E2/E9 variable properties

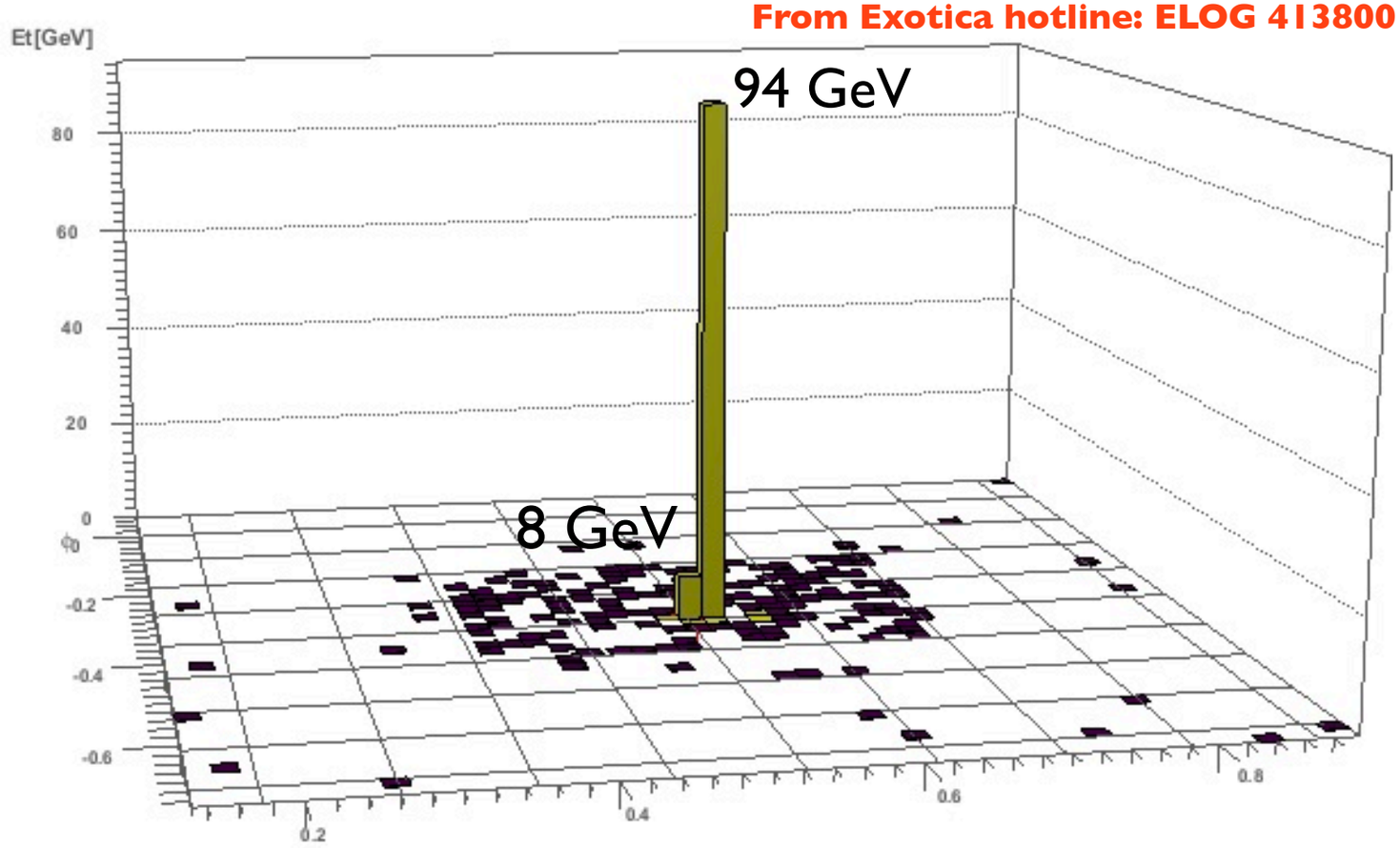
D.A. Petyt Nov 1st 2010

- E2/E9 is an extended isolation variable designed to reject non-isolated or "double" spikes that contain significant energy in 2 neighbouring xtals

- This talk:
  - summary of E2/E9 properties (rechit based analysis)
  - E2/E9 spike-killing efficiency/residual contamination

Run: 139458 Event: 163767467

From Exotica hotline: ELOG 413800



~100 GeV in 2 adjacent xtals. Little surrounding activity.

Fails Swiss cross cut (1-E4/E1~0.91)

E2/E9 references

- Discussed extensively in ASC task force and ECAL DPG meetings

Non-isolated spike properties:

detailed presentations:
http://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=97399
http://indico.cern.ch/getFile.py/access?resId=3&materialId=slides&confId=102650

summary at 26th August 2010 ECAL DPG:
http://indico.cern.ch/getFile.py/access?subContId=4&contribId=1&resId=0&materialId=slides&confId=105063

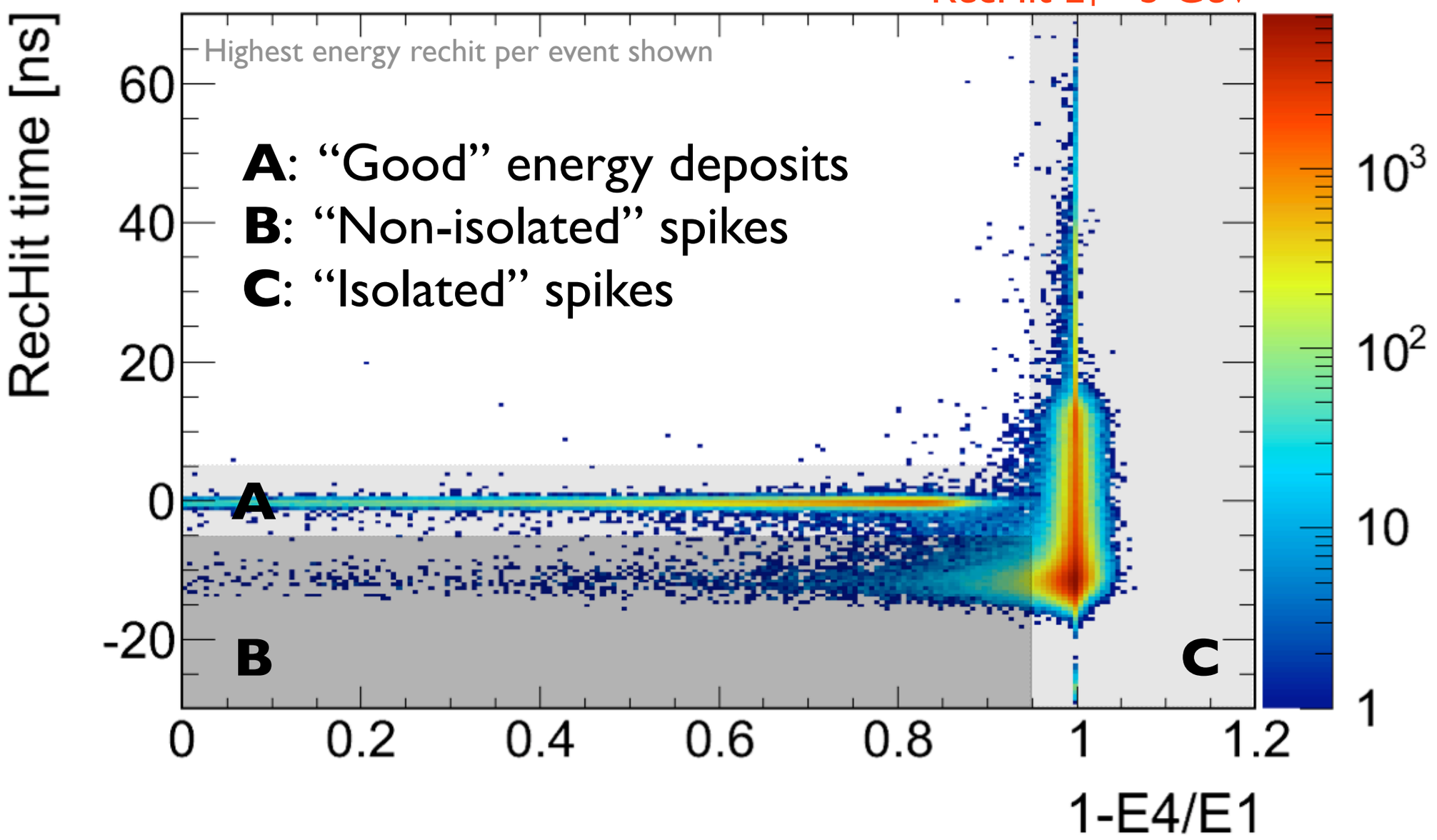
Ancilliary info:
Performance of E2/E9 and E2/E25 variables on data and MC:
http://homepages.spa.umn.edu/~petyt/spikes/e2e9_mc.pdf

Spike summary at Bodrum:
http://indico.cern.ch/getFile.py/access?contribId=100&sessionId=29&resId=0&materialId=slides&confId=101503

E2/E9 code implemented in EcalSeverityLevelAlgo (in 3_9_0_pre5):
http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/CMSSW/RecoLocalCalo/EcalRecAlgos/src/EcalSeverityLevelAlgo.cc?view=markup&pathrev=CMSSW_3_9_0_pre5

Time vs topology phase space

MINBIAS Dataset - NO spike cleaning

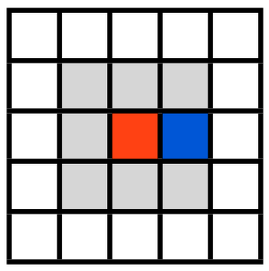


Fraction of non-isolated spikes:

	E_T> 5 GeV	E_T> 10 GeV	E_T> 15 GeV	E_T> 20 GeV	E_T> 30 GeV
relative to all spikes: **B/(B+C)**	2.59%	0.64%	0.34%	0.23%	0.14%
relative to all non-isolated events: **B/(A+B)**	10.01%	11.30%	13.44%	16.15%	20.65%

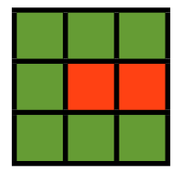
E2/E9 variable definition

- find highest energy hit, then search for second-highest hit in the surrounding 3x3. Call this summed energy **E2**
- calculate energy in a 3x3 surrounding the highest energy xtal



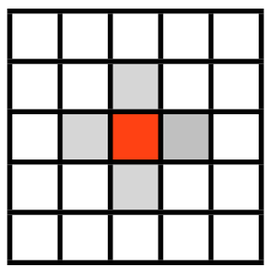
E^{1st} E^{2nd}

$E2=E^{1st}+E^{2nd}$



E2/E9

Swiss-cross reminder:



E1

E4

Swiss-cross= 1-E4/E1

Datasets and cuts

- Datasets

/EG/Run2010A-PromptReco-v4/RECO
/Photon/Run2010B-PromptReco-v2/RECO

- HLT

Photon10* || Photon20* || Photon30* || Photon50*

- Cuts

bit0, (bit40||bit41) no beam halo (bits36-39)
good vtx (!isfake && |z|<15 cm && ndof>5)

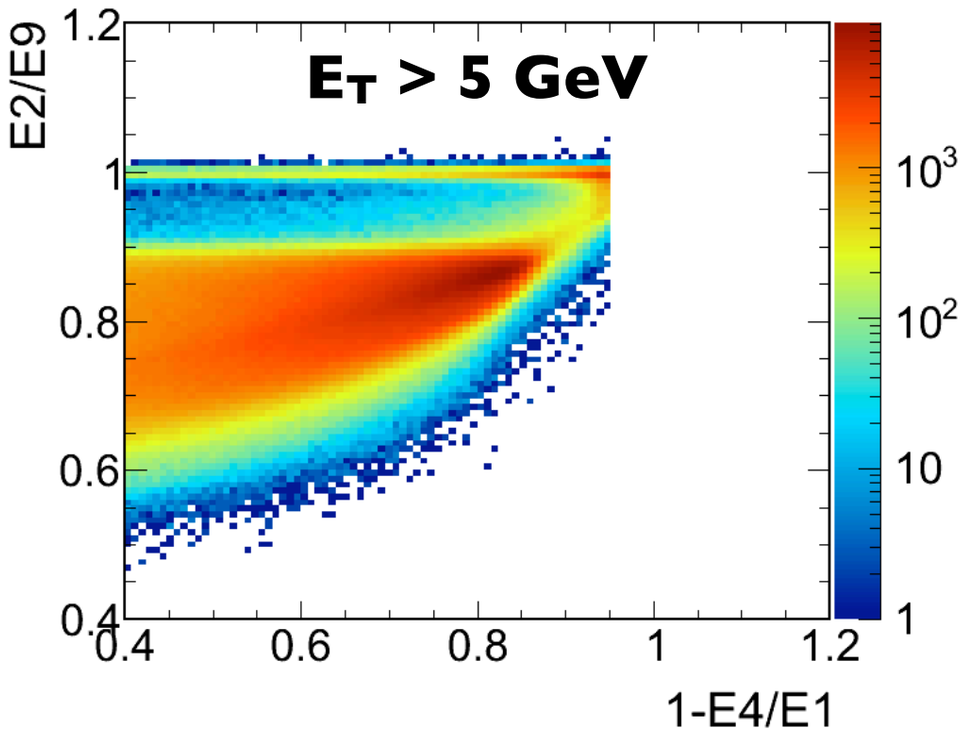
- RecHit selection

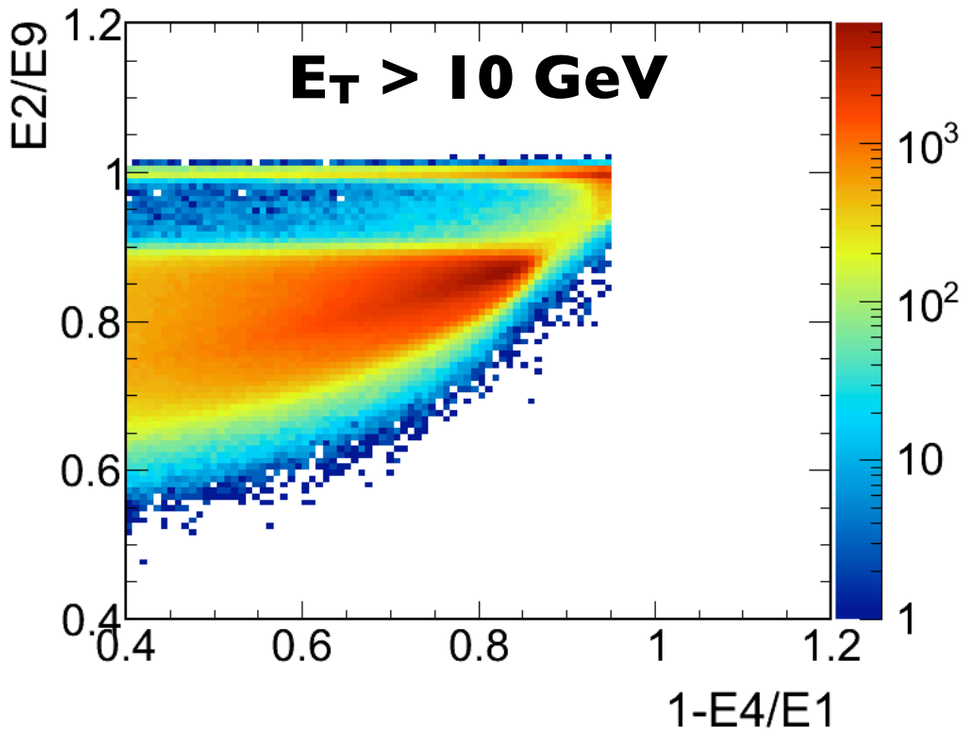
Highest energy EB rechit in each event

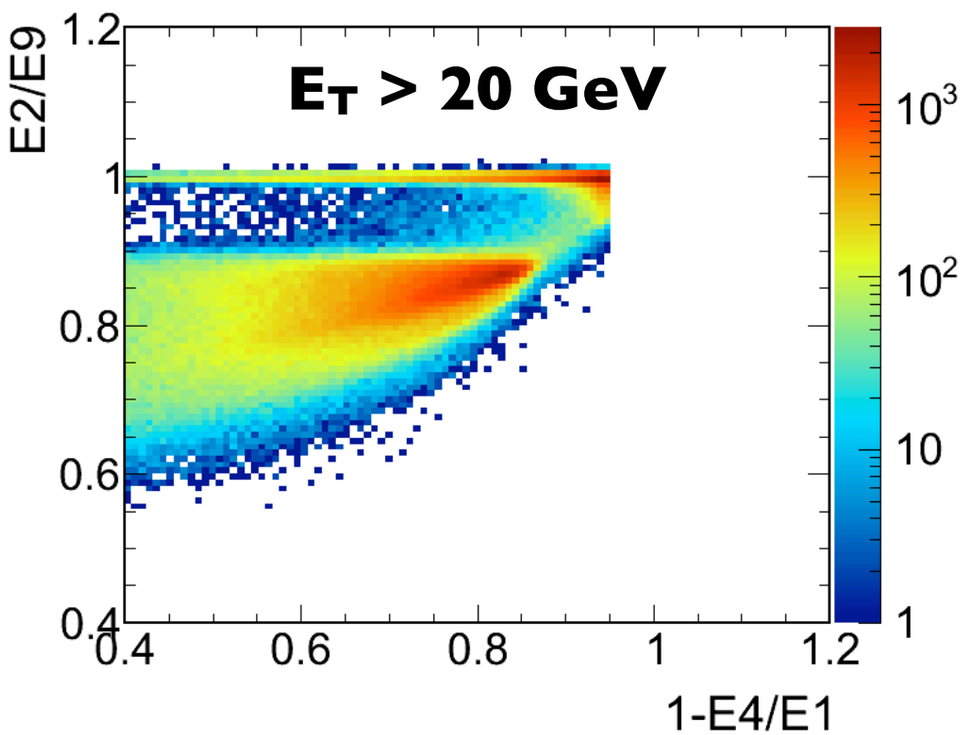
photon object-based analysis in progress. More sophisticated studies performed by others in Egamma group.

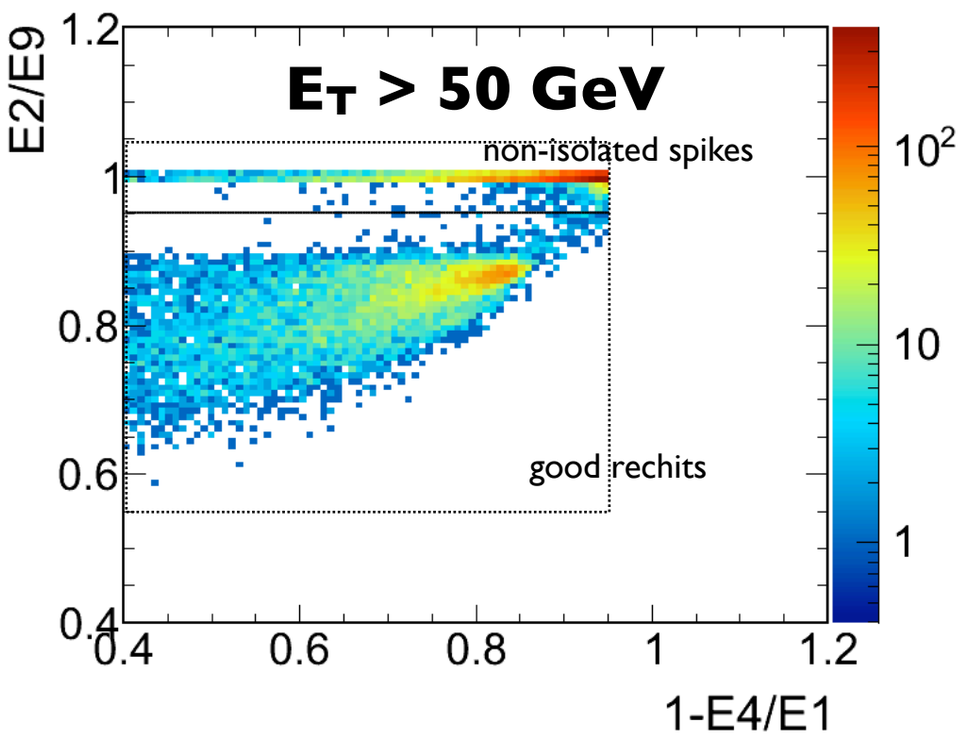
E2/E9 vs Swiss-Cross

1-E4/E1<0.95









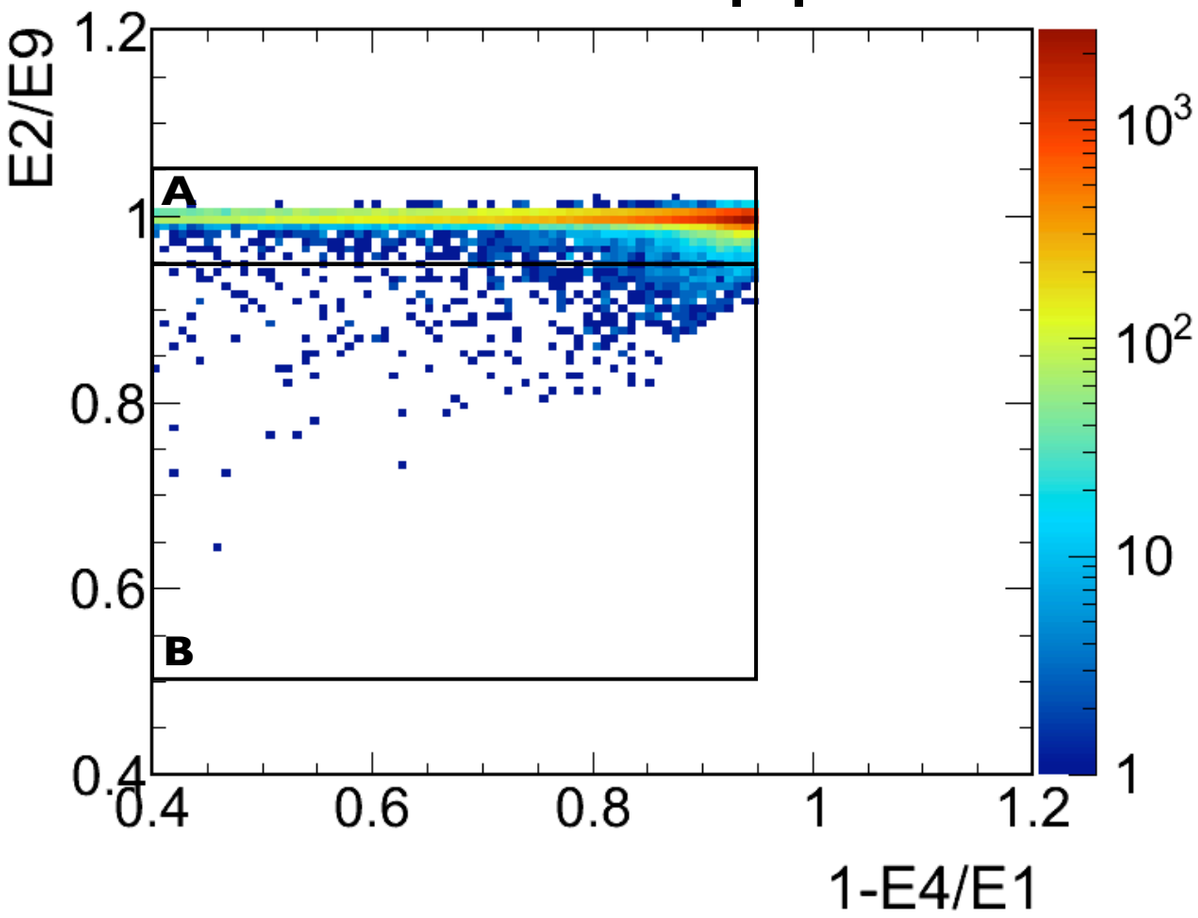
Fairly distinct band of events that are non-isolated using Swiss-cross, but isolated using E2/E9.

Not observed in MC - suggested origin: non-isolated spikes

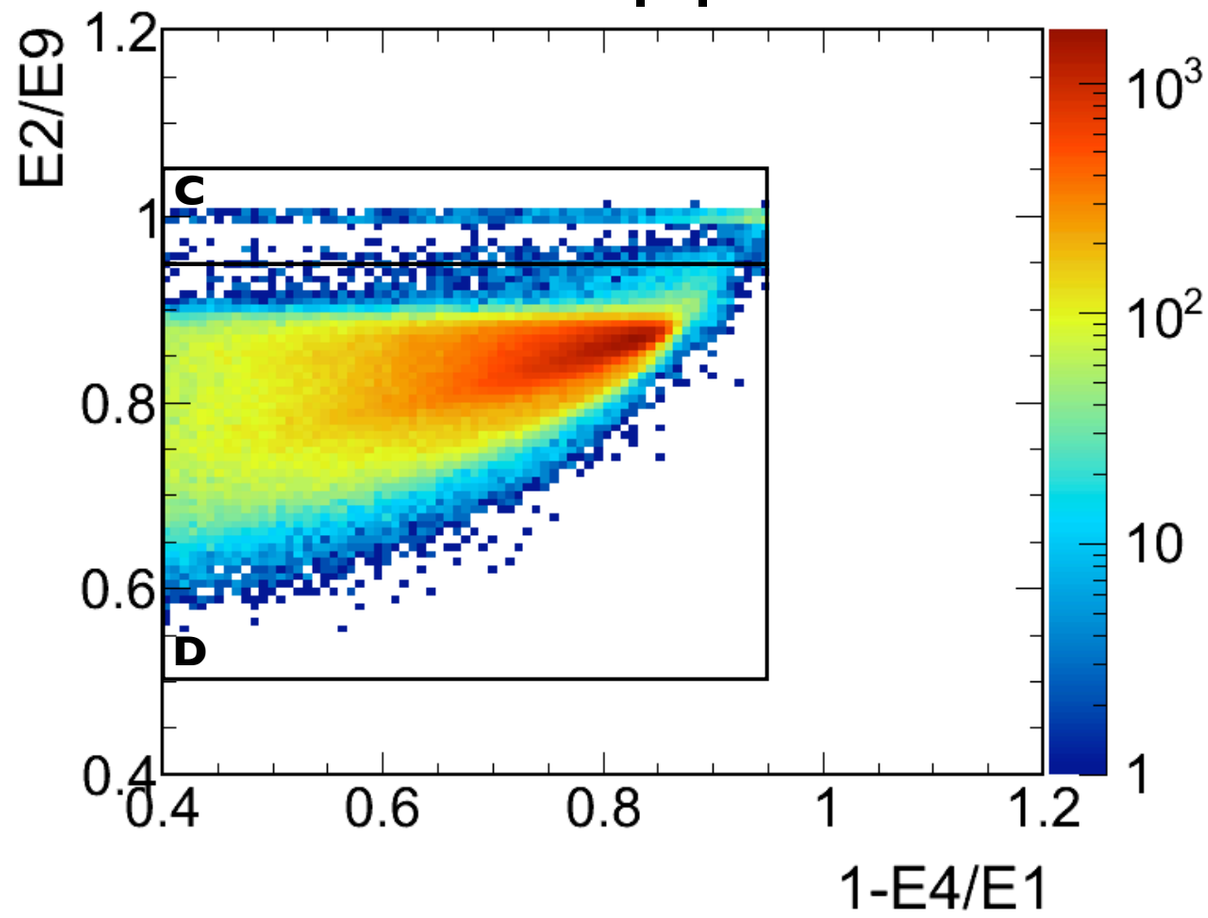
In/out of time

E_T>20 GeV

Out-of-time: |t|>3ns



In-time: |t|>3ns



Events with E2/E9<0.95 generally out-of-time. Supports spike-like origin

Confusion matrices

E_T>20 GeV

A 54704	C 1451
B 4368	D 324144

E_T>30 GeV

A 25125	C 543
B 1108	D 78377

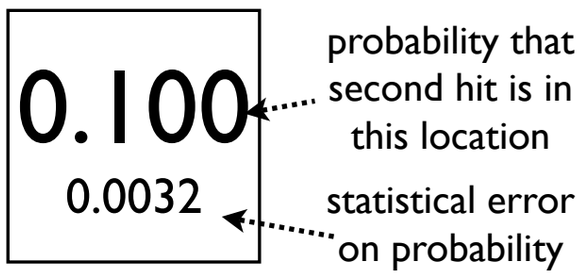
E_T>50 GeV

A 6112	C 118
B 246	D 10856

E2 most probable position

- Calculate difference in eta and phi between highest hit, E1 and second highest hit, E1a
  - **non-isolated: E1/E9<0.95, 1-E4/E1<0.95**
- Compare for 2 cases:
  - **both hits in-time (|t|<3 ns)**
  - **both hits out-of-time (|t|>3 ns)**

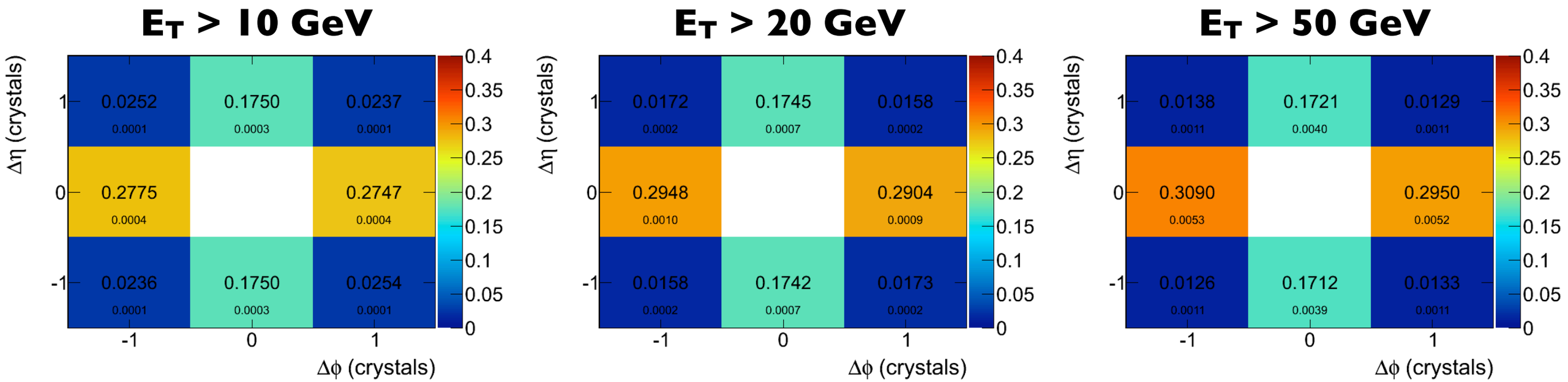
Key to maps on next 2 slides:



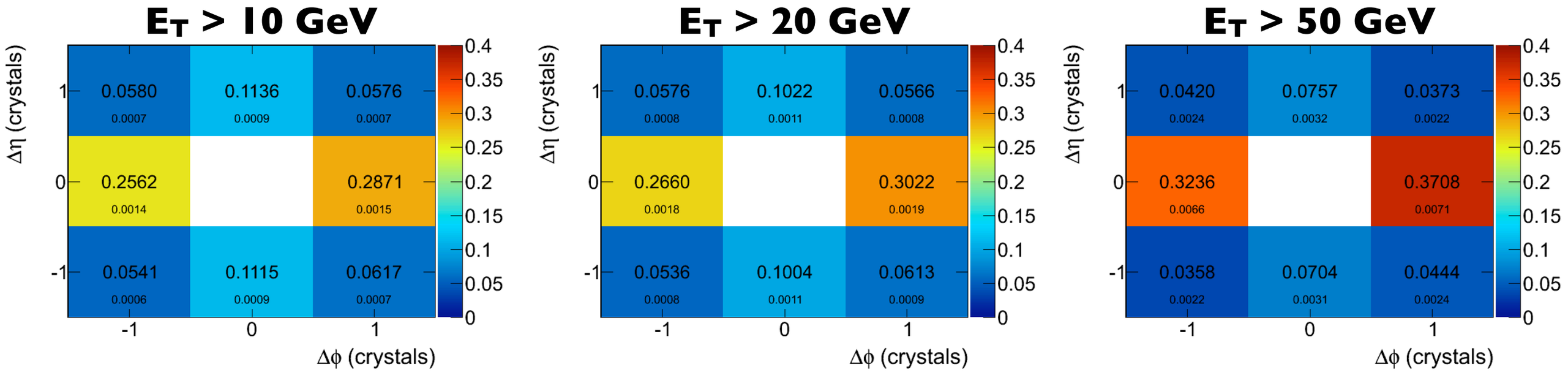
Position of 2nd highest RecHit relative to most energetic hit

Non-isolated: E1/E9<0.95

- in-time (|t|<3 ns)



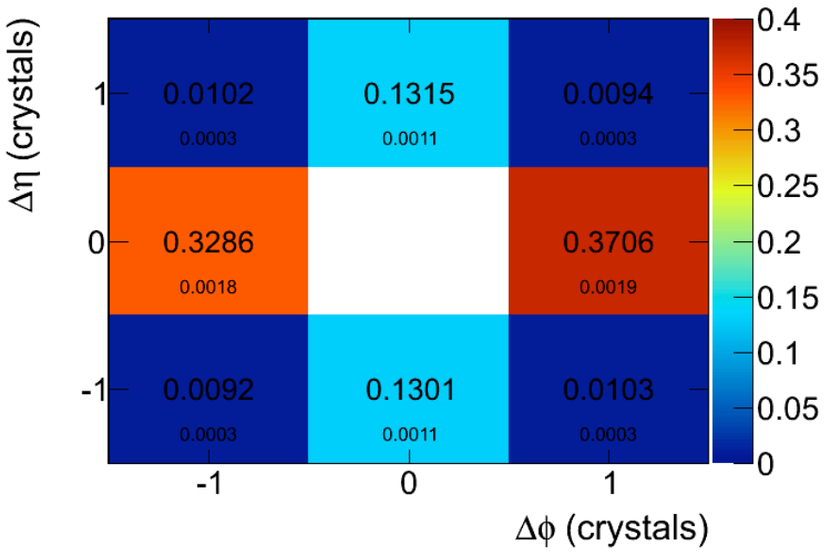
- out-of-time (|t|>3)

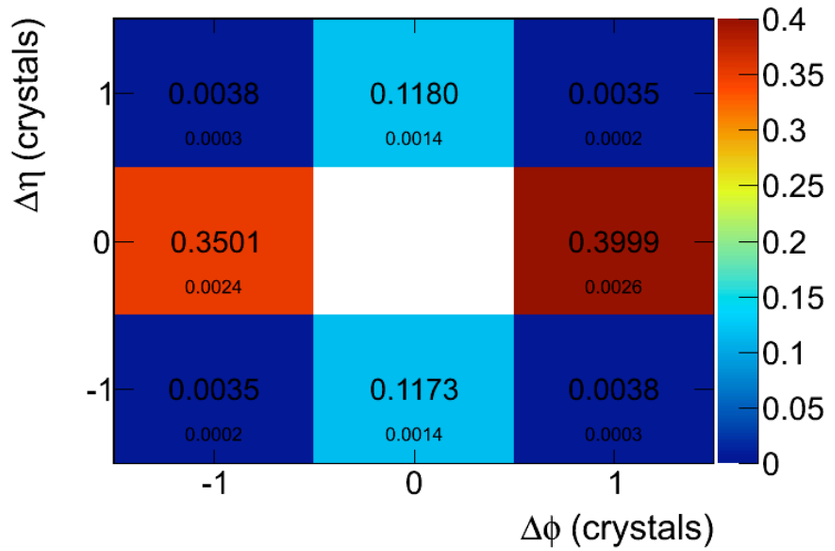


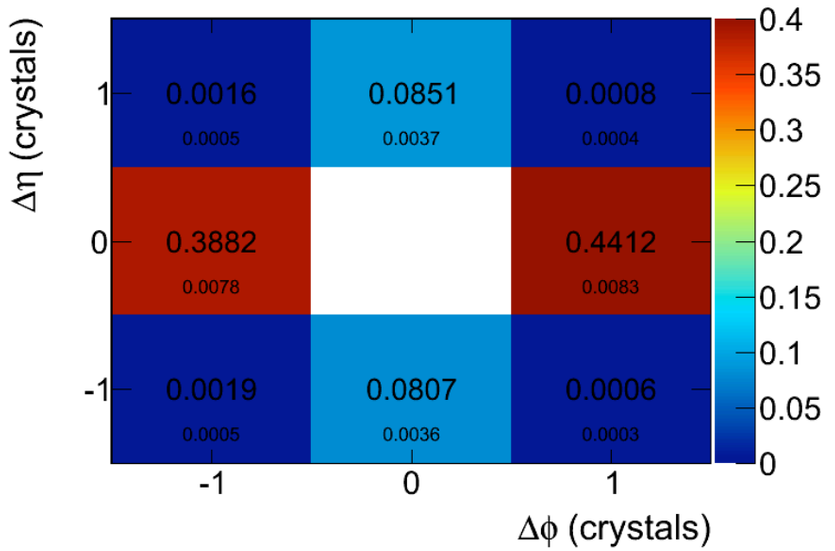
Position of 2nd highest RecHit relative to most energetic hit

Non-isolated: 1-E4/E1<0.95

- out-of-time (|t|>3 ns)







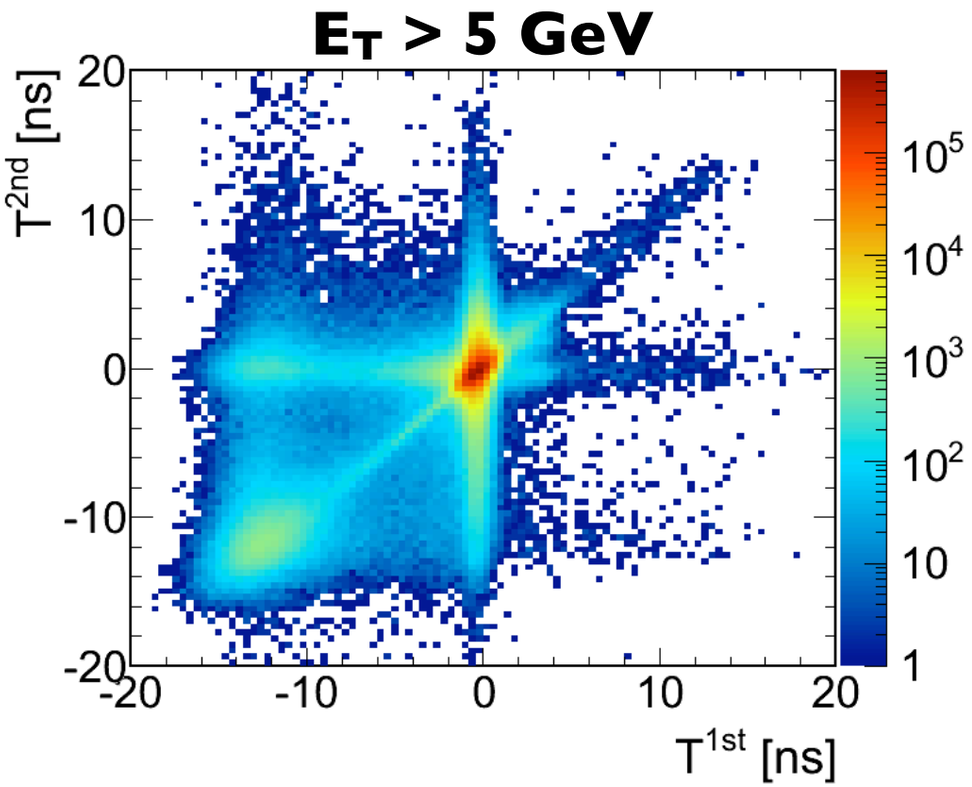
Applying a cut on 1-E4/E1 biases the 2nd highest hit to the 4 Swiss-cross positions

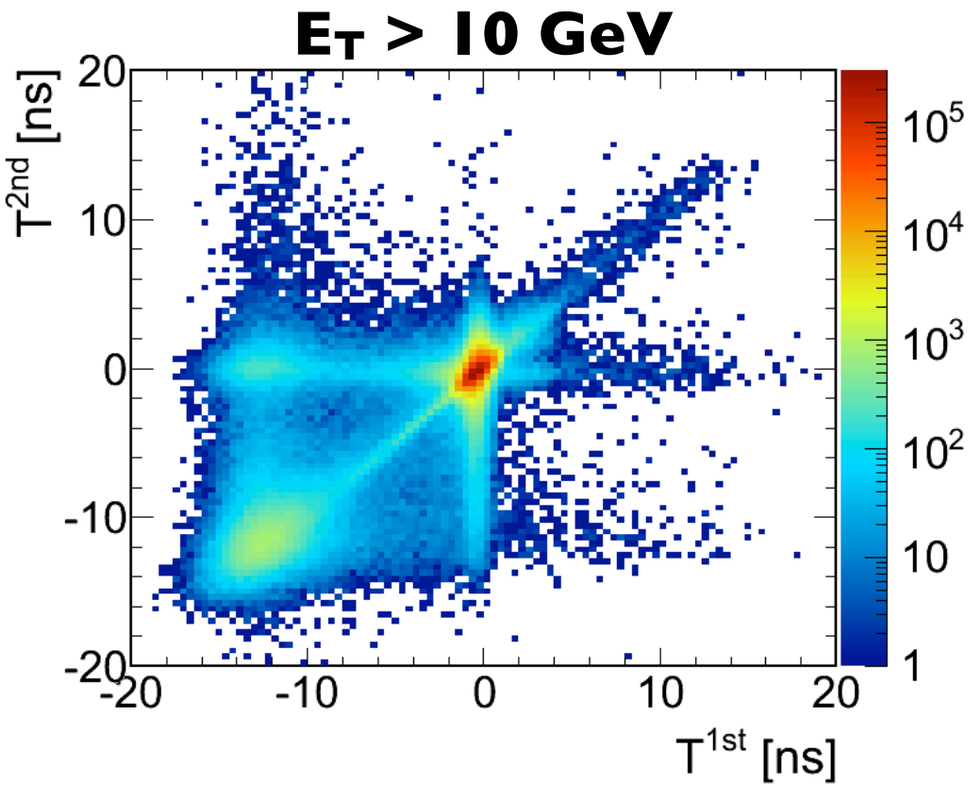
2nd hit is 2-3x more likely to be at the same value of η as the central xtal as it is to be at the same value of ϕ. This ratio appears to increase with E_T.

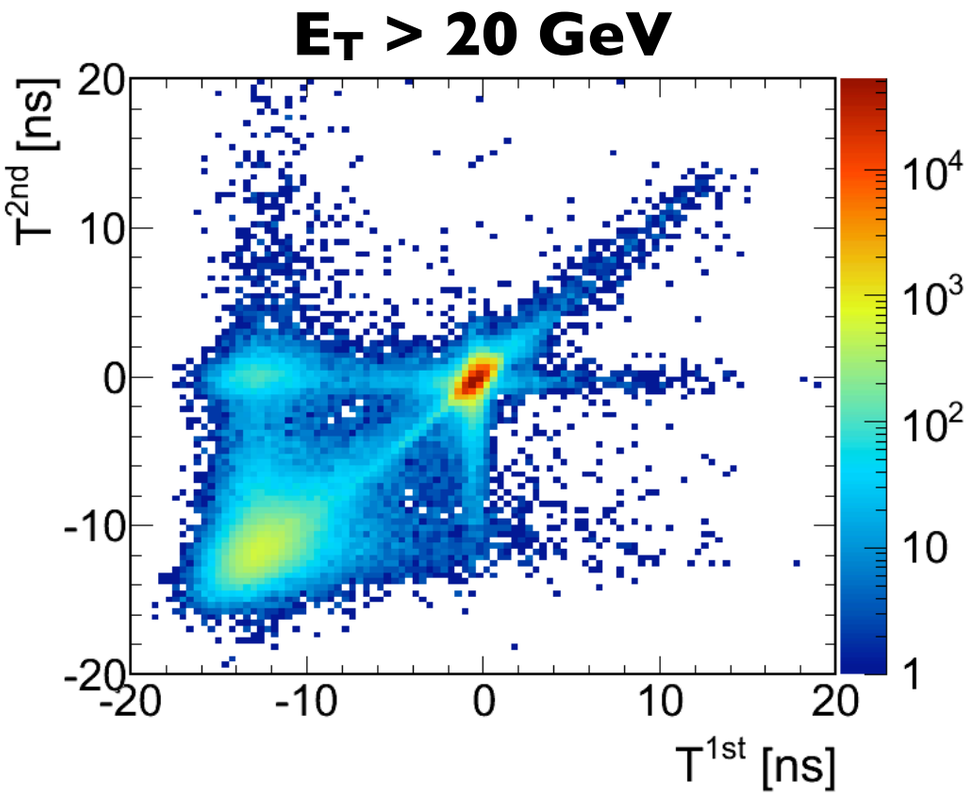
Implies that $\sigma_{i\eta i\eta}$ should be more effective than $\sigma_{i\phi i\phi}$ at removing spikes

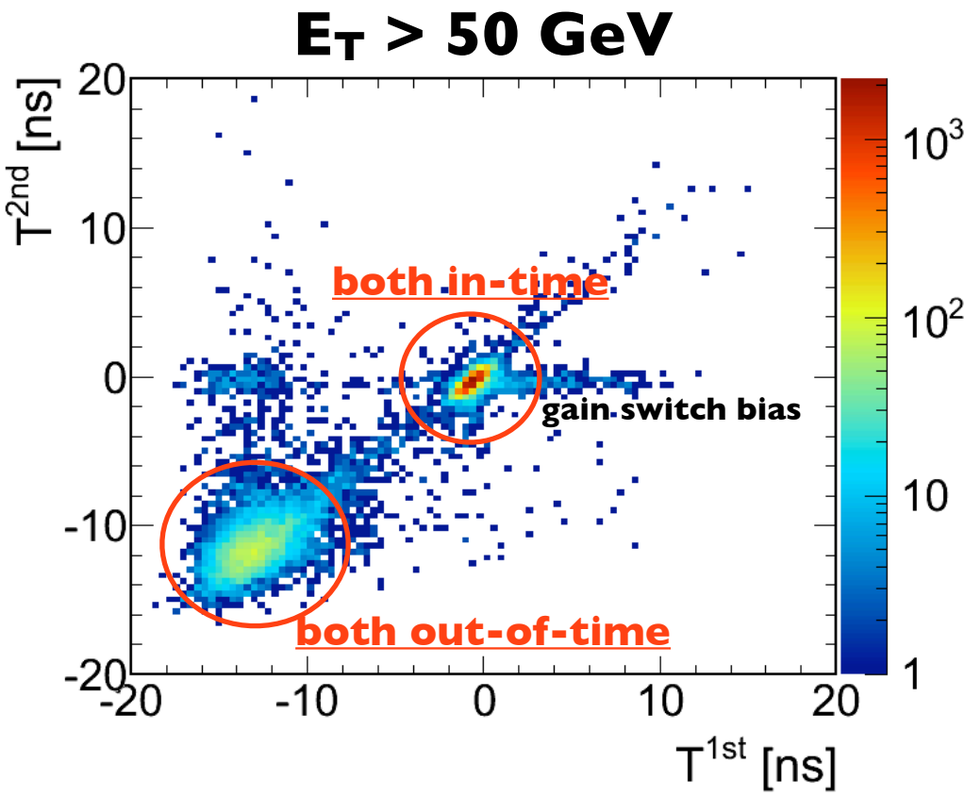
Cutting on E2/E9 roughly equivalent to a cut on ($\sigma_{i\eta i\eta}$ AND $\sigma_{i\phi i\phi}$)

E1 vs E2 timing









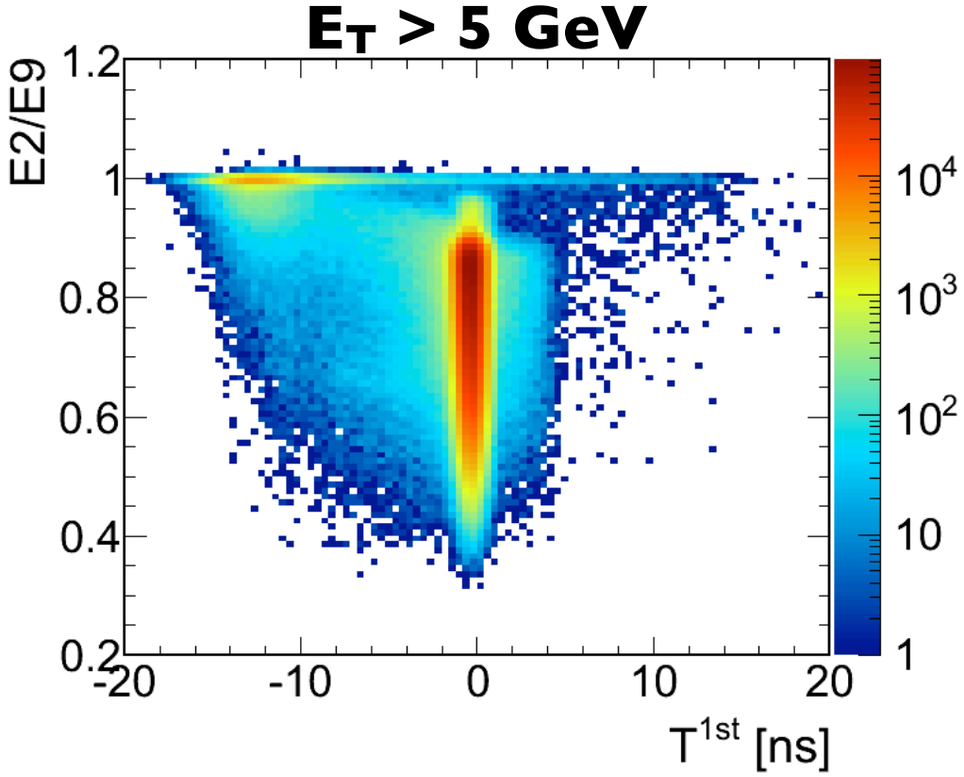
Complex structure. Becomes cleaner as E_T threshold is raised

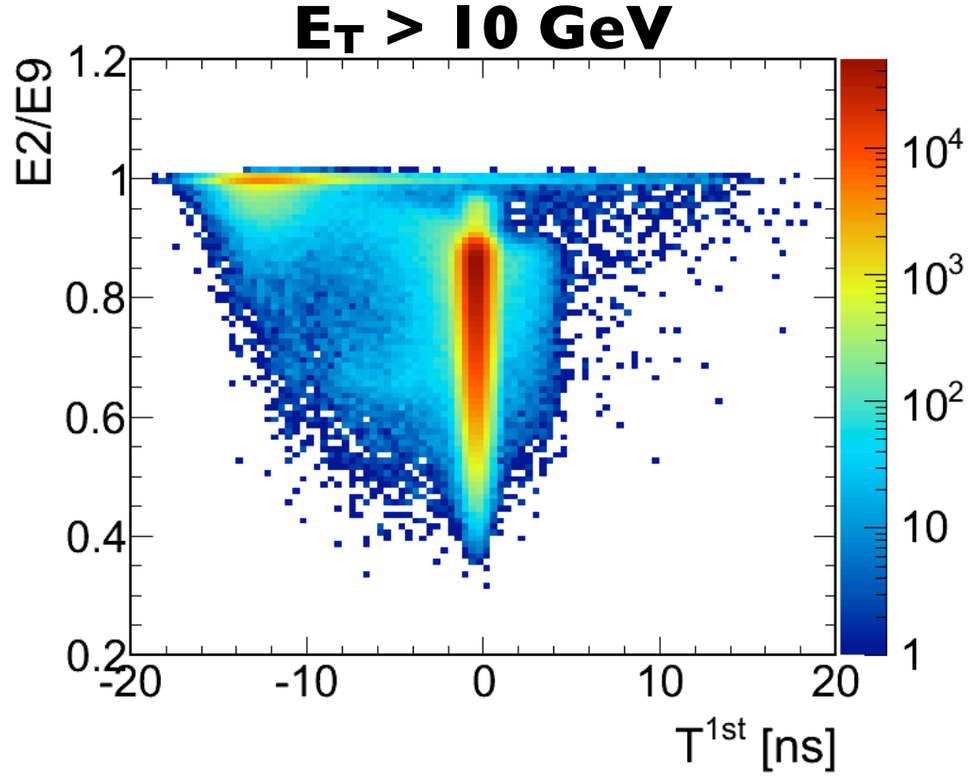
Time of E_{1st} and E_{2nd} is clearly correlated.

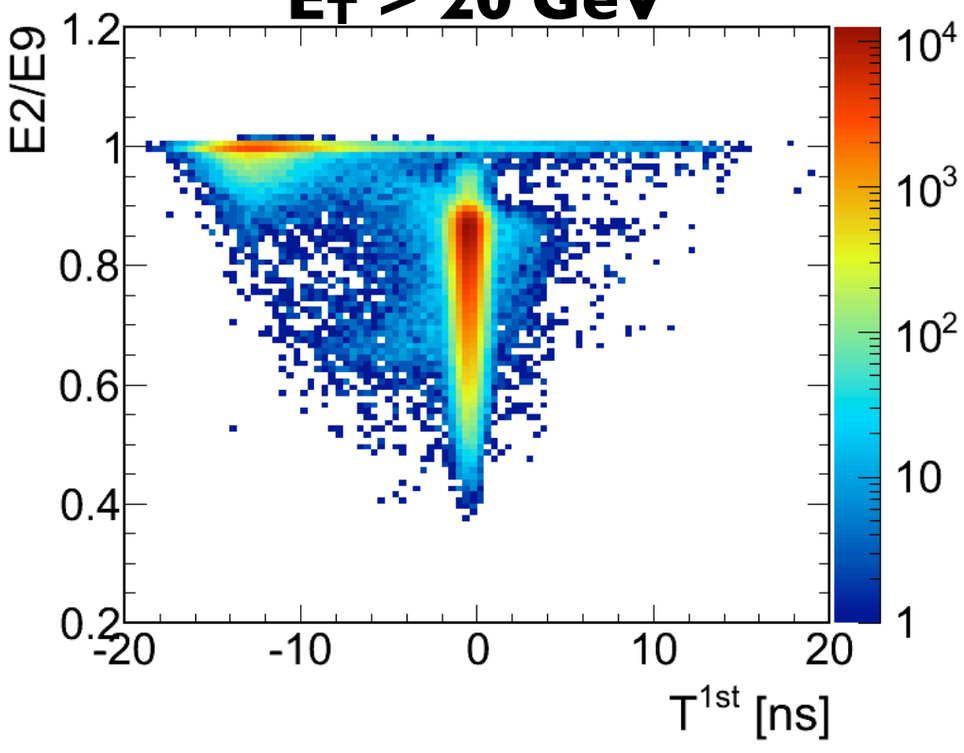
Suggests a common origin for the two hits. Spike-like origin for t~-10ns

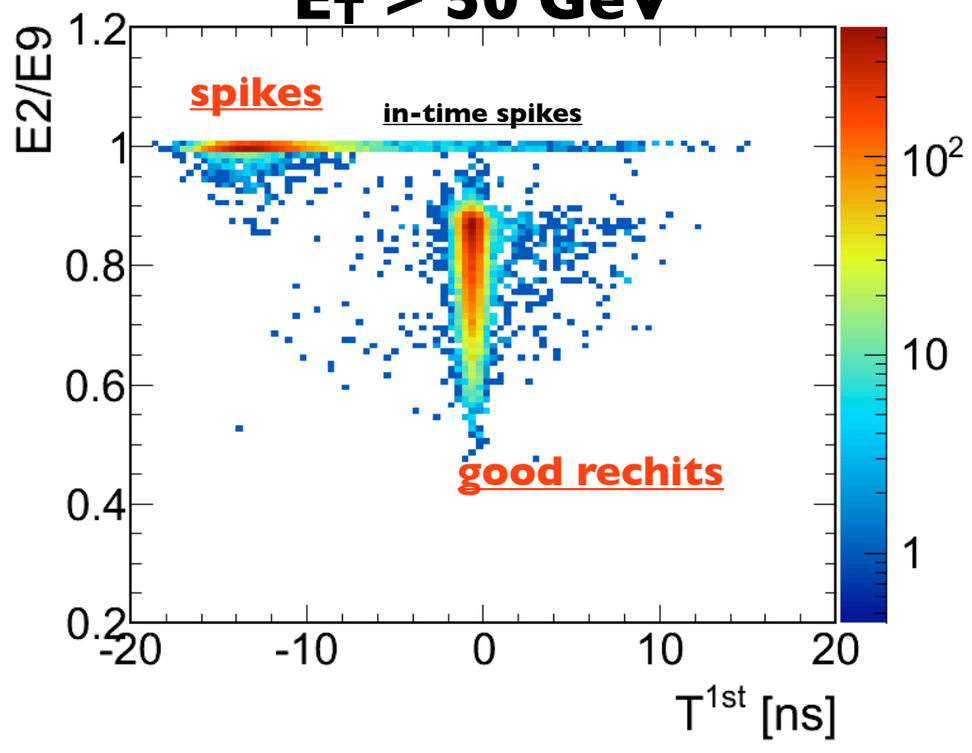
E1 timing vs E2/E9

1-E4/E1<0.95



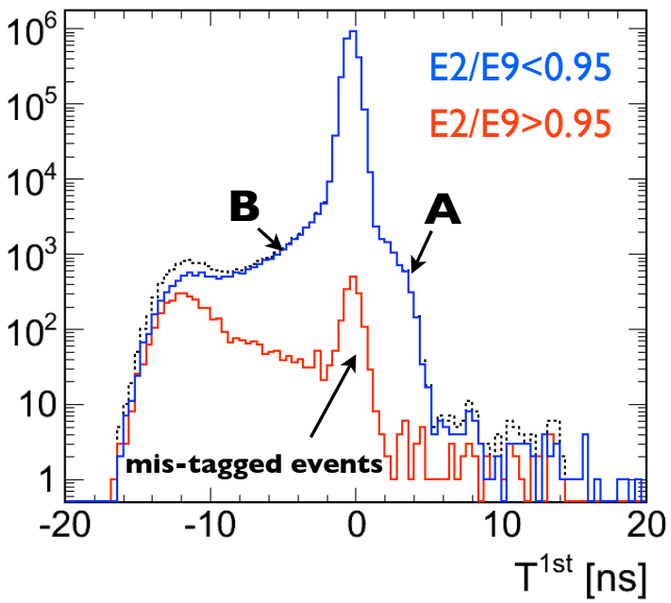




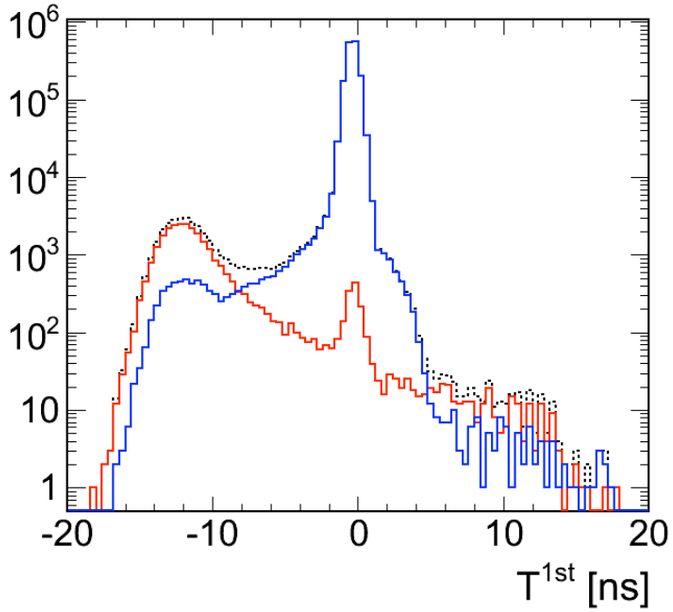


E2 contamination

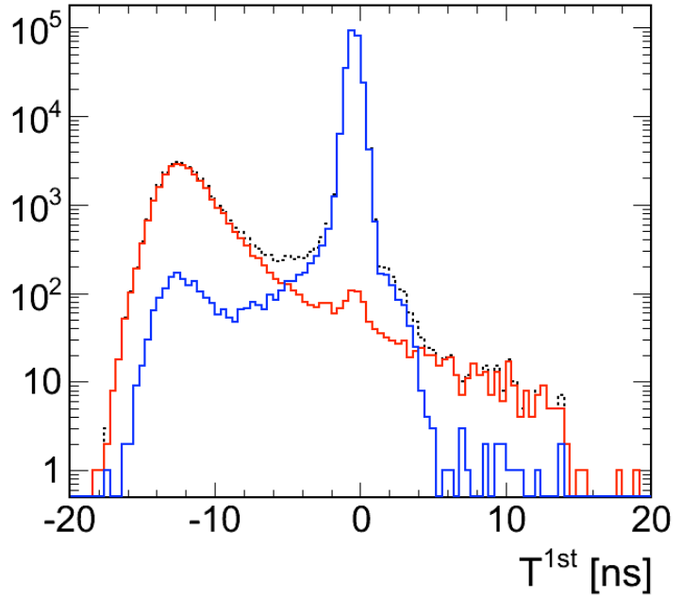
5< E_T < 10 GeV



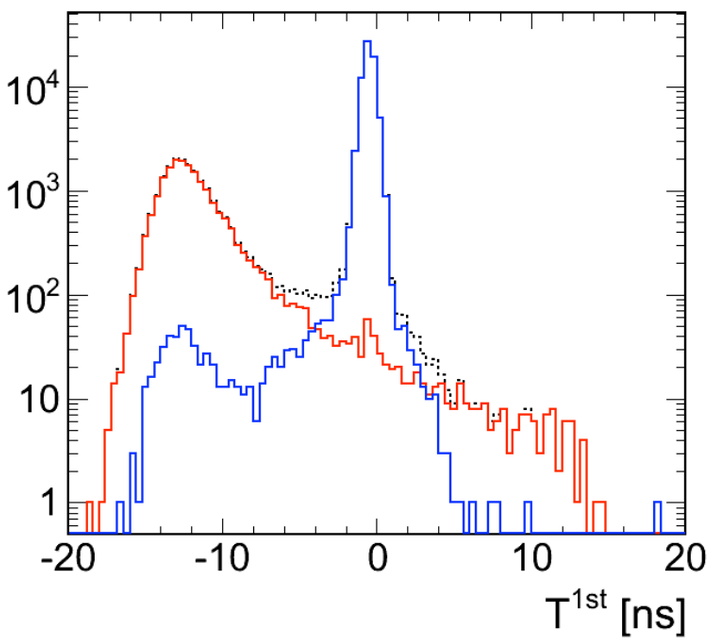
10< E_T < 20 GeV



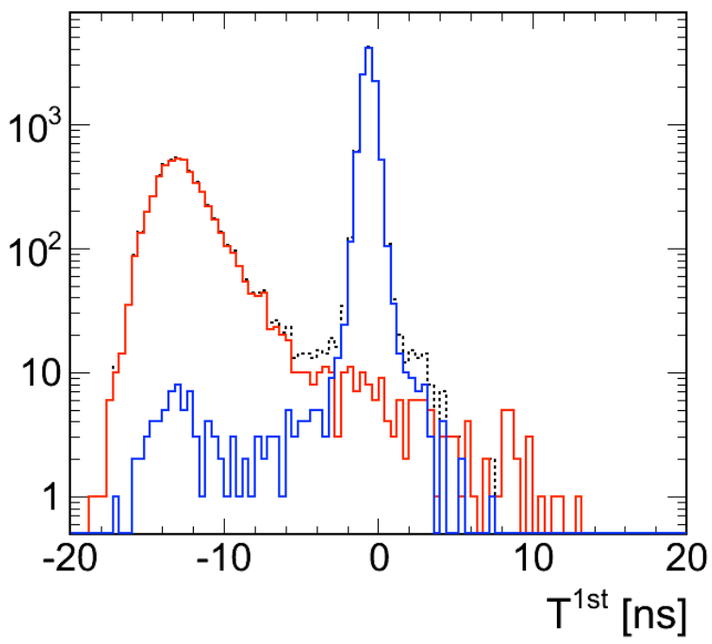
20< E_T < 30 GeV



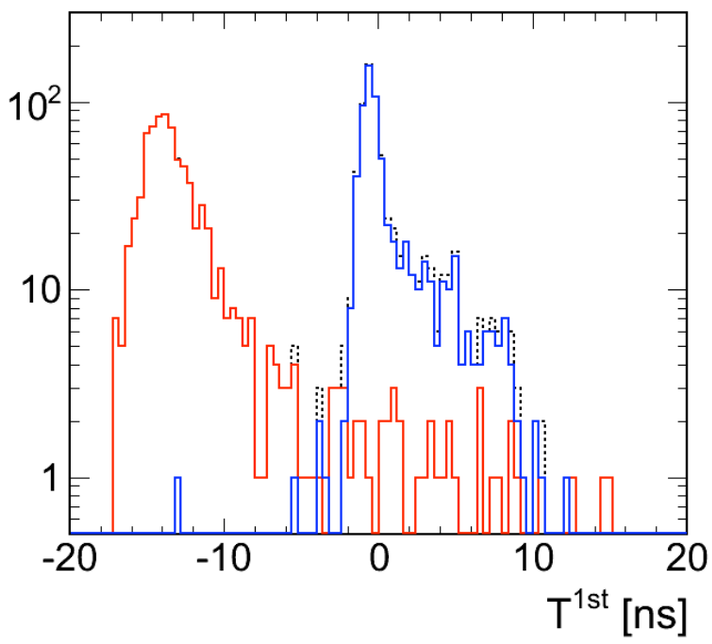
30< E_T < 50 GeV



50< E_T < 100 GeV



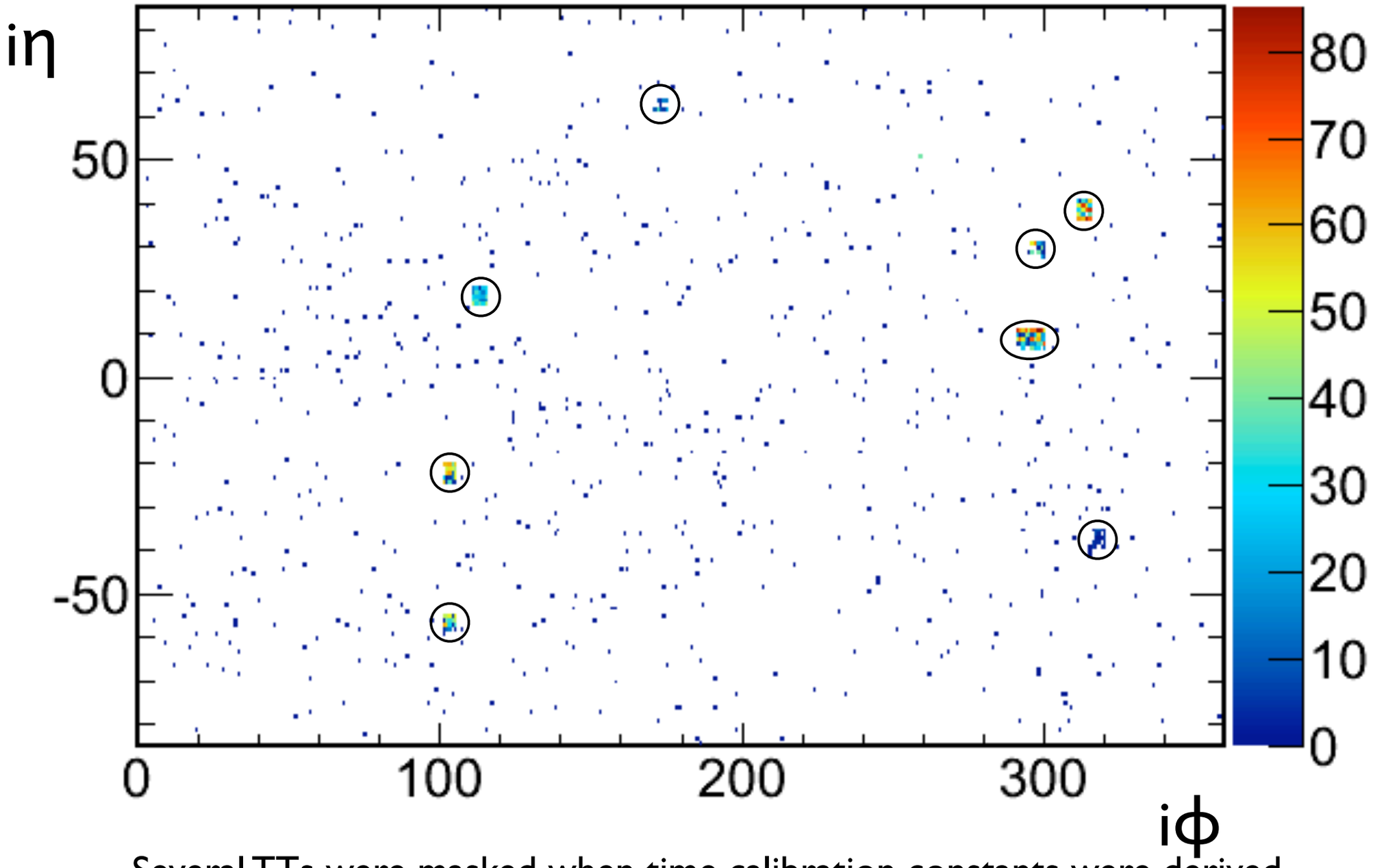
E_T > 100 GeV



Structures **A** and **B** investigated on next slide

Structure "A"

Rechits with +2<t<+5 ns



Several TTs were masked when time calibration constants were derived

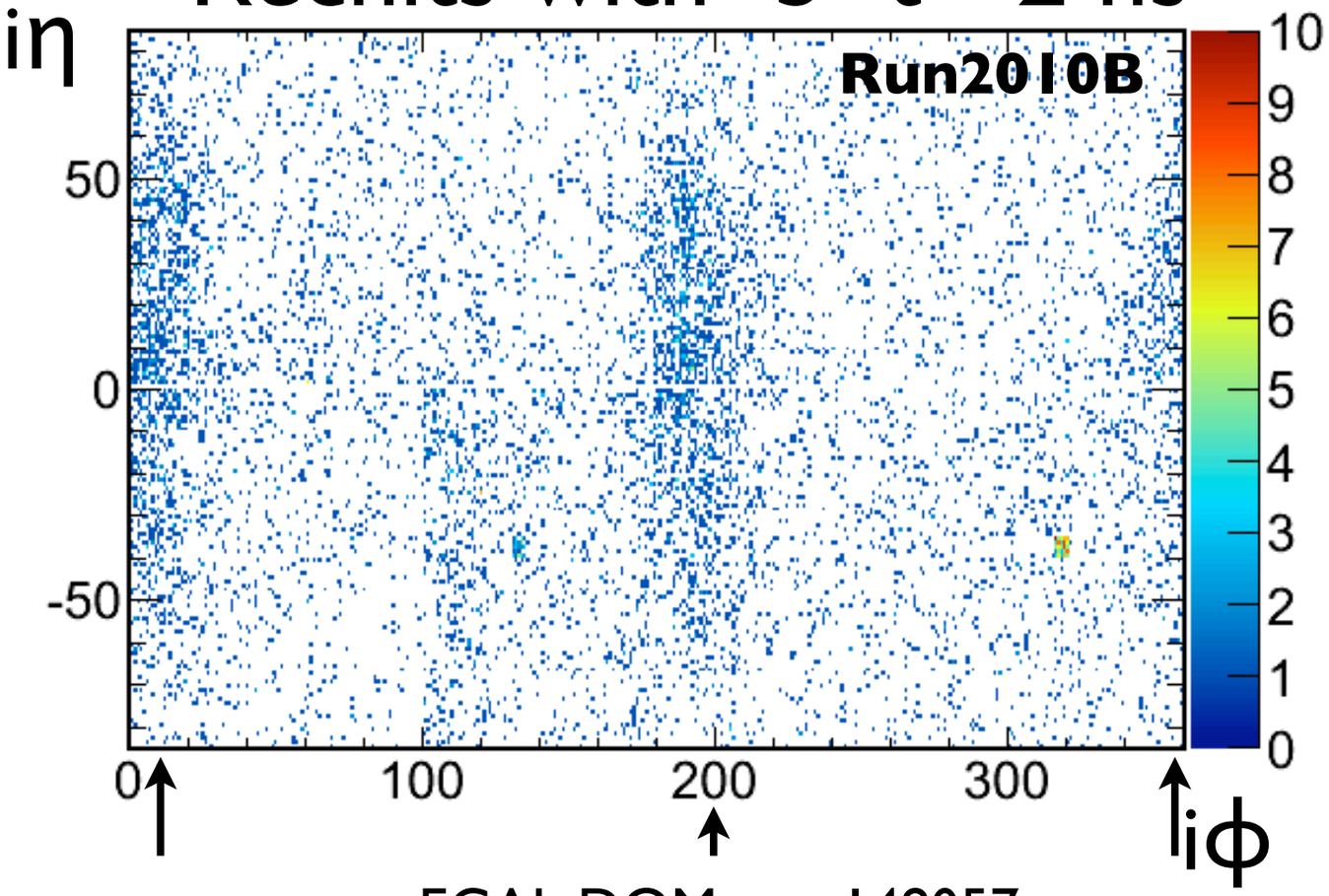
(based on Splash 09 data + early Collisions10 data).

ECAL r/o phase changed to +2.08ns for 2010 running

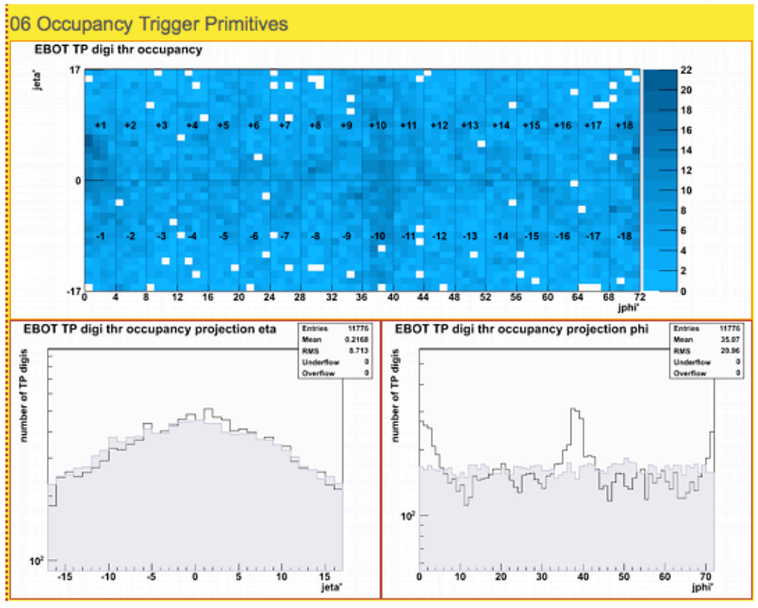
Constants for these towers will be corrected for 2010 collisions data reprocessing.

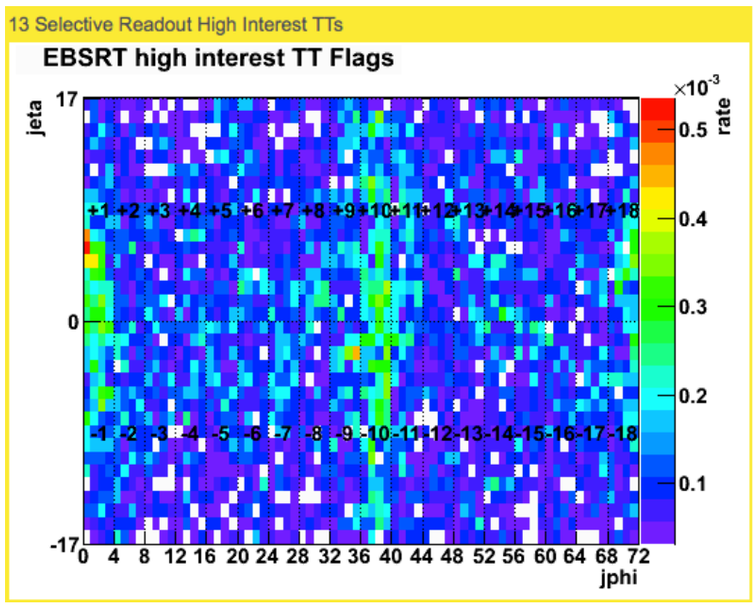
Structure "B"

Rechits with -5<t<-2 ns



ECAL DQM, run 148057





Interesting structure seen for hits with -5<t<-2 ns. Two "bands" evident at $\phi=0,\pi$

Beam-related background?

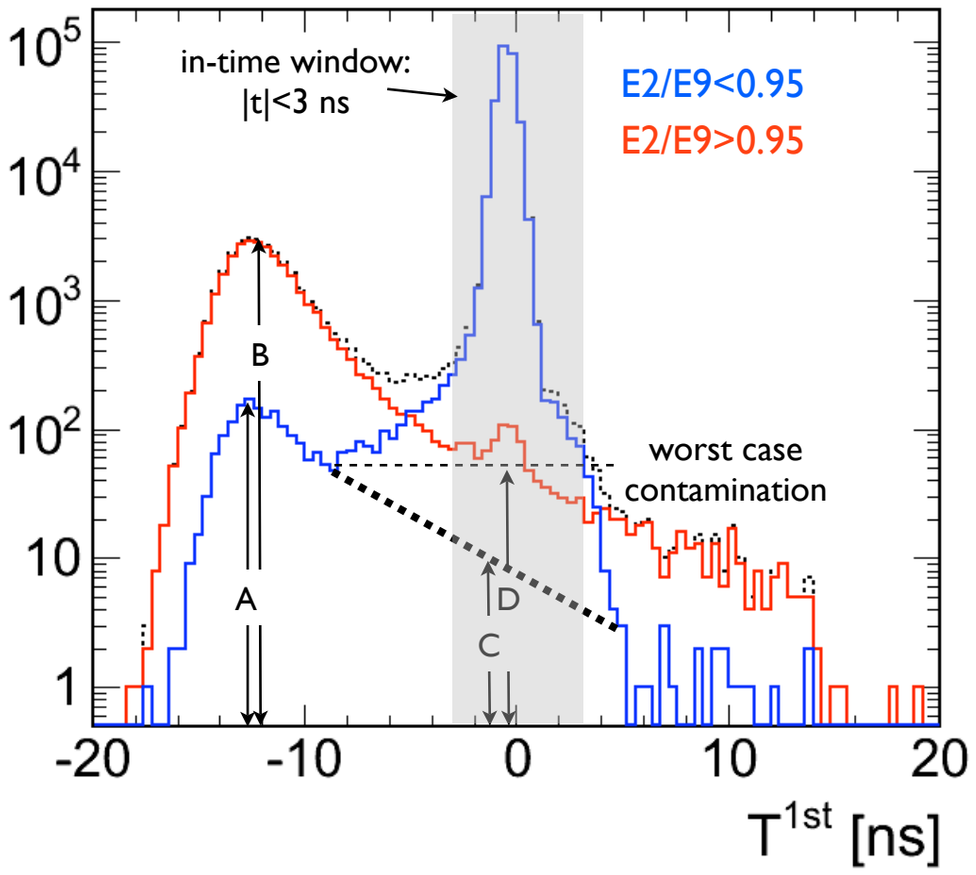
More prominent in Run2010B data than 2010A

This effect also clearly seen in relevant ECAL DQM plots for recent data (see bottom left)

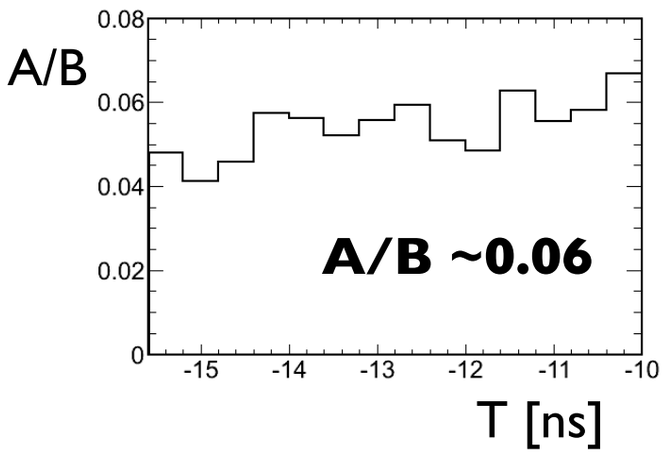
Interesting to see if such an effect is visible in photon-selected samples

It's possible I am missing a cleaning cut in my analysis...

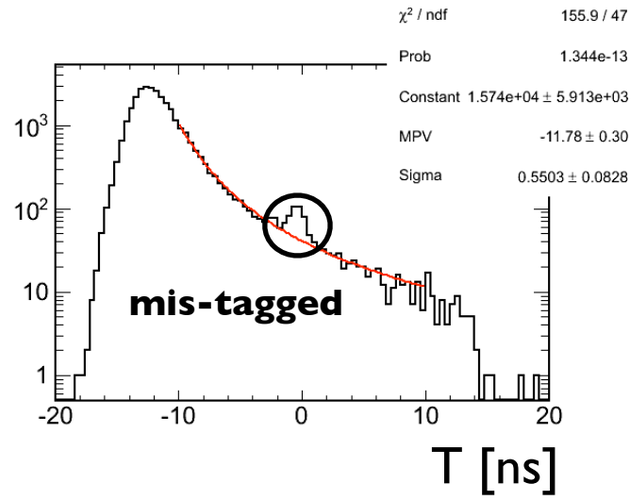
$20<E_T<30$ GeV example



Ratio of blue/red



Fit to red spike-tagged dist



A rough (illustrative) example

Assume shape of spike-tagged sample (red) is representative of residual spike component of blue distribution.

(modulo peak of mis-tagged good hits at zero)

i.e **A/B=C/D**

1) Evaluate A/B for T<-10ns

2) Define |t|<3 ns as "in-time" window

Integral of spike-tagged dist (|t|<3)= 651

Estimated in-time spike contamination (|t|<3)=**39 events**

Integral of non-spike-tagged dist (|t|<3) = 2.46e5 events

spike fraction ~ 1.6e-4

worst case scenario (flat non-tagged dist) spike fraction ~0.3%

mis-tagged events (|t|<3): **257**

Spike fraction vs E_T

V. Preliminary

RecHit E_t cut

	5-10	10-20	20-30	30-50	50-100
A/B	1.6	0.2	0.06	0.03	~0.02
Spike tagged (|t|<3)	241	646	651	333	65
estimated untagged spikes (|t|<3)	386	129	39	10	~2
Non-spike |t|<3	2.46E+06	1.56E+06	2.46E+05	6.75E+04	1.03E+04
Spike fraction (worst case)	1.6e-4 (0.003)	8.3e-5 (0.003)	1.6e-4 (0.003)	1.5e-4 (0.004)	~2e-4 (0.004)
Mis-tagged events |t|<3	1354	1032	257	92	31

Estimated spike contamination: 1-2 x10^{-4}

worst case contamination (assuming flat spike time dist): 0.3%

Summary

- E2/E9 properties explored using RecHit based analysis
  - E2/E9 variable is effective in removing non-isolated spikes (failing Swiss-cross cut)
  - A cut on E2/E9 is roughly related to cuts on $\sigma_{i\eta i\eta}$ and $\sigma_{i\phi i\phi}$
    - position of "2nd highest hit" is adjacent to 1st hit. More often at same value of η than ϕ.
    - Clear correlation between the timing of the two hits.
  - Apply non-isolated spike cleaning using E2/E9<0.95
    - rough attempt to estimate residual in-time spike contamination after this cut.
    - Result depends on assumed shape of non-tagged spike timing distribution. However residual contamination is clearly small (<<1%)
      - *Assume same shape as tagged spike timing dist*: in-time contamination ~ **1-2x10^{-4}** (largely independent of E_T)
      - *Assume flat timing distribution (pessimistic)*: in-time contamination ~ **0.3%**
      - **Should be checked using more sophisticated analyses with photon selection**
      - **Efficiency on MC should also be qualified with further studies. See back-up slides for a brief summary of my RecHit-based analysis.**

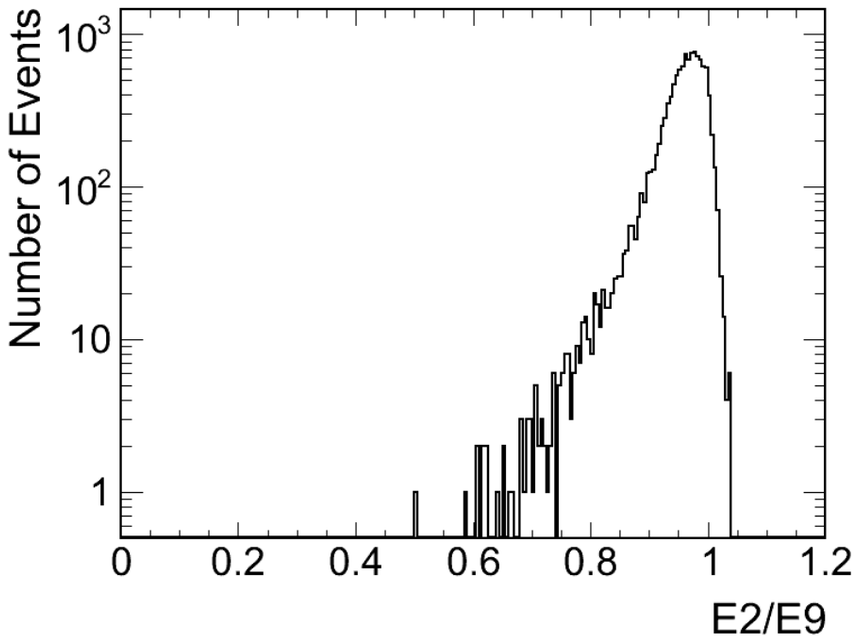
Backup slides

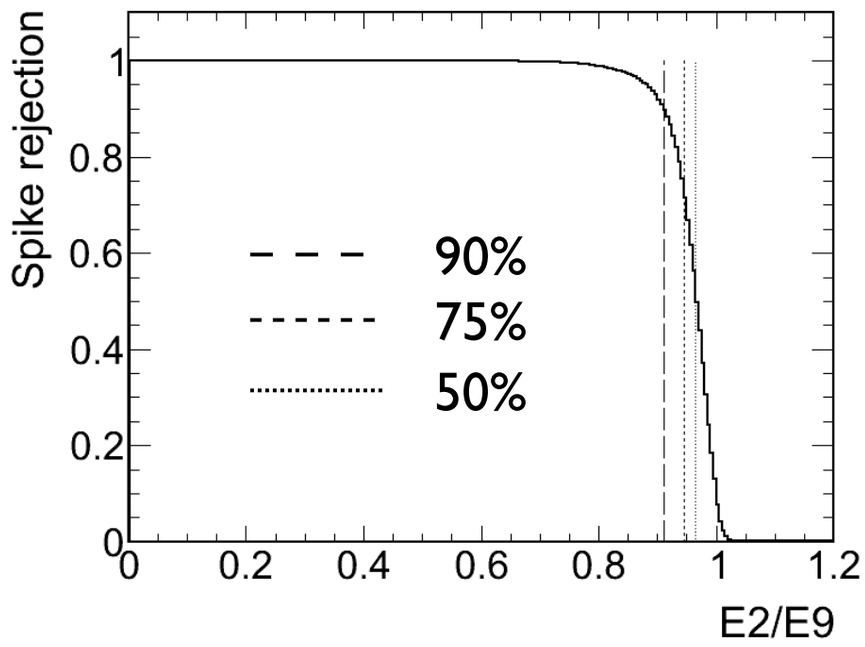
E2 efficiency/rejection: Data and MC

Define control sample of non-isolated spikes based on Swiss-Cross and timing

E_T^{1st}>5 GeV, 1-E4/E1<0.95 t^{1st}/σ<-5

MinBias data





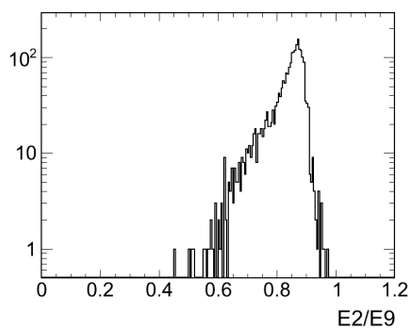
calculate threshold values of E2/E9 for 90,75,50% rejection efficiency

Find cut values corresponding to 90,75,50% spike rejection efficiency

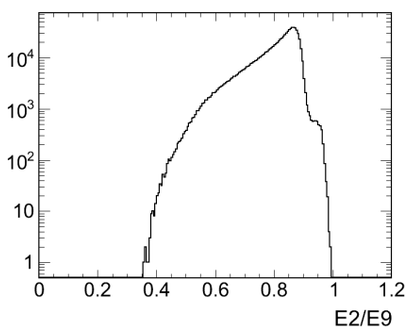
E2/E9 distributions, MC

E_T^{1st}>5 GeV

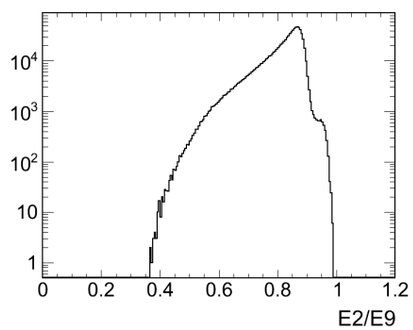
MC: minbias



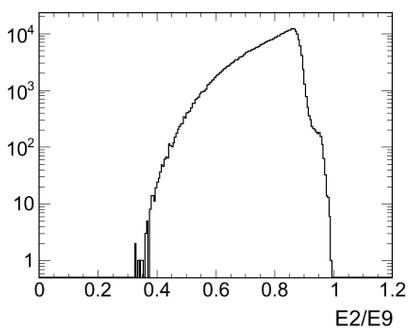
MC: wenu



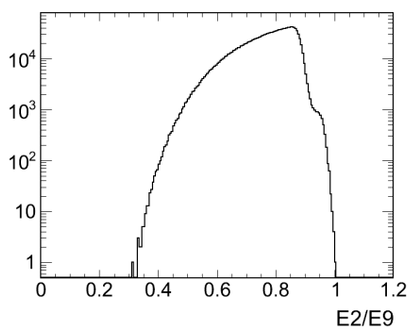
MC: zee



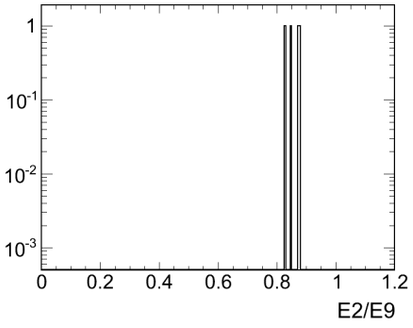
MC: ttbar



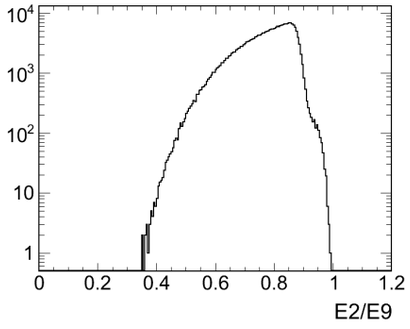
MC: qcd_em_80_170



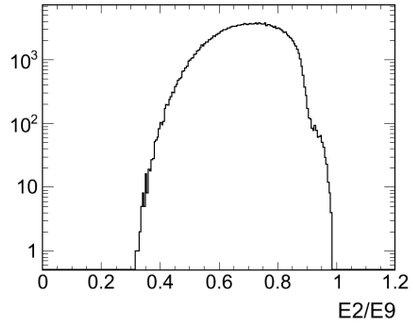
MC: qcd_0_15



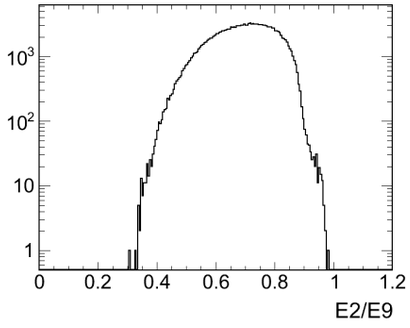
MC: qcd_80_120



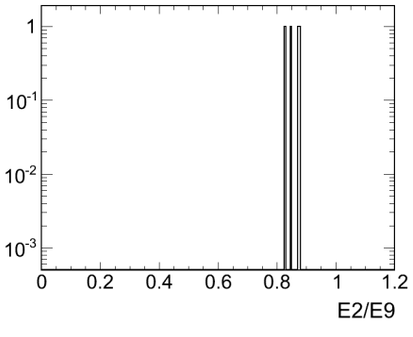
MC: qcd_300_380



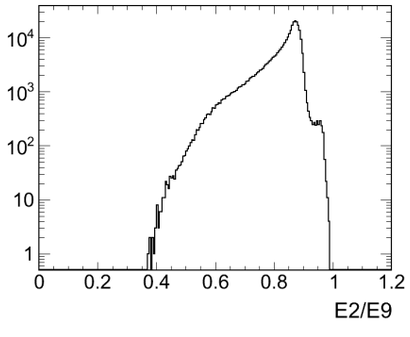
MC: qcd_600_800



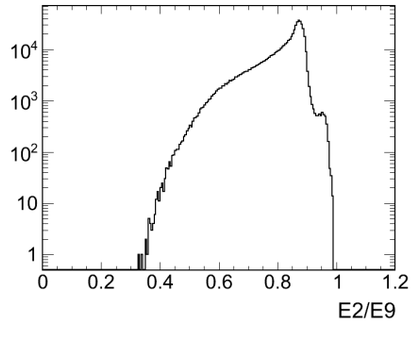
MC: photon15



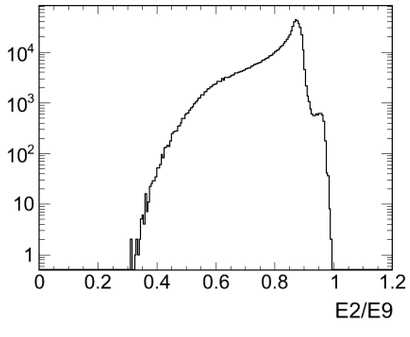
MC: photon30



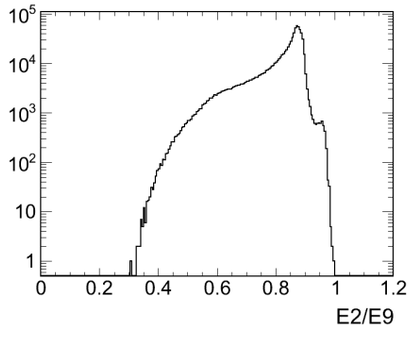
MC: photon80



MC: photon170



MC: photon300



photon samples have largest fraction of events with E2/E9~1

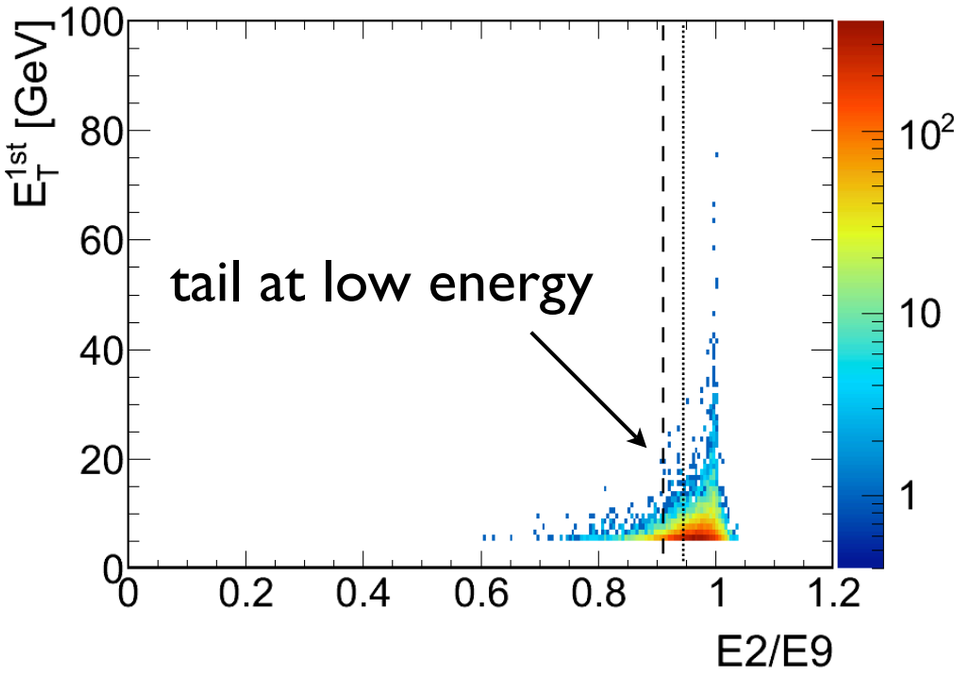
E2/E9 vs energy

Plot E2/E9 vs energy of the leading hit.

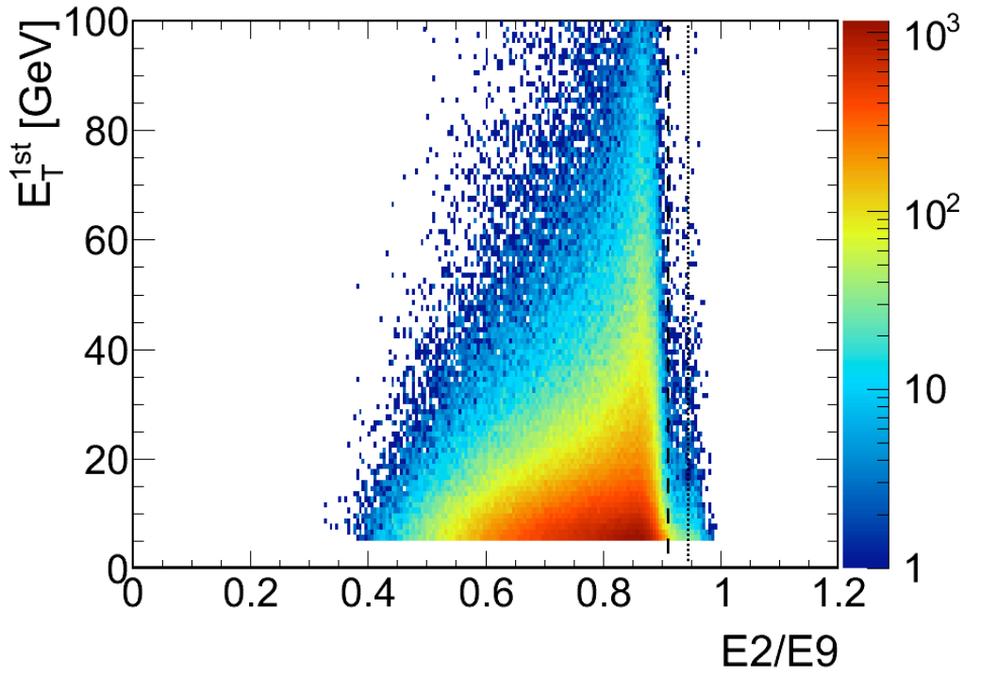
– – – – 90% non-isolated spike rejection

..... 75% non-isolated spike rejection

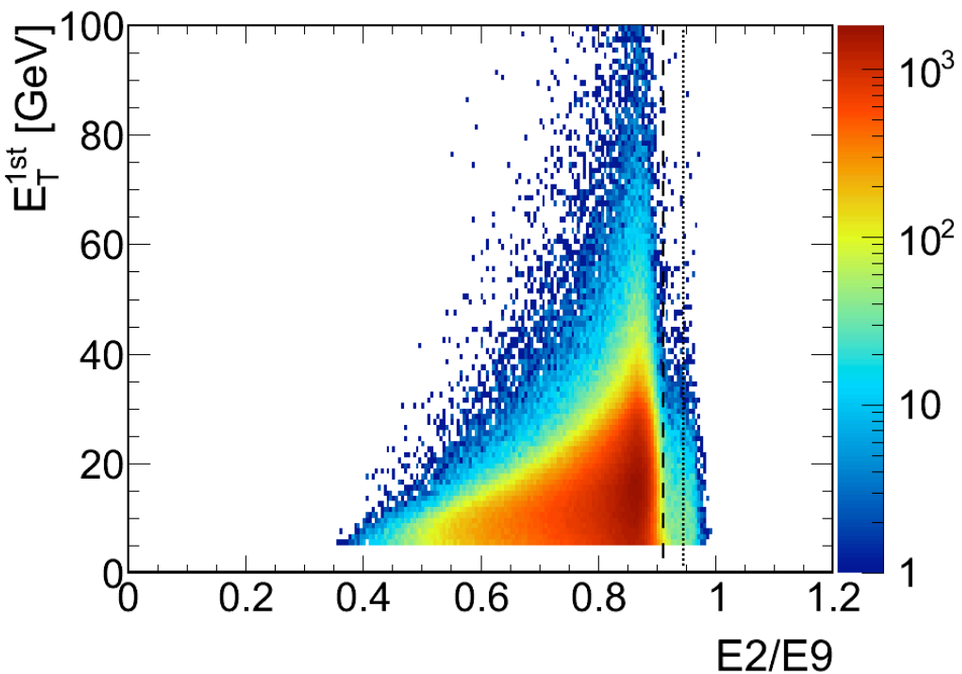
DATA: MinBias



MC: TTbar

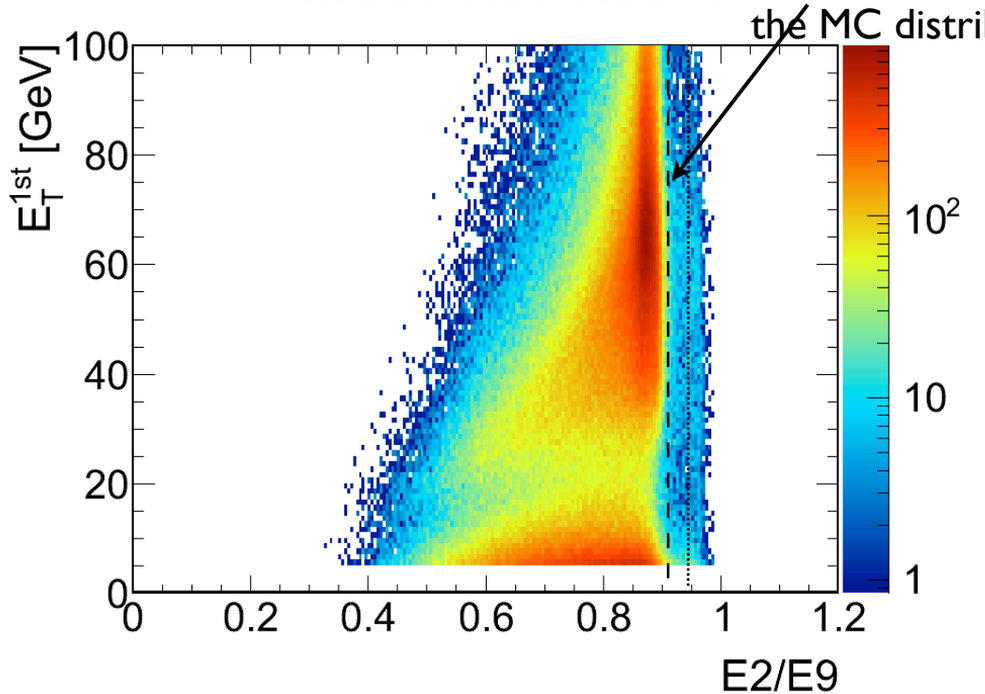


MC: W→eν



MC: Photon80

for photons, 90% efficiency spike cut is on the edge of the MC distribution



Efficiency table: 10 GeV threshold

Cut applied=spike_etmax1>10.0
Events in spike sample=973

	90% efficiency (E2/E9>0.940) (E2/E25>0.895)		**75% efficiency** (E2/E9>0.970) (E2/E25>0.945)		**50% efficiency** (E2/E9>0.990) (E2/E25>0.975)	
	E2/E9	**E2/E25**	**E2/E9**	**E2/E25**	**E2/E9**	**E2/E25**
minbias	9.99e-08	2.00e-07	<9.99e-08	9.99e-08	<9.99e-08	<9.99e-08
ttbar	6.42e-04	7.69e-04	2.53e-05	5.54e-05	<1.58e-06	<1.58e-06
wenu	8.79e-04	1.05e-03	3.92e-05	6.60e-05	<4.79e-07	4.79e-07
zee	8.55e-04	9.96e-04	4.60e-05	7.59e-05	<3.93e-07	1.18e-06
photon15	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06
photon30	1.54e-03	2.24e-03	8.24e-05	3.59e-04	<1.21e-06	2.42e-06
photon80	2.71e-03	3.71e-03	2.08e-04	6.55e-04	<9.19e-07	1.29e-05
photon170	3.61e-03	5.12e-03	2.53e-04	8.99e-04	<1.05e-06	9.45e-06
photon300	3.65e-03	5.74e-03	2.63e-04	9.04e-04	2.95e-06	1.18e-05
qcd_0_15	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06
qcd_80_120	2.72e-04	3.18e-04	2.67e-05	4.87e-05	<1.57e-06	1.57e-06
qcd_300_380	7.78e-04	6.79e-04	5.31e-05	7.59e-05	<3.79e-06	<3.79e-06
qcd_600_800	4.30e-04	3.58e-04	1.45e-05	1.45e-05	<4.83e-06	<4.83e-06
qcd_em_80_170	3.99e-04	4.66e-04	2.92e-05	5.65e-05	1.21e-06	2.66e-06

Efficiency/rejection improves when E_T cut is increased from 5 to 10 GeV

E2/E9 better than E2/E25

10% of non-isolated spikes in MinBias sample are above 10 GeV

Cutting harder...

Cut applied=spike_etmax1>10.0 && spike_etmax2>1.0
Events in spike sample=481

	90% efficiency (E2/E9>0.965) (E2/E25>0.930)		**75% efficiency** (E2/E9>0.990) (E2/E25>0.975)		**50% efficiency** (E2/E9>0.995) (E2/E25>0.990)	
	E2/E9	**E2/E25**	**E2/E9**	**E2/E25**	**E2/E9**	**E2/E25**
minbias	<9.99e-08	<9.99e-08	<9.99e-08	<9.99e-08	<9.99e-08	<9.99e-08
ttbar	5.54e-05	1.52e-04	<1.58e-06	<1.58e-06	<1.58e-06	<1.58e-06
wenu	8.57e-05	2.32e-04	<4.79e-07	<4.79e-07	<4.79e-07	<4.79e-07
zee	9.75e-05	2.33e-04	<3.93e-07	3.93e-07	<3.93e-07	<3.93e-07
photon15	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06
photon30	2.33e-04	8.13e-04	<1.21e-06	2.42e-06	<1.21e-06	<1.21e-06
photon80	4.92e-04	1.49e-03	<9.19e-07	9.19e-06	<9.19e-07	<9.19e-07
photon170	6.70e-04	2.16e-03	<1.05e-06	9.45e-06	<1.05e-06	<1.05e-06
photon300	6.70e-04	2.28e-03	<1.97e-06	1.08e-05	9.85e-07	9.85e-07
qcd_0_15	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06	<1.30e-06
qcd_80_120	2.99e-05	7.08e-05	<1.57e-06	<1.57e-06	<1.57e-06	<1.57e-06
qcd_300_380	9.49e-05	1.52e-04	<3.79e-06	<3.79e-06	<3.79e-06	<3.79e-06
qcd_600_800	3.38e-05	6.77e-05	<4.83e-06	<4.83e-06	<4.83e-06	<4.83e-06
qcd_em_80_170	4.95e-05	1.24e-04	2.42e-07	1.21e-06	<2.42e-07	<2.42e-07

require E_T^{1st}> 10 GeV and E_T^{2nd}>1 GeV

cuts down spike sample by a factor of 2, relative to 10 GeV cut on E_T^{1st} only. Factor of ~5 reduction of signal MC inefficiency

Effect of any spike rejection cut should also be qualified on MC, as well as data

Exercise should be repeated using photon candidates, rather than RecHits