CMS Experience with Delphes

Seth Zenz, Princeton University
For the CMS Collaboration

2nd Fast Monte Carlo Workshop in HEP
DESY, Zeuthen
15 January 2013

Overview

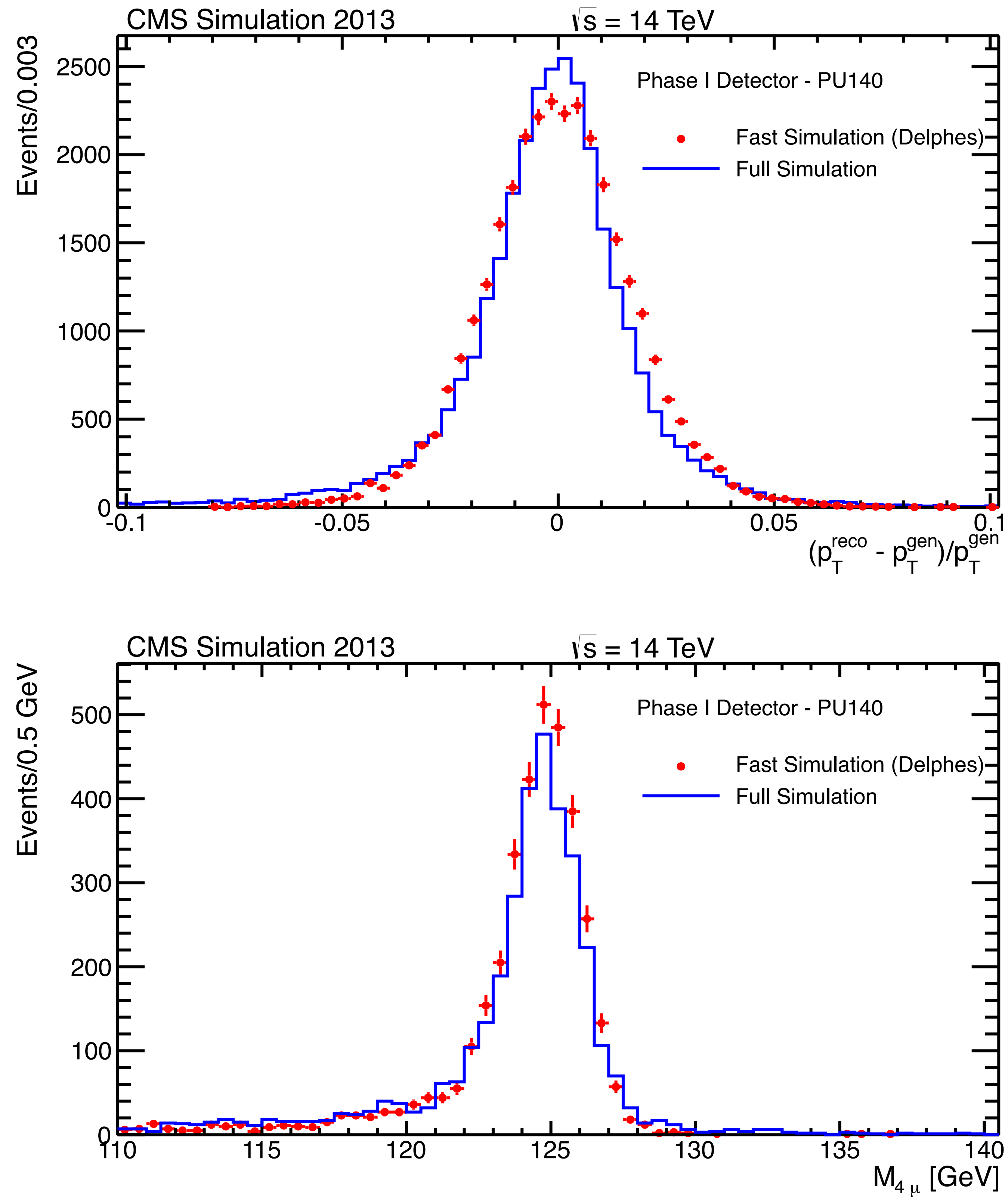
- Why use Delphes?
  - What it's good for
  - What it's not
  - Validation, parameterization outside Delphes
- CMS Physics studies with Delphes
  - Snowmass & ECFA
  - N.B. relatively recent history
- Future plans

Why use Delphes?

Warning: some personal opinion ahead

What Delphes Does Well

- Quantities that arise directly from objects with validated parameterization
- $H \to ZZ \to 4\mu$ (CMS-PAS-FTR-13-003)



What Delphes Does Parametrically

- Ultimately, all particle behavior in Delphes is parameterized, so comes out right only if the parameterizations are correct for the applicable detector environment
- Analysis quantities determined by simple functions of eta and pT:
  - Electron/muon reconstruction efficiency
  - Electron/muon resolution
  - B-tagging and Tau ID efficiency
- Example: changing tracking efficiency and amount of pileup do not impact b-tagging: parameterization must be updated

What Delphes Does Quasi-Realistically

- Individual particles smeared and passed through simple calorimeter model
- Effects dominated by large numbers of hadrons thus modeled quasi-realistically:
  - Jets
  - Missing Energy
  - Object isolation
  - Impact of Pileup on the above (e.g. resolution)
- N.B. Output still likely to require correction and comparison with full simulation!

What Delphes Doesn't Do: Fakes

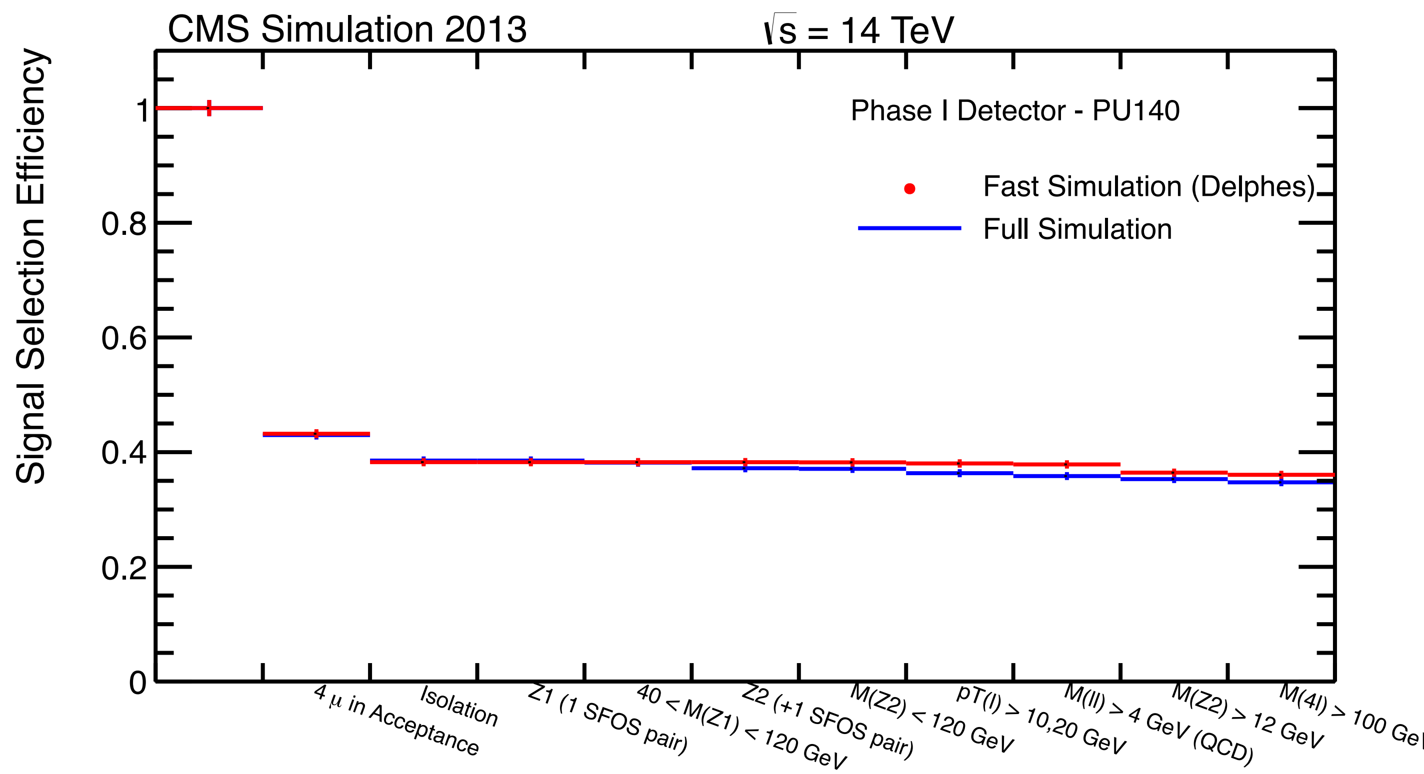
- Example challenge for Delphes analysis: how to optimize object isolation cuts when the background you're trying to exclude doesn't exist?
- How can we do analyses without fakes?
  - Make sure fakes are negligible or estimate and include fakes outside of Delphes
  - Derive realistic isolation efficiencies from Full Simulation or Data, port parameterization to Delphes

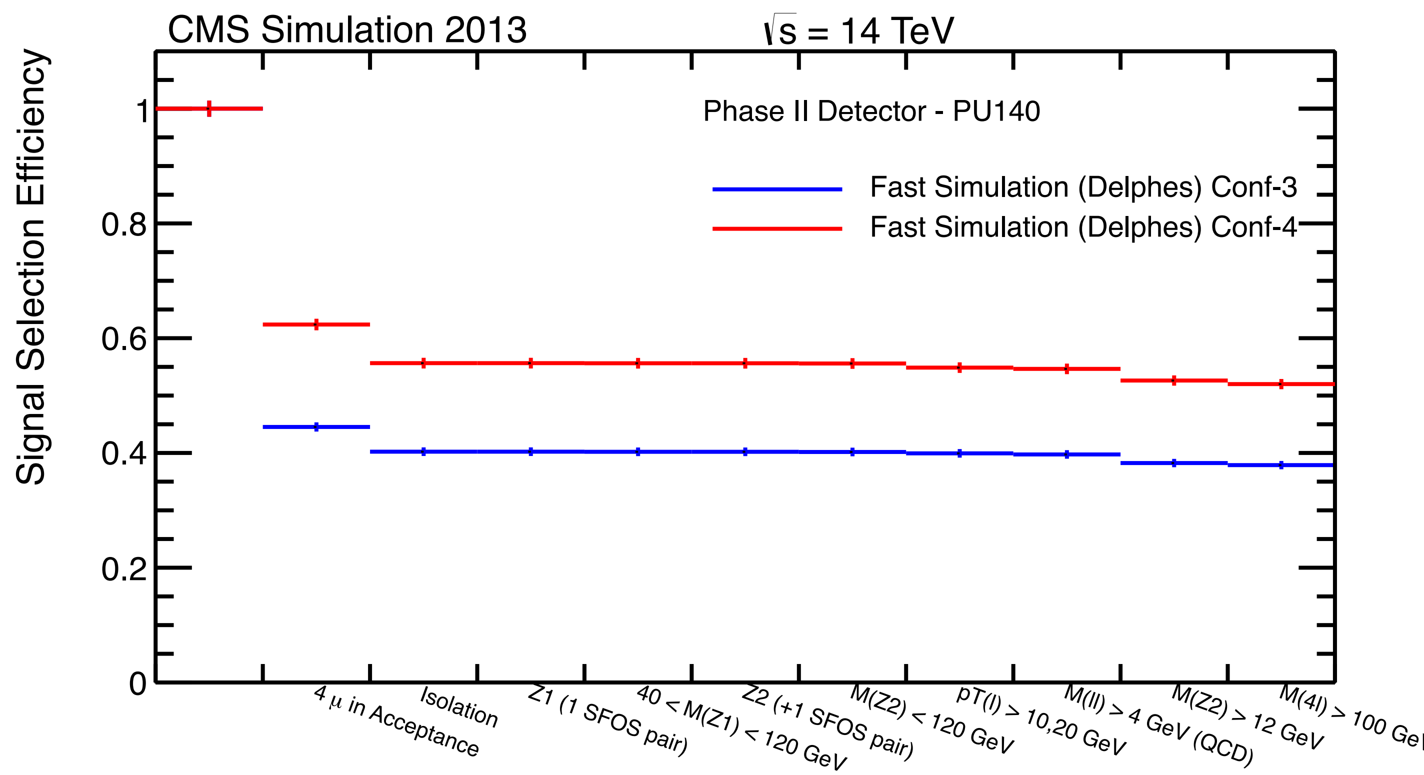
So why use Delphes?

- Full- or ordinary Fast-Sim Monte Carlo samples for future studies may not be available because:
  - New features being added (e.g. timing)
  - Many possible future geometries to be considered, not all can be simulated
  - Quick turnaround required
- Can give good results when:
  - Relevant quantities for analysis are validated with respect to full simulation for some samples and some geometries
  - Extrapolate to other configurations with modest differences
- Flexible configuration system, easy to add modules
  - We add features to Delphes as needed inside CMS!
  - Convenient ROOT output with examples for simple analysis

Extrapolated from validated simulation

- H → ZZ → 4μ
- Tune parameterization and demonstrate consistency, then change configuration
- Lower-right plot: tracker and muon system acceptance increased
  - Conf. 3: |η| < 2.4
  - Conf. 4: |η| < 4.0

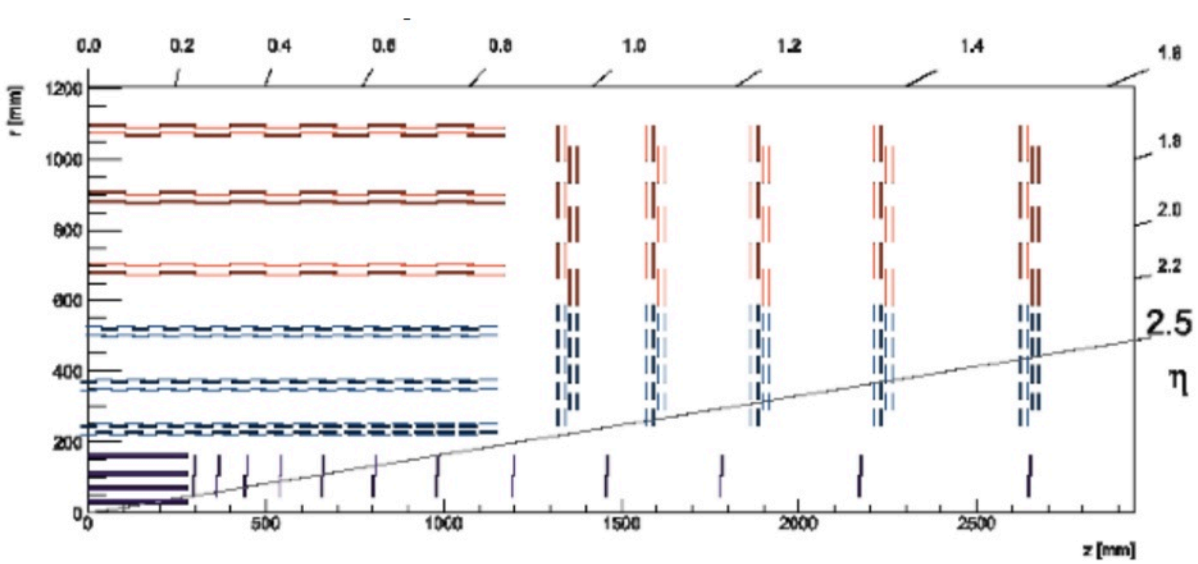




Delphes Studies in CMS

CMS ECFA Production

- Recent CMS Delphes results prepared for ECFA 2013 workshop, building on Snowmass energy-frontier common detector samples
- Complete set of electroweak and Higgs backgrounds targeted to, e.g., di-Higgs and SUSY studies: 300 fb^{-1} and 3000 fb^{-1}



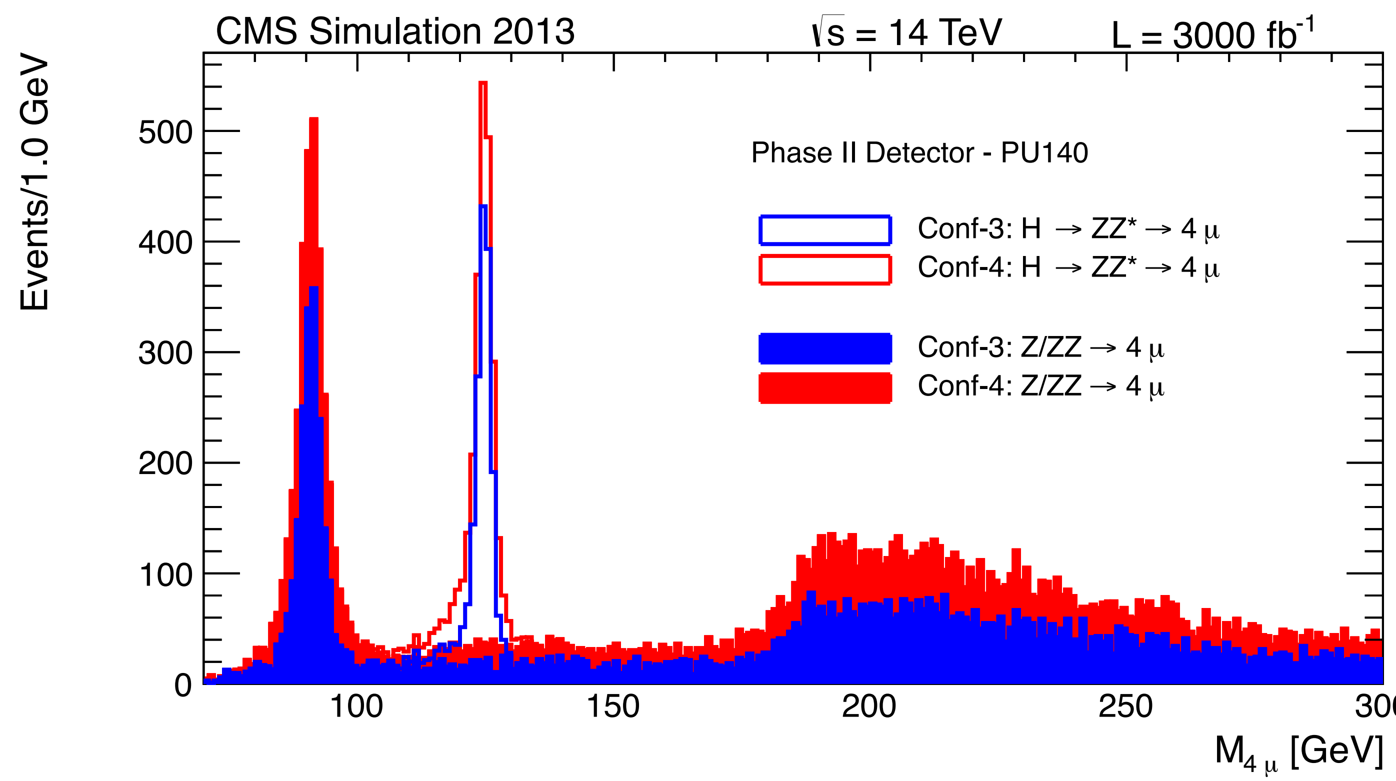
$|\eta| < 2.4 \rightarrow |\eta| < 4.0$?

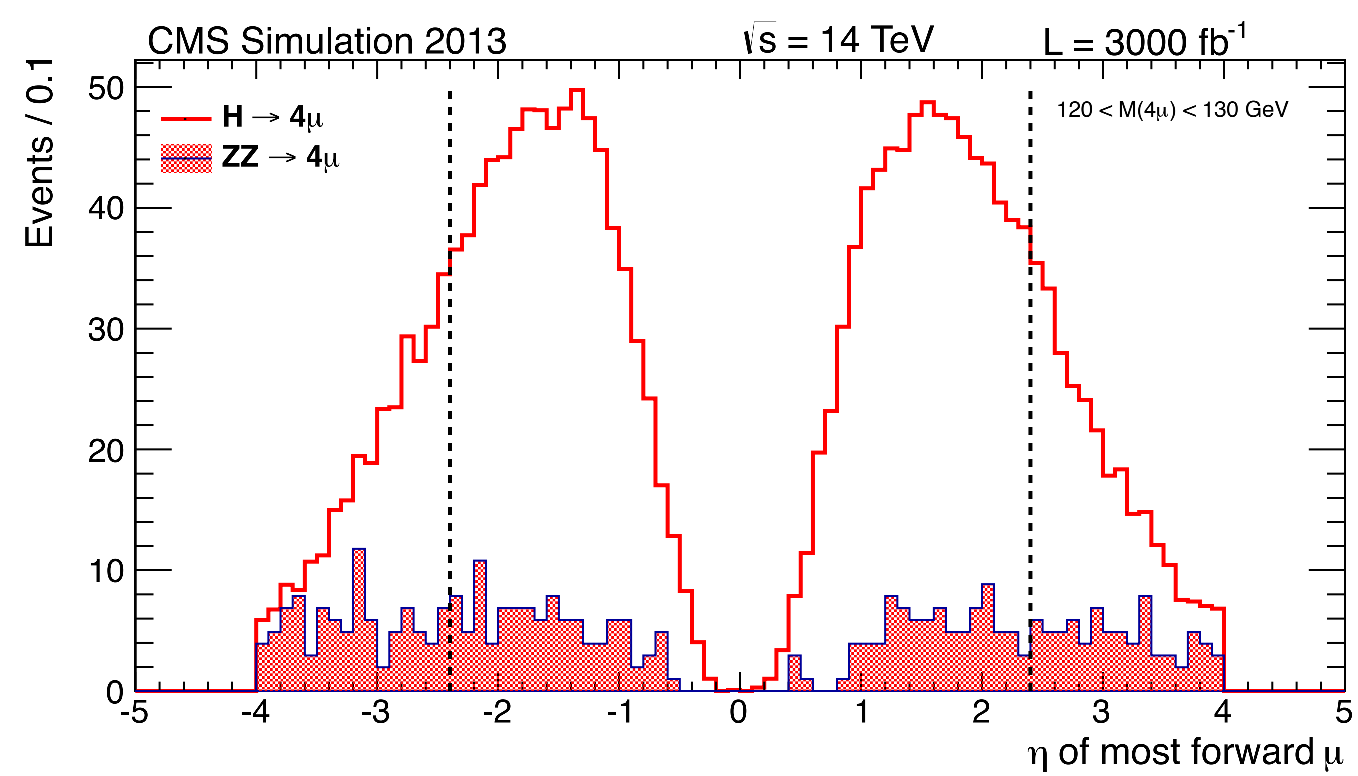
Dataset name	Physics process	Number of recoil jets
B-4p	γ or on-shell W, Z	0
Bj-4p	γ or on-shell W, Z	1-3
Bjj-vbf-4p	γ or off-shell W, Z, H in VBF topology	2-3
BB-4p	Diboson (γ, W, Z) processes	0-2
BBB-4p	Tri-boson (γ, W, Z) processes including BH	0-1
LL-4p	Non-resonant dileptons (including neutrinos) with $m_{ll} > 20$ GeV	0-2
LLB-4p	Non-resonant dileptons with an on-shell boson, $m_{ll} > 20$ GeV	0-1
H-4p	Higgs	0-3
tj-4p	Single top (s- and t-channel)	0-2
tB-4p	Single top associated with a boson	0-2
tt-4p	$t\bar{t}$ pair production	0-2
ttB-4p	$t\bar{t}$ associated with γ, W, Z, H	0-1

Table 1-2. *Table of background processes. All processes include the particles in the dataset name plus additional recoil jets up to four generated particles. On-shell vector bosons, off-shell dileptons, Higgs bosons, top quarks, and jets are denoted B, LL, H, t, and j, respectively. In the Bjj-vbf-4p case, B includes Higgs. In the BBB-4p case, BBB includes BH. Samples are generated in bins of H_T^* for $\sqrt{s} = 14$, 33, and 100 TeV.*

H → ZZ → 4μ Results

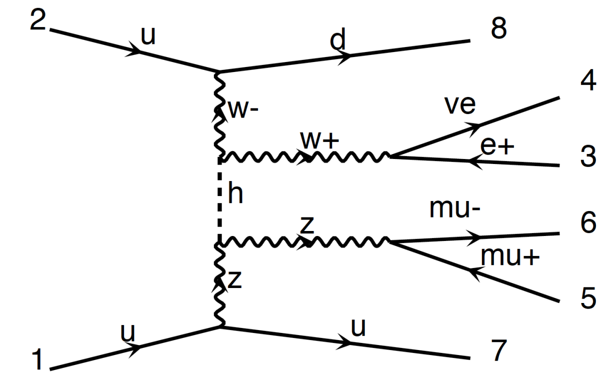
- p_T > 20,10,5,5 GeV
- Form opposite-charged pairs
- Higgs mass requirements [GeV]
  - 40 < m_{Z1} < 120
  - 12 < m_{Z2} < 120
- Selection efficiency increase > 40% with extended tracker





VBF + Quartic Gauge Coupling (1)

- Search for anomolous Quartic Gauge Couplings in pp → ZWjj
  - Require 3 leptons with pT > 20 GeV, $|\eta| < 2.4$, well-separated from each other and from jets
  - Z mass window
  - $d\eta_{jj} > 4.0$, $m_{jj} > 600$ GeV
- CMS-PAS-FTR-13-006



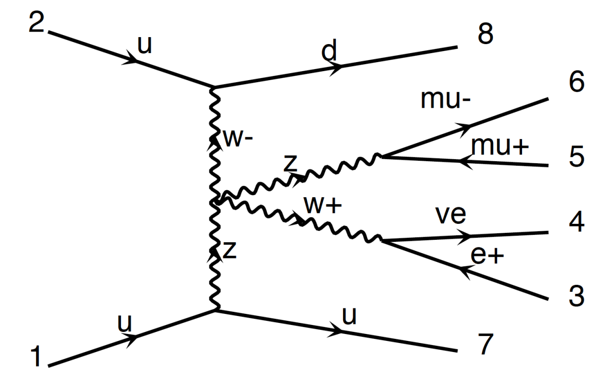
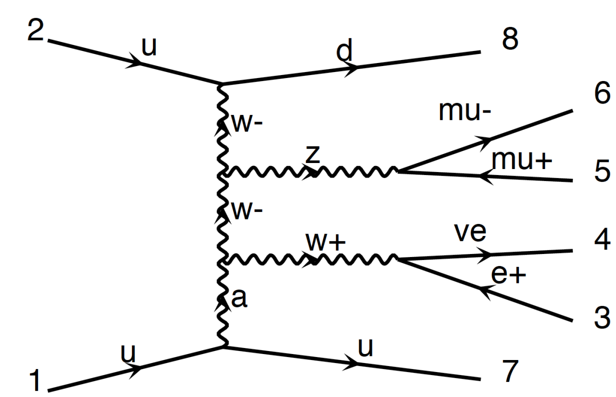


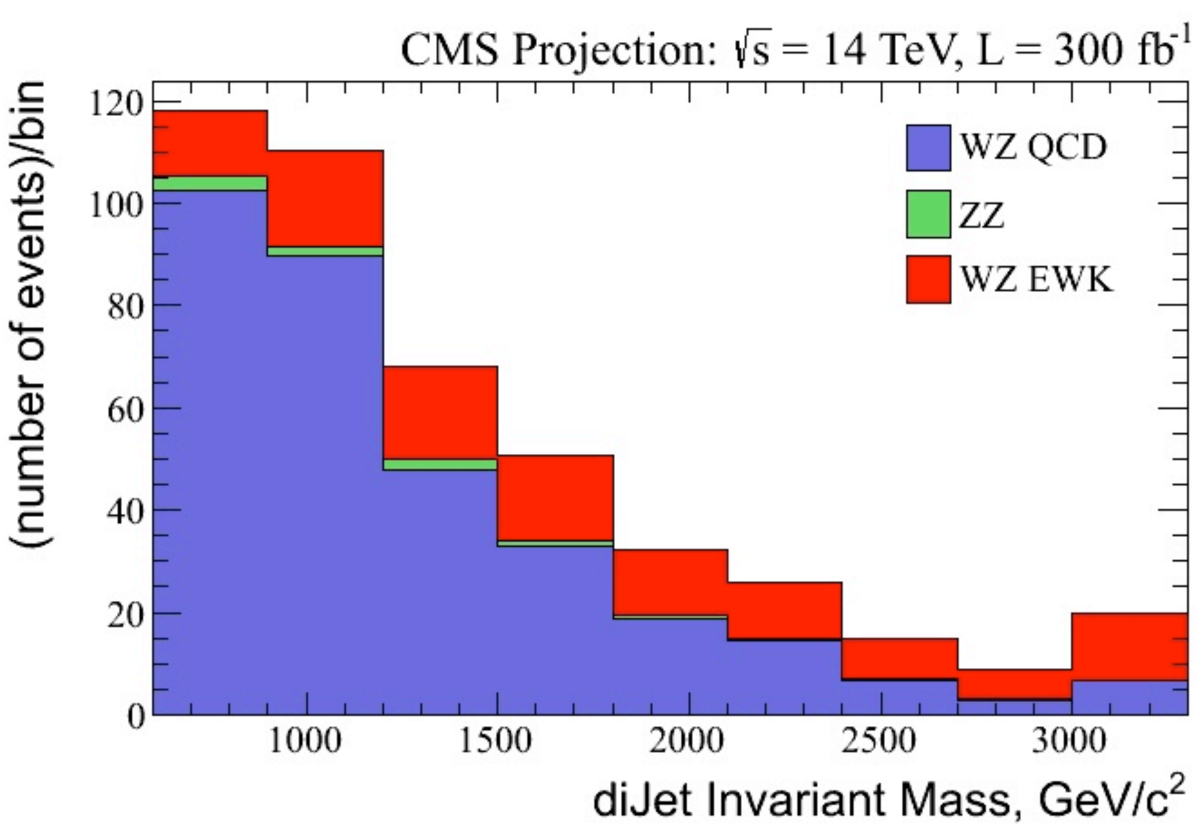
Table 1: Cross sections in fb at 14 TeV. The L_{T1} cross section is the additional cross section compared to the SM rate for an aQGC with $f_{T1}/\Lambda^4 = 1.0$.

	WZ EWK	WZ QCD	ZZ	L_{T1}
Total	7.7	270	16	3.1
Fiducial	0.69	0.96	0.038	0.57

$$L_{T1} = (f_{T1}/\Lambda^4) Tr[\hat{W}_{\alpha\nu}\hat{W}^{\mu\beta}] Tr[\hat{W}_{\mu\beta}\hat{W}^{\alpha\nu}]$$



VBF + Quartic Gauge Coupling (2)



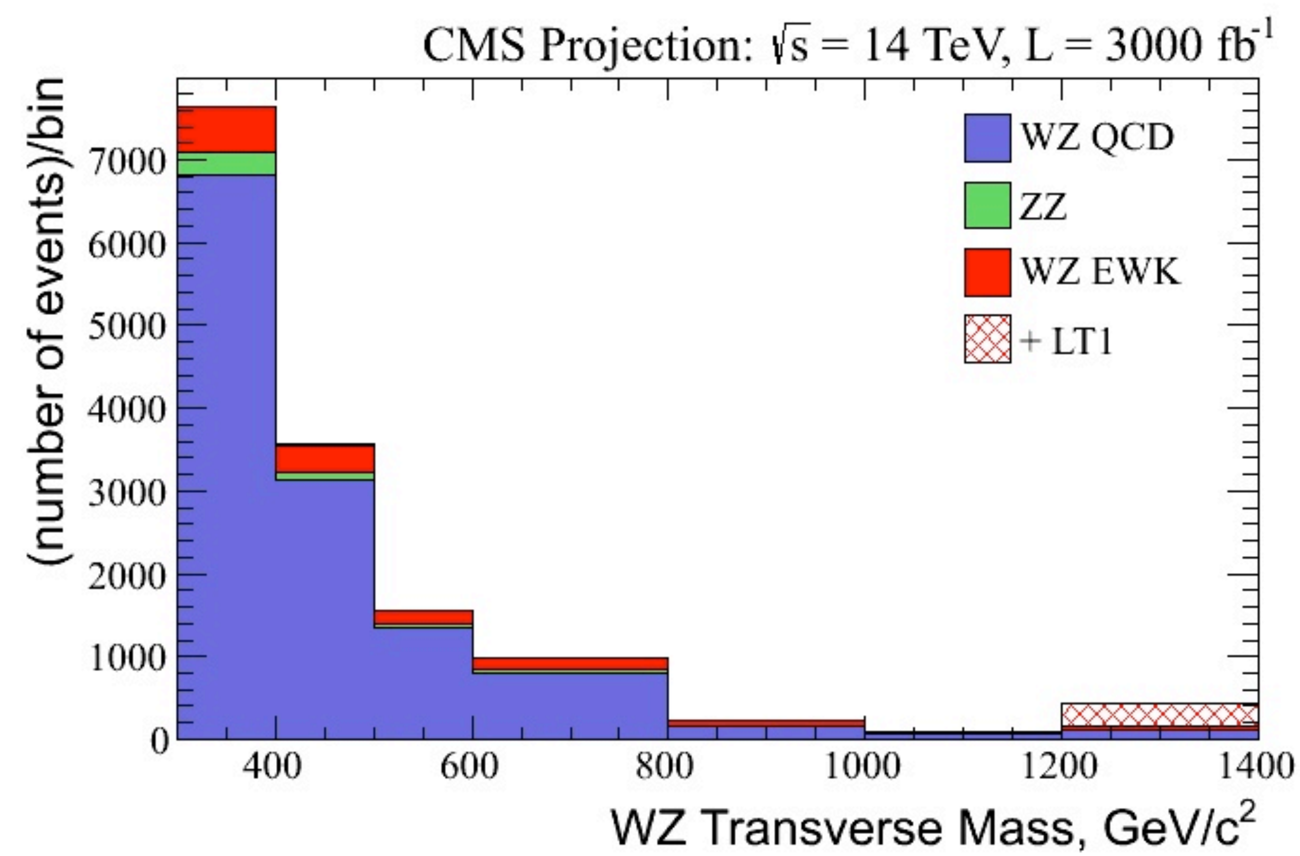
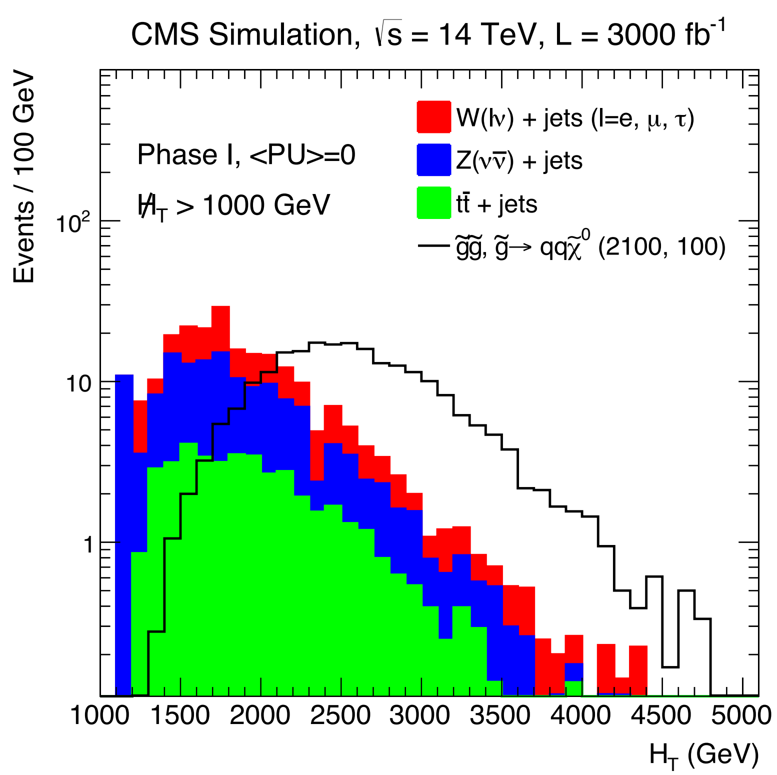
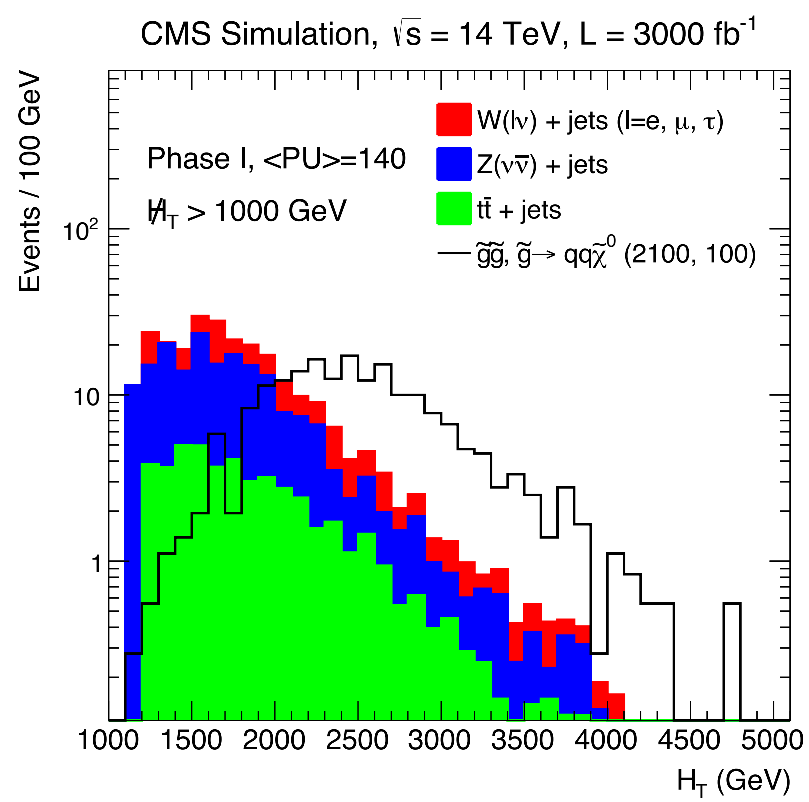


Table 2: Sensitivities for SM EWK scattering discovery and aQGC. The integrated luminosities for SM EWK discovery at 3σ and 5σ are reported while aQGC prospects for discovery are given in terms of the L_{T1} operator coupling constant f_{T1}/Λ^4.

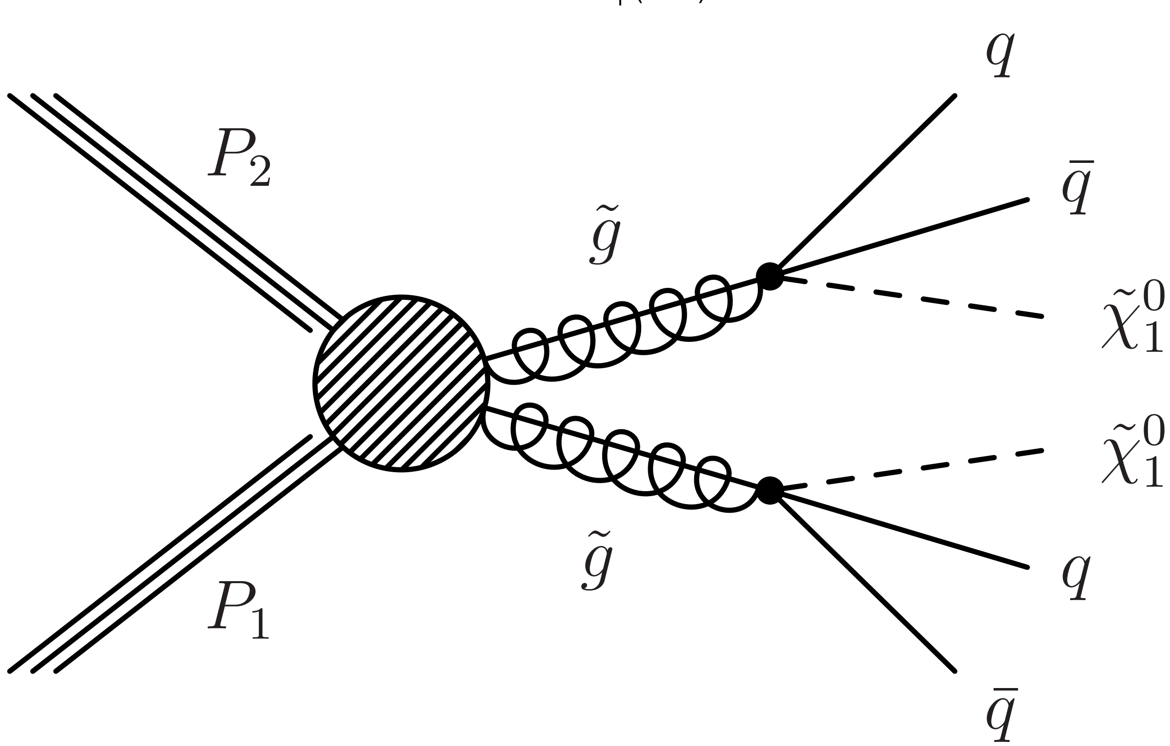
Significance	3σ	5σ
SM EWK scattering discovery	75 fb^{-1}	185 fb^{-1}
f_{T1}/Λ^4 at 300 fb^{-1}	0.8 TeV^{-4}	1.0 TeV^{-4}
f_{T1}/Λ^4 at 3000 fb^{-1}	0.45 TeV^{-4}	0.55 TeV^{-4}

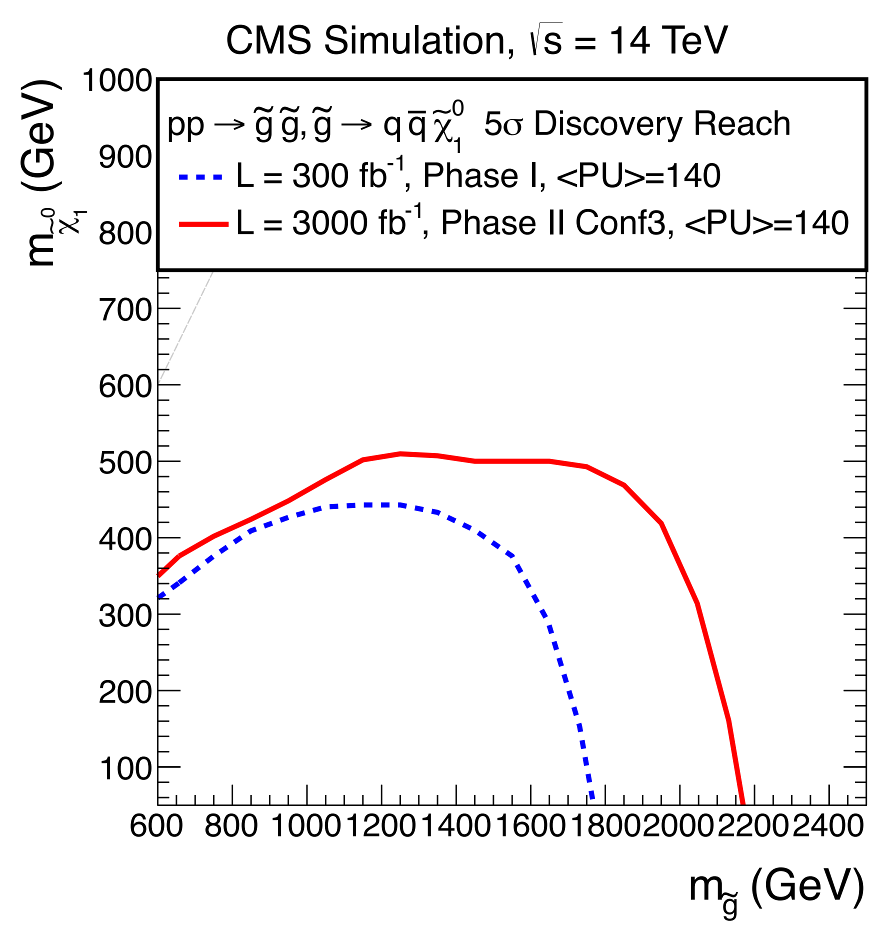
SUSY Searches: Jets + MET



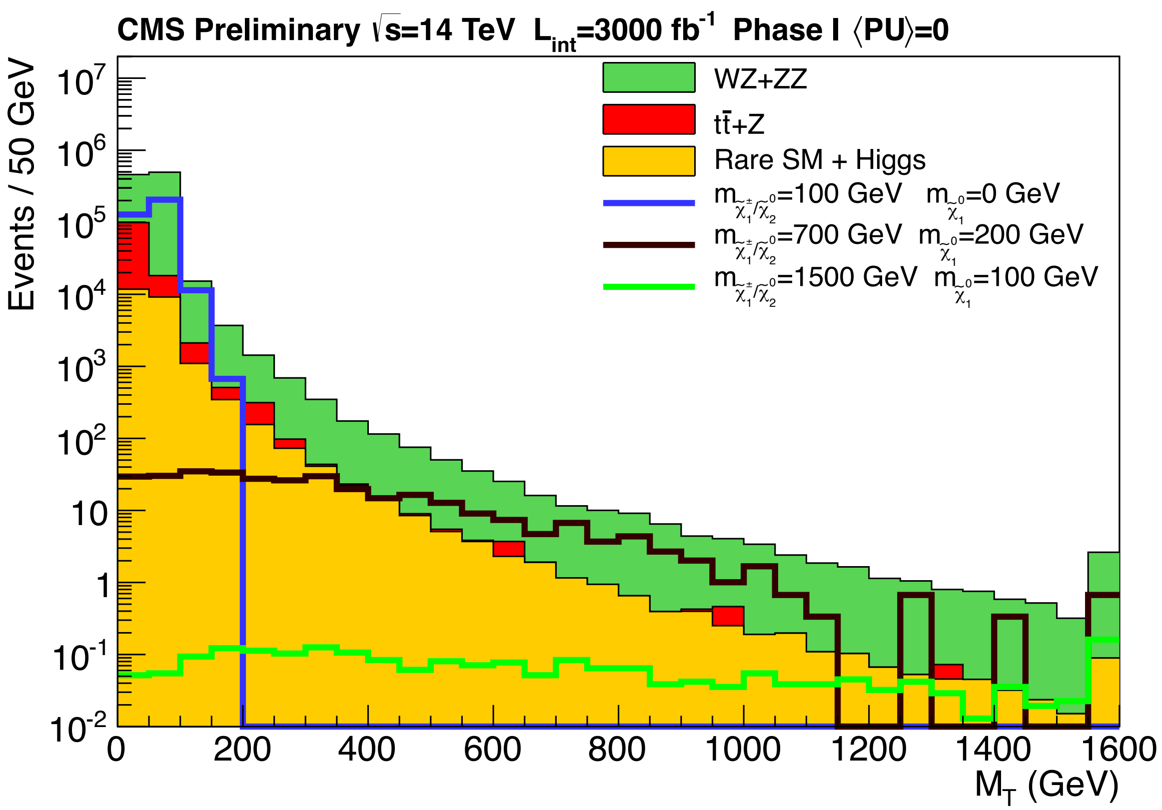


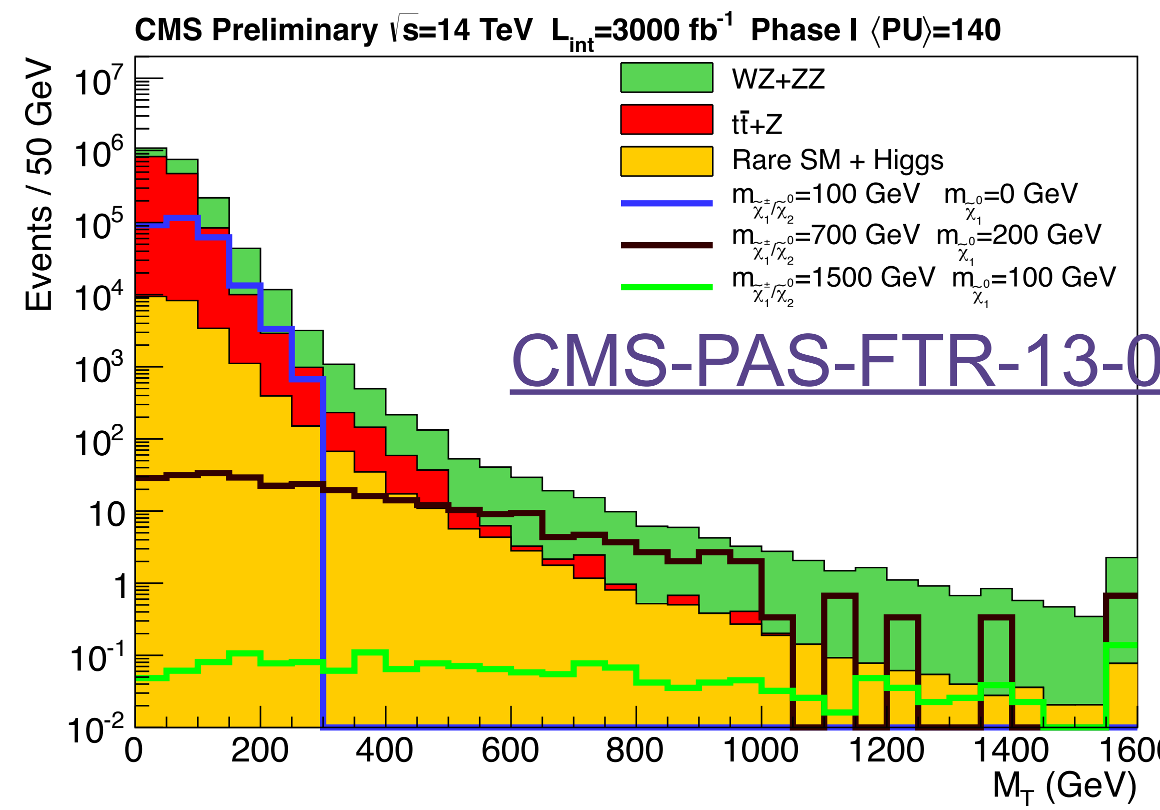
CMS-PAS-FTR-13-014



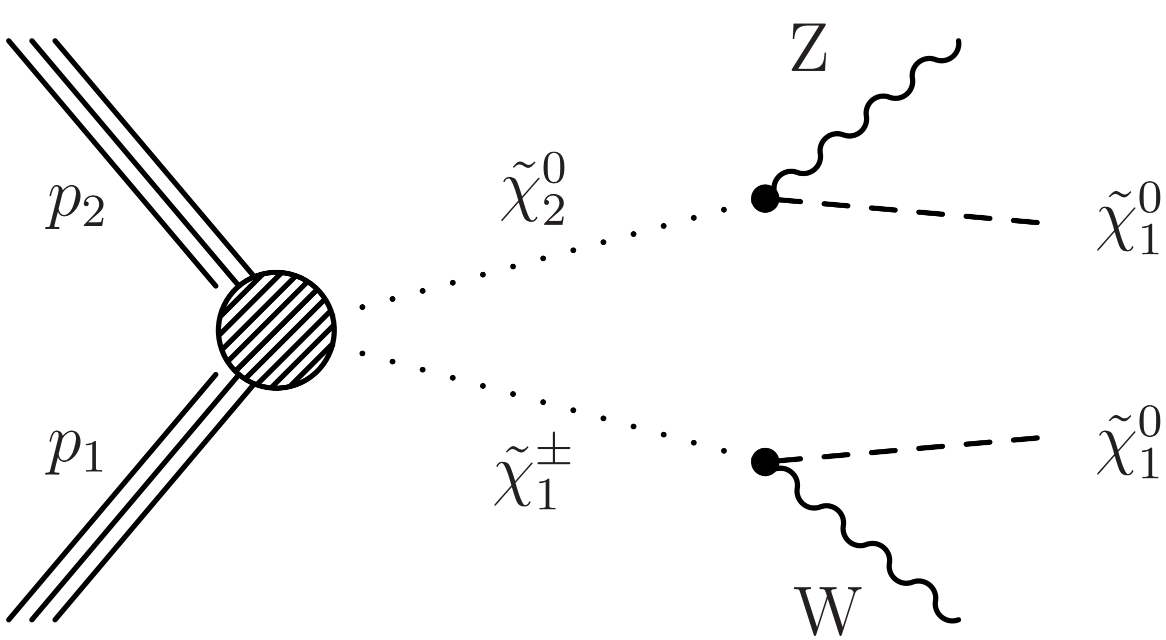


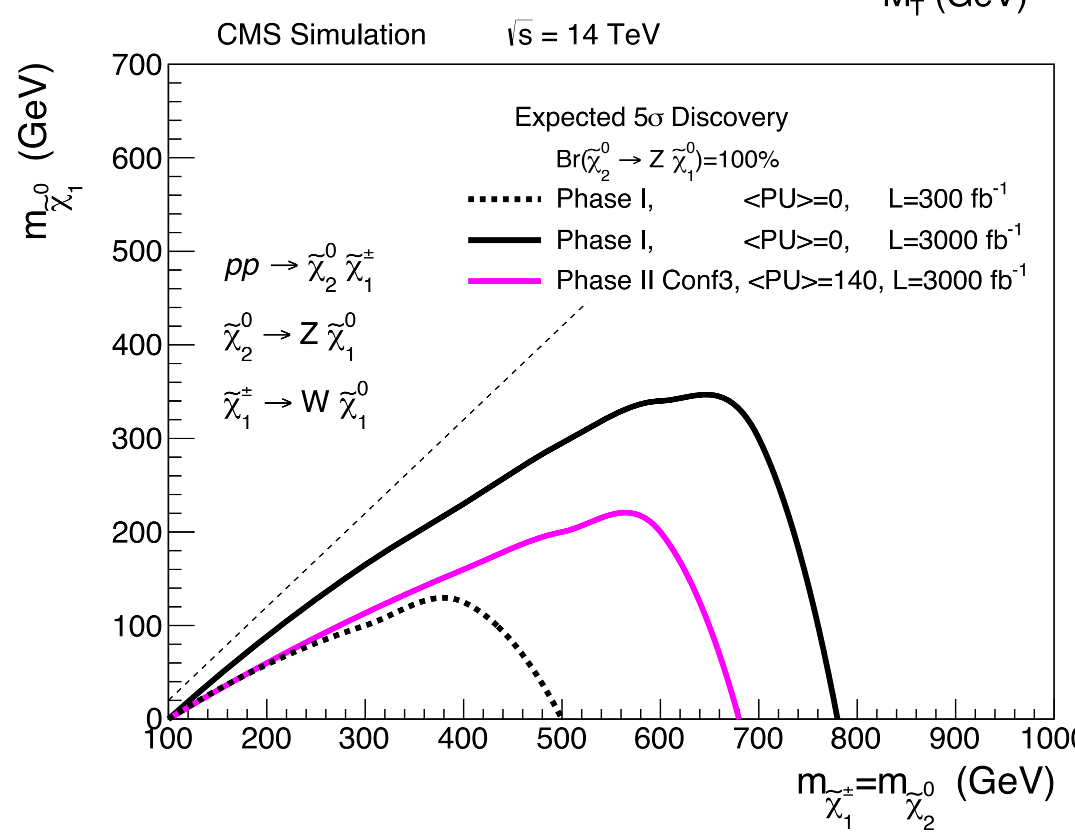
SUSY Searches: EWKino → W+Z+MET



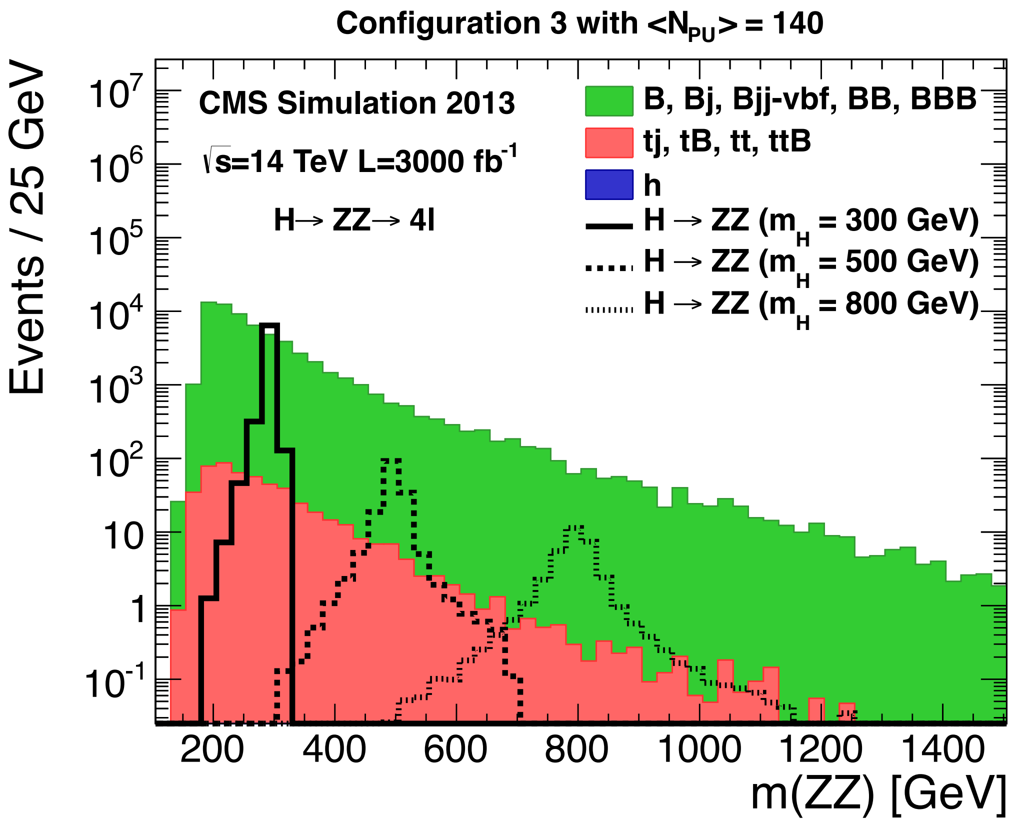


CMS-PAS-FTR-13-014

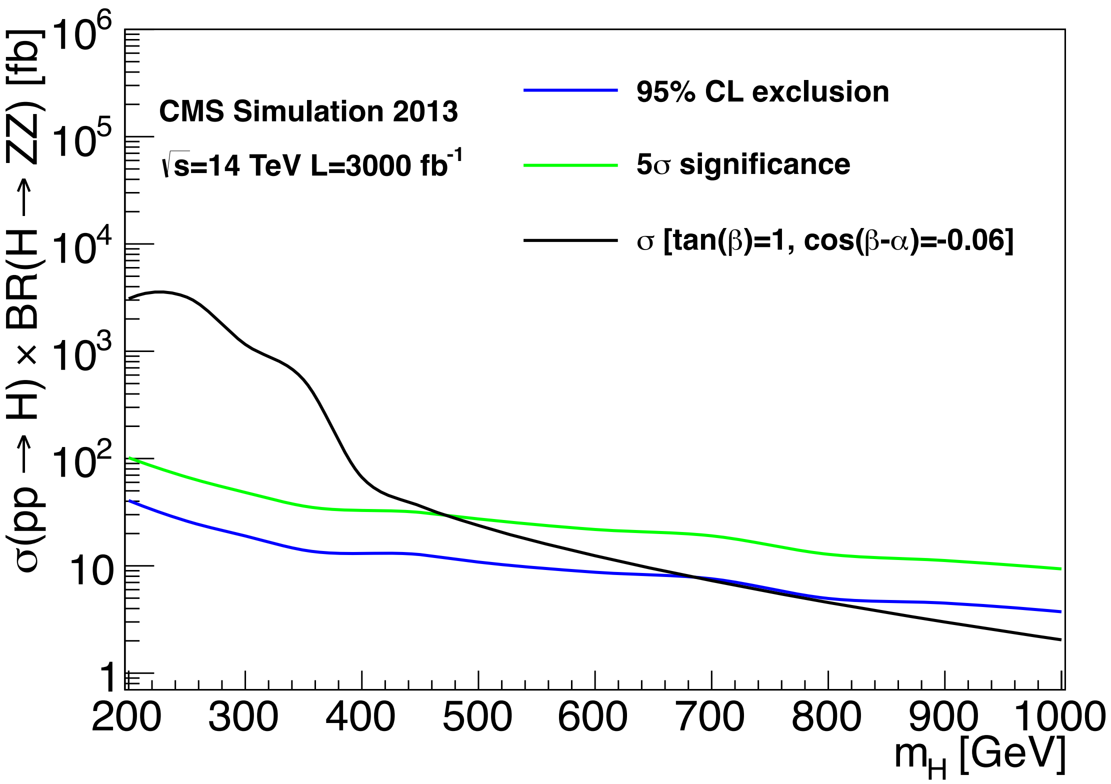


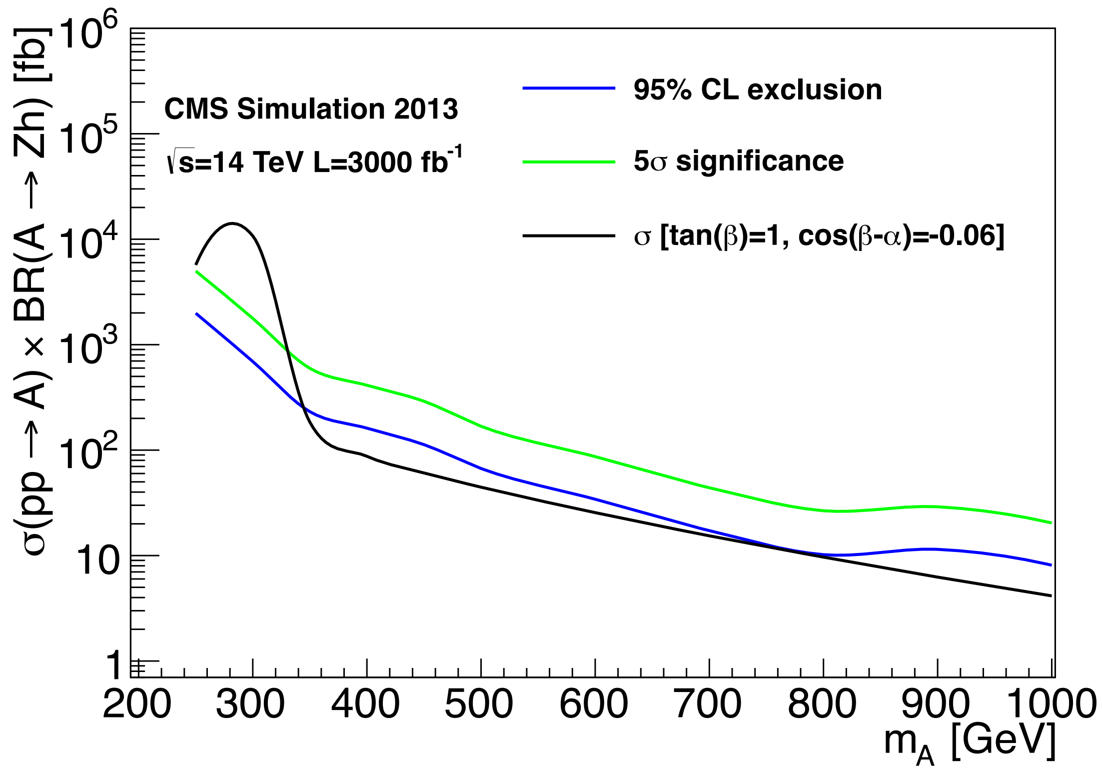


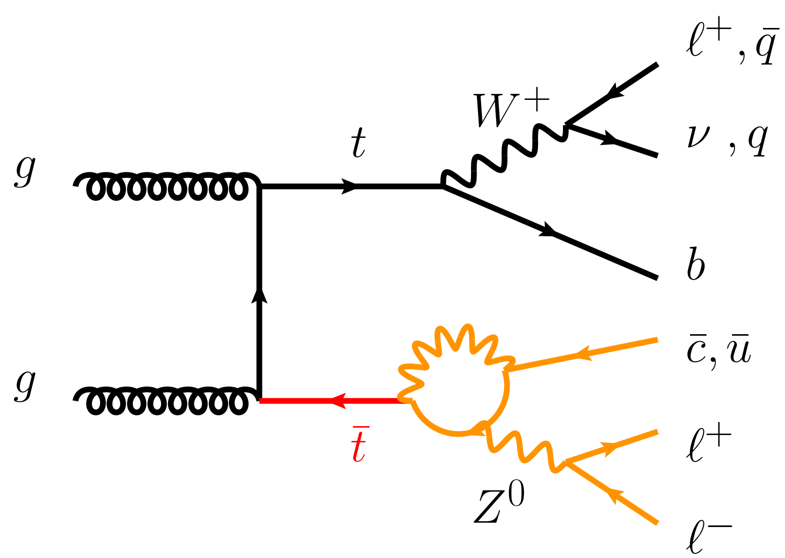
2 Higgs Doublet Models



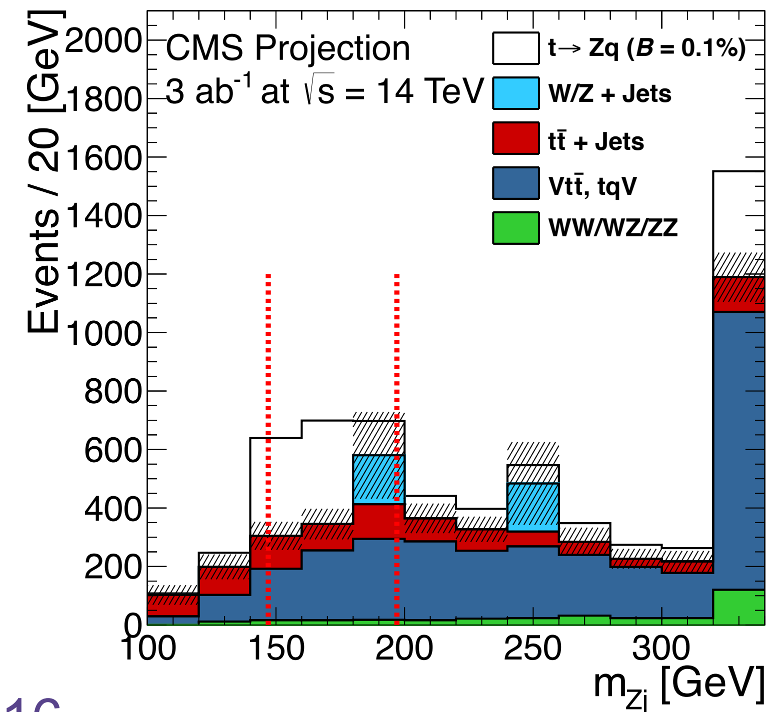
- CMS-PAS-FTR-13-024
  - $H \to ZZ \to 4\ell$
  - $A \to Zh \to \ell\ell bb$







FCNC signal	Events	Efficiency
All events	663313	—
At least two leptons	323051	48.7%
One Z	246968	37.2%
Three leptons	84780	12.8%
$\not\!E_T > 30$ GeV	77115	11.6%
At least two jets	20444	3.08%
One b-tagged jet	9241	1.39%
Top mass	2036	0.31%
Expected at 3000 fb^{-1}	578 ± 13	



Process	Cross section (pb)	Order	Expected @ 3000 fb^{-1}
top FCNC signal	1.91	NNLL	578 ± 13
vector boson + jet	338×10^3	NLO	< 55
di-boson + jets	412	NLO	40 ± 31
top pair + jets	954	NNLL	70 ± 57
single top + jets	323	NLO	< 8
boson + top + jets	97.3	NLO	15 ± 15
boson + top pair + jets	3.97	NLO	144 ± 14

CMS-PAS-FTR-13-016

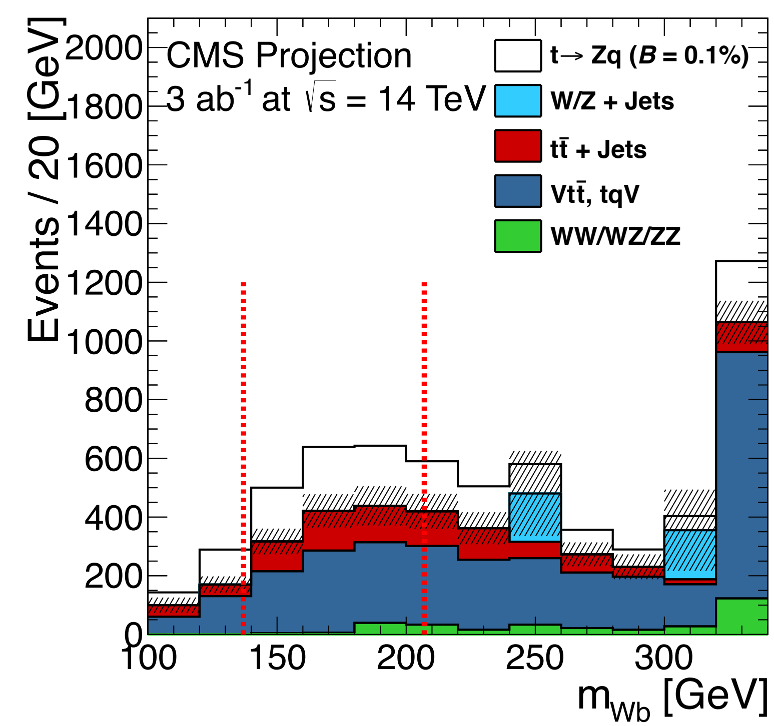


Table 5: Upper limits at the 95% CL for $\mathcal{B}(\mathrm{t} \to \mathrm{Zq})$, including the results obtained in 8 TeV data, which corresponds to an integrated luminosity of 19.5 fb^{-1} and the projections made for 14 TeV with an integrated luminosity up to 3000 fb^{-1} using DELPHES samples.

$\mathcal{B}(\mathrm{t} \to \mathrm{Zq})$	19.5 fb^{-1} @ 8 TeV	300 fb^{-1} @ 14 TeV	3000 fb^{-1} @ 14 TeV
Exp. bkg. yield	3.2	26.8	268
Expected limit	$< 0.10\%$	$< 0.027\%$	$< 0.010\%$
1 σ range	$0.06 - 0.13\%$	$0.018 - 0.038\%$	$0.007 - 0.014\%$
2 σ range	$0.05 - 0.20\%$	$0.013 - 0.051\%$	$0.005 - 0.020\%$

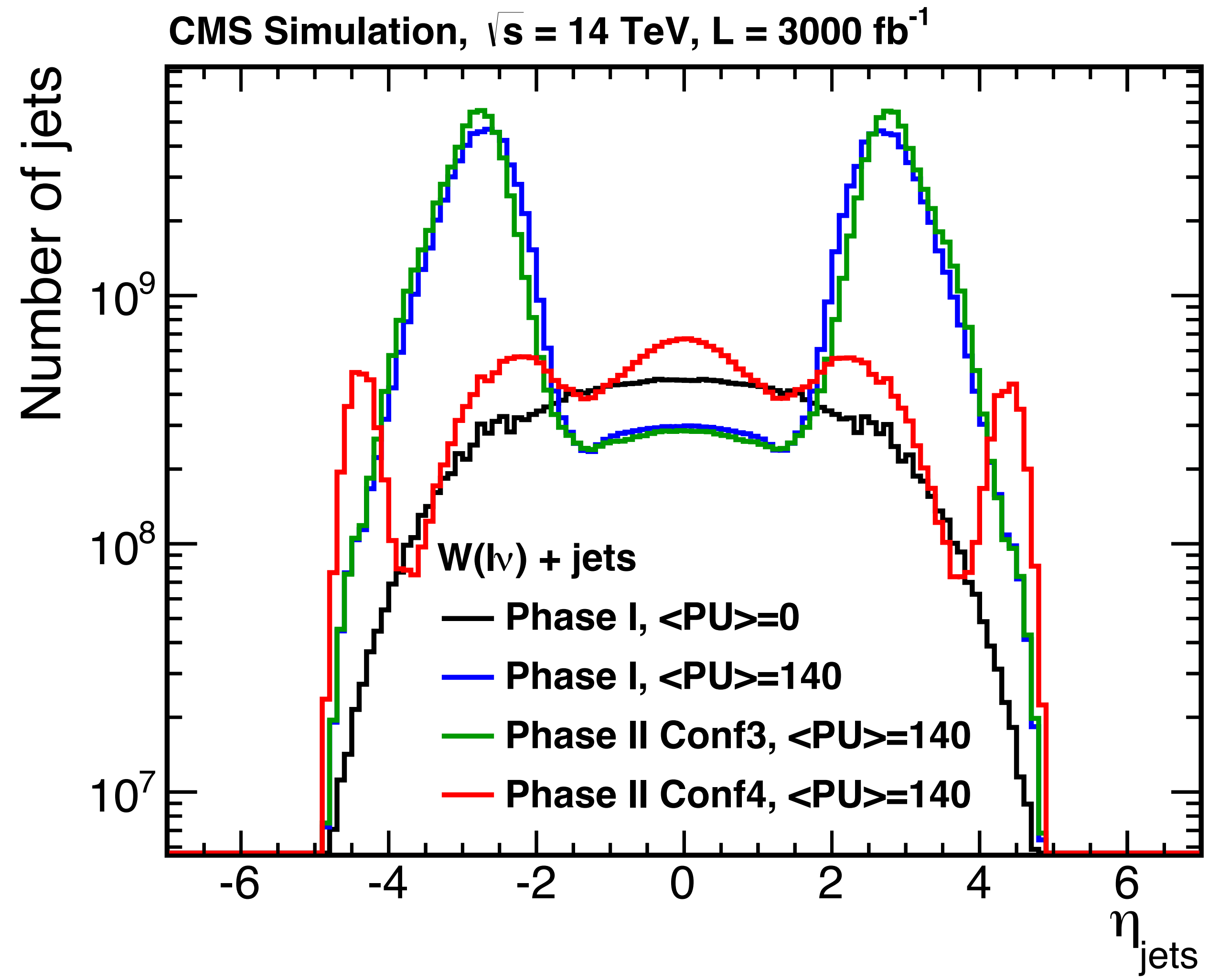
Future Plans

Planned Delphes Updates

- Next round of Delphes studies being planned, and will likely play some role in decision-making for Phase II Technical Proposal foreseen for this year
- Better tools, particularly for jets
  - Pileup Jet ID
  - Improved jet energy corrections for pileup
  - Timing information
- Goals: final conclusions for existing studies, new studies where Delphes will give valuable first-pass results
- Scope of production still under discussion in CMS

Rho Corrections for Jets

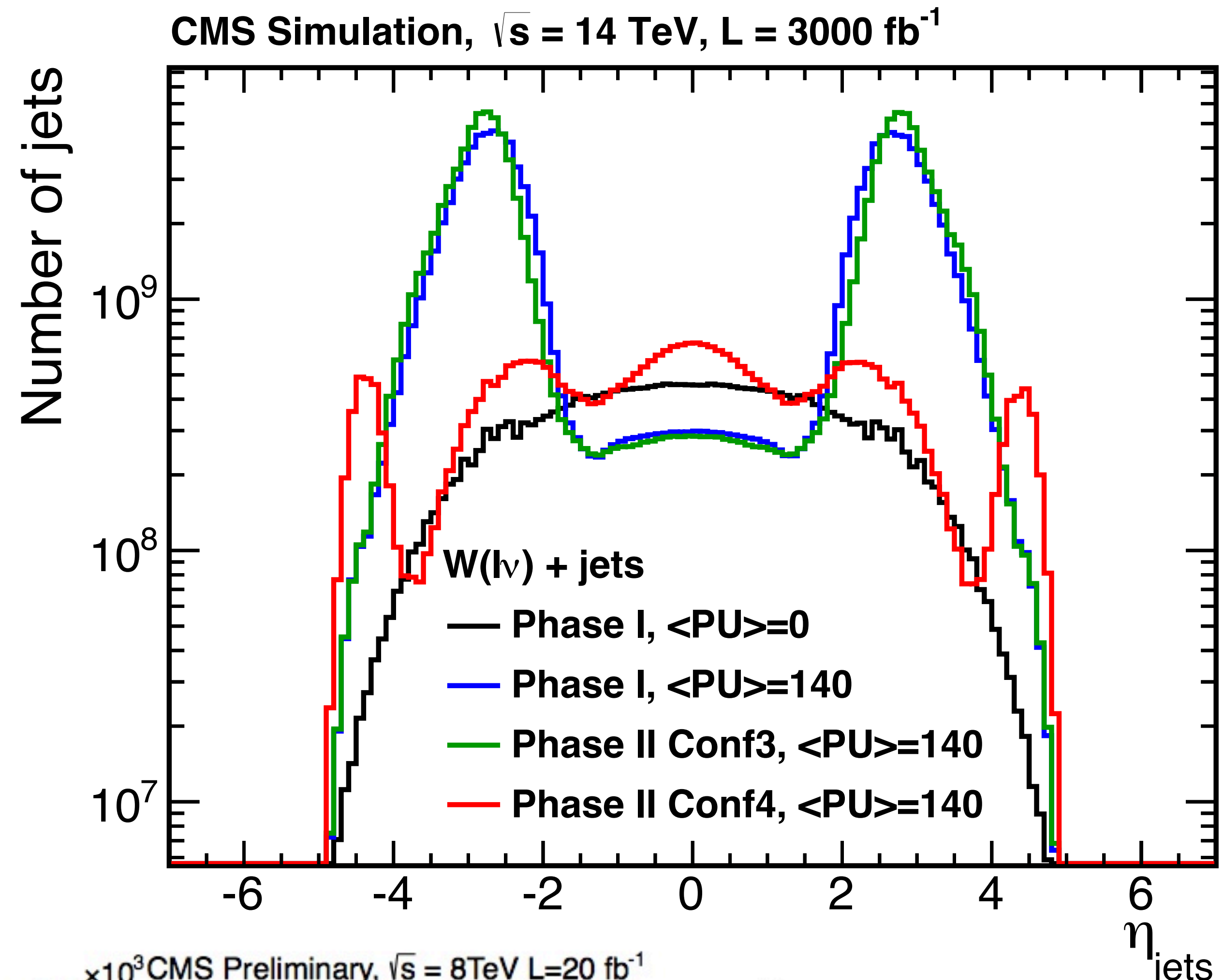
- Delphes versions up to 3.0.10 (including ECFA studies) calculated ρ*A subtraction for jets using constituents from entire detector volume
- Problem: pileup tracks included only for |η| > 2.5

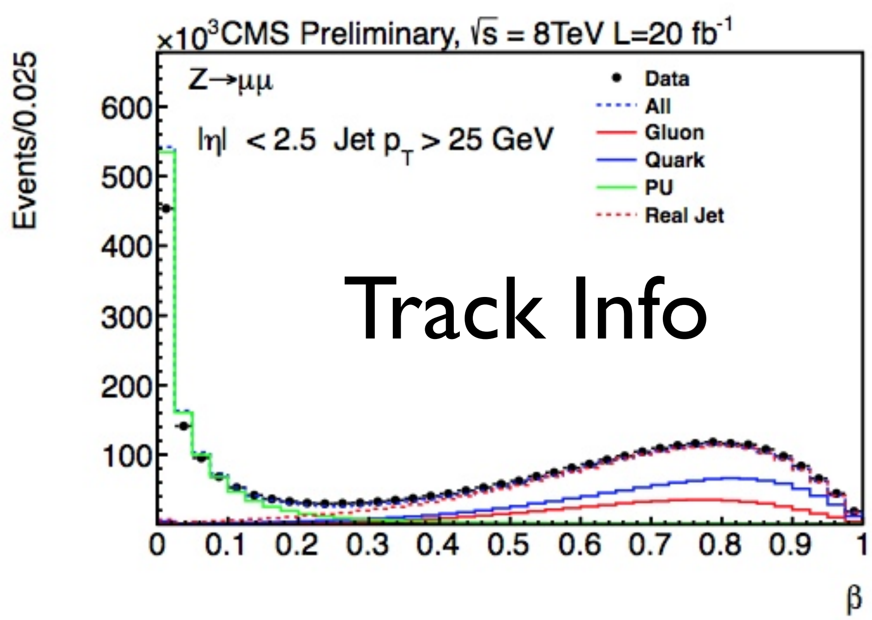


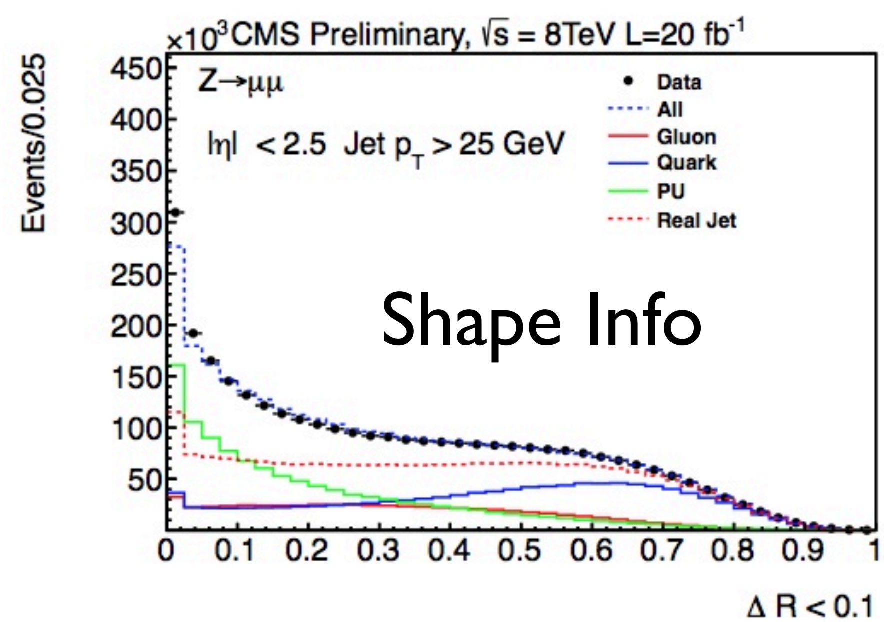
- Central jet energy overcorrected, high-rapidity jet energy undercorrected
- Delphes 3.0.11: ρ can be calculated in bins of eta
- Use of CMS standard binning/parameterization under study

PileUp Jet ID

- CMS has a toolkit for identifying pileup-like jets
  - Track information
  - Shape information
  - PAS-JME-13-005
- Now implemented in Delphes







Calorimeter Model Development

- Being worked on by Delphes authors based on discussion with CMS jet experts: update of calorimeter model
- Current: one grid of "calorimeter cells" with EM & hadronic energy values based on particles hitting that cell, simple model of PF in which electrons and muons assumed to be reconstructable and "subtracted" from cells before PF objects produced
- New features: parameterizations of charge sharing, showering and depth segmentation, different cell sizes for ECAL and HCAL grids
- More sophisticated version of energy flow will be required on the CMS side to take advantage of multiple segmentation
- Illustrates a general question: how far should we go in adding "realism" to parameterized simulation?

Timing Information

- Use of timing information, e.g. from ECAL cells, is under study for use in pileup rejection after Phase II Upgrade
- Modifications to Delphes made to set vertex timing, extrapolate particle time to calorimeter surface, smear appropriately, and attach information to calorimeter cells and jets
- Very coarse estimate – but this feature is not yet available in full or fast simulation, so this will be the *first* information we have on what timing information can add

Development in CMS ↔ Delphes authors

- Significant development of Delphes features, modules, and configurations inside CMS
- Tendency for "CMS version" of Delphes to diverge from official releases!
- Regular contact and exchange of ideas with Delphes authors is critical
  - Authors can develop broadly-needed features and complex core changes
    - Eta-dependent rho corrections (3.0.11)
    - Multiple calorimeter grids (in progress)
  - Some CMS tools may be ported into the release, e.g. PileUp Jet ID (in progress) – others CMS-specific, e.g. timing info
- Is "full synchronization" of Delphes versions necessary/possible/desirable?

Conclusions

- Delphes is a fast, easily-configured tool
- Valuable for studies when:
  - Large samples are needed fast
  - Parameterization of relevant quantities can be validated against full simulation
- Significant use of Delphes for ECFA 2013 results
- Ongoing Delphes development in CMS is targeted to studies where new information can be gained before Full/Fast simulation is available, e.g. jet timing
- Communication with Delphes authors very important to avoid duplicated effort, use latest features, and keep code "relatively" synchronized

Extras

PileUp Jet ID Variables

- Variables documented in PAS-JME-13-005
- Track-dependent variables
- Shape variables
- MVA overkill for Delphes?

- β
- β^*
- d_Z
- $n_{vertices}$

$$\beta = \frac{\sum_{i \in PV} p_{Ti}}{\sum_i p_{Ti}}$$

$$\beta^* = \frac{\sum_{i \in other PV} p_{Ti}}{\sum_i p_{Ti}}$$

- $\langle \Delta R^2 \rangle$
- $A < (\Delta R) < A + 0.1$
- $N_{charged}$
- $N_{neutrals}$
- p_T^D

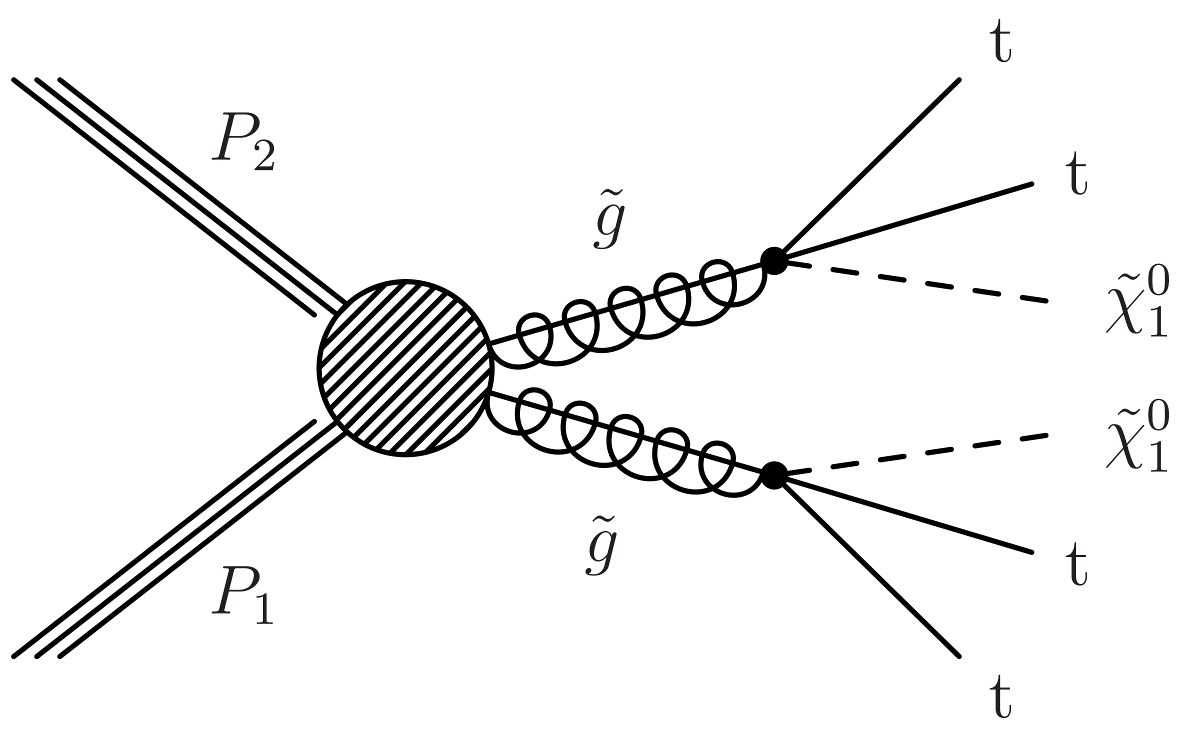
$$\langle \Delta R^2 \rangle = \frac{\sum_i \Delta R_i^2 p_{Ti}^2}{\sum_i p_{Ti}^2}$$

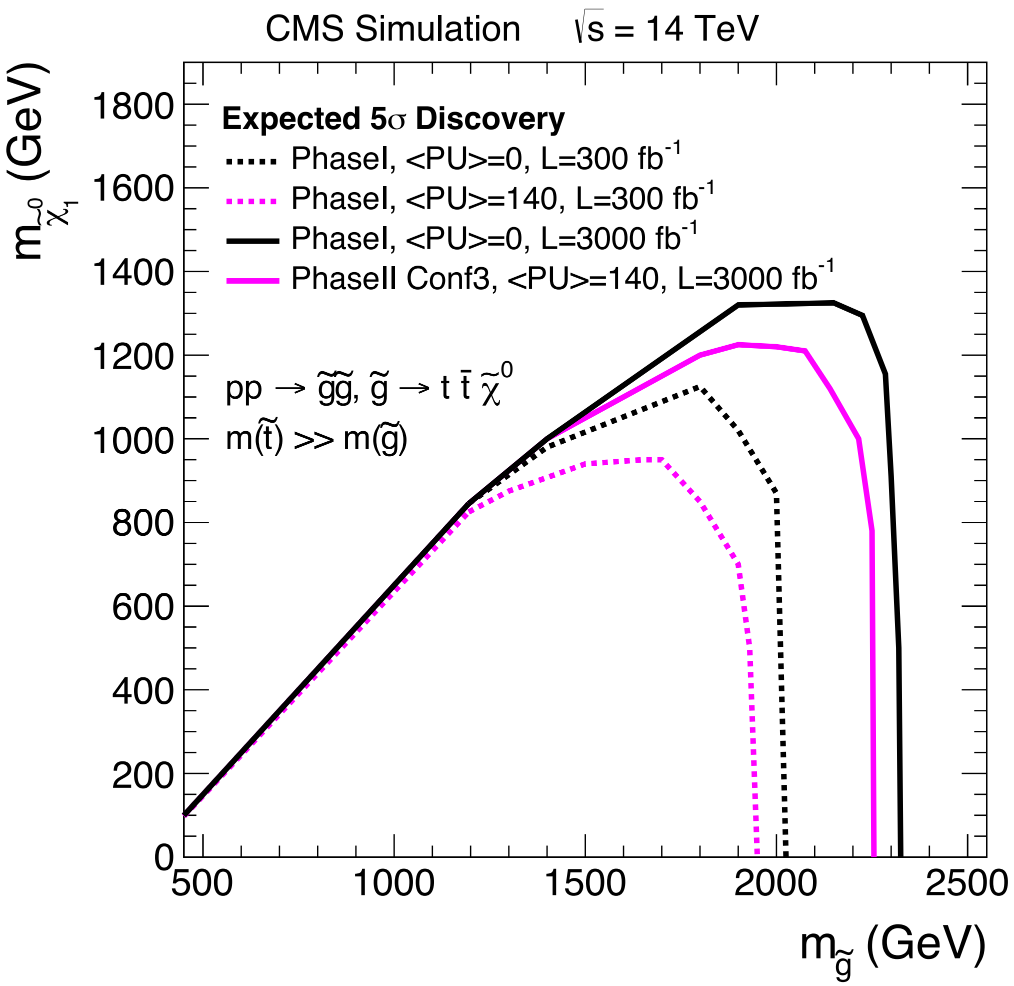
$$p_T^D = \frac{\sqrt{\sum_i p_{Ti}^2}}{\sum_i p_{Ti}}$$

$$A < (\Delta R) < A + 0.1 = \frac{1}{p_T^{jet}} \sum_{i \in A < \Delta R < A+0.1} p_{Ti}$$

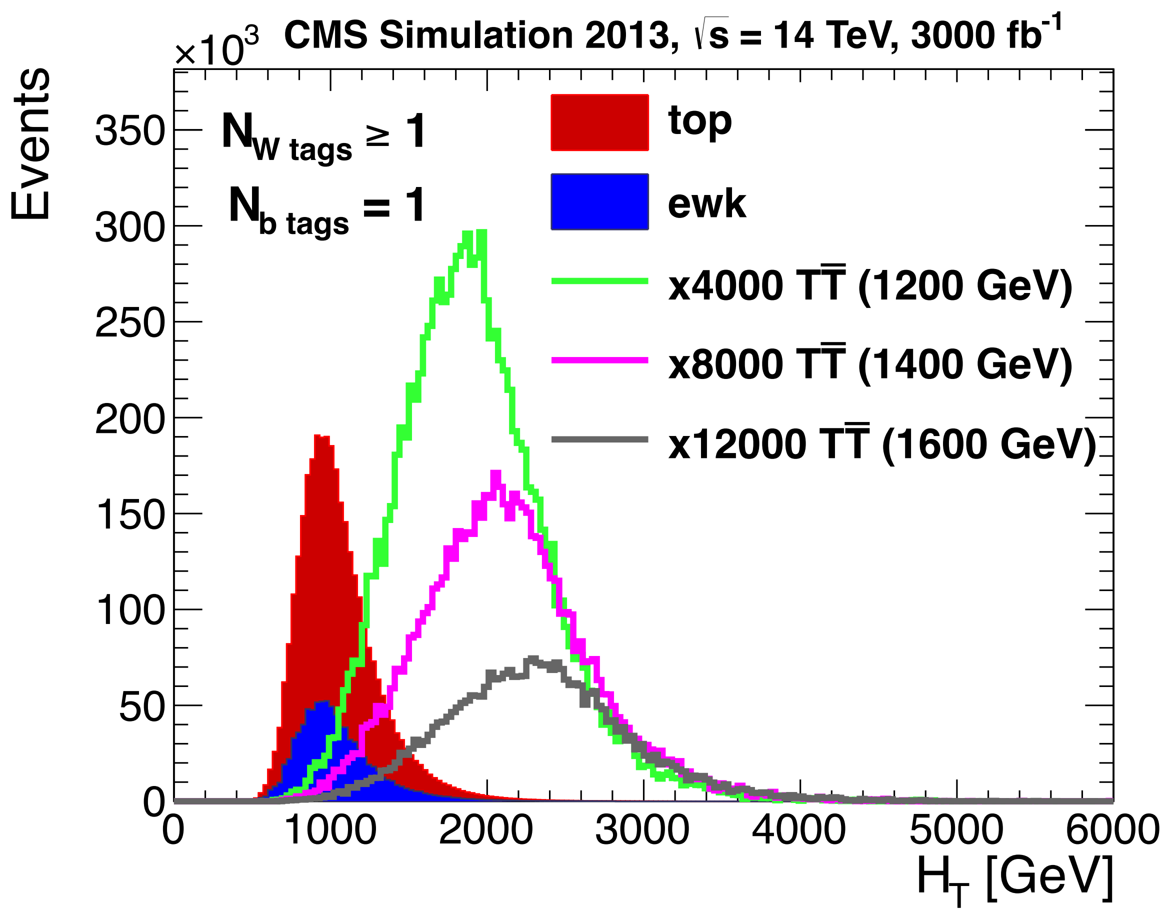
SUSY Searches: 4 top quarks + MET

CMS-PAS-FTR-13-014





Heavy vector-like charged T quarks



CMS-PAS-FTR-13-026