Search for supersymmetry two photons + jet events in pp collisions at sqrt(s)= 8 TeV

with a brief discussion:

Motivation and trigger strategies for displaced jet searches in LHC Run 2 data

In fulfillment of the Princeton Physics PhD Pre-Thesis Requirement:

Joshua Hardenbrook

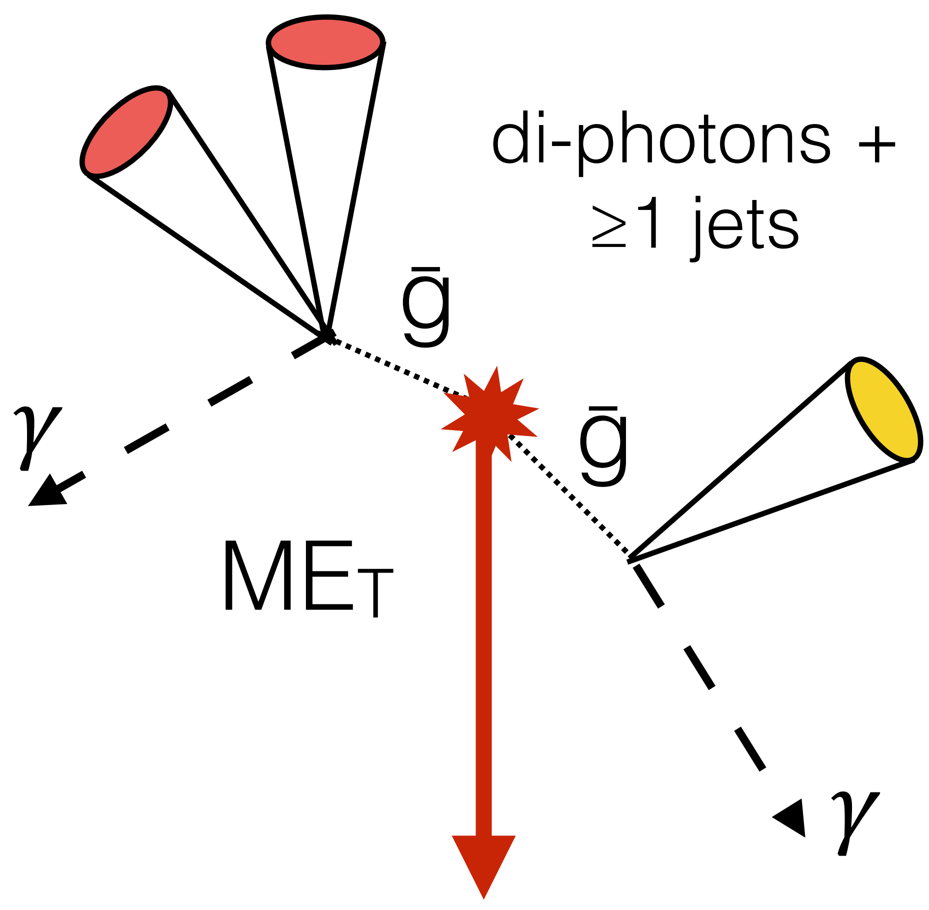
Outline

- Introduction
- Theoretical motivations
- The Razor Kinematic Variables
- Analysis
- Interpretation
- Analysis Conclusions
- Brief Discussion of Displaced Jets

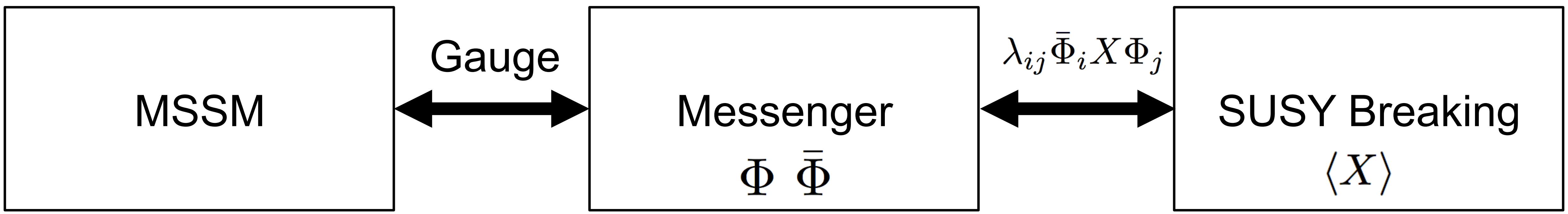
https://twiki.cern.ch/twiki/bin/viewauth/CMS/PhysicsResultsSUS14008

Introduction

- We present a search for SUSY in diphoton + jet(s) + missing energy final states using the razor variables
- Our search is motivated by GMSB-like signals with heavy pair production of squarks or gluinos
- We interpret our results in terms of an simplified model T5gg, and a full model inspired by GMSB



Theoretical Motivations: Vanilla GMSB



When supersymmety is broken at the ~~SUSY~~ scale, the associated goldstino acquires a scalar and auxiliary vev M and F respectively:

$$\langle X \rangle = M + \theta^2 F$$

The SUSY breaking is communicated to the MSSM via a messenger sector which interacts with the SUSY scale via a simple super-potential term with the chiral superfield phi:

$$W \supset \lambda_{ij}\bar{\Phi}_i X \Phi_j$$

“General” Gauge Mediation

2008 “General Mediation”. Meade, Seiberg, Shih arXiv:0801.3278v3

Definition: “In the limit that the MSSM gauge couplings go to zero the theory decouples into the MSSM and a separate Hidden Sector”

General Predictions:

1. Flavor universality
2. Small trilinear couplings at the messenger scale
3. **Gravitino LSP (most important for this study)**

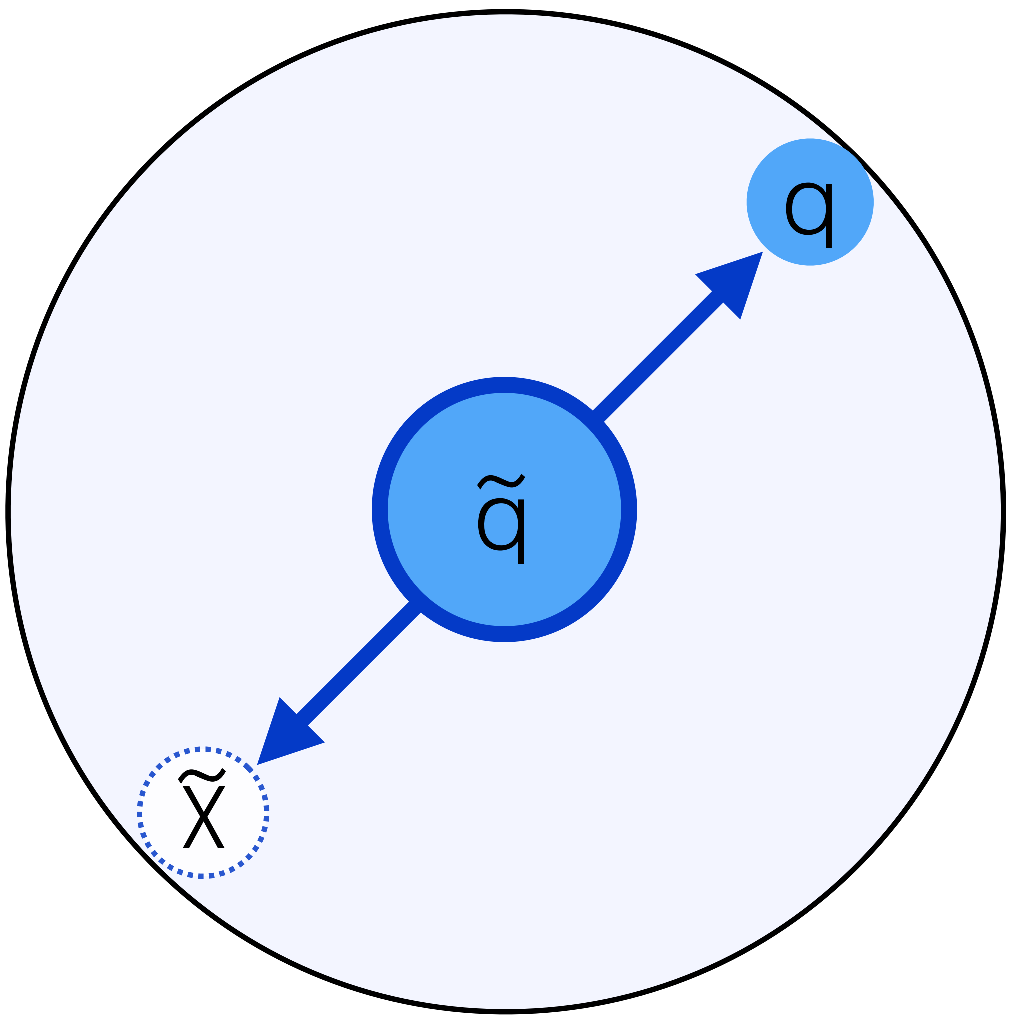
$$m_{\tilde{G}} \approx \langle F \rangle / M_{Pl} = M_{\cancel{SUSY}} / M_{Pl} \approx 10^{-3} eV$$

Searching for GMSB

- Because the gravitino is the LSP, the typically (as it is studied) stable LSP (typically a neutralino DM candidate) is promoted to being the NLSP leaving an additional decay open.
- In the case of a bino-like neutralino NLSP we expect the bino to predominately decay to gravitino + photon. Since both particles are effectively massless, we expect a **high p_t photon** and **missing energy**
- In the case of heavy squark/gluino pair production we expect cascading decays down to the NLSP generating **many jets**
- To leverage the large mass splittings required to reach the massless final state from heavy gluino/squark pair production we utilize the **razor variable framework**

The Razor Kinematic Variables

Squark Rest Frame



- The canonical razor is derived from pair production of squarks each decay to quark + LSP
- Examining the squarks rest frame the LSP and jet have equal momentum
- In the decay frame we see that the jet momentum is a measure of the scale of the new physics M_Δ

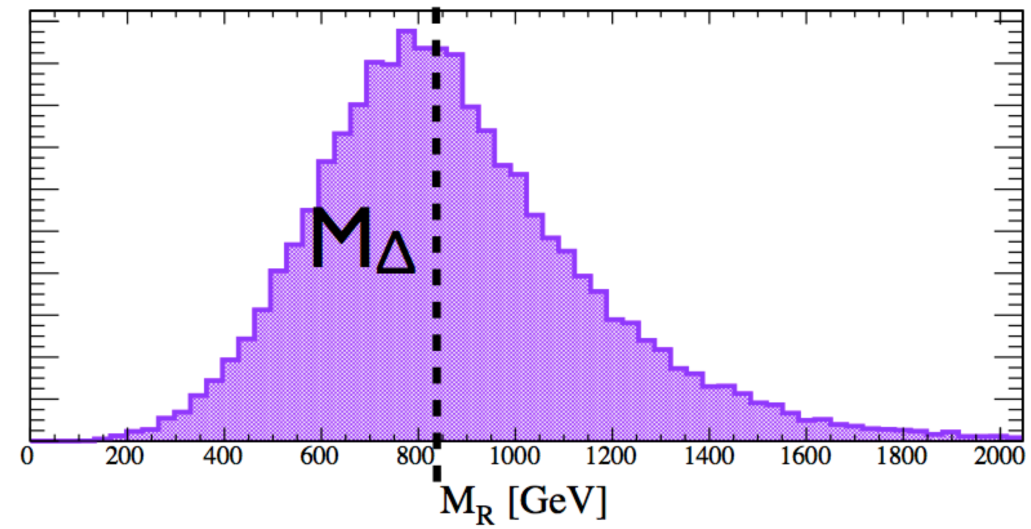
$$2|p_j| = M_\Delta = \frac{m_{\tilde{q}}^2 - m_{\tilde{\chi}^0}^2}{m_{\tilde{q}}}$$

- A series of approximations is made to reach this frame (the razor frame, R) and the jet momentum is evaluated

The Razor Variables

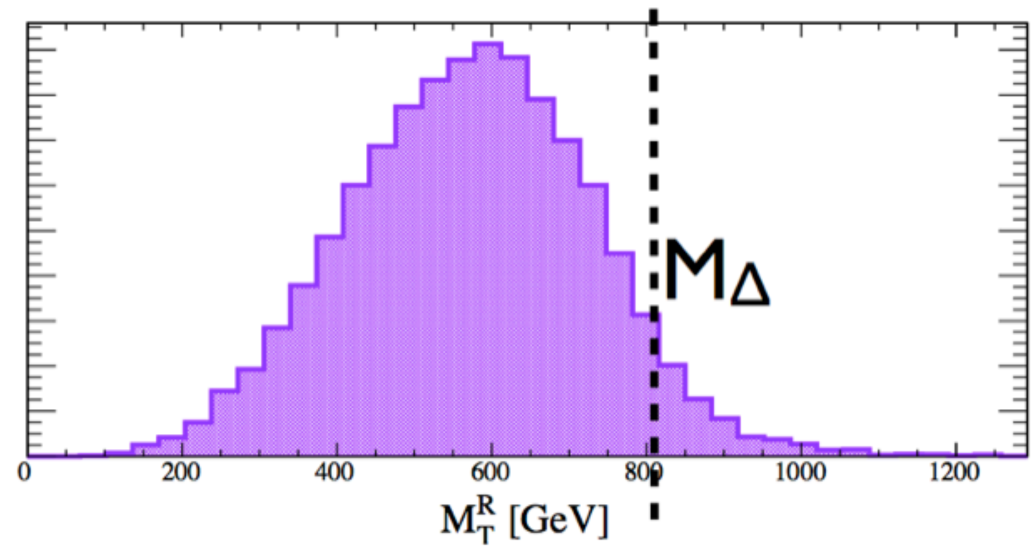
- M_R peaks at the mass scale M_Δ:

$$M_R = \sqrt{(|\vec{p}_{j_1}| + |\vec{p}_{j_2}|)^2 - (p_z^{j_1} + p_z^{j_2})^2}$$



- M_T^R is related to the missing energy and approximates M_Δ as an edge:

$$M_T^R = \sqrt{\frac{E_T^{miss}(p_T^{j_1} + p_T^{j_2}) - \vec{E}_T^{miss} \cdot (\vec{p}_T{}^{j_1} + \vec{p}_T{}^{j_2})}{2}}$$

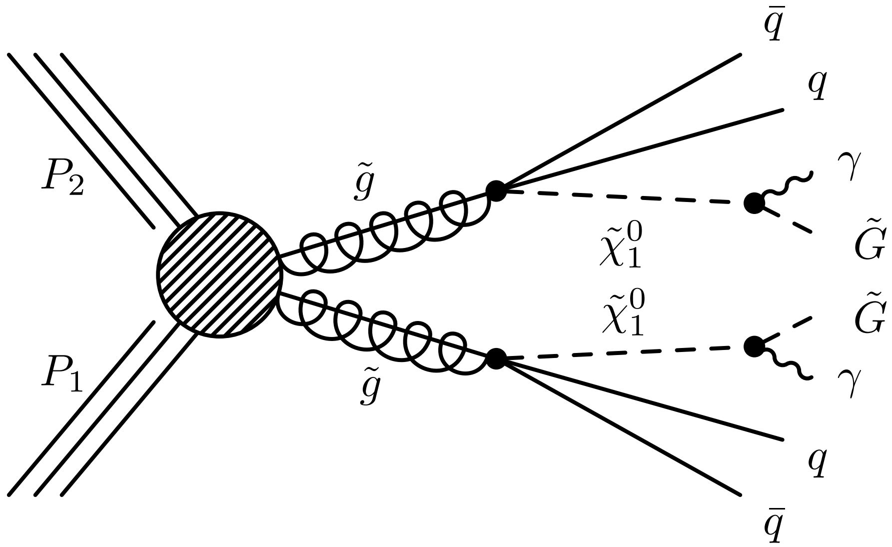


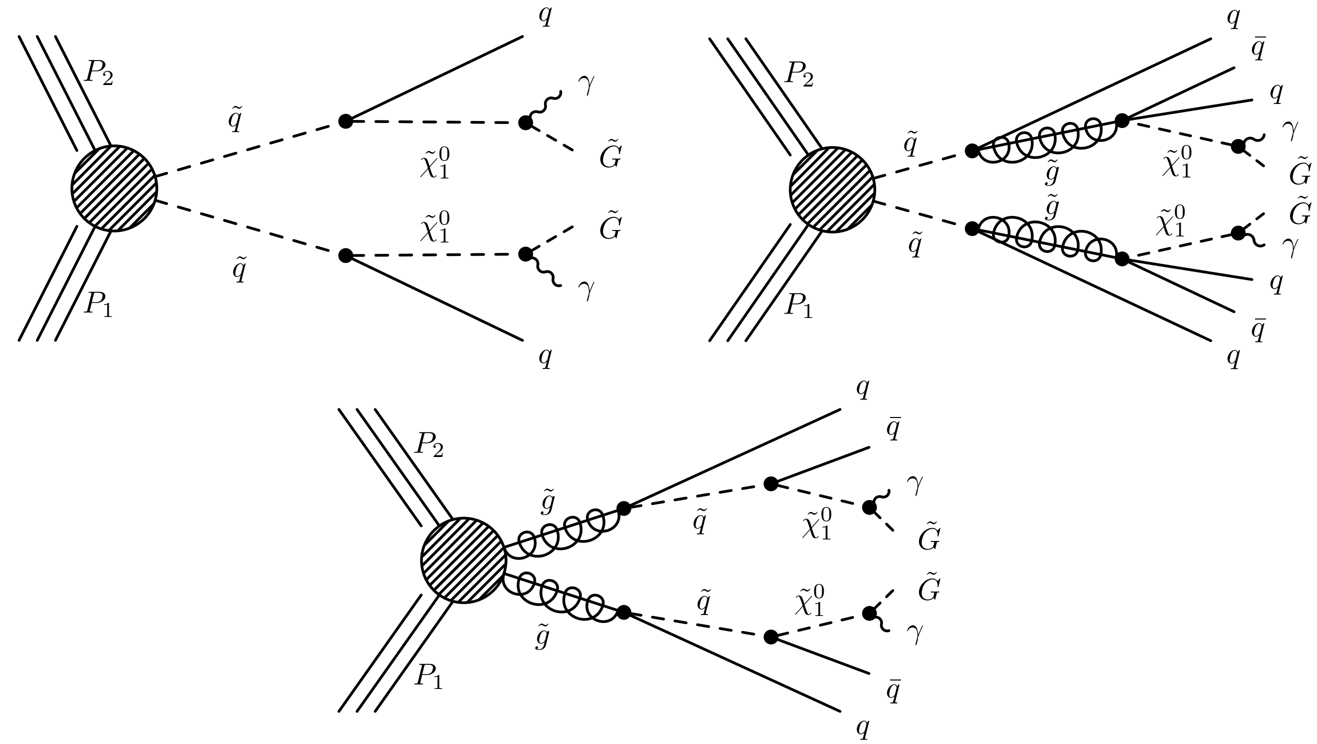
- R is their ratio: $R = \dfrac{M_T^R}{M_R}$

Analysis Strategy

1. Background shape is derived in a control region of R^2 with minimal signal contamination using a 1D fit in the variable M_R
2. The shape is then normalized to the signal region to derive our expectation in bins of $M_{R.}$
3. This procedure is checked for closure on a data driven control sample which is designed to represent the dominant backgrounds to the selection sample.
4. Interpret prediction as upper limits or the cross section of the signal process

Interpretations / Signal Samples

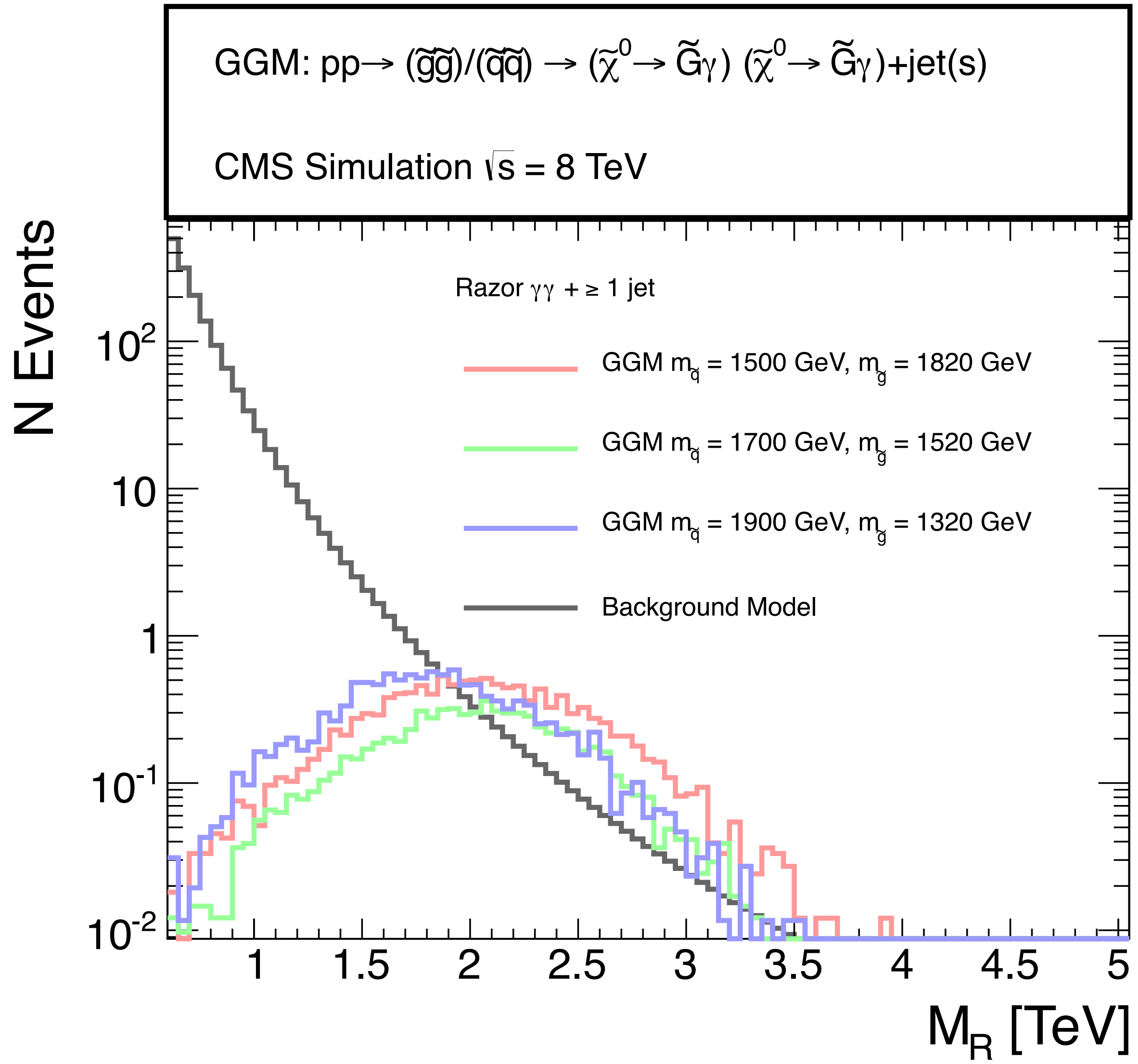


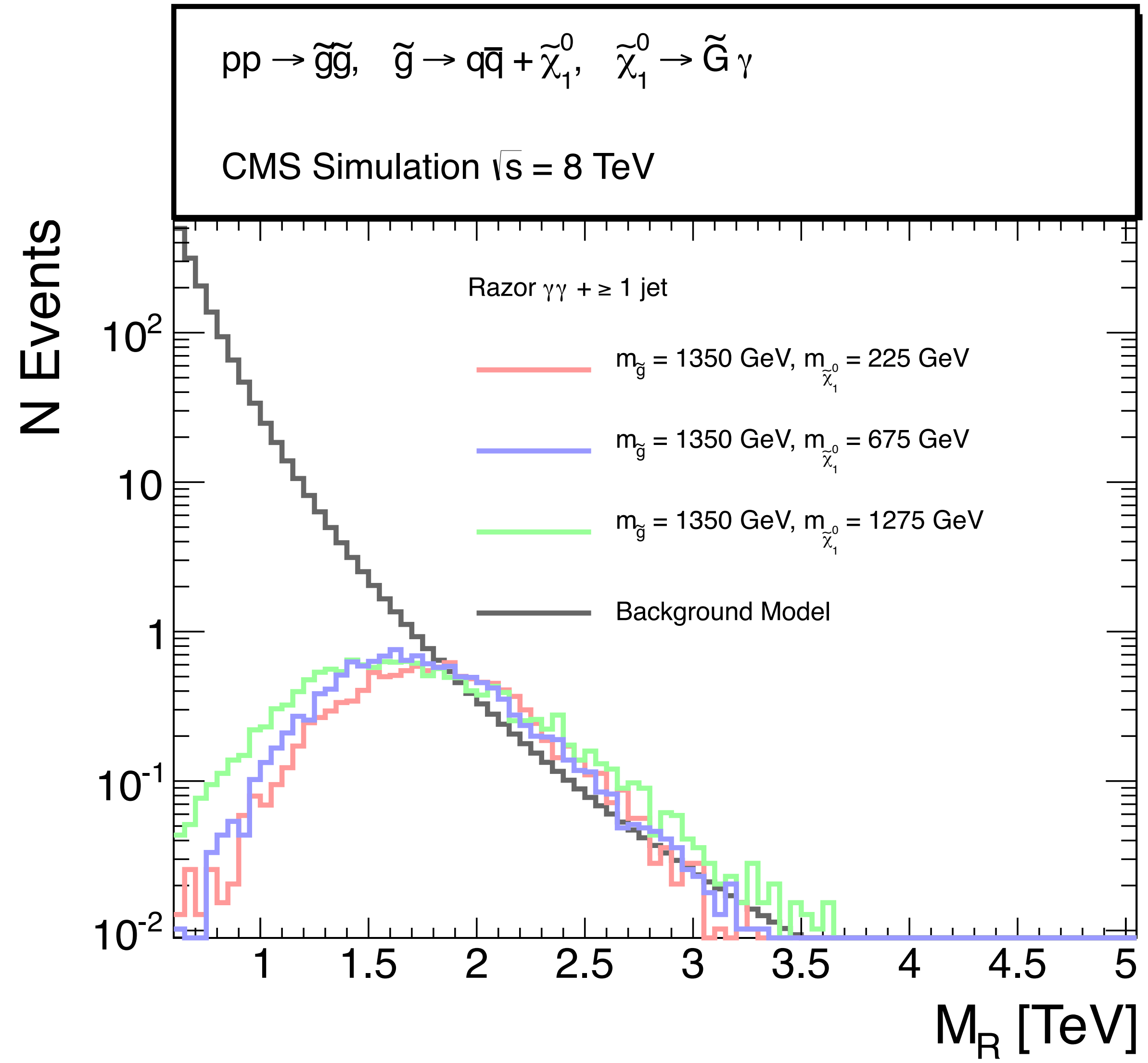


T5gg SMS (madgraph)
(gluino, bino-like neutralino)
800 ≦ m_g ≦ 1500 GeV
25 ≦ $m_χ$ ≦ m_g - 25 GeV
50 GeV steps

GGM Sample (pythia)
(gluino, squark)
1220 ≦ m_g ≦ 2020 GeV
1200 ≦ m_q ≦ 2000 GeV
$m_χ$ = 375 GeV (fixed)
100 GeV steps

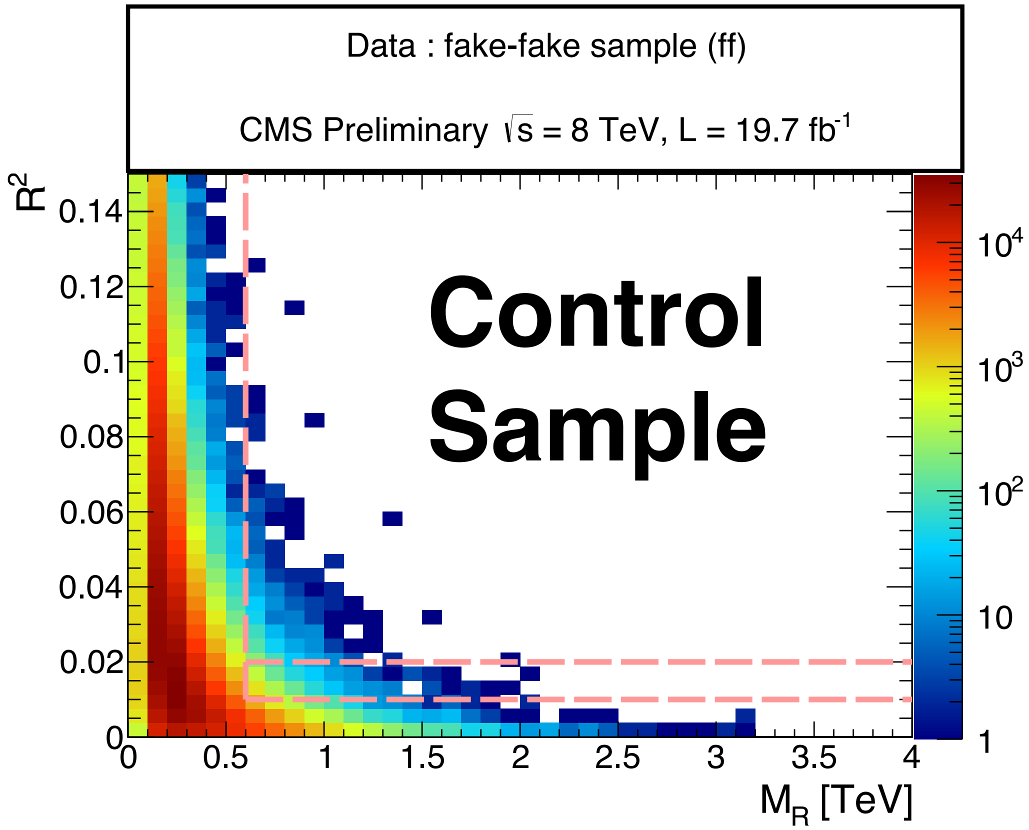
The Razor in diphoton GMSB

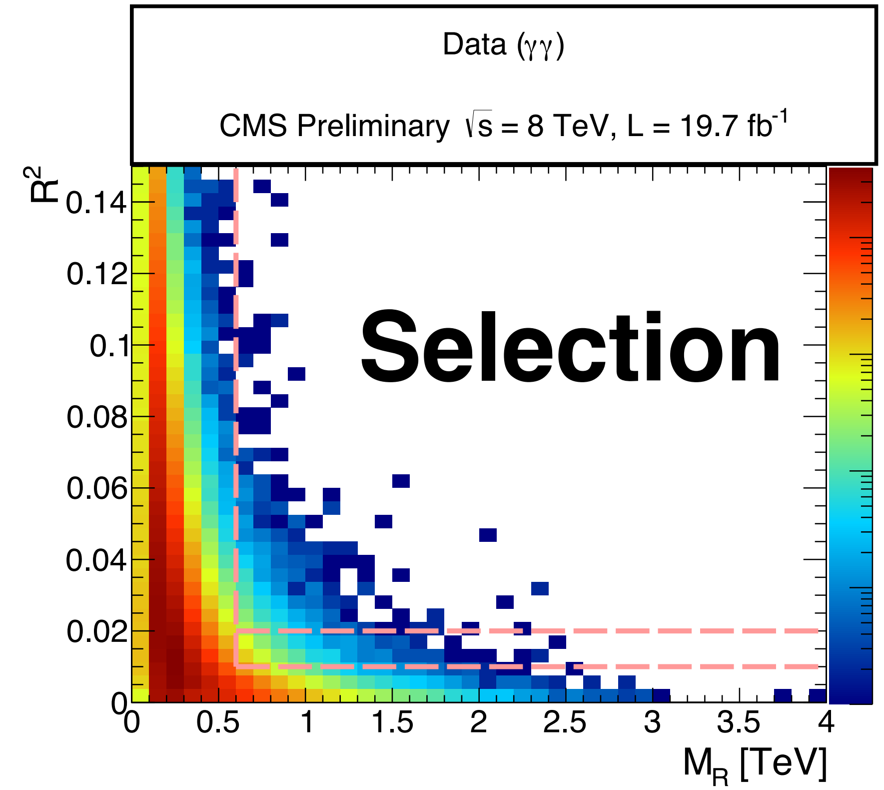


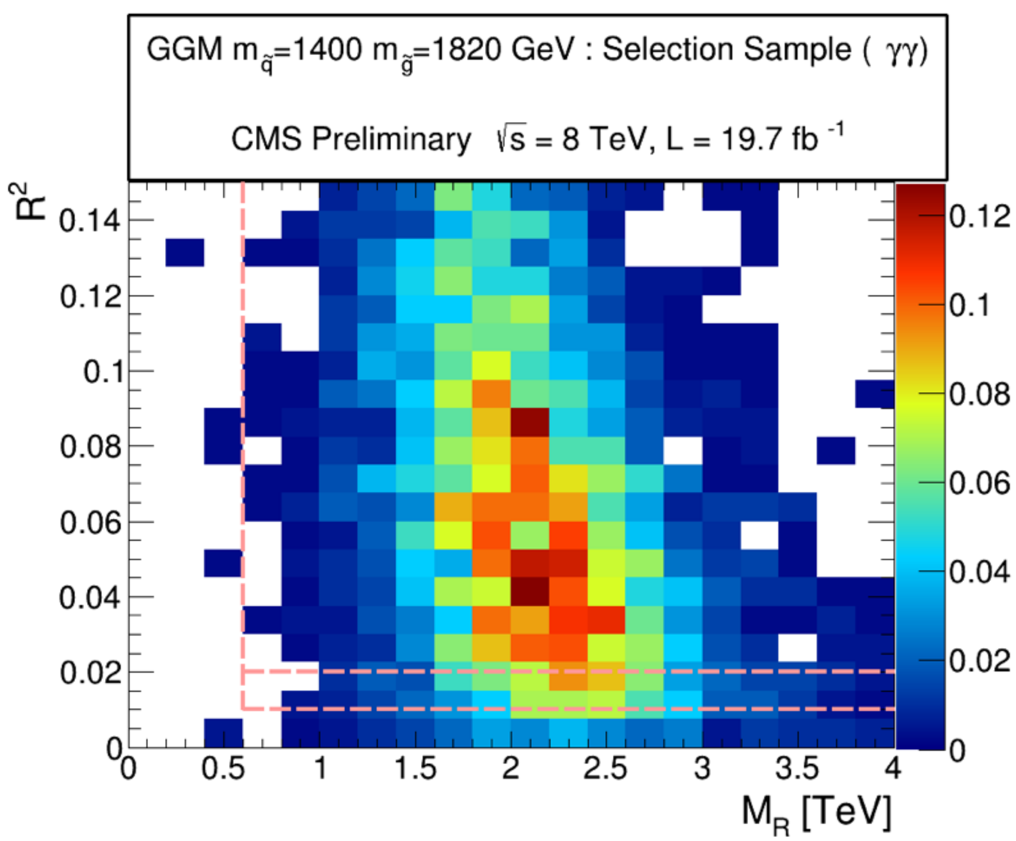


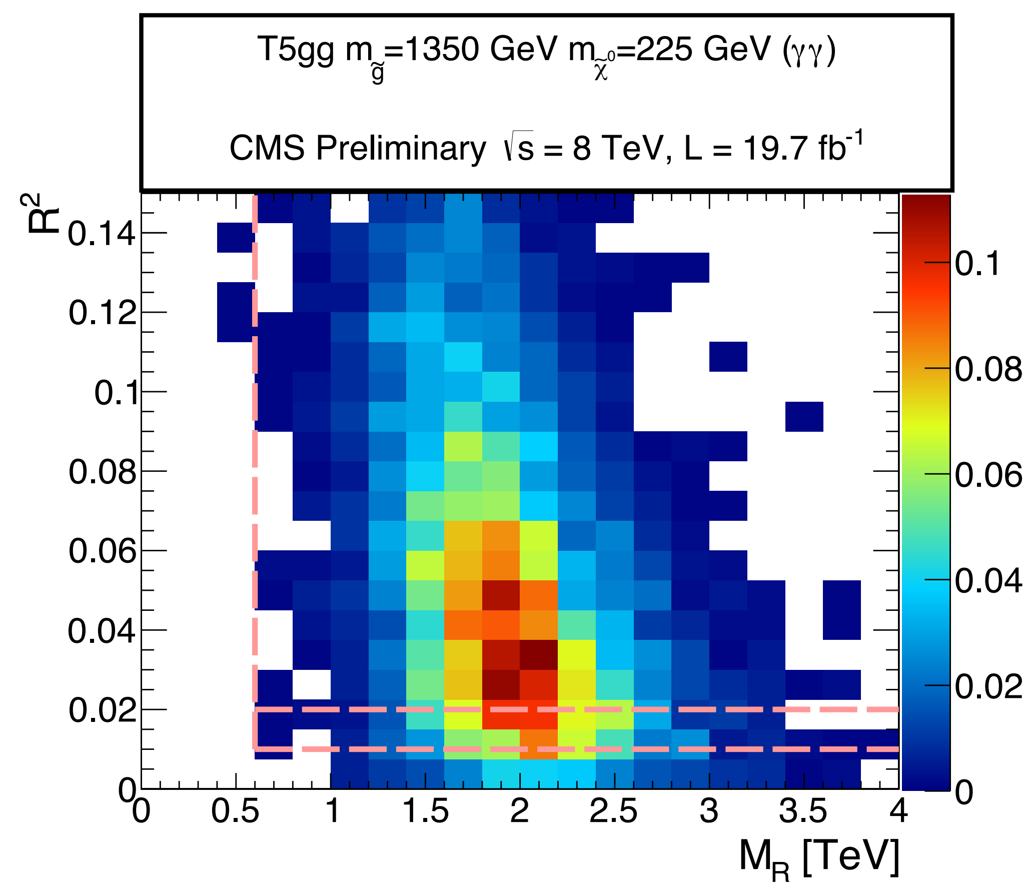
- The final decay to diphotons involves a number of mass splittings (and a mixture of different processes for the GGM sample)
- The GGM sample tends to peak at higher values of M_R than for the T5gg model.

Signal and Control Regions

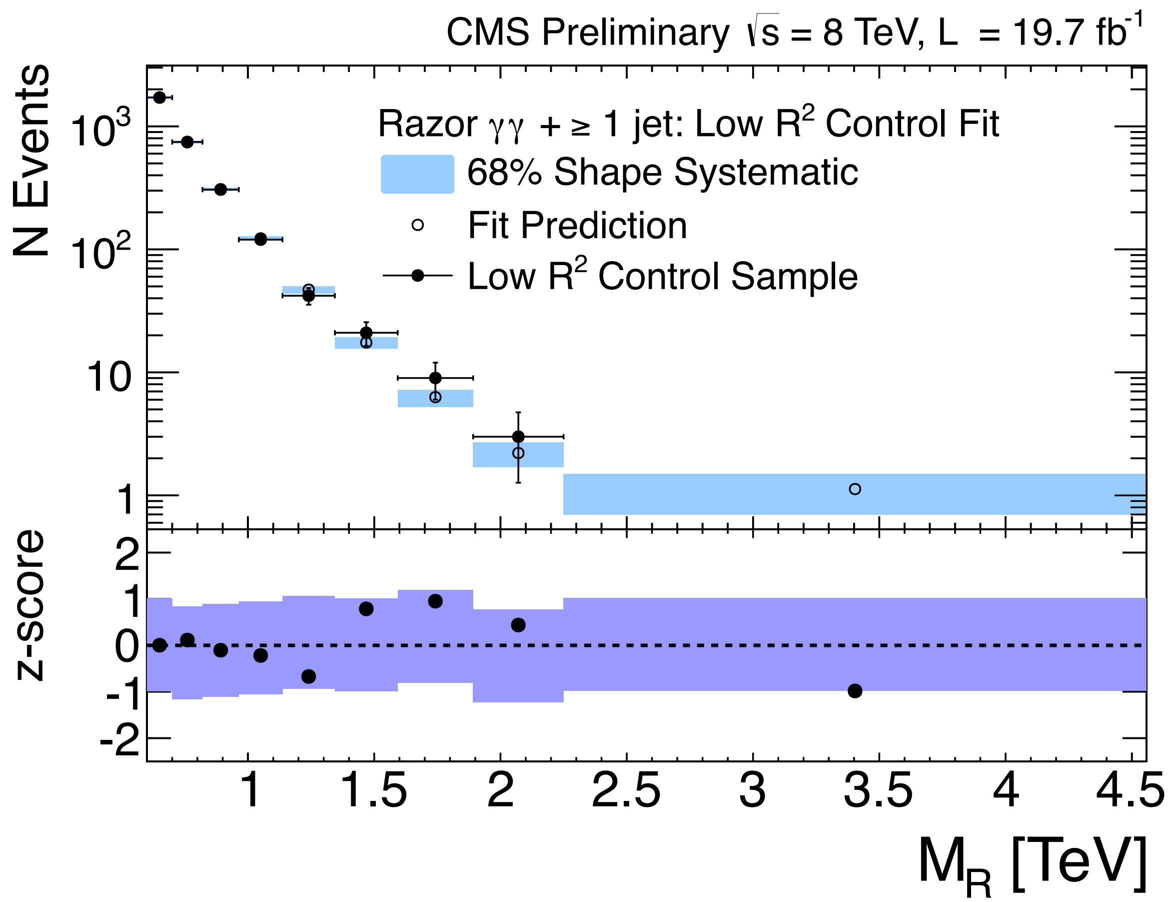


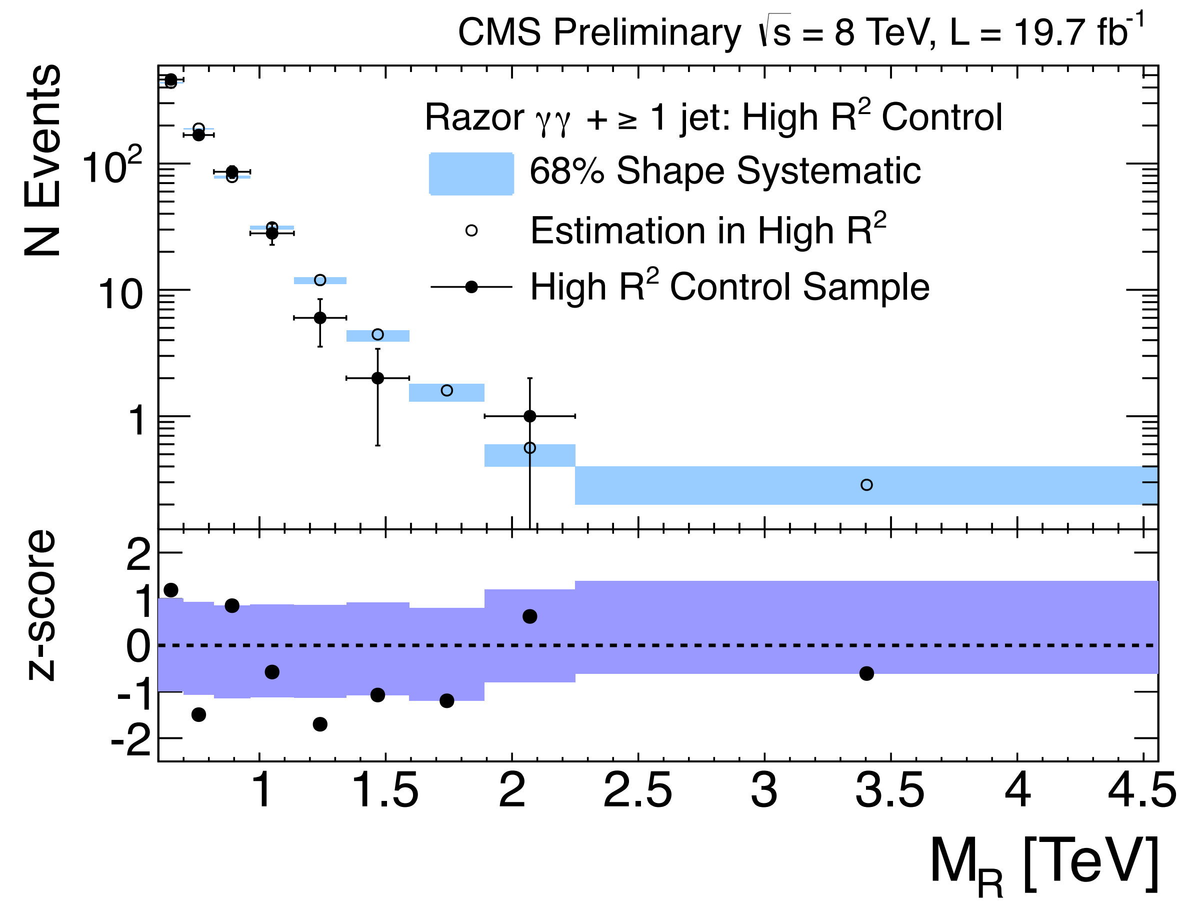






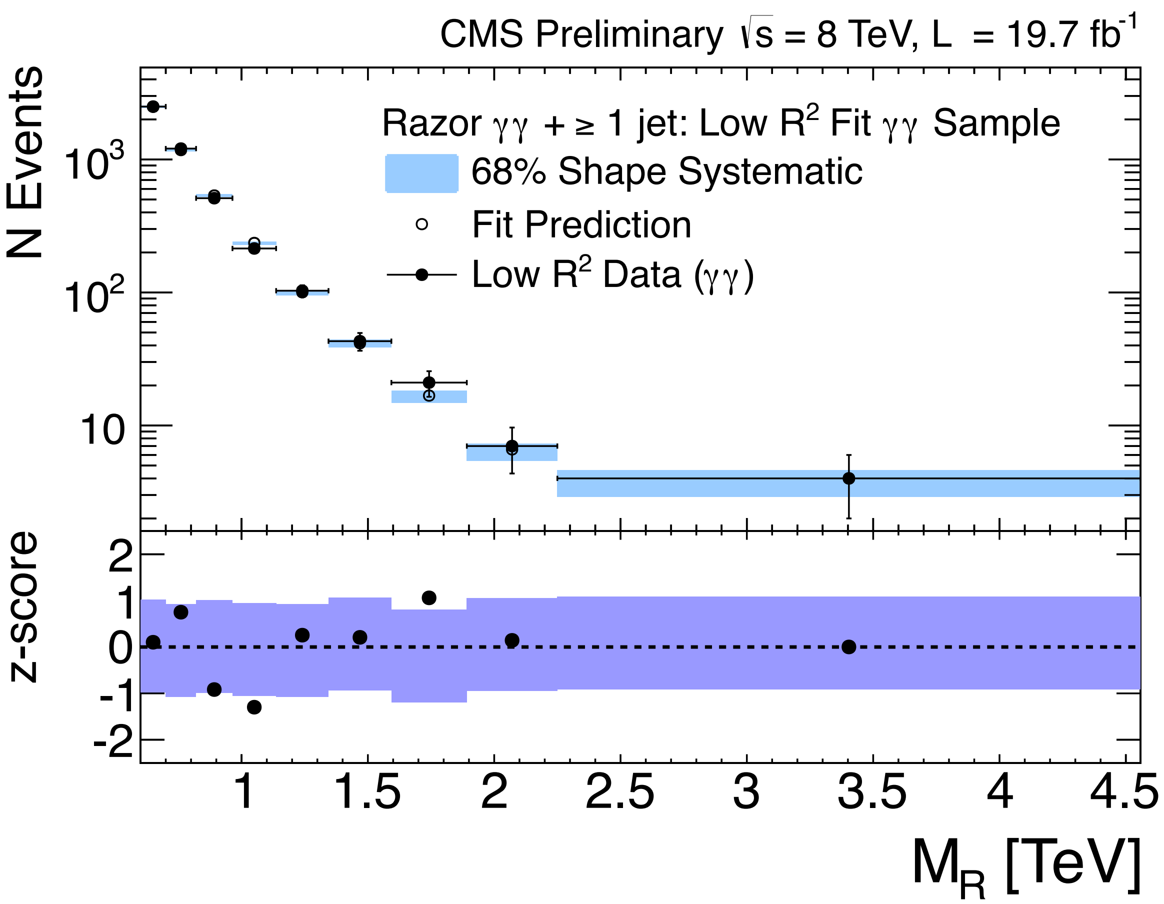
Control-Sample Closure Test

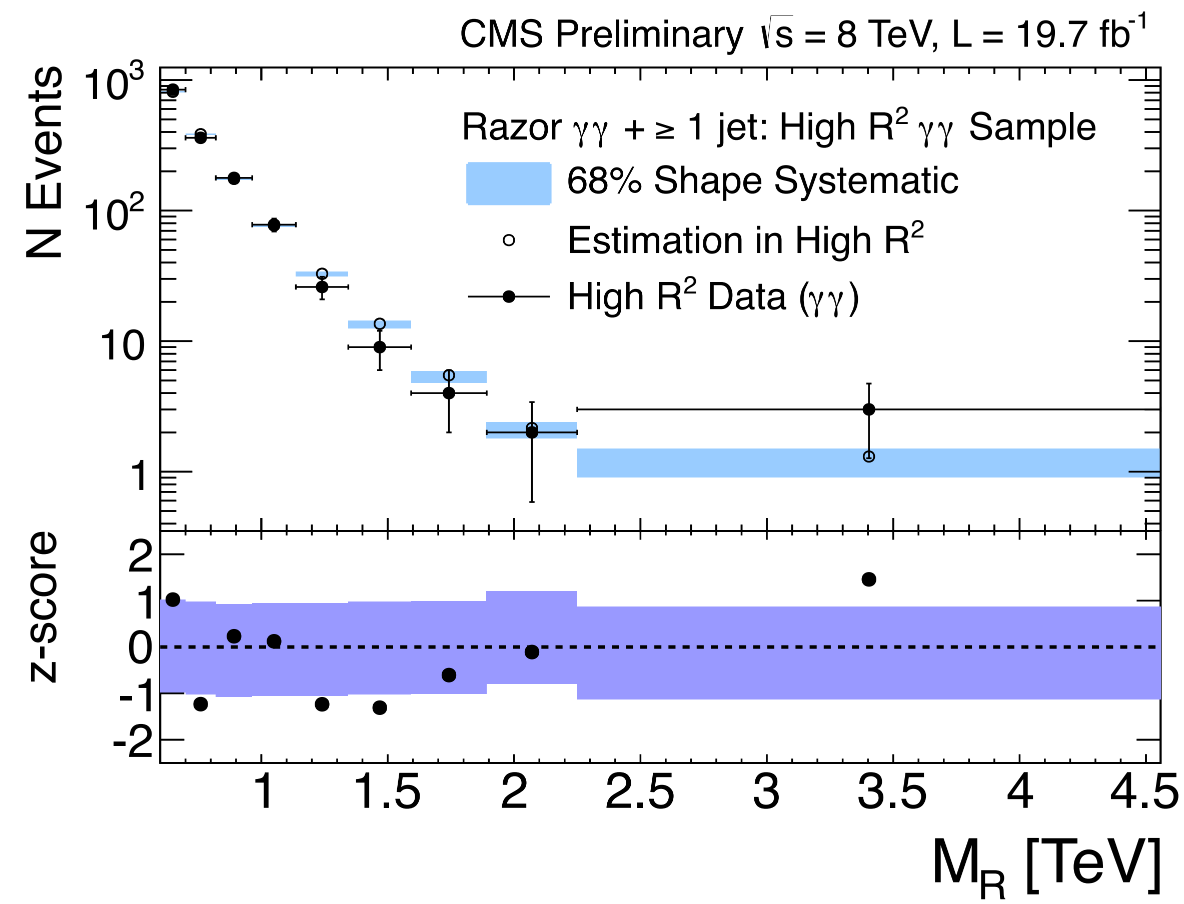




- Using the control sample fake-fake we perform the fit in the low Rsq region to derive the shape (left)
- The shape is the normalized to the yield in the signal region on the (right)
- The systematic on the closure is done bin by bin as the % difference between the observed and the expected in High Rsq

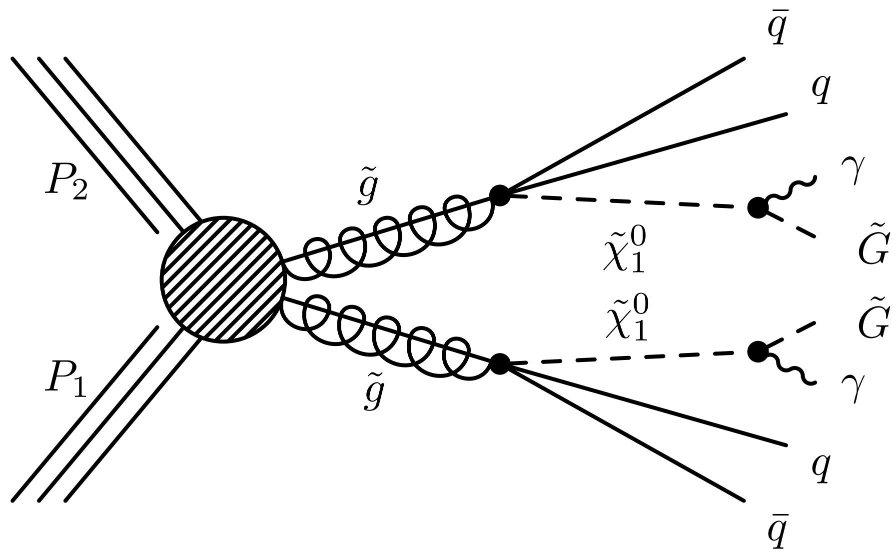
Selection Sample Fit and Extrapolation

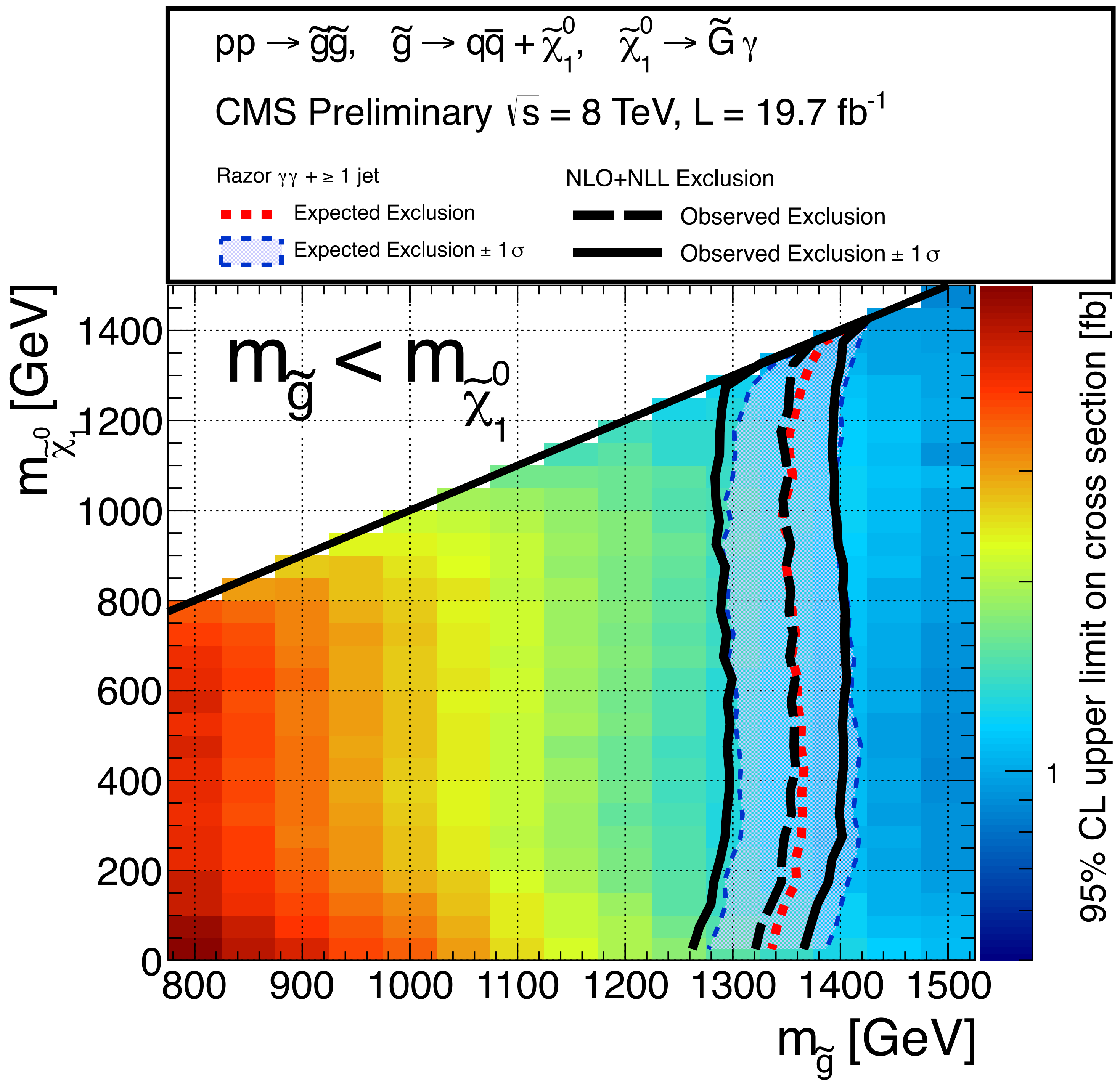




- As performed with the closure test, the low Rsq region is fit and the shape derived
- The shape is the normalized to the yield in the signal region on the (right)
- No significant excess is observed. The systematic overestimate is bin to bin is correlated due to shape systematic

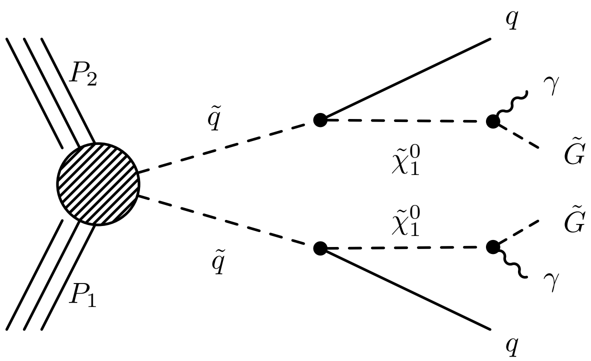
Limit T5gg

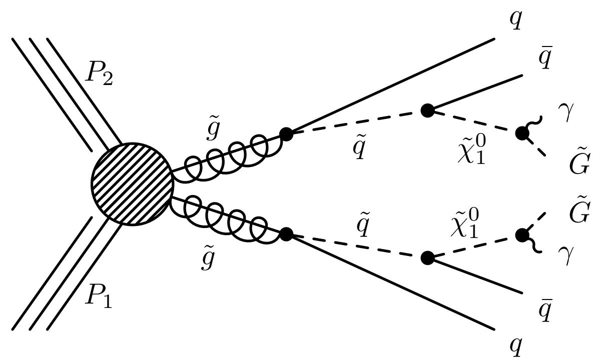


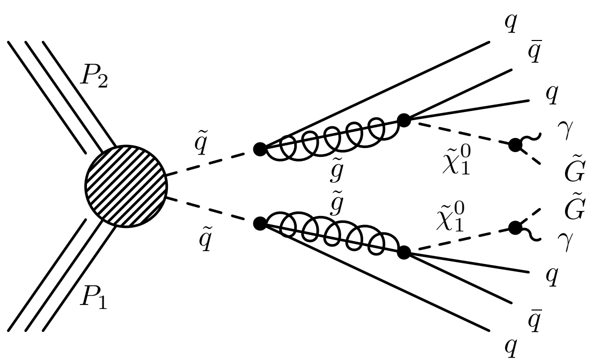


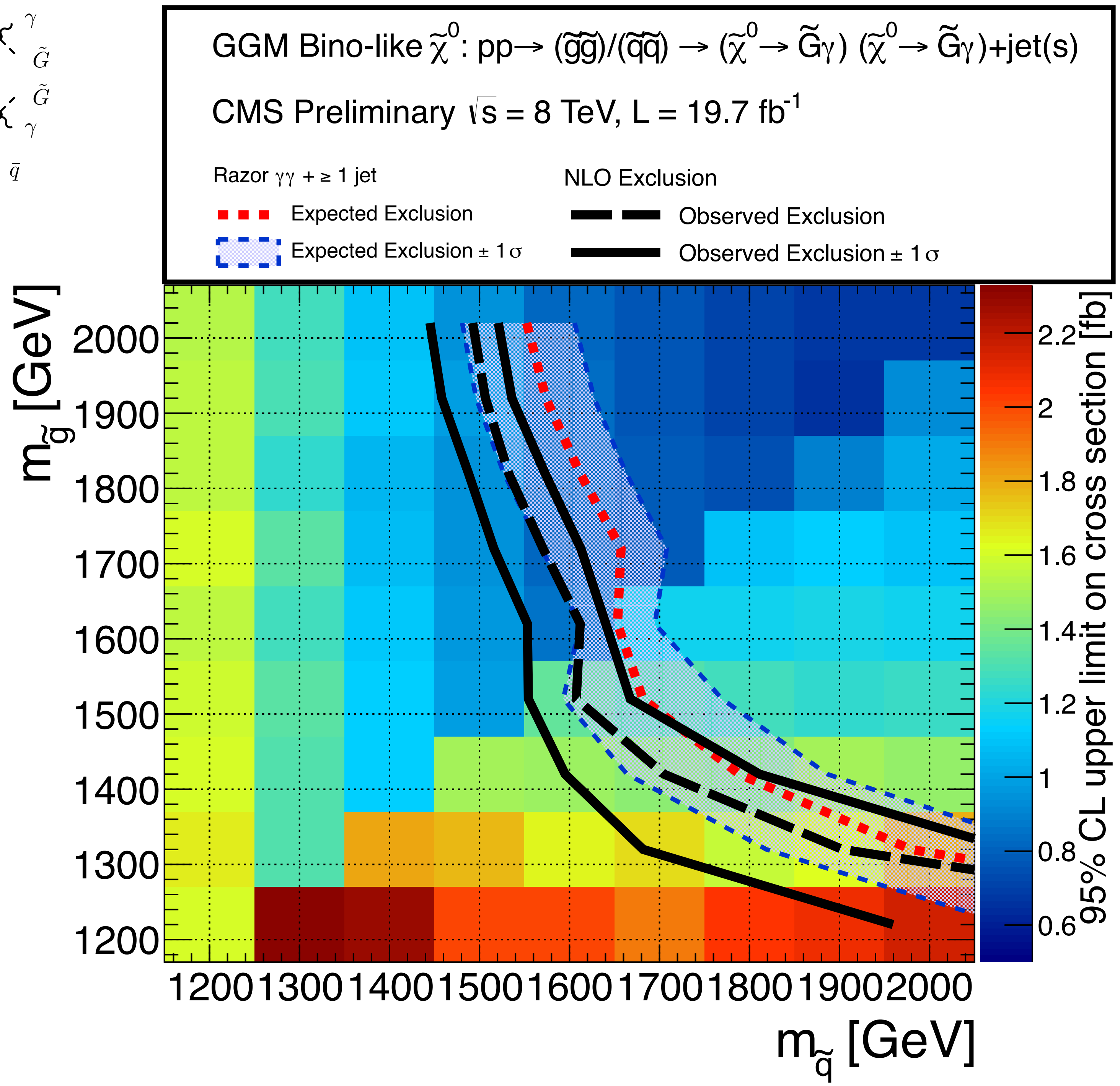
- We first compute the asymptotic limits and scan the signal strengths between the +/- 2 sigma asymptotic limits using 30k toy pseudo-experiments
- Good agreement between expected and observed bands

Limit GGM







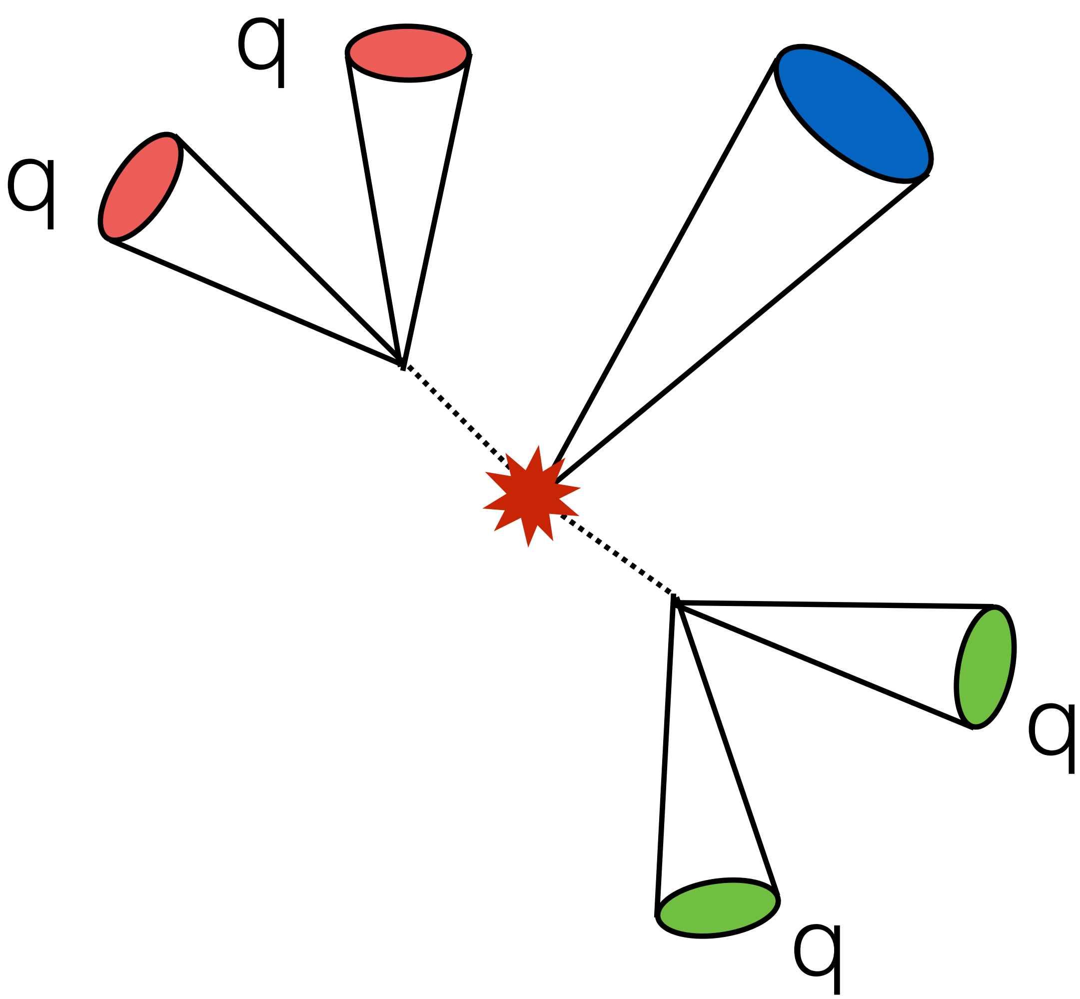


As the squark mass decouples ~2 TeV we see similar agreement with the T5gg limit

Conclusions

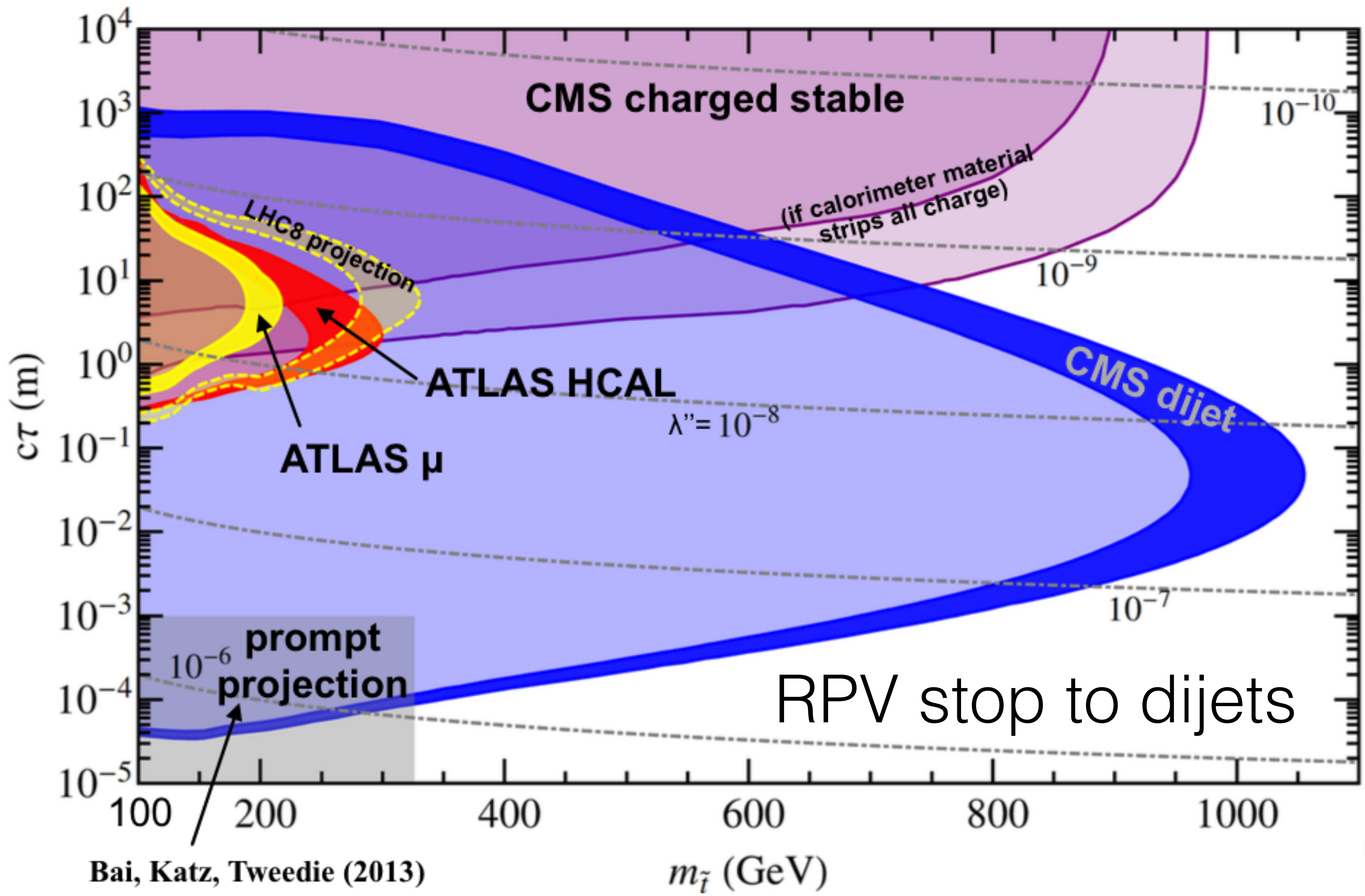
- We have performed a search for SUSY using 19.7 fb^{-1} of 8 TeV data inspired by heavy pair production with resulting decays to jets, photons, and missing energy
- Background estimation is performed using a kinematic fit to the razor variables with a data-driven closure check
- In the absence of an excess, the estimation is interpreted as limits for a simplified model T5gg as well as a previously interpreted squark-gluino model
- Results have been physics approved by CMS to be shown publicly with a paper in preparation for journal submission.

A Brief Discussion on Displaced Jets



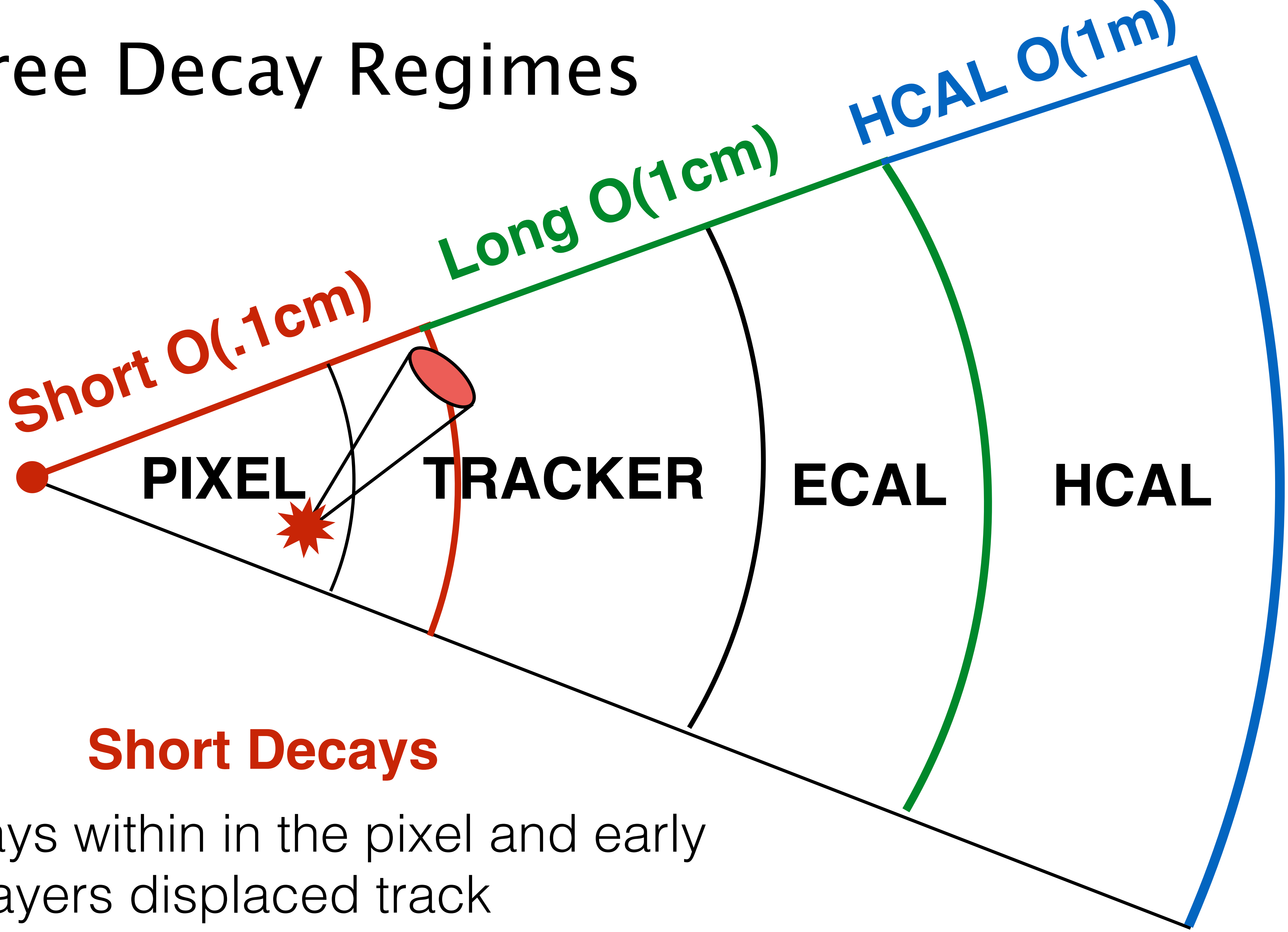
- An array of models predict physics with long lived particles. These particles decay measurably displaced from the collision point
  - RPV SUSY
  - GMSB
  - Hidden Valleys
  - Split SUSY
- As nearly all SM processes are prompt these signatures enjoy small backgrounds
- The discovery and lifetime measurement could would identify the new physics and relevant new scales

Recast CMS displaced dijet search



When the search is recast, the reach of the jet approach extends across large mass, lifetime, and topological space

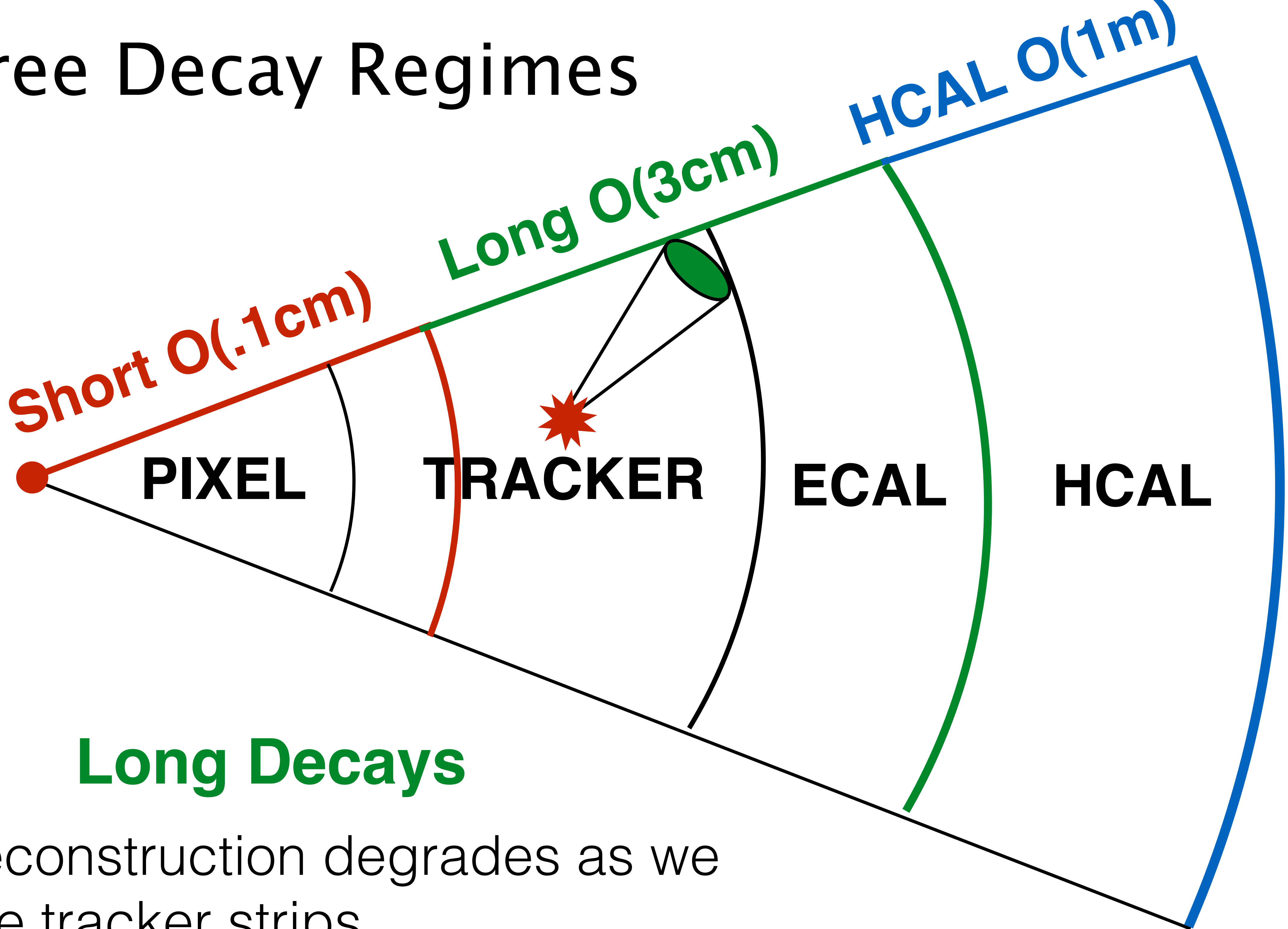
The Three Decay Regimes



Short Decays

- For decays within in the pixel and early tracker layers displaced track reconstruction is very efficient
- Generate secondary vertices similar to B-tagging algorithms

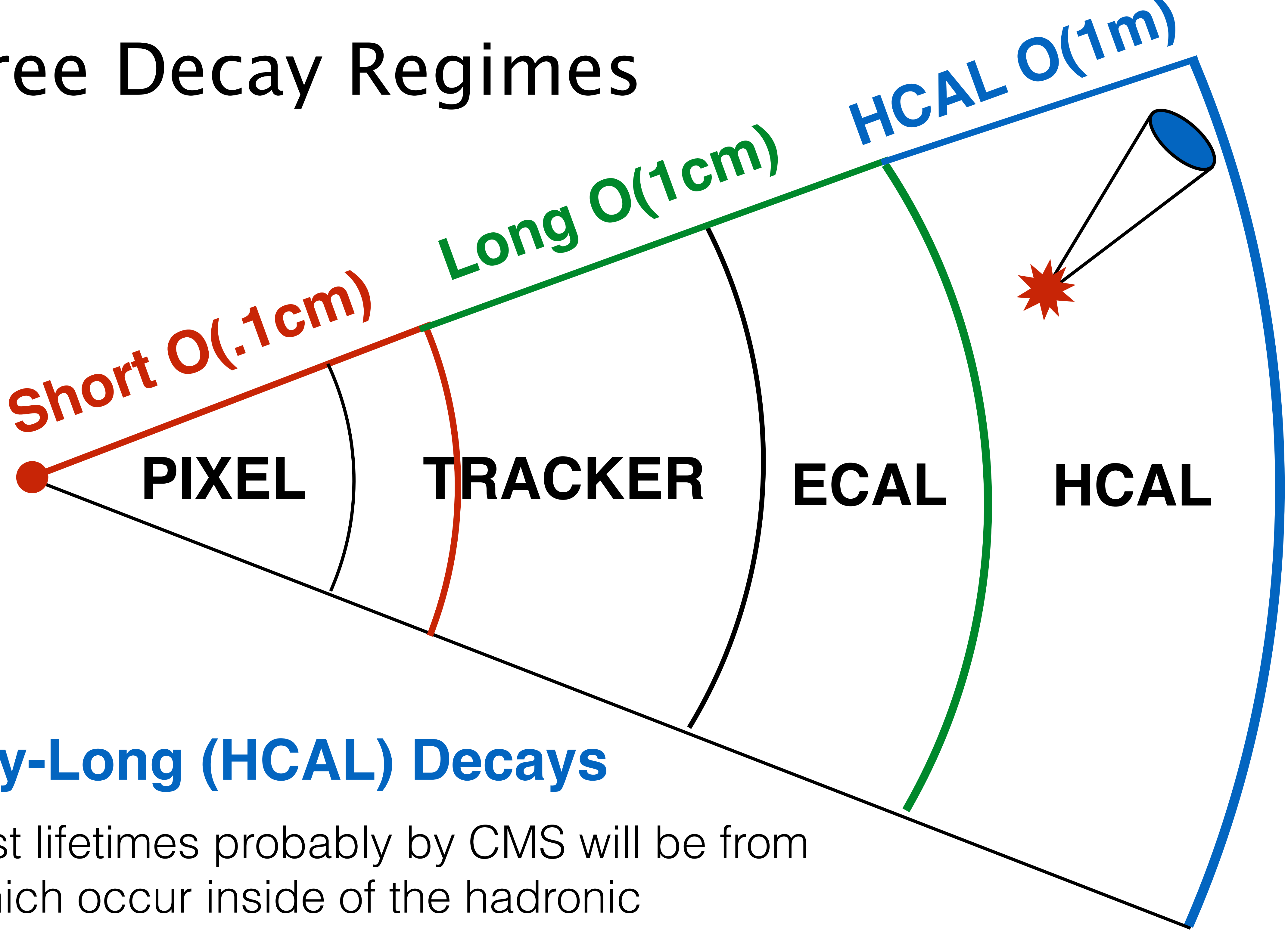
The Three Decay Regimes



Long Decays

- Track reconstruction degrades as we enter the tracker strips
- Utilize small prompt track contribution to jet energy

The Three Decay Regimes



Very-Long (HCAL) Decays

- The longest lifetimes probably by CMS will be from decays which occur inside of the hadronic calorimeter
- These jets have high hadronic energy fraction and no tracks pointing to the energy deposit

Displaced Dijet Triggers

HLT_HT650_DisplacedDijet80_Inclusive_v1

Inclusive Trigger: tight in HT with **loose tracking requirements** of the displaced jets. Few prompt tracks.

- High signal efficiency past the kinematic thresholds.
- Works well when lifetimes are long and tracks are not reconstructed well

HLT_HT350_DisplacedDijet80_DisplacedTrack_v1

Track Tagging Trigger: as low as the L1 turn-on in H_T, but **require displaced tracks**.

- Can target softer kinematics.
- Works best when the decay still occurs within the pixels layers.

Displaced Jet + VBF

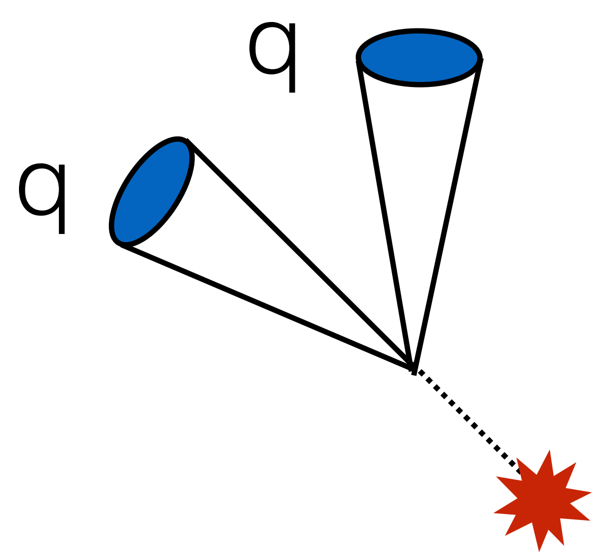
HLT_VBF_HadronJet40

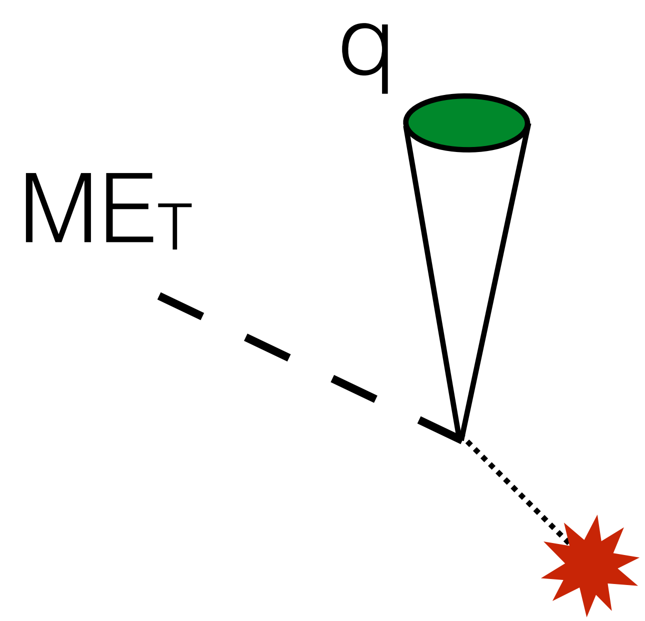
- **Long lifetimes** with decays inside the HCAL ctau ~1m
- **1 narrow calo jet** with **high hadronic energy** fraction
- Few prompt tracks

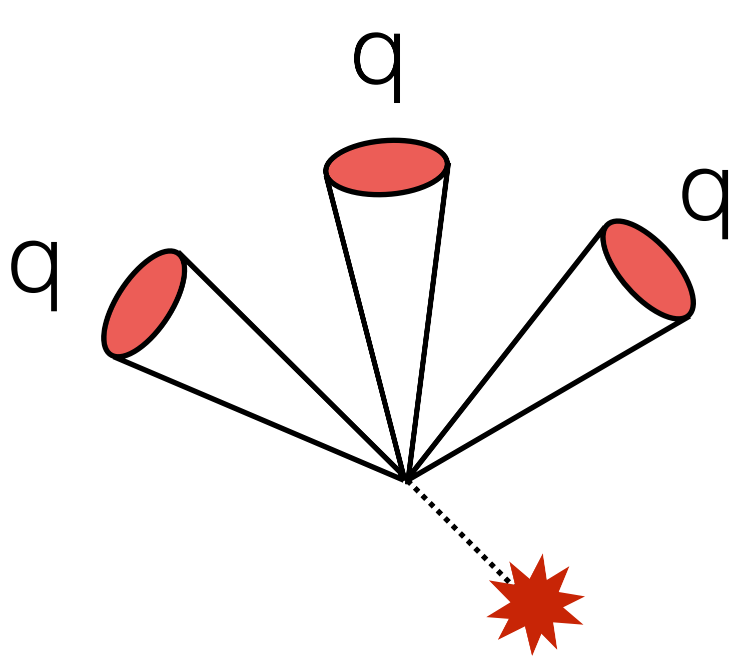
HLT_VBF_DisplacedJet40

- **Short lifetimes** ctau ~ 3 cm
- **1 calo jet** with **displaced tracks** (high IP significance)
- Few prompt tracks

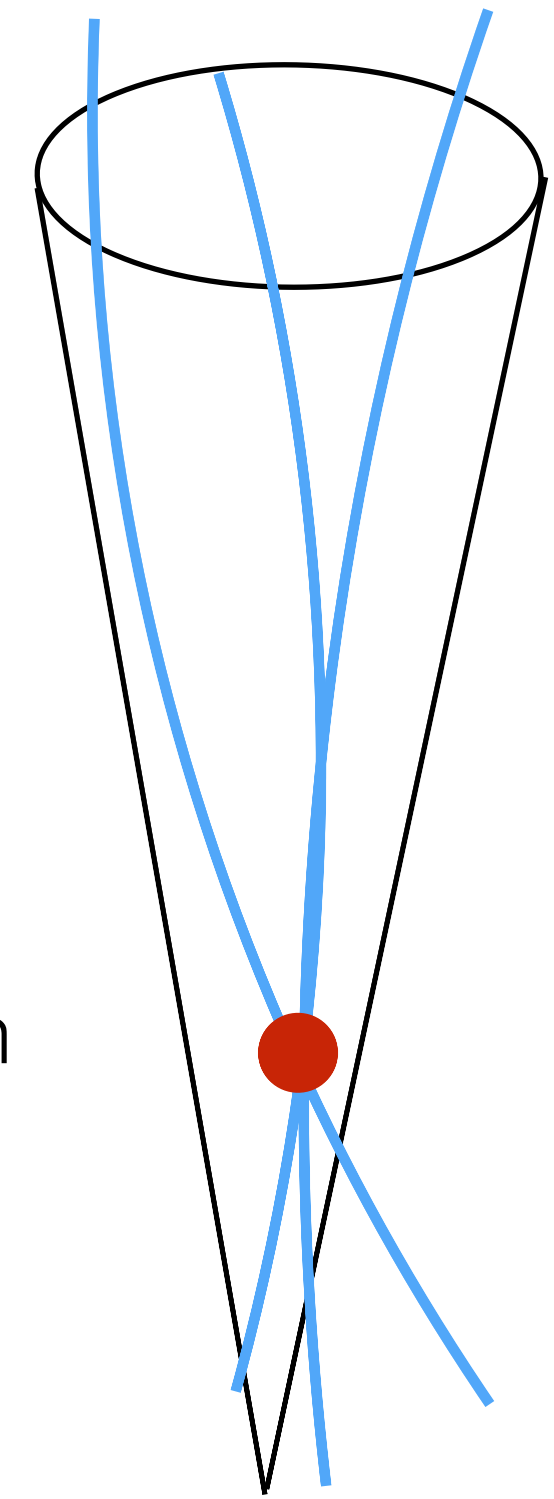
Combined… Combined… Secondary Vertex



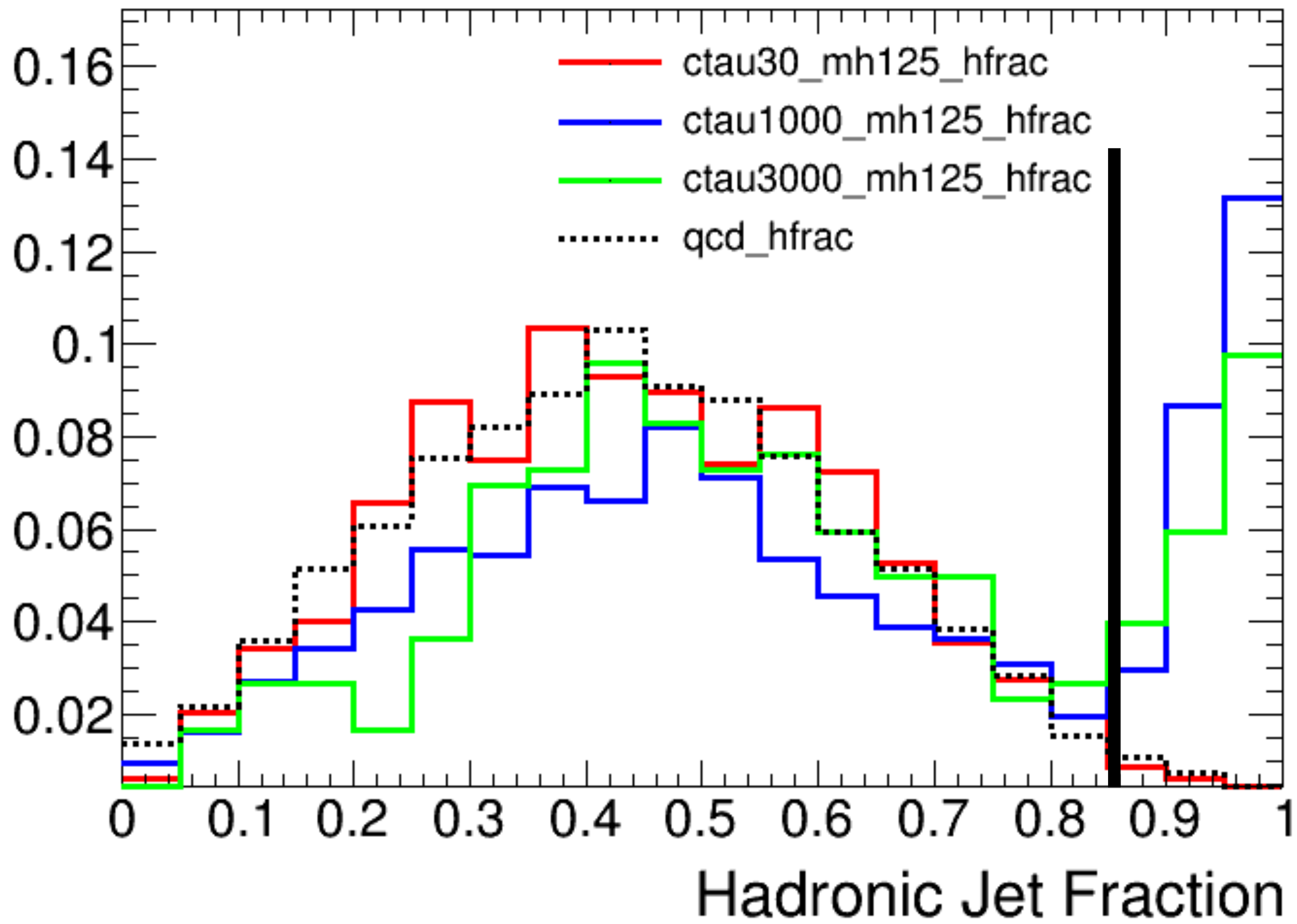




Utilize 3D clustering/isolation of secondary jet vertices to have an estimator of 1,2, and 3+ …. pronged long lived decays



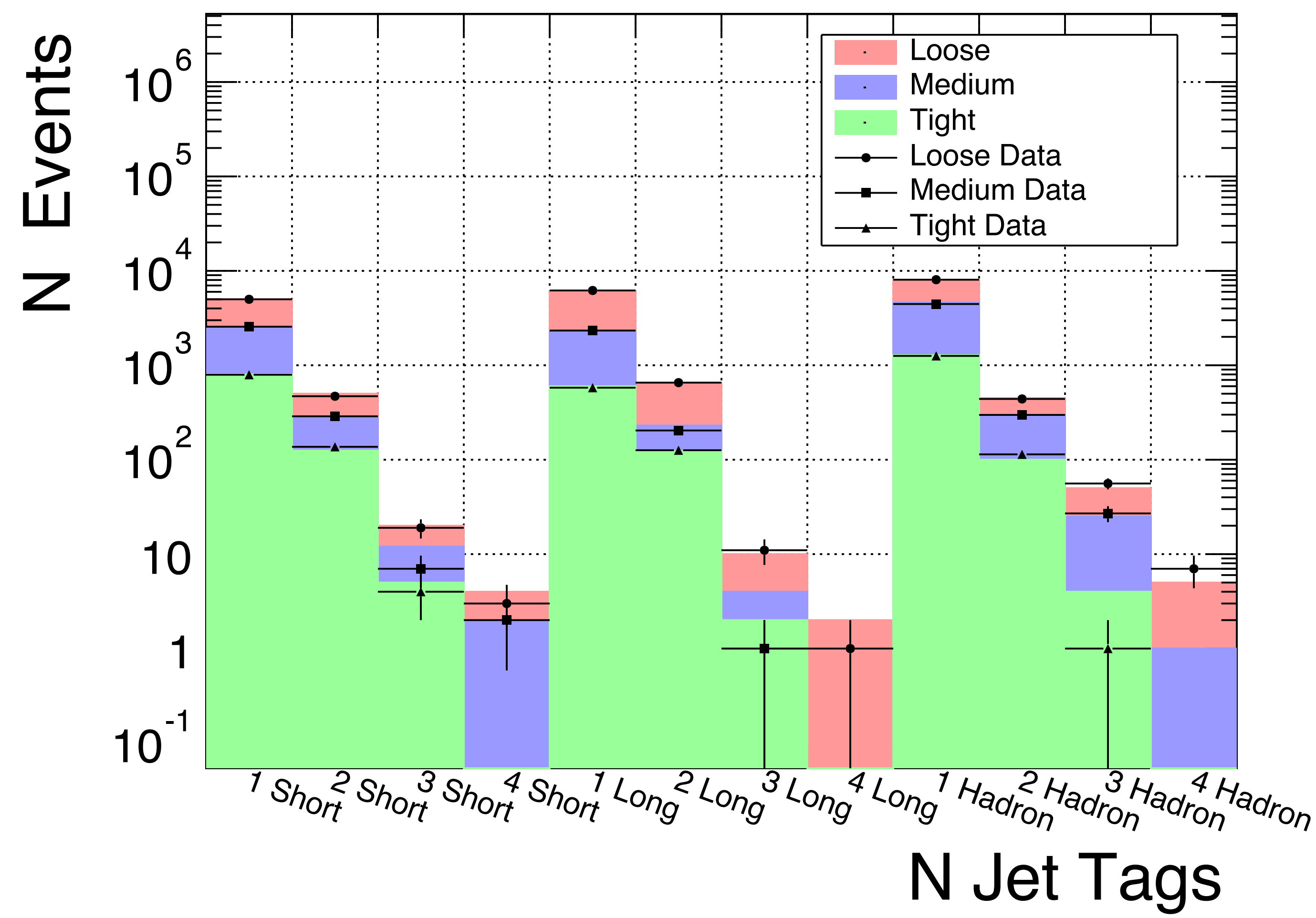
Very Long Decays



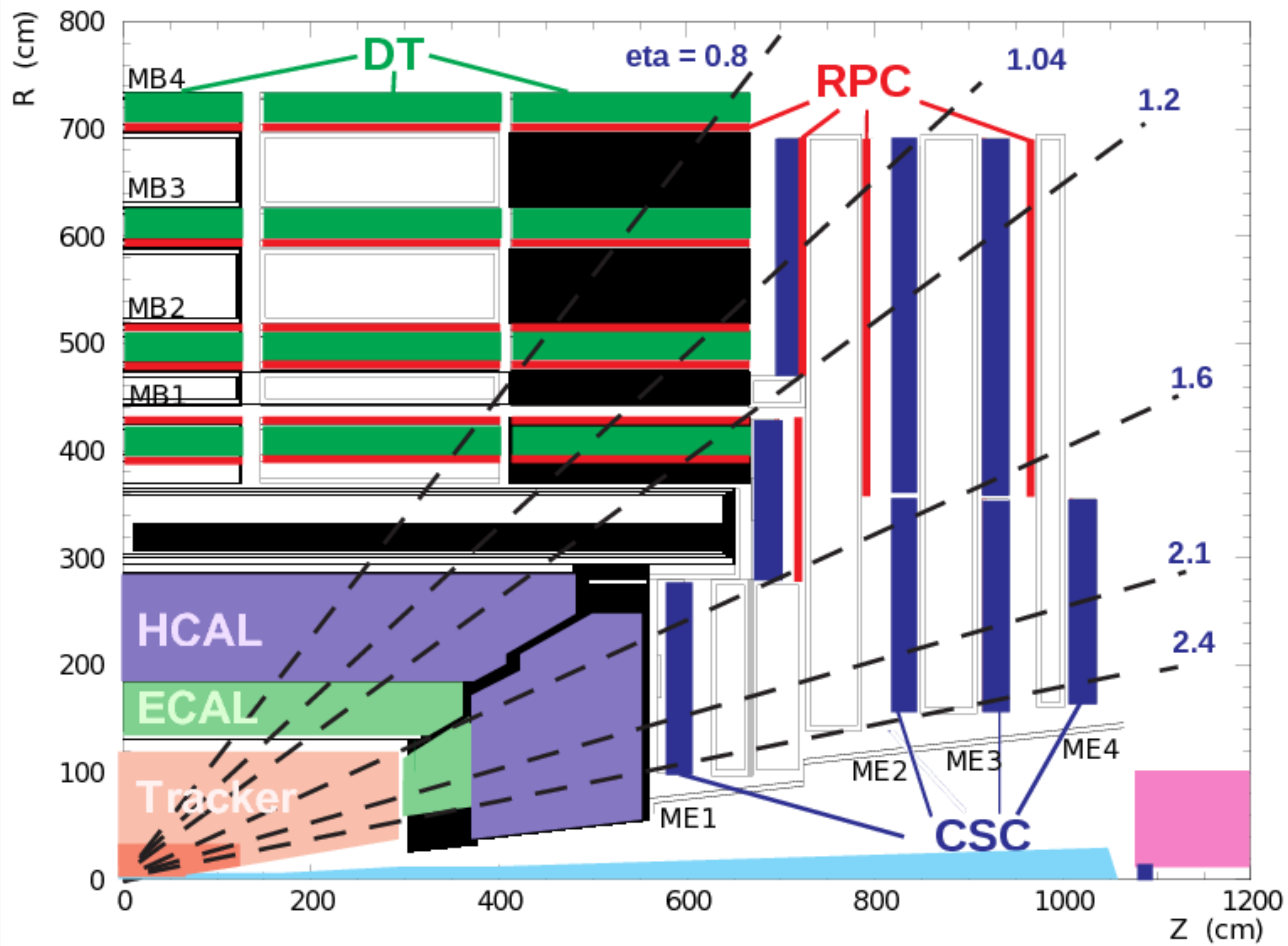
- Plotted are all central calo jets in events passing the L1 condition.
- Multiple signal lifetimes 30mm, 1000mm, and 3000mm
- Long lifetimes have high hadronic fraction from decaying inside the HCAL

The Big Picture / Conclusions

- Keep the analysis as inclusive as possible
- Develop displaced jet tags in the three decay regimes and performing counting experiments
- Use a background prediction minimally dependent on the trigger and run the study over multiple paths (with higgs associated production in mind)
- Plenty room for creativity and discovery!



BACKUP



R (cm)
800
700
600
500
400
300
200
100
0
0
200
400
600
800
1000
1200
Z (cm)
DT
MB4
MB3
MB2
MB1
eta = 0.8
1.04
1.2
1.6
2.1
2.4
RPC
HCAL
ECAL
Tracker
ME1
ME2
ME3
ME4
CSC

The Broken Mass Spectrum in the Messenger Sector

$$W \supset \lambda_{ij} \bar{\Phi}_i X \Phi_j \qquad\qquad \langle X \rangle = M + \theta^2 F$$

Expanding about the newly acquired VEV we have new mass terms for the scalar ϕ after collecting the D term of the kinetic term and F term of the potential term

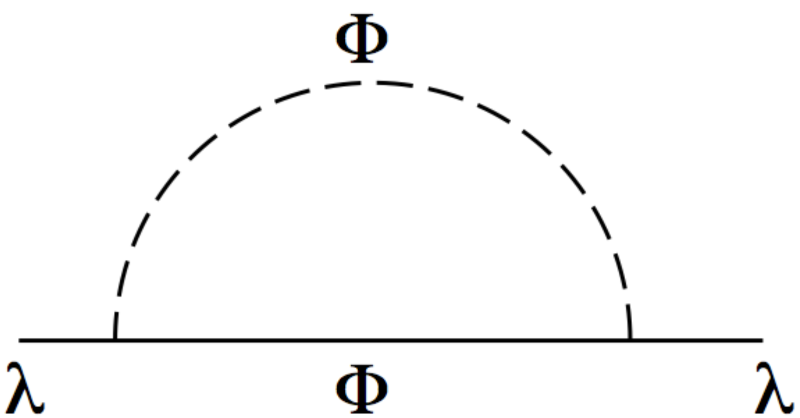
$$\mathcal{L} \supset \begin{pmatrix} \phi^\dagger & \bar{\phi} \end{pmatrix} \begin{pmatrix} |\lambda M|^2 & -(\lambda F)^\dagger \\ -\lambda F & |\lambda M|^2 \end{pmatrix} \begin{pmatrix} \phi \\ \bar{\phi}^\dagger \end{pmatrix}$$

The eigenvalues give us the broken mass spectrum in the messenger spectrum and F is a measure of how broken:

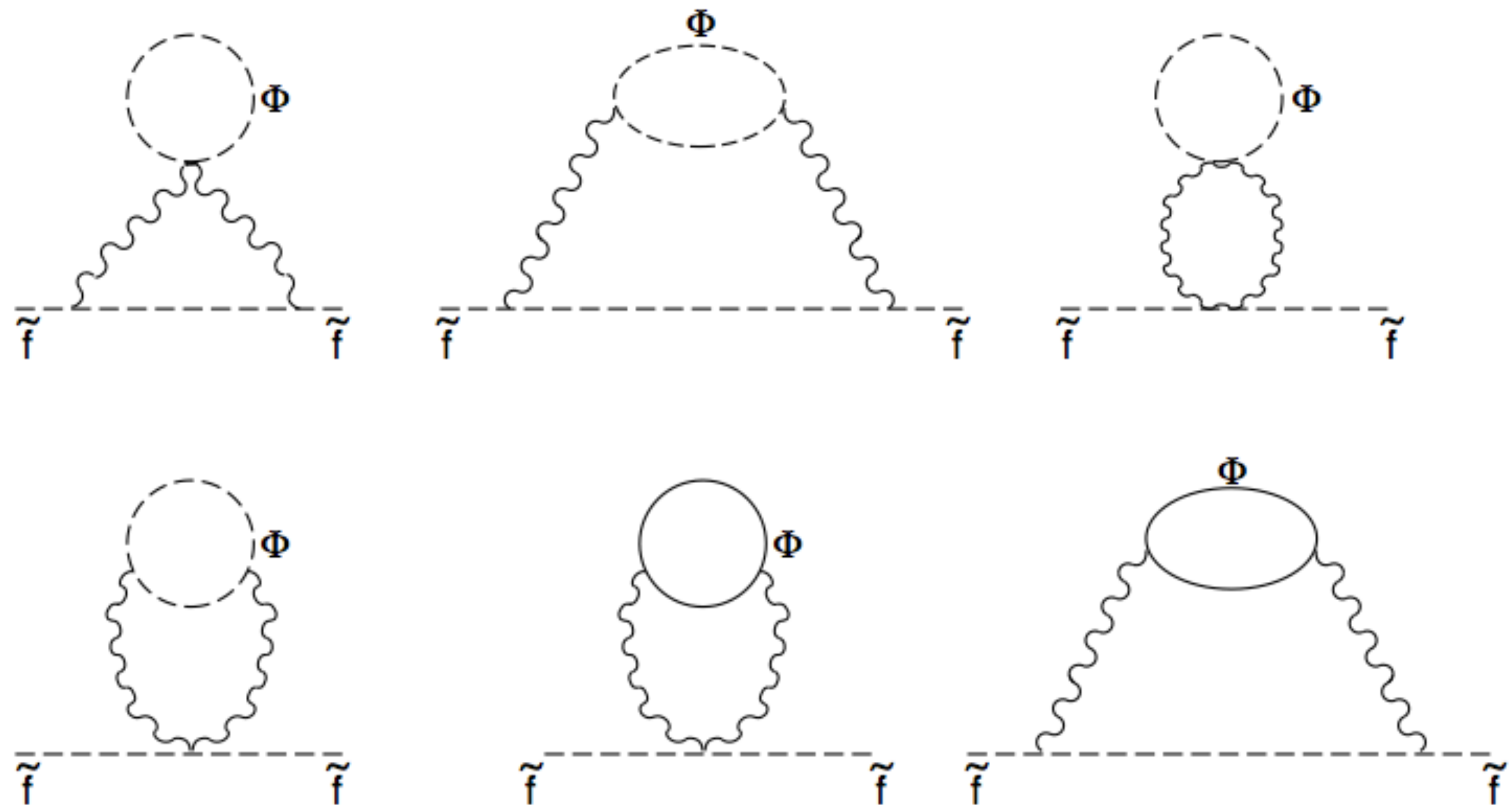
$$m^2_{\phi_\Phi} = |\lambda M|^2 \pm |\lambda F|$$

MSSM Mass Generation

gaugino masses



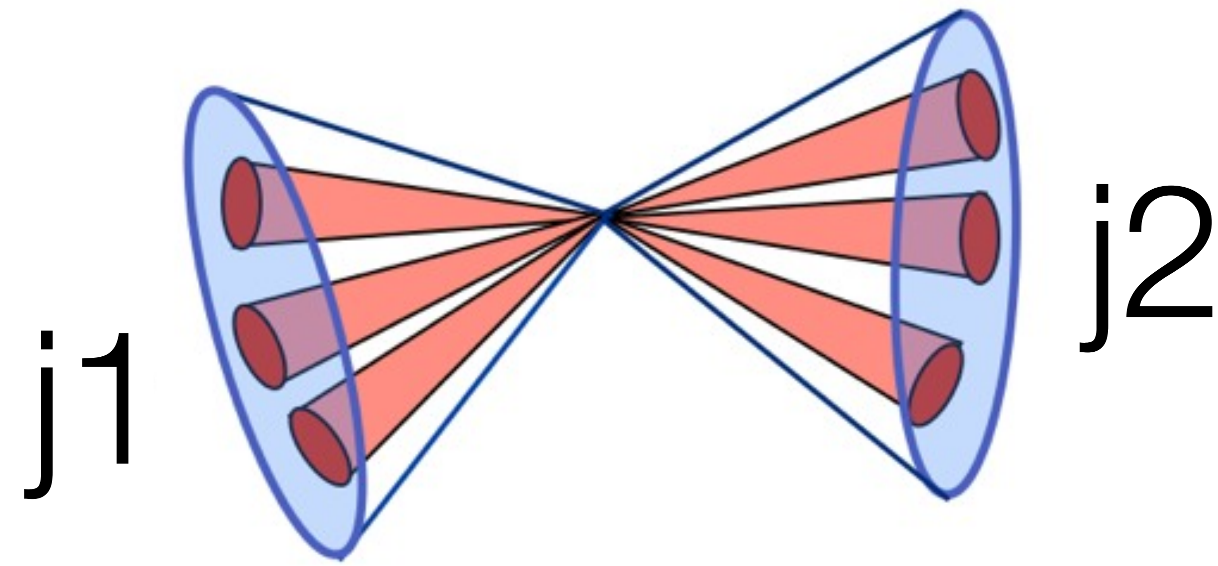
sfermion masses



- Masses in the MSSM are generated via gauge interactions between the messenger sector
- sfermion masses are generated at two loops
- gaugino masses are generated at one loop

Razor Variables: Hemisphere Formation

- Since there are usually more than 2 objects in an event we must force a di-jet topology to evaluate the razor variables
- The collection of objects are partitioned into two hemispheres such that the sum of the two hemisphere's invariant mass is minimized



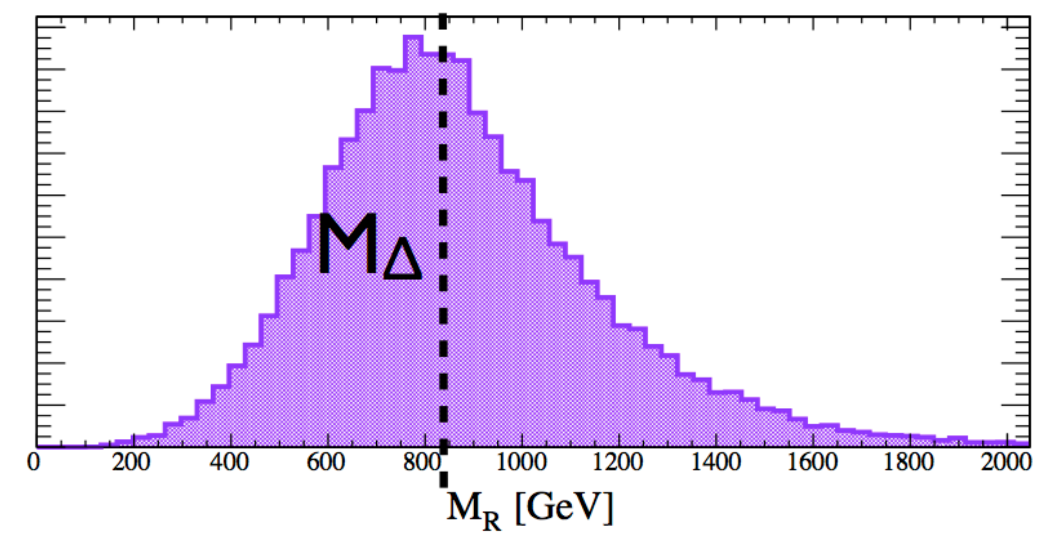
$$\mathrm{M_R} \; = \; \sqrt{(|\vec{\mathrm{p}}_{j_1}| + |\vec{\mathrm{p}}_{j_2}|)^2 - (\mathrm{p}_z^{j_1} + \mathrm{p}_z^{j_2})^2}$$

- This allows the kinematic variables to be calculated for any pair produced topology with a visible final state by interpreting the hemispheres are the visible component of two pair produced heavy particle decays

The Razor Variables

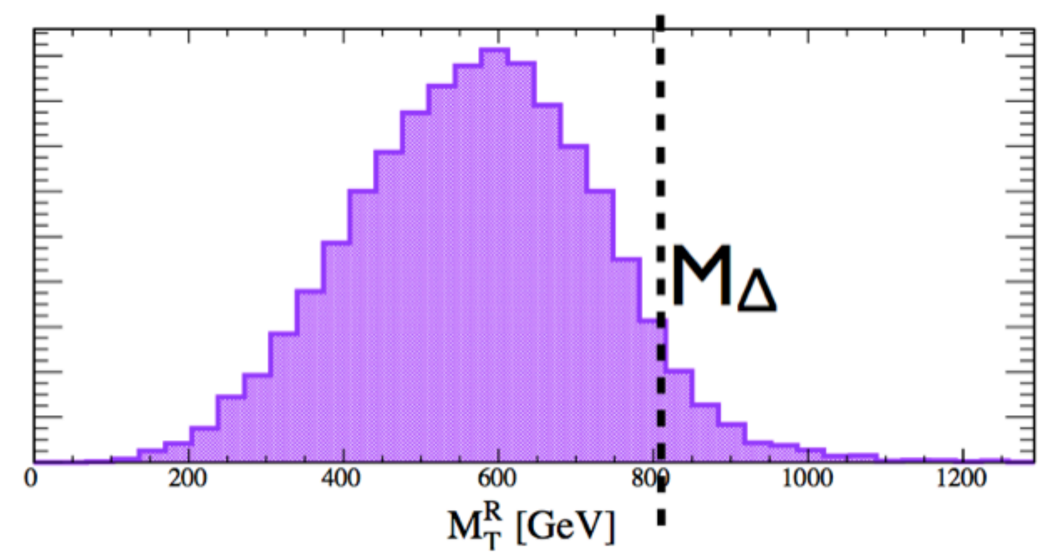
- M_R peaks at the mass scale M_Δ:

$$M_R = \sqrt{(|\vec{p}_{j_1}| + |\vec{p}_{j_2}|)^2 - (p_z^{j_1} + p_z^{j_2})^2}$$



- M_T^R is related to the missing energy and approximates M_Δ as an edge:

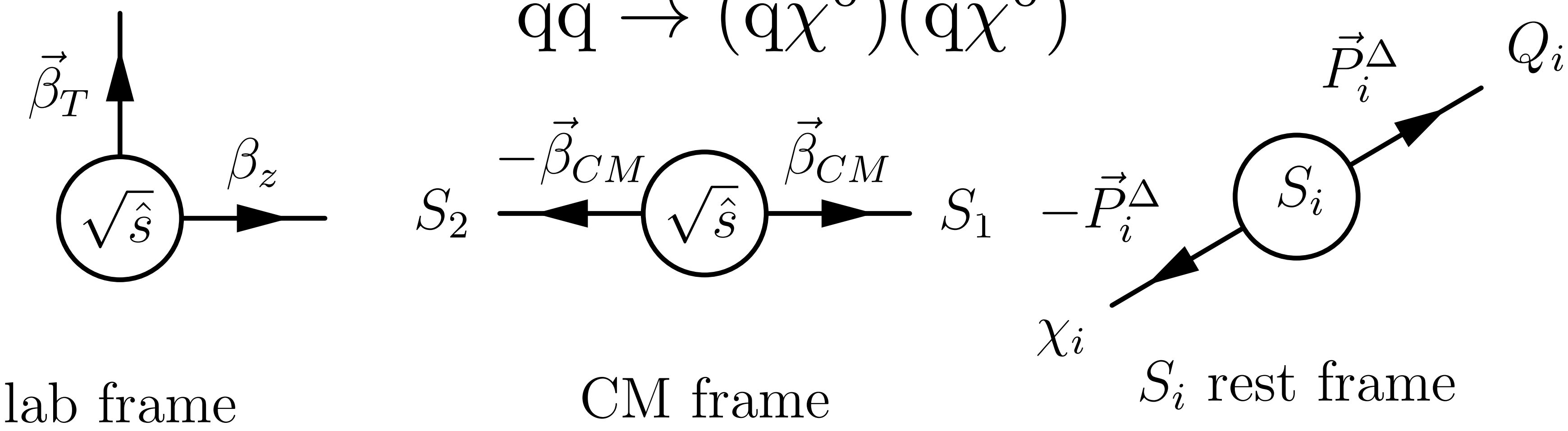
$$M_T^R = \sqrt{\frac{E_T^{miss}(p_T^{j_1} + p_T^{j_2}) - \vec{E}_T^{miss} \cdot (\vec{p}_T^{\,j_1} + \vec{p}_T^{\,j_2})}{2}}$$



- R is their ratio: $R = \dfrac{M_T^R}{M_R}$

Razor Variables: DOF Counting

$$\tilde{q}\tilde{q} \rightarrow (q\tilde{\chi}^0)(q\tilde{\chi}^0)$$



2 x $\tilde{\chi}_1^0$ — (E, p_x, p_y, p_z)

8 : 2 four vectors of the missing neutralinos

Constraints

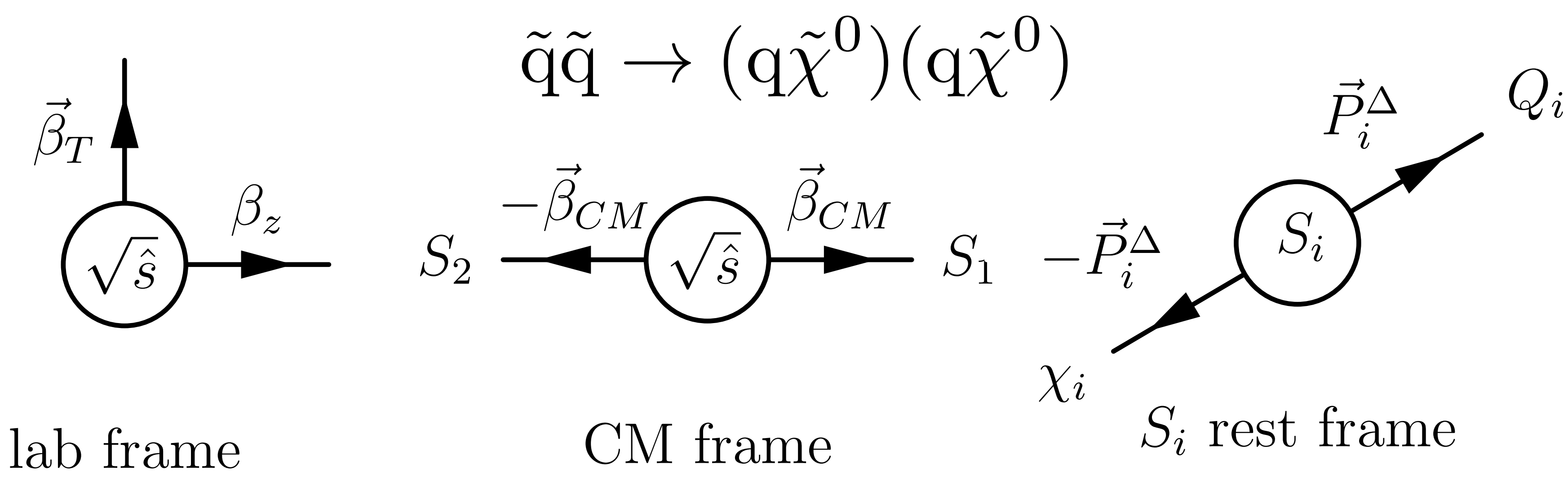
-2 : Conservation of Transverse Momentum (p_x, p_y)

-2 : The squarks and neutralinos must have the same mass

-1 : Both jets must have the same momentum in the decay frame

3 Remaining DOF to solve for the decay boost

Razor Variables: DOF Counting



	3	Remaining DOF after applying constraints
Approx.	-1	The jet energy (M_R) is invariant under longitudinal boosts
	-2	The transverse components of the boost to the CM frame are chosen to minimize the jet energy (sparticles are produced ~approx. rest)
	0	Solved

Sample Definitions

Triggers: Both samples are required to pass the OR of Hgg triggers:

`HLT_`**`Photon26`**`_R9Id85_OR_CaloId10_Iso50_`**`Photon18`**`_R9Id85_OR_CaloId10_Iso50_`**`Mass70`**

`HLT_`**`Photon36`**`_R9Id85_OR_CaloId10_Iso50_`**`Photon22`**`_R9Id85_OR_CaloId10_Iso50`

These two triggers have been shown to have 99.5% efficiency to T&P Zee decays that pass preselection with MVA score > .05 and pt (30, 22.5) [AN-12-160]

Selection (gg) Sample:

- 2 Photons 30, 22 passing egamma loose WP
- 1 or more jets with pt > 40 passing jet ID
- Hemispheres are clustered with jets passing ID and pt > 40

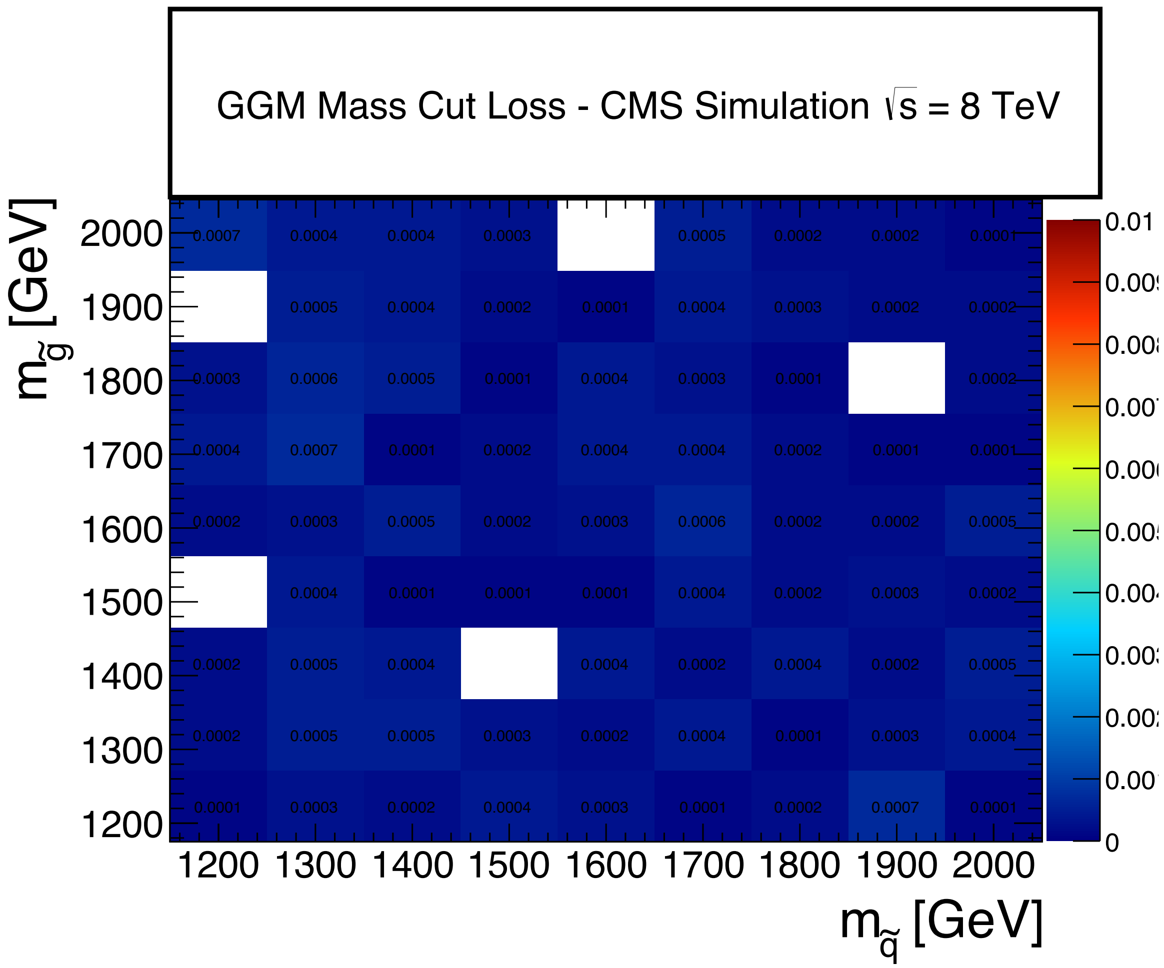
Fake-Fake (ff) Sample:

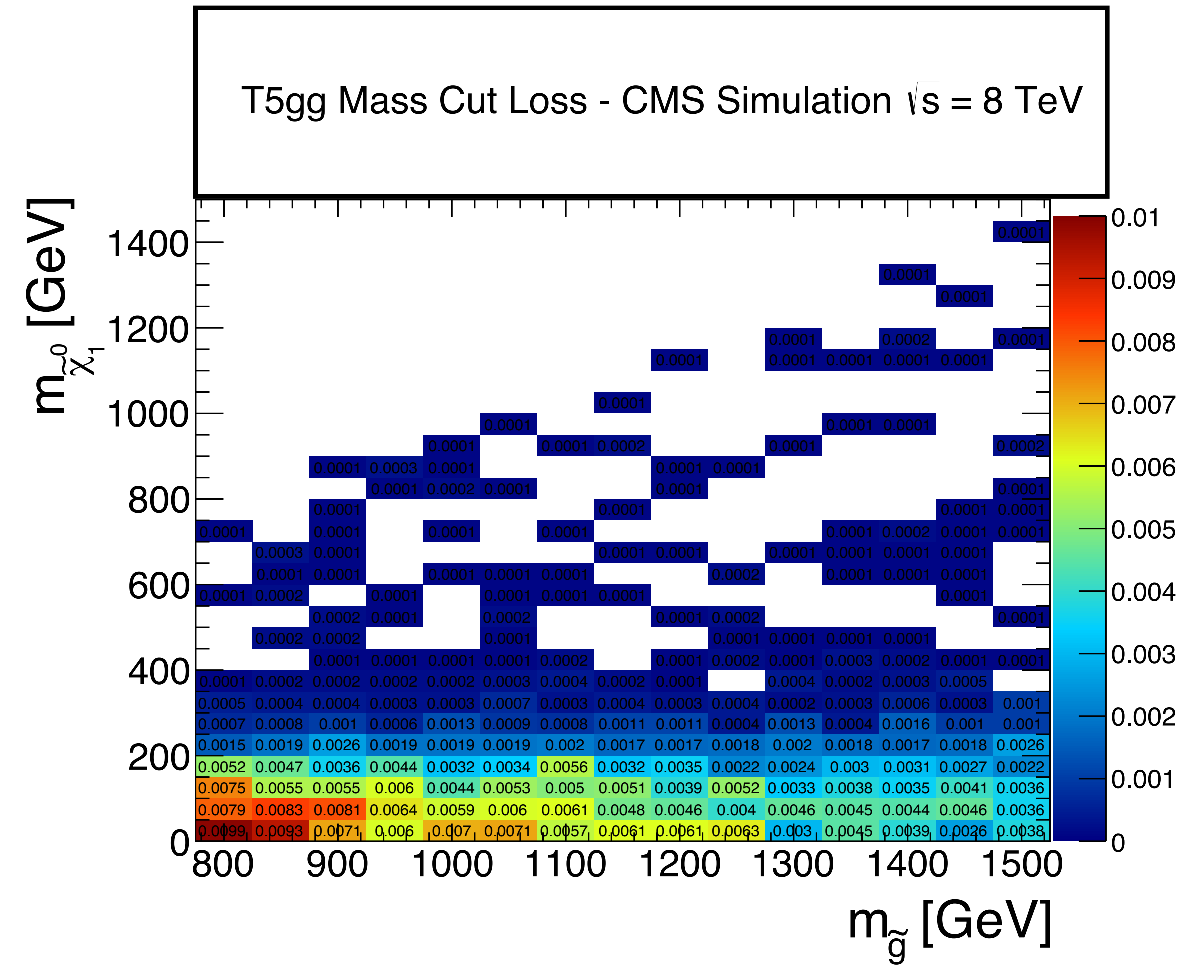
- The same as the selection sample, but the photon ID is modified
- Fake photons either fail the $\sigma_{ietaieta}$ cut or exactly 2 of the 3 isolation variables

Resonant Trigger Mass Cut Inefficiency

HLT_**Photon26**_R9Id85_OR_CaloId10_Iso50_**Photon18**_R9Id85_OR_CaloId10_Iso50_**Mass70**

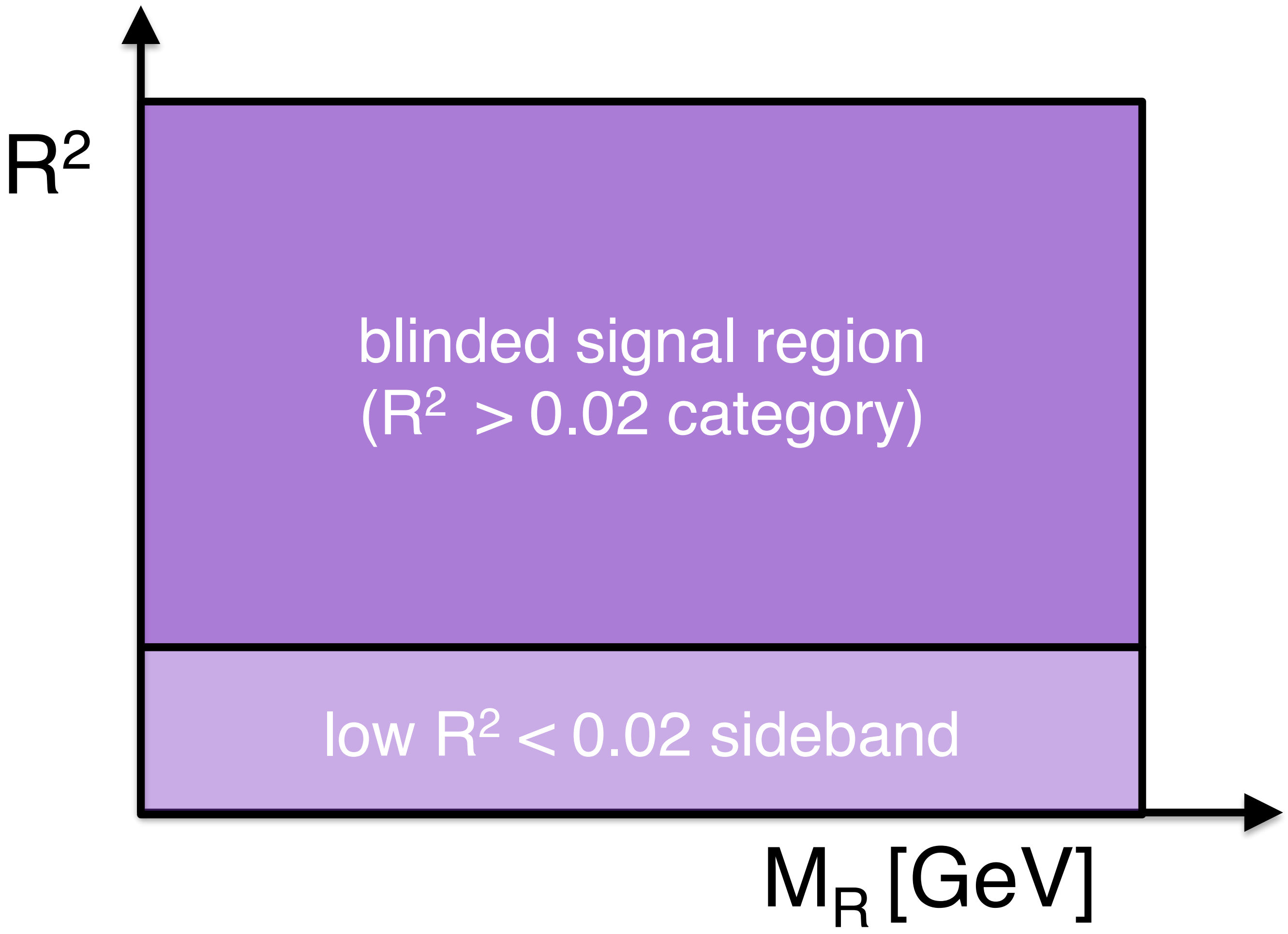
HLT_**Photon36**_R9Id85_OR_CaloId10_Iso50_**Photon22**_R9Id85_OR_CaloId10_Iso50





- Plotted is the fraction of σ x L with diphoton mass < 100 GeV that fails a 40 (25) leading (sub-leading) cut
- The conclusion is that signal events with low invariant mass are generally picked up by the non-resonant trigger

Baseline Cuts and Categories

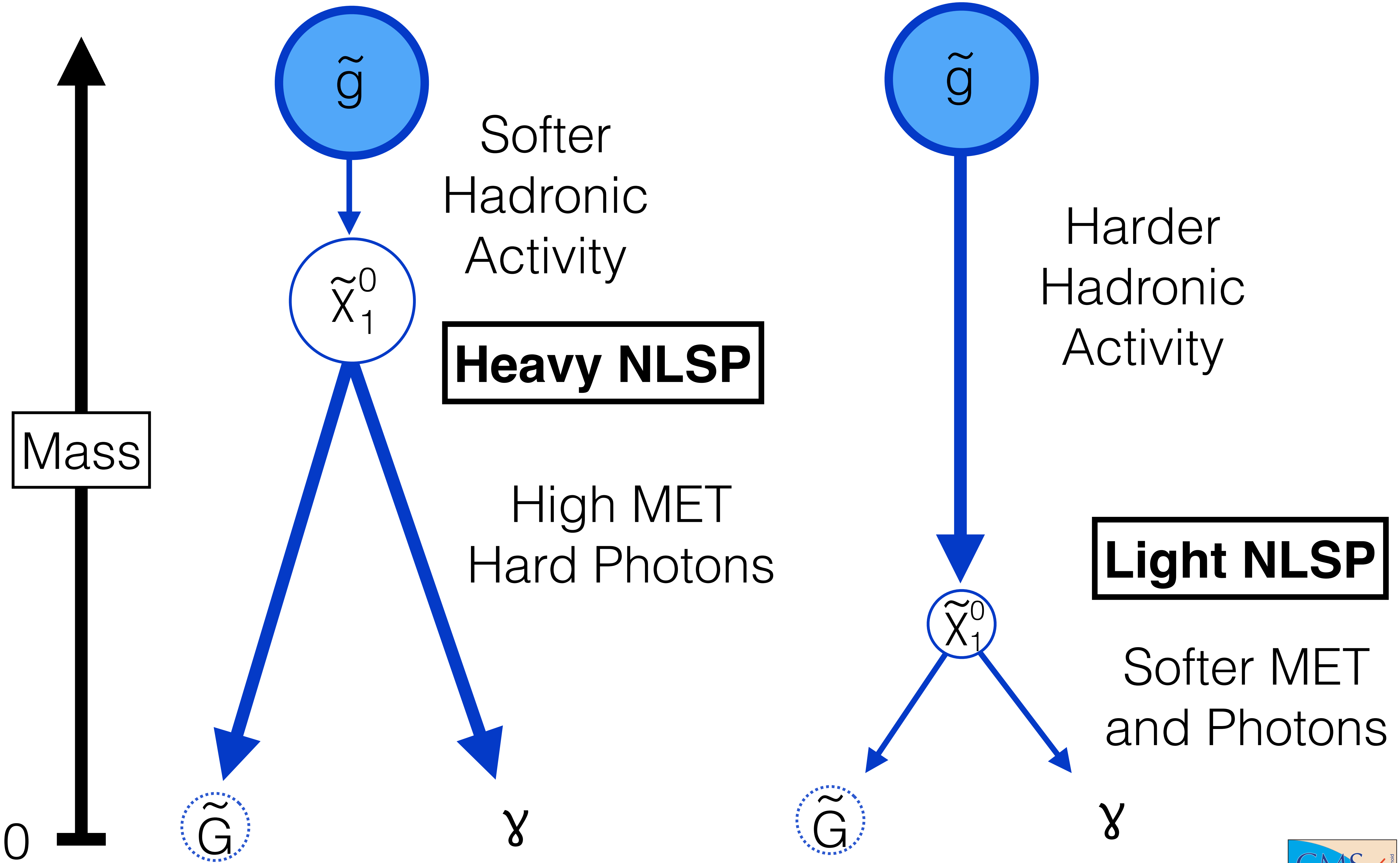


- Baseline cuts are made at $R^2 > 0.01$ and $M_R > 600$
- The plane is further divided into two kinematic regions:
  - Control region of low $R^2 < .02$
  - Signal region of high $R^2 > .02$
- Squeezing the control band limits our signal contamination

The guidelines for how these categories were chosen:

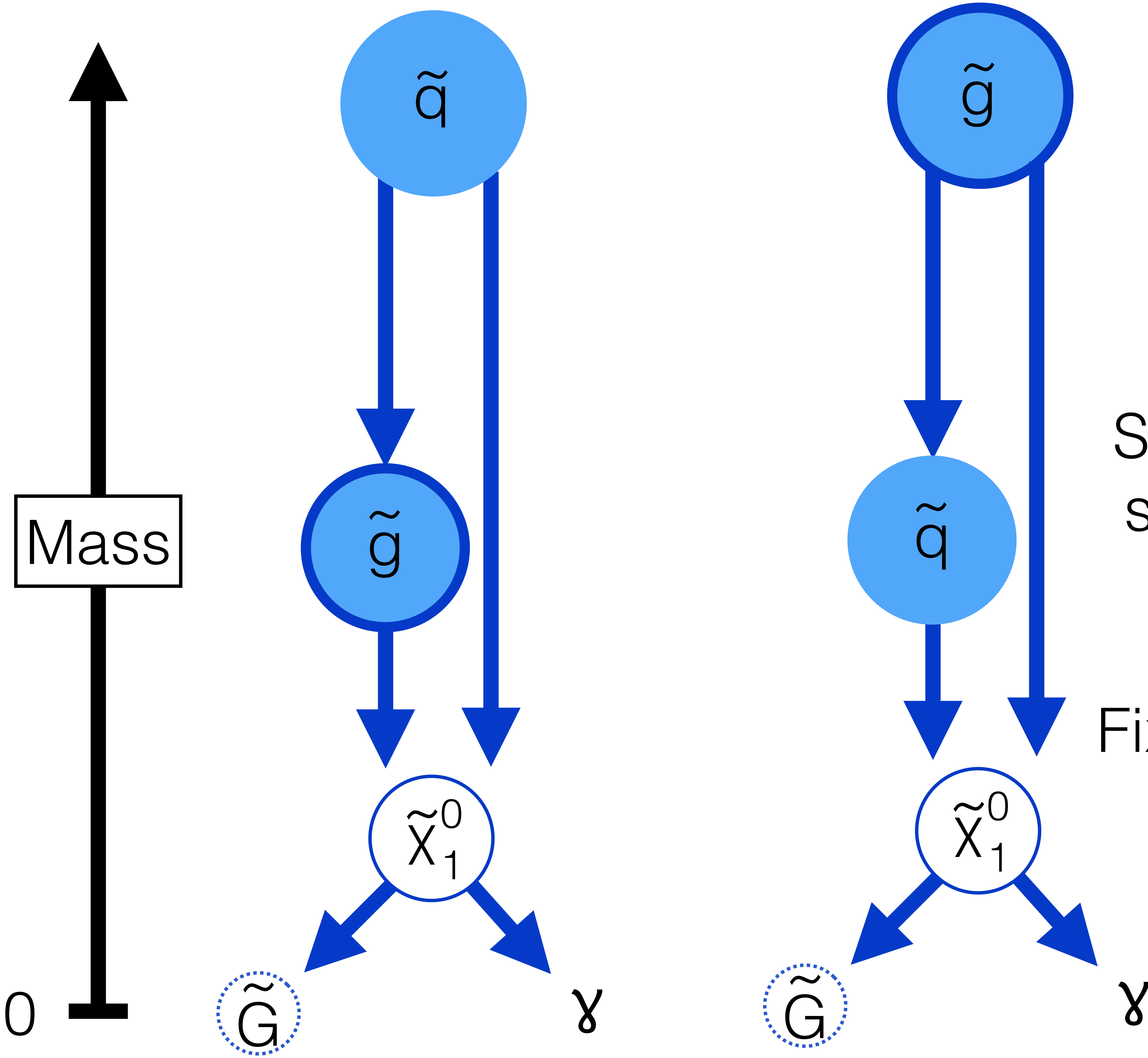
- less than 10% signal contamination
- $>$ 1k events in the fit region for each sample
- ~1k events in the for each sample extrapolation region
- signal region maximally efficient to signal

Kinematic Regimes (T5gg)



Kinematic Regimes (GGM)

Heavy Squark



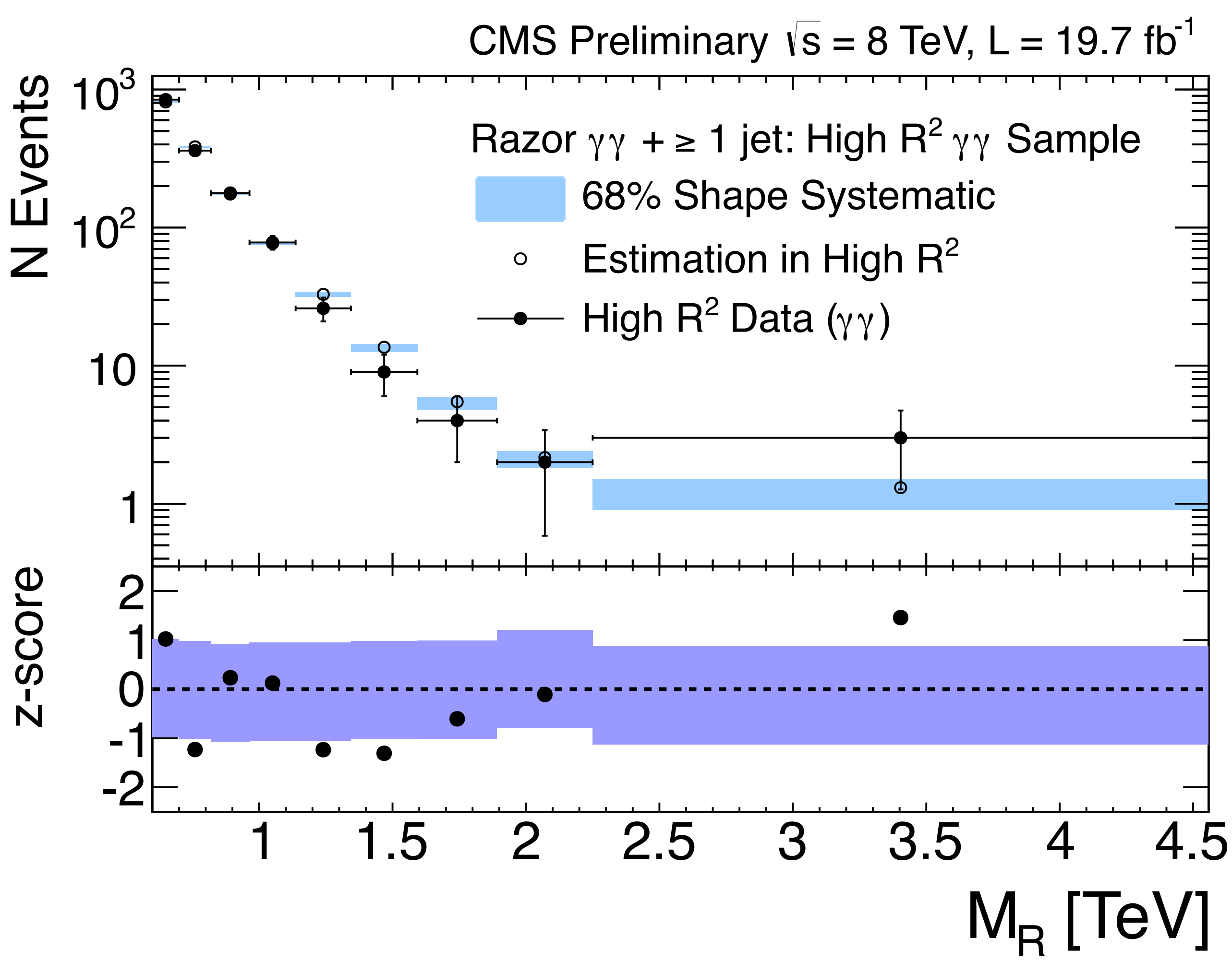
Heavy Gluino

On or Off
Shell decays through
squarks and gluinos

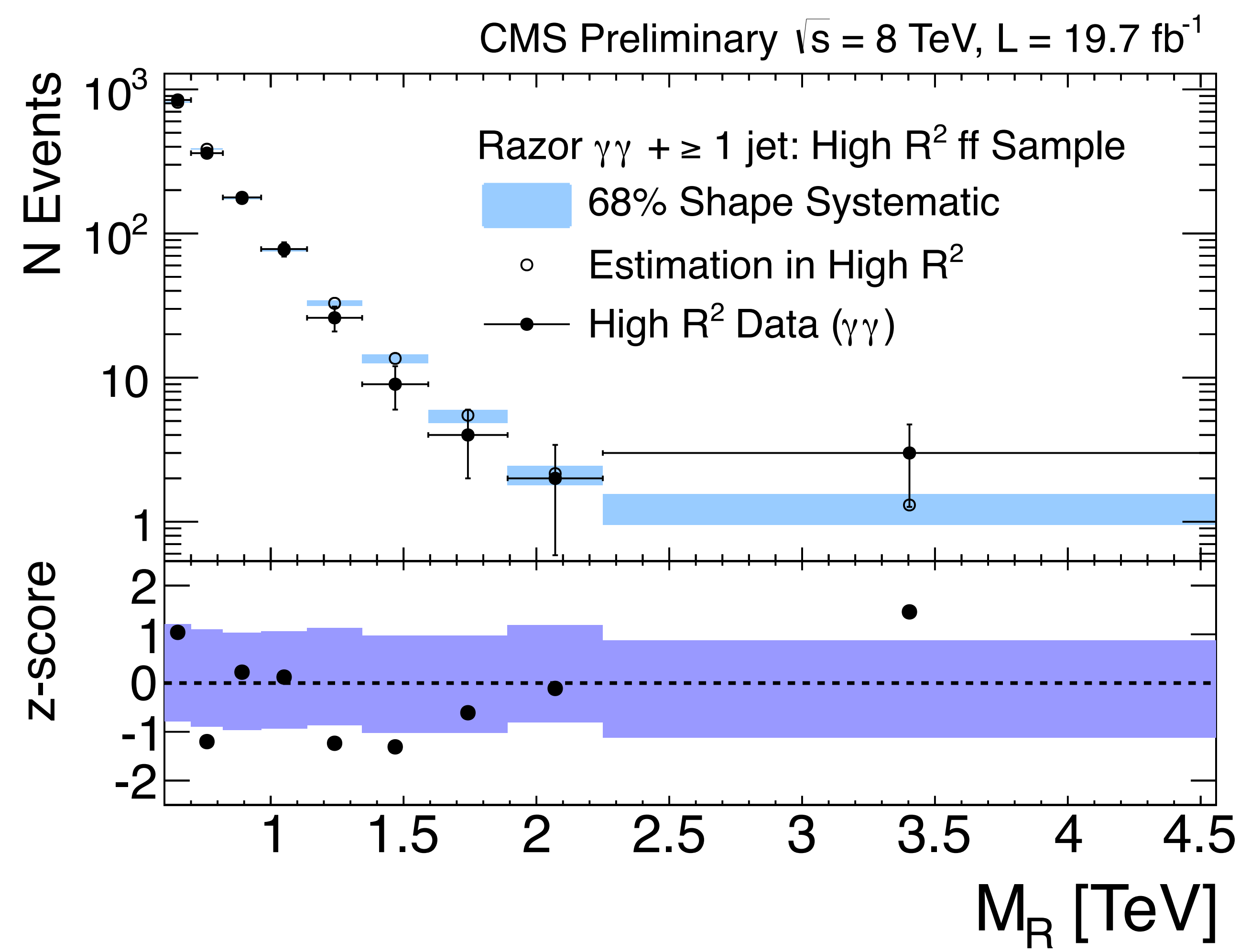
Fixed Neutralino Mass
of 375 GeV
(varied branching
fractions)

Extrapolation Normalization

Normalization fixed
to signal region

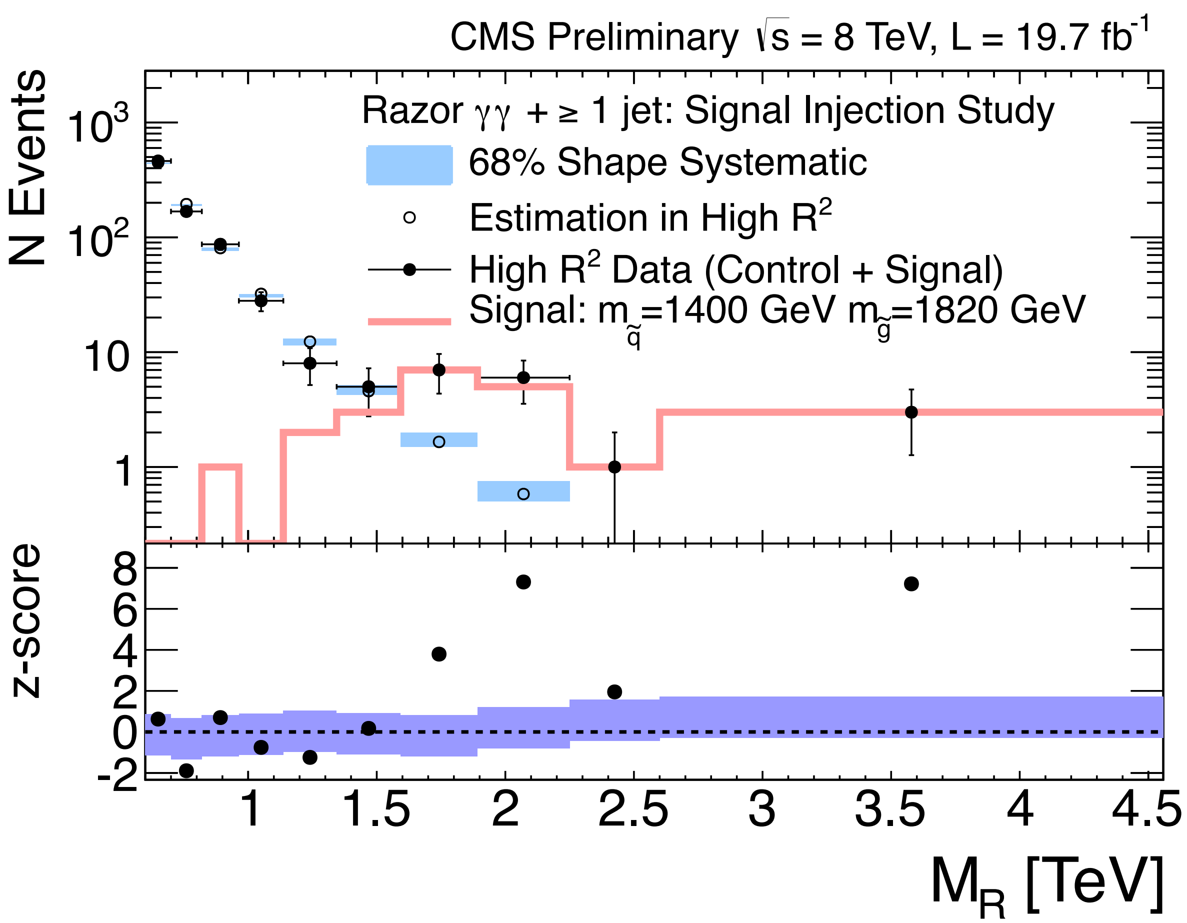


Normalization fixed to first
four bins of M_R in signal region

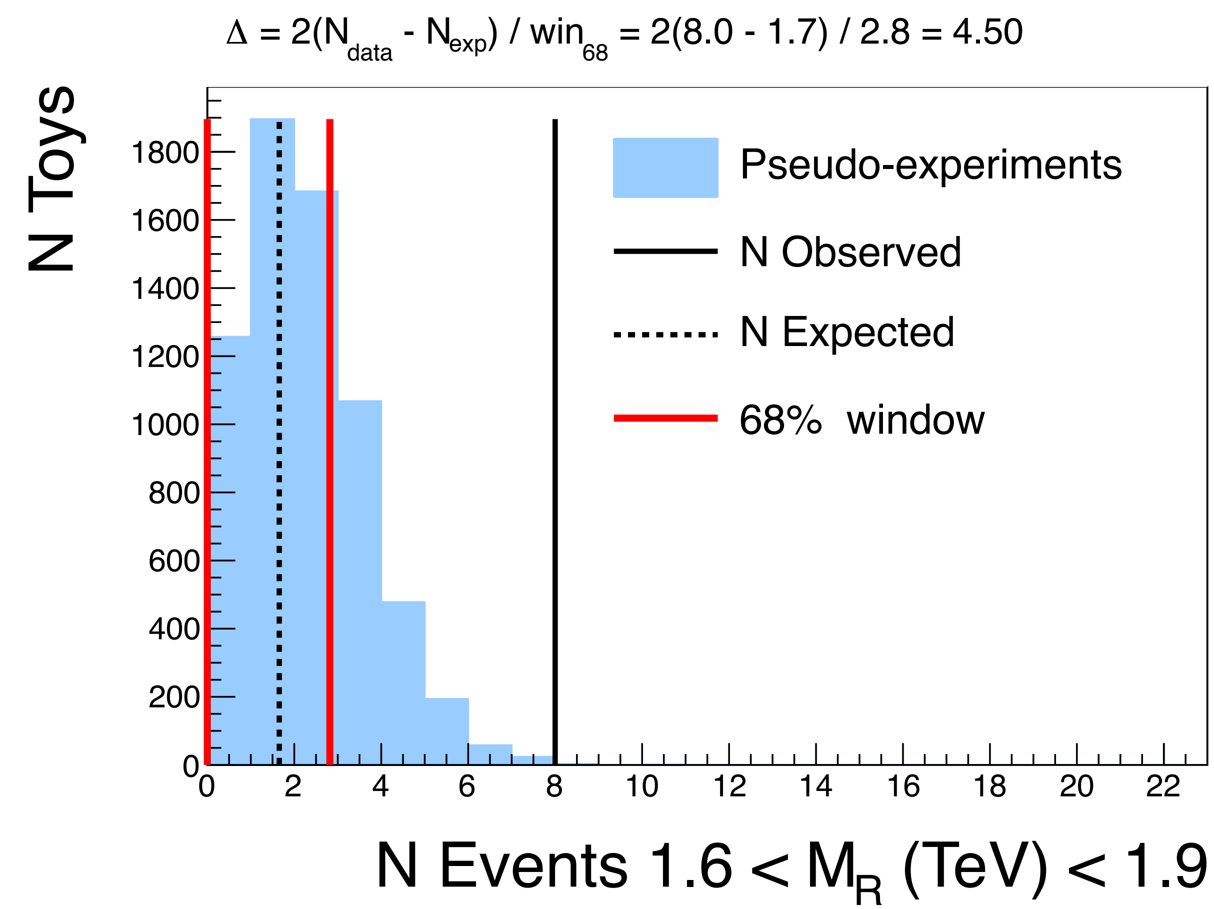


Nearly identical results are achieved in both cases

Signal Injection Test

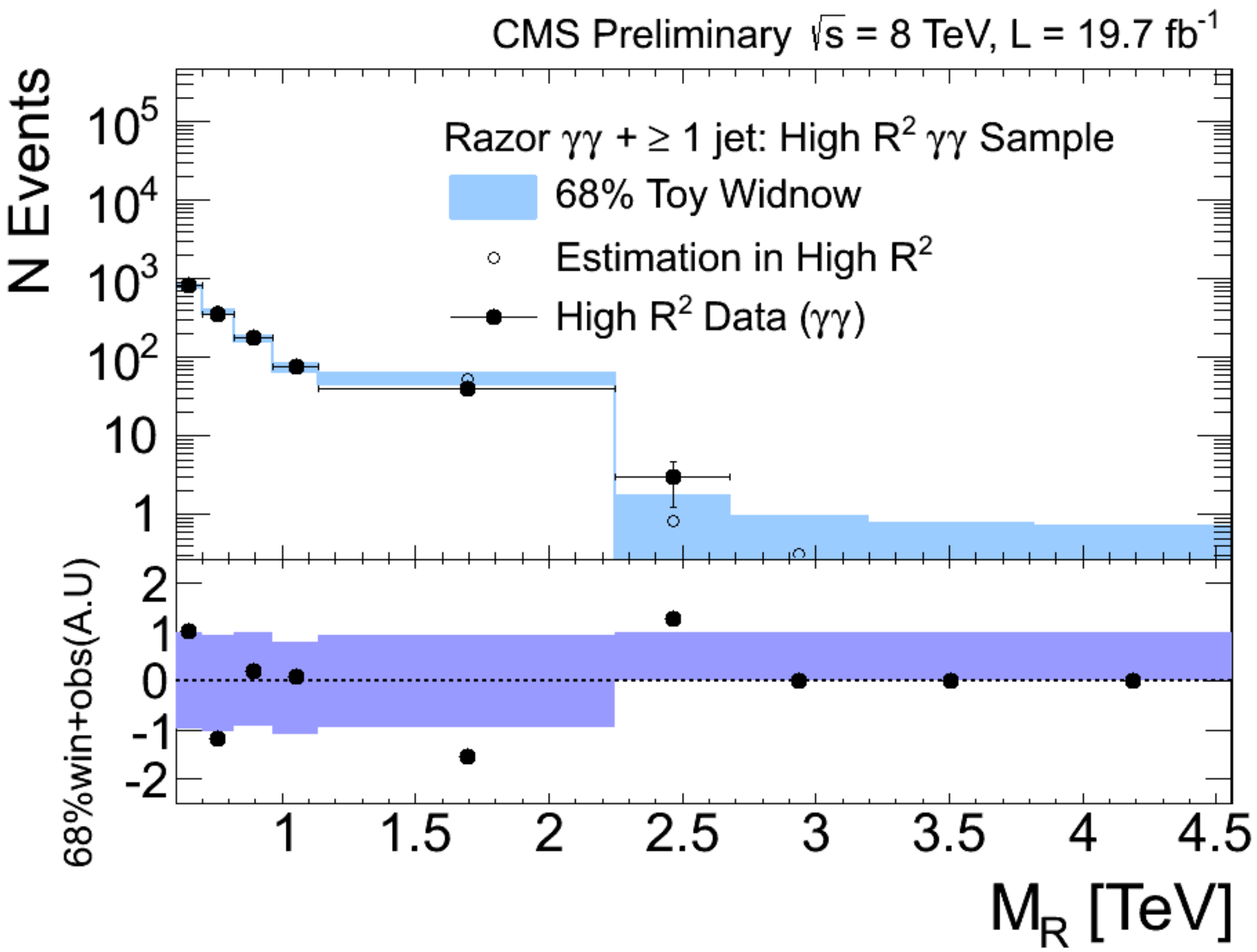


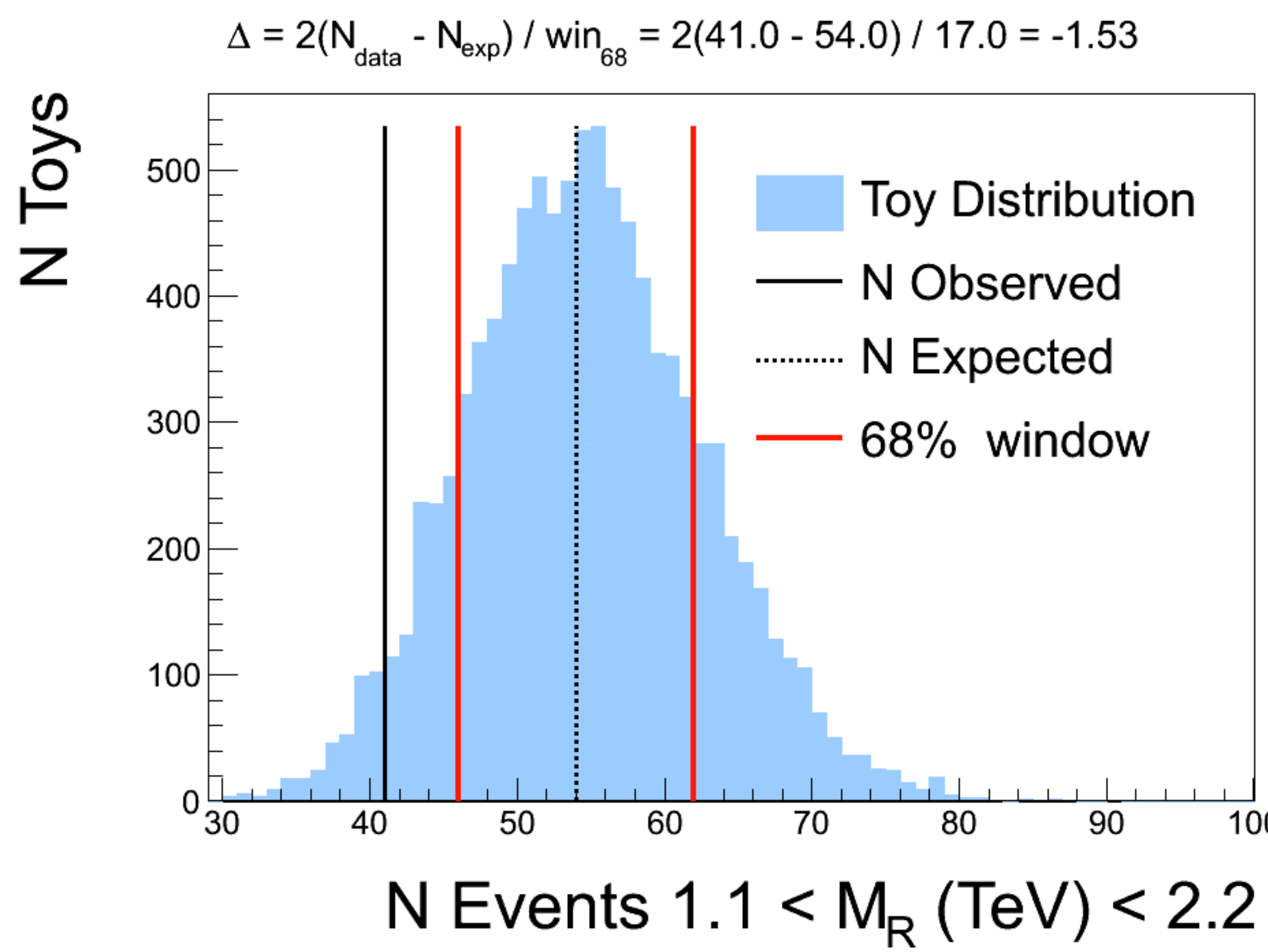
signal like fluctuation



- A cocktail of events of fake-fake and signal injected to represent the theoretical cross section xsec = 2.7 fb
- The closure procedure is performed again, showing our sensitivity at high M_R values to signal.

Bin to Bin Correlation of Underestimation



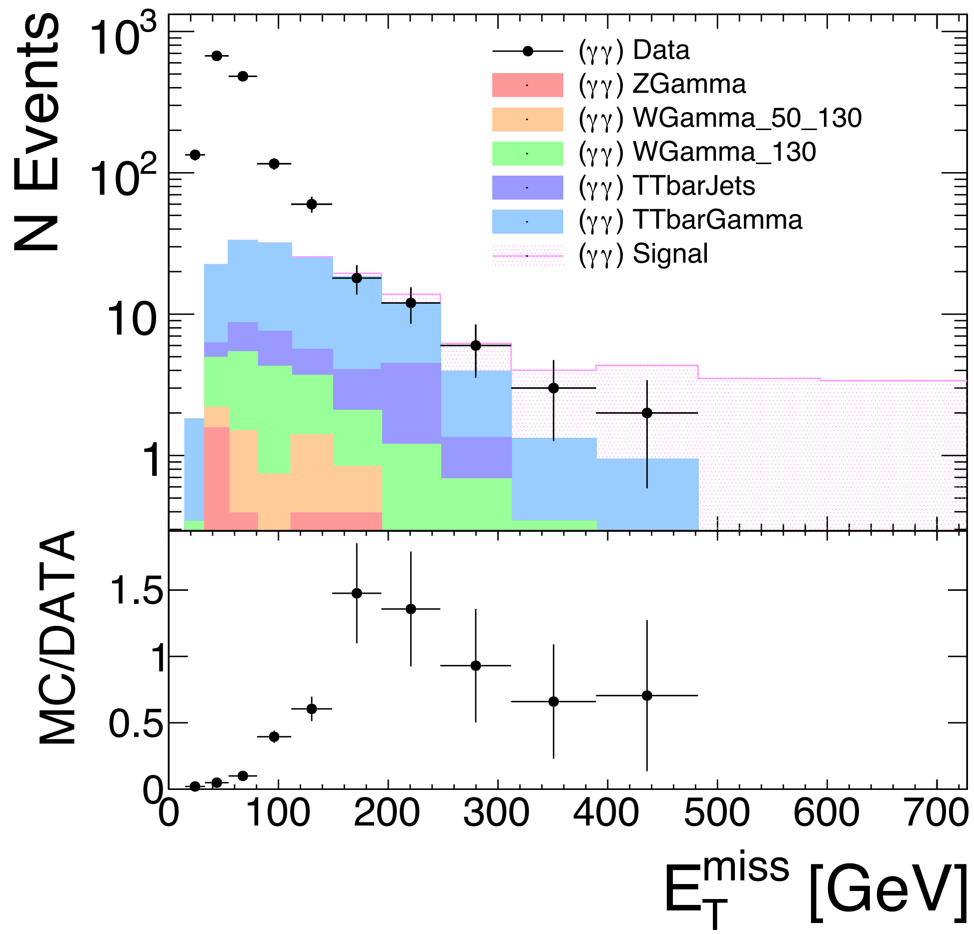


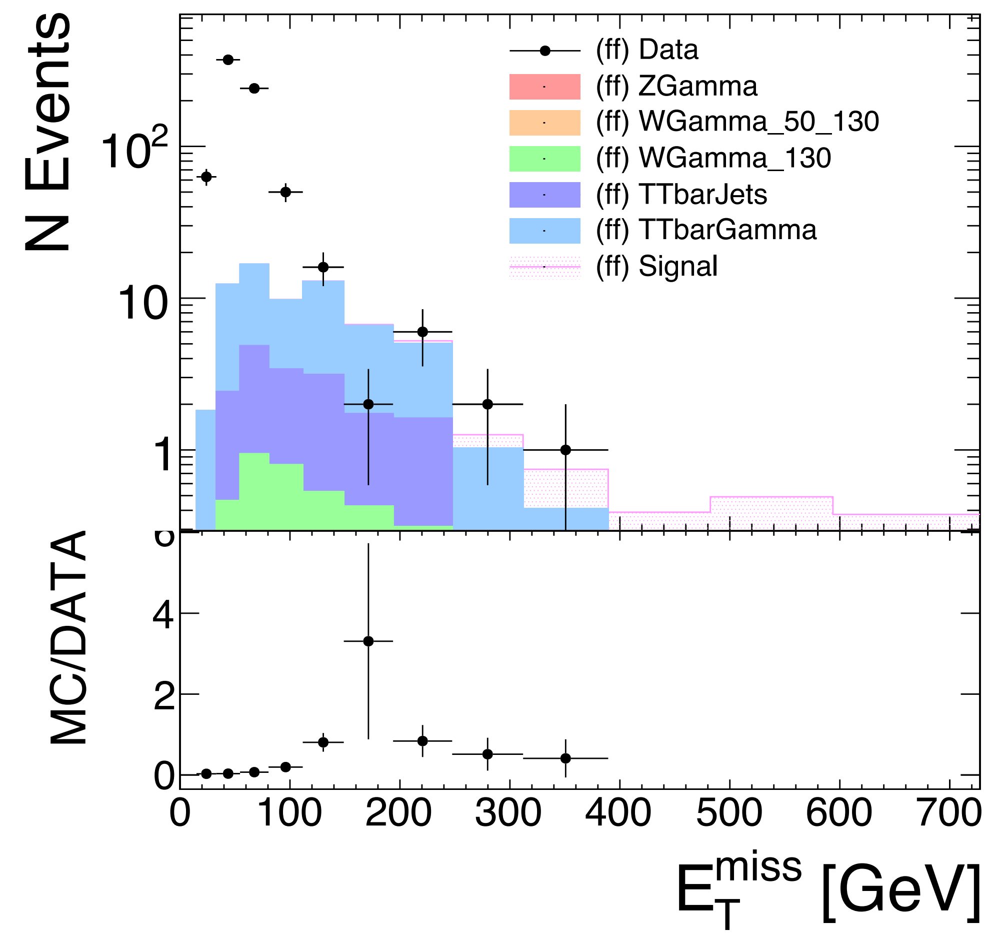
- The over-estimate is combined into bin
- Band has shape error convoluted with poisson
- Using the half 68% window as our one sigma definition, this is a 1.5 sigma **local** significance

Process	MC Sample for EWK Study	Order	xsec	#	Event Weight @ 20fb-1	k-factor***
TTbar+ gamma + jets	/TTGJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V19-v1/AODSIM	LO	14 pb	1.7M	0.16	1.5
W+gamma (pt: 130 - inf)	/WGToLNuG_PtG-130_8TeV-madgraph-pythia6_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	LO	0.25 pb	470k	0.01	1.5
W+gamma (pt: 50 - 130)	/WGToLNuG_PtG-50-130_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	LO	1.1 pb	1.1M	0.01	1.5
TTBar + jets	TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	NLO	225.2 pb	6.9M	0.65	None
Z+gamma	/ZG_Inclusive_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	LO	123 pb	6.3M	0.39	1.5

***A flat k-factor of 1.5 is used for LO cross sections

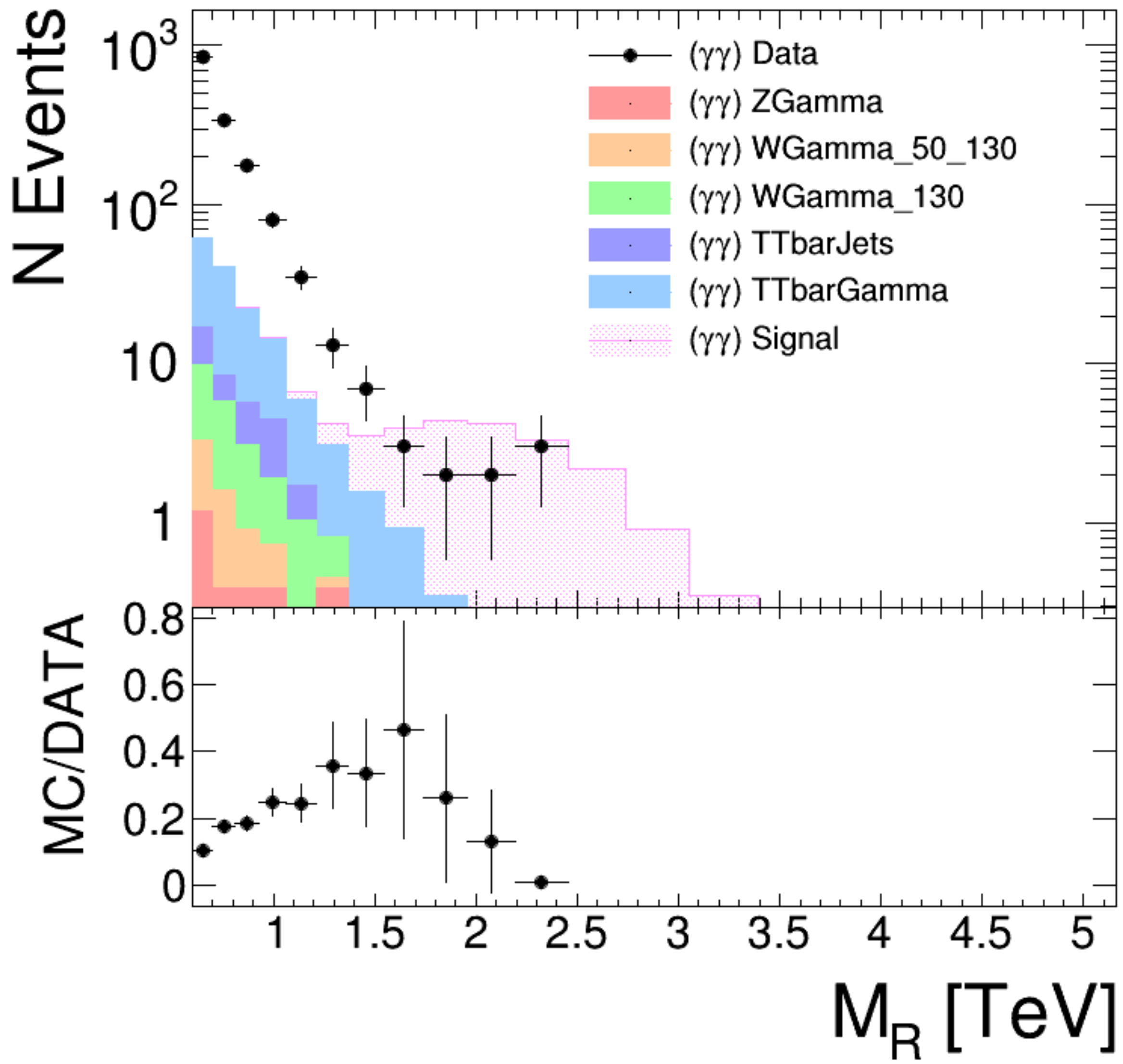
MET Distribution (Signal Region)

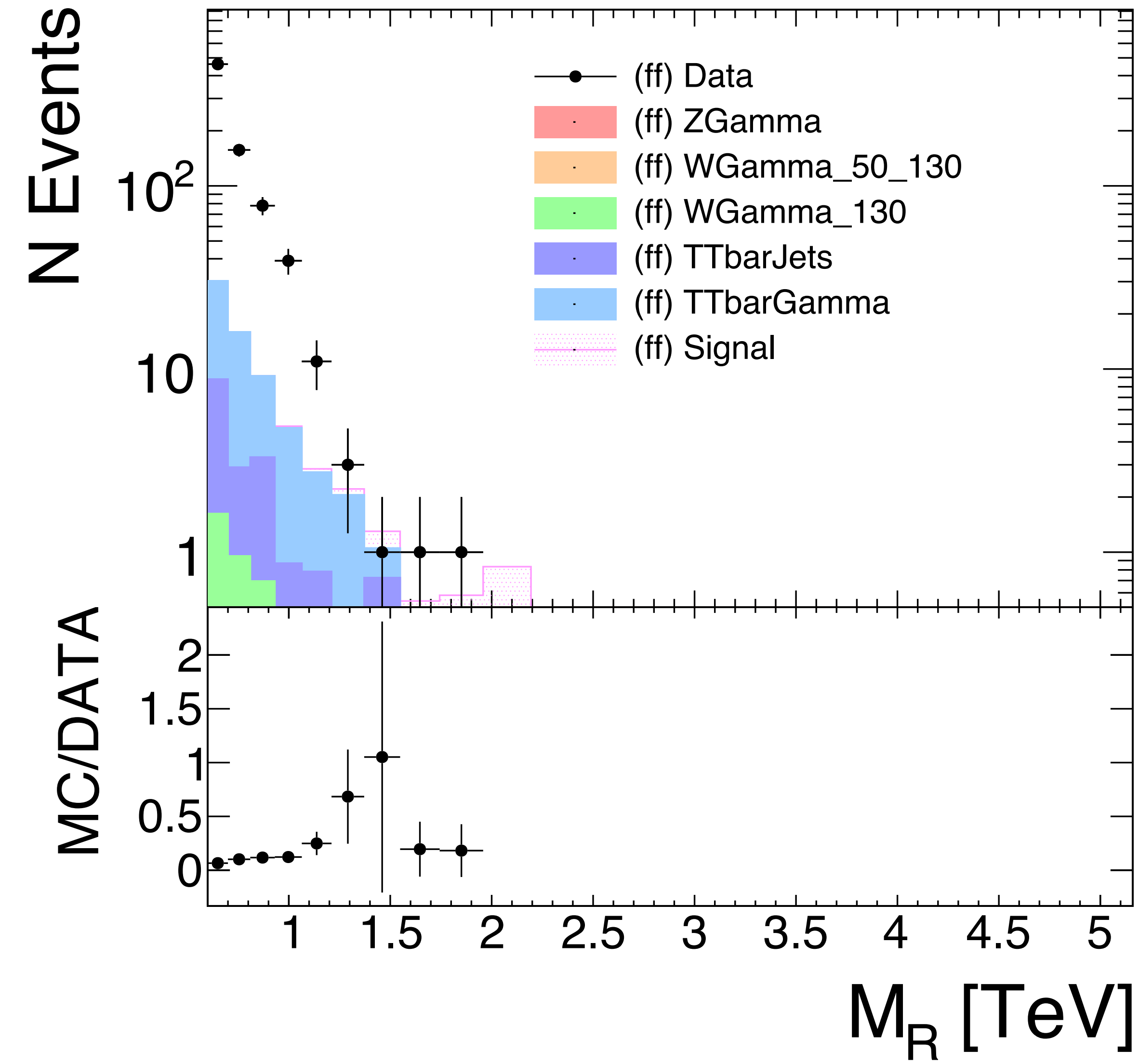




- Fake MET backgrounds j+j, γ+j, and γ γ there is no sample large enough for 20 fb^{-1} and are not included
- Checking the signal region of the razor analysis $M_R > 600$ and $R^2 > .02$ we see that the high MET distribution is dominated by EWK bkg
- We also see a similar effect in the fake-fake sample. This shows the fake definition is inclusive to relevant the EWK backgrounds.

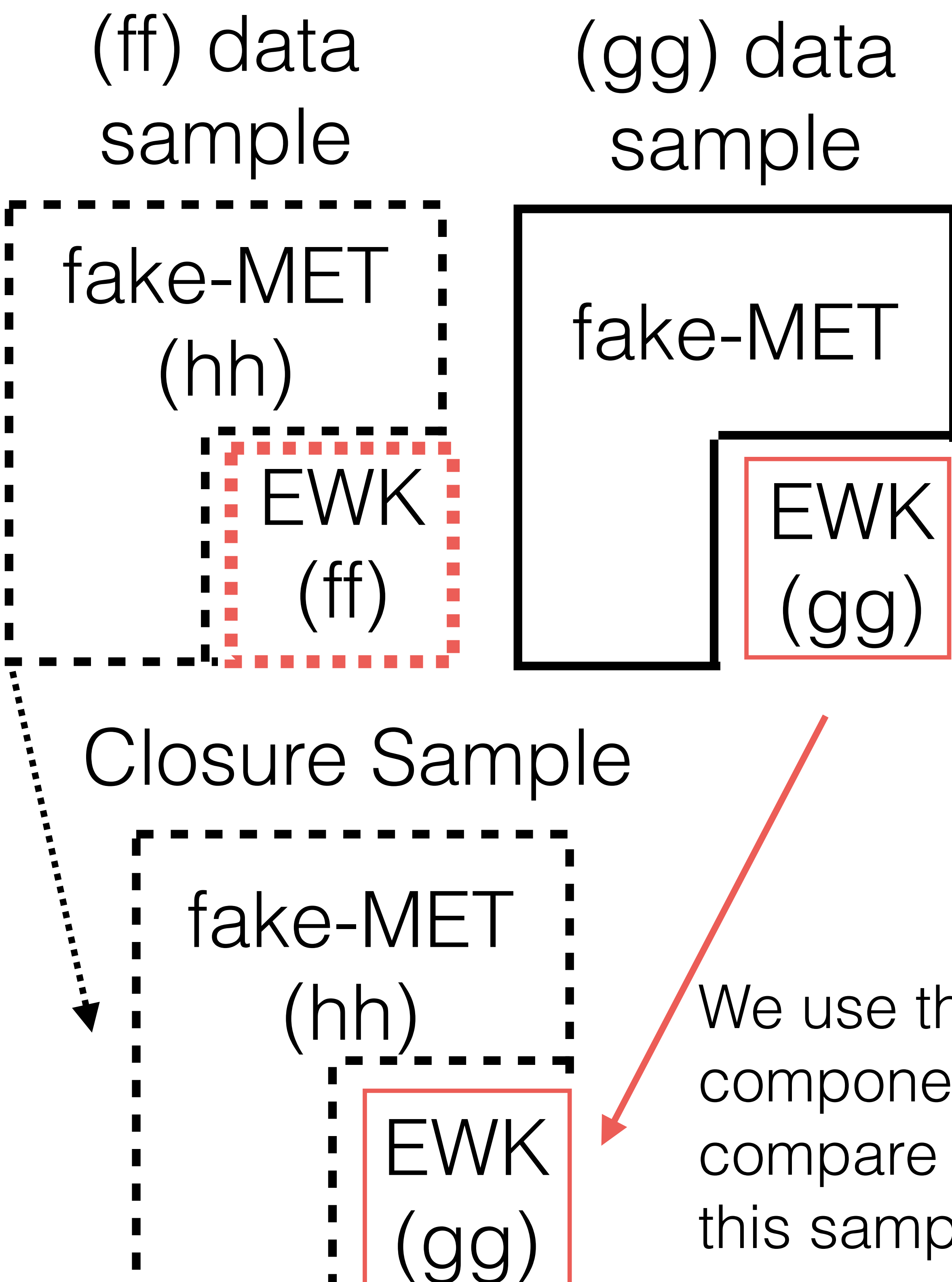
MR (Signal Region)

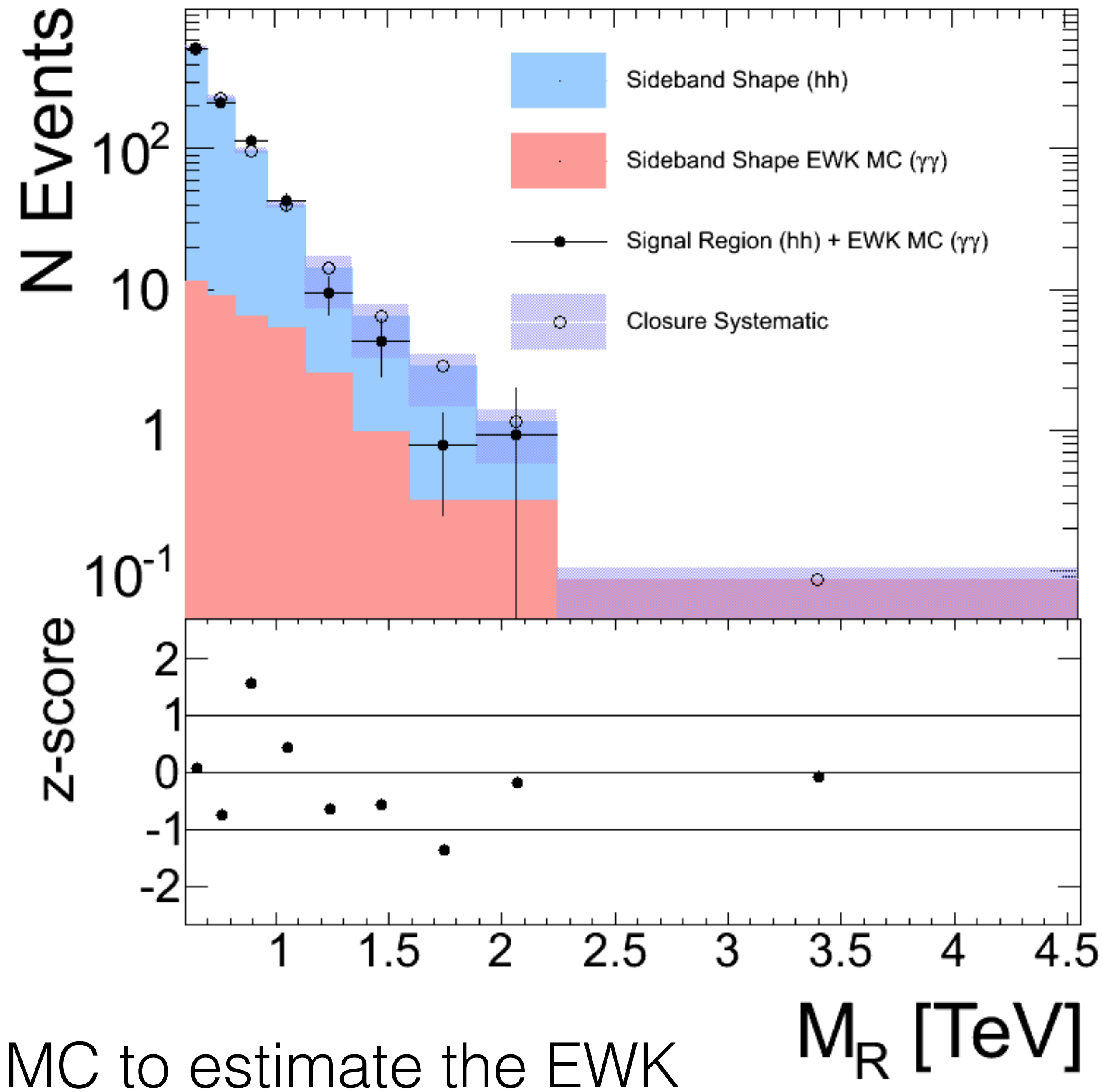




- When we project these same events into M_R we see the EWK backgrounds are subdominant to the fake-MET backgrounds across the full distribution
- A similar small contribution is seen in the fake-fake sample

EWK (real MET bkg) Closure





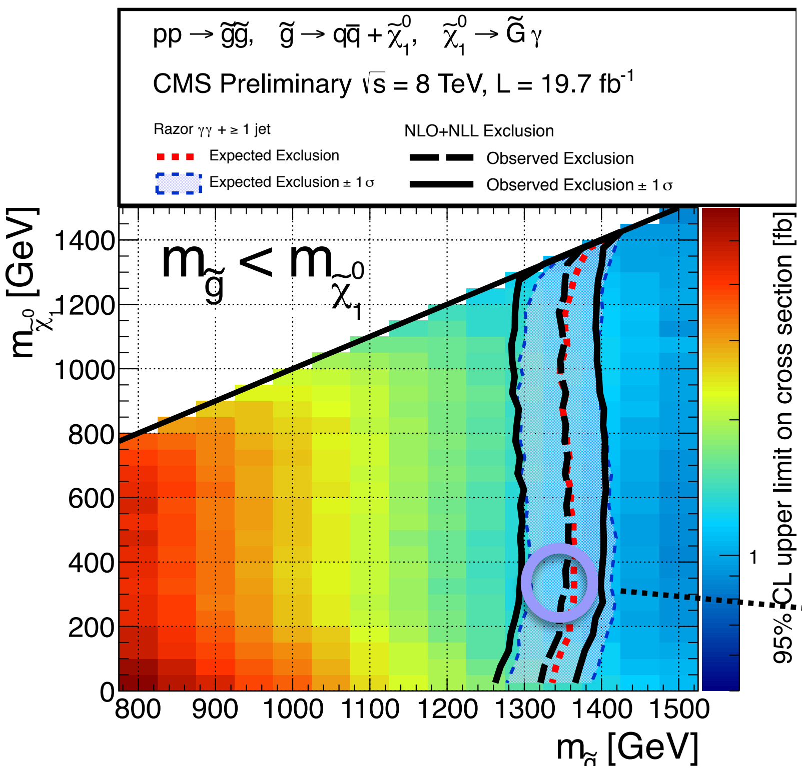
We use the MC to estimate the EWK component of each sample. Then compare the shape of the sideband of this sample to that of the signal region

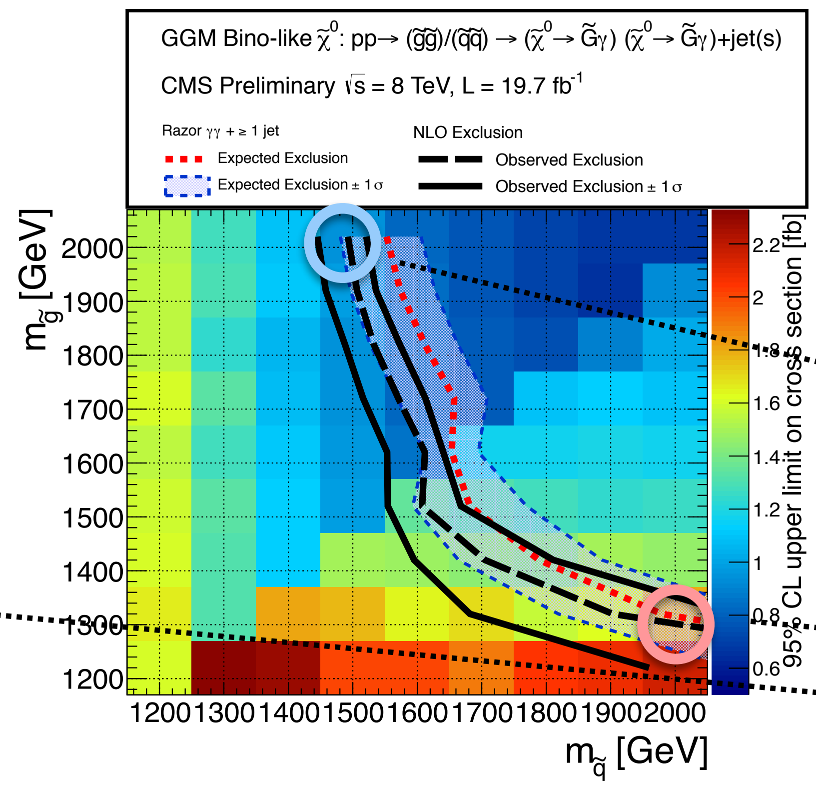
Systematics Summary

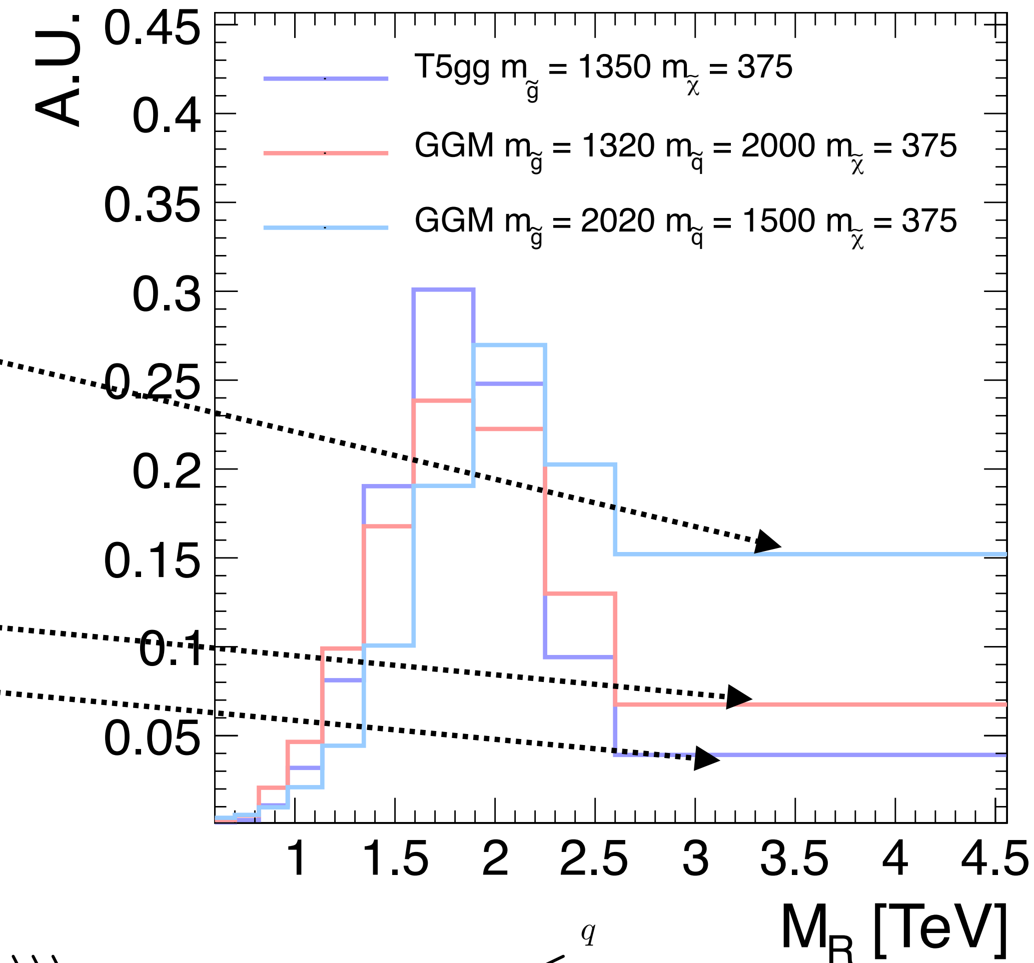
	Systematic	Typical Size of Systematic T5gg	GGM
Signal Systematics	PDF Acceptance Error	1%	1-3%
	PDF Rate Error	inc. in SMS xsec	2-50%
	Data/MC Scale Factors	1%	
	Luminosity	2.6%	
	Jet Energy Scale (shape)	2-5%	
	Trigger Efficiency	1%	
Bkg	Fit Parameter Uncertainty (shape)	2-40%	
	Fit Function Uncertainty (shape)	5-50%	

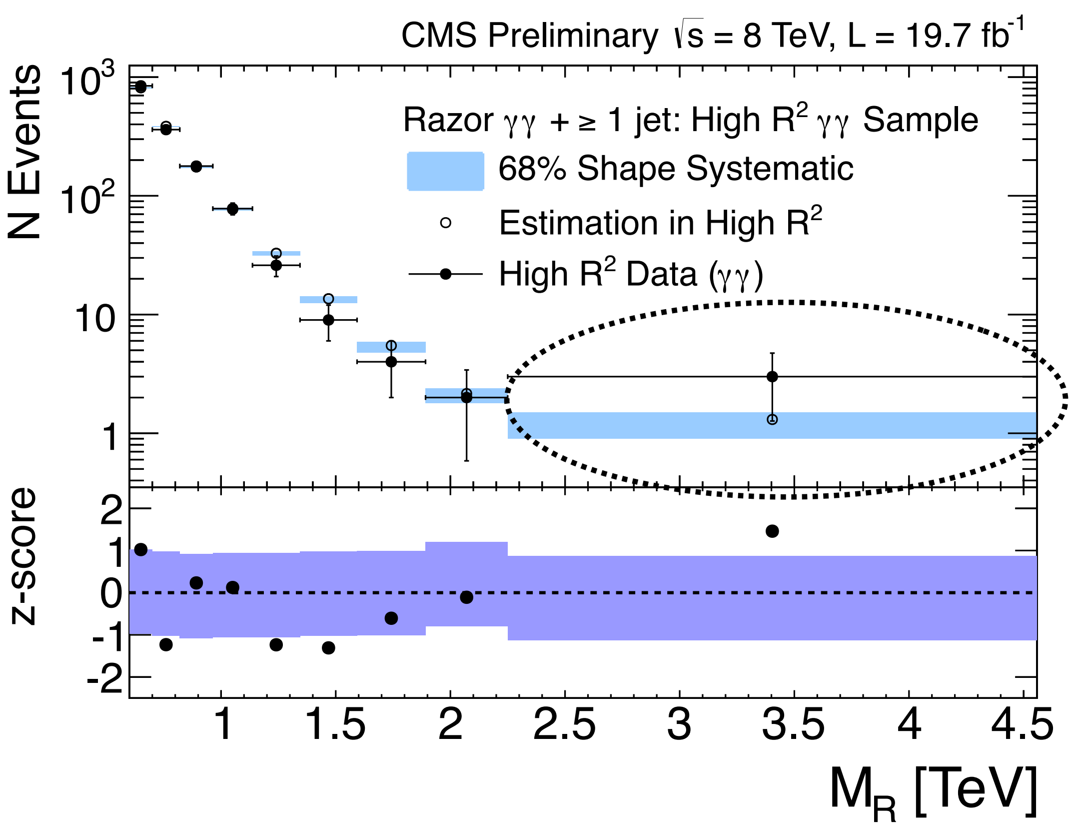
Systematics done bin by bin are labeled with (shape)

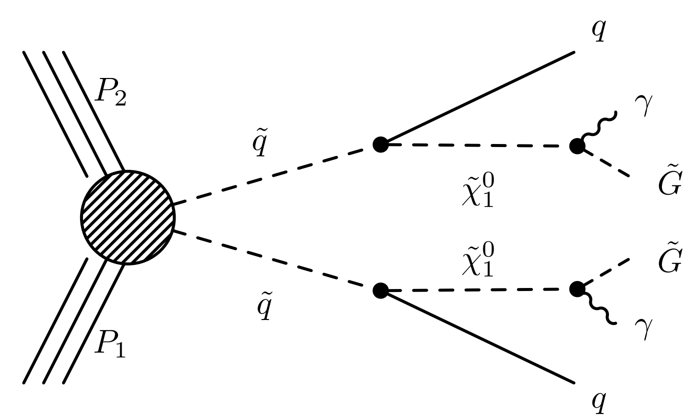
Differences in Limit





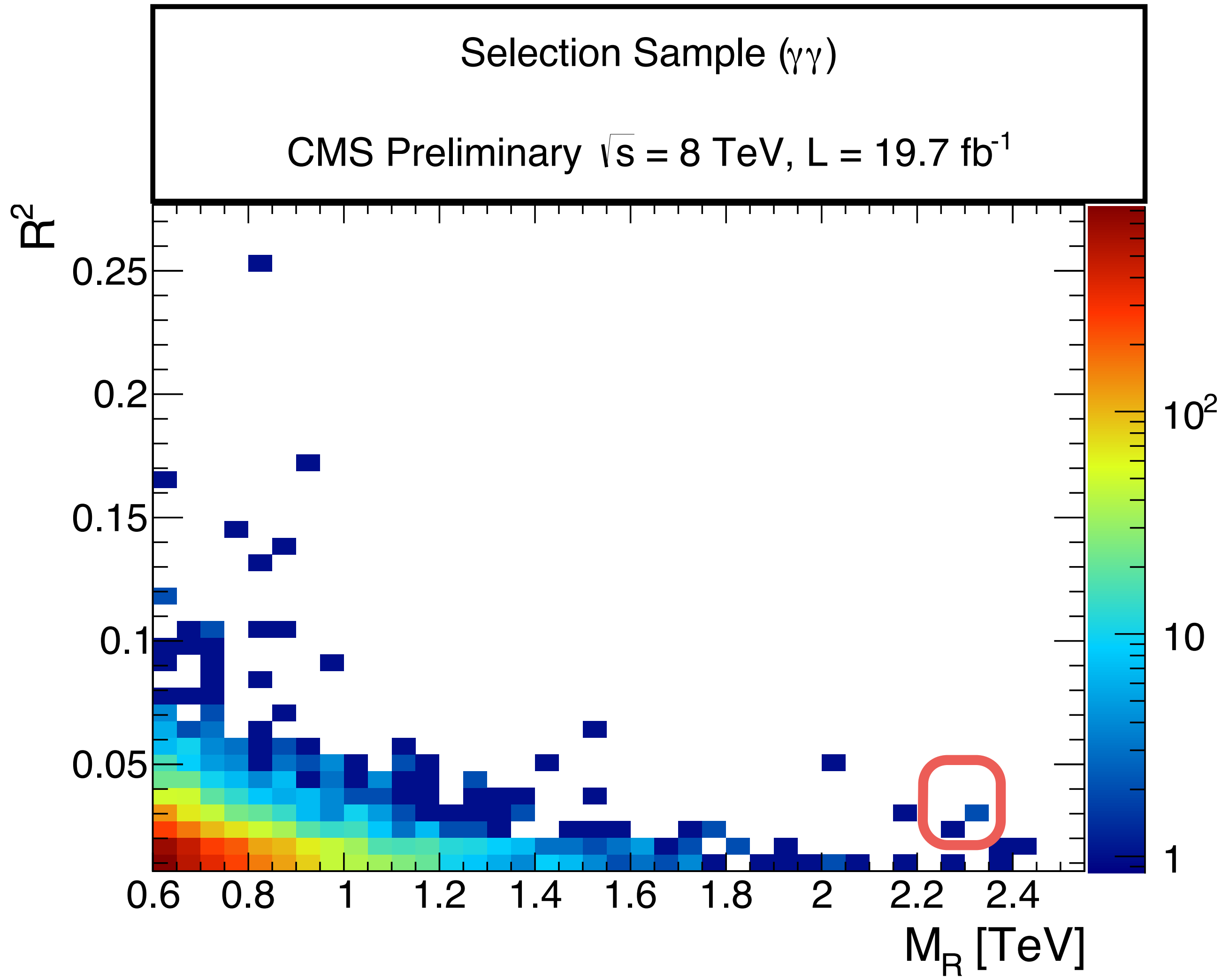


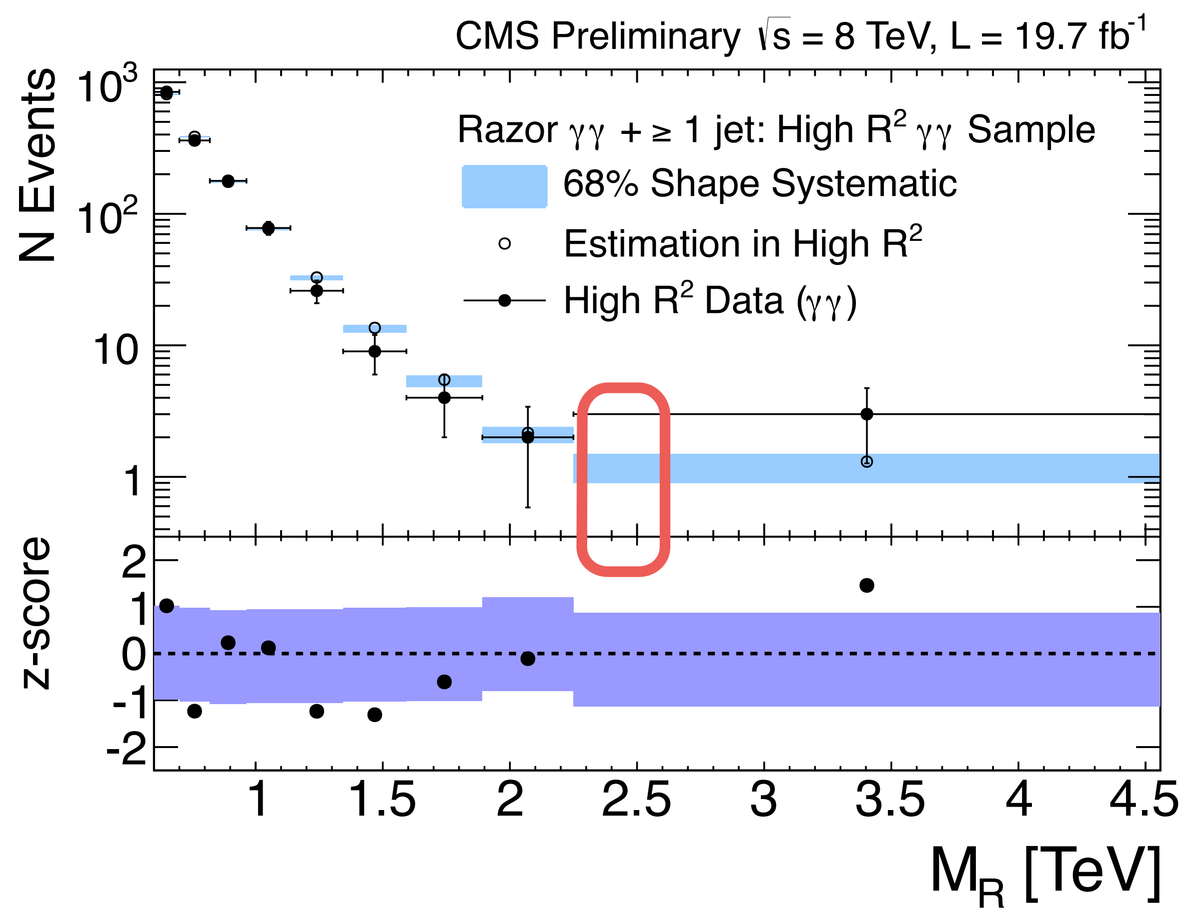




Signal like fluctuation favors the GGM light squark scenario where events have higher M_R

Signal-like Events

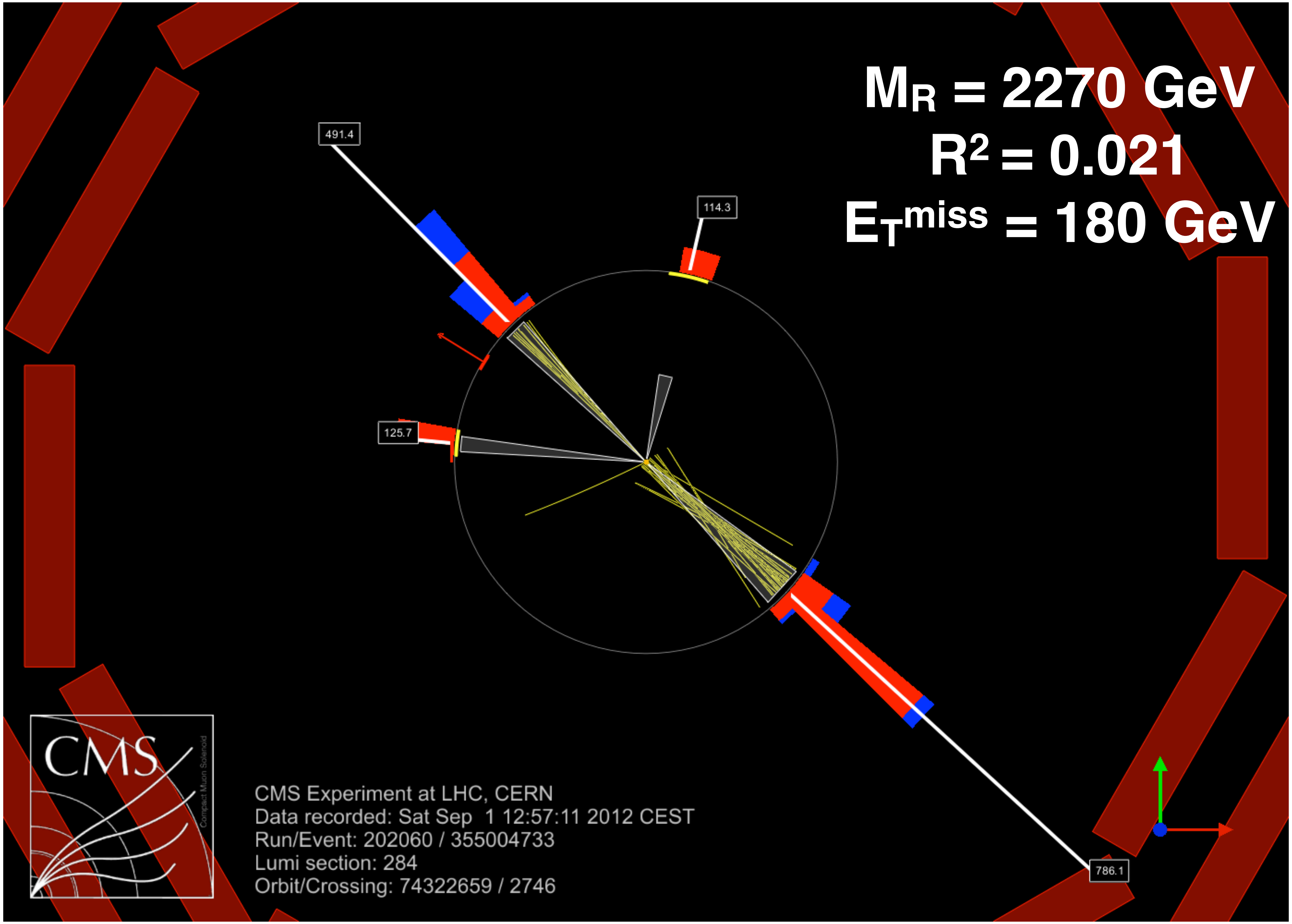




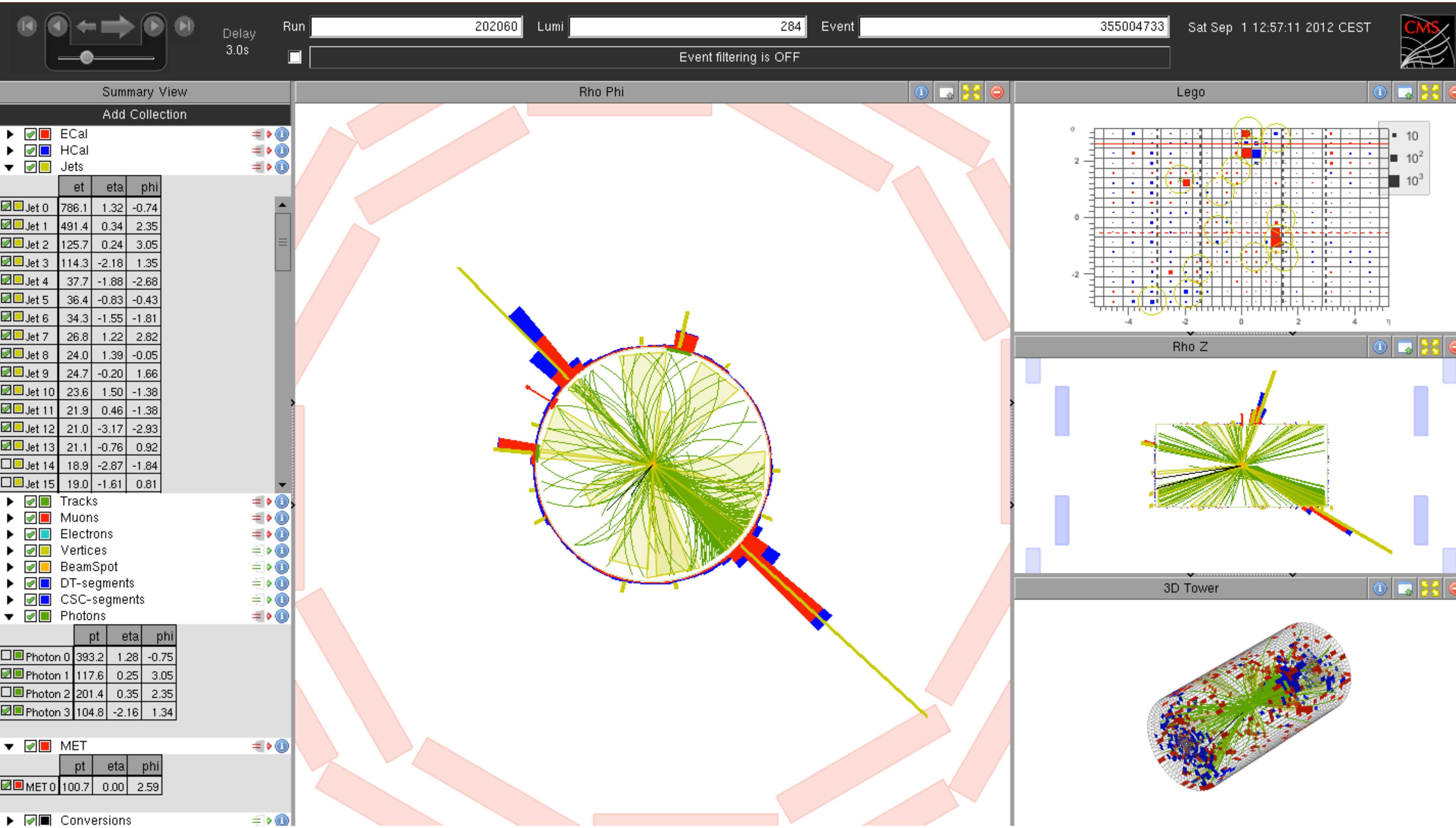
Event	MR	Rsq	MET
1	2320	0.03	360
2	2270	0.021	180
3	2300	0.03	260

Event displays are in the back up. Nothing pathological about the events is observed.

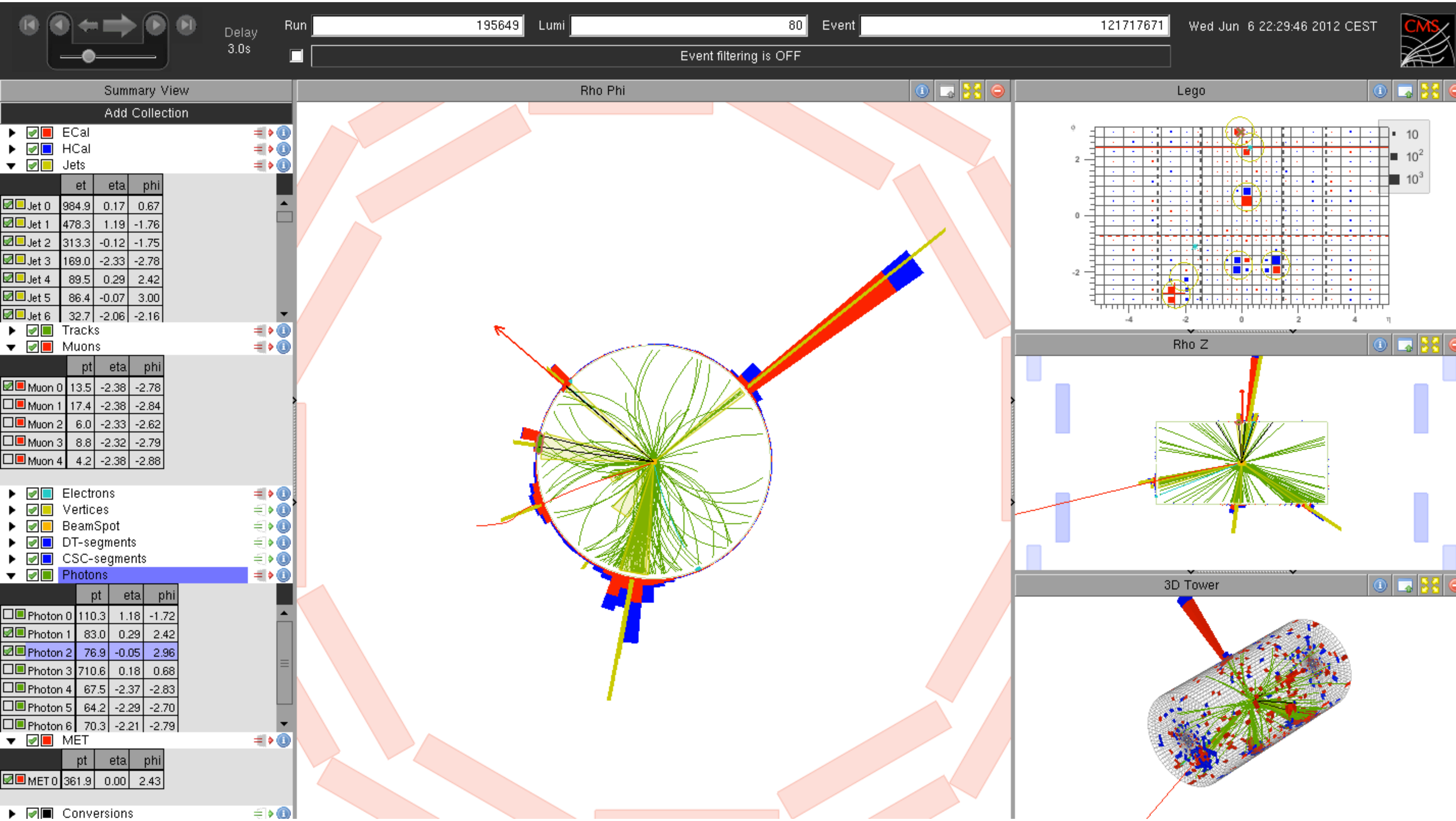
Event Display of Signal Candidate



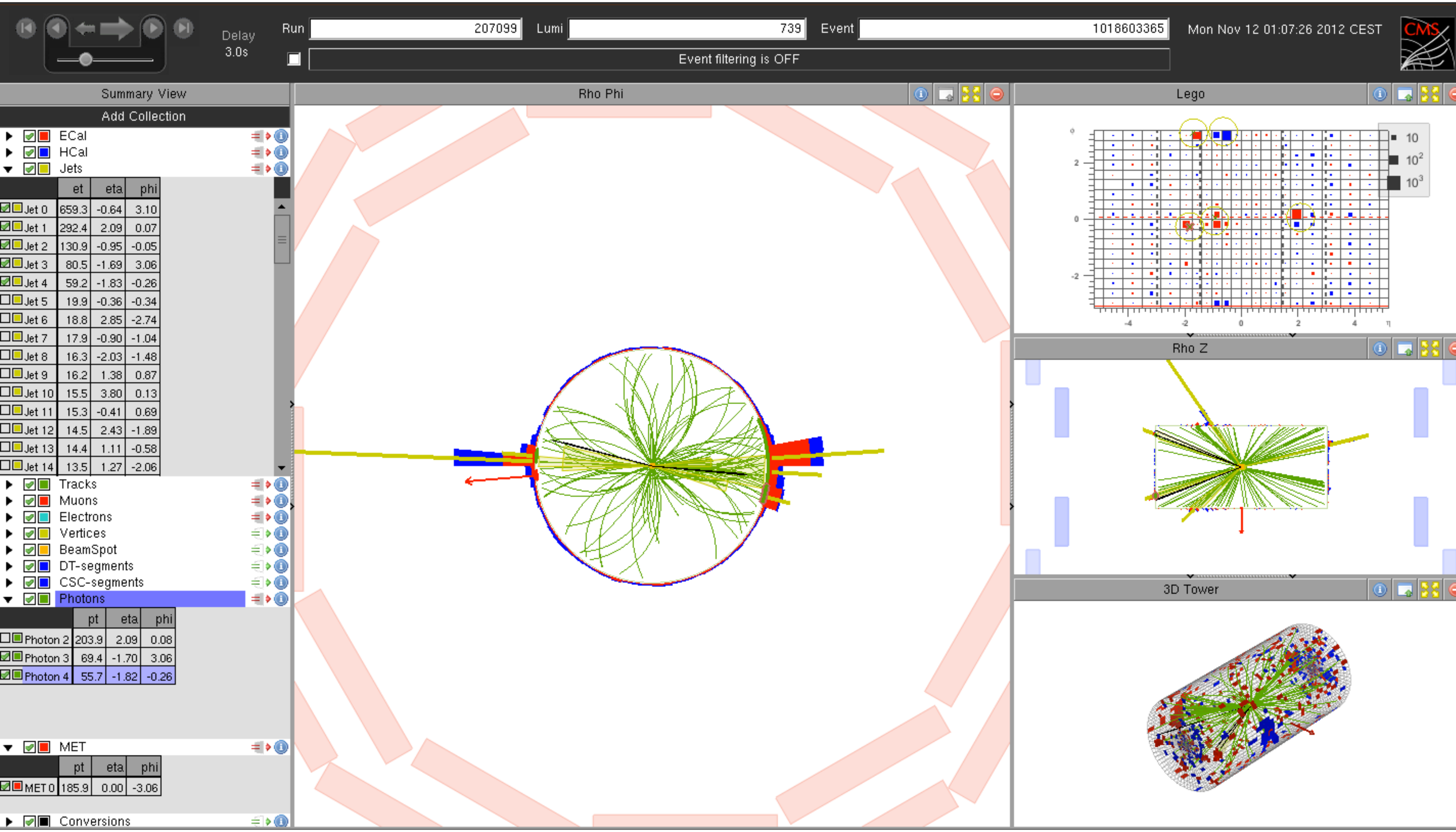
3 jets passing ID — photons 117, 105 —MR 2270 Rsq 0.021 MET 180



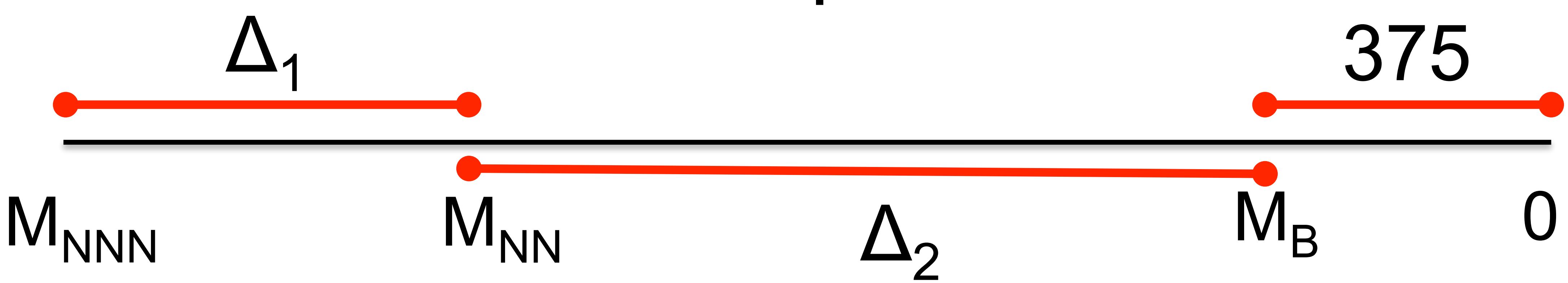
5 jets passing ID — photons 83,77 — MR 2320 Rsq 0.03 MET 360



3 jets passing ID — photons 67,54 — MR 2300 Rsq 0.03 MET 260



Understanding the Mass Splitting For RA3 Sample



M_g	M_s	M_B	Δ_1	Δ_2	M_R peak
1500	1320	375	280	945	1500
1800	1020	375	680	645	1100
2000	1620	375	380	1245	1900
1200	2020	375	820	825	1400

- M_R peaks higher for the processes with the larger splitting: Max(Δ_1, Δ_2)
- M_R peaks at least as high as M_{NN}
- The third split M_{NLSP} → Gravitino is always fixed at 375 GeV