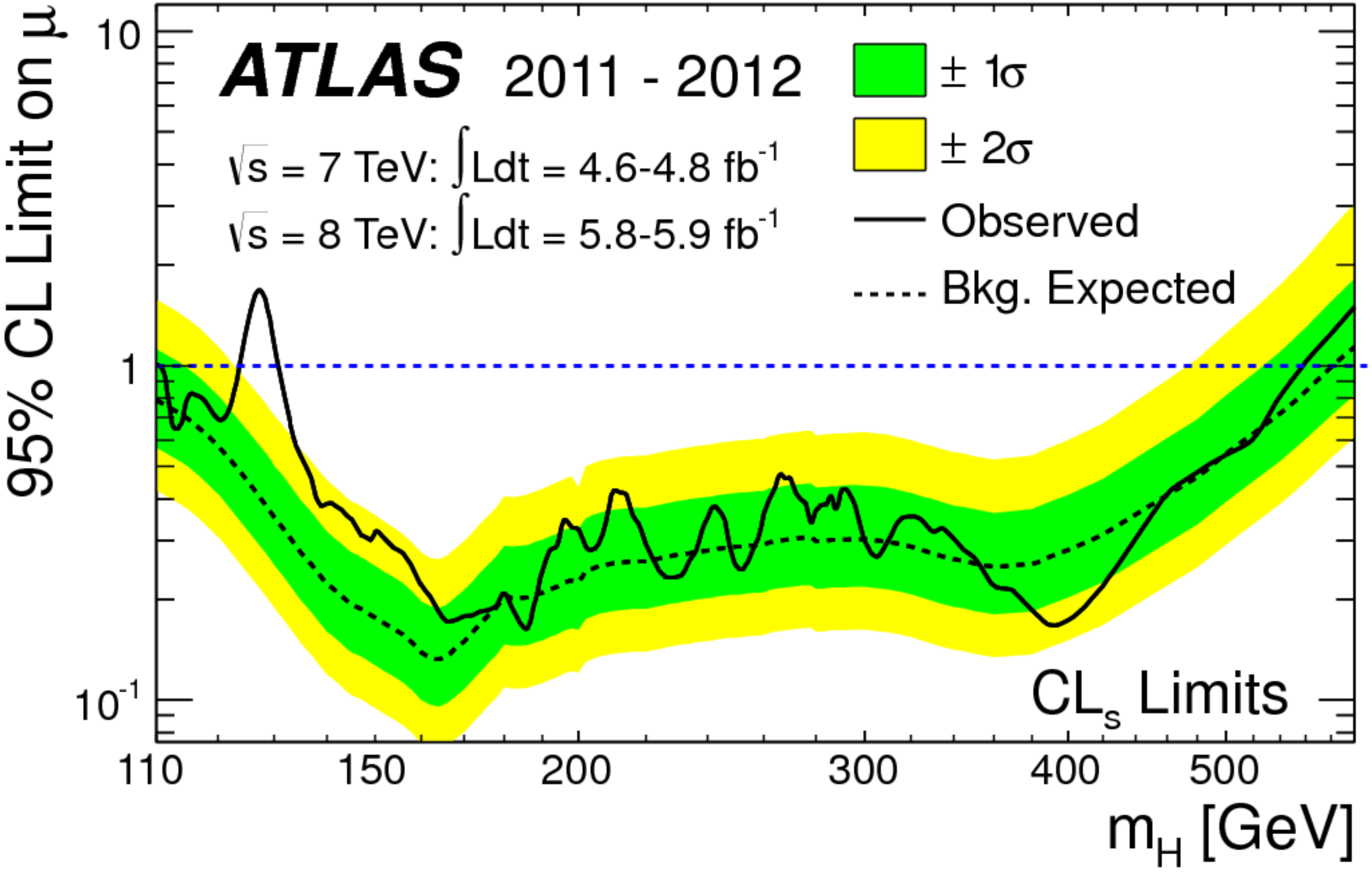
STATISTICAL TOOLS

- The full setup of statistical procedure was setup at Higgs discovery time
  - (asymptotic) CLs emerged as the limit-setting procedure
  - one-sided p-values, converted to number of sigmas, are quoted for evidence of an excess
- These procedures result in plots like these
- In case their meaning is not clear, I will go through the procedure to get them



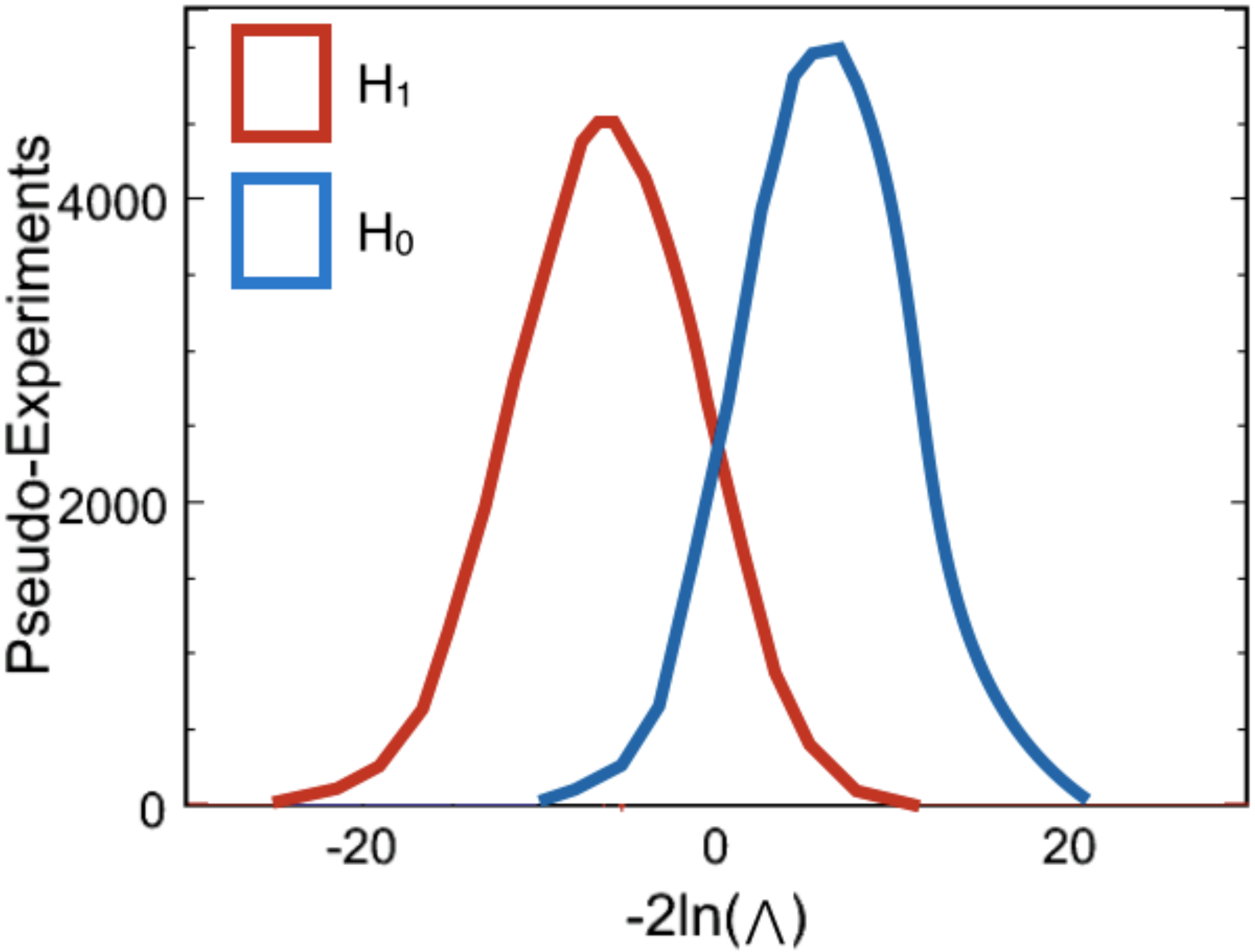


LIMIT SETTING



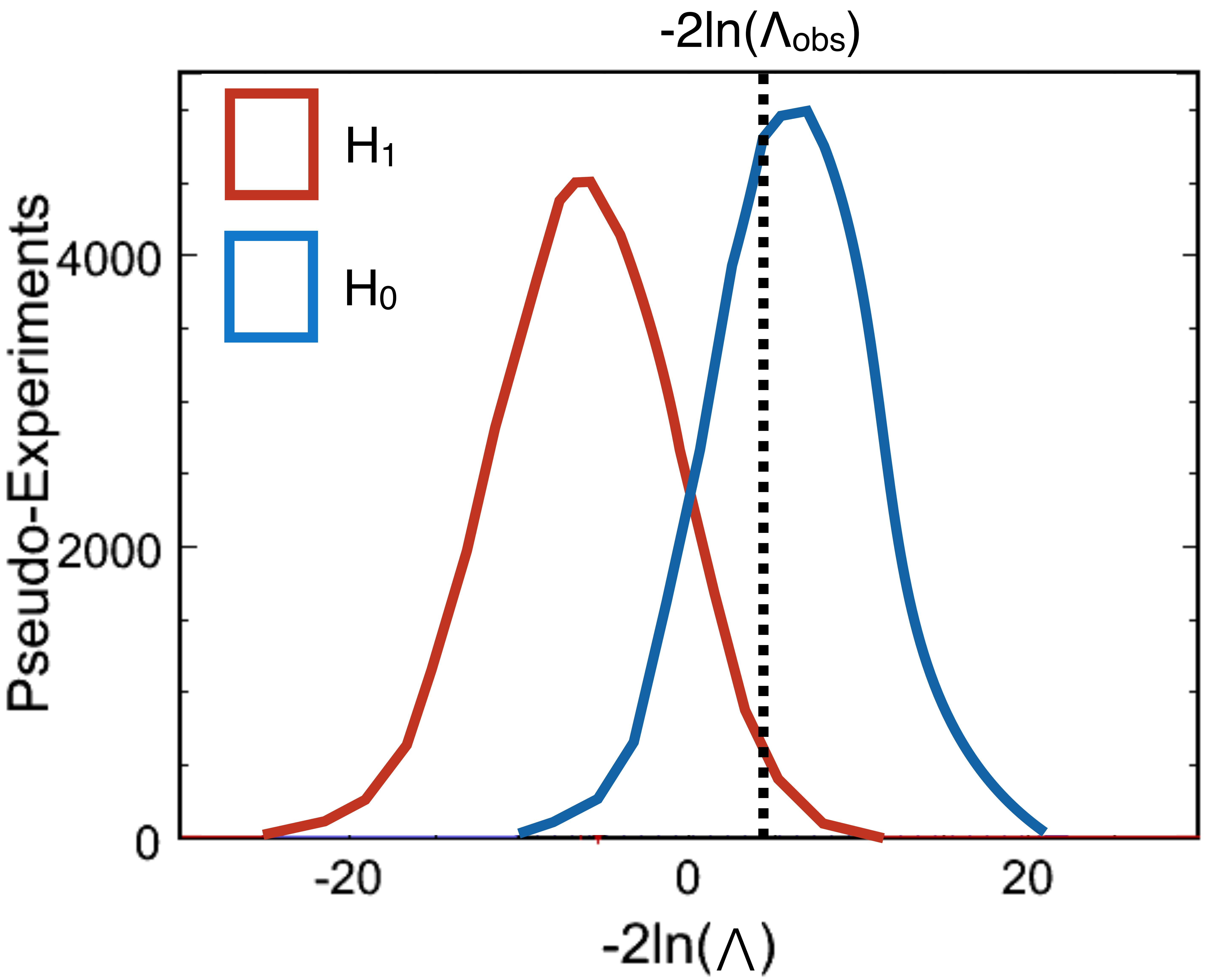
LIMIT SETTING

- *Let's say you are looking for the Higgs boson. You have two hypotheses:*
  - *H0: no Higgs*
  - *H1: Higgs somewhere*
- *Assume a mass value*
- *For each mass value, assume a cross section and construct the two distributions for some discriminating quantity $\Lambda^{(*)}$ under H0 and H1*
  - *generate toy MC with σ=0 (H0)*
  - *generate toy MC with σ=σ* (H1)*

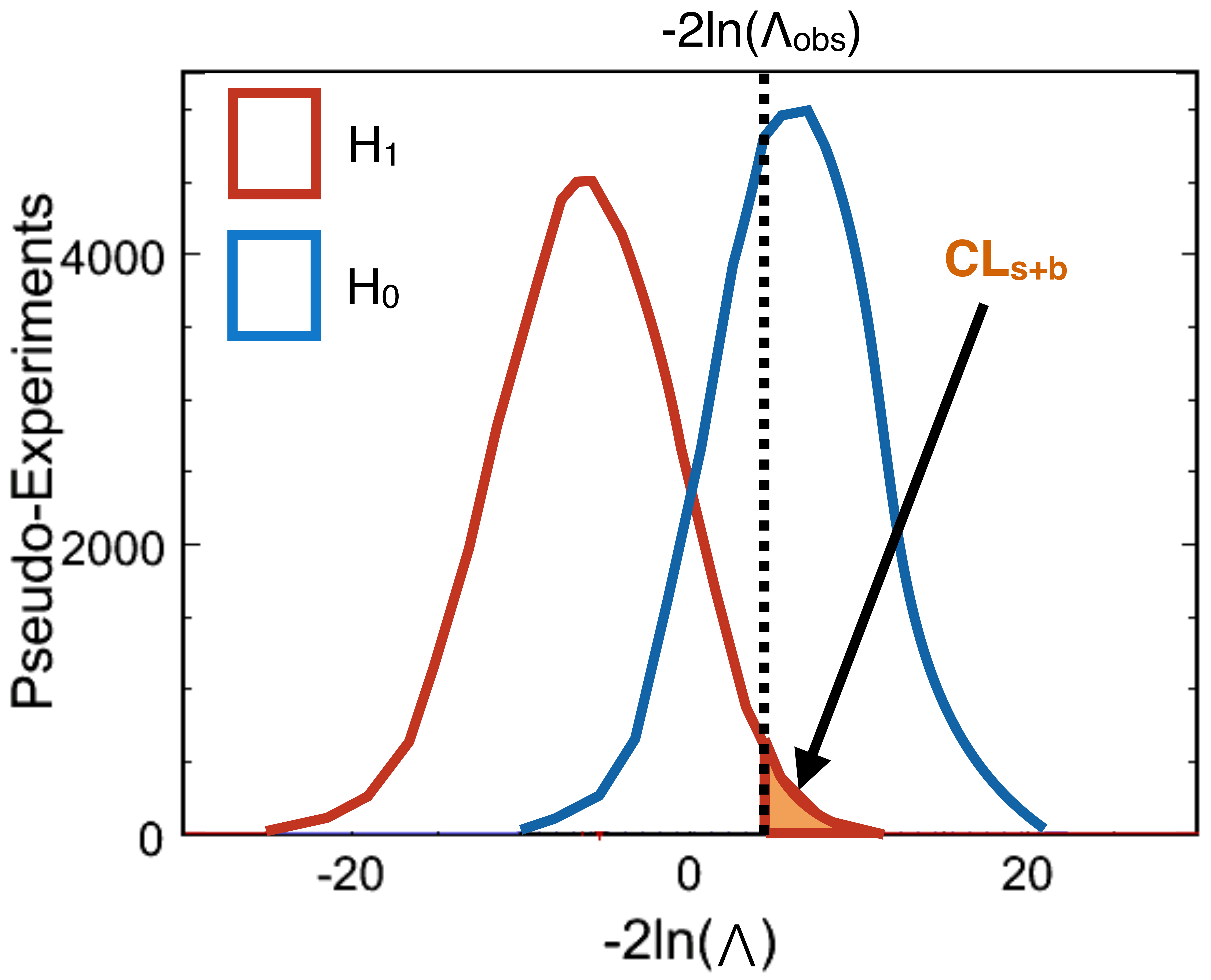


(*) Neyman Pearson lemma says "use the ratio of the likelihood under the two hypotheses

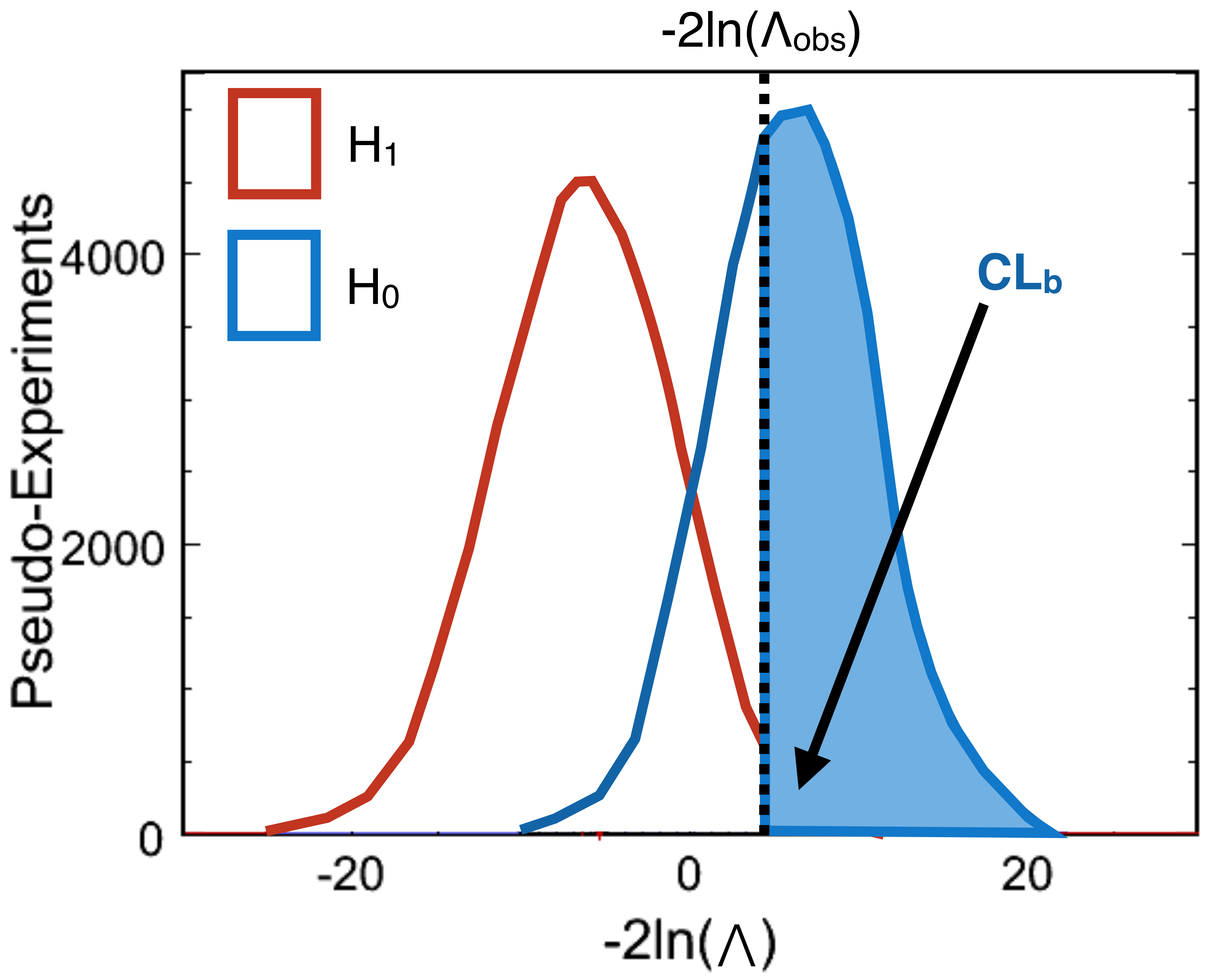
YOUR OBSERVATION



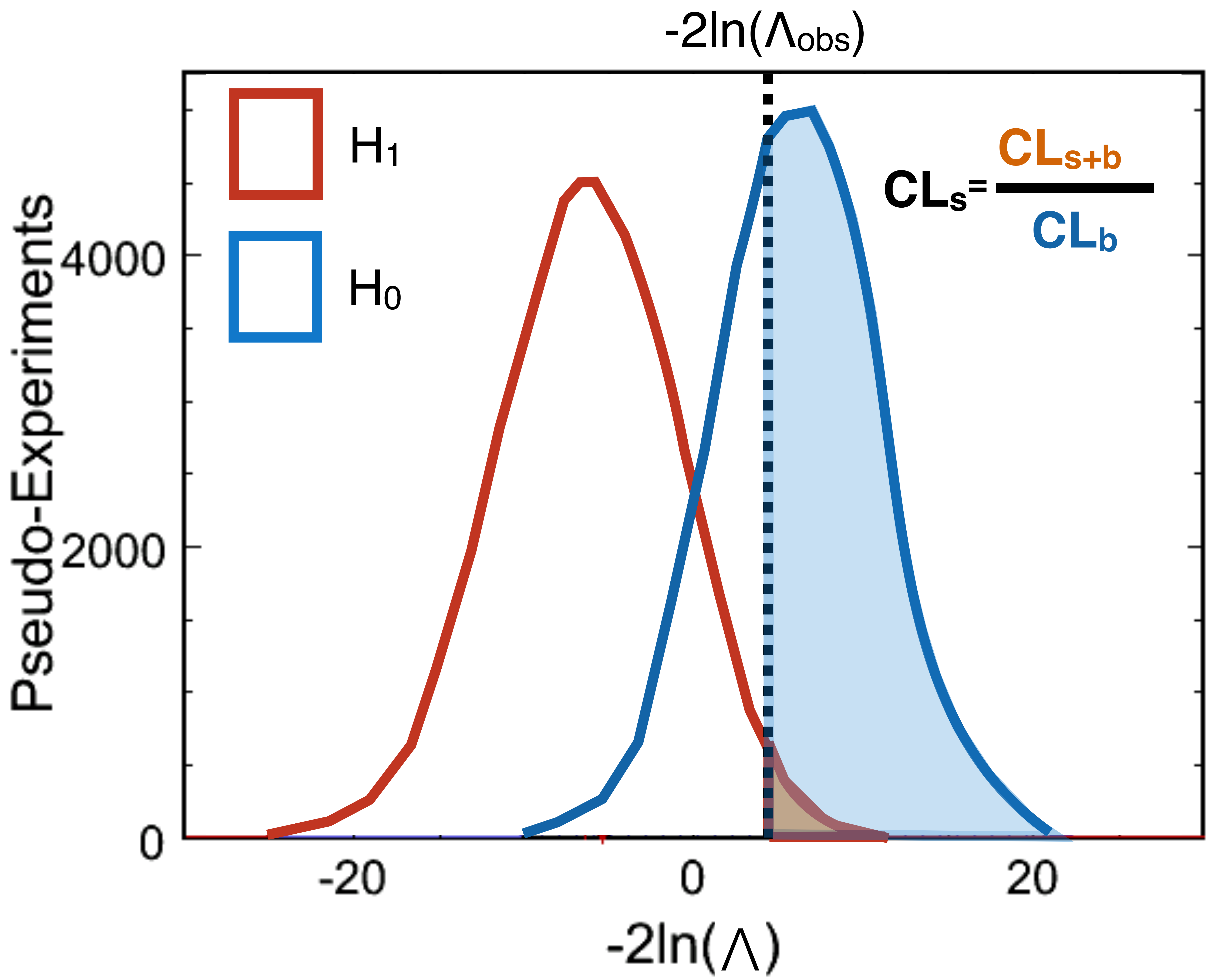
OBSERVED CL_{S+B}



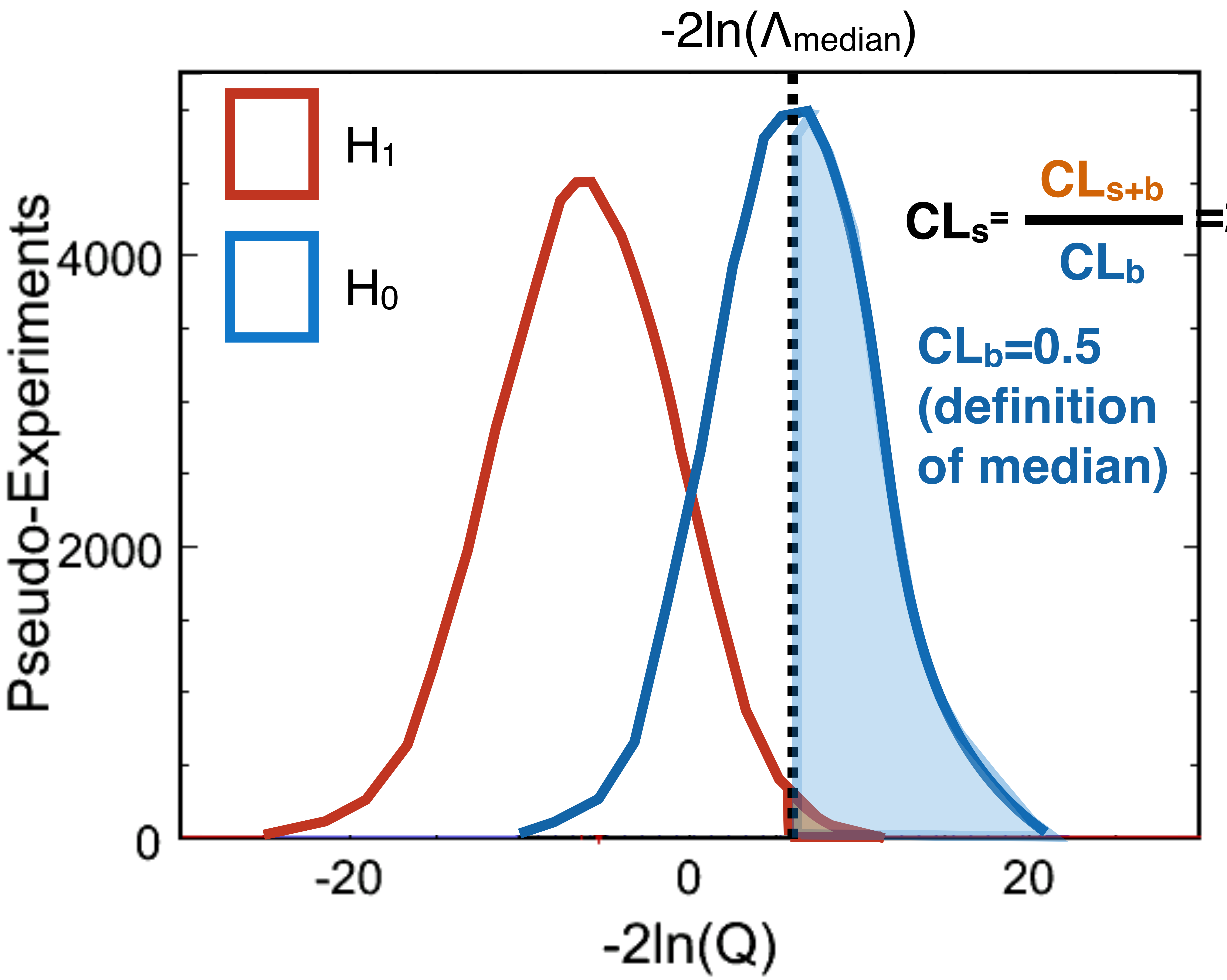
OBSERVED CL_B



OBSERVED CLs



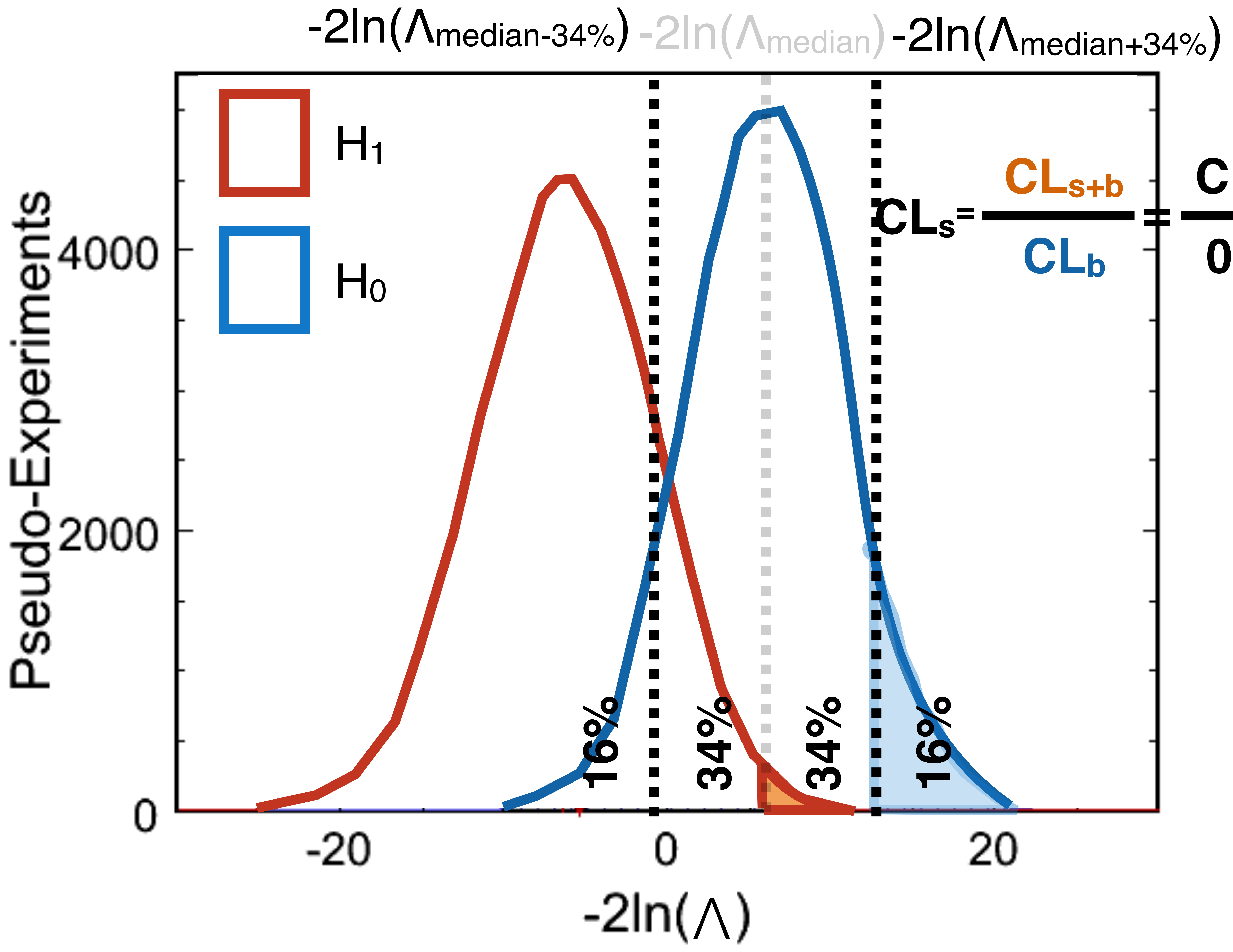
EXPECTED CLs



$$CL_s = \frac{CL_{s+b}}{CL_b} = 2CL_{s+b}$$

$CL_b = 0.5$ (definition of median)

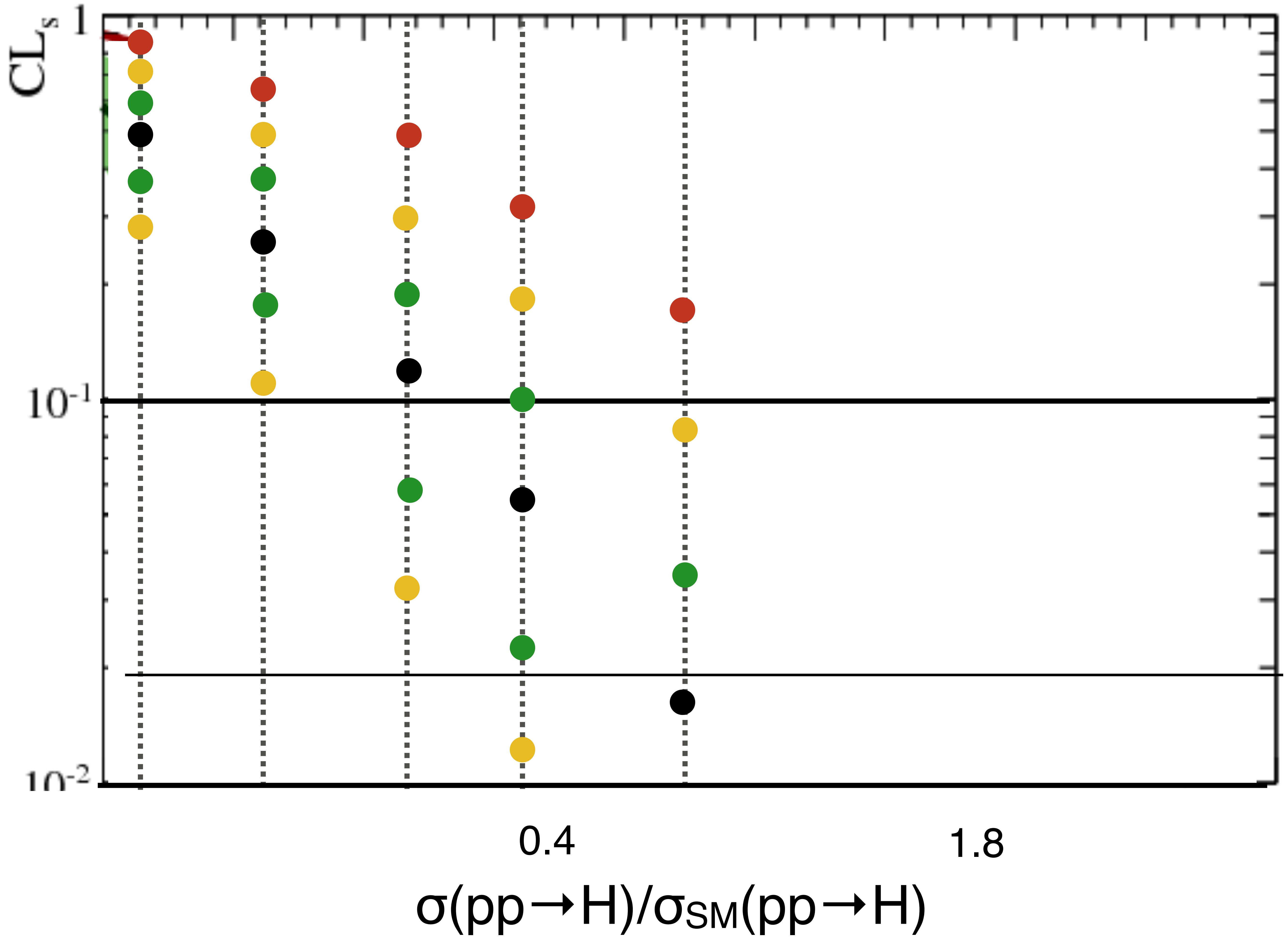
EXPECTED CLs



$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{CL_{s+b}}{0.16}$$

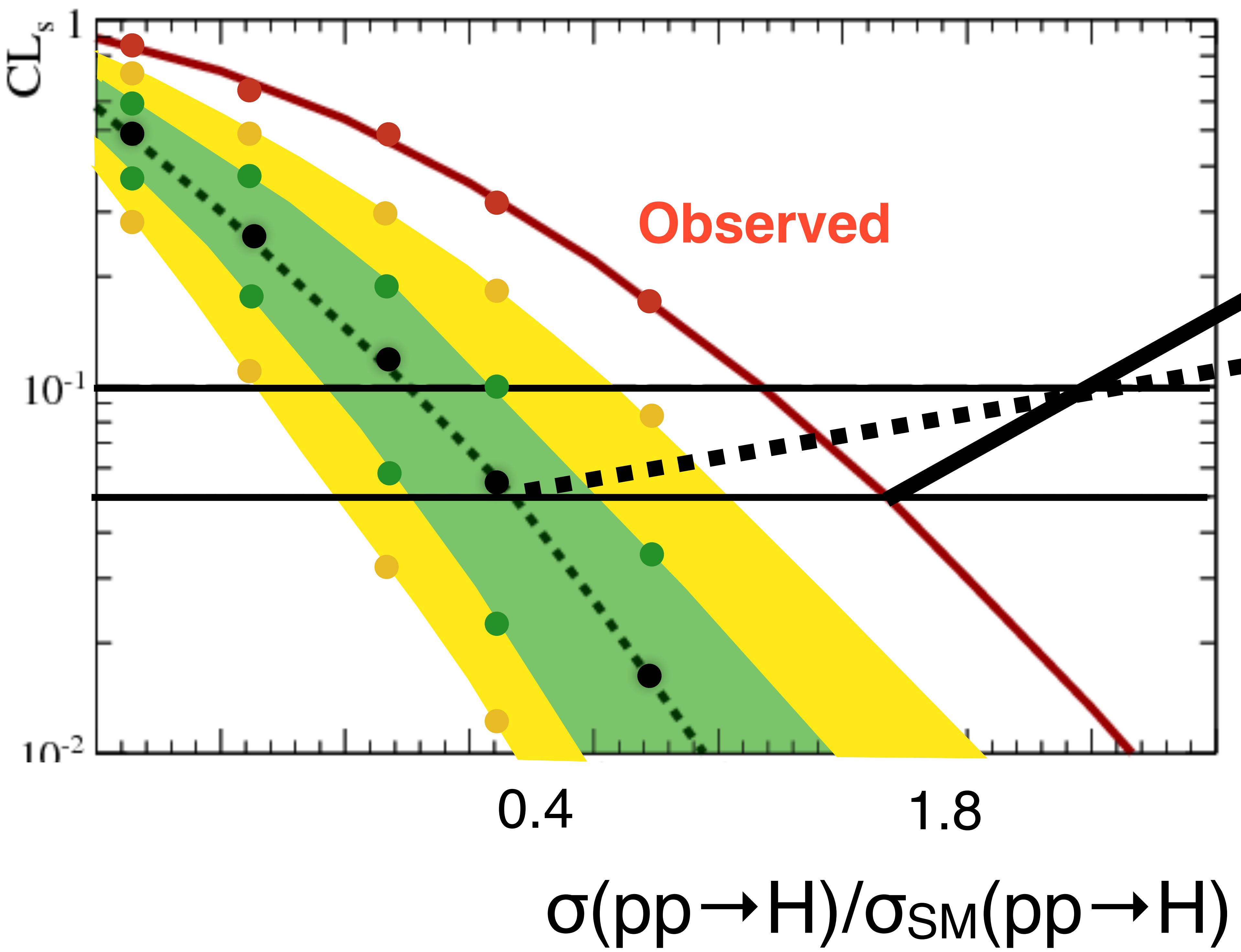
THE BRAZILIAN FLAG PLOT

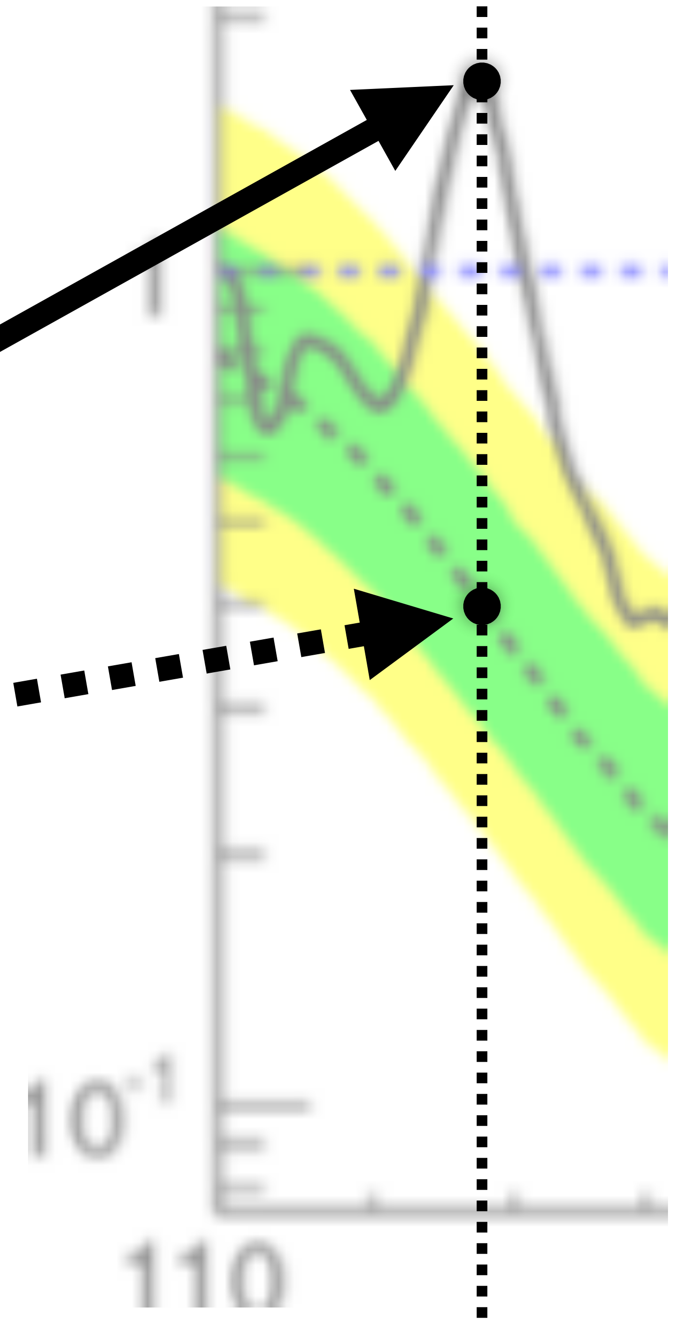
- *Repeat the procedure above for several values of σ/σ_{SM}*



THE BRAZILIAN FLAG PLOT

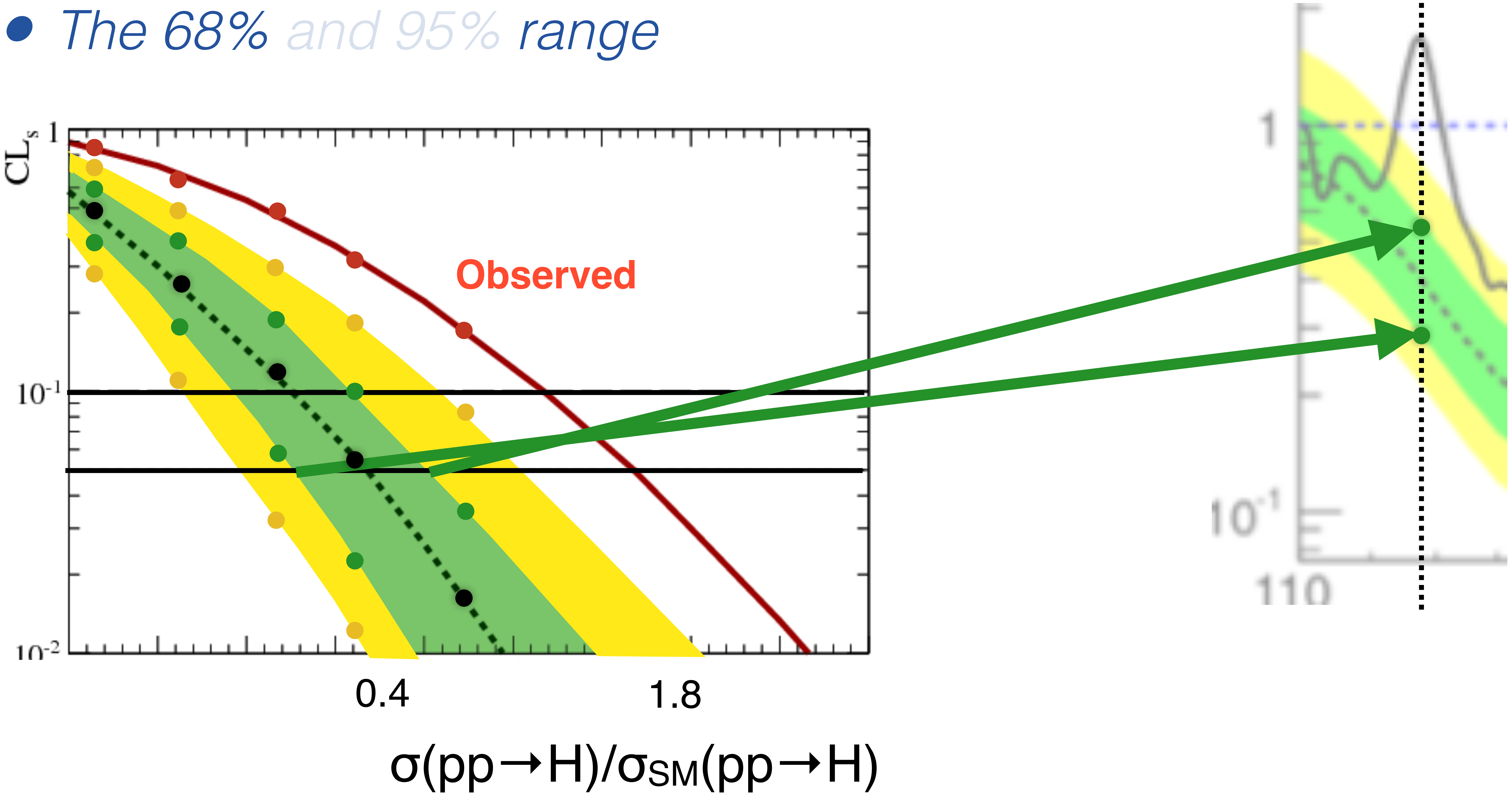
- *Each line intercept CLs = 0.05. The intersection gives you*
- *The expected and observed limit*
- *The 68% and 95% range*





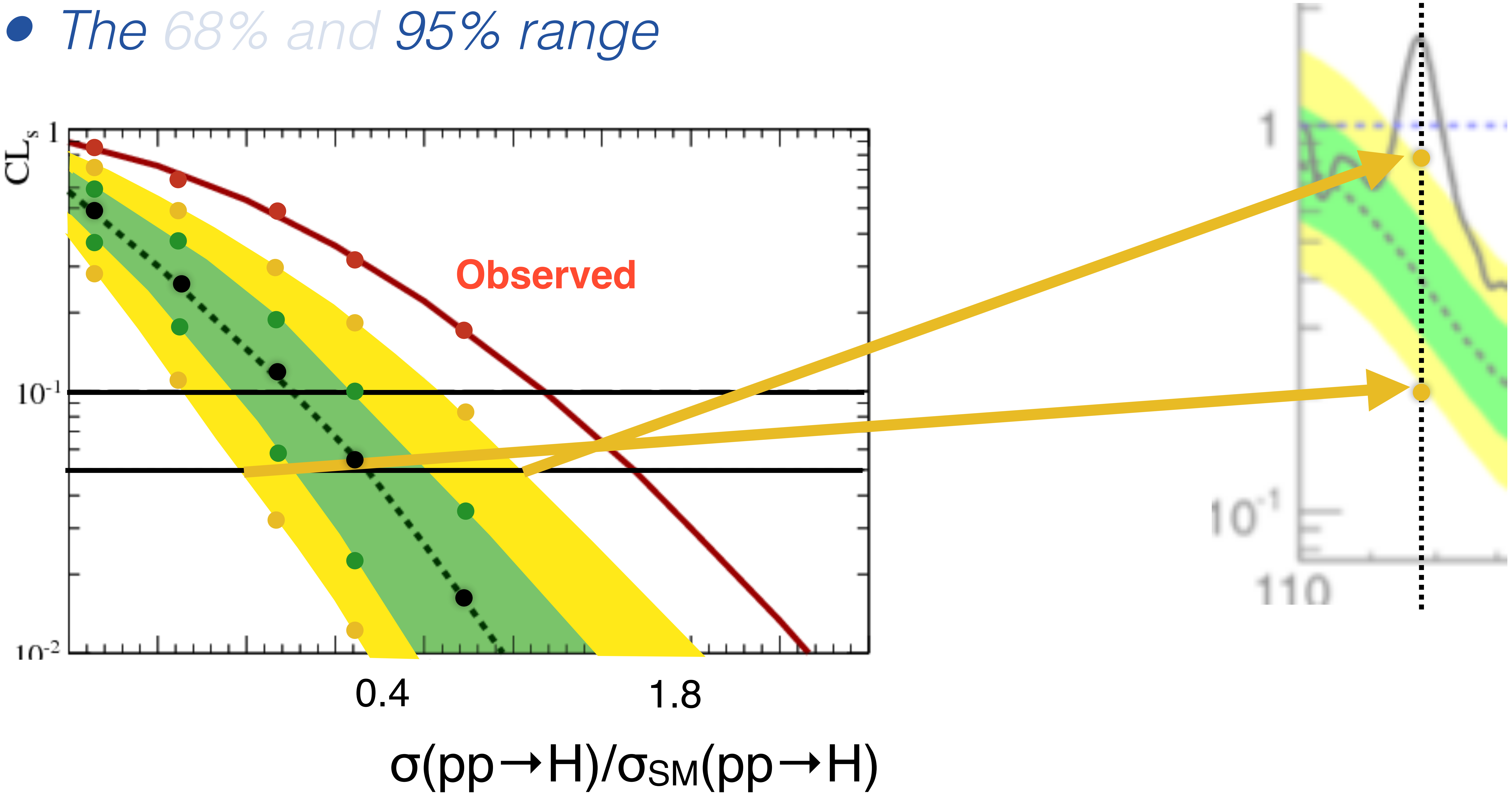
THE BRAZILIAN FLAG PLOT

- *Each line intercept CLs = 0.05. The intersection gives you*
- *The expected and observed limit*
- *The 68% and 95% range*



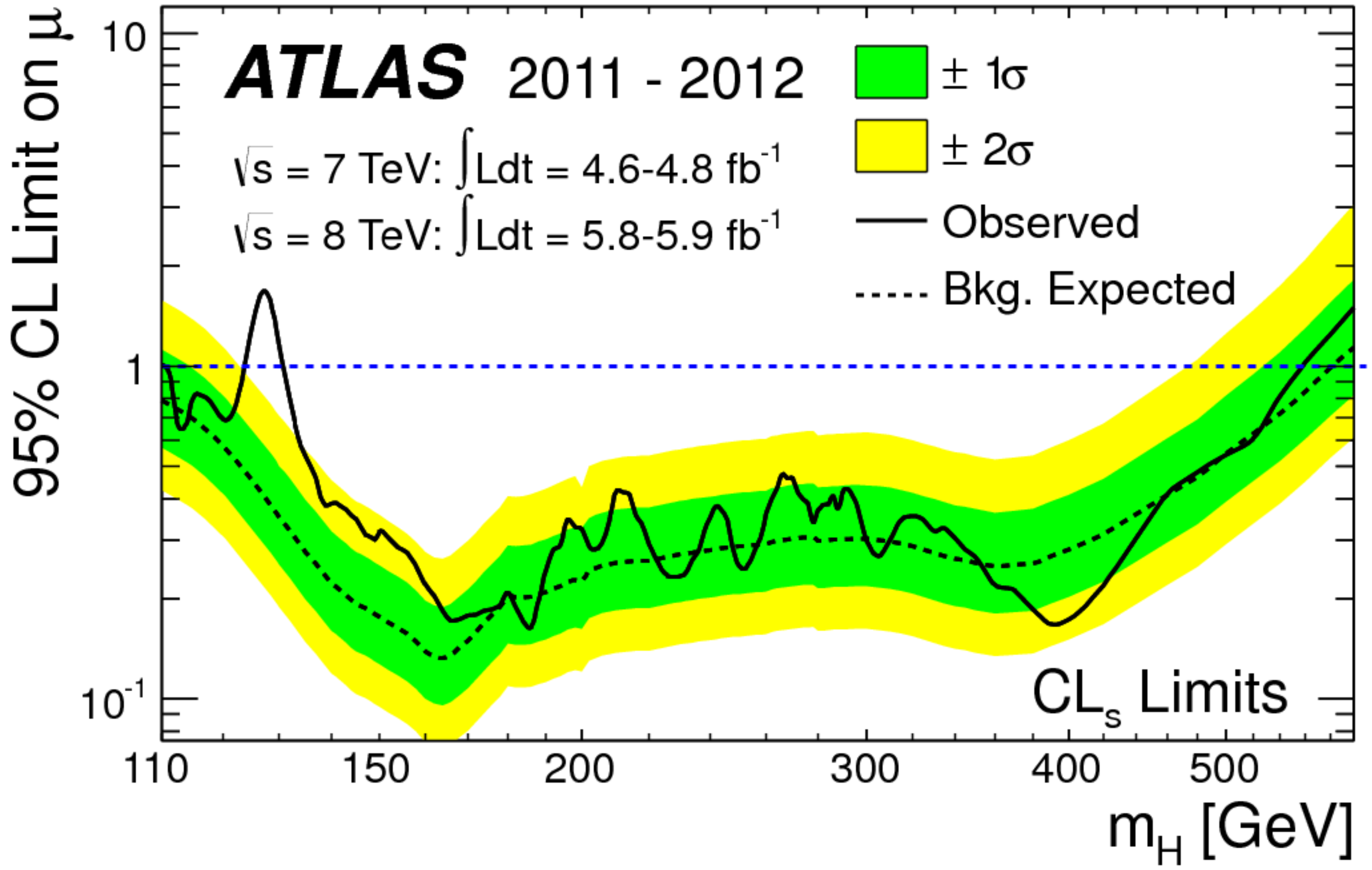
THE BRAZILIAN FLAG PLOT

- *Each line intercept CLs = 0.05. The intersection gives you*
- *The expected and observed limit*
- *The 68% and 95% range*



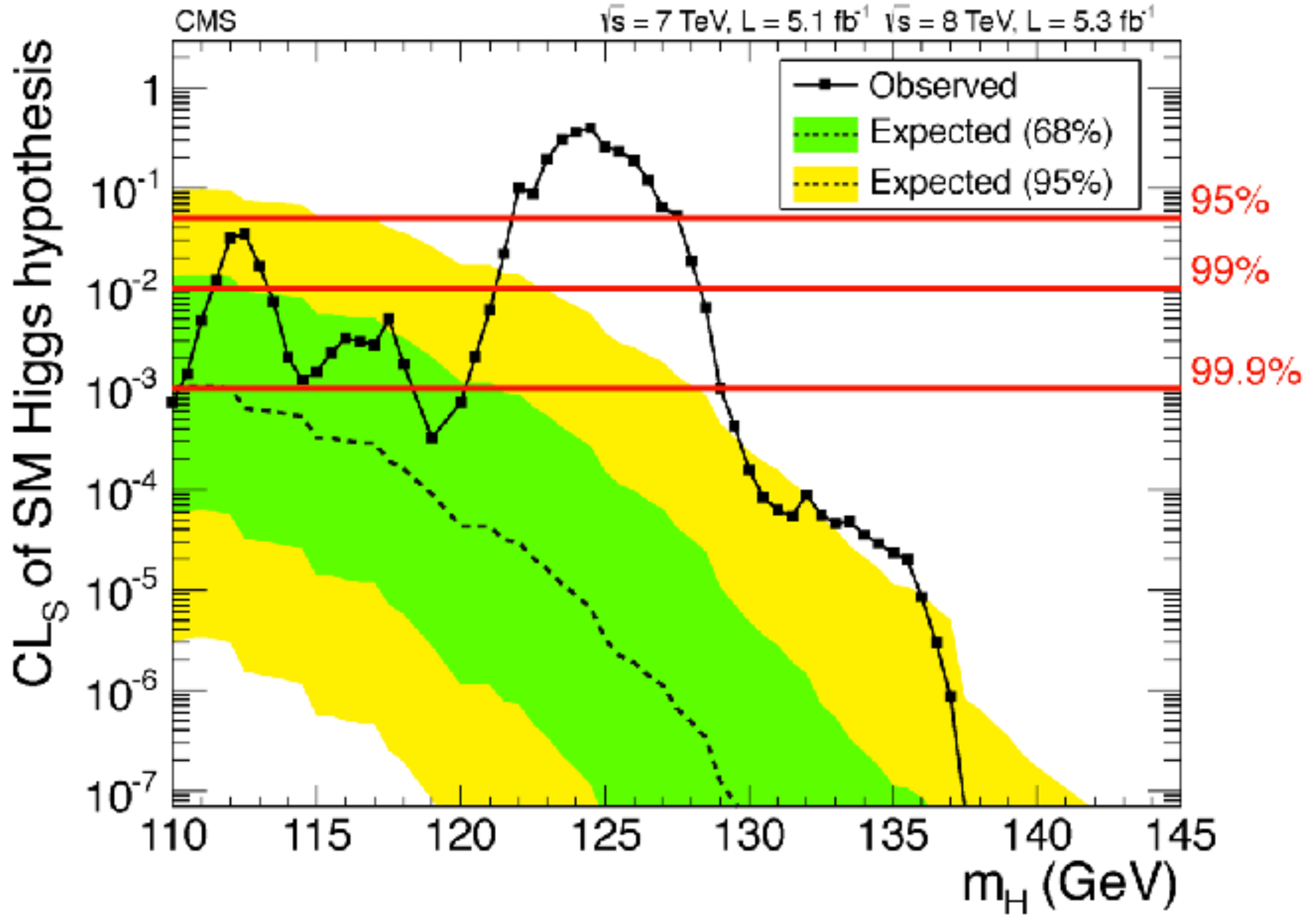
THE BRAZILIAN FLAG PLOT

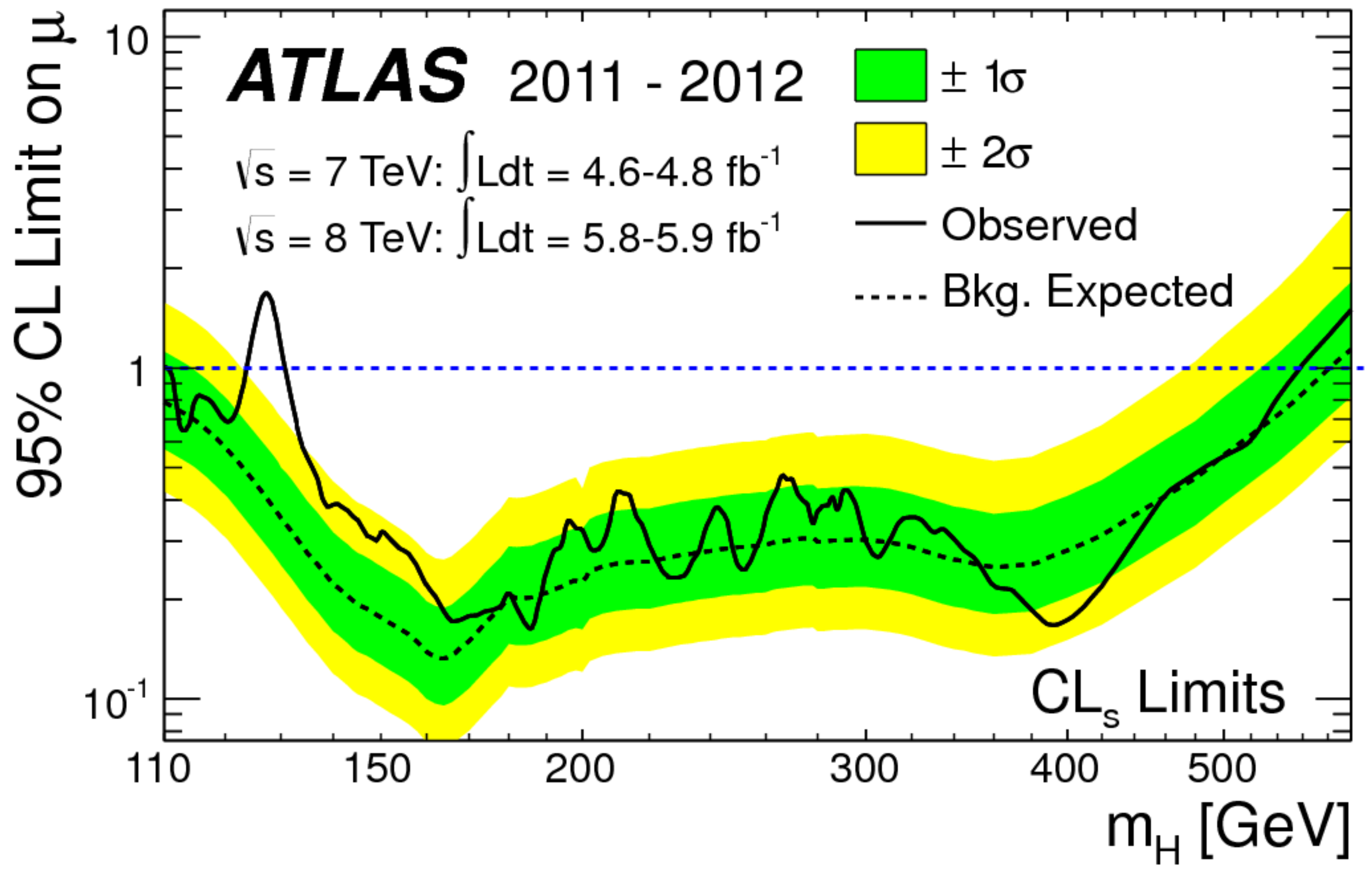
- *Now repeat the procedure for any value of mH and connect the dots*



WHEN EXCLUSION GOES WRONG

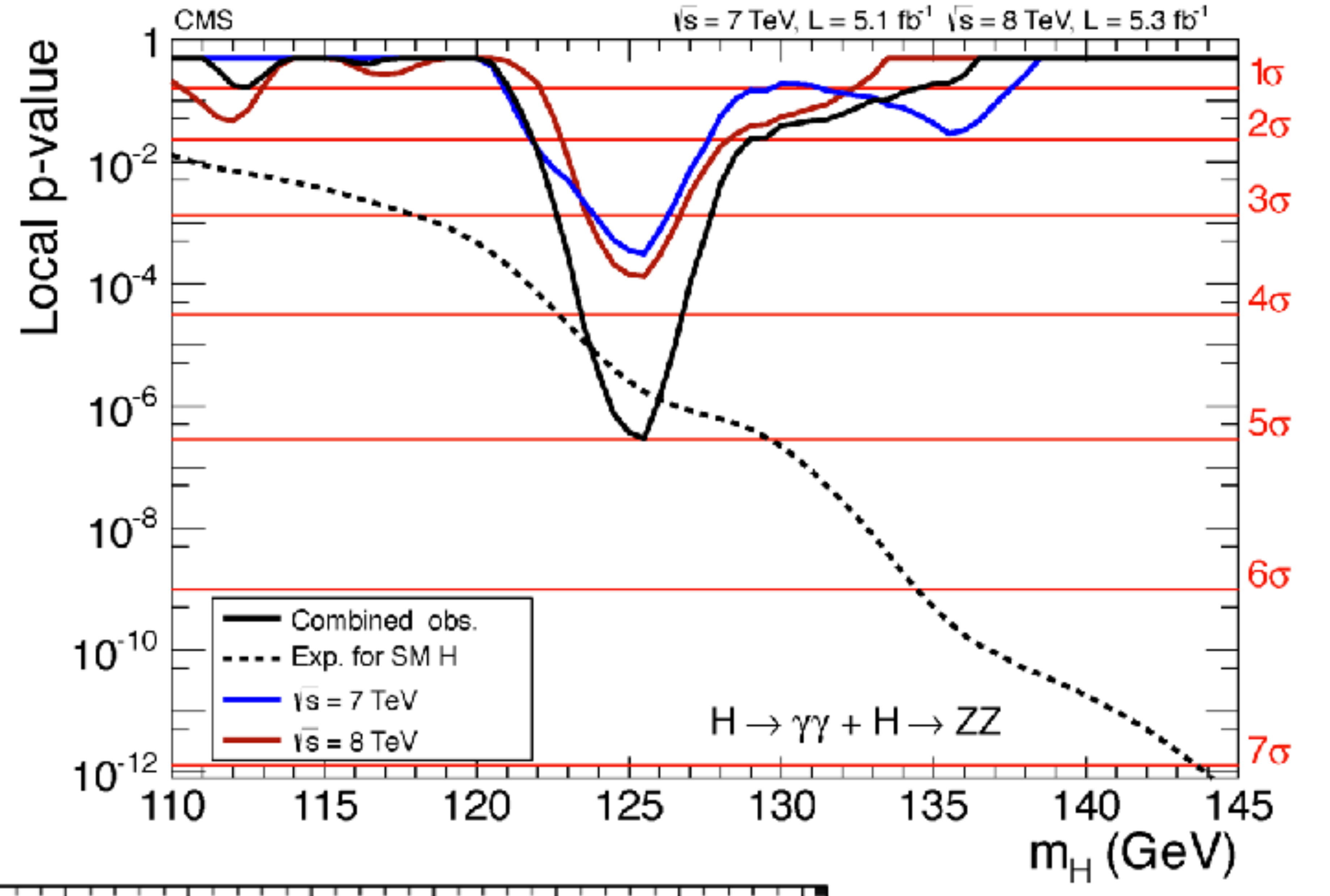
- *When you don't know if you have a signal, you first try to exclude it*
- *If the signal is there, your limit will be poor (and worse than expectation)*
- *If it is much worse, you might have discovered a signal…*
- *… or you might have discovered that your analysis is terrible*
- *these plots are not the right plots to establish the presence of a signal*

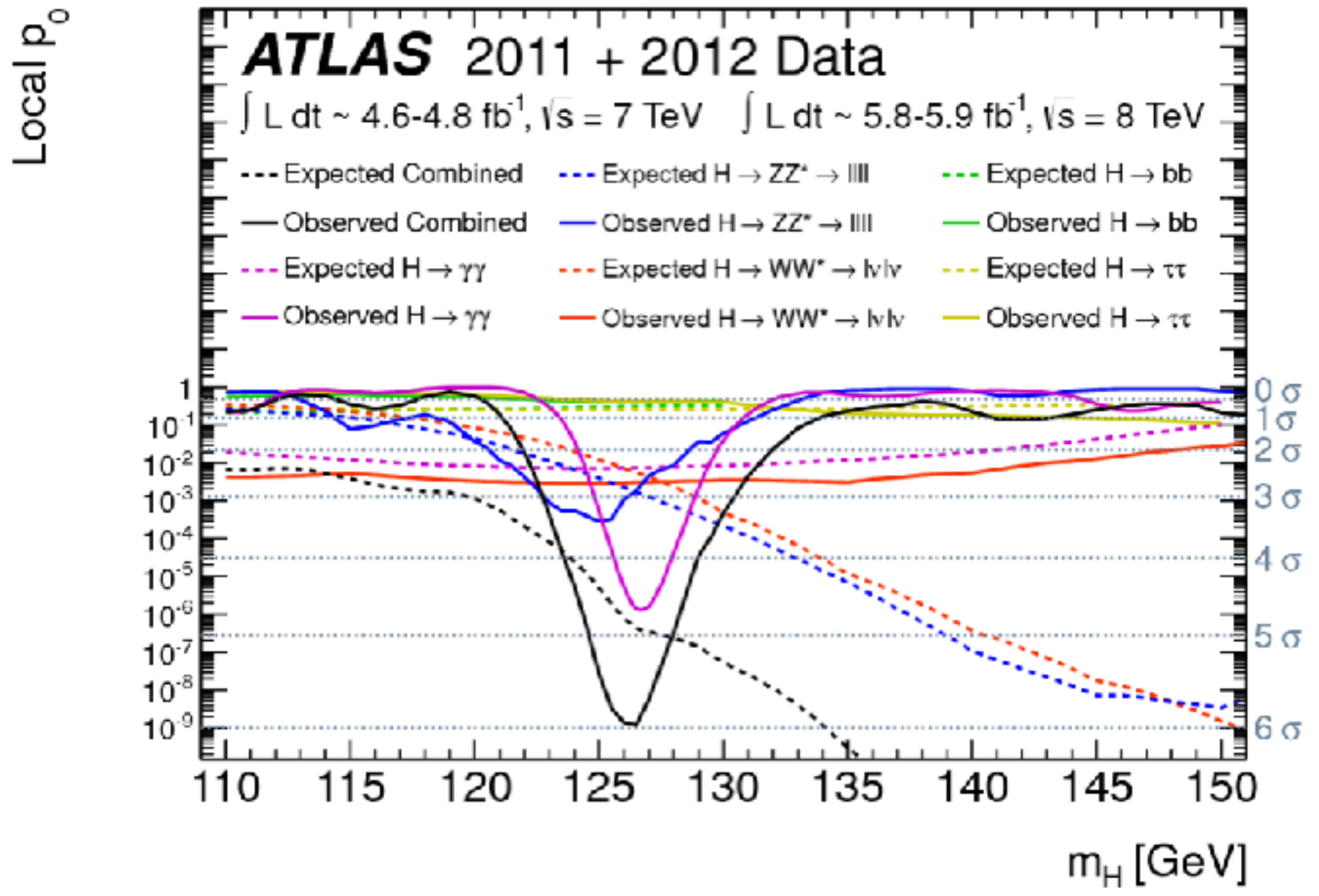




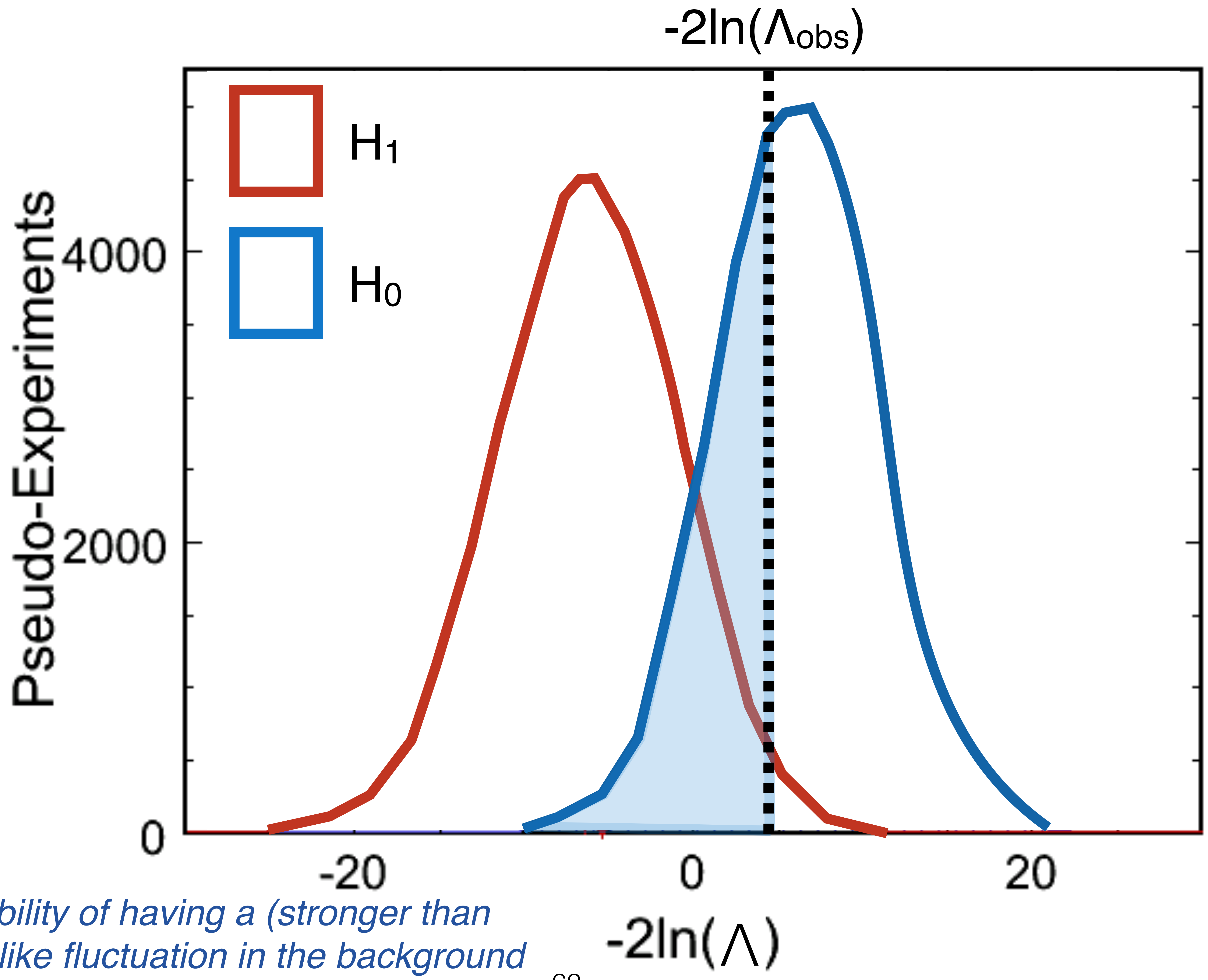
ASSESSING A DISCOVERY

- *To claim a discovery, you need to exclude the possibility that your background could mimic a signal*
- *To do so, you measure (with toy experiments) the probability that a bkg-only sample gives a result as signal-like as what you see on data*
- *The signal is stringer than the conventional 5σ threshold so…*



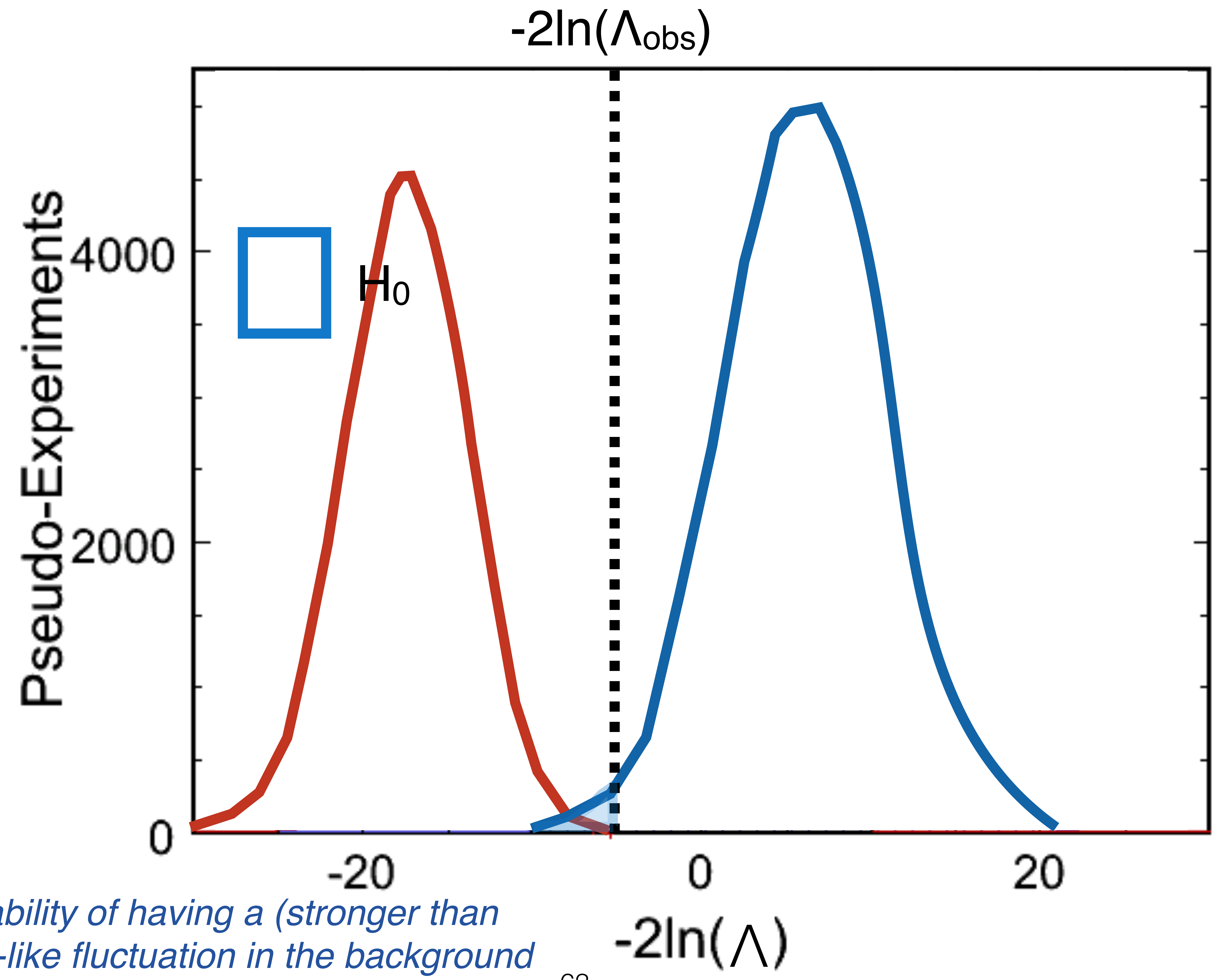


BACKGROUND P-VALUE



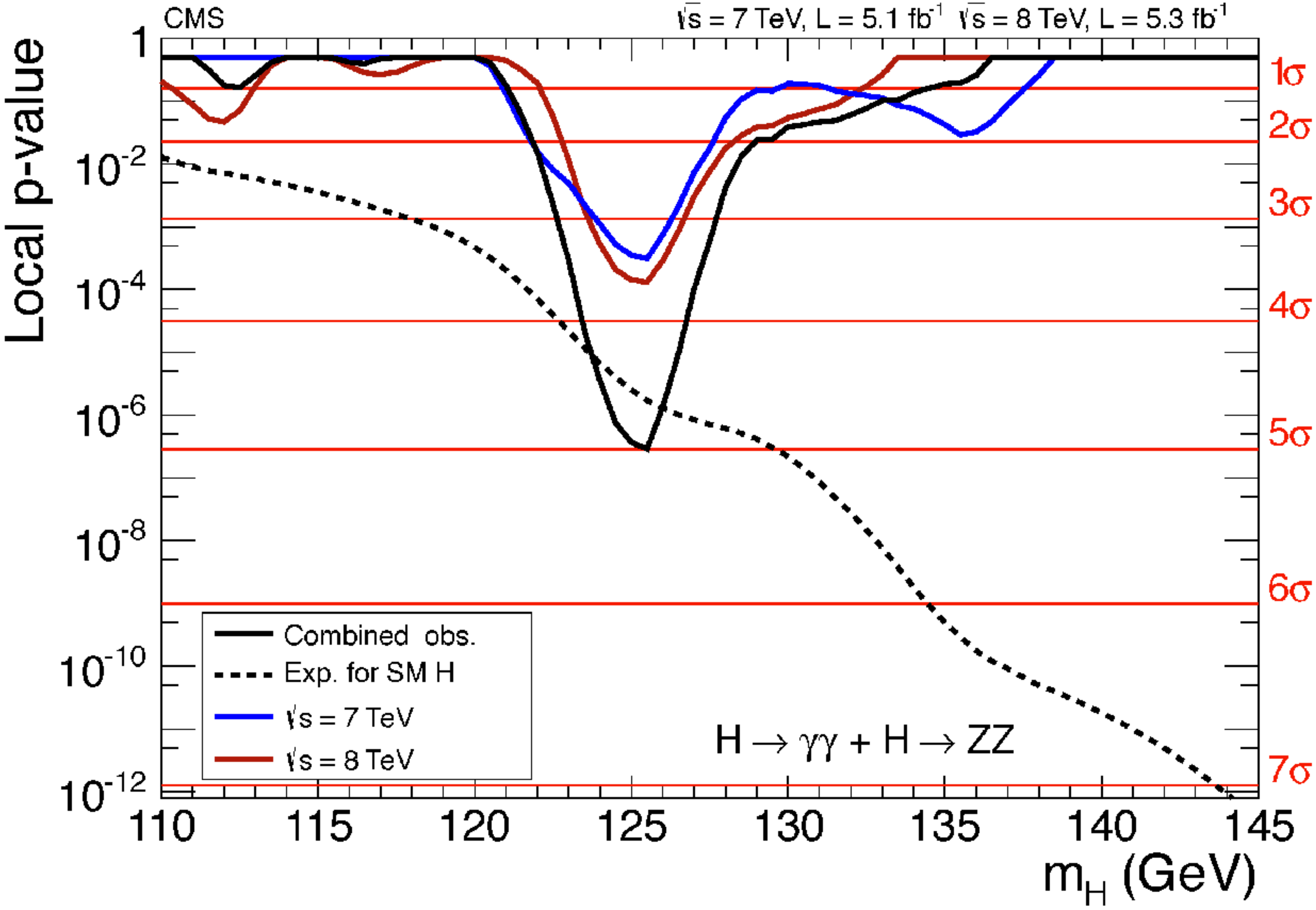
p-value =probability of having a (stronger than observed) signal-like fluctuation in the background

BACKGROUND P-VALUE



p-value =probability of having a (stronger than observed) signal-like fluctuation in the background

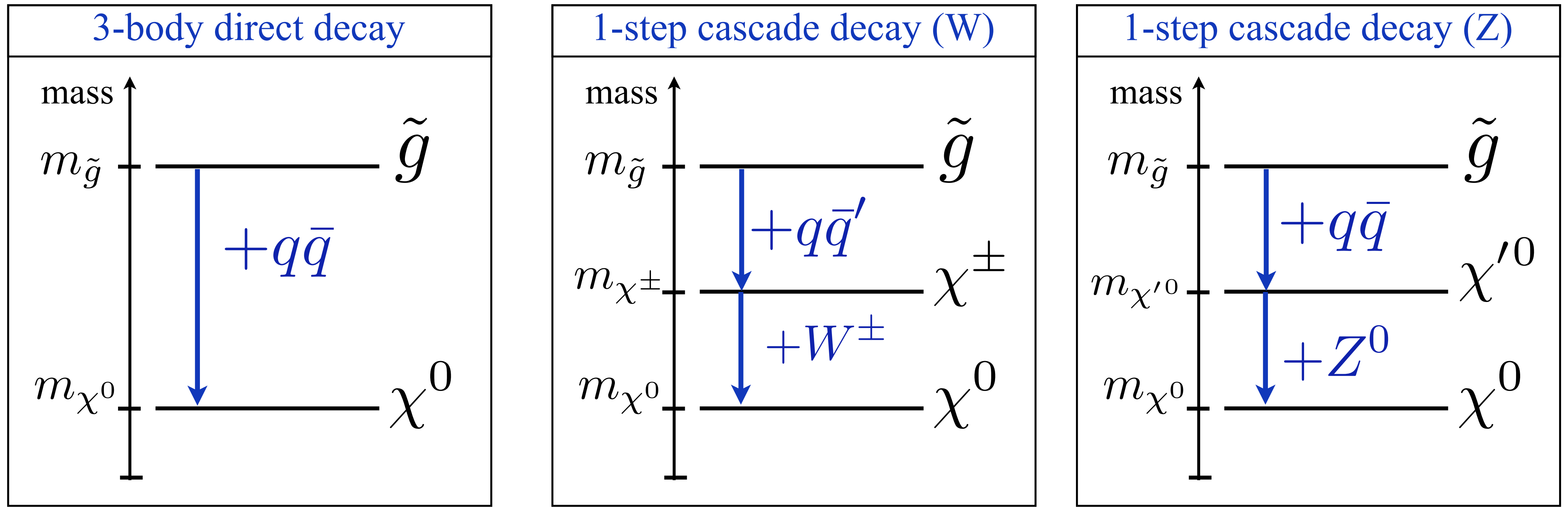
NUMBER OF SIGMAS

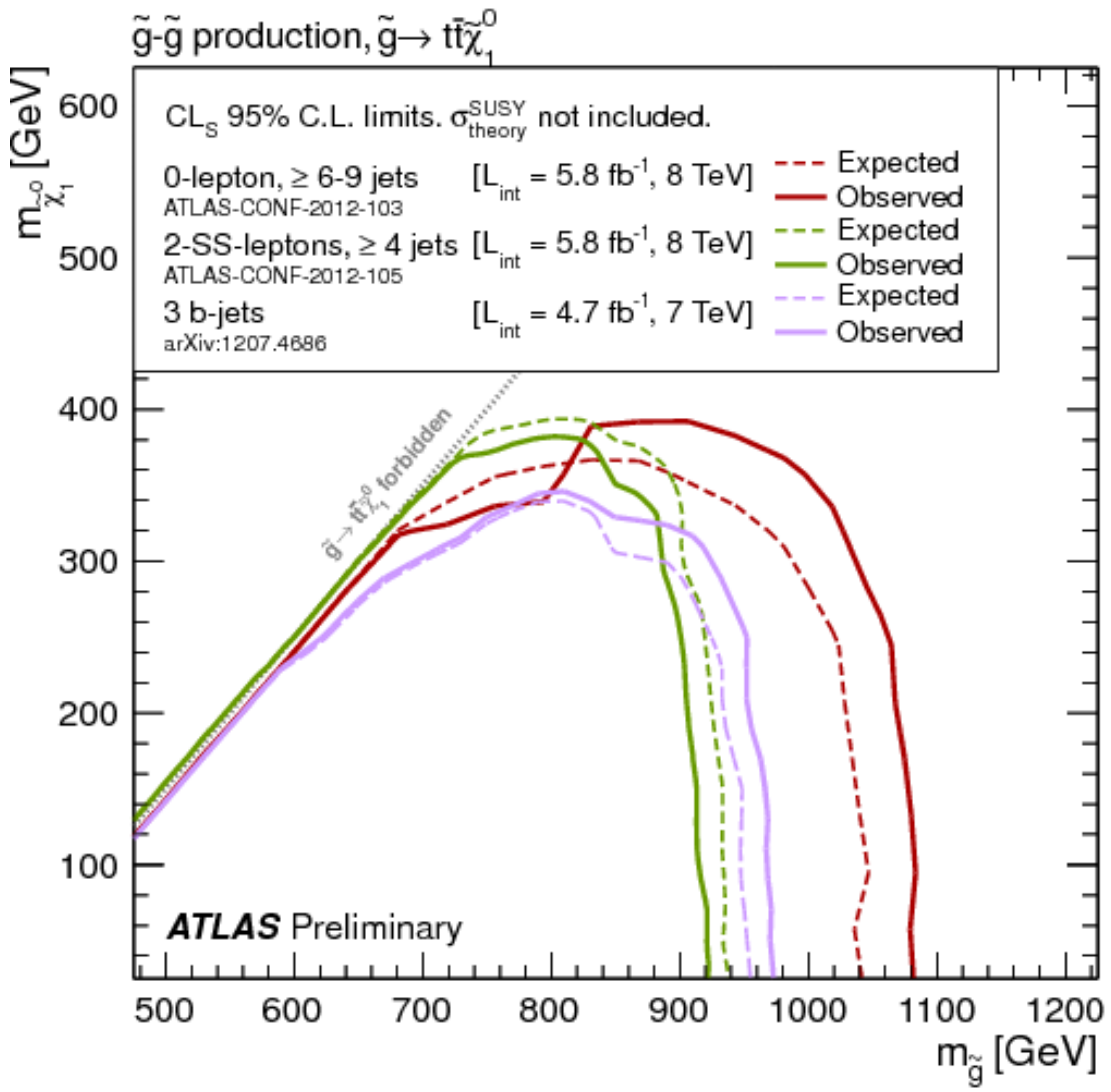



CHECK LIST FOR
MY SUSY ANALYSIS
SIGNATURE
TRIGGER
RECONSTRUCTION
ANALYSIS STRATEGY
STATISTICAL ANALYSIS
INTERPRETATION

SIMPLIFIED MODELS

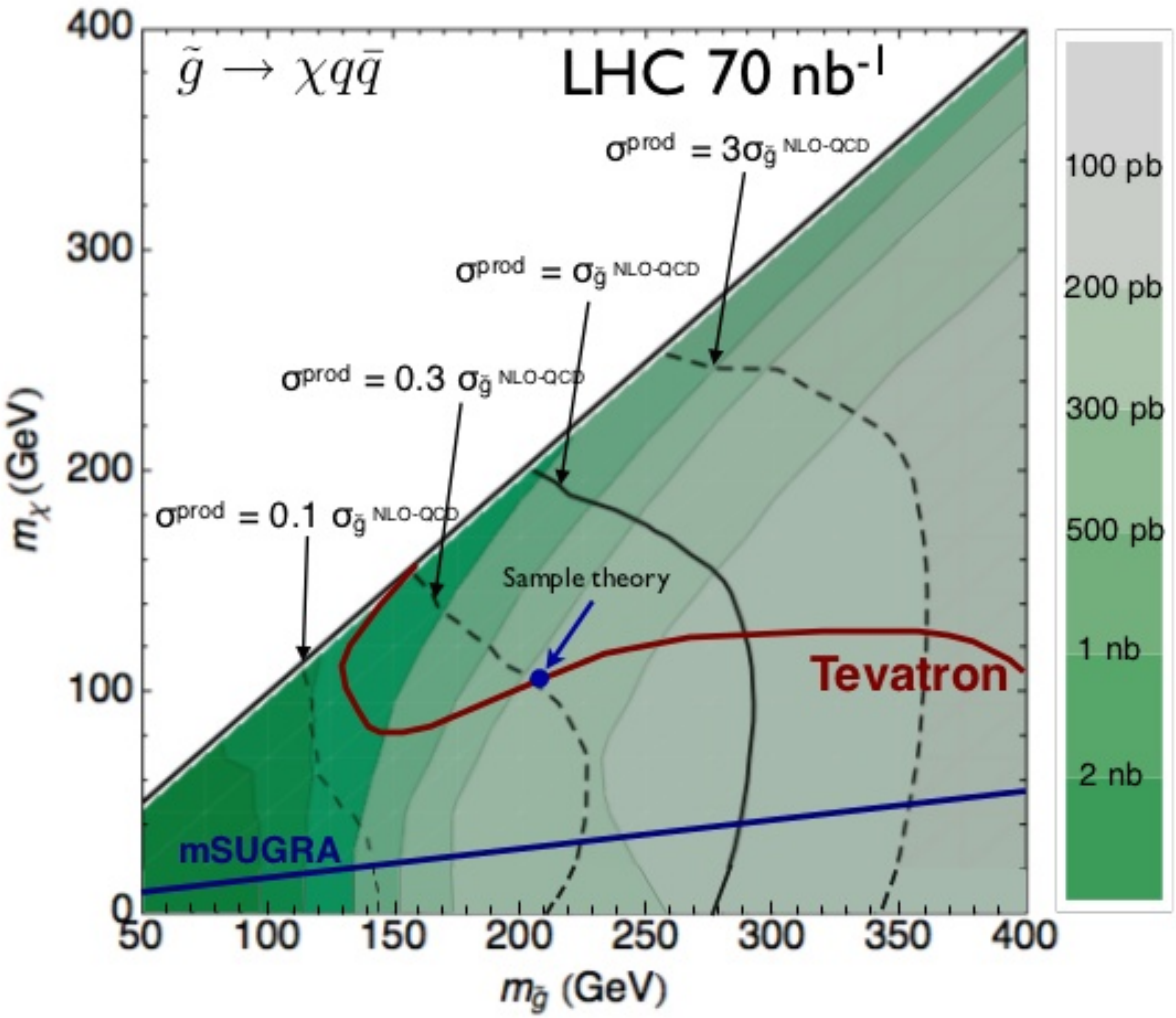
- Before the LHC, searches were entered on full models
- This changed during Run I. Simplified models became the standard
  - Focus on a specific process x decay chain
  - Interpret the analysis in this context





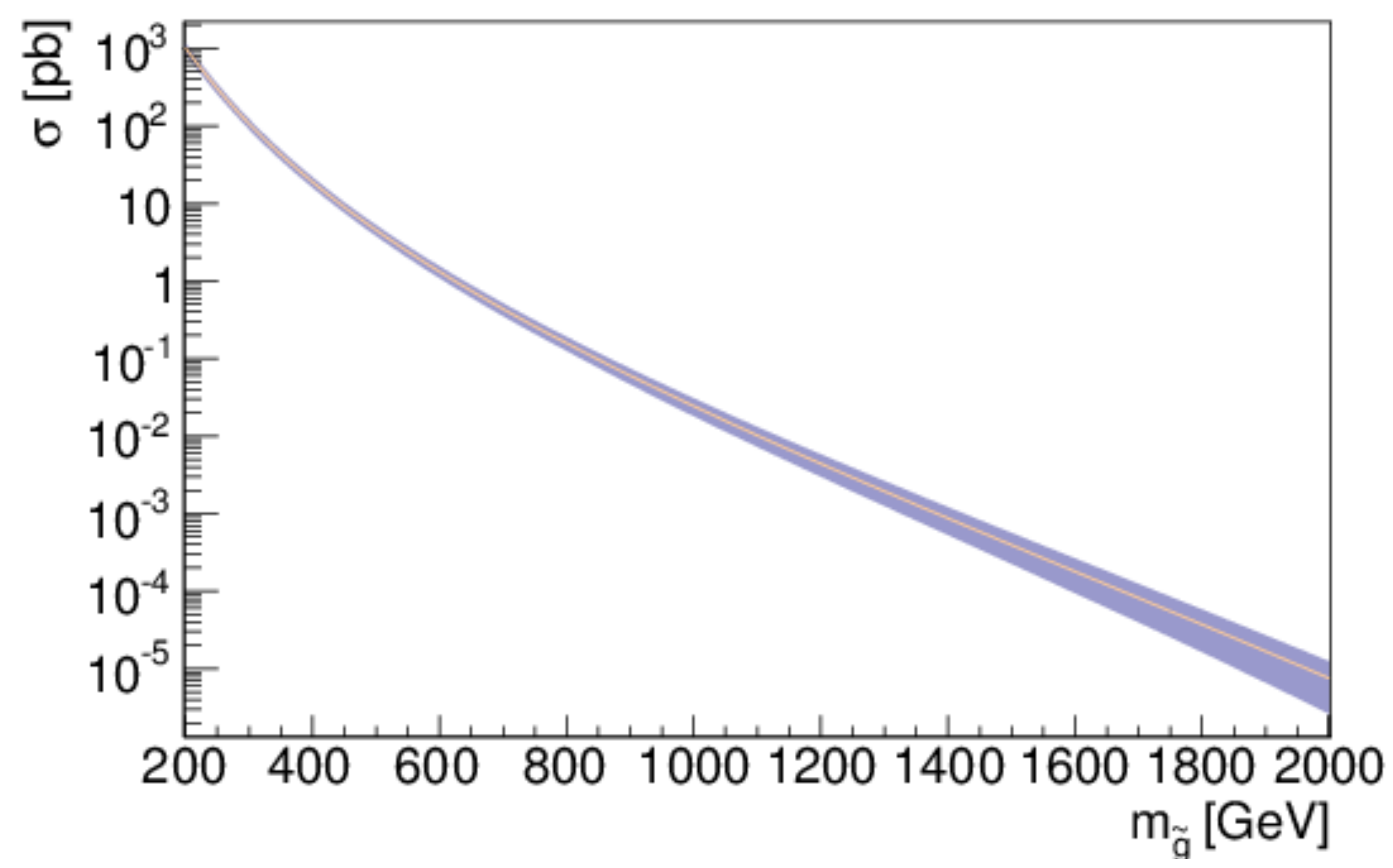
WHY IS THIS NEEDED? YES!

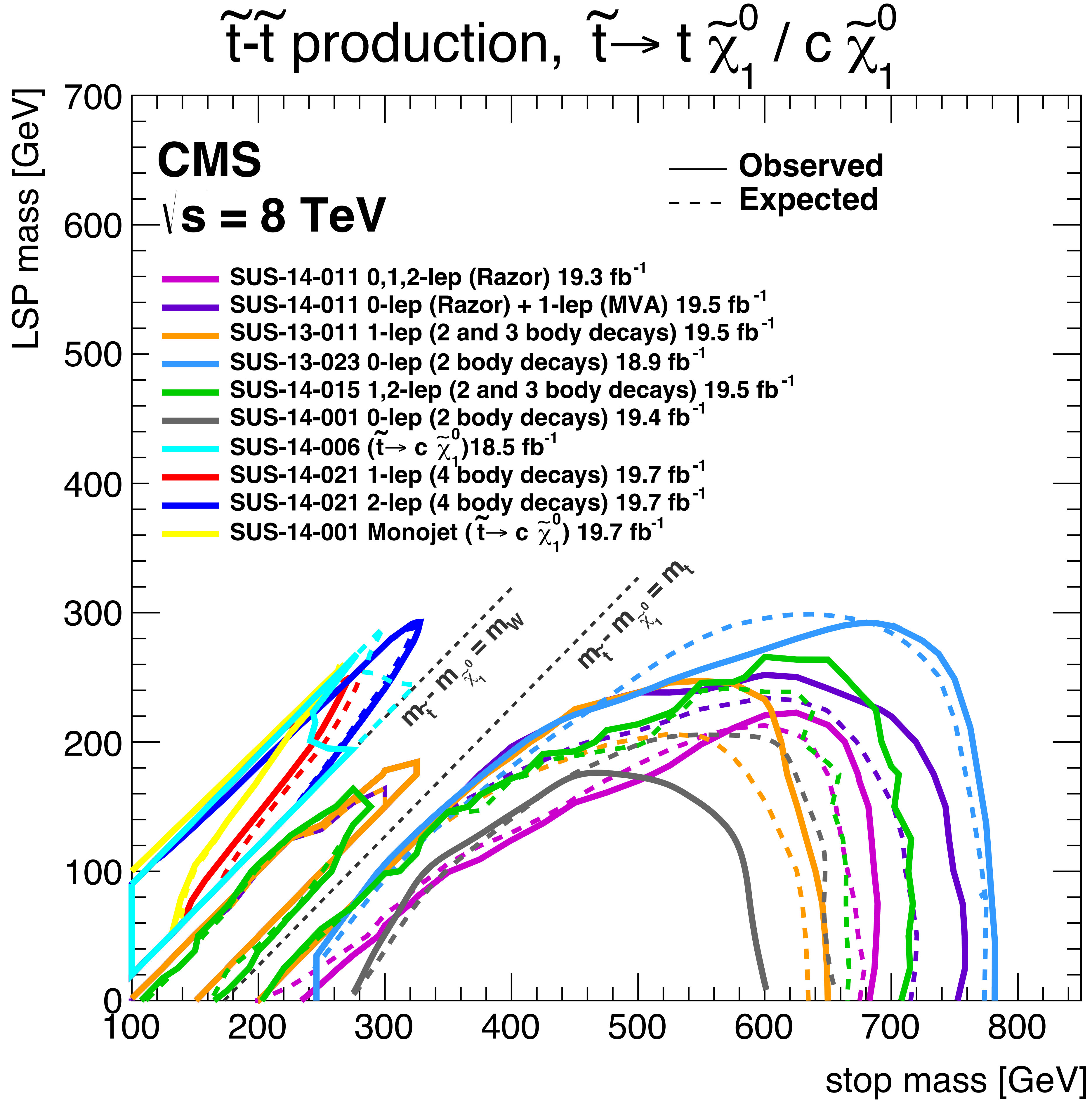
- *At the beginning of the LHC, many pre-LHC-data analyses were actually found to be too much tailored on the benchmark models*
- *Simplified models allowed to go beyond certain implicit assumptions*
- *This new paradigm allowed to discover weakness in the search program and design a next-generation set of analyses*
- *In general, the use of simplified models made our search strategy more robust*



MIND THE HIDDEN ASSUMPTIONS

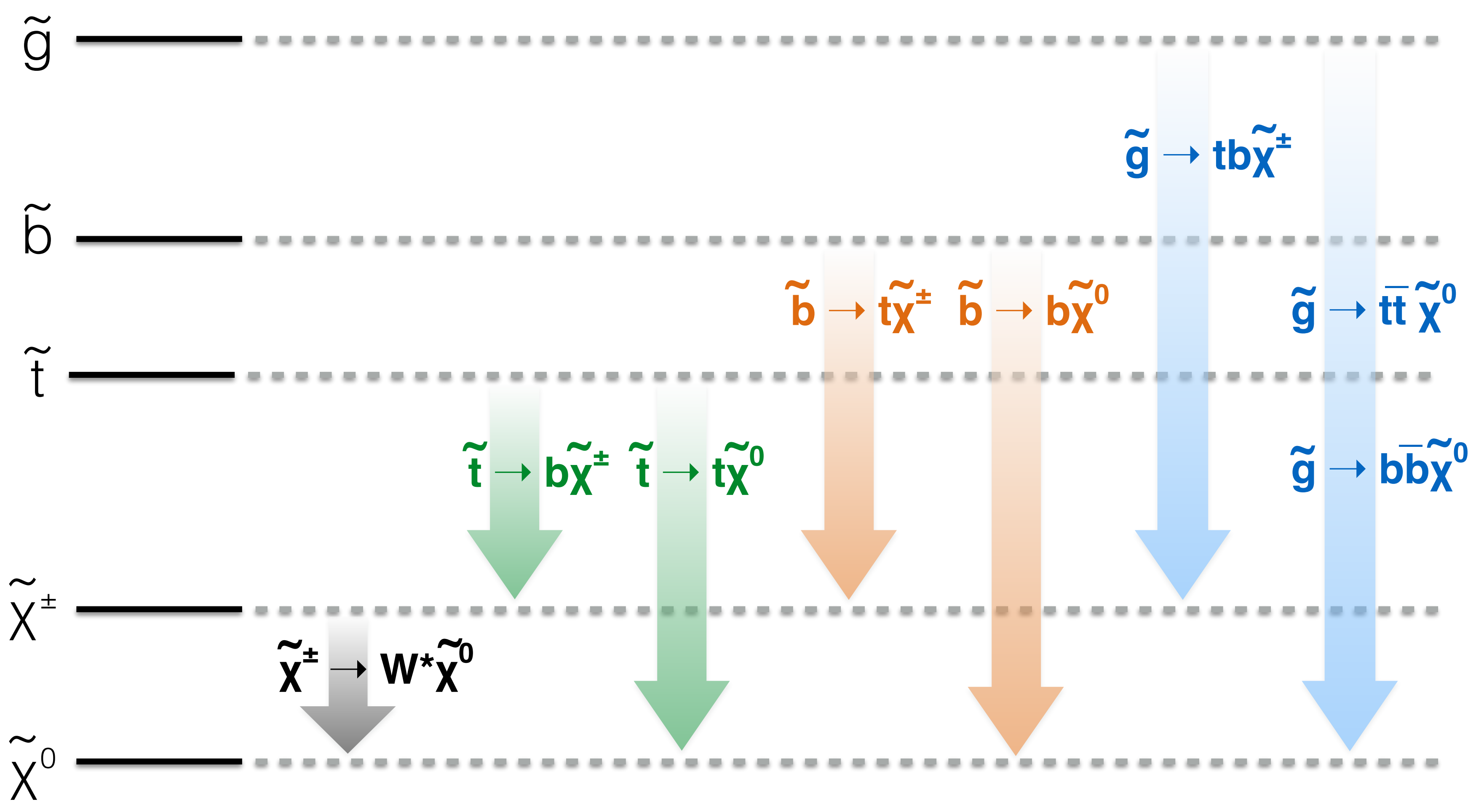
- *BRs are usually assumed to be 100%. This means that every line in a summary plot is implicitly excluding the others*
- *Cross sections are sometimes computed under special assumptions (e.g., decoupling limit) and don't hold in general*

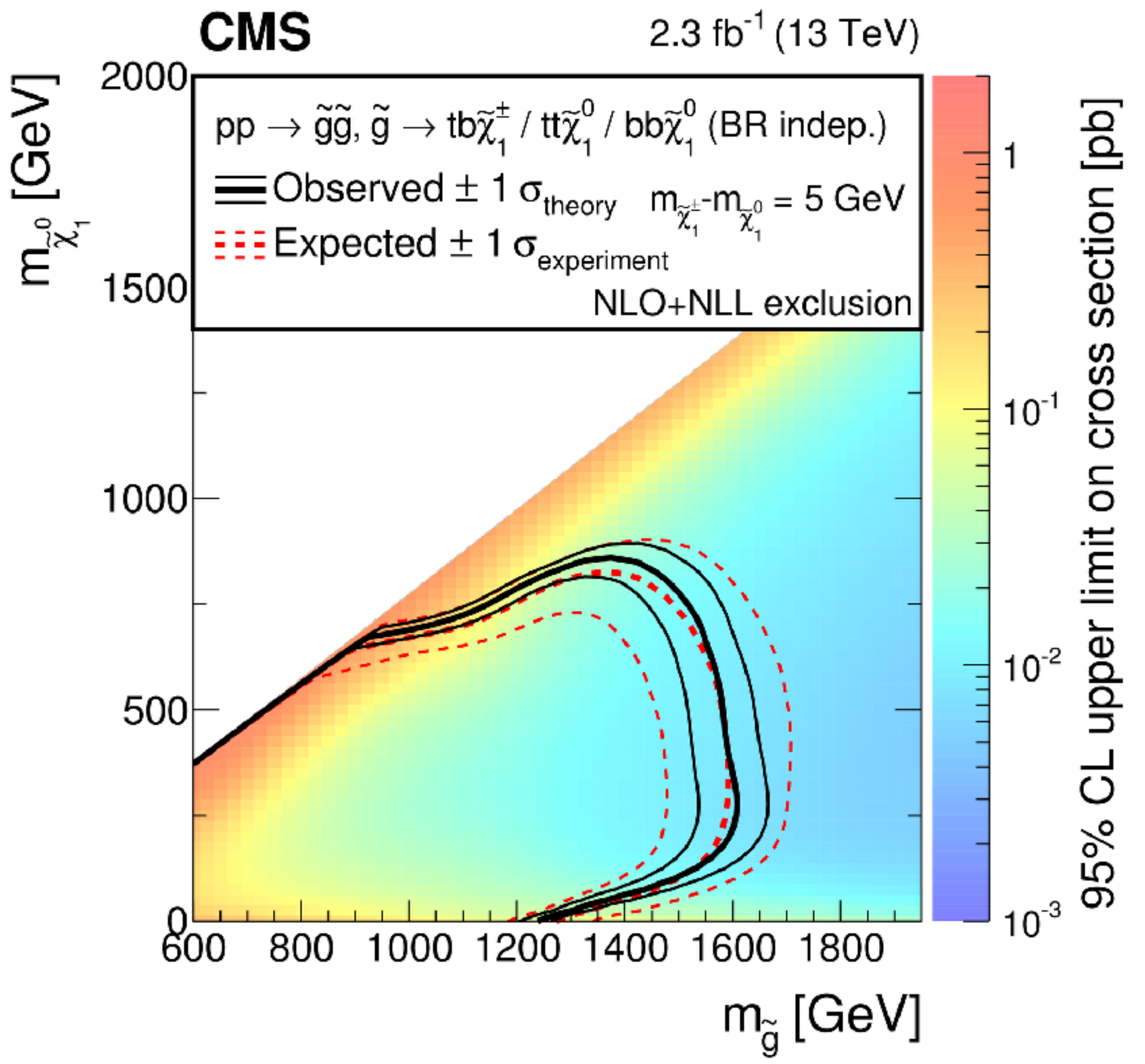




GOING BEYOND ASSUMPTION

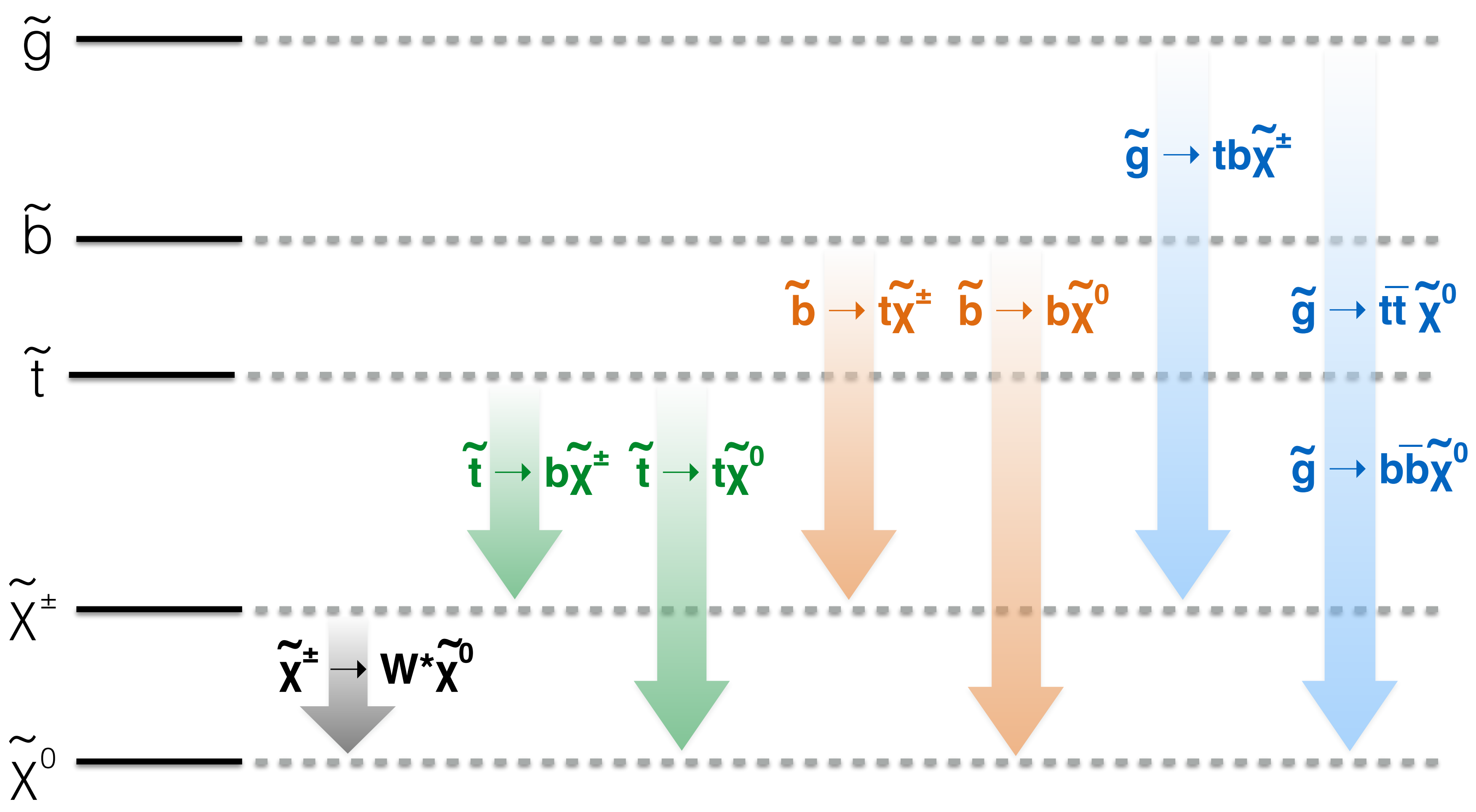
- Recent tendency is to present BR-independent results
  - easy in specific scenarios like Natural SUSY
  - e.g., done by scanning the BR space and quoting the worst result

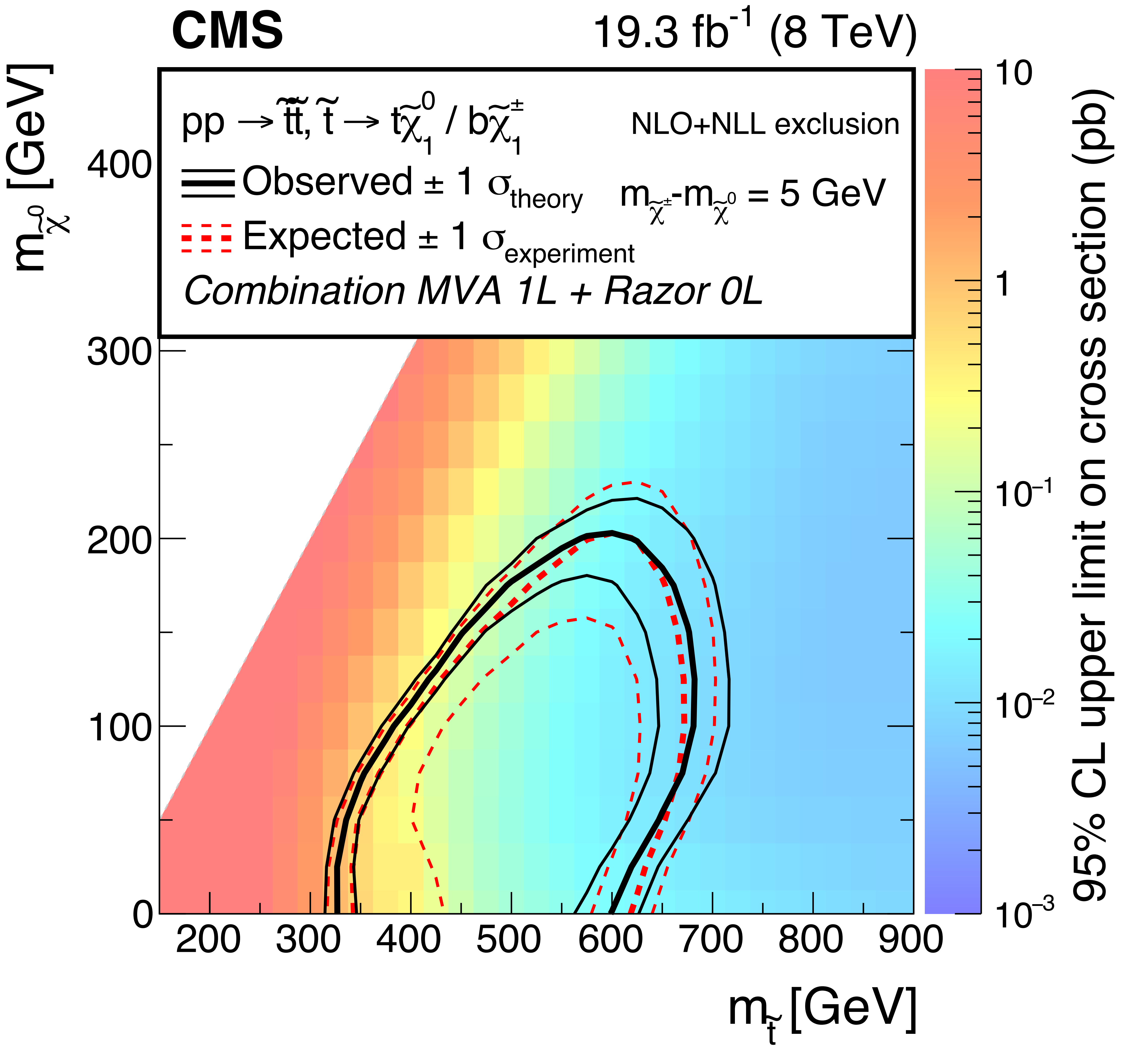




GOING BEYOND ASSUMPTION

- *Recent tendency is to present BR-independent results*
  - *easy in specific scenarios like Natural SUSY*
  - *e.g., done by scanning the BR space and quoting the worst result*





SUMMARY OF EPISODE I

- LHC has all it needs to be a SUSY discovery machine
  - can produce full spectra of particles
  - can observe many final states for any particle
- Practical limitations (e.g., trigger) should come into consideration when designing the analysis
- Data control samples are a key ingredient (a 100% MC-based background prediction would not be considered acceptable at LHC)
- Statistical tools in place from Higgs discovery
- Simplified models great guidance to interpret and improve searches, when taken with a grain of salt

BACKUP SLIDES