Experiment: Presentation of scientific results

Maurizio Pierini
CERN

Why This Talk

- I was asked to guide the discussion on this topic, summarizing what has been done so far
- I am not saying anything new and I am not the right person to do that
- I will take the opportunity to introduce a few points that we are all familiar with
- From this premises we will get to the recent Les Houches recommendation to standardize the presentation of experimental searches
- Most probably you all know this. I hope I will not be too boring

Interpreting a counting experiment

- The result of a counting experiment is an observed yield n given an expected background b
- Expectation comes with uncertainty db. ATLAS and CMS write the likelihood as

$$\boxed{L(n) = P(n|s+b)\mathrm{LogN}(b|bexp,db)}$$

- If n, bexp, and db are not quoted in a paper, then the paper is not describing the measurement. Experiments should be more rigorous there, at least submitting extra information with the paper
- When the numbers are provided, you don't have to care about the background. We are telling you how much is there
- Your only concern is how much signal your model would produce in my detector

The Signal Yield

- The observed yield is the product of three numbers

$$s = \text{eff} * \text{Lumi} * \text{xsec}$$

This you should know (it comes with the model you consider)

This is quoted (with it's error) in the papers

This is quoted for benchmark models, sometimes. For the generic case you need to know how the detector behaves

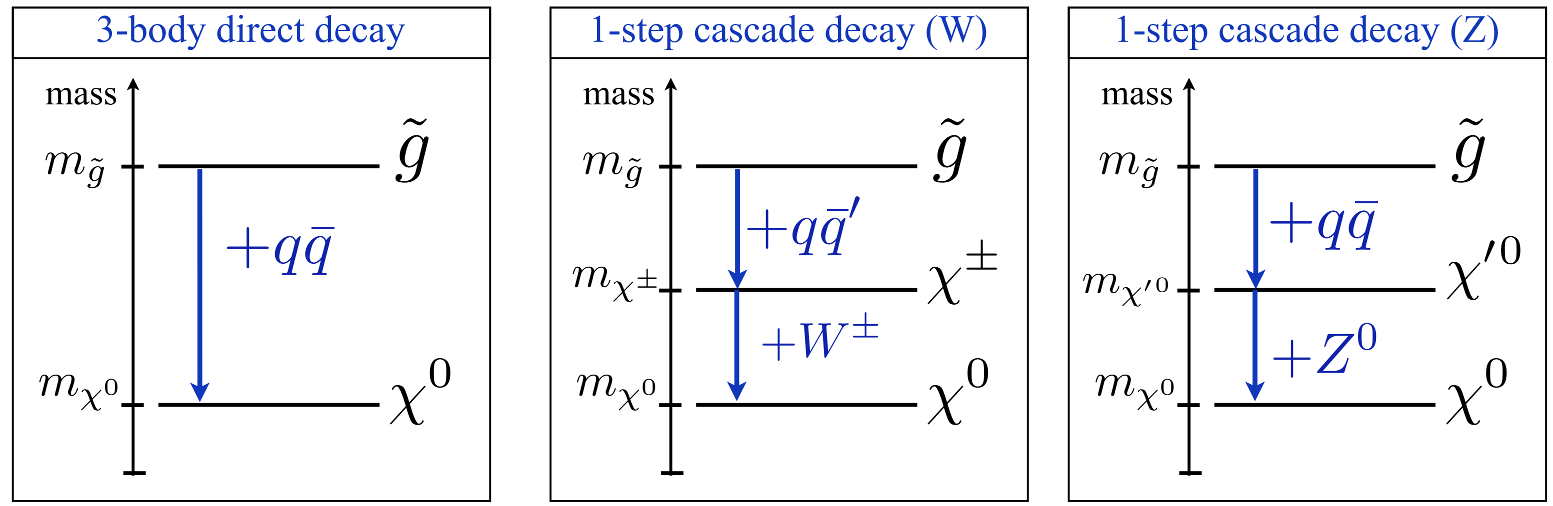
Getting the Efficiency

- The signal efficiency is usually the big mystery for someone external to an experimental collaboration
- Experiments are trying to be more open, adopting the idea of Simplified Models
- Whenever your generic model is the superposition of a limited set of models, you can get the answer right
- Whenever this is not the case, you can at least use the simplified models to get a calibration of your own mock detector
- And sometimes the generator-level information is enough

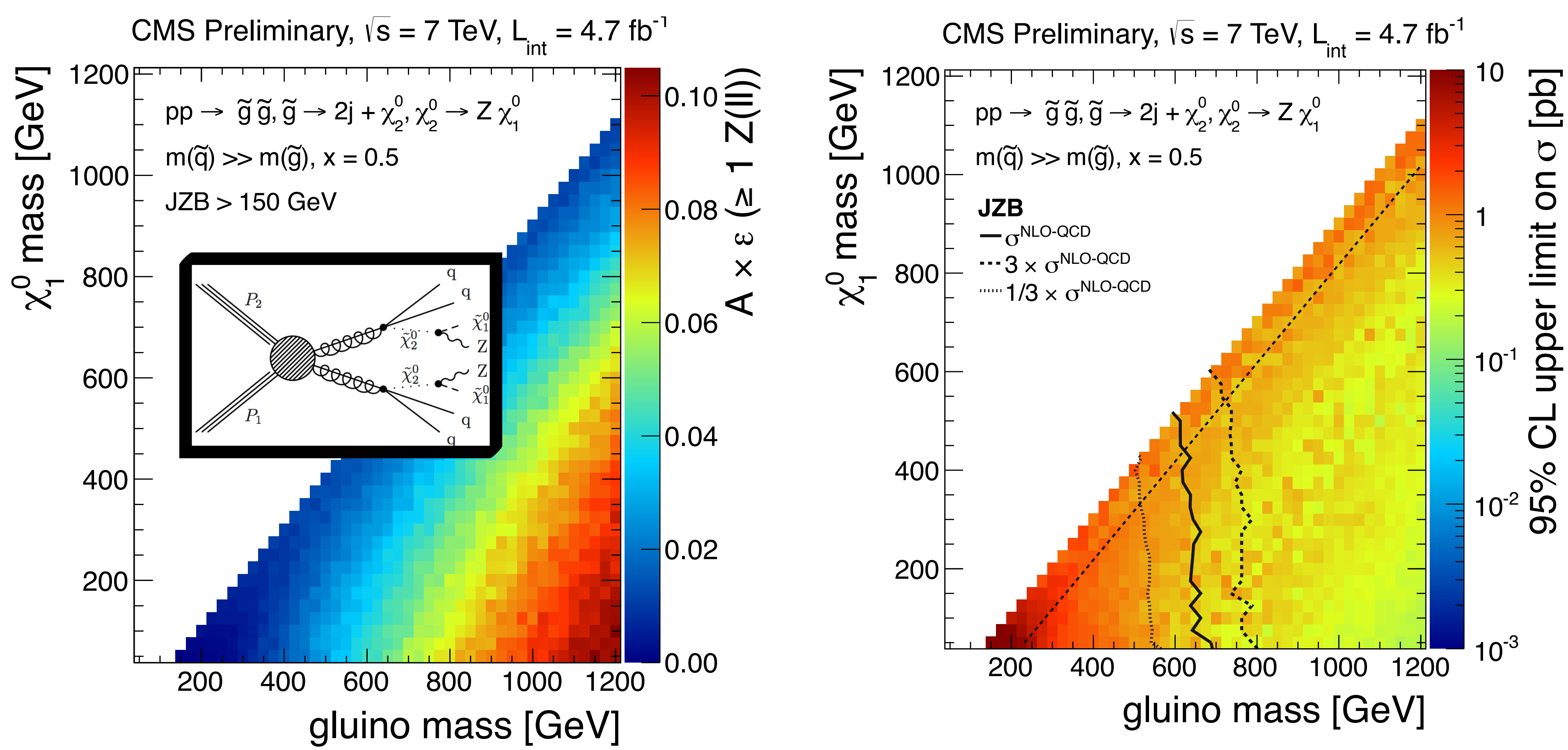
Simplified Model

- A simplified model is a model in which a few particles (e.g. of a SUSY spectrum) actually contribute to the LHC phenomenology, the others being heavier

http://arxiv.org/abs/1105.2838

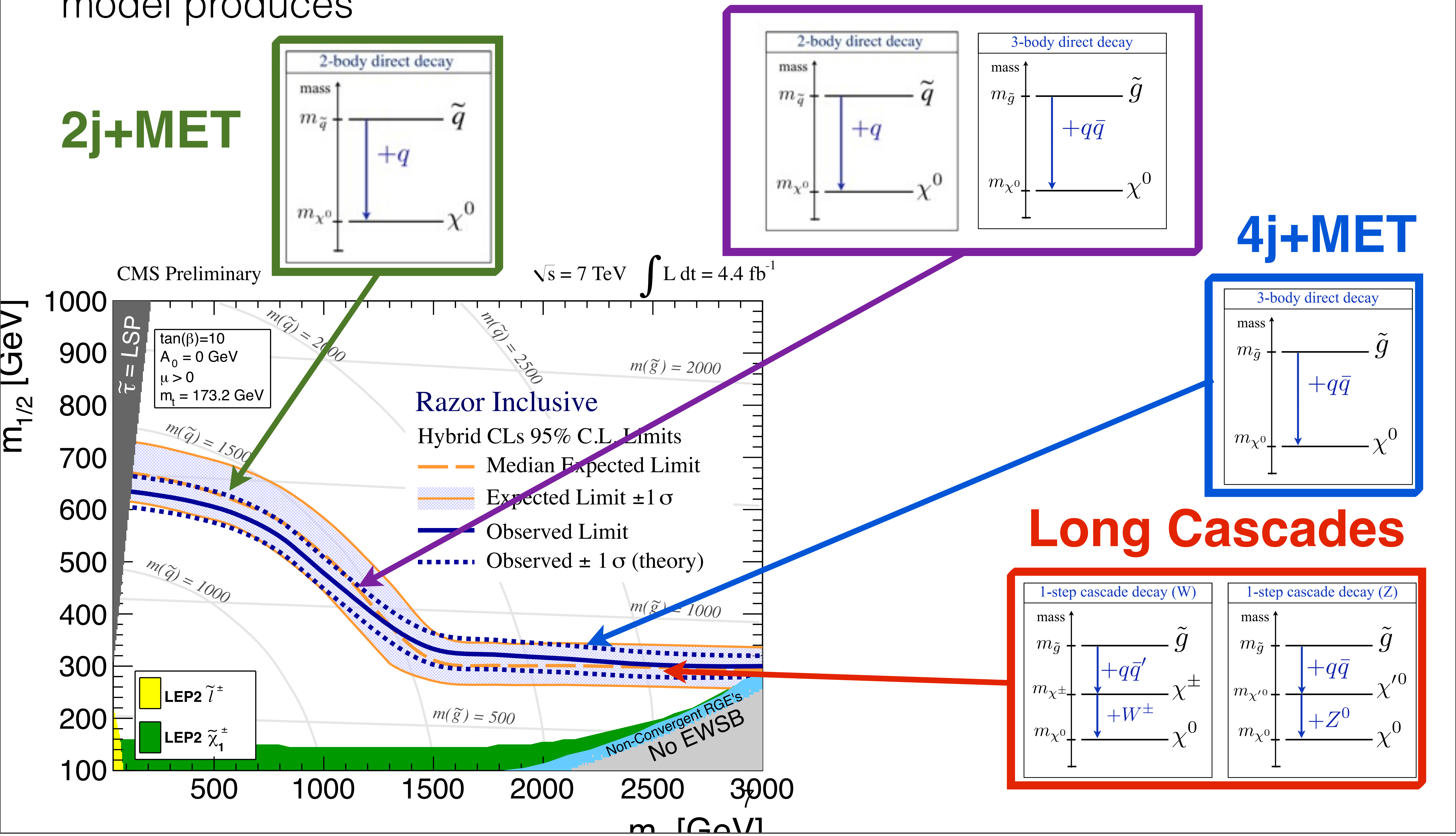


- Result on a simplified model is given with two plots: efficiency and excluded cross section



Combining Simplified Models

- We take CMSSM as an example of a “full model” which we want to study
- Rather than scanning the model parameters, we can scan the signatures this model produces

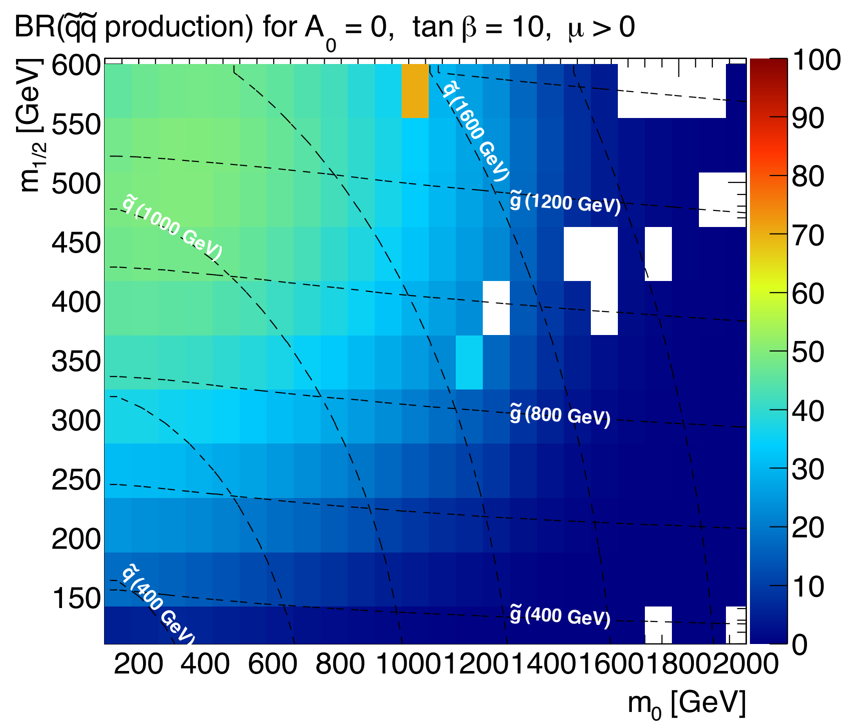


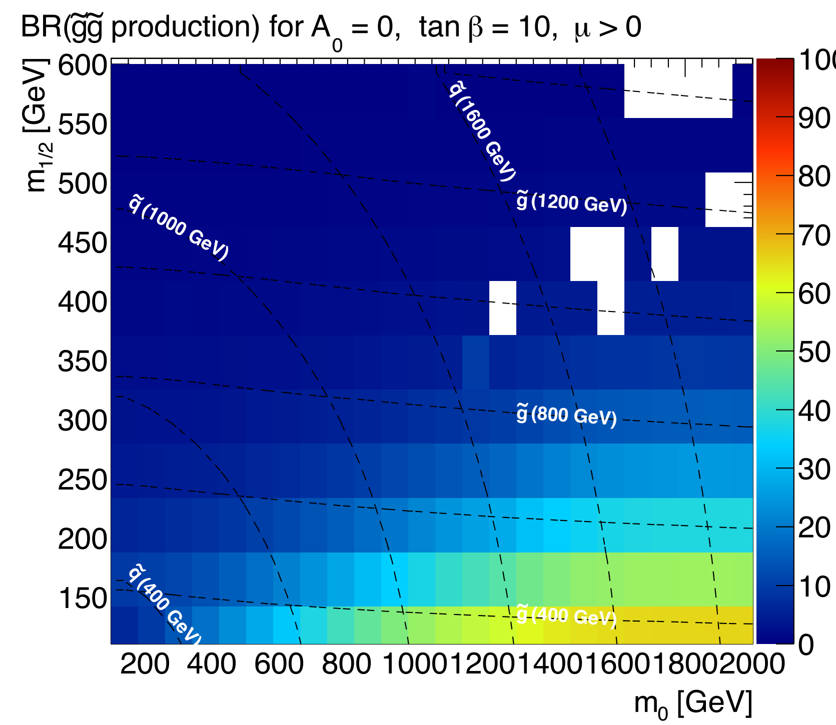
Combining Simplified Models

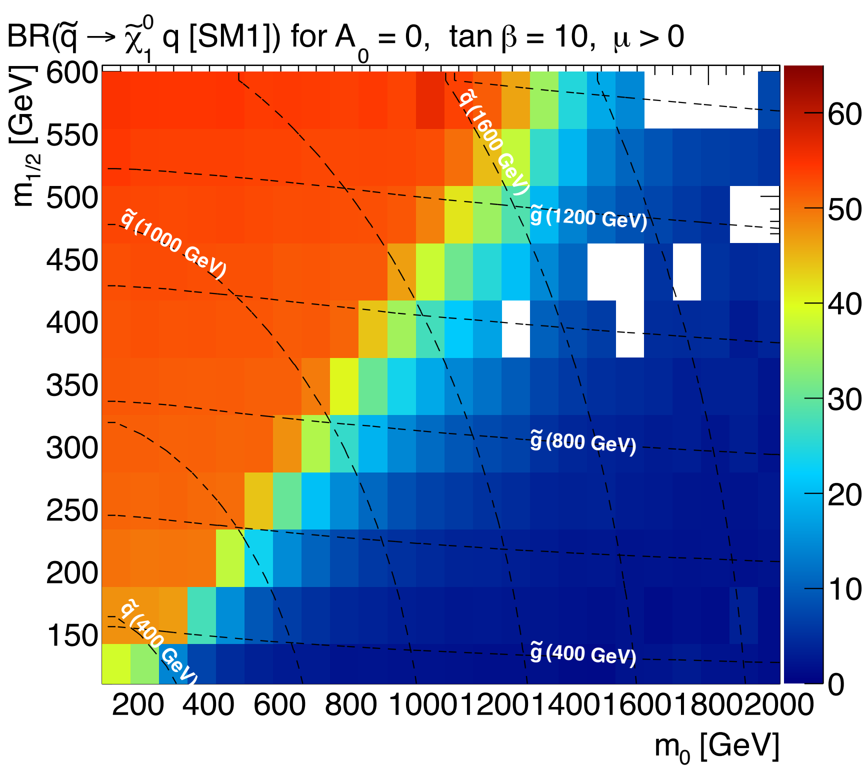
- One can then deduce the CMSSM limit by combining a limited set of SMSs

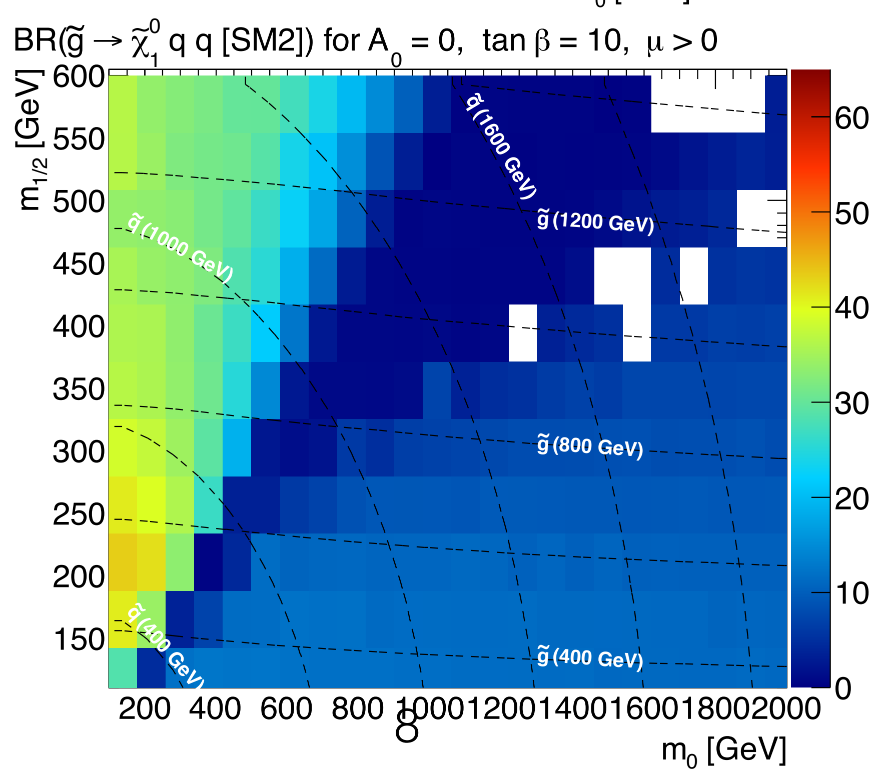
C. Gütschow, Z. L. Marshall
http://arxiv.org/pdf/1202.2662.pdf

different cross sections for different processes





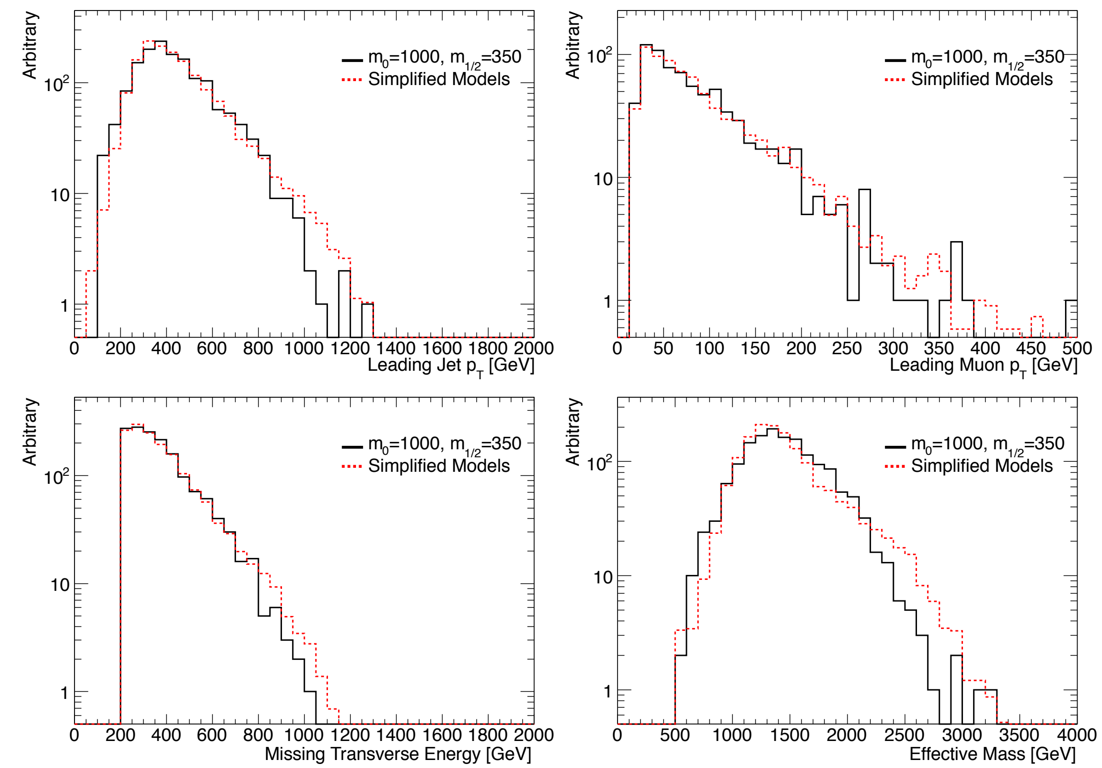


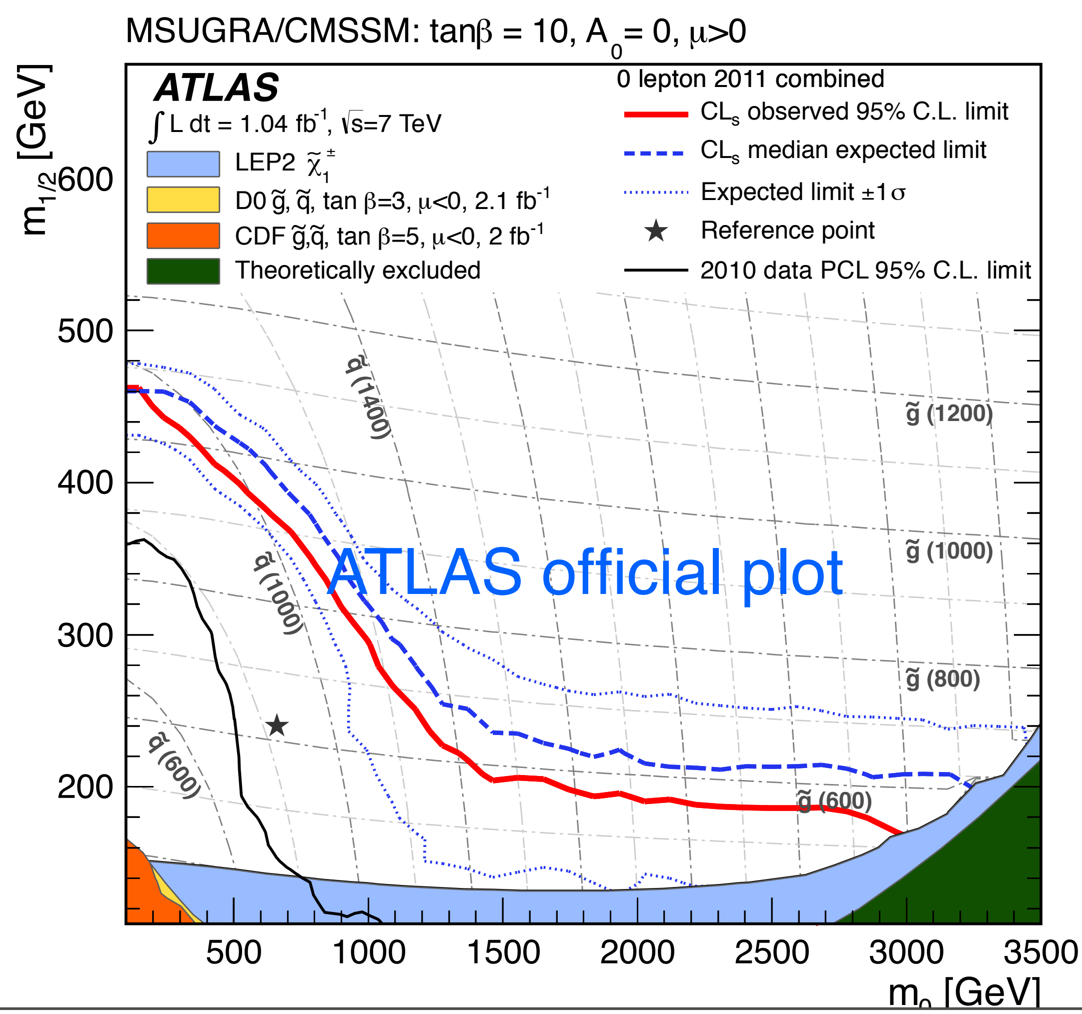


SMSs need to be reweighted by the BRs

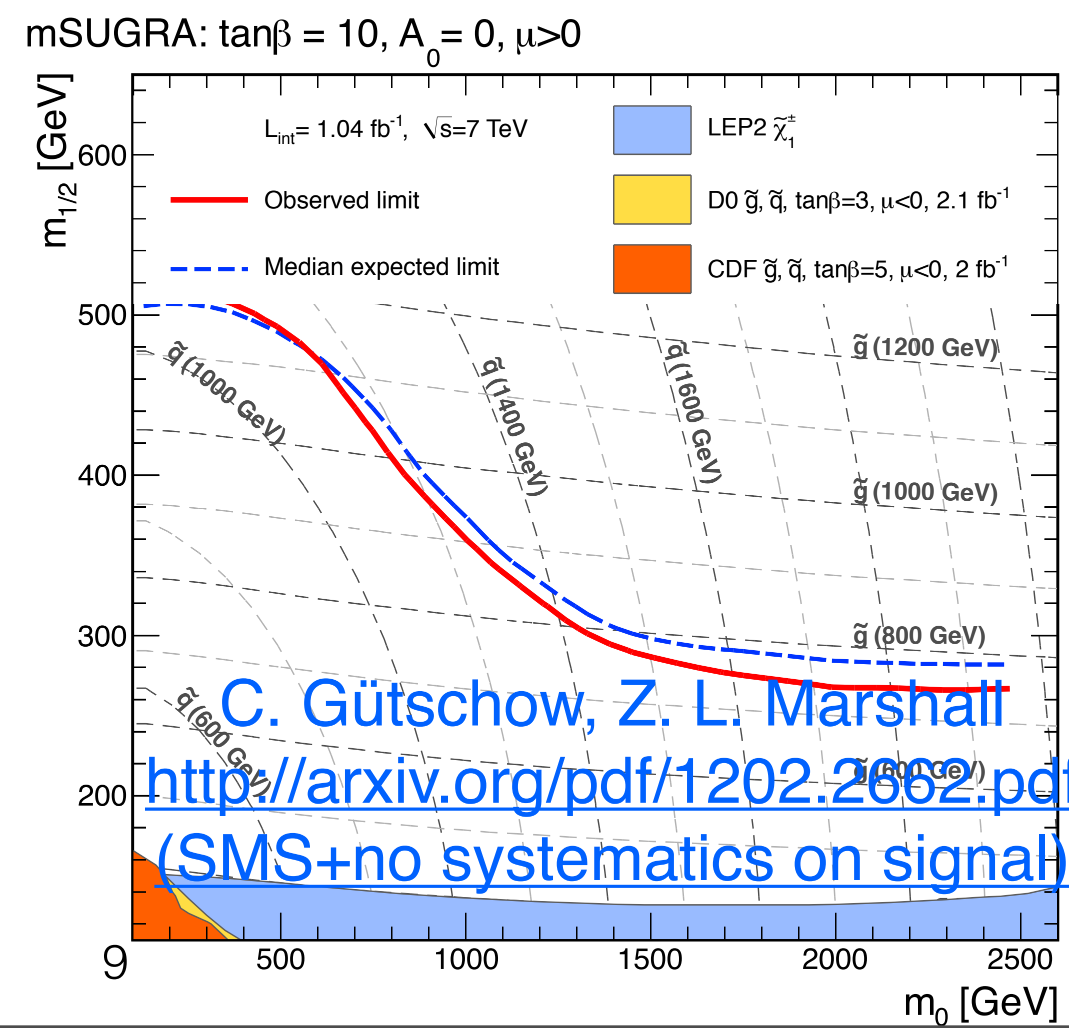
Combining Simplified Models

- The combination of the different SMSs allows to reproduce the distribution of kinematic variables to a good level of accuracy





ATLAS official plot



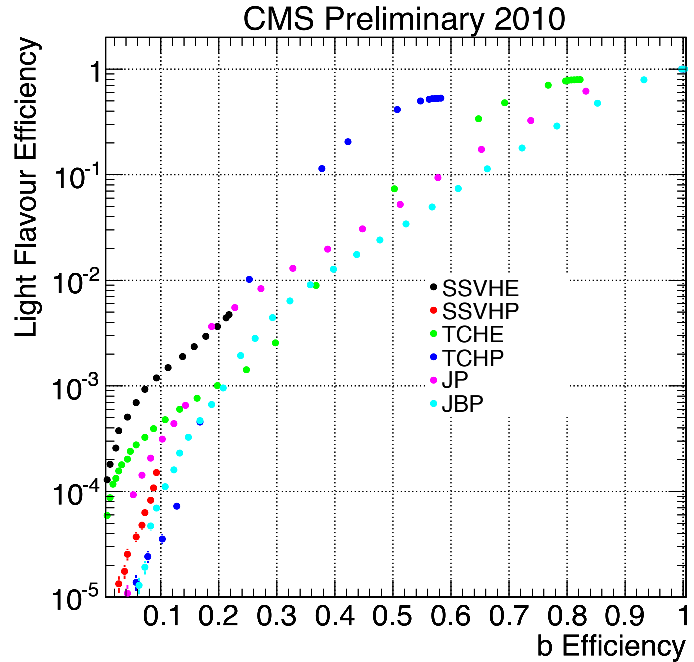
C. Gütschow, Z. L. Marshall
http://arxiv.org/pdf/1202.2662.pdf
(SMS+no systematics on signal)

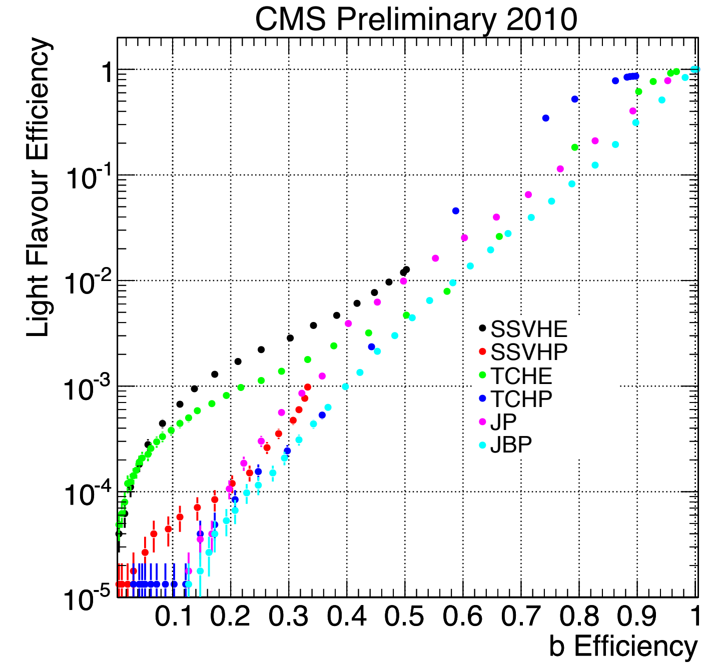
A Few Remarks

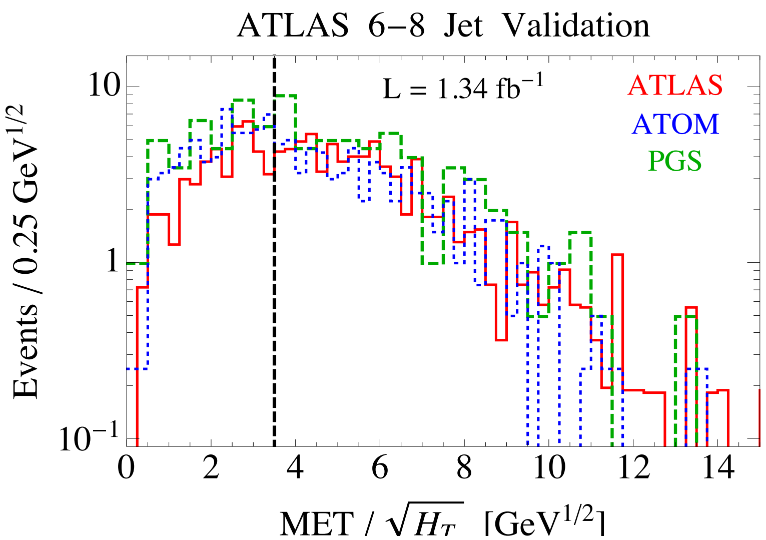
- CMSSM is a very special case of over-simplified model
- The same exercise with a more generic case might result in a number of SMSs which is bigger than the free parameters of the model
- Bottom line, the simplified model can be a good solution to avoid the details of the detector simulation, if and only if the relevant ingredients are provided by the collaborations
- Imagining that ATLAS and CMS can generate the needed plots for ALL the possible decay chains is just not possible, so the simplified model cannot be the ultimate solution, even if it is a good step in the right direction
- For sure the SMSs allow good-willing theorists to do more

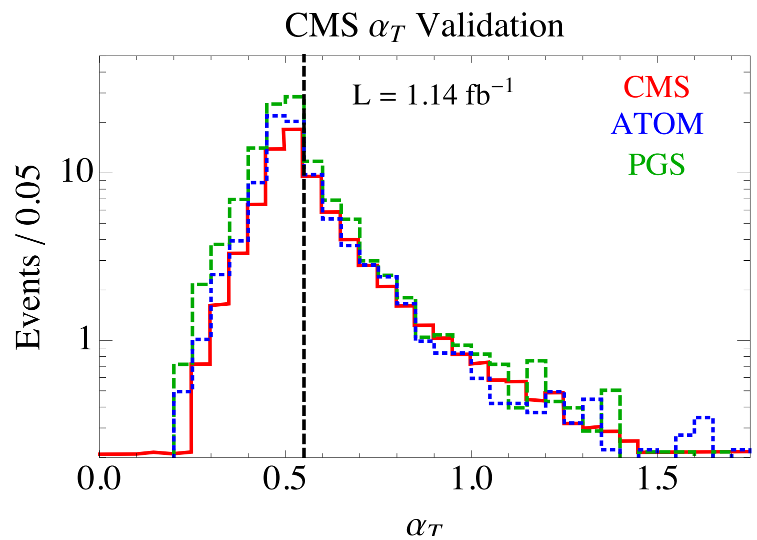
Simplified Models for Calibration

Detector performances have been presented for basically all objects (jets, met, etc). This can be used to write a detector simulation



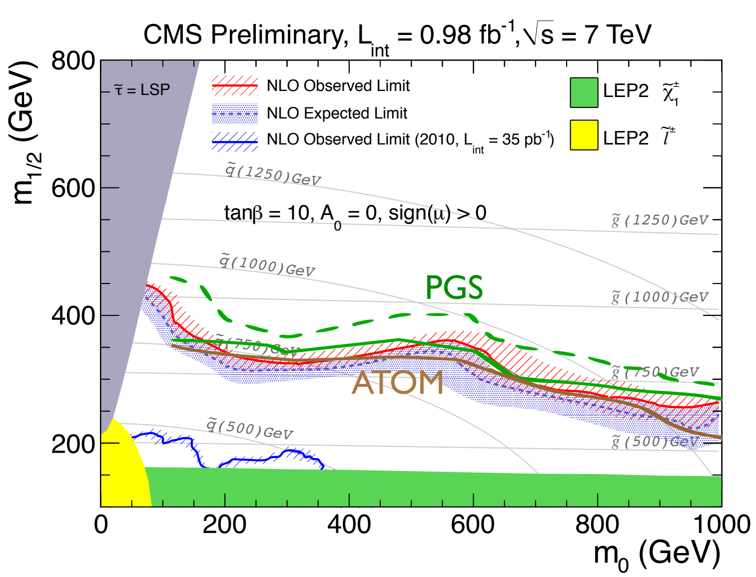


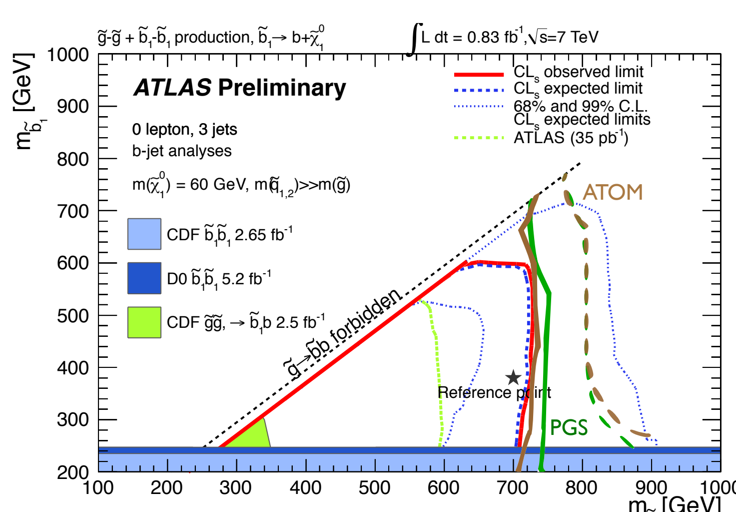




Papers come with plots of kinematic variables that one can use validate a detector simulation

M. Papucci, J. T. Ruderman, A. Weiler
http://arXiv.org/abs/arXiv:1110.6926





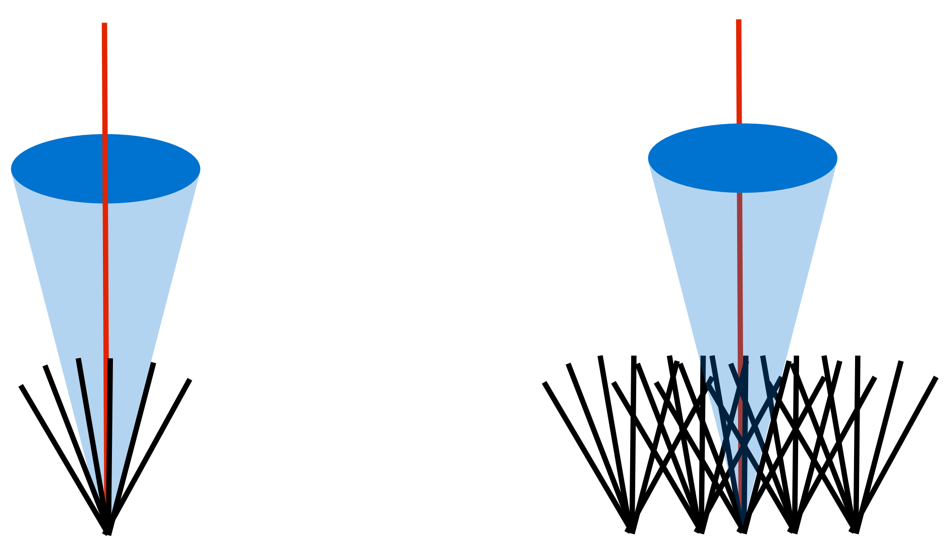
One can then try to reproduce the limits, as a validation of the detector simulation (and analysis implementation)

What Can The Experiments do?

~~Rather than waiting for 1000 theorists in the world to develop 1000 simulations of the detector, ATLAS and CMS could try to provide one that could be used by everybody~~

“NO WAY!!!!!!”

The experiments could guide the theorists in their exercise with some outreach study. Detector resolution and efficiency is not the end of the story. Particularly difficult is to emulate an isolation. In this case one needs extra information



Things like pileup are not easy to get in a simplified detector simulation
These things affect the performances
The effect is analysis-dependent
More information has to be provided

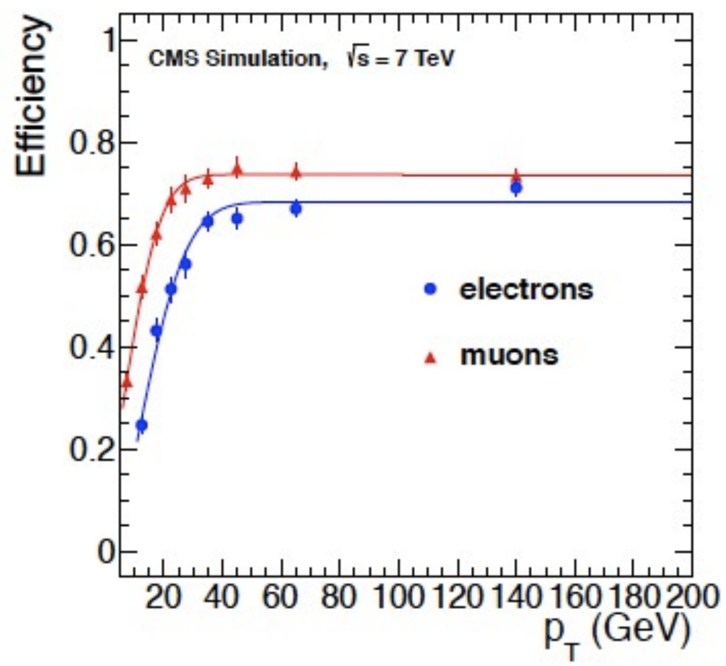
Interpretation Outreach

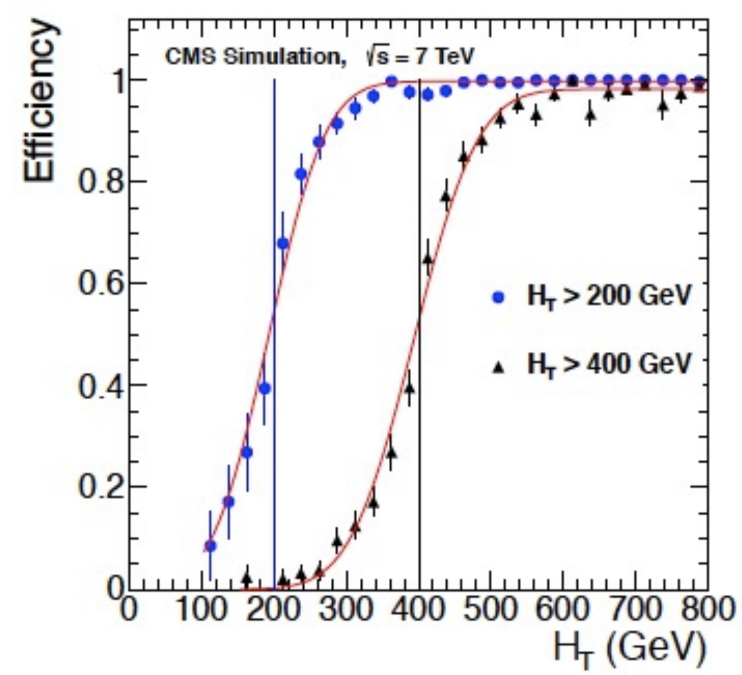
With a few lines of text and some plot, experimentalists can guide the theorists to understand the analysis

CMS SUS-11-011

9 Additional Information for Model Testing

Other models of new physics in the dilepton final state can be confronted in an approximate way by simple generator-level studies that compare the expected number of events in 0.98 fb^{-1} with the upper limits from Section 8. The key ingredients of such studies are the kinematic requirements described in this note, the lepton efficiencies, and the detector responses for H_T and E_T^{miss}. The trigger efficiencies for events containing ee, $e\mu$ or $\mu\mu$ lepton pairs are 100%, 95%, and 90%, respectively. The muon identification efficiency is $\approx$ 96%; the electron identification efficiency varies approximately linearly from $\approx$ 60% at $p_T = 10\,\mathrm{GeV}/c$ to 90% for $p_T > 30\,\mathrm{GeV}/c$. The lepton isolation efficiency depends on the lepton momentum, as well as on the jet activity in the event. In $t\bar{t}$ events, it varies approximately linearly from $\approx$ 73% (muons) and $\approx$ 82% (electrons) at $p_T = 10\,\mathrm{GeV}/c$ to $\approx$ 97% for $p_T > 60\,\mathrm{GeV}/c$. In LM1 (LM3) events, this efficiency is decreased by $\approx$5–10% ($\approx$10%,$\approx$5%)over the whole momentum spectrum. The average detector responses (the reconstructed quantity divided by the generated quantity) for H_T and E_T^{miss} are consistent with 1 within the 7.5% jet energy scale uncertainty. The experimental resolutions on these quantities are 9% and 12%, respectively.





Compute R and MR

Given the four momenta of the two hemispheres, MR is computed as:

```
double CalcMR(TLorentzVector ja, TLorentzVector jb){
  double A = ja.P();
  double B = jb.P();
  double az = ja.Pz();
  double bz = jb.Pz();
  TVector3 jaT, jbT;
  jaT.SetXYZ(ja.Px(),ja.Py(),0.0);
  jbT.SetXYZ(jb.Px(),jb.Py(),0.0);
  double ATBT = (jaT+jbT).Mag2();
  double temp = sqrt((A+B)*(A+B)-(az+bz)*(az+bz)-
           (jbT.Dot(jbT)-jaT.Dot(jaT))*(jbT.Dot(jbT)-jaT.Dot(jaT))/(jaT+jbT).Mag2());
  double mybeta = (jbT.Dot(jbT)-jaT.Dot(jaT))/sqrt(ATBT*((A+B)*(A+B)-(az+bz)*(az+bz)));
  double mygamma = 1./sqrt(1.-mybeta*mybeta);
  //gamma times MRstar
  temp *= mygamma;
  return temp;
}
```

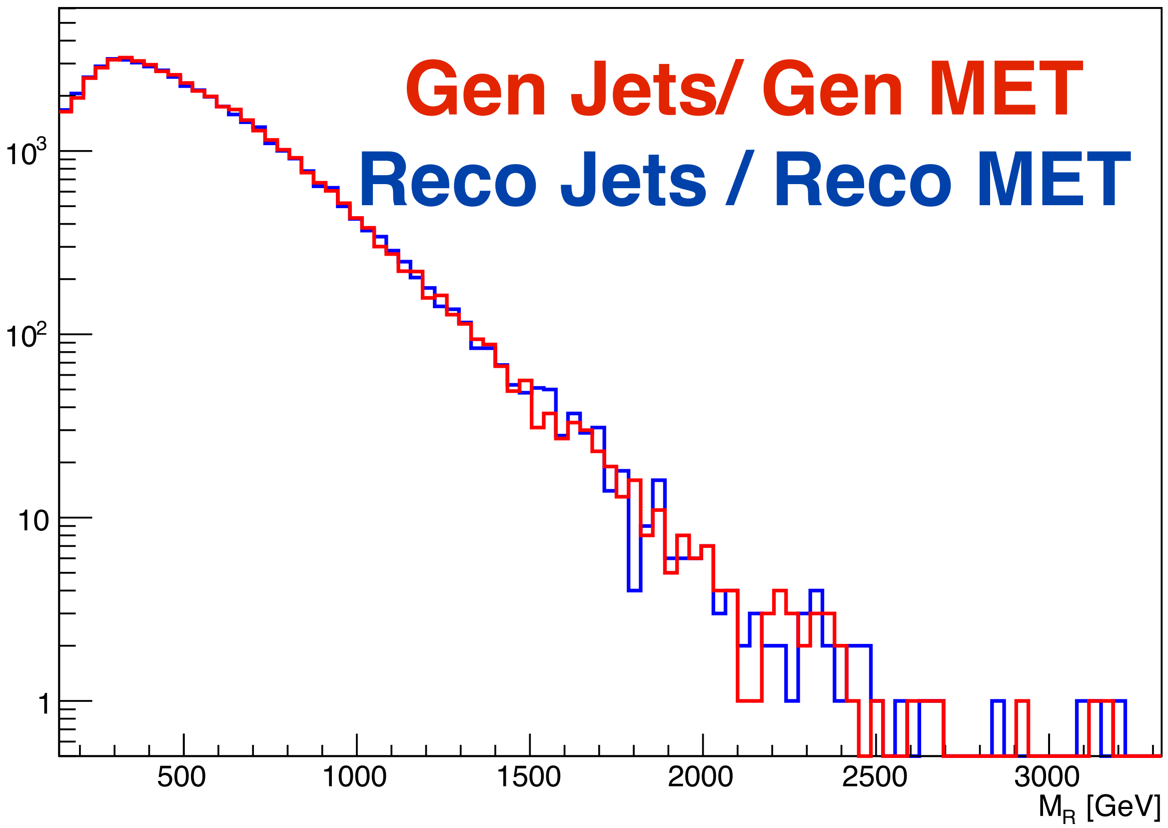
CMS SUS-12-005

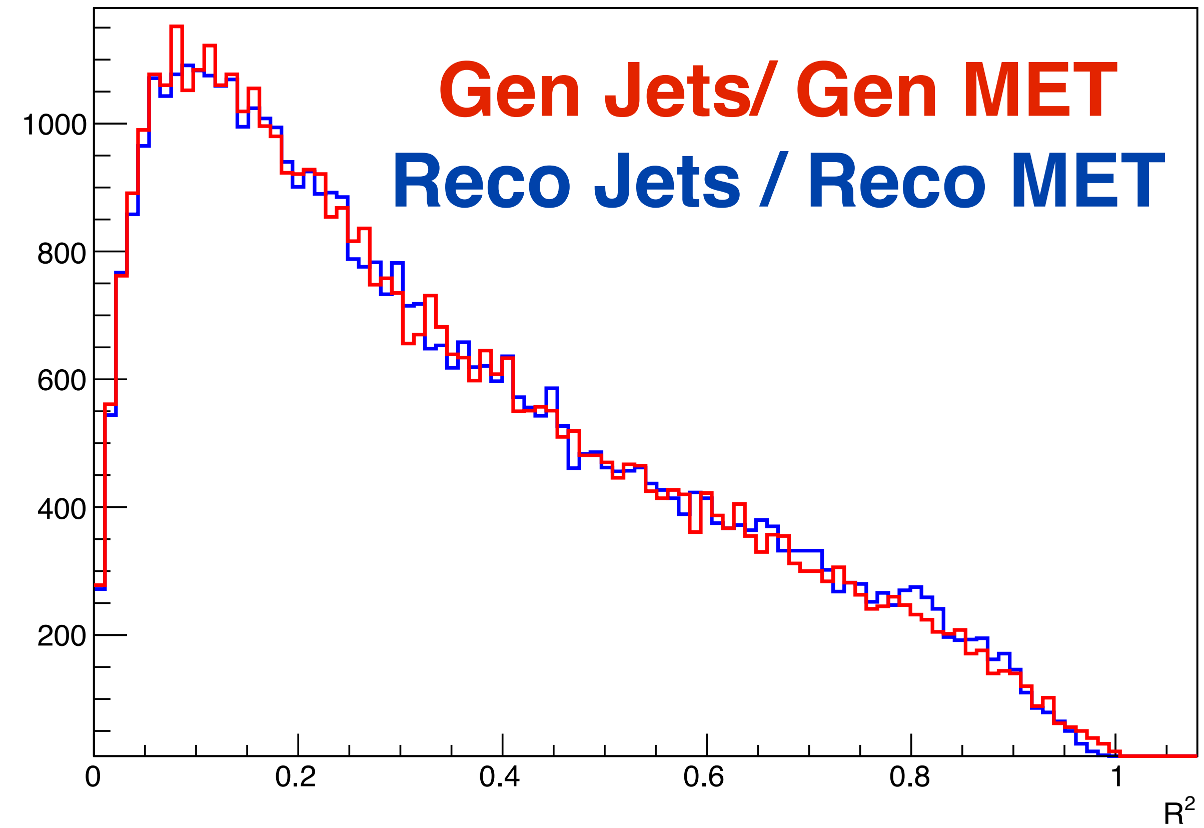
Code is sometimes published to help theorists to implement the kinematic variables used in the analysis

https://twiki.cern.ch/twiki/bin/view/CMSPublic/RazorLikelihoodHowTo

Is This Always Needed?

- Sometimes we make a big deal about detector effects when they are not so important to describe the signal





- We should focus on detector effects when they are needed (e.g. lepton efficiency) and understand that sometimes the resolution is an issue already at generator level (e.g. 2LSPs in the final state)

Possible Complications

- THE SHAPE ANALYSIS: Analyses are getting more and more complicated, because we want them to be more and more sensitive to a signal. The complication is the price to pay. The efficiency of an analysis is the end of the story if the analysis is a counting experiment. What about a shape analysis?

- NOT-STANDARD OBJECTS: Signatures may be weird (slow particles being detected out of time, non-pointing decays in the middle of the tracker giving noise-looking events, etc)

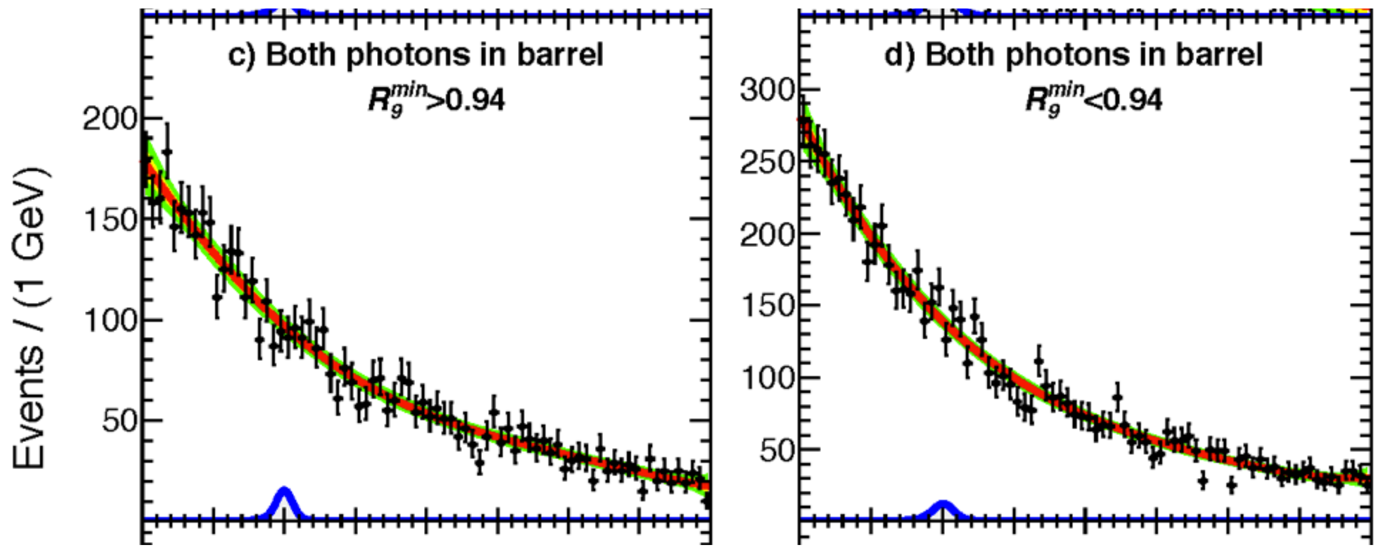
  The better you can do is to propose the signature, weird signatures are problematic even with our fast simulation

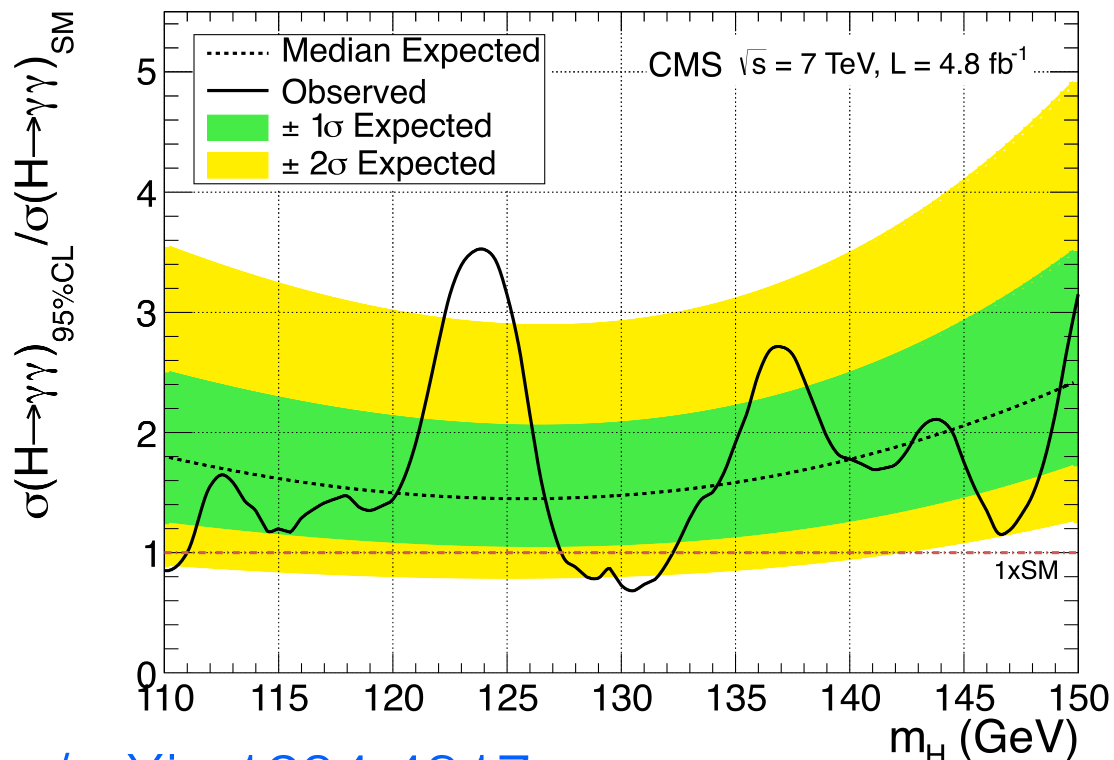
- MISSING INFORMATION: The needed information is not provided, and the interpretation of the result (e.g. The CMSSM plot) is given as the result

  Papers should be rejected by the journals: the main information (the experimental result) is missing!!!!

The shape Analysis

Even in the easiest case (Higgs search) when signal is specified and a list of (s,n,b) values could be provided, experiments failed to provide the needed information

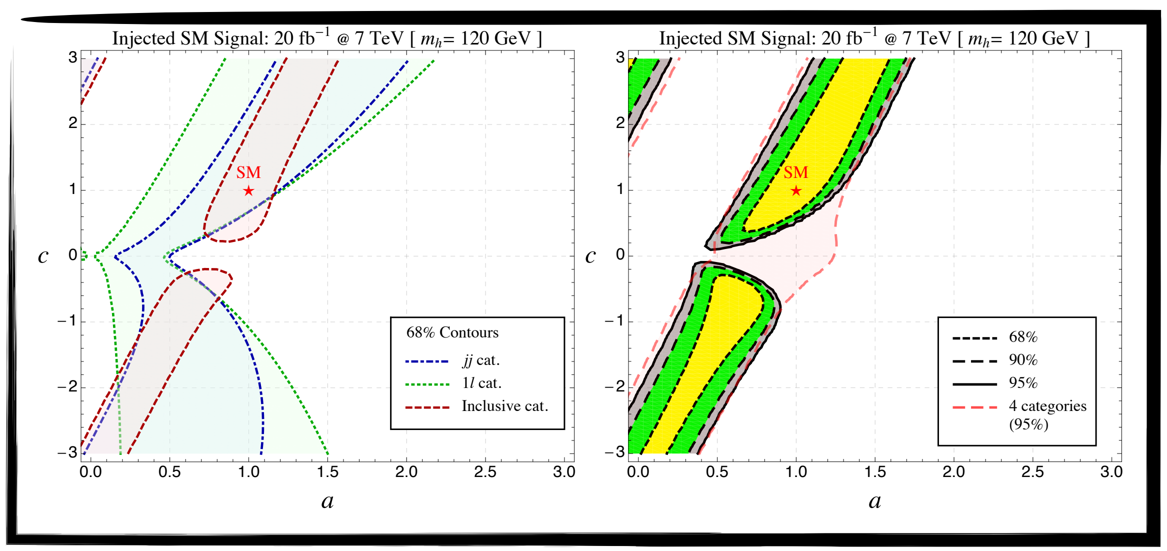




People had to do some gymnastic from the green/ yellow plots to "guess" the likelihood (under Gaussian assumption)

http://arXiv.org/abs/arXiv:1204.4817

http://arXiv.org/abs/arXiv:1203.4254



$$\sigma = \frac{R_{\text{expected}}}{1.96}, \quad \mu = R_{\text{observed}} - R_{\text{expected}},$$

The shape Analysis

If instead we had quoted bin by bin the observed yield and the expected signal and background (with errors) people could have reconstructed the likelihood

$$L = \prod P(n_i|s_i,\bar{b}_i)G(\bar{b}_i|b_i,\delta b_i)$$

i = bin number

n_i = observed yield

s_i = signal yield

$\bar{b}_i$ = background yield

b_i = expected background yield

δb_i = error on expected background yield

P = Poisson function

G = Gaussian function

this is usually a log-normal

Even for an unbinned fit, one can find a binning that approximate the likelihood (matching the bin size to a fraction of the resolution

Particularly for the Higgs, the information should be given by process (as in SUSY simplified model we quote the efficiency by process)

In SUSY, this might become impractical for signal (one efficiency SMS plot be bin might be too much information to digest) but it should be done for the background

The LH recommendation

http://arxiv.org/abs/1203.2489

i. e. we should make our papers clear!!!!

1. (a) Provide a clear, explicit description of the analysis in publications. In particular, the most crucial information such as basic object definitions and event selection should be clearly displayed in the publications, preferably in tabular form, and kinematic variables utilised should be unambiguously defined. Further information necessary to reproduce the analysis should be provided, as soon as it becomes available for release, on a suitable common platform.

   (b) The community should identify, develop and adopt a common platform to store analysis databases, collecting object definitions, cuts, and all other information, including well-encapsulated functions, necessary to reproduce or use the results of the analyses, and as required by other recommendations.

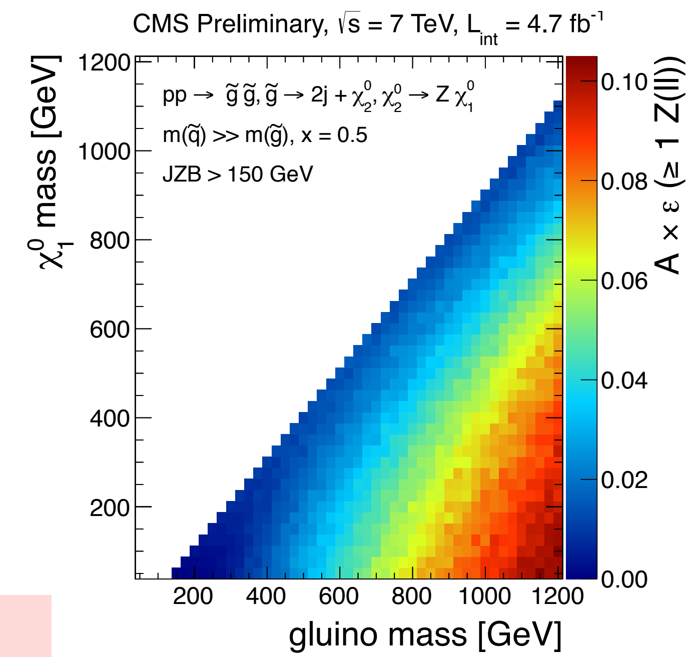
Rivet vs inspires vs HEPdata vs ...

(a) "crucial" recommendations, defined as actions that we believe should be undertaken immediately, and

(b), (c) "desirable steps", i.e. actions that would help, but whose implementation is recognized as requiring major efforts and a longer timescale.

The LH recommendation

http://arxiv.org/abs/1203.2489



2. (a) Provide histograms or functional forms of efficiency maps wherever possible in the auxiliary information, along with precise definitions of the efficiencies, and preferably provide them in standard electronic forms that can easily be interfaced with simulation or analysis software.

   (b) The community should take responsibility for providing, validating and maintaing a simplified simulation code for public use, reproducing the basic response of the LHC detectors. The validation and tuning of this tool should be based on comparisons with actual performance plots, and/or other inputs, made available by the experiments along the lines of Recommendation 2a. Limits of validity should be investigated and clearly documented.

Workshop at CERN LPCC on June 11-12
https://indico.cern.ch/conferenceDisplay.py?confId=187127

(a) "crucial" recommendations, defined as actions that we believe should be undertaken immediately, and

(b), (c) "desirable steps", i.e. actions that would help, but whose implementation is recognized as requiring major efforts and a longer timescale.

The LH recommendation

http://arxiv.org/abs/1203.2489

3. (a) Provide all crucial numbers regarding the results of the analysis, preferably in tabulated form in the publication itself. Further relevant information, like fit functions or distributions, should be provided as auxiliary material.
   Addendum:
   For multi-bin results, provide an ensemble of sets of the numbers B, δB, $\mathcal{L}$, $\delta\mathcal{L}$, Q, k, etc in the auxiliary information. These would be created by sampling from the various experiment-specific systematic effects, such as the jet energy scale, jet energy resolution, etc. Results should be quoted without inclusion of systematic/theoretical uncertainties external to the experiment.

   (b) When feasible, provide a mathematical description of the final likelihood function in which experimental data and parameters are clearly distinguished, either in the publication or the auxiliary information. Limits of validity should always be clearly specified.

   (c) Additionally provide a digitized implementation of the likelihood that is consistent with the mathematical description.

(a) "crucial" recommendations, defined as actions that we believe should be undertaken immediately, and

(b), (c) "desirable steps", i.e. actions that would help, but whose implementation is recognized as requiring major efforts and a longer timescale.

The LH recommendation

http://arxiv.org/abs/1203.2489

4. In the interpretation of experimental results, preferably provide the final likelihood function (following Recommendations 3b/3c). When this is not possible or desirable, provide a grid of confidence levels over the parameter space. The expected constraints should be given in addition to the observed ones, and whatever sensitivity measure is applied must be precisely defined. Modeling of the acceptance needs to be precisely described.

5. For Higgs searches, provide all relevant information on a channel-by-channel basis for both production and decay processes.

6. When relevant, design analyses and signal regions that are based on disjoint sets of events.

This is really complicated sometimes: my signal region is your control sample

(a) "crucial" recommendations, defined as actions that we believe should be undertaken immediately, and

(b), (c) "desirable steps", i.e. actions that would help, but whose implementation is recognized as requiring major efforts and a longer timescale.