Searching for SUSY with CMS

Sezen Sekmen, CERN

IPMLHC2013, 7-11 Sept 2013
IPM, Tehran

This talk is intended as a lecture that gives an overview of why and how we do SUSY searches at CMS along with a few examples of current searches. A more thorough description of the current searches will be given in 5 other talks.

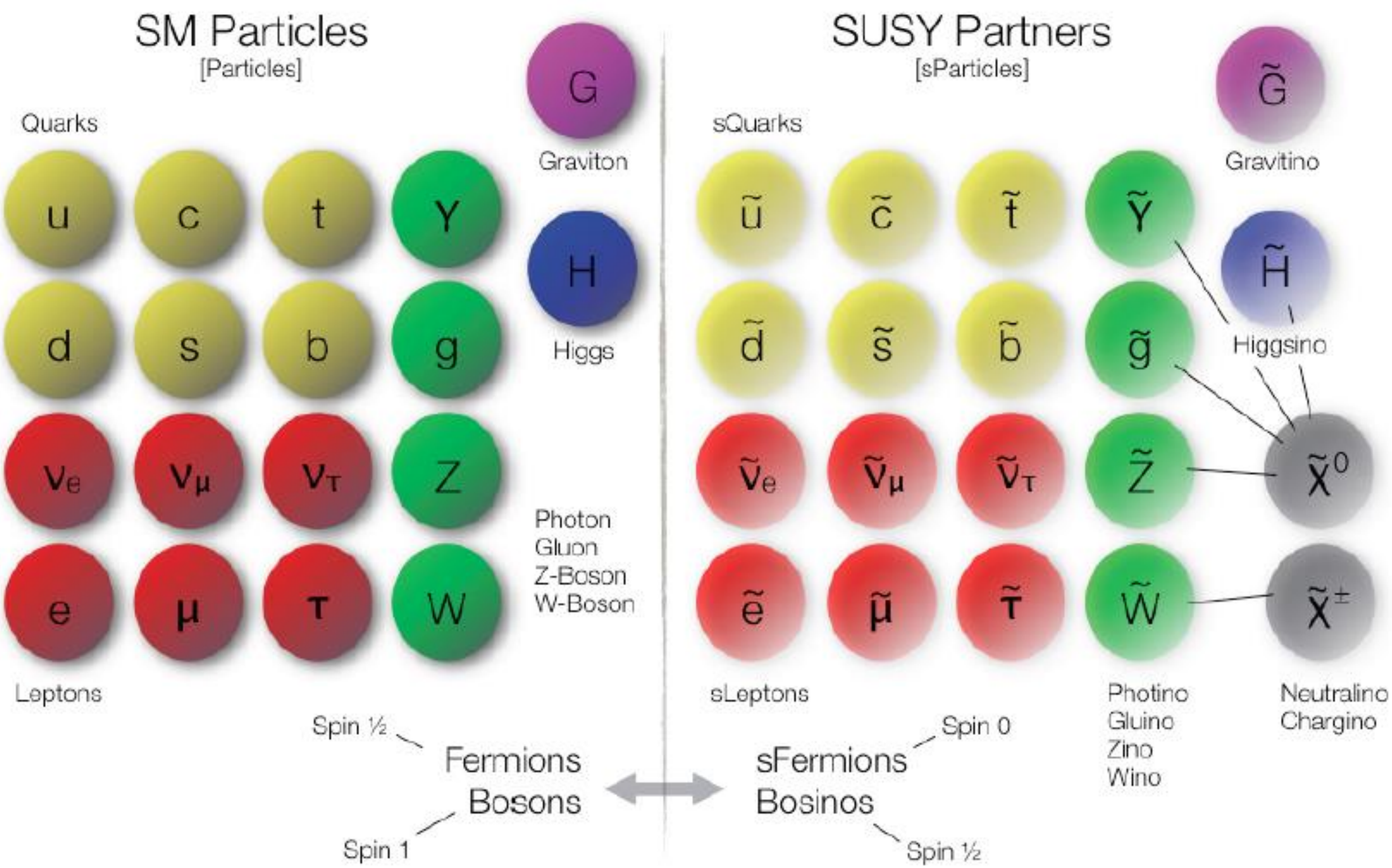
Outline:

- **Supersymmetry and LHC:** definition, motivation, production, models and final states
- **Elements of a SUSY search:** signal characterization, trigger, objects, background estimation, statistical analysis, systematics
- **SUSY searches at CMS:** a few inclusive searches and summary of current interpretation

SUSY and the LHC

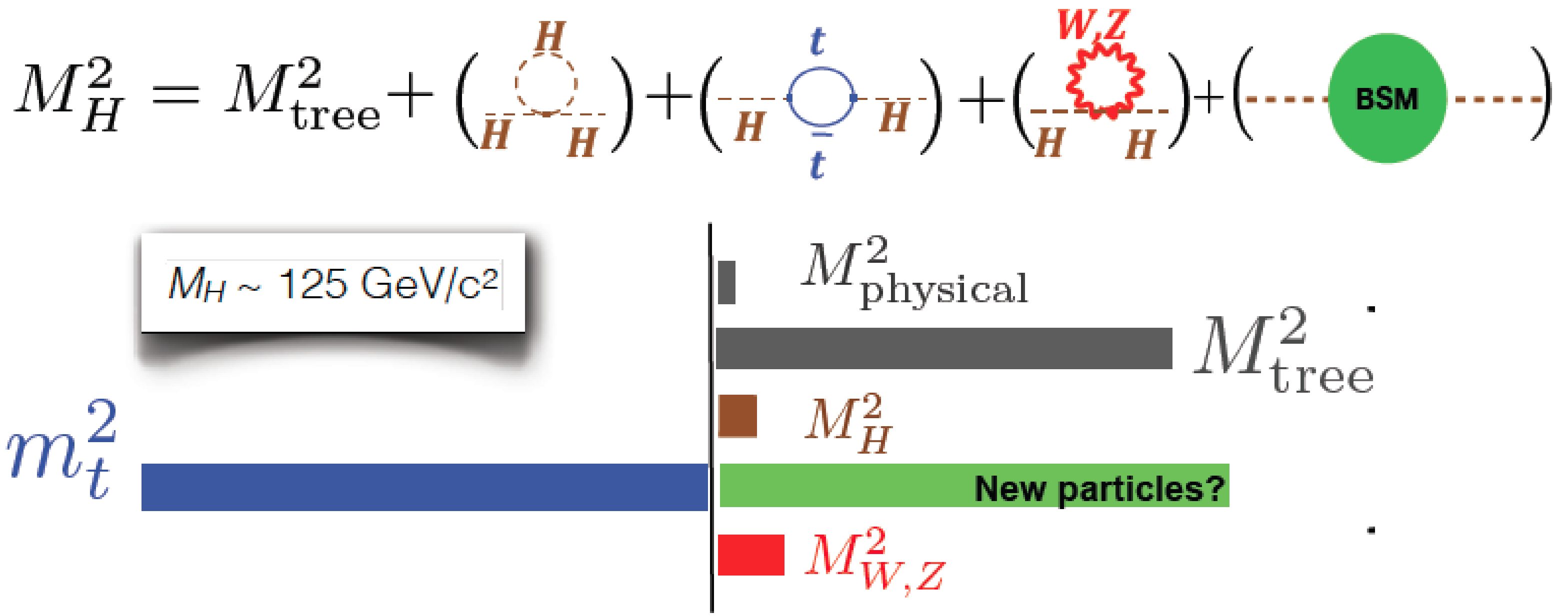
What is SUperSYmeetry?

Supersymmetry (SUSY) is a symmetry between fermions and bosons, between matter and force. It predicts the existence of new particles. For every SM particle, there is a superpartner with ½ spin difference.



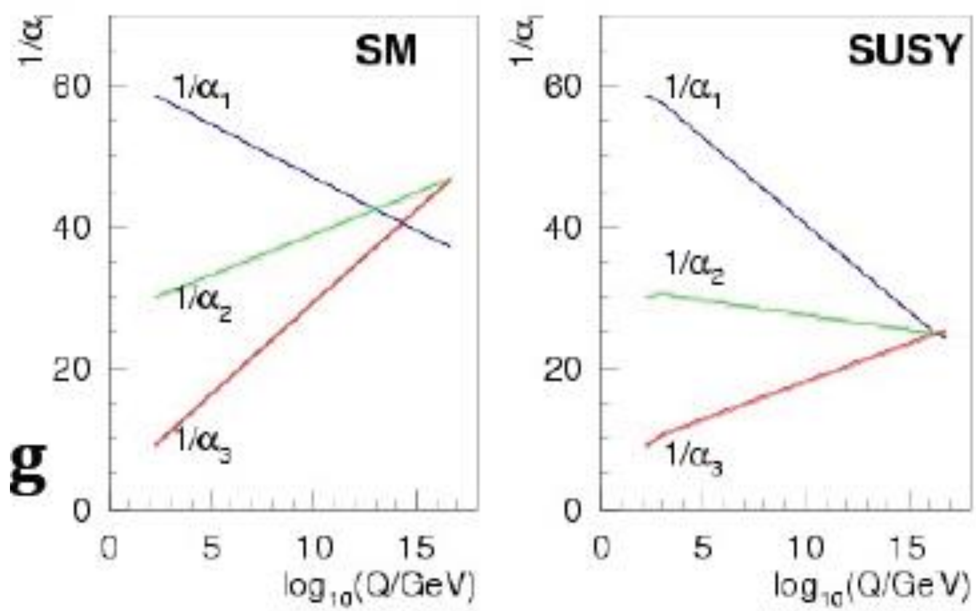
Why supersymmetry? - I

- Standard Model is an effective theory. We would like to understand physics in a more generic framework which completes the missing pieces.
- SM does not incorporate gravity.
- Fine tuning in the corrections to the Higgs mass can be resolved by adding new particles with opposite spin. SUSY contributions to Higgs mass cancel SM contributions.

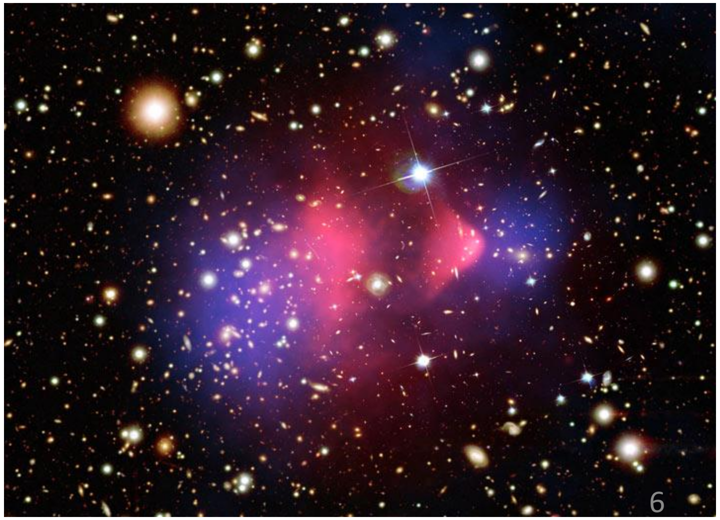


Why supersymmetry? - II

- SUSY unifies gauge couplings at the GUT scale, because contributions from new particles modify running of the gauge couplings.

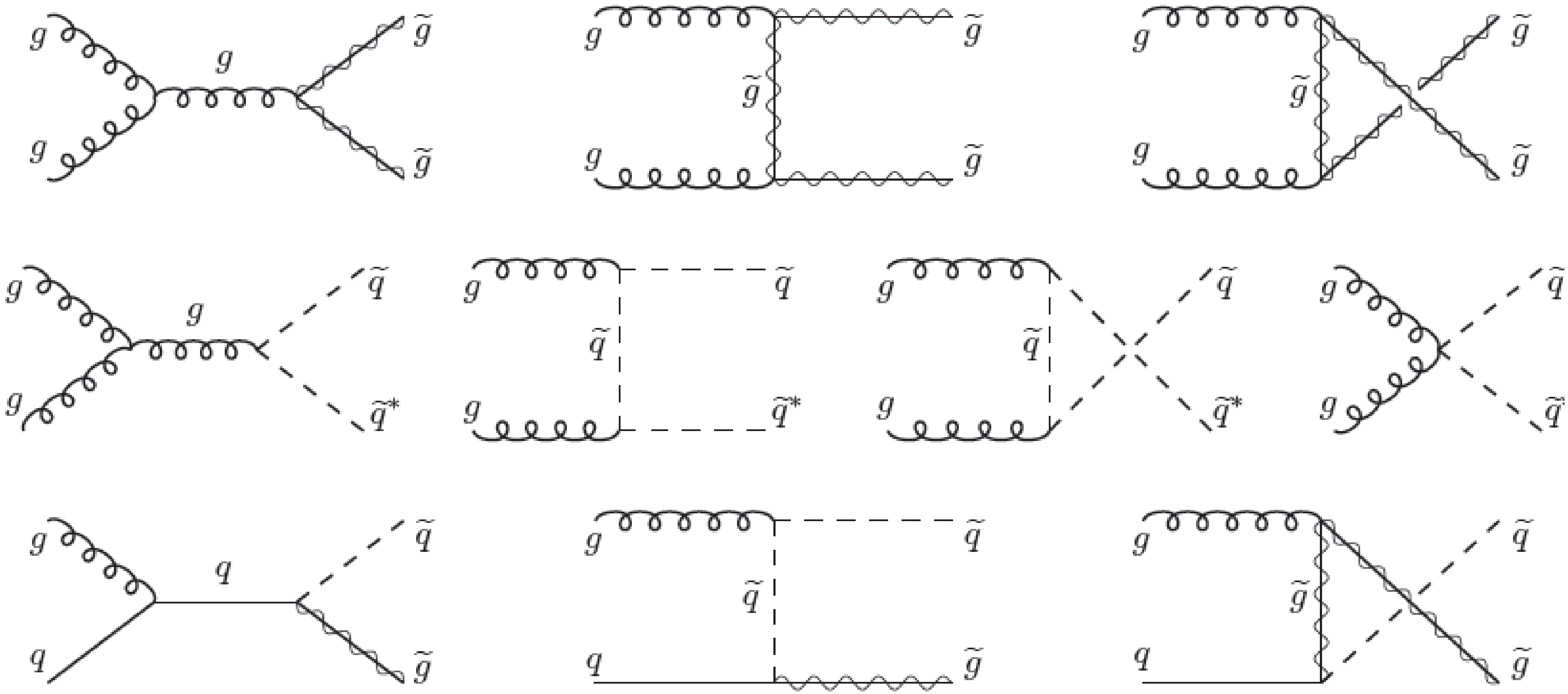


- SUSY offers a dark matter candidate. Lightest supersymmetric particle (LSP) can be heavy, neutral and stable.



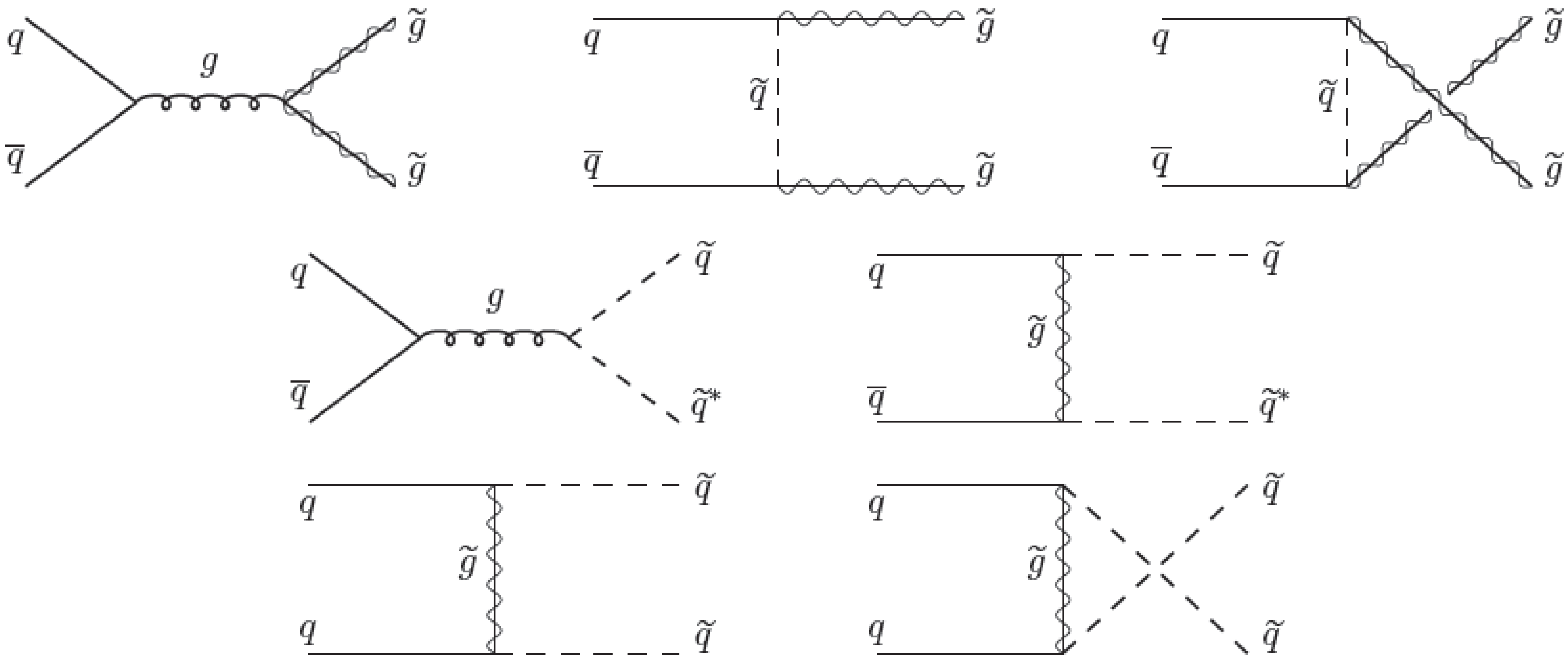
Sparticle production @ LHC - I

Gluino and squark production via gluon-gluon and quark-gluon processes.
Gluon-gluon processes are dominant at LHC energies.
Squark and gluino production is dominant for light gluinos/squarks (<TeV).



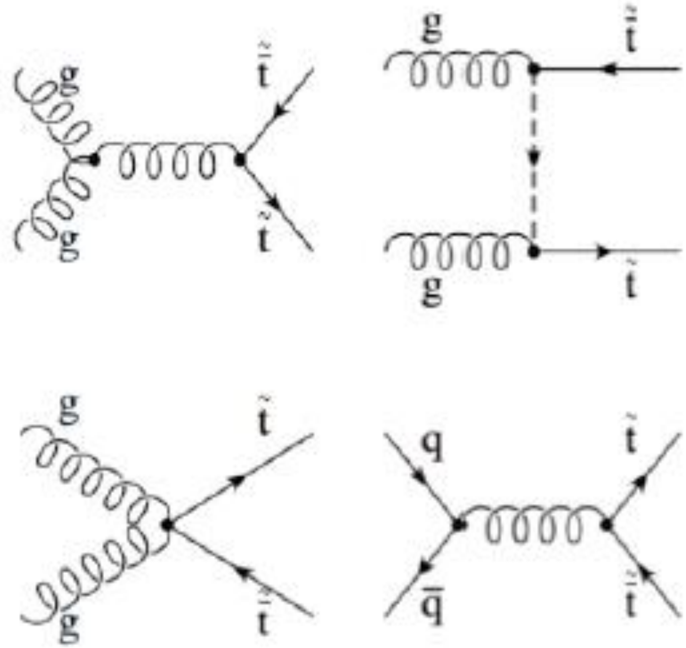
Sparticle production @ LHC - II

Gluino and squark production via quark-antiquark annihilation and quark-quark scattering at the LHC.

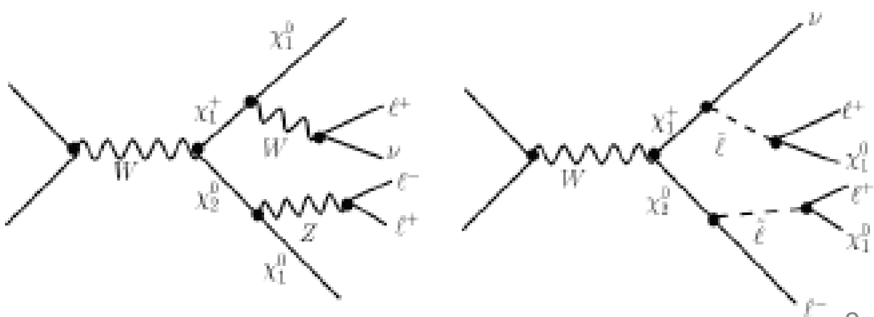


Sparticle production @ LHC - III

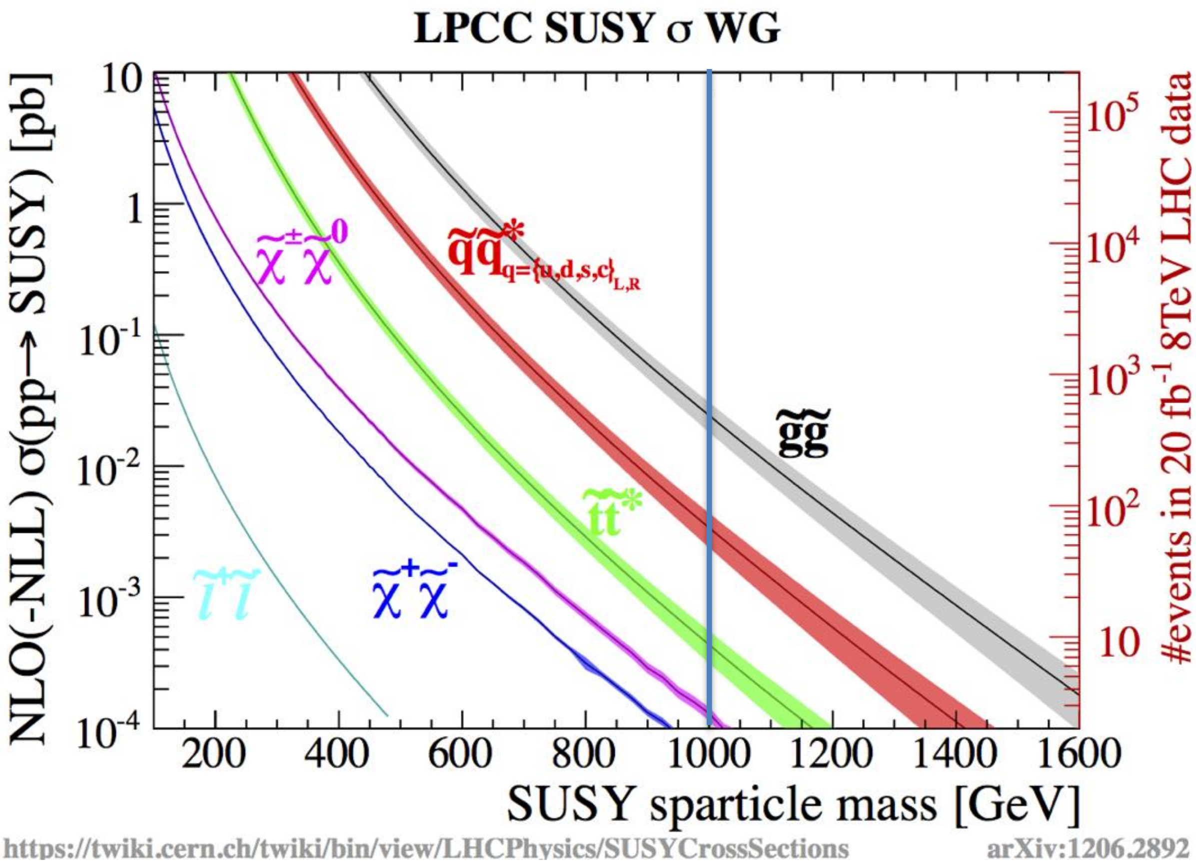
Stop-stop production at the LHC. Stop1-stop1 or stop2-stop2 production dominate over stop1-stop2 processes since gluon-gluon processes are dominant.



Example chargino-neutralino production at LHC from quarks.



Sparticle production cross sections @ LHC

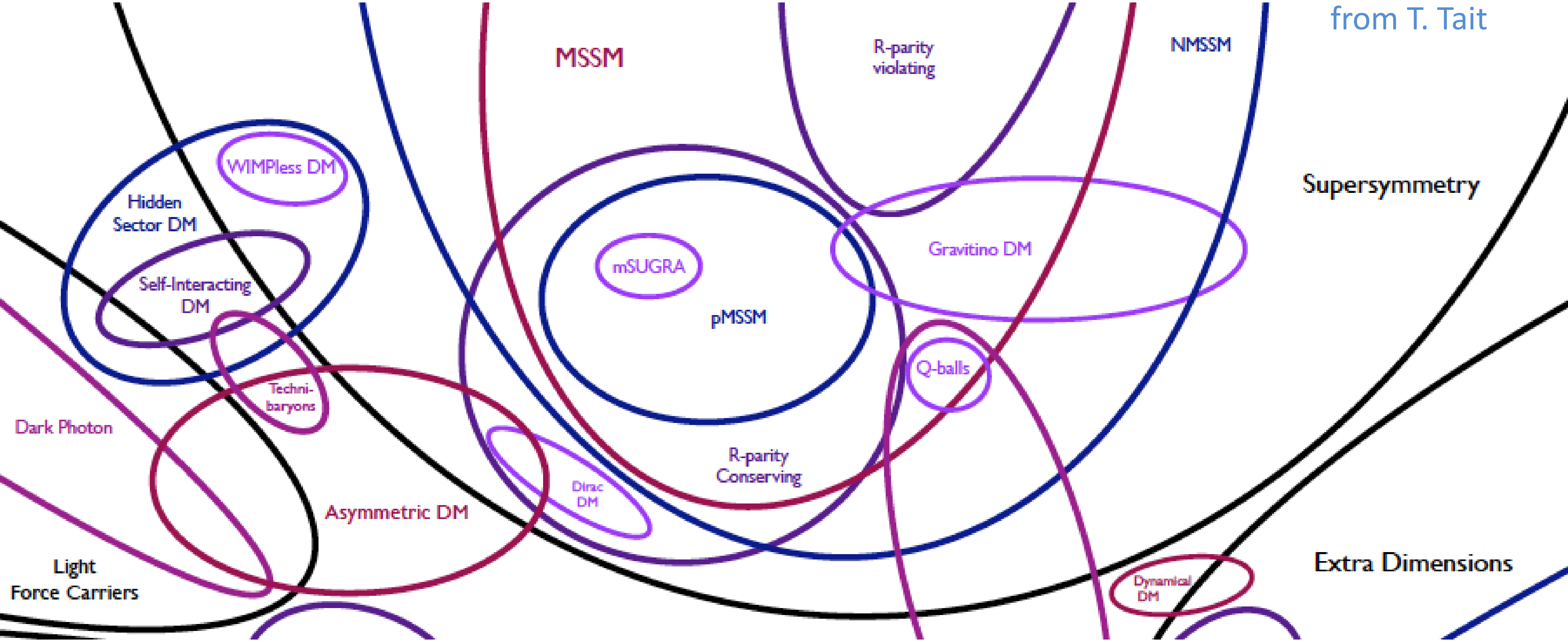


- Gluinos, 1st and 2nd generation squarks (when they are degenerate) – high cross sections.
- 3rd generation squarks (stops, sbottoms) – moderate cross sections.
- Charginos, neutralinos, sleptons – small cross sections, but feasible.

Which supersymmetry?

Supersymmetry is a wide framework with diverse realizations → diverse final states at the LHC.
In its most generic form, it is defined by >100 free parameters.

from T. Tait



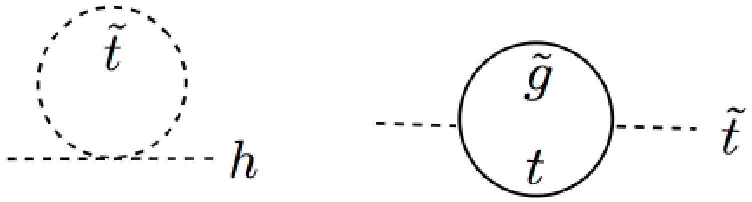
Sparticles are heavier than particles. SUSY is a broken symmetry.
We don’t know the nature of SUSY breaking yet – but there are many models.
SUSY breaking models define SUSY phenomenology.

Which supersymmetry?

SUSY proposes diverse realizations. We need to search every direction.
But it is nice to have some well-motivated principles to guide our searches.

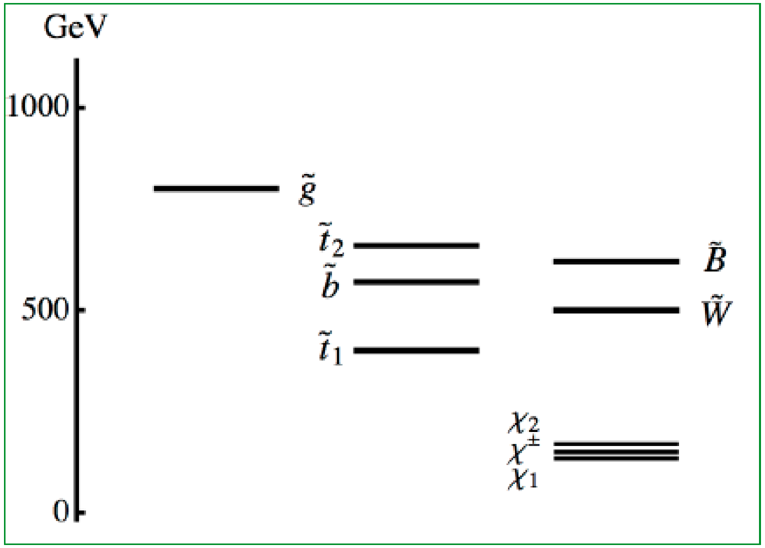
Natural SUSY?

Hierarchy problem: Higgs mass is 125GeV despite the divergent corrections from the top loop. The divergencies can be cancelled by introducing SUSY particles – but this imposes requirements to the SUSY mass spectrum.



- Leading contribution to the Higgs mass comes from Higgsinos → ≤ few hundred GeV
- Stops contribute to Higgs mass via 1-loop corrections → ≤ few hundred GeV
- Sbottom left is tied to stop left → ≤ few hundred GeV.
- Gluinos contribute to Higgs mass via 2-loop corrections → ≤ few TeV

- Rest of the spectrum can be decoupled / heavy.



R.Barbieri & D.Pappadopulo JHEP 0910:061,2009

A generic framework: pMSSM

Given that we are not convinced by a theorietical motivation, we can consider a more generic framework.

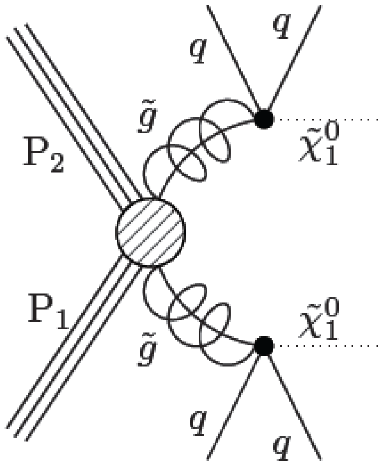
- p(henomenological)MSSM is a 19-dimensional parameterization of MSSM at the SUSY scale.
- pMSSM is defined by
  - 3 gaugino mass parameters
  - 10 sfermion mass parameters
  - 3 trilinear couplings
  - ratio of Higgs VEVs $\tan\beta$, Higgsino mass parameter μ and pseudoscalar Higgs mass m_A
  - plus a set of minimal assumptions.
- It is a full model with no assumptions on the nature of SUSY breaking mechanism and no correlations between the sparticle masses. It allows to make generic statements on sparticle masses.

Simplified models

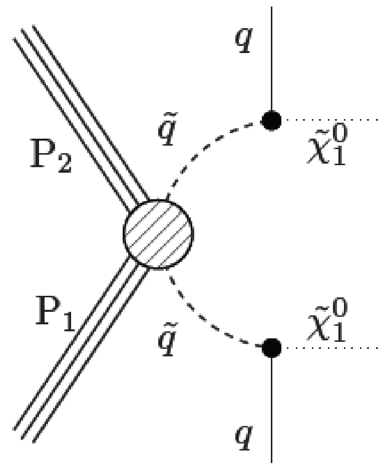
Or maybe we would like to simply have a way of modeling SUSY-like final states one-by-one in terms of an effective framework?

- A simplified model is defined by a set of hypothetical particles and a sequence of their production and decays.
- For each simplified model, values for the product of the experimental acceptance and efficiency (A X e) are calculated to translate a number of signal events into a signal cross section.
- From this information, a 95% confidence level upper limit on the product of the cross section and branching fraction is derived as a function of the particle mass.
- Only the production process of two particles is considered.
- Each particle decays directly or via a cascade to particles X + a neutral, undetected particle (i.e., the LSP.)

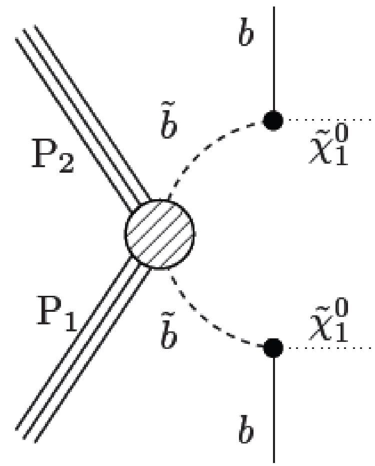
Simplified models



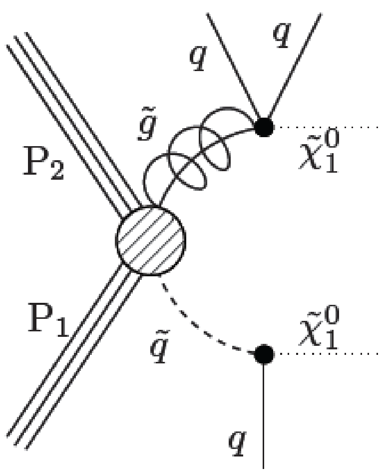
(a) T1



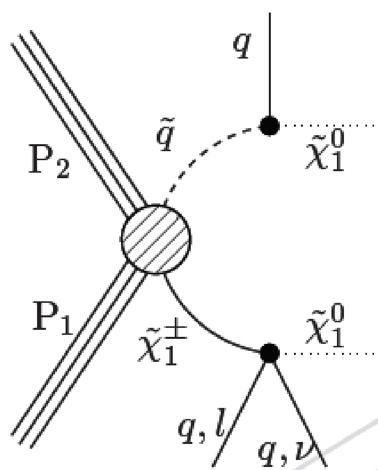
(b) T2



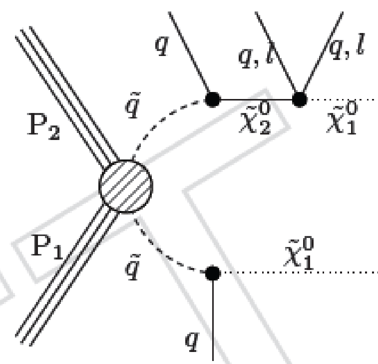
(c) T2bb



(d) TGQ

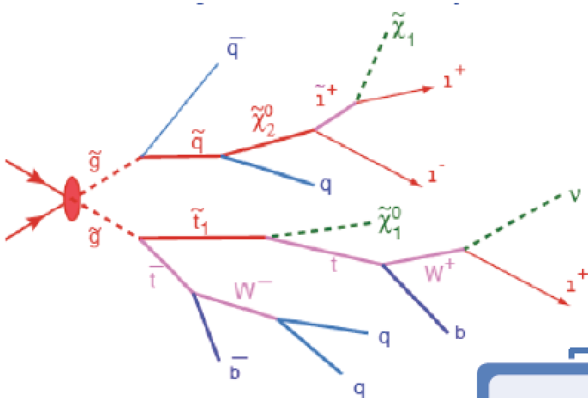


(e) TNS



(f) T4N

What are we searching for?



- Classifying searches for SUSY
  - Prompt?
    - RPC?
      - Strong?
        - 1st and 2nd generation squarks and gluinos
        - 3rd generation squarks (stops, sbottoms)
      - Weak?
        - EWK gauginos, sleptons
    - RPV?
  - Long-Lived?
    - How long lived?

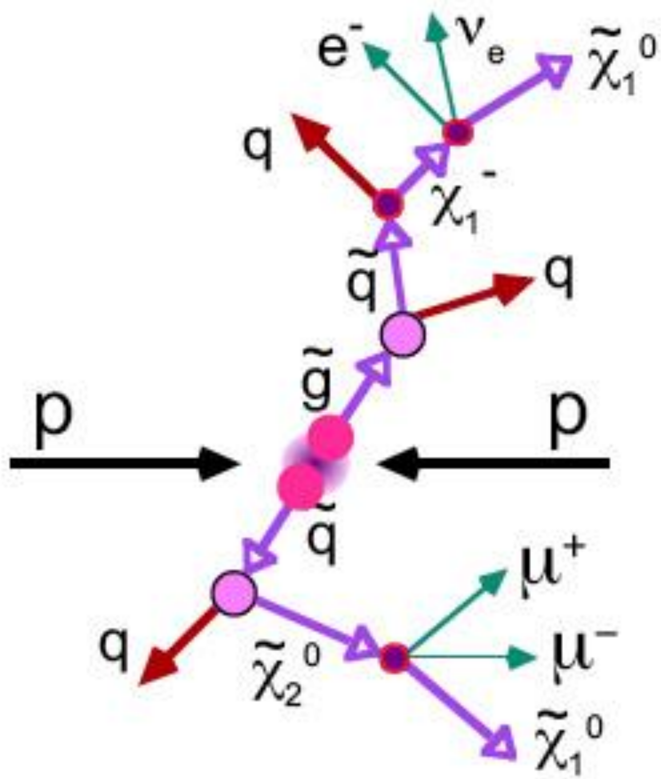
from E. Halkiadakis

Elements of a SUSY search

Elements of SUSY (new physics) search

- Signal characterization and search strategy
- Designing the triggers
- Object reconstruction and identification
- Signal characterization and event selection
- Background estimation
- Statistical analysis
- Systematic uncertainties
- Results
- Interpretation

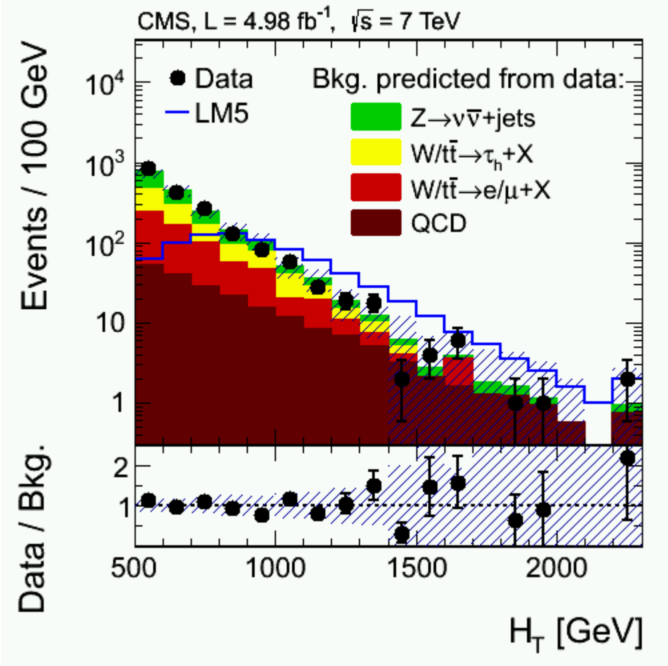
Characterizing the signal



- SUSY can appear in diverse final states, but we can still classify and investigate some characteristic SUSY topologies.
- Some classical topologies with missing ET:
  - Dijets
  - Multijets
  - 1 lepton + jets
  - 2 leptons (same sign/opposite sign) + jets
  - Multileptons + jets
  - photons + jets
  - 3rd generation (tau/b/top) versions of the above

- We use some variables to characterize the SUSY signals and to distinguish them from the SM backgrounds: HT, MT2, alphaT, razor, endpoints, etc.

Characterizing the signal – kinematic variables
Global variables: HT and MHT



Hadronic transverse energy is the scalar sum of the momenta of all jets in the event.

$$H_T = \sum_{i}^{n\,jets} p_T^{jet_i}$$

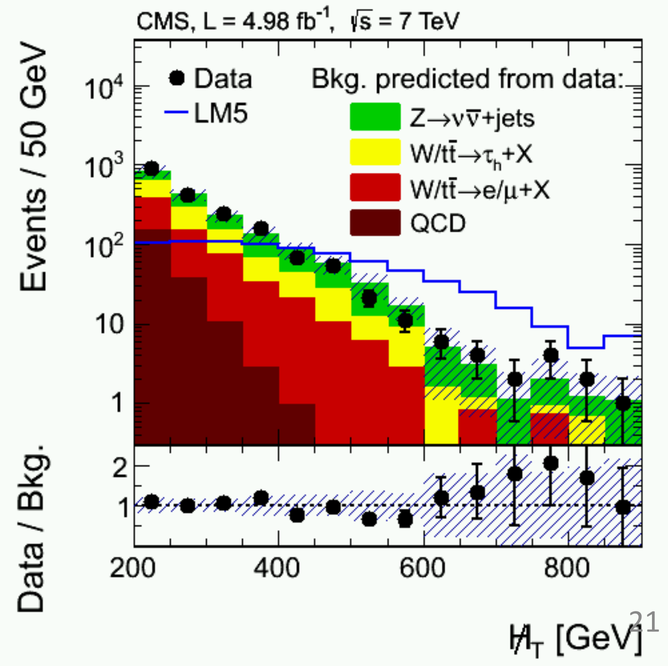
Conventional SUSY events are supposed to have high hadronic transverse activity and high HT.

CMS-SUS-12-011
PRL 109, 171803 (2012)

Missing hadronic transverse momentum is the negative vectorial sum of momenta of all jets in the event:

$$\not{H}_T = H_T^{miss} = -\sum_{i}^{n\,jets} \vec{p}_T^{\,jet_i}$$

Conventional SUSY events have energetic missing particles, and hence high HTmiss.



Characterizing the signal – kinematic variables
Transverse mass

Sometimes, we'd like to get information on the heavy particles produced. When all decay products are visible, we can reconstruct its invariant mass.

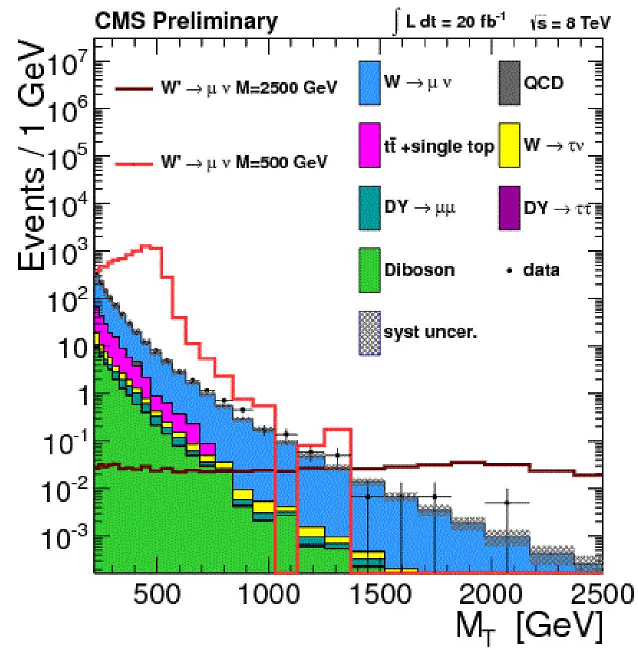
BUT, sometimes some decay products are invisible, and we don't have access to full 4-momenta of the final state particles.

For example, in W → lν decays, invisible neutrinos escape the detector. If there is only one ν in the event, we can approximate ν transverse momentum p_T^{ν} by the MET. We define the transverse mass for W as:

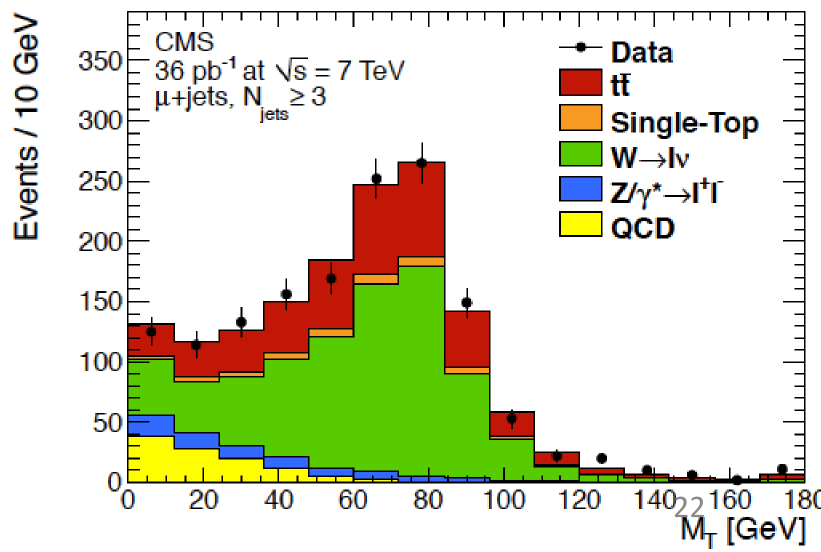
$$m_{T,W}^2 = m_\ell^2 + m_\nu^2 + 2(p_T^\ell p_T^\nu - \vec{p}_T^{\,\ell} \vec{p}_T^{\,\nu})$$

$$(m_\ell, m_\nu \sim 0 \to) \simeq 2p_T^\ell p_T^\nu (1 - \cos\Delta\phi(\ell,\nu))$$

where $m_{T,W}^{max}$ gives m_W because $m_{T,W} < m_W$.



Top plot: W M_T used in new physics searches. M_T distribution fc hypothetical W' particles where W' → eν.



Bottom plot: W M_T is used extensively in top searches and searches for new physics with top-like particles as a discriminating variable in the event selection (Right: from ttbar cross section measurement in leptons+jets channel).

Characterizing the signal – kinematic variables
"s"transverse mass

BUT…what if we have more than one invisible particles in the final state? Take the typical case

$$pp \to \tilde{q}_1\tilde{q}_2 \to j_1\tilde{\chi}_1 j_2\tilde{\chi}_2$$

where ~χs are invisible. Two invisible particles make up the MET. The stransverse mass

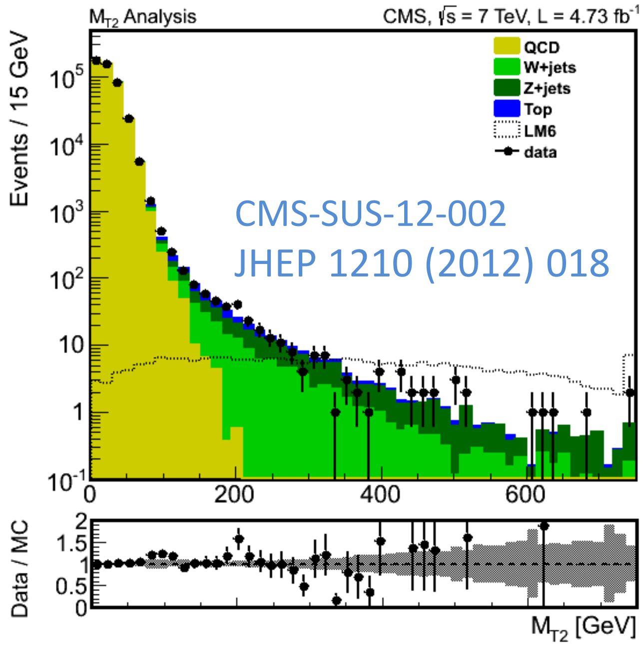
$$m_{T2}(m_{\tilde{\chi}}) = \min_{\vec{p}_T^{\tilde{\chi}1}+\vec{p}_T^{\tilde{\chi}2}=\vec{p}_T^{miss}} \left[\max\left(m_T(\vec{p}_T^{j1},\vec{p}_T^{\tilde{\chi}1}), m_T(\vec{p}_T^{j2},\vec{p}_T^{\tilde{\chi}2})\right)\right] \leq m_{\tilde{q}}^2$$

suggests a way to decompose the MET into these particles.

more on E. Eskandari's talk

The minimization is over all possible partitions of the measured MET.

However, for massive ~χ, we need the ~χ mass for calculating m_{T2}. It is shown that for different input $m_{\sim\chi}$ values, endpoint of the corresponding m_{T2} distributions makes a kink at the correct $m_{\sim\chi}$ value.



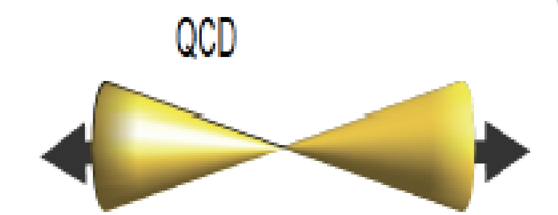
CMS-SUS-12-002
JHEP 1210 (2012) 018

MT2 is used as a selection variable in SUSY searches in ATLAS and CMS

Characterizing the signal – kinematic variables alphaT

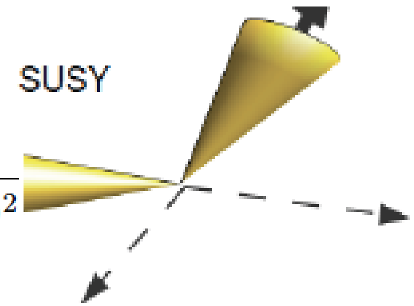
Can we distinguish events with genuine MET from events with misreconstructed MET?

We quantify the unbalance caused by the nature of MET in dijet events with the alphaT variable



$$\alpha_T = E_T^{j_2}/M_T$$

$$M_T = \sqrt{\left(\sum_{i=1}^{n} E_T^{j_i}\right)^2 - \left(\sum_{i=1}^{n} p_x^{j_i}\right)^2 - \left(\sum_{i=1}^{n} p_y^{j_i}\right)^2 - \left(\sum_{i=1}^{n} p_z^{j_i}\right)^2} = \sqrt{H_T^2 - (H_T^{miss})^2}$$

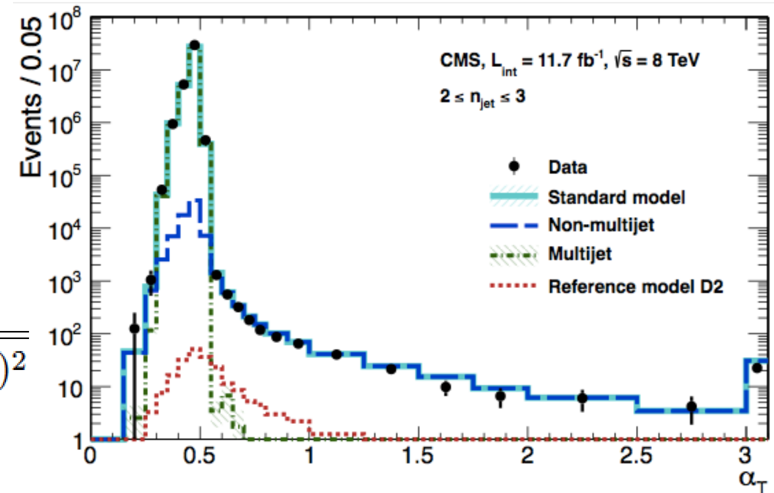


If njets > 2, we reconstruct two pseudojets by combining all jets in the event.

alphaT can also be generalized as

$$\alpha_T = \frac{1}{2}\frac{H_T - \Delta H_T}{M_T} = \frac{1}{2}\frac{H_T - \Delta H_T}{\sqrt{H_T^2 - (H_T^{miss})^2}} = \frac{1}{2}\frac{1 - \Delta H_T/H_T}{\sqrt{1 - (H_T^{miss}/H_T)^2}}$$

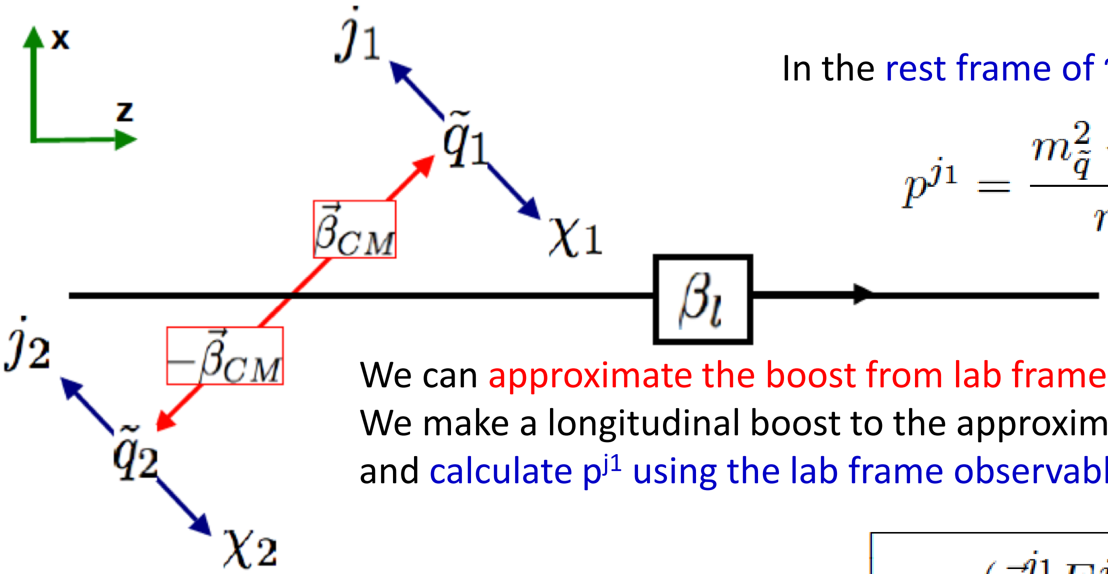
$$\Delta H_T = p_T^{J1} - p_T^{J2}$$



CMS-SUS-13-028
EPJC 73 (2013) 256

Characterizing the signal – kinematic variables
Mighty razor variables - I

C. Rogan, arXiv:1006.2727



In the rest frame of ~q, we have

$$p^{j1} = \frac{m_{\tilde{q}}^2 - m_{\tilde{\chi}}^2}{m_{\tilde{q}}^2} = \frac{m_\Delta}{2}$$

We can approximate the boost from lab frame to q rest frame to be a longitudinal boost. We make a longitudinal boost to the approximate q rest frame (we call it "the R frame") and calculate p^{j1} using the lab frame observables as:

$$M_R = 2p^{j_1(R\,frame)} = \sqrt{\frac{(\vec{p}_z^{j1} E^{j_2} - \vec{p}_z^{j2} E^{j_1})^2}{(\vec{p}_z^{j1} - \vec{p}_z^{j2})^2 - (E^{j_1} - E^{j_2})^2}} = m_\Delta$$

And there is a second (MT2-like) way to approximate the m_Δ distribution using the transverse components of the lab frame objects, whose kinematic endpoint gives m_Δ:

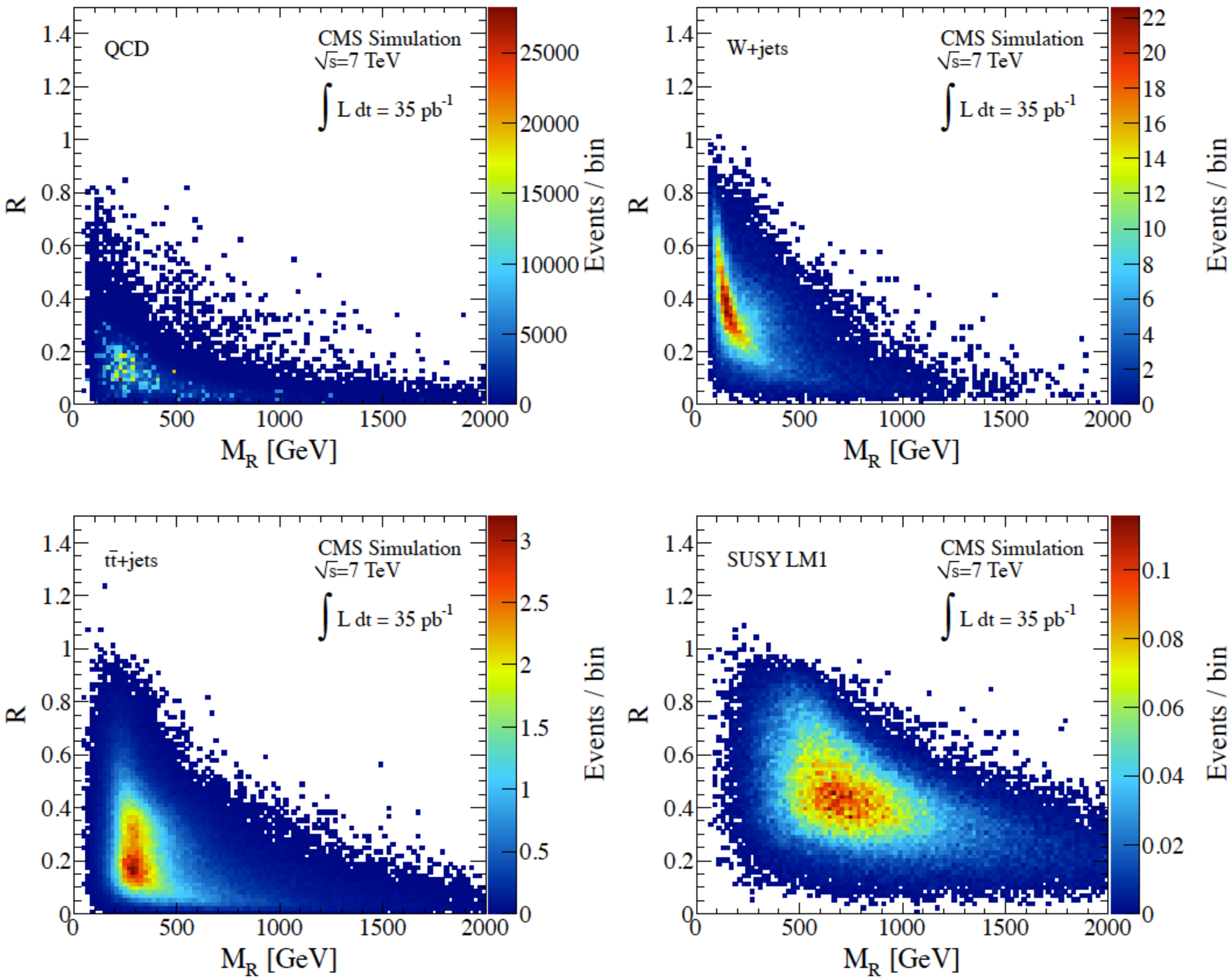
$$M_T^R = \sqrt{\frac{E_T^{miss}}{2}(p_T^{j1} + p_T^{j2}) - \frac{1}{2}\vec{E}_T^{miss} \cdot (\vec{p}_T^{j1} + \vec{p}_T^{j2})} < m_\Delta$$

Then, the ratio

$$R \equiv M_T^R / M_R$$

is a dimensionless quantity that combines two different ways of measuring the same thing.

Characterizing the signal – kinematic variables
Mighty razor variables - II



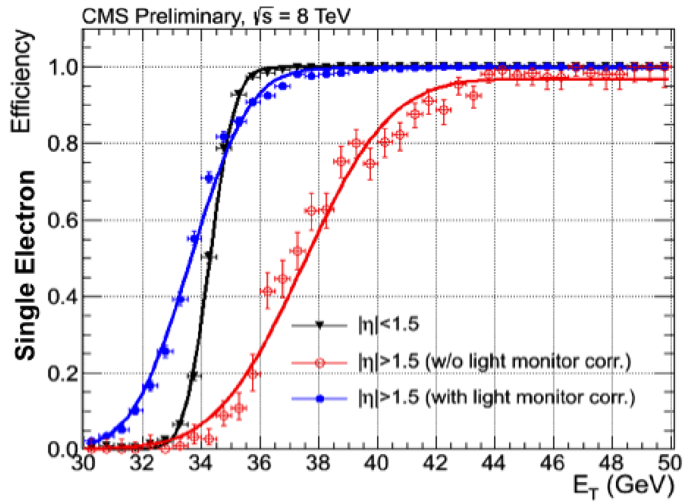
Most kinematic discriminators give an excess in the tails (e.g. MET), but razor variables define a “bump”, hence they provide very good signal-BG discrimination.

more on S. Paktinat’s talk

CMS and ATLAS use razor extensively for new physics searches.

Trigger

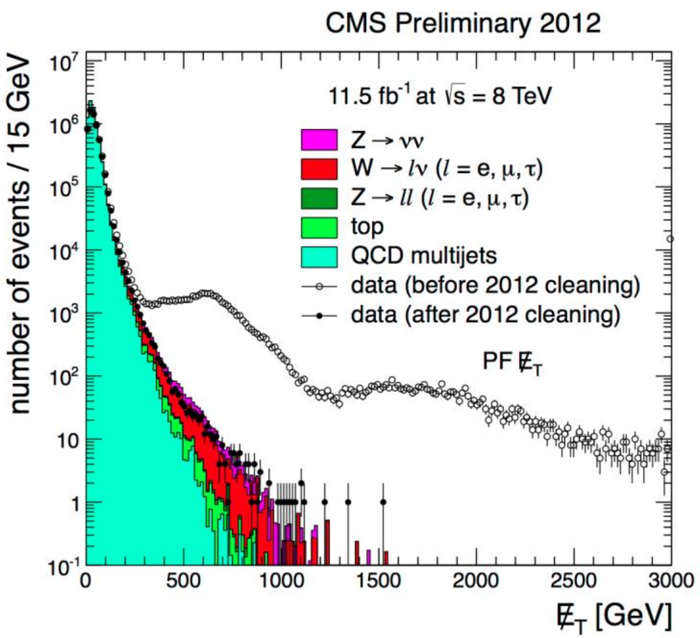
- Triggers are fast online filters that select the most interesting events during data taking, and store them for the offline analysis – if not stored, events are lost forever!
- Trigger is a rough sketch of the offline analysis: we select events with final states representative of the physics we're looking for.
  - Trigger on object kinematics and multiplicities
  - Trigger on kinematic variables: HT, HTmiss, alphaT, razor
- Two important trigger tasks for new physics searches:
  - Design the triggers that would cover the target signals.
  - Triggers are not fully efficient with respect to the offline cuts. Estimate the trigger efficiency given by:

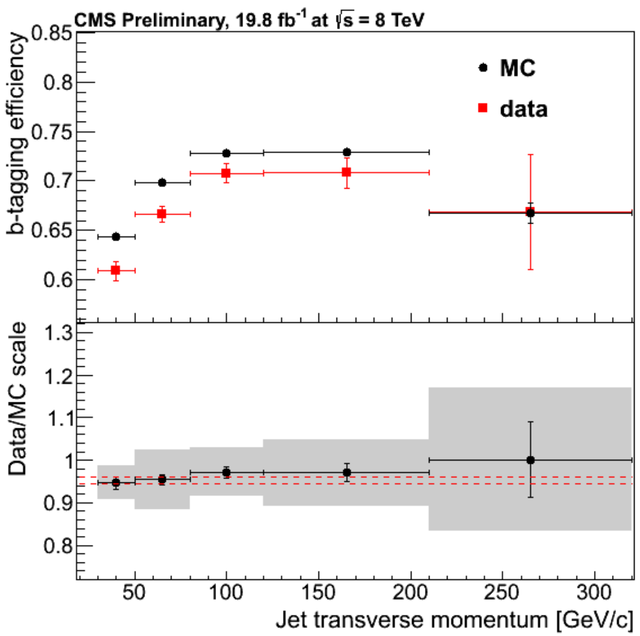


$$\epsilon_{trigger} = \frac{\text{number of events that pass the trigger}}{\text{number of total events}}$$

Objects

- Information from subdetectors is combined to reconstruct objects (jets, electrons, muons, taus, photons, missing transverse energy, b-tagged jets, boosted objects (Ws and tops)).
  - CMS uses particle flow (PF) which combines information from all subdetectors to reconstruct particles.
- Objects are then required to pass some identification and isolation criteria. We must find the optimum criteria that reflect our final state best.



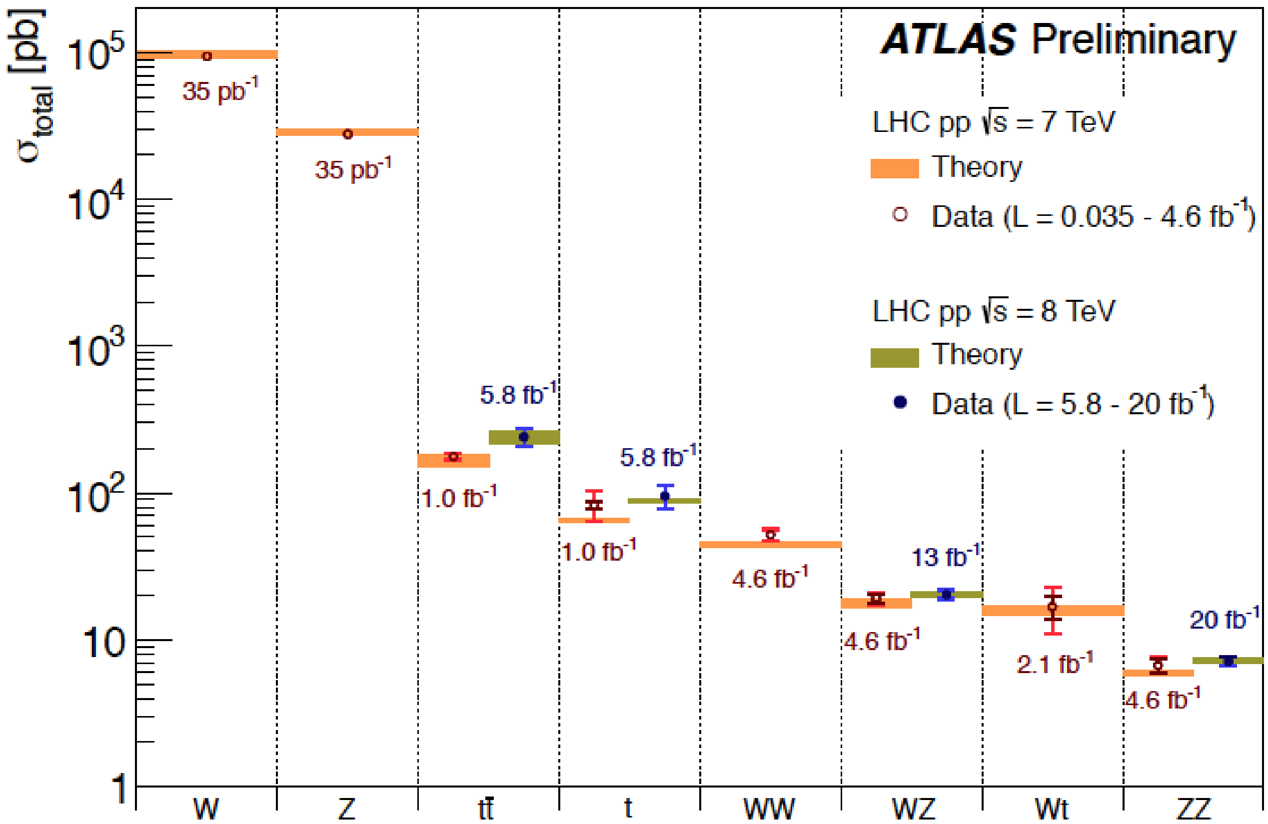


Event selection: principles

- Characterize the signal. Find final state topologies and kinematic variables that discriminate the signal from the backgrounds. Multijets? Opposite-sign dileptons? b-rich? Discriminating kinematic properties?
- Look for statistically significant signal regions. There should be sufficient number of events, and sufficient number of predicted signal events over the expected background.
- Make sure that there is a way to estimate the expected background in the signal region.
- Make sure that the offline selection corresponds to a region where the trigger efficiency is well-modeled.
- Make sure that the selection variables are reconstructed and identified in well-defined regions of the detector (not feasible to design a search with forward electrons).
- Numerous multivariate methods exist for selection optimization: rectangular cuts, fisher discrimination, likelihoods, neural networks, decision trees, support vector machines, …

BAAACKGROOOOOUNDSSS!!!

SM backgrounds in SUSY searches



https://twiki.cern.ch/twiki/bin/view/AtlasPublic/CombinedSummaryPlots

Inclusive SM cross sections are measured with precision and over many orders of magnitude.

Precise measurements of kinematical distributions are crucial for SUSY searches.

Greatest background sources are QCD (not shown), Ws, Zs and ttbar.

Background estimation

We need to estimate the amount (and shape) of the irreducible backgrounds that remain in the signal region after the event selection.

A crucial part of the analysis – numerous methods available and are being devised.

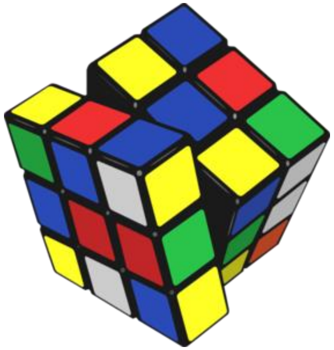
Use predictions from Monte Carlo simulations!

- Contains all our knowledge on the theory and on our detector.
- It is a long, but persistent way from roughness to precision.

Devise data-driven estimation methods:

A common principle: Use control regions

- Find a region in the cut phase space which is background enriched and signal depleted (the control region).
- Obtain the information on BG and extrapolate it to the signal region.

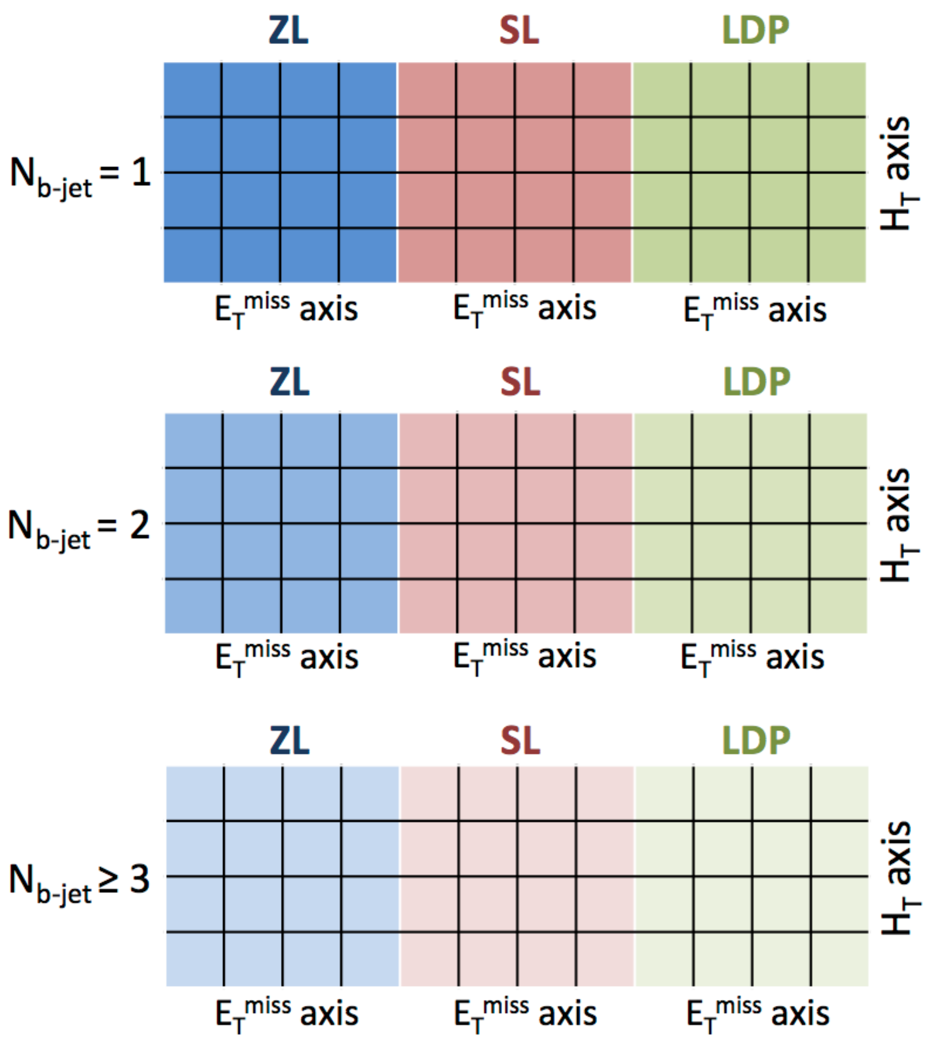


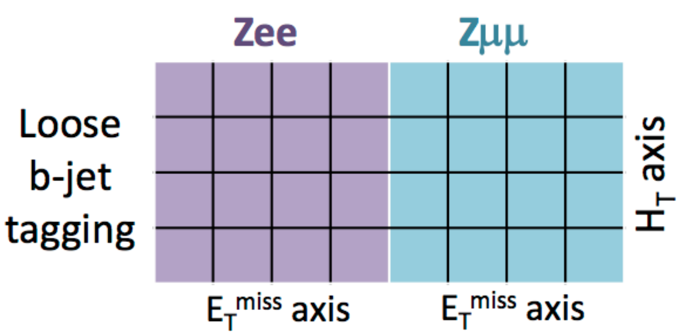
Data and MC work together:

- Data is used for tuning MC parameters
- For well-described kinematic variables, MC shapes are used in BG estimation.

Example signal and control sample definition
Final state: ≥3j, ≥1b, MET - I

ZL = Zero Lepton; signal sample	**SL** = Single Lepton; top & W+jets control sample	**LDP** = low $\Delta\hat{\phi}_{min}$; QCD control sample	**Zee** = $Z \to e^+e^-$; Z to $\nu\bar{\nu}$ control sample	**Zμμ** = $Z \to \mu^+\mu^-$; Z to $\nu\bar{\nu}$ control sample





Bin	H_T (GeV)	E_T^{miss} (GeV)
1	400 – 500 (HT1)	125 – 150 (MET1)
2	500 – 800 (HT2)	150 – 250 (MET2)
3	800 – 1000 (HT3)	250 – 350 (MET3)
4	> 1000 (HT4)	> 350 (MET4)

Example signal and control sample definition
Final state: ≥3j, ≥1b, MET - II

Build an expression that links expected background yield for each background in each bin with the observed yield in the control sample for the relevant background for each bin.

scale factor common to all bins

Expected BG in bin ijk of signal sample

$$\mu^{\mathrm{ttWj}}_{\mathrm{ZL};i,j,k} = S^{\mathrm{ttWj}}_{i,j,k} \cdot R^{\mathrm{ttWj}}_{\mathrm{ZL/SL}} \cdot \mu^{\mathrm{ttWj}}_{\mathrm{SL};i,j,k}$$

Observed yield in the control sample

bin-by-bin MC-based scale factor which accounts for the shape difference between signal and control samples.

There are many more methods. A few more are in the backup slides.

Statistical modeling and likelihood analysis

- The statistical model of an analysis provides the complete mathematical description of that analysis.
- It relates the observed quantities x to the parameters θ through the probability density $p(x|\theta)$.
- The likelihood $L(\theta) = p(X_0|\theta)$ is the probability density $p(x|\theta)$ evaluated at the observed values X_0 of the observables x.
- A likelihood is the starting point of any serious interpretation.

Example model – multibin Poisson

We count events. The probability of counting/observing N events for an expected average n = s (:signal) + b (:background) is given by a Poisson distribution:

$$Pois(N|s+b) = \frac{(s+b)^N e^{-(s+b)}}{N!}$$

Generally, s and b are given in terms of some parameters:

$$Pois(N|\sigma,\epsilon,\mathcal{L},\beta_j) = \frac{(s(\sigma,\epsilon,\mathcal{L}) + b(\beta_j))^N e^{-(s(\sigma,\epsilon,\mathcal{L})+b(\beta_j))}}{N!}$$

σ: cross section, L: luminosity, ε: efficiency and b_j: some BG shape parameters. When we have I disjoint bins, we can take the product of the Poisson for each bin:

$$p\left(\sum_i N_i|\sigma,\epsilon_i,\mathcal{L},\beta_j\right) = \prod_i \frac{(s_i(\sigma,\epsilon_i,\mathcal{L}) + b_i(\beta_j))_i^N e^{-(s_i(\sigma,\epsilon_i,\mathcal{L})+b_i(\beta_j))}}{N_i!}$$

We insert the observed counts N_i to get the likelihood, and estimate the parameters from the likelihood using dedicated statistical methods.

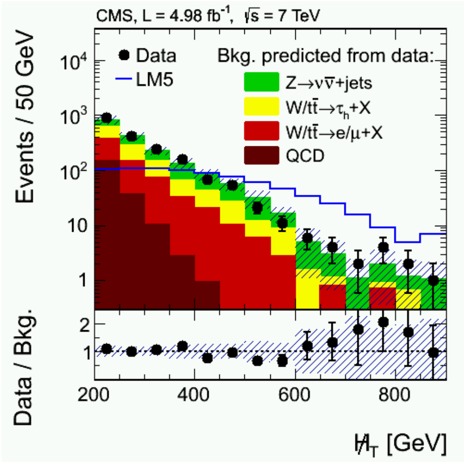
Results and interpretation

- An experimental result is the empirical outcome of the experiment, such as an event count, or the measurement of some physical quantity, such as mass, cross section, spin, charge asymmetry, kinematic edges, etc.
- Given an experimental result, we can find its effect on a theoretical model.
- Interpretation is the act of comparing the experimental results to theoretical model predictions. Beware - it is NOT the experimental result!
- We use likelihoods that incorporate signal predictions to evaluate the impact of the searches on the candidate models.

Systematic uncertainties

Systematic uncertainties are those that cause a shift in the mean of a measurement from the true value.

Systematics are calculated for background estimates, derived measurements (mass, cross section, endpoint, etc.) and for MC predictions of signals (which are used for interpretation).



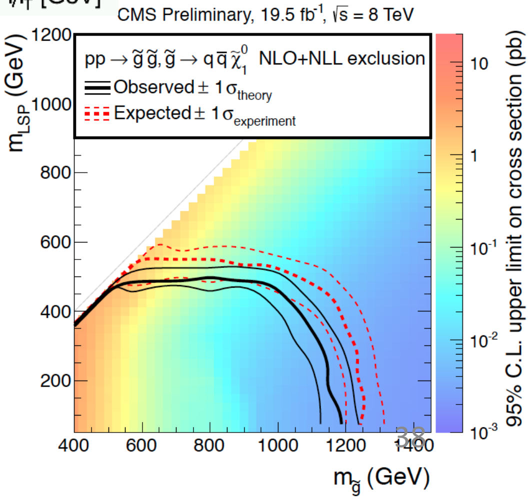
Typical sources of systematics are:

Experimental:

- Luminosity calculation
- Trigger efficiencies
- Jet energy scale, jet energy resolution
- Lepton, photon, b-tag, W-tag, top-tag, etc. efficiencies

Theoretical:

- Cross section and branching ratio calculations
- Parton distribution functions
- ISR/FSR, renormalization scale/factorization scale



SUSY searches at CMS

in other words "how we couldn't yet find SUSY yet"

What is going on in CMS SUSY now?

Most up-to-date public CMS SUSY results are listed here:
https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS

- Inclusive / generic searches: target mostly gluinos and $1^{st}/2^{nd}$ generation squarks
  - in this talk, S. Paktinat's and E. Eskandari's talks
- Naturalness-inspired searches: targeting light stops/sbottoms and light gluinos with 3^{rd} generation decay modes
  - in S.Paktinat's talk
- Search for pair production of electroweak gauginos and sleptons
  - in H. Bakshian's and A. Fahim's talks
- Search for Higgs in SUSY decays
  - in this talk
- Search for multi-leptons and R-parity violating signatures
  - in B. Safarzadeh's talk

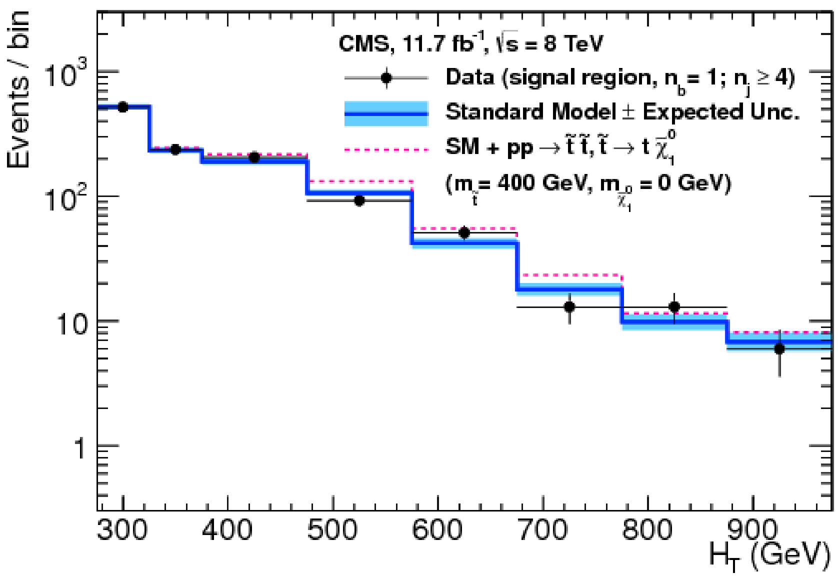
Jets, b-jets, E_T^{miss} with H_T, α_T (8 TeV, 11.7 fb^{-1}) Inclusive search

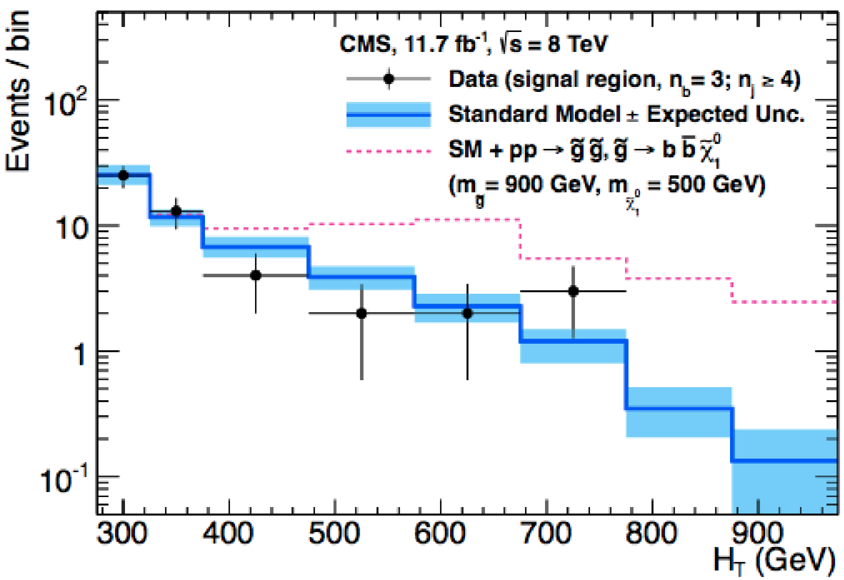
CMS-SUS-12-028, EPJC 73 (2013) 2568

Signal selection: ≥2jets, jet1,2 p_T > 100 GeV, no isolated leptons, α_T > 0.55

Signal final state: binned in jet multiplicity (sensitivity to ~q~q, ~q~g and ~g~g), b jet multiplicity (sensitive to 3rd generation) and HT (probe models with large mass splitting range)

Likelihood analysis using a multibin Poisson. No excess over SM observed.





Jets, H_T^{miss} (8 TeV, 11.7 fb^{-1}) Inclusive search

CMS-SUS-13-012

Signal selection: ≥3jets, HT > 500 GeV, H_T^{miss} > 200 GeV, no isolated leptons, $\Delta\Phi(jet_{1,2,3}, H^T_{miss}) > (0.5, 0.5, 0.3)$

Signal final state: 36 bins in jet multiplicity, H_T and H_T^{miss}.

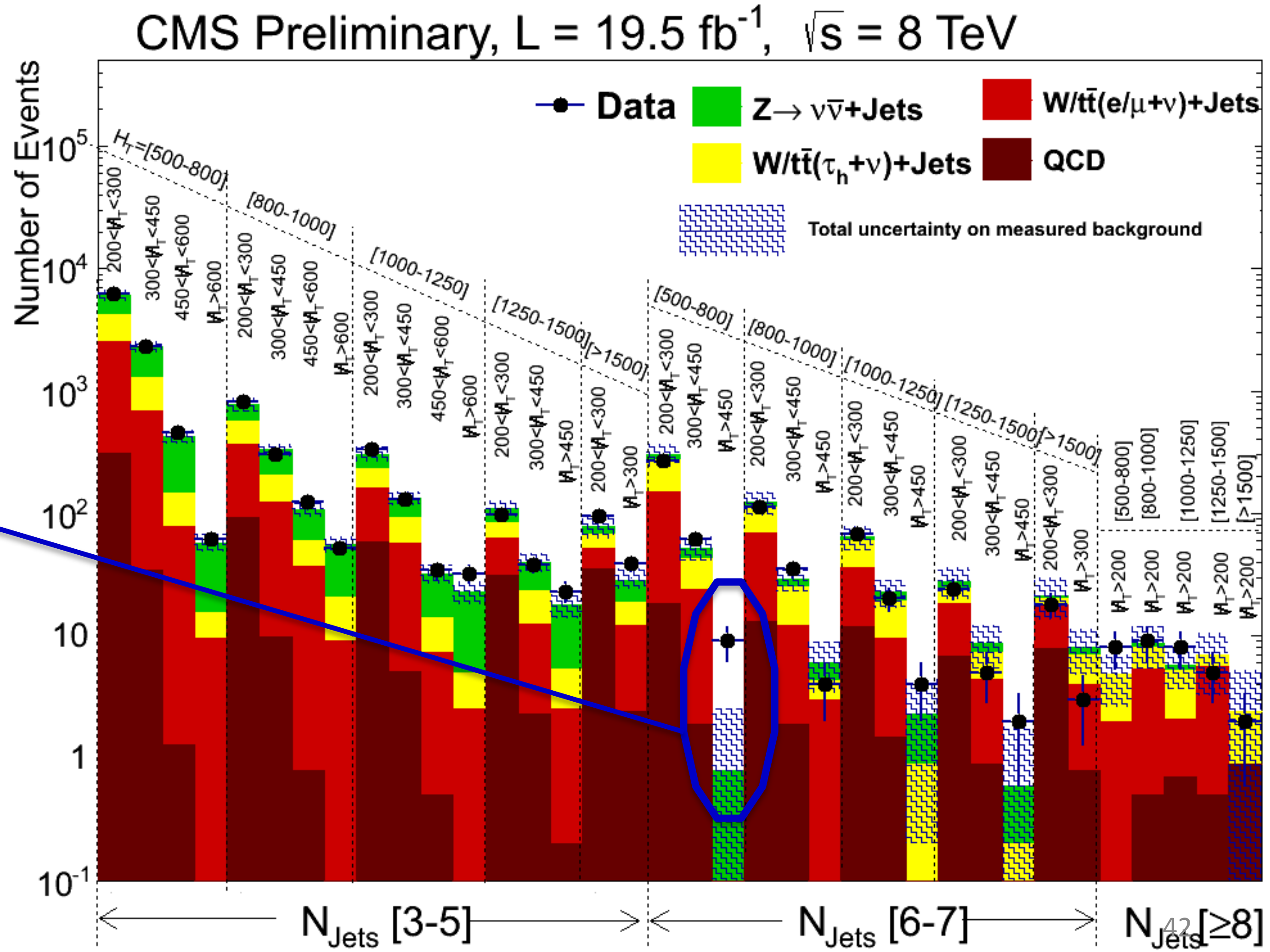
N = 9

b = 0.8 ± 1.7 (2.7σ)

Could this be interesting?

The effect reduces when we include the impact of doing the analysis simultaneously in 36 bins.

More data will tell.



Jets, b-jets, E_T^{miss} with H_T, $\Delta\Phi_{min}$ (8 TeV, 19.4 fb^{-1})
Inclusive search, 3rd generation

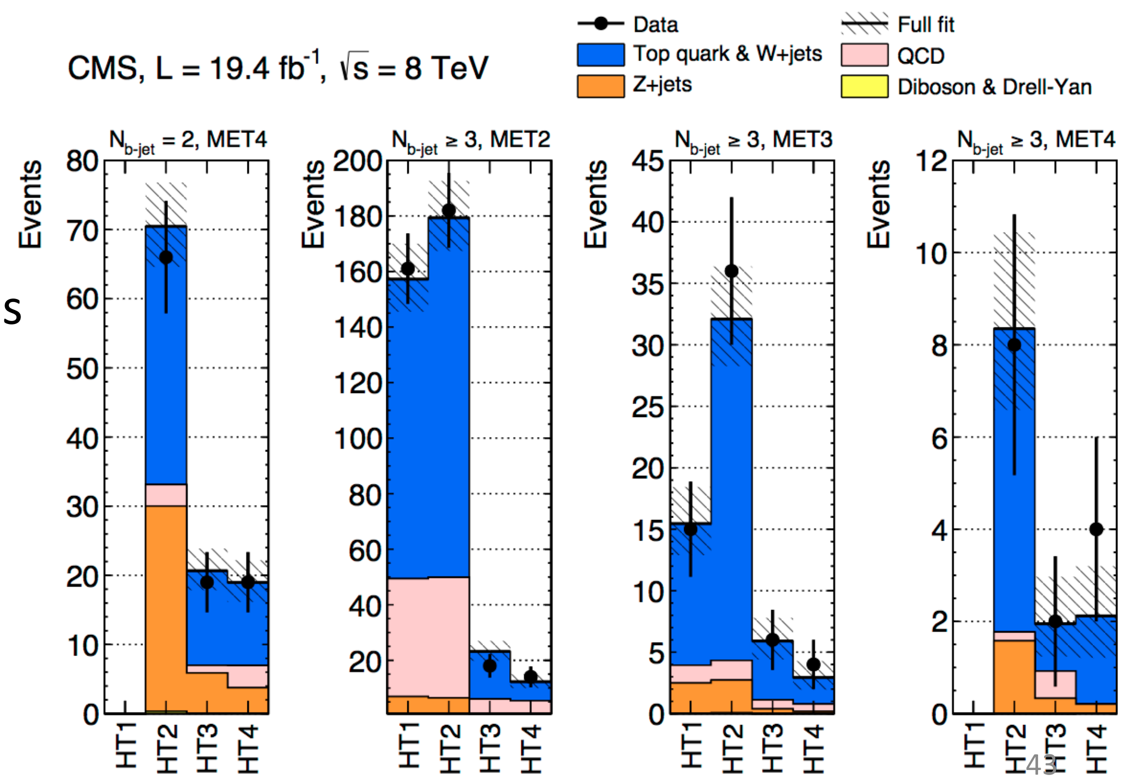
CMS-SUS-12-024, PLB 725 243 (2013)

Signal selection: ≥3jets, jet1,2 p_T > 70 GeV, ≥1 b-tagged jets, no isolated leptons, H_T > 400 GeV, MET > 125 GeV, $\Delta\Phi_{min}$ (jet_i, MET) > 4.0

Signal final state: binned in jet multiplicity, HT and MET.

Likelihood analysis using a multibin Poisson. No excess over SM observed.

Comparison of data with the SM prediction in the 14 most sensitive bins to new physics, as found in the likelihood fit with SUSY signal strengths set to 0. Data consistent with the SM.



Anomolous production of multileptons - I
Inclusive search

CMS-SUS-13-002

Signal selection:

- 3 or 4 leptons (e/μ) with possibly one tau with $p_T > 20$ GeV among them
- 0 or >1 b tagged jet
- HT < 200 GeV or > 200 GeV
- 0, 1 or 2 opposite-sign-same-flavor (OSSF) lepton pairs
- If OSSF exist: dilepton invariant mass m_{ll} below/on/above Z mass.
- Reject events with $m_{ll} < 12$ GeV to avoid low mass resonances.
- Reject events with both $|m_{l+l-} - m_Z| > 15$ GeV and $|m_{l+l-l'} - m_Z| < 15$ GeV to avoid photon conversion from final state radiation.

Signal final states: MET distributions in 64 bins of number of leptons, OSSF pairs, b-tagged jets, taus; dilepton mass wrt Z-mass and H_T.

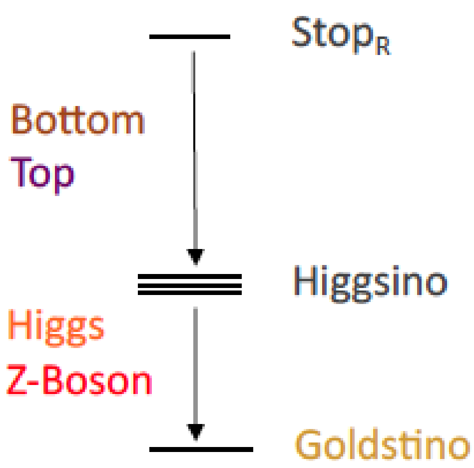
Backgrounds:

- ttbar, WZ → smear MC MET distributions using data
- non-prompt leptons or taus → find conversion factor from data
- asymmetric internal photon conversions → find conversion factor from data

Anomolous production of multileptons - II
Natural Higgsino LSP scenario

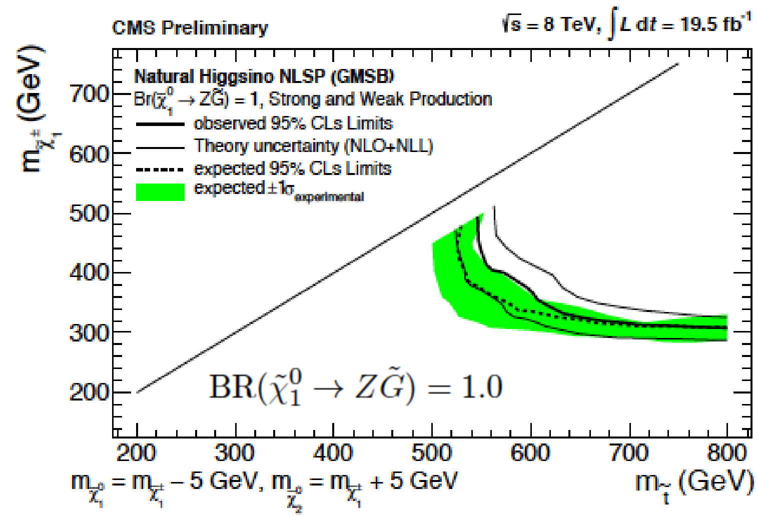
CMS-SUS-13-002

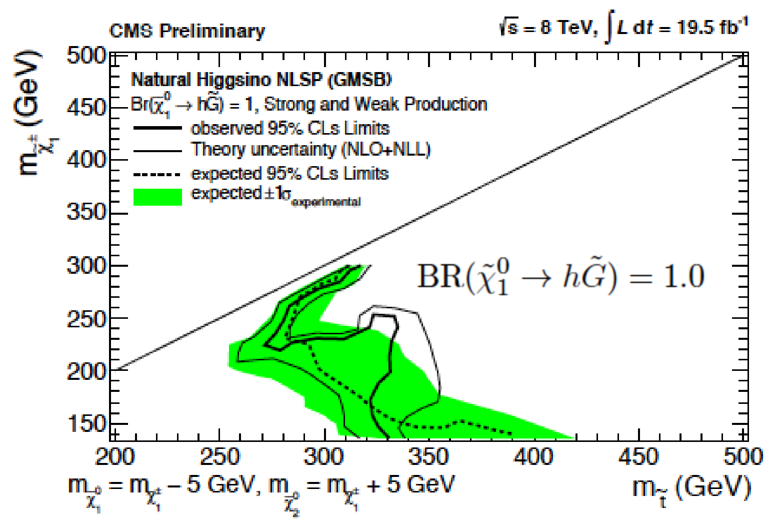
Natural SUSY in GMSB models:

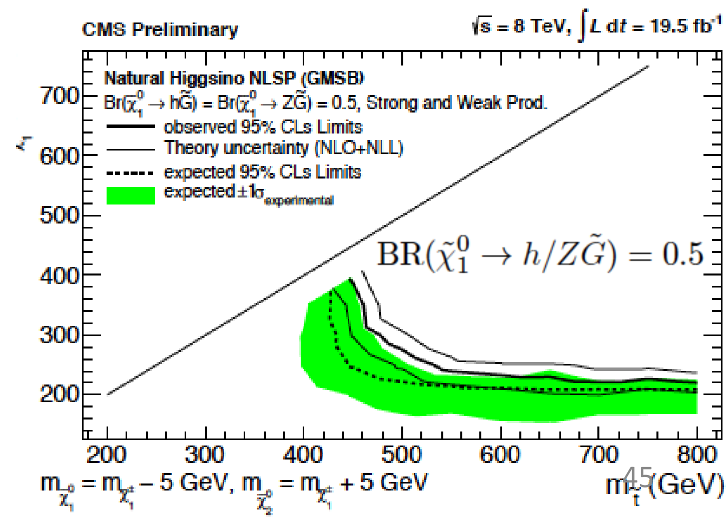


$$\tilde{\chi}_1^0 \rightarrow h\tilde{G} \text{ or } Z\tilde{G}$$

large lepton multiplicity from hh, hZ, ZZ decays

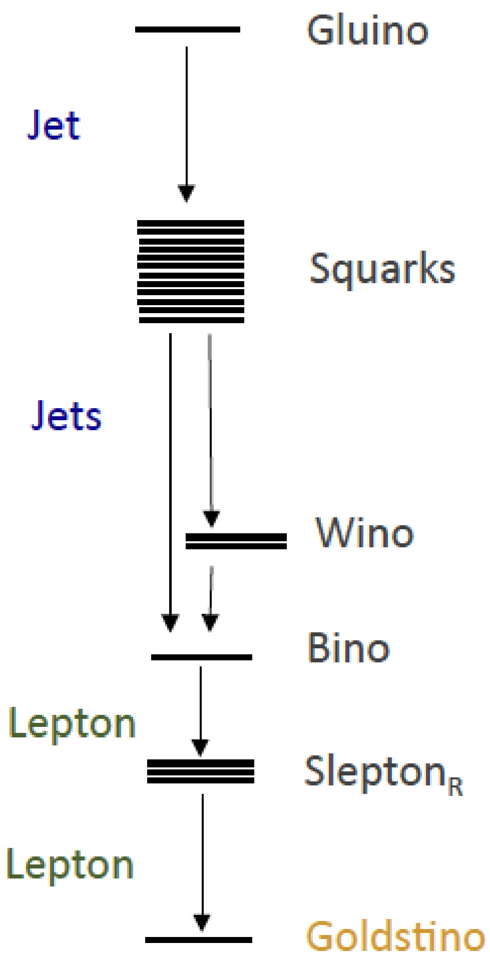




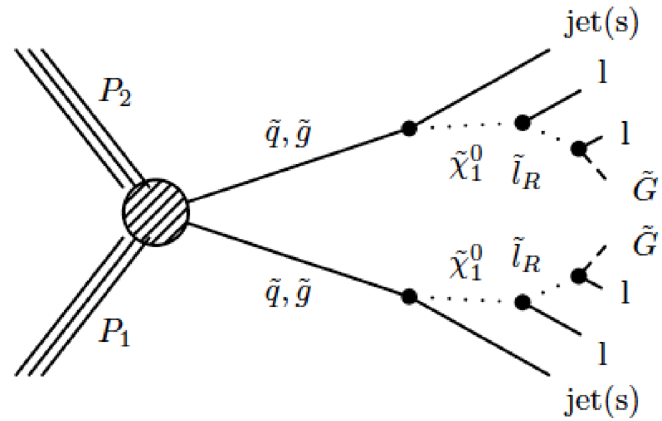


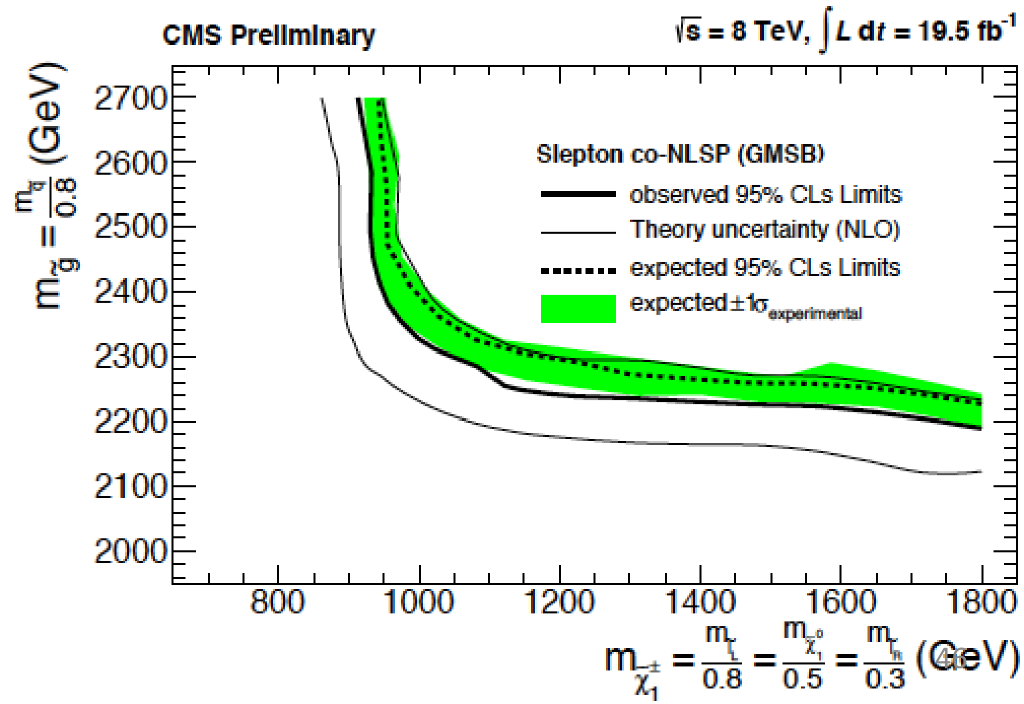
Anomolous production of multileptons - III
Slepton co-LSP

CMS-SUS-13-002



Light sleptons from bino decays lead to multilepton signatures





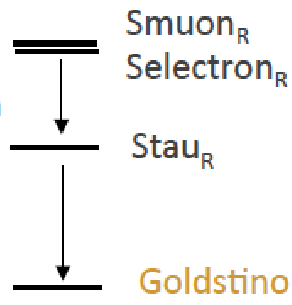
Anomolous production of multileptons - III
Stau-(N)NLSP scenario

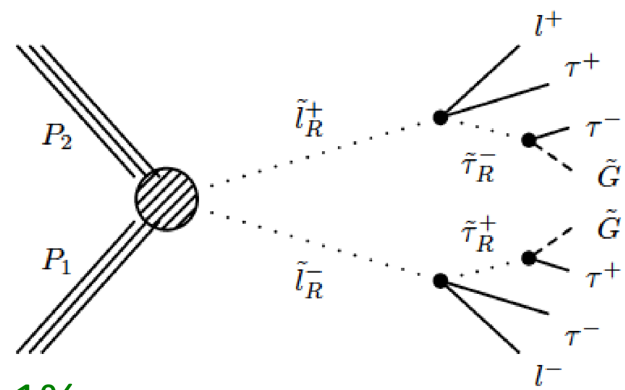
CMS-SUS-13-002

Taus will come from stau decays.

4l (1τ), 1OSSF (off-Z), no b, low HT bin is sensitive to this model.

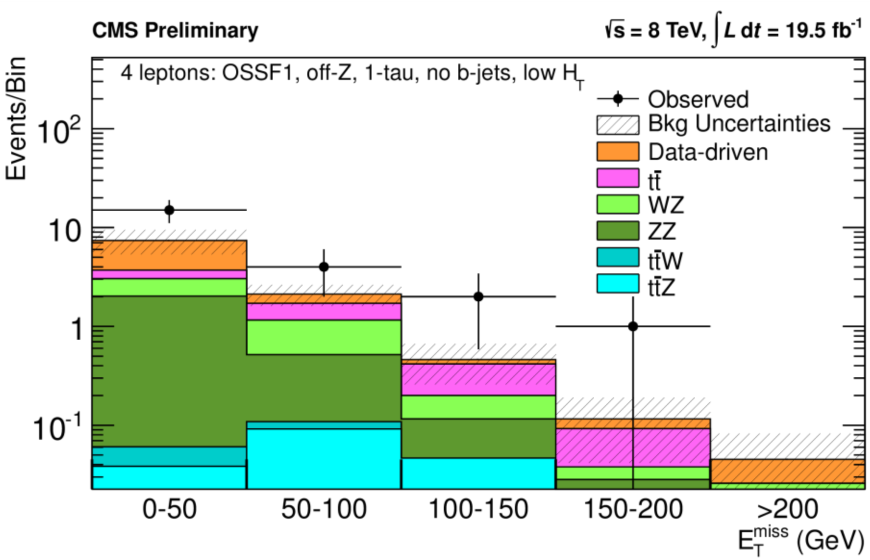
An excess has been observed:
N = 22, b = 10 ± 2.4

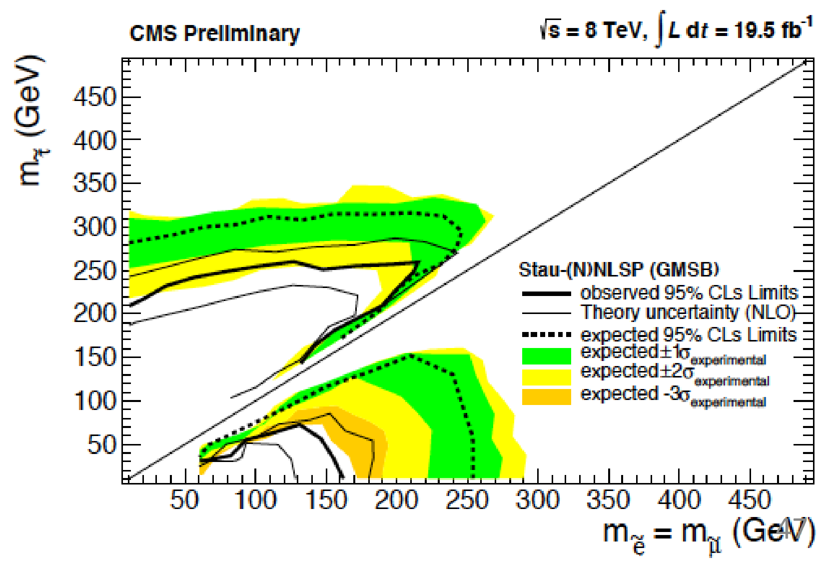




Probability of observing such an excess in a single bin is 1%.

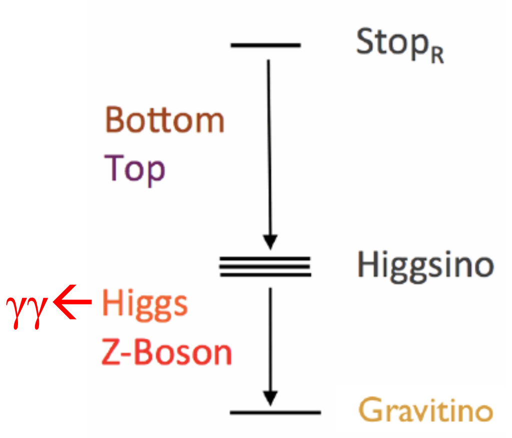
Probability of observing such an excess for this analysis looking at 64 bins simultaneously is 50%.





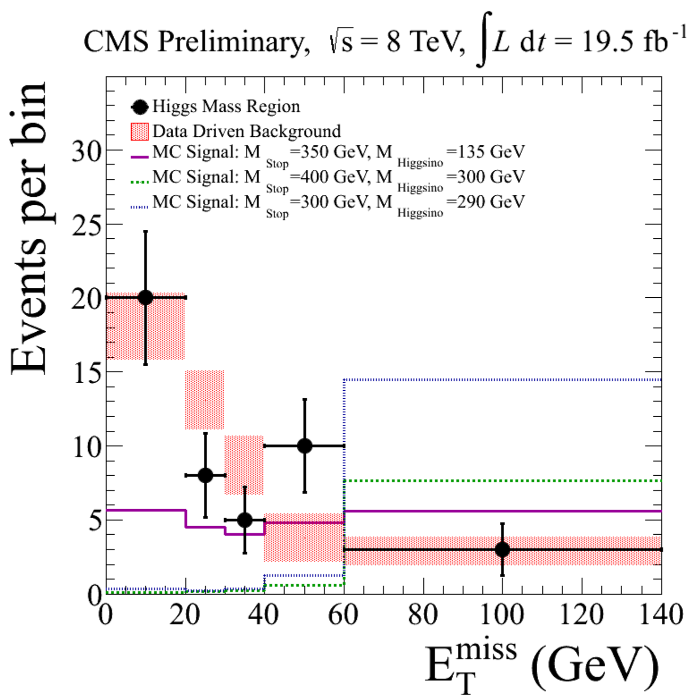
Search for stops and higgsinos in $H \rightarrow \gamma\gamma$ decays
SUSY Higgs search – naturalness-motivated GMSB

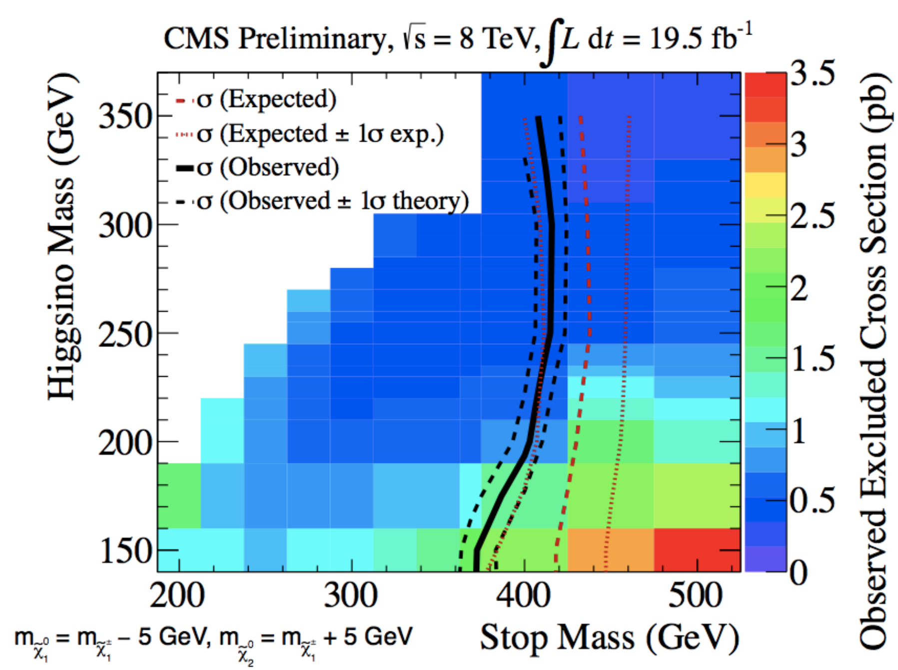
CMS-SUS-13-014



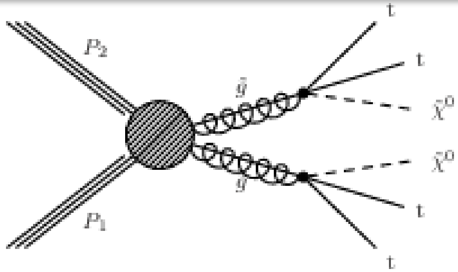
Signal selection: 2 isolated photons (E_T > 40, 25 GeV), ≥2b-jets (p_T > 30 GeV)

BG estimation: Fit a function to $m_{\gamma\gamma}$ in sidebands and extrapolate the fit to the signal region $120 < m_{\gamma\gamma} < 131$. Take the MET shape from the sidebands, normalize to the BG fitted in the signal region and compare with data.

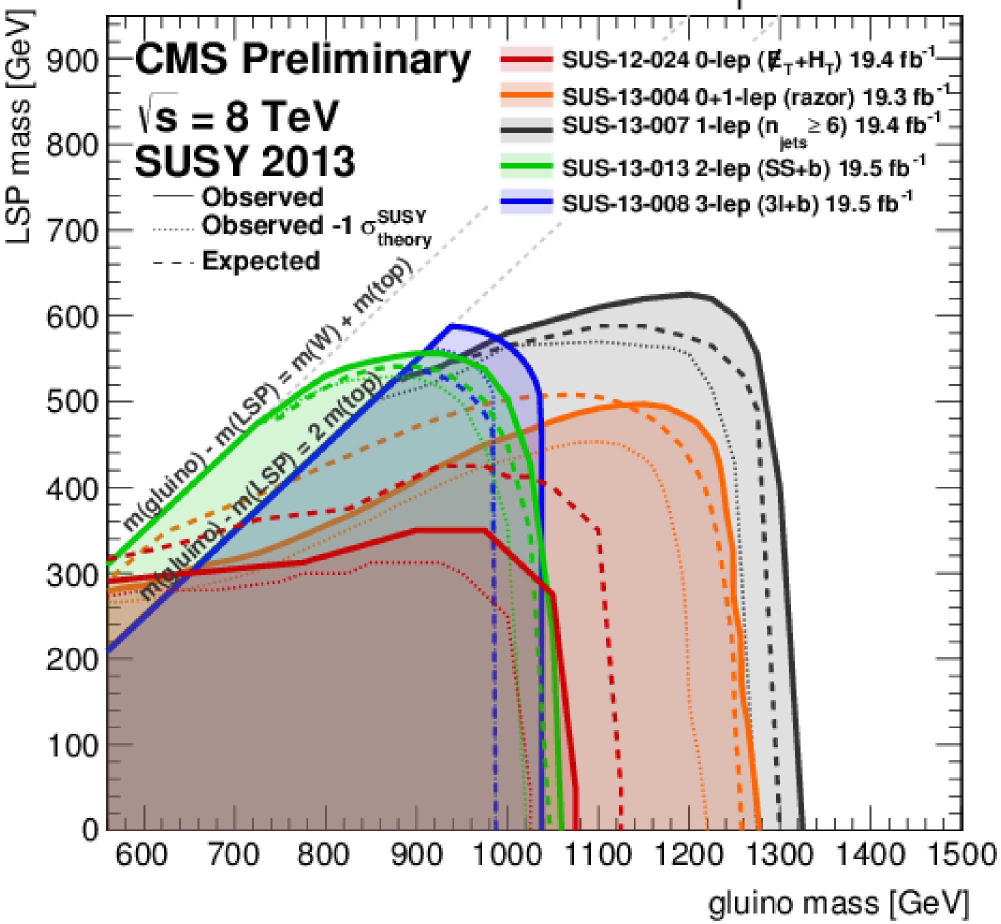


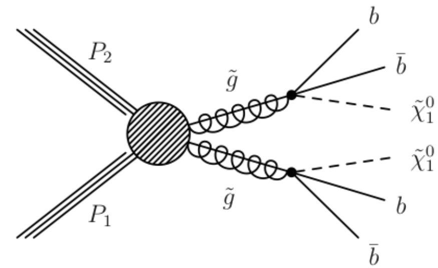


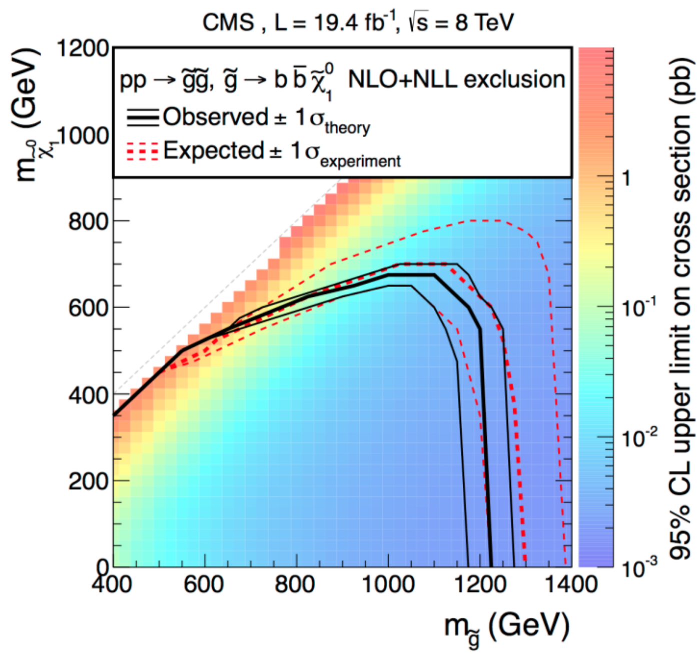
Gluino-neutralino mass reach summary



$\tilde{g}$-$\tilde{g}$ production, $\tilde{g} \rightarrow t\,\bar{t}\,\tilde{\chi}^0_1$

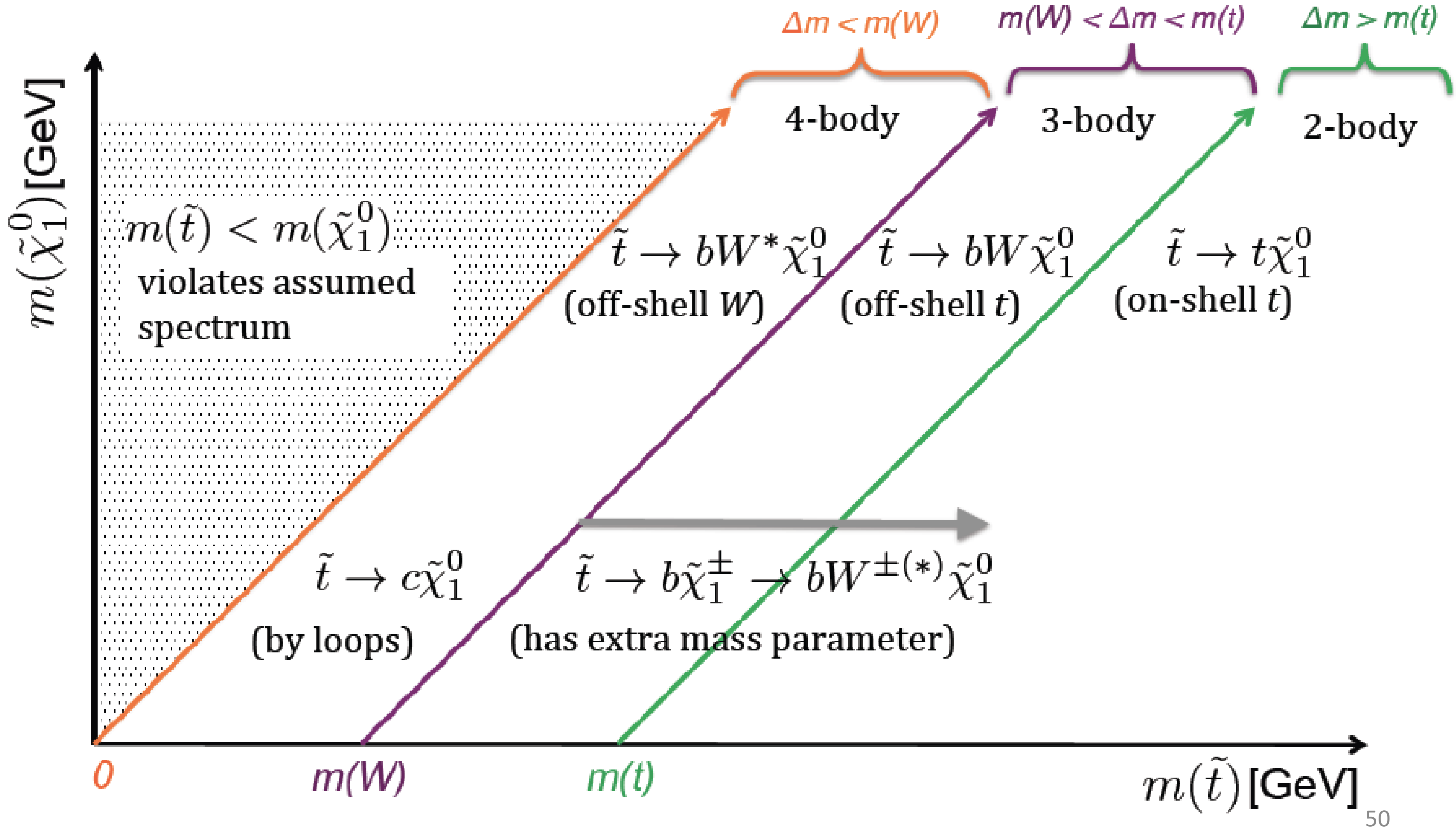




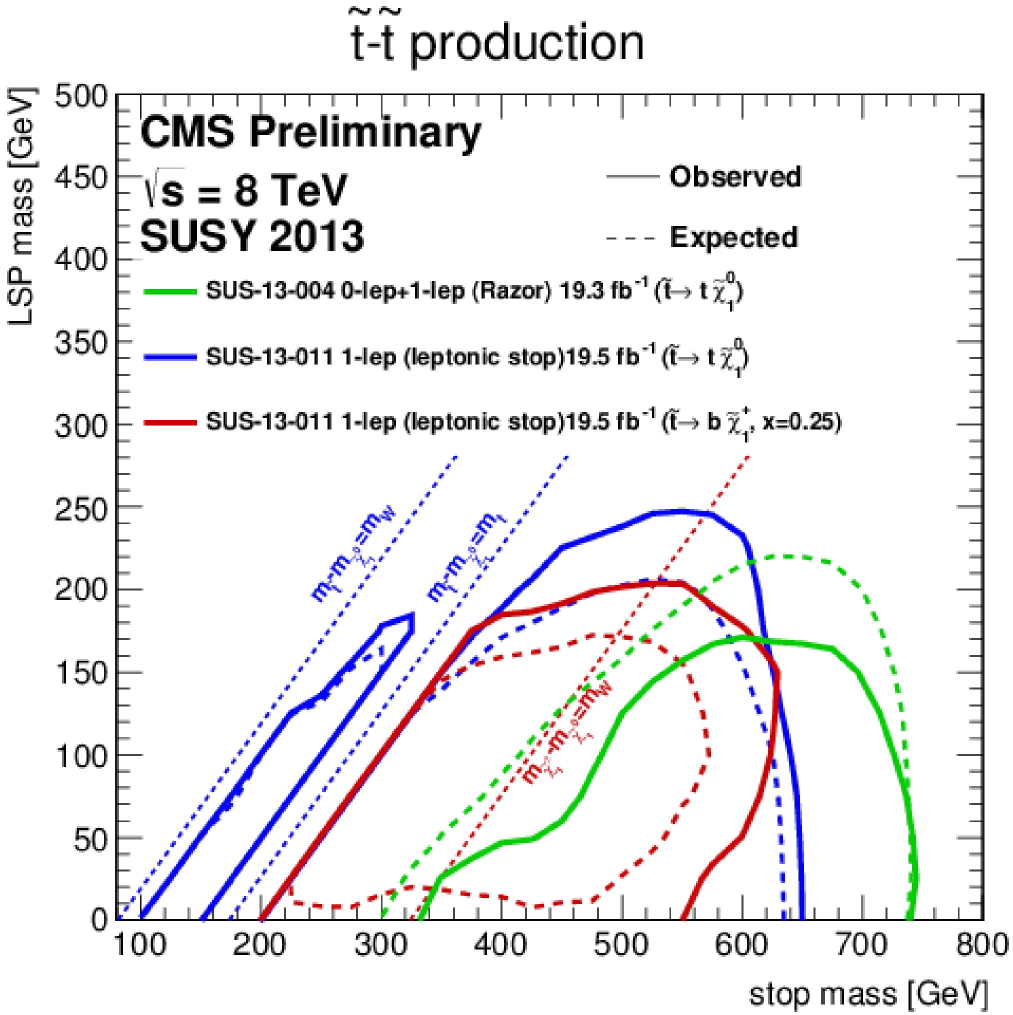


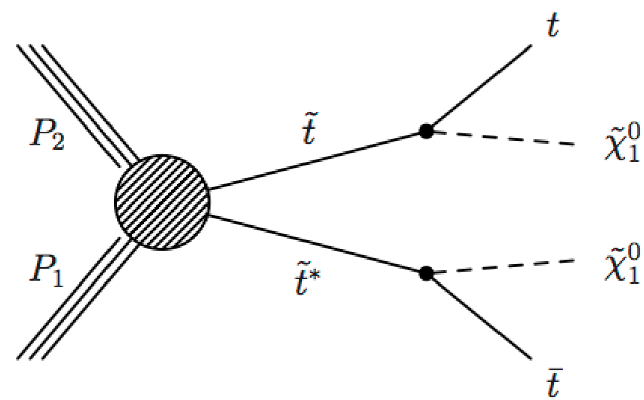
Stop decays and final states

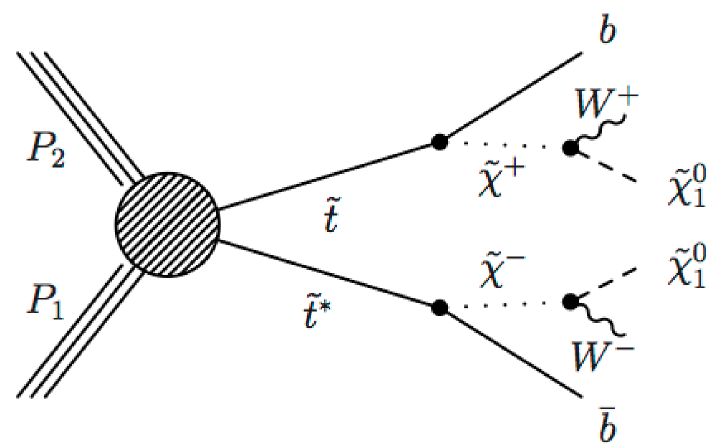
$$\Delta m = m(\tilde{t}) - m(\tilde{\chi}_1^0)$$



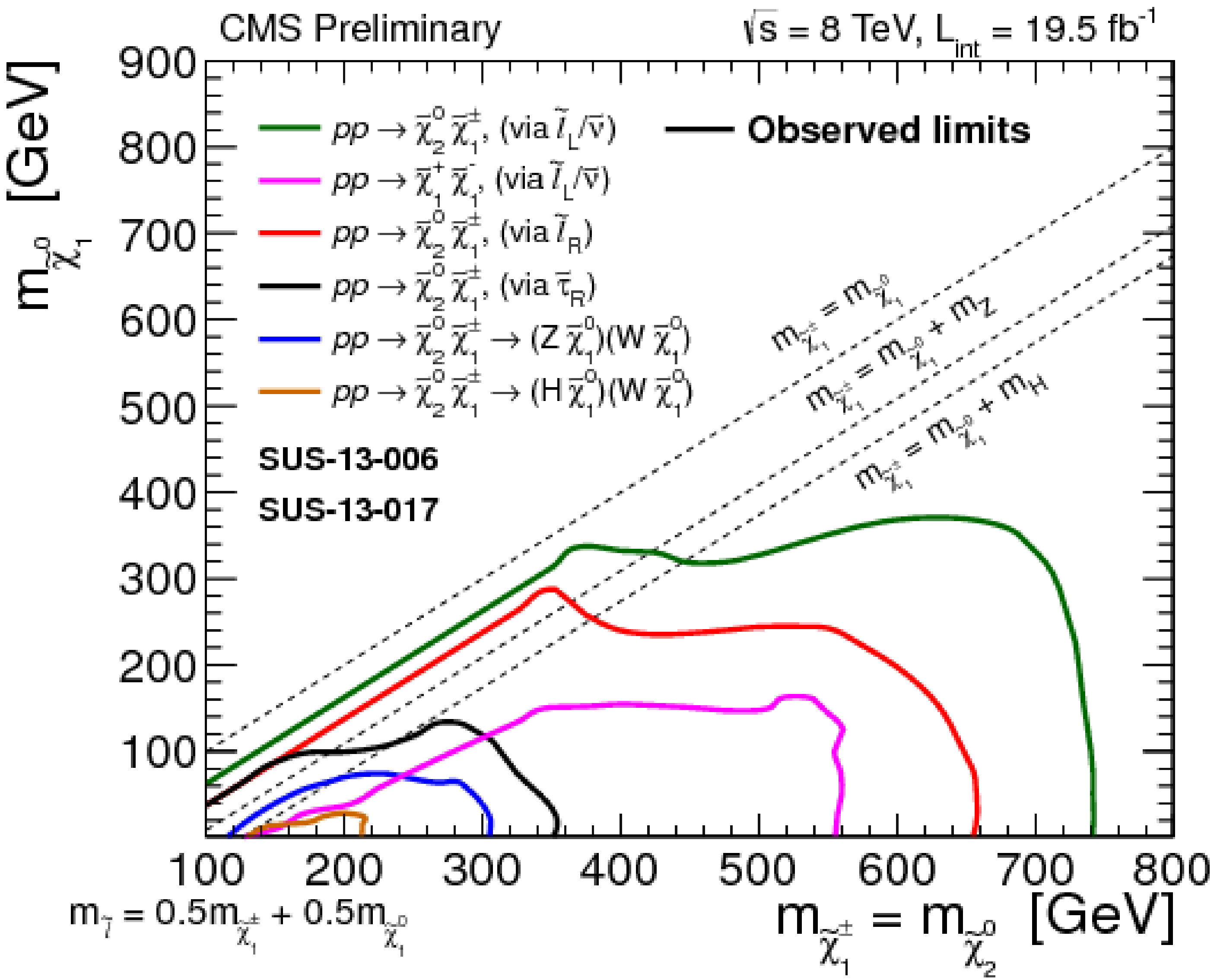
Stop-neutralino mass reach summary



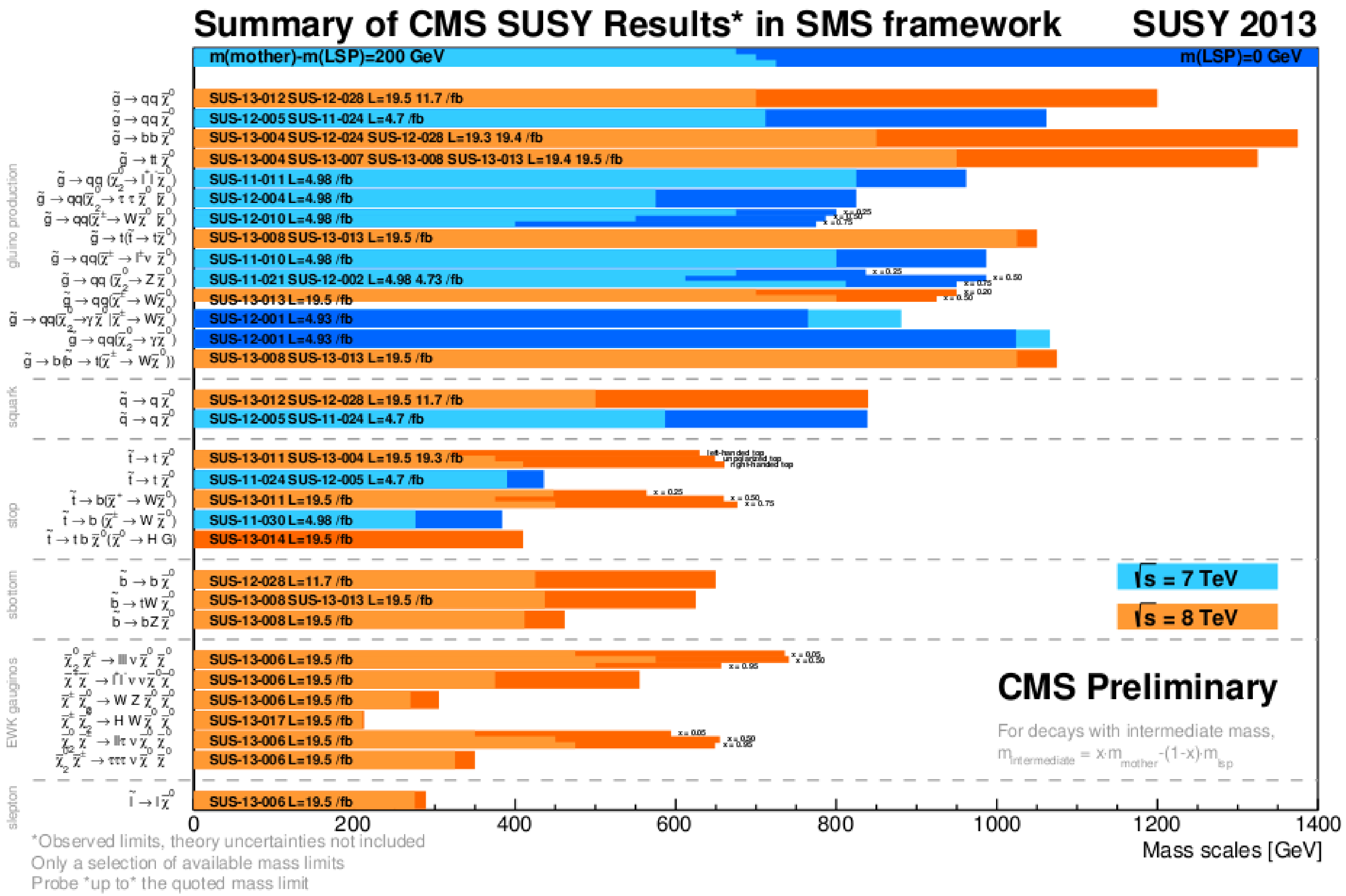




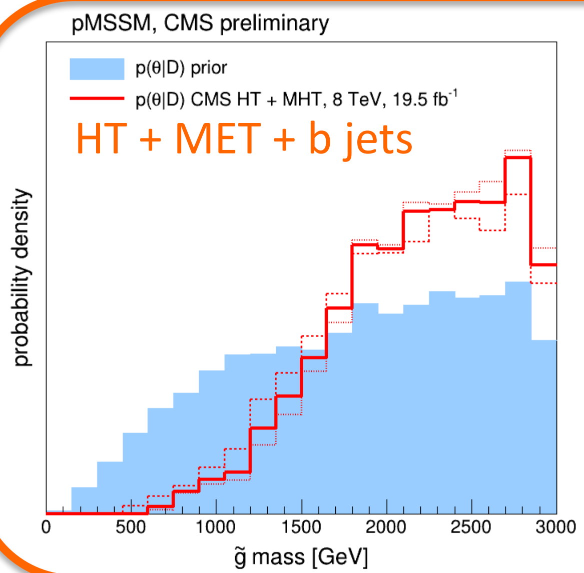
Chargino-neutralino mass reach summary

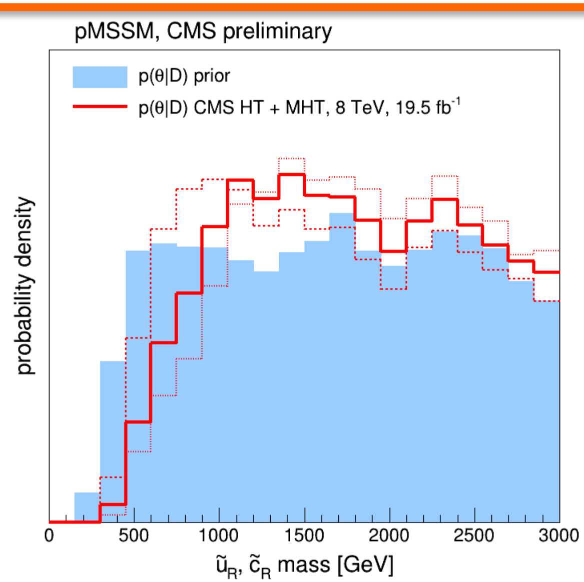


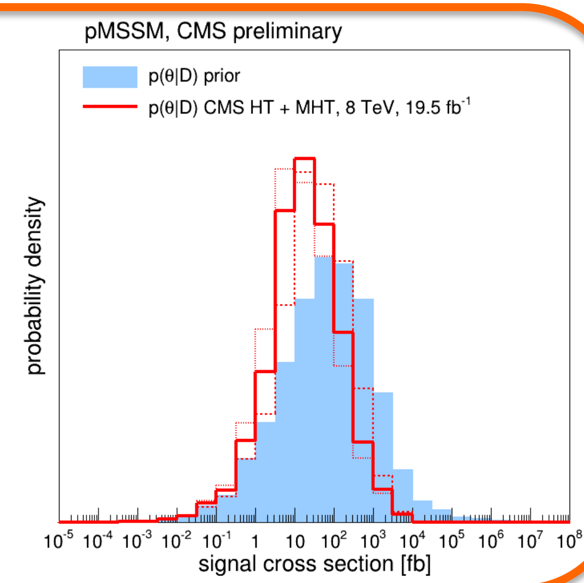
Summary of SMS mass limits

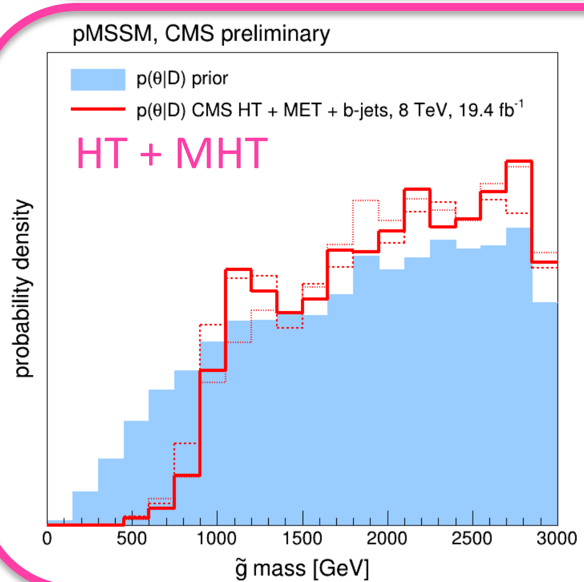


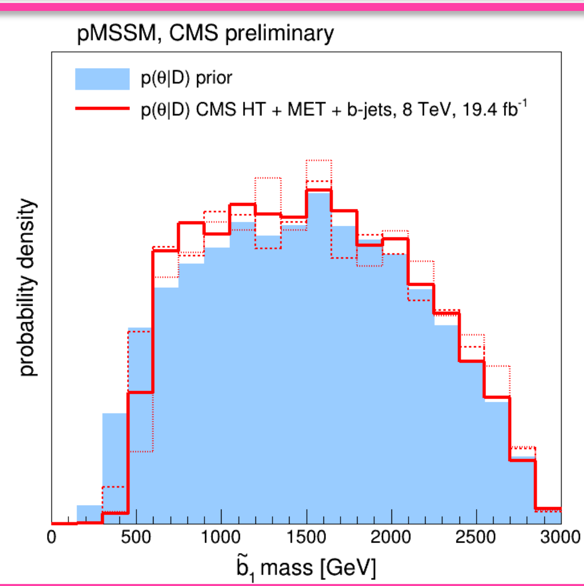
pMSSM interpretation of inclusive searches

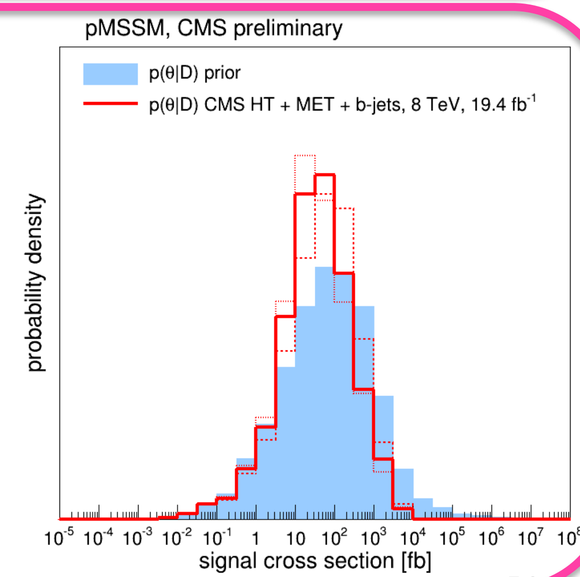








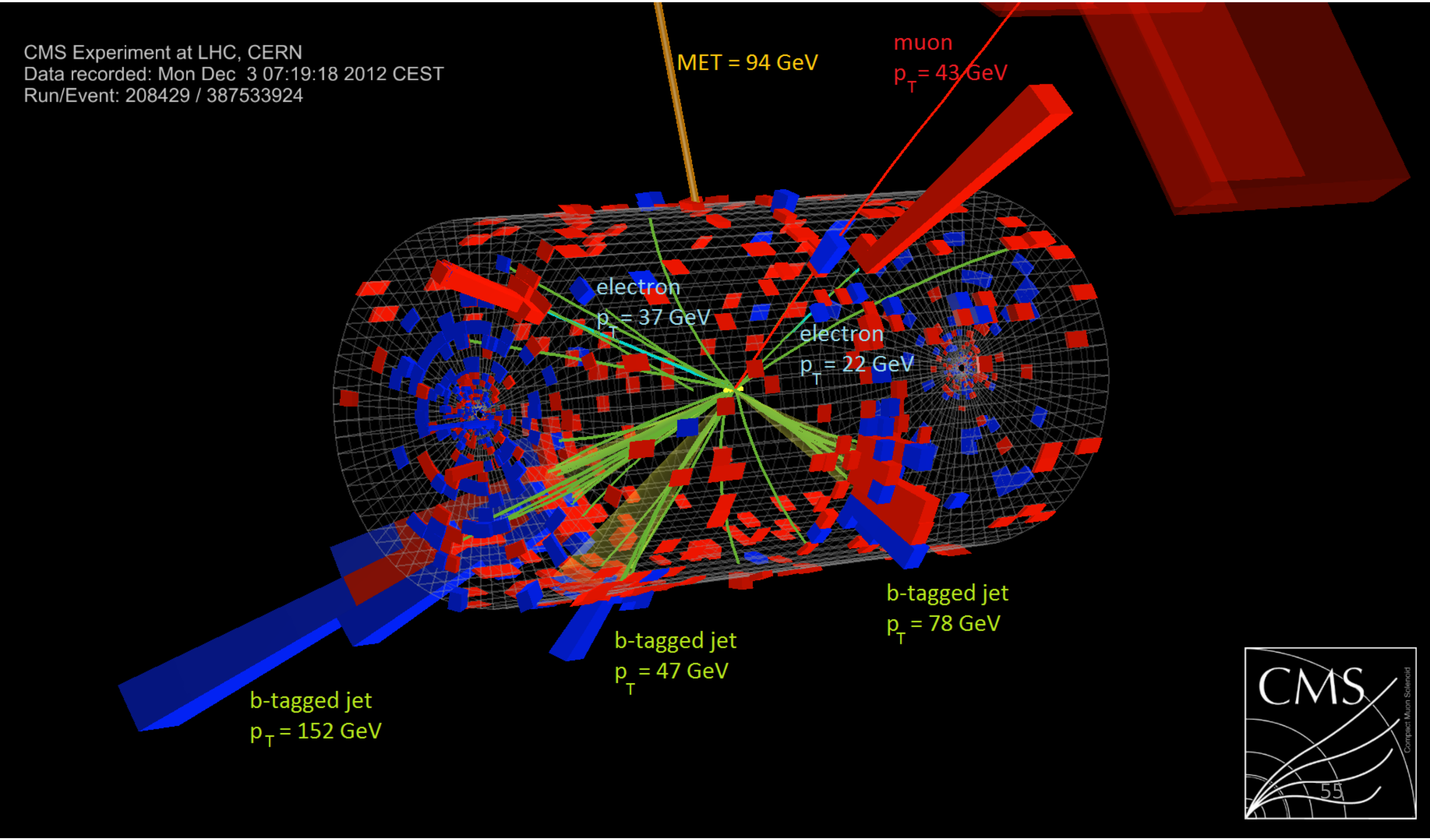




Could this be SUSY?

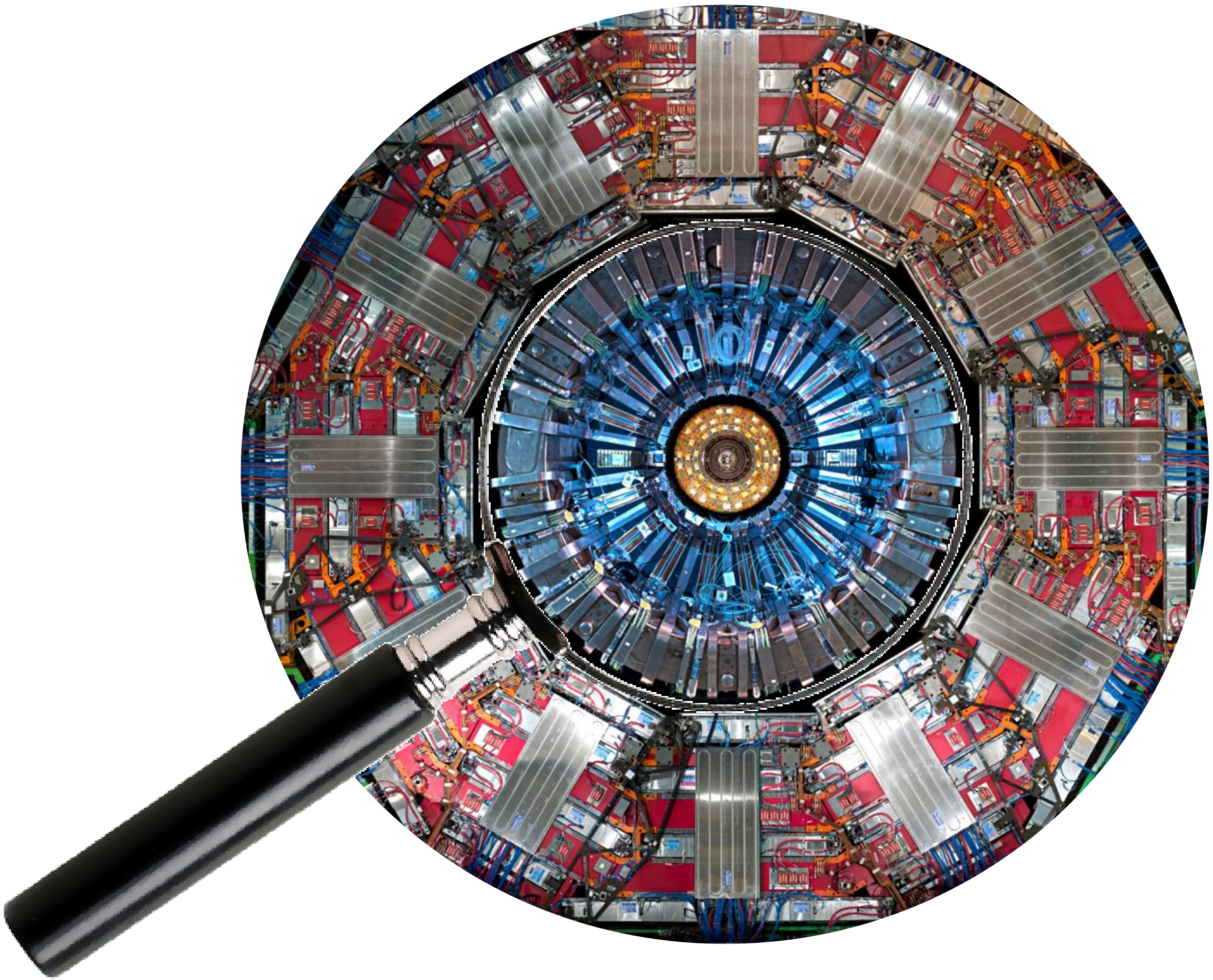
A spectacular 3 leptons + 3-b jets + high missing E_T event at CMS.

CMS-PAS-SUS-13-008



Summary

- CMS has conducted a rich variety of searches with up to 19.5 fb^{-1} of 8 TeV proton proton data – but data is (more or less!) consistent with the SM.
- So we entertain ourselves with disfavoring models and setting limits.
- We've probed gluinos up to 1.3 TeV, squarks up to 800 GeV and stops up to 750 GeV.
- We are focusing more on difficult scenarios with low cross sections, low MET, with compressed spectra (soft objects) and with kinematics resembling the SM.
- We are trying to see Higgs being born from SUSY decays.
- And we are getting ready for the 13-14 TeV run in 2015!



No SUSY yet?

KEEP CALM AND SEARCH ON

BACKUP SLIDES

Background estimation methods
Sideband method

Used in searches for resonances, where the BG has a smooth, well-described shape, and the signal peaks over the BG.

- Define a signal region, and the signal-free control regions, i.e. the sideband regions around the signal.
- Deduce the shape of the BG from the sidebands (polynomial, exponential, etc.?)
- Extrapolate the BG in sidebands to the signal region.
- Either count the extrapolated events under the signal peak – or -- fit the data distribution to BG shape + signal shape and extract the parameters of the BG function.

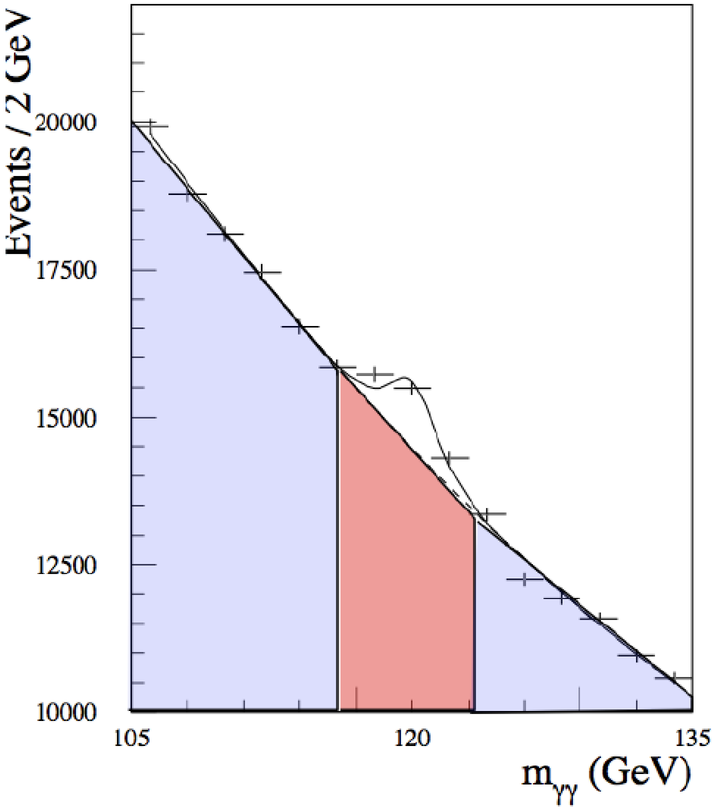
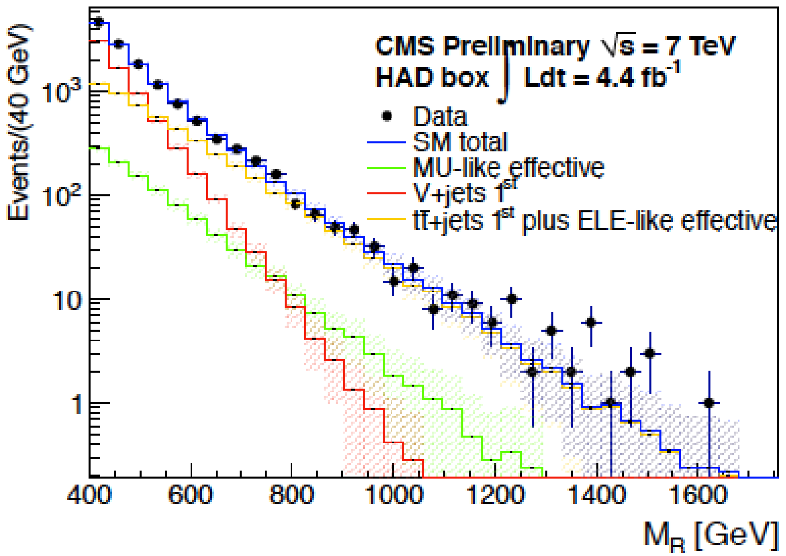


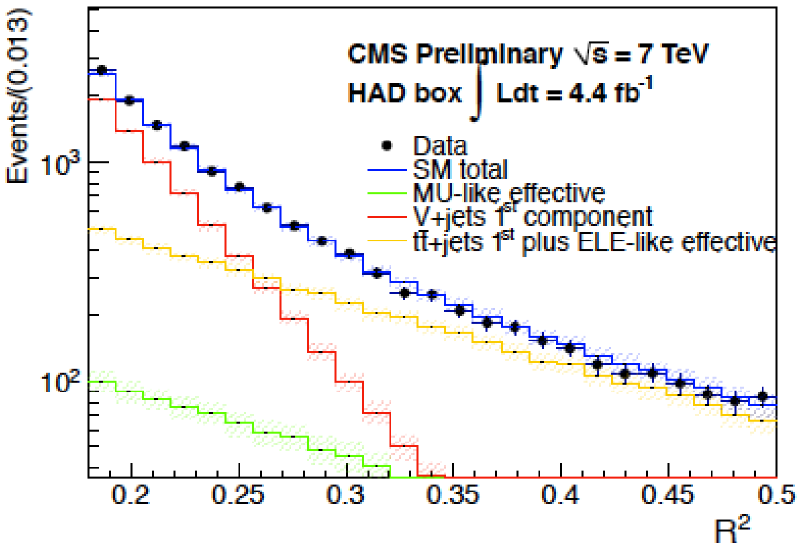
Figure from P. Govoni HCP2011 lectures

Background estimation methods
Fitting to an analytical function

Sometimes the BG is well-described by an analytical function. In these cases

- Find a control region dominated by the BG.
- Find an analytical function that describes the BG well.
- Fit the data to this analytical function in the control region and find the parameters of the analytical function.
- Extrapolate the fit to the signal region.





CMS razor analysis employs a fit to a 2D exponential-like function.

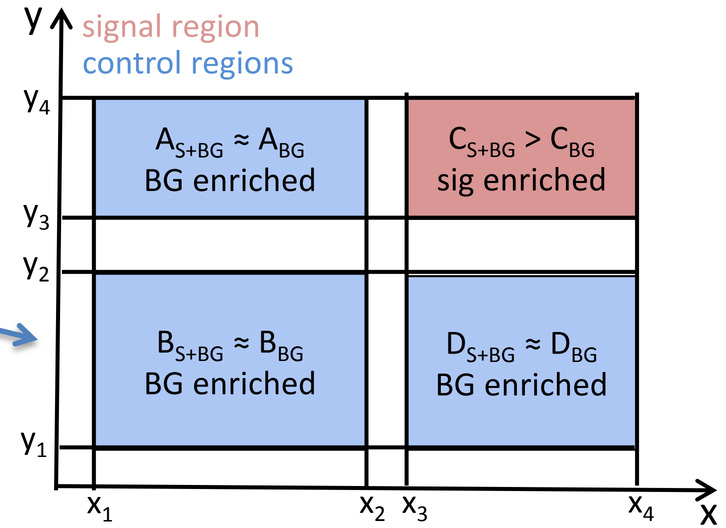
Background estimation methods
The matrix – or ABCD - method

When there exist two variables x and y for which the BG is uncorrelated, i.e. factorizable:

$$f^{BG}(x,y) = f^{BG}(x) \cdot f^{BG}(y)$$

- Apply all cuts except those on x and y on data
- Divide the x-y plane into 4-regions:
- When there is no signal, we have

$$\frac{N_A^{BG}}{N_C^{BG}} = \frac{N_B^{BG}}{N_D^{BG}}$$



- In the presence of signal, C will be contaminated by the signal. But we can estimate the number of BG events in C from

$$N_A^{BG} = \frac{N_B^{BG} \cdot N_C^{BG}}{N_D^{BG}}$$

Note: Always beware the signal contamination in the control regions. Add it as a systematic.

Background estimation methods
Fake rates method

The ratios of objects found by a tight identification over objects found by a loose identification is widely used as a BG estimation tool.

Suppose we would like to estimate QCD in a signal region that has leptons. Real leptons come from the signal and fake leptons come from QCD (jets faking leptons). We define two event selections with loose and tight lepton ID criteria, which can be decomposed as:

$$
\begin{aligned}
N_{loose} &= N_{loose}^{real} + N_{loose}^{fake} \\
N_{tight} &= N_{tight}^{real} + N_{tight}^{fake} \\
\epsilon^k \equiv N_{tight}^k / N_{loose}^k \rightarrow \quad &= \epsilon^{real} N_{loose}^{real} + \epsilon^{fake} N_{loose}^{fake}
\end{aligned}
$$

Background estimation methods
Fake rates method

The ratios of objects found by a tight identification over objects found by a loose identification is widely used as a BG estimation tool.

Suppose we would like to estimate QCD in a signal region that has leptons. Real leptons come from the signal and fake leptons come from QCD (jets faking leptons). We define two event selections with loose and tight lepton ID criteria, which can be decomposed as:

Get these counts from data (N_{loose}, N_{tight})

$$N_{loose} = N^{real}_{loose} + N^{fake}_{loose}$$

$$N_{tight} = N^{real}_{tight} + N^{fake}_{tight}$$

$$\epsilon^k \equiv N^k_{tight}/N^k_{loose} \rightarrow \quad = \epsilon^{real} N^{real}_{loose} + \epsilon^{fake} N^{fake}_{loose}$$

Background estimation methods
Fake rates method

The ratios of objects found by a tight identification over objects found by a loose identification is widely used as a BG estimation tool.

Suppose we would like to estimate QCD in a signal region that has leptons. Real leptons come from the signal and fake leptons come from QCD (jets faking leptons). We define two event selections with loose and tight lepton ID criteria, which can be decomposed as:

Get these counts from data

$$N_{loose} = N_{loose}^{real} + N_{loose}^{fake}$$

$$N_{tight} = N_{tight}^{real} + N_{tight}^{fake}$$

$$\epsilon^k \equiv N_{tight}^k / N_{loose}^k \rightarrow \quad = \epsilon^{real} N_{loose}^{real} + \epsilon^{fake} N_{loose}^{fake}$$

Find the ID efficiency, using e.g. tag and probe method.

Find this efficiency, i.e. the fake rate using e.g. a QCD control sample (e.g. low MET)

Background estimation methods
Fake rates method

The ratios of objects found by a tight identification over objects found by a loose identification is widely used as a BG estimation tool.

Suppose we would like to estimate QCD in a signal region that has leptons. Real leptons come from the signal and fake leptons come from QCD (jets faking leptons). We define two event selections with loose and tight lepton ID criteria, which can be decomposed as:

Get these counts from data

$$N_{loose} = N_{loose}^{real} + N_{loose}^{fake}$$

$$N_{tight} = N_{tight}^{real} + N_{tight}^{fake}$$

$$\epsilon^k \equiv N_{tight}^k / N_{loose}^k \rightarrow \quad = \epsilon^{real} N_{loose}^{real} + \epsilon^{fake} N_{loose}^{fake}$$

Solve the two equations simultaneously to get these numbers.

Find the ID efficiency, using e.g. tag and probe method.

Find this efficiency, i.e. the fake rate using e.g. a QCD control sample (e.g. low MET)

Background estimation methods
Fake rates method

The ratios of objects found by a tight identification over objects found by a loose identification is widely used as a BG estimation tool.

Suppose we would like to estimate QCD in a signal region that has leptons. Real leptons come from the signal and fake leptons come from QCD (jets faking leptons). We define two event selections with loose and tight lepton ID criteria, which can be decomposed as:

$$N_{loose} = N_{loose}^{real} + N_{loose}^{fake}$$

$$N_{tight} = N_{tight}^{real} + N_{tight}^{fake}$$

$$\epsilon^{k} \equiv N_{tight}^{k} / N_{loose}^{k} \rightarrow \quad = \epsilon^{real} N_{loose}^{real} + \epsilon^{fake} N_{loose}^{fake}$$

Get these counts from data

Solve the two equations simultaneously to get these numbers.

Find the ID efficiency, using e.g. tag and probe method.

Find this efficiency, i.e. the fake rate using e.g. a QCD control sample (e.g. low MET)

Finally obtain the number of BG events from $\epsilon^{fake} N_{loose}^{fake} = N_{tight}^{fake} = N_{BG}$

Background estimation methods
Replacing particles: Z → νν from Z → μμ

Z → νν is a troublesome irreducible BG for the hadronic searches that use high MET. But we can use the Z → μμ events to estimate the BG contribution from Z → νν, because Z → νν and Z → μμ events have same kinematic characteristics.

- Select a $\mu^+\mu^-$ sample with $m(\mu^+\mu^-)$ in the Z mass range (we assume this sample is signal-free).
- Count the muons as MET, i.e.: add muon momenta to MET and recalculate the MET.
- Apply the MET cut and count the observed events. The Z → μμ can be estimated from

$$N^{estm}_{Z\to\nu\nu} = \frac{N^{obs}_{\mu\mu} - N^{BG}_{\mu\mu}}{A^{GEN}_{\mu\mu} \cdot \epsilon_{\mu}} \cdot R\left(\frac{BR(Z\to\nu\nu)}{BR(Z\to\mu\mu)}\right)$$

Number of observed μμ events

non-Z → μμ BG in the μμ sample

ratio of branching ratios

Z → mumu selected / total Z → mumu in generator level

muon reconstruction efficiency

Background estimation methods
Using flavors – opposite flavor subtraction

Suppose we have a signal and a BG with dilepton final state where

- for the signal, flavors of the two leptons are correlated (i.e. decays to same flavor (SF), ee or μμ alone, but not to opposite flavor (OF), eμ) - e.g. Z/Z' decays, neutralino decays, etc.
- for the BG, flavors of the two leptons are uncorrelated (i.e. decays to ee, μμ and emu) - e.g. ttbar, WW.

The amount of BG in the SF and OF regions can be related via branching ratios (for ttbar, $N_{SF} = N_{OF}$).

Thus, to estimate the BG in SF region:

- Count the events in the OF region
- Correct the number for branching ratios and lepton ID efficiencies.