Hadron Colliders: Background simulation, Snowmass 2013

Snowmass Energy Frontier Workshop

University of Washington, Seattle, June 30 – July 3rd, 2013

Sanjay Padhi

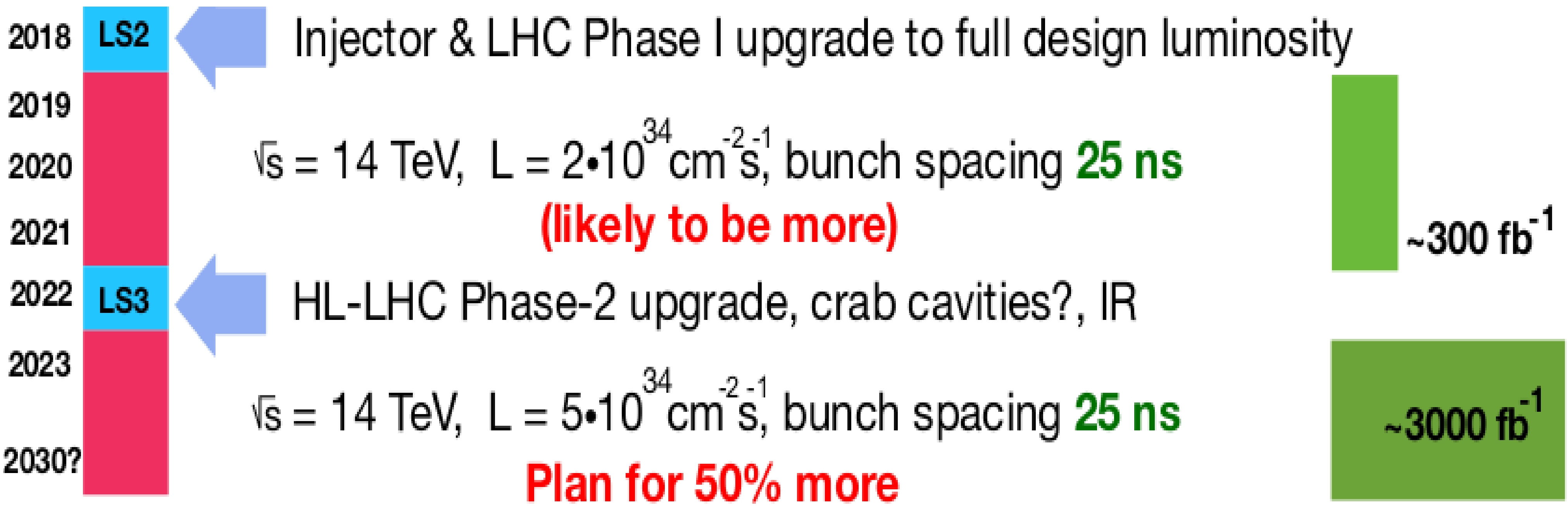
University of California, San Diego

with

Aram Avetisyan, Kevin Black, Raymond Brock, Sergei Chekanov, Timothy Cohen, James Hirschauer, Kiel Howe, Ashutosh Kotwal, Tom LeCompte, Sudhir Malik, Patricia Mcbride, Kalanand Mishra, Meenakshi Narain, Jim Olsen, Michael E. Peskin, Marko J. Slyz, John Stupak III, Alexandre Vaniachine, and Jay G. Wacker

Outline

- LHC evolution
- Simulation framework
- Common backgrounds for Snowmass
  - Event weights and cross section
  - Reconstructed objects
- OSG infrastructure and usage
- Summary and Conclusion

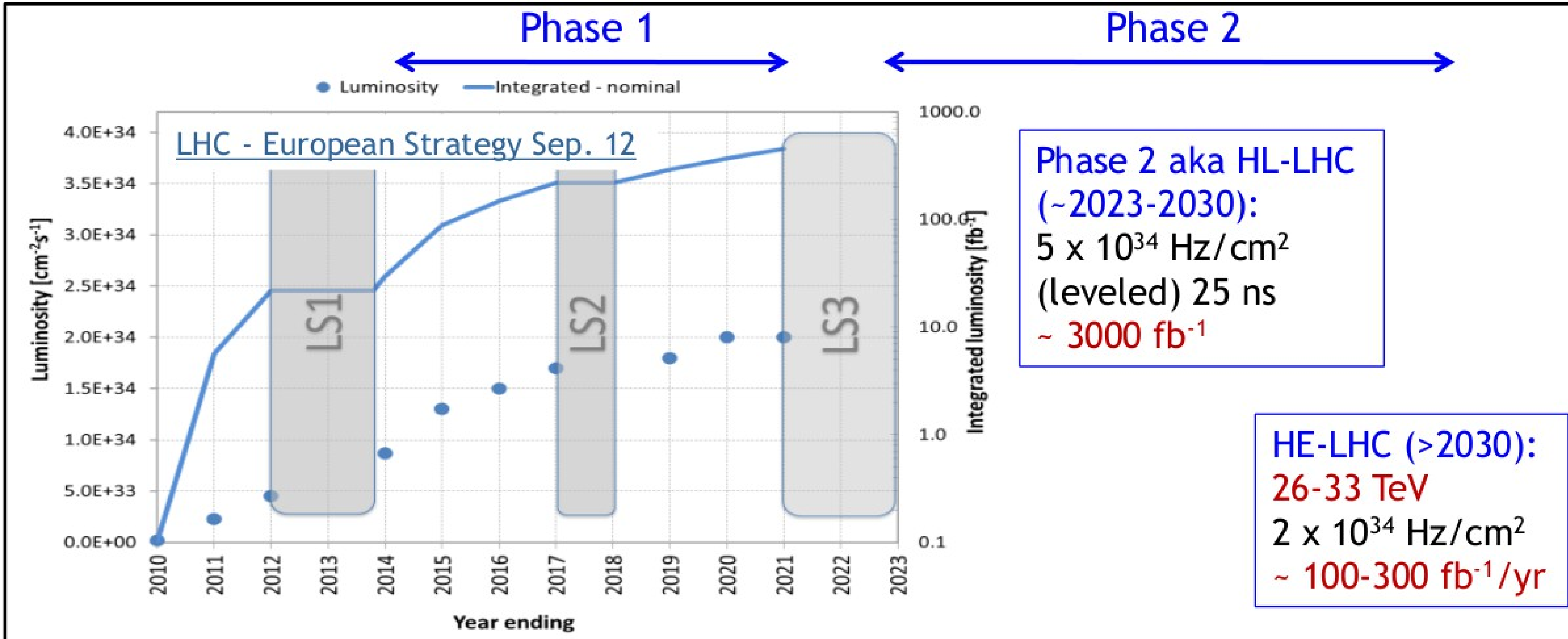


In addition for future energy frontier the following facilities are planned:

- 33 TeV pp collisions with >= 140 Pileups

- 100 TeV VLHC with 40 Pileups

LHC Evolution



LHC Phase-I: 13/14 TeV pp collisions with 50 – 80 pileup events

LHC Phase-II (HL-LHC): 13/14 TeV pp collisions with ~140 pileup events

LS1-LS2 baseline: 0.8 → 1.7 x 10^{34} Hz/cm^2 at 25 ns. ~300 fb^{-1} by LS2 @ 13-14 TeV

- Alternative with 1.8 x10^{34} Hz/cm^2 at 50 ns with lumi-leveling.

After LS2 injection chain upgrades: 25 ns will allow ≥ 2 x 10^{34} Hz/cm^2

Simulation framework for Snowmass

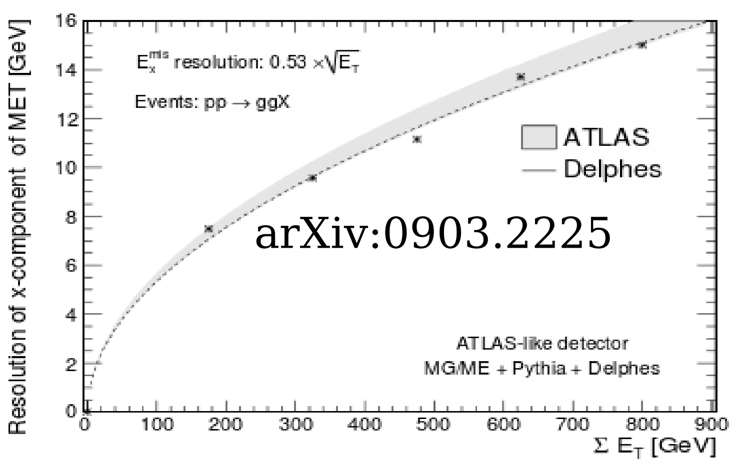
For long range physics planning at Snowmass, we need to make a physics case

- with high luminosity running, higher energy, etc.

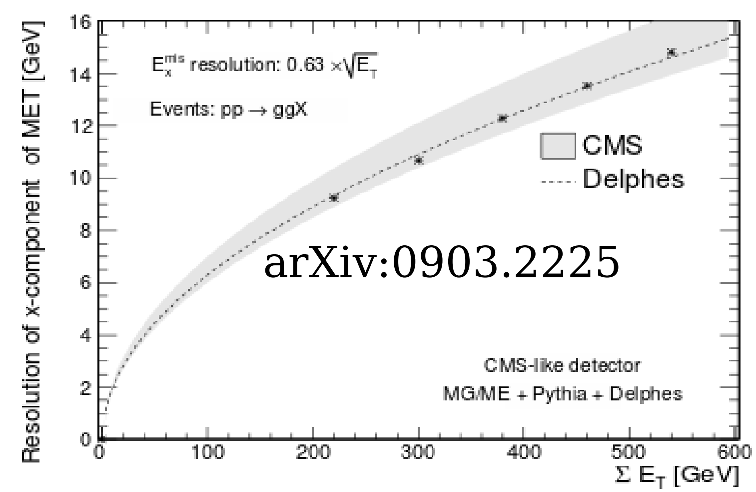
<u>Experiments are currently re-evaluating their full detector simulation framework to accommodate the expected HL performance with large pileup events.</u>

For last EU strategy meeting:

- ATLAS: Simulates the present analyses with parametrized (smeared?) Phase-II detector response with large pileups
- CMS : Extrapolates present results assuming Phase-II detector
  - Data at HL-LHC ~ Same as 2012 data
  - The goal was to retain acceptance, resolution, background and fake rates.



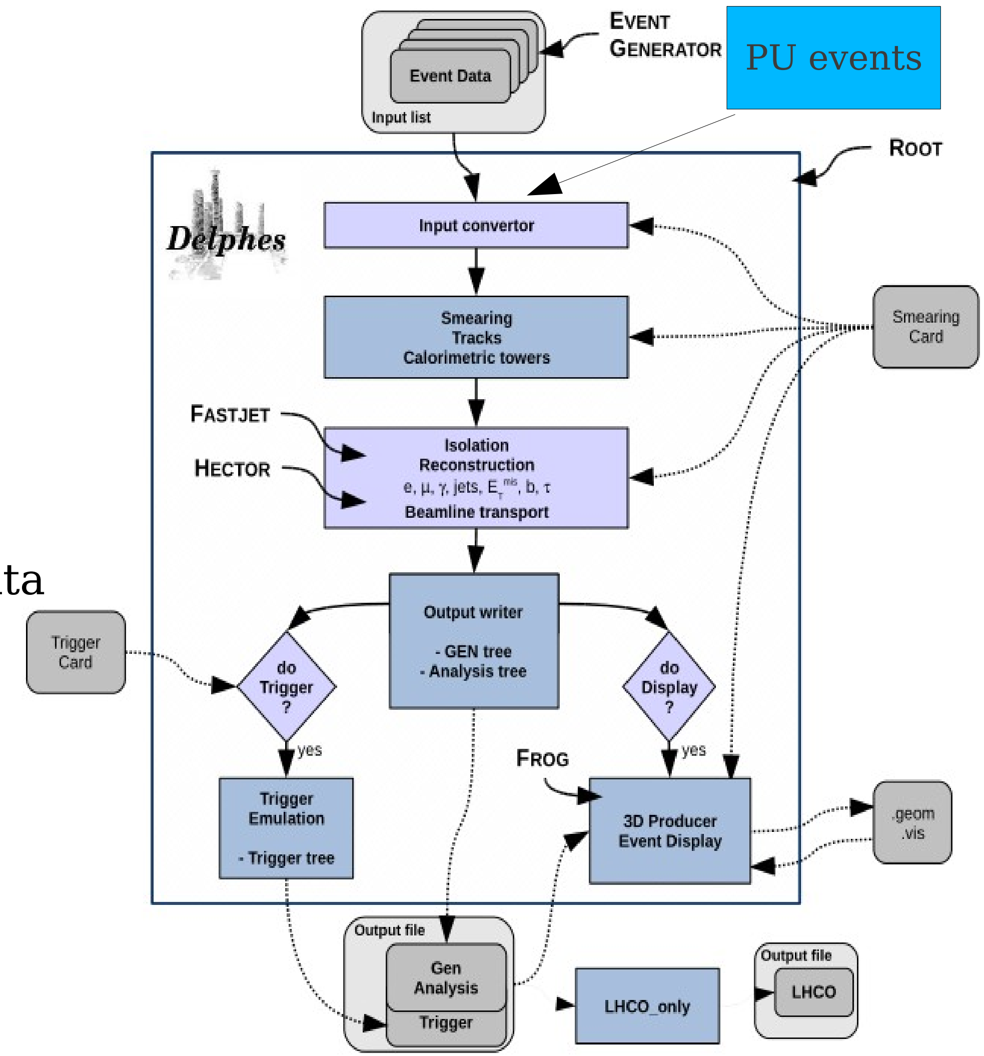
← No PU events →



Simulation framework for Snowmass

Delphes-3 fast simulation (https://cp3.irmp.ucl.ac.be/projects/delphes)

- Delphes3 supports addition of PU events
- Many improvements were motivated based on current studies



For Phase-I studies:

We use Delphes3 framework with:

- realistic detector performance with PU =50
- parameterize using available full simulation
- retain object performance as obtained using data
- use best of both ATLAS/CMS performance (if publicly available)

For Phase-II studies:

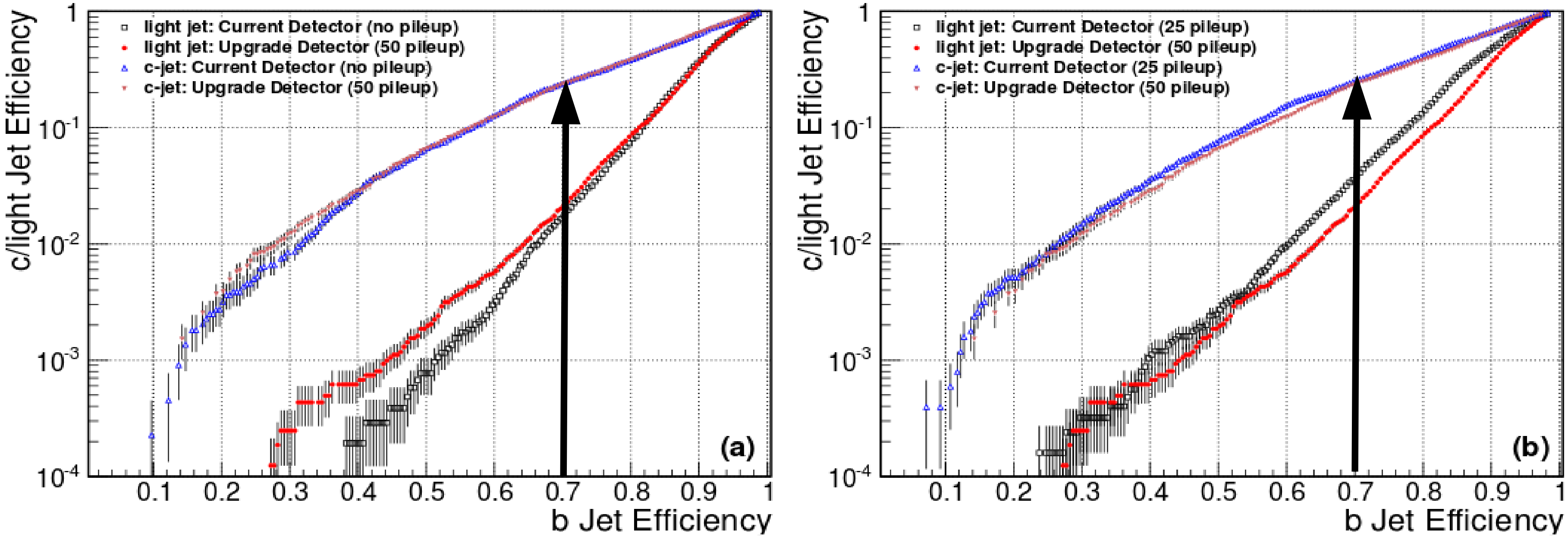
- use higher pileups – 140
- assume the upgraded detector with best available performance
- use best of both ATLAS/CMS expected performance

Validation is crucial for all of these to work

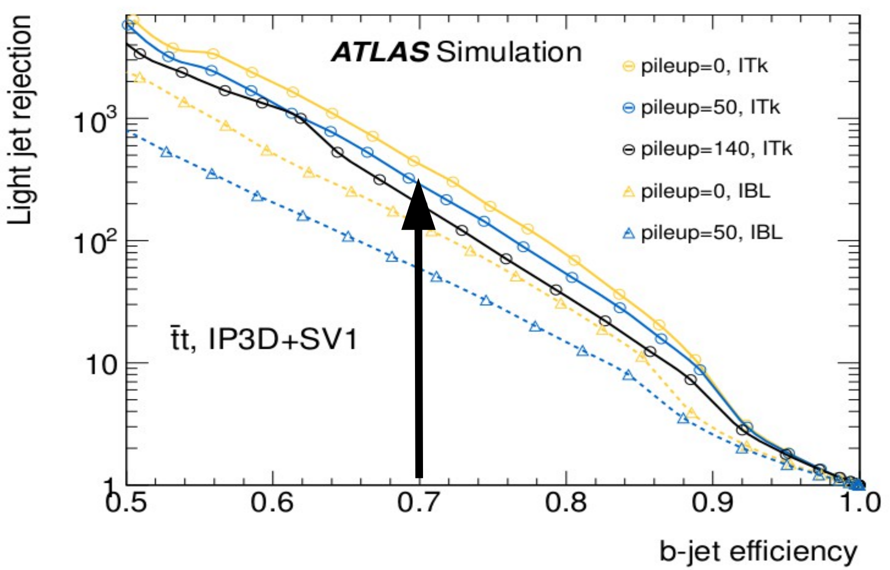
- pileup subtraction is done ala particle flow (for tracks), and jet Rho method for neutrals

Object performance (Example: btag jets)

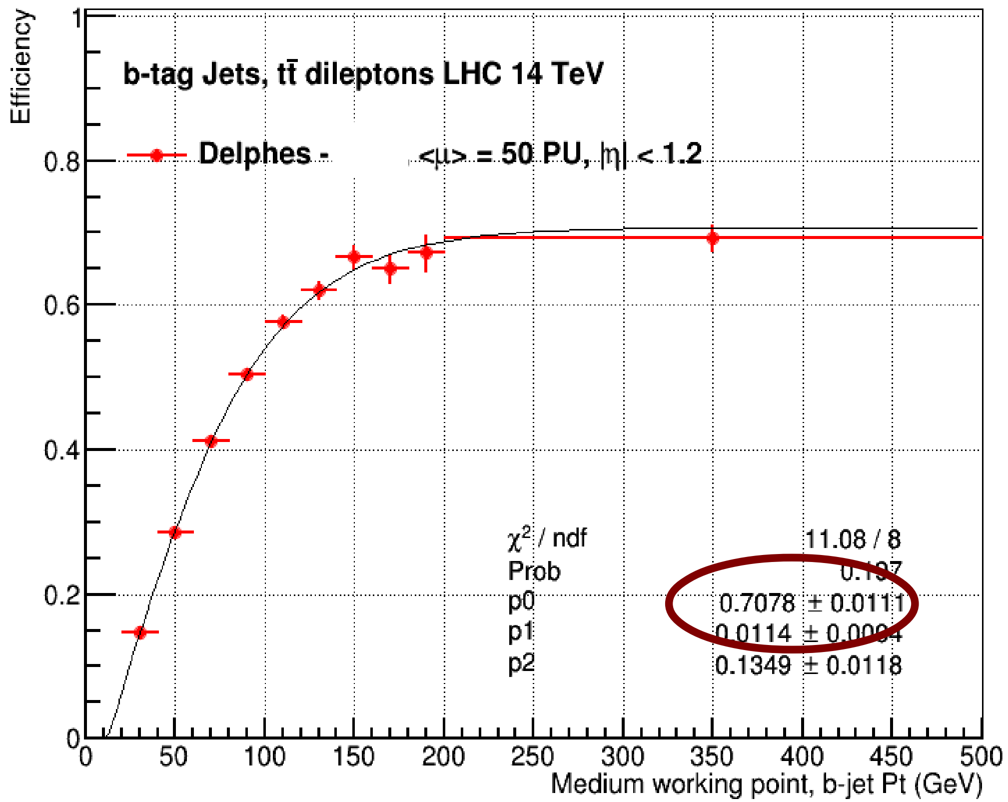
CMS-TDR-011

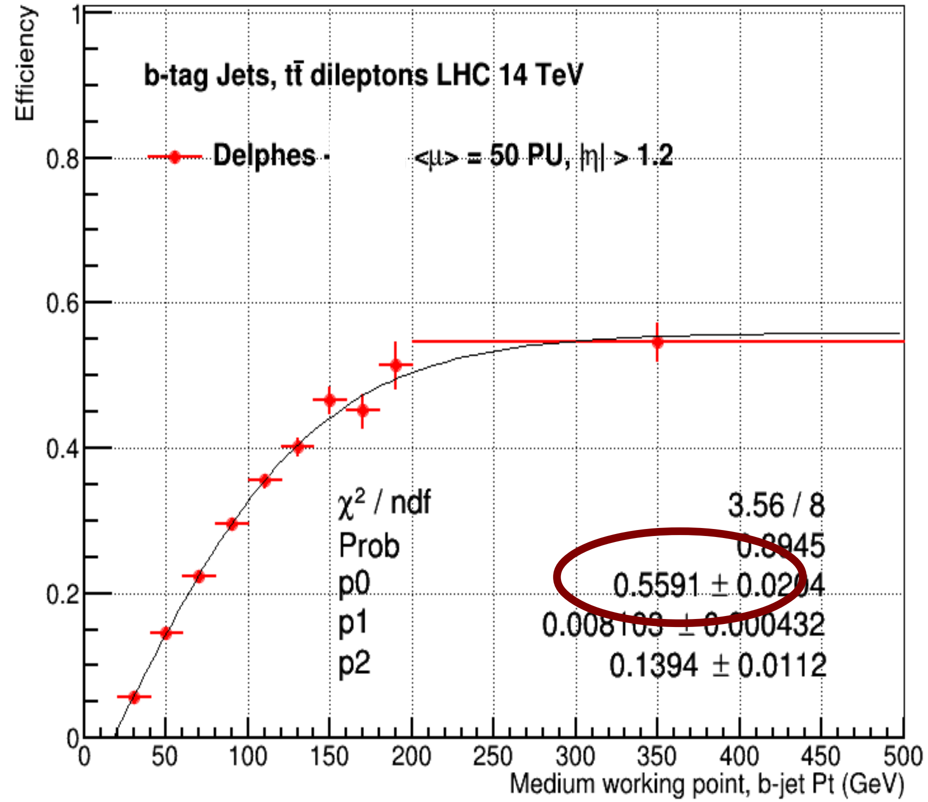


attribute	ATLAS	CMS	Snowmass
B tagging efficiency	70 (65) %	70 (65) %	70 (65) %
Charm mistag rate	15%	15%	15%
Light jet mistag rate	2% (1) %	2.5 (1) %	2% (1%)
Light jet mistag rate (upgrade)	0.5 (0.25) %	1 (0.5) %	0.5 (0.25) %



Object performance (Example: b-tag jets)





Overall b-tag efficiency: ~65% (barrel and endcap) after pileup subtraction:

- PU = 50, Mistag = 1%, btag rate = 65%, c-fake ~ 10%

Efficiency in barrel = 70%, Mistag = 1%, c-fake ~ 15%

Efficiency in endcap = 56%, Mistag = 1%, c-fake ~ 15%

More details:

https://indico.bnl.gov/getFile.py/access?contribId=51&sessionId=2&resId=0&materialId=slides&confId=571

Background samples using the framework

Inclusive samples using 13 and 33 TeV pp collisions are available

- http://red-gridftp11.unl.edu/Snowmass/Inclusive/Delphes-3.0.9.1/

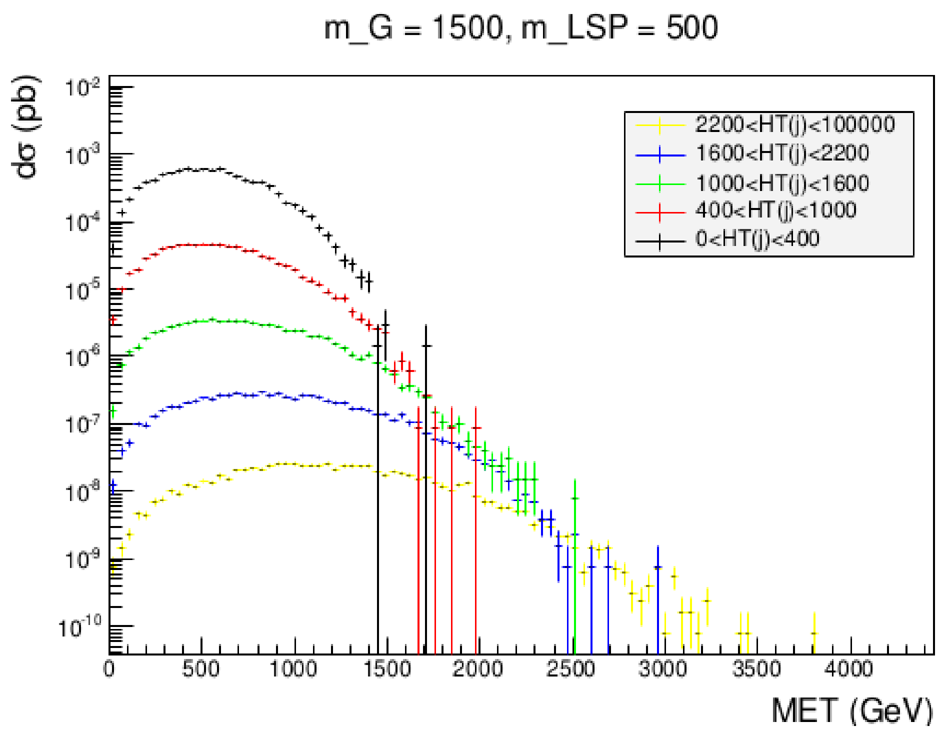
33 TeV

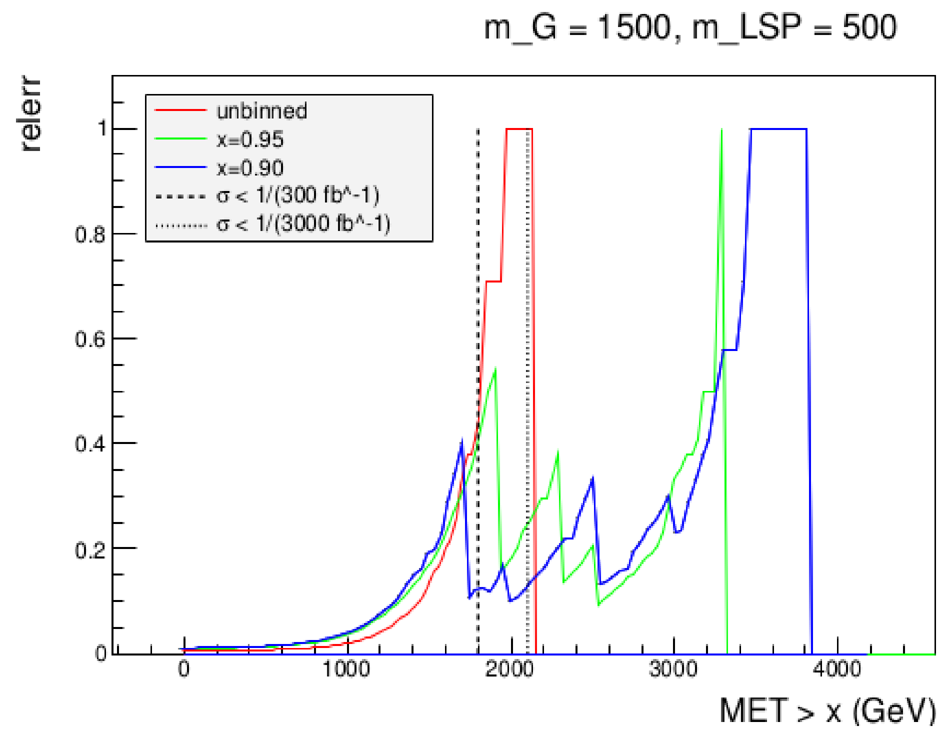
			NoPileUp	50PileUp	140PileUp
Background	QCUT	Sigma pb	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)
DIPHOTONS_33TEV	40	249	7.06 (28.3)	7.06 (28.3)	4.18 (16.8)
PHOTONJETS_33TEV	40	2.23e+05	1.56 (0.00701)	1.55 (0.00697)	0.932 (0.00418)
SSWWJETS_33TEV	40	0.162	7.46 (4.62e+04)	7.56 (4.68e+04)	5.79 (3.58e+04)
TTBARW_33TEV	80	3.32	6.47 (1.95e+03)	6.61 (1.99e+03)	4.92 (1.48e+03)
TTBARZ_33TEV	80	6.21	0.0413 (6.66)	0.0399 (6.44)	0.0399 (6.44)
TTBARJets_33TEV	80	4.07e+03	5.79 (1.42)	5.78 (1.42)	3.42 (0.84)
WGJETS_33TEV	40	617	10.8 (17.5)	10.7 (17.3)	6.39 (10.4)
wjetsmad_33TEV	40	8.26e+04	7.75 (0.0938)	7.66 (0.0927)	4.89 (0.0592)
WWW_33TEV	40	1.05	8.87 (8.46e+03)	8.93 (8.51e+03)	5.42 (5.17e+03)
WWZ_33TEV	40	0.812	10.5 (1.3e+04)	10.4 (1.28e+04)	5.93 (7.31e+03)
WW_33TeV	40	159	11.4 (71.3)	11.2 (70.3)	6.65 (41.8)
WZ3LNUJETS_33TEV	40	3.46	7.58 (2.19e+03)	7.59 (2.19e+03)	4.34 (1.25e+03)
WZZ_33TEV	40	0.276	6.14 (2.23e+04)	6.12 (2.22e+04)	3.58 (1.3e+04)
WZ_33TeV	40	47.8	11.1 (233)	11.1 (231)	6.53 (136)
ZGJETS_33TEV	40	167	10.3 (62)	10.4 (62.1)	6.28 (37.6)
ZJETS_33TEV	40	7.87e+03	0.533 (0.0677)	0.536 (0.0682)	0.551 (0.07)
ZZ4LJETS_33TEV	40	0.277	8.32 (3e+04)	8.36 (3.02e+04)	4.92 (1.78e+04)
ZZJETS_33TEV	40	5.02	8.4 (1.67e+03)	8.38 (1.67e+03)	5.03 (1e+03)
ZZZ_33TEV	40	0.0462	6.38 (1.38e+05)	6.41 (1.39e+05)	3.83 (8.28e+04)

Difficult to simulate full statistics for samples with large cross-sections

H_T binned samples (with weights):

- 1:1 MC generation will require large statistics (beyond the current scope)
- $\sigma(t\bar{t}) \sim 900\ pb$; one needs 2.7 (27) billion events for 300 (3000) fb^{-1} of lumi
- For 30% relative uncert. in tails, one needs 10x the equivalent stat.





Weighted sample reduces the error drastically in the tails

Background samples using the framework

H_T binned samples (with weights):

$$J = \{g, u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, b, \bar{b}\}$$

$$L = \{e^+, e^-, \mu^+, \mu^-, \tau^+, \tau^-\}$$

$$B = \{W^+, W^-, Z^0, h^0, \gamma\}$$

$$T = \{t, \bar{t}\}$$

Binned in H_T:

- Scalar sum p_T at gen-level

Dominant: 4-7 bins of H_T

Sub-dominant: 2-3 bins in H_T

> 5M events per bin

Rare backgrounds have

- No H_T binning

Process	Order	14 TeV	33 TeV	100 TeV
Dominant Backgrounds				
$B + n_4 J$	$\mathcal{O}(\alpha_s^{n_4}\alpha_w)$	SGLKp	SGk	sk
$TT + n_3 J$	$\mathcal{O}(\alpha_s^{2+n_3})$	SGlk	SGk	sk
$BB + n_2 J$	$\mathcal{O}(\alpha_s^{n_2}\alpha_w^2)$	SGlk	SGk	sk
$TB + n_2 J$	$\mathcal{O}(\alpha_s^{n_2+1}\alpha_w)$	SGlk	SGk	sk
$T + n_3 J$	$\mathcal{O}(\alpha_s^{n_3-1}\alpha_w^2)$	SGlk	SGk	sk
$LL + n_3 J$	$\mathcal{O}(\alpha_s^{n_3}\alpha_w^2)$	SGlk	SGk	sk
Subdominant Backgrounds				
$TTB + n_1 J$	$\mathcal{O}(\alpha_s^{2+n_1}\alpha_w)$	SG	SG	
$BLL + n_1 J$	$\mathcal{O}(\alpha_s^{n_1}\alpha_w^3)$	SG	SG	
$B + n_3 J$	$\mathcal{O}(\alpha_s^{n_3}\alpha_h)$	SG	SG	
$B + n_3 J$	$\mathcal{O}(\alpha_s^{n_3-2}\alpha_w^3)$	SG	SG	

Background samples using the framework

H_T binned samples (with weights):

http://www.snowmass2013.org/tiki-index.php?page=Energy_Frontier_FastSimulation

14 TeV (HT binned samples)

		NoPileUp	50PileUp	140PileUp		
Background	QCUT	Sigma pb	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)	Comments
tt-4p-0-600-v1510_14TEV	80	531	2.19 (4.13)	0.961 (1.81)	0.653 (1.23)	ttbar + jets, 0 < HT (GeV) < 600
tt-4p-600-1100-v1510_14TEV	80	42.5	14.3 (337)	5.93 (139)	4.93 (116)	ttbar + jets, 600 < HT (GeV) < 1100
tt-4p-1100-1700-v1510_14TEV	80	4.48	13.6 (3.05e+03)	5.71 (1.27e+03)	4.94 (1.1e+03)	ttbar + jets, 1100 < HT (GeV) < 1700
tt-4p-1700-2500-v1510_14TEV	80	0.528	13.7 (2.6e+04)	5.72 (1.08e+04)	4.87 (9.21e+03)	ttbar + jets, 1700 < HT (GeV) < 2500
tt-4p-2500-100000-v1510_14TEV	80	0.0545	13.8 (2.52e+05)	5.46 (1e+05)	5.19 (9.51e+04)	ttbar + jets, 2500 < HT (GeV) < 100000

Statistics up to several ab^{-1} have been generated for 14 TeV collisions

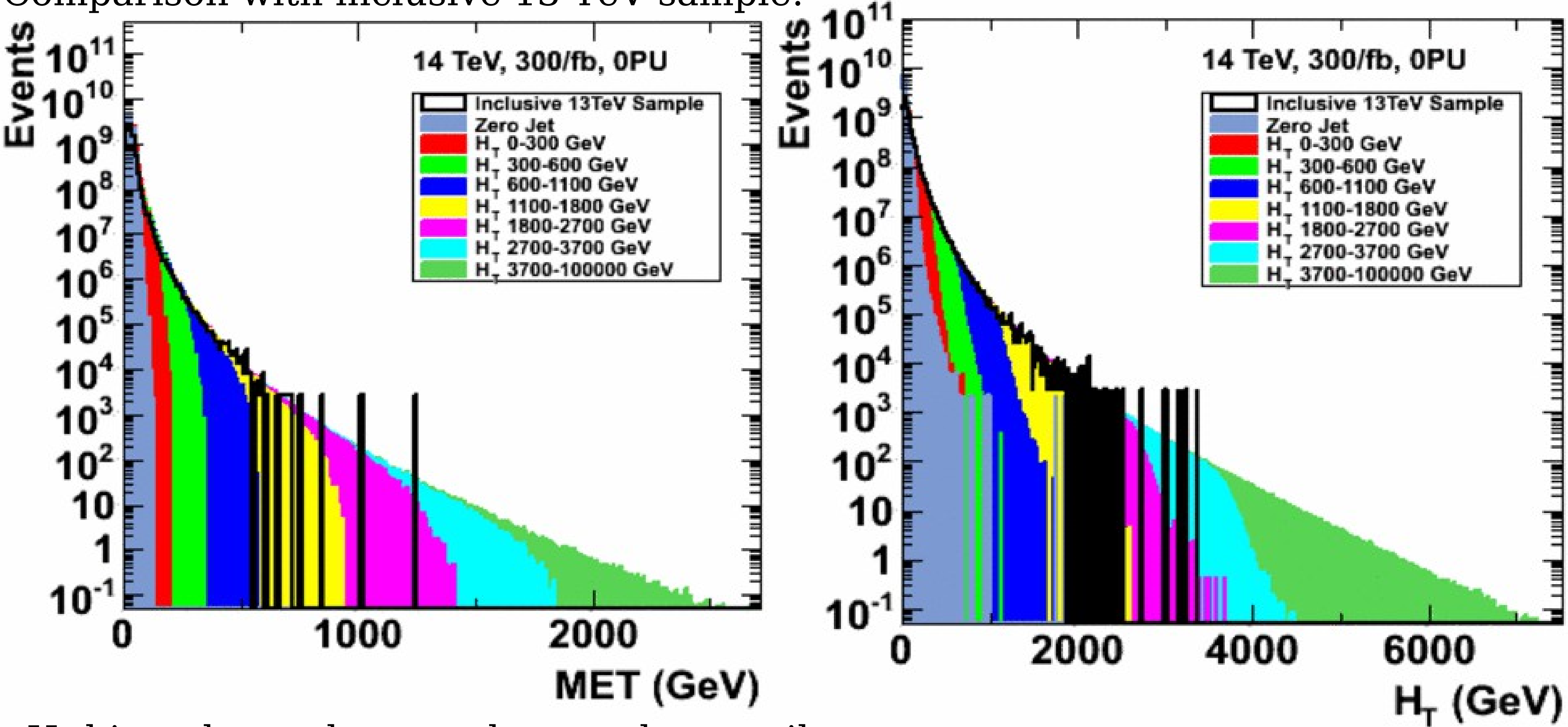
Internal weights include NLO kfactors.

33 TeV and 100 TeV samples are in process (expect to be done within a week)

Background samples using the framework

Location: http://red-gridftp11.unl.edu/Snowmass/HTBinned/Delphes-3.0.9.1/

Comparison with inclusive 13 TeV sample:



H_T binned sample extends up to large tails

Important for new physics searches with large MET and/or H_T

- Extrapolations based on low stat or (14/8) energy ratio's can have huge errors

Background samples using the framework

http://www.snowmass2013.org/tiki-index.php?page=Energy_Frontier_FastSimulation

14 TeV (HT binned samples)

		NoPileUp	50PileUp	140PileUp			
Background	QCUT	Sigma pb	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)	Events Mil (Lumi 1/fb)	Comments	
B-4p-0-1-v1510_14TEV	40	2.05e+05	30.2 (0.147)	11.3 (0.055)	10.7 (0.0521)	Boson (photon, W, Z) + Jets	
Bj-4p-0-300-v1510_14TEV	40	3.44e+04	57.3 (1.67)	8.7 (0.253)	10.7 (0.31)	Boson (photon, W, Z) + Jets, 0 < HT (GeV) < 300	
Bj-4p-300-600-v1510_14TEV	40	2.64e+03	41.4 (15.7)	6.32 (2.4)	7.25 (2.75)	Boson (photon, W, Z) + Jets, 300 < HT (GeV) < 600	
Bj-4p-600-1100-v1510_14TEV	40	294	36.4 (124)	5.7 (19.4)	6.36 (21.6)	Boson (photon, W, Z) + Jets, 600 < HT (GeV) < 1100	
Bj-4p-1100-1800-v1510_14TEV	40	25.9	35.9 (1.39e+03)	5.49 (212)	6.66 (257)	Boson (photon, W, Z) + Jets, 1100 < HT (GeV) < 1800	
Bj-4p-1800-2700-v1510_14TEV	40	2.42	33.1 (1.37e+04)	5.07 (2.09e+03)	6.46 (2.67e+03)	Boson (photon, W, Z) + Jets, 1800 < HT (GeV) < 2700	
Bj-4p-2700-3700-v1510_14TEV	40	0.227	32.9 (1.45e+05)	4.85 (2.14e+04)	6.02 (2.65e+04)	Boson (photon, W, Z) + Jets, 2700 < HT (GeV) < 3700	
Bj-4p-3700-100000-v1510_14TEV	40	0.0276	33.6 (1.22e+06)	5.11 (1.85e+05)	5.91 (2.14e+05)	Boson (photon, W, Z) + Jets, 3700 < HT (GeV) < 100000	
Bjj-vbf-4p-0-700-v1510_14TEV	40	86.5	18 (208)	7.07 (81.8)	6.47 (74.8)	Boson (photon, W, Z, h) + >= 2Jets, 0 < HT (GeV) < 700	
Bjj-vbf-4p-700-1400-v1510_14TEV	40	4.35	13.7 (3.16e+03)	5.81 (1.34e+03)	5.26 (1.21e+03)	Boson (photon, W, Z, h) + >= 2Jets, 700 < HT (GeV) < 14000	
Bjj-vbf-4p-1400-2300-v1510_14TEV	40	0.325	11.6 (3.56e+04)	5.02 (1.54e+04)	4.57 (1.41e+04)	Boson (photon, W, Z, h) + >= 2Jets, 1400 < HT (GeV) < 2300	
Bjj-vbf-4p-2300-3400-v1510_14TEV	40	0.0303	4.12 (1.36e+05)	1.81 (5.98e+04)	1.78 (5.87e+04)	Boson (photon, W, Z, h) + >= 2Jets, 2300 < HT (GeV) < 3400	

Accessing Samples

Browse metadata and ROOT files in http (slow sometimes):

```
http://red-gridftp11.unl.edu/Snowmass/HTBinned/Delphes-3.0.9.1/140PileUp/tt-4p-0-600-v1510_14TEV/*.txt

http://red-gridftp11.unl.edu/Snowmass/HTBinned/Delphes-3.0.9.1/140PileUp/tt-4p-0-600-v1510_14TEV/*.root
```

Using storage resource manager (srm) from FNAL (need grid certificate)

```
lcg-ls "srm://srm.unl.edu:8443/srm/v2/server?SFN=/mnt/hadoop/user/Snowmass/

HTBinned/Delphes-3.0.9.1/NoPileUp/tt-4p-0-600-v1510_14TEV/"
```

Access XRootD:

```
TFile* file0 = TFile::Open("root://red-gridftp11.unl.edu//mnt/hadoop/user/

Snowmass/HTBinned/Delphes-3.0.9.1/50PileUp/tt-4p-0-600-v1510_14TEV/

tt-4p-0-600-v1510_14TEV_50PileUp_114845928.root”)
```

Transfer all root files in directory:

```
wget --no-check-certificate -r -l1 -H -t1 -nd -N -np -A root -E

http://red-gridftp11.unl.edu/Snowmass/HTBinned/

Delphes-3.0.9.1/50PileUp/tt-4p-0-600-v1510_14TEV/
```

Event weights and cross sections

Event weights are in the ntuples (Event.Weight):

This can be combined using: $$\sum_{i}^{N_\alpha} \frac{w_i}{N_\alpha} = \sigma_{\mathrm{NLO},\alpha} \times \mathrm{BR}.$$

Accounts for NLO k-factor and event-by-event BR

Delphes->("Event.Weight") * (LO cross section) / (Number of Delphes events processed)

Analyst knows this number

http://www.snowmass2013.org/tiki-index.php?page=Energy_Frontier_FastSimulation#HT_binned_samples

B-4p-0-1-v1510_14TEV	:	200944.68129	+-	25.15587
Bj-4p-0-300-v1510_14TEV	:	34409.92339	+-	10.65700
Bj-4p-300-600-v1510_14TEV	:	2642.85309	+-	1.13482
Bj-4p-600-1100-v1510_14TEV	:	294.12311	+-	0.13649
Bj-4p-1100-1800-v1510_14TEV	:	25.95000	+-	0.01344
Bj-4p-1800-2700-v1510_14TEV	:	2.42111	+-	0.00129
Bj-4p-2700-3700-v1510_14TEV	:	0.22690	+-	0.00016
Bj-4p-3700-100000-v1510_14TEV	:	0.02767	+-	0.00001

Software version and Cards for signal simulation

We use Delphes 3.0.9, with the following configuration along with 14/33/100 TeV pileup files

1. The src code can be found at:

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/spadhi/Snowmass/Delphes/Delphes-3.0.9.tar.gz?view=log

2. Pileup files at 14, 33 and 100 TeV

    - http://uaf-2.t2.ucsd.edu/~spadhi/Snowmass/data/

3. Optimized simulation cards for 0 , 50, 140 and VLHC (cvs head version)

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/spadhi/Snowmass/Cards/delphes_card_Snowmass_NoPileUp.tcl?view=log

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/spadhi/Snowmass/Cards/delphes_card_Snowmass_50PileUp.tcl?view=log

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/spadhi/Snowmass/Cards/delphes_card_Snowmass_140PileUp.tcl?view=log

http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/spadhi/Snowmass/Cards/delphes_card_Snowmass_VLHCPileUp.tcl?view=log

The detector performance remains the same (as shown during the BNL meeting)

https://indico.bnl.gov/getFile.py/access?contribId=51&sessionId=2&resId=0&materialId=slides&confId=571

MC Truth information

By default we do not store all generated particles.

- Only status = 3 particles are stored (space reasons)

Change this behavior for your signal by editing the cards:

Change this:

```
module TreeWriter TreeWriter {
...
#  add Branch Delphes/allParticles Particle GenParticle
   add Branch StatusPid/filteredParticles Particle GenParticle
...
}
```

to this:

```
module TreeWriter TreeWriter {
...
   add Branch Delphes/allParticles Particle GenParticle
#  add Branch StatusPid/filteredParticles Particle GenParticle
...
}
```

Reconstructed Objects

- Electrons

Electrons after isolation

```
std::vector<electron_s> selectron;
for (unsigned int i=0; i<electron.size(); i++) {
 if (!(electron[i].PT > 10)) continue;
 if (!(fabs(electron[i].Eta) < 2.5)) continue;
```

- Muons

Muons after isolation

```
std::vector<muon_s> smuon;
for (unsigned int i=0; i<muon.size(); i++) {
  if (!(muon[i].PT > 10)) continue;
```

- Tau jets

taus

```
std::vector<jet_s> taus;
for (unsigned int i=0; i<jet.size(); i++) {
  if (!(jet[i].PT > 20)) continue;
```

Reconstructed Objects

- Photons

Photons after isolation and PU subtraction

```
std::vector<photon_s> sphoton;
for (unsigned int i=0; i<photon.size(); i++) {
  if (!(photon[i].PT > 20)) continue;
```

- Jets: kt (with cone = 0.5), CA (with cone = 0.8)

Jets after PU subtraction

```
std::vector<jet_s> sjet;
 for (unsigned int i=0; i<jet.size(); i++) {
    if (!(jet[i].PT > 30)) continue;
```

- btag jets

Tight and loose btags

```
if( jet->BTag & (1 << 0) ) { # passes standard b-tagging }
   if( jet->BTag & (1 << 1) ) { # passes loose b-tagging }
```

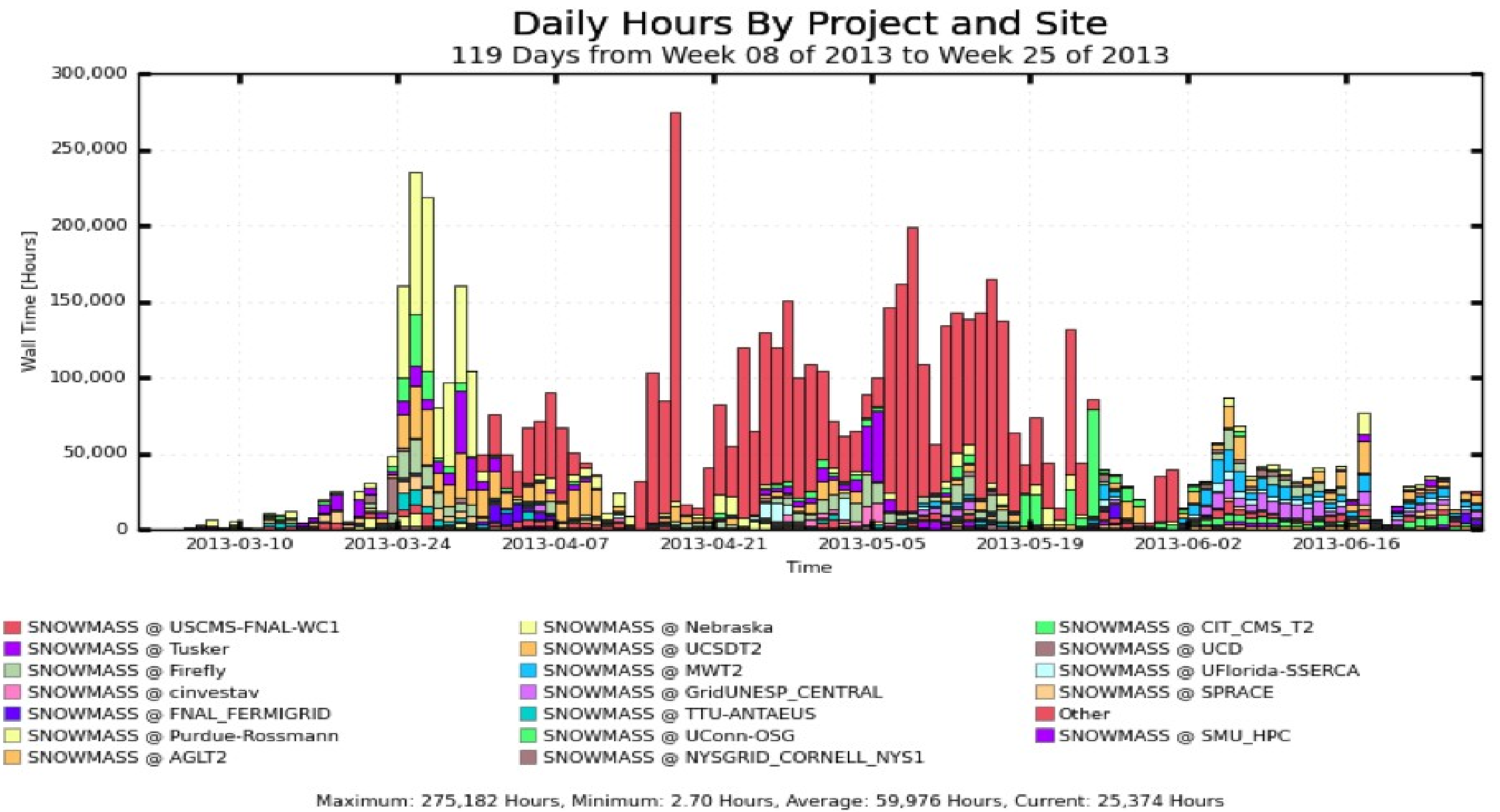
- MET before correction

MET before subtraction

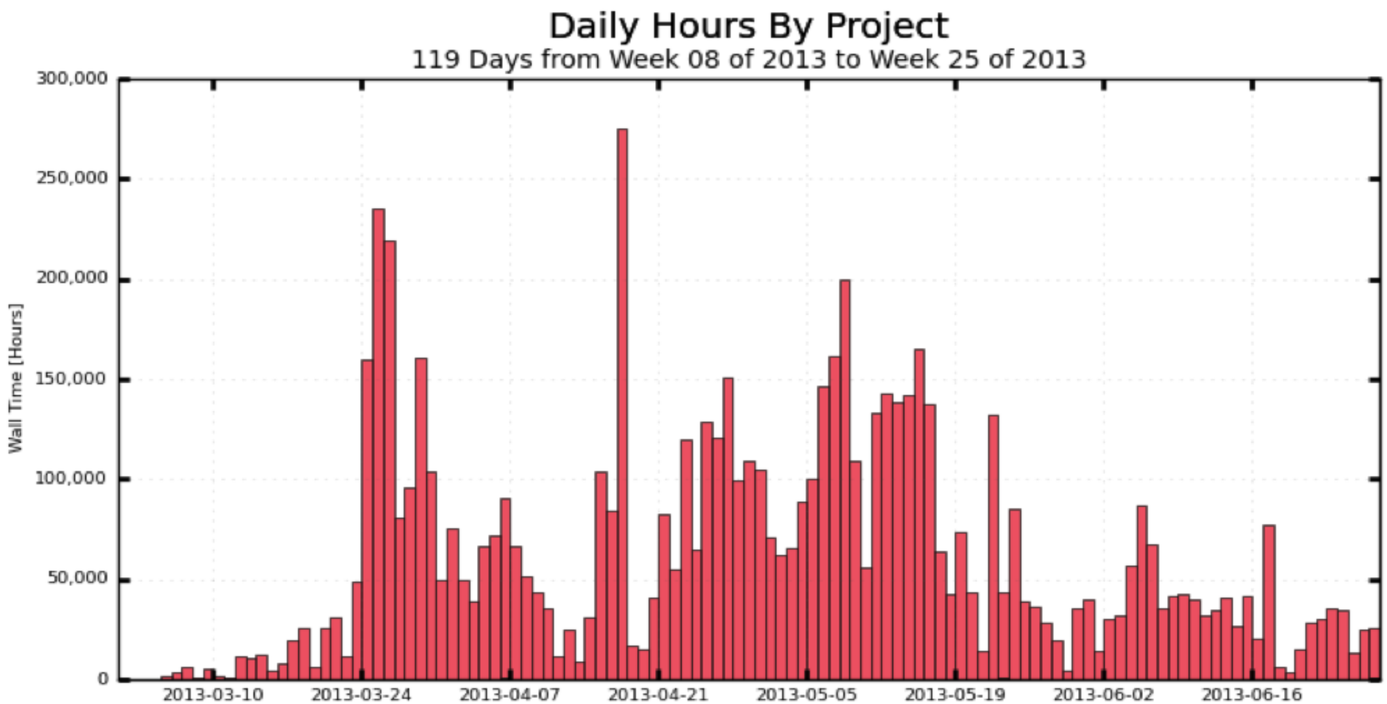
```
double met = missinget[0].MET;
```

OSG infrastructure and usage

We thank the Open Science Grid (OSG) and facilities such as ANL, BNL, FNAL and University of Nebraska (UNL) for providing their infrastructure and resources for the simulation.



OSG infrastructure and usage



Month	Fermi dCache (TB)	UNL (TB)
June	65.0	46.4
May	12.4	5.2
April	189.7	10.8
March	1.1	0.0
Total	268.3	62.5

Total of: ~ 268 TB used at FNAL

Reduced size: 62.5 TB used at UNL

We plan for “compact samples”:

- with ~few 10s of GB per dataset.

Summary and Outlook

It was a huge effort to get ~3000 fb^{-1} of events simulated using 14 TeV LHC

We use parameterized simulation framework with pilups for snowmass studies

All of 14 TeV simulations are completed

We expect to finalize the 33 and 100 TeV simulation within a week (ASAP)

Details on background samples can be found at:

http://www.snowmass2013.org/tiki-index.php?%20page=Energy_Frontier_FastSimulation

Object performance (with pileup) can be found in my BNL Snowmass talk:

https://indico.bnl.gov/getFile.py/access?contribId=51&sessionId=2&resId=0&materialId=slides&confId=571

Please cite the following articles in your white paper studies (we plan to submit the final paper to the arXiv soon):

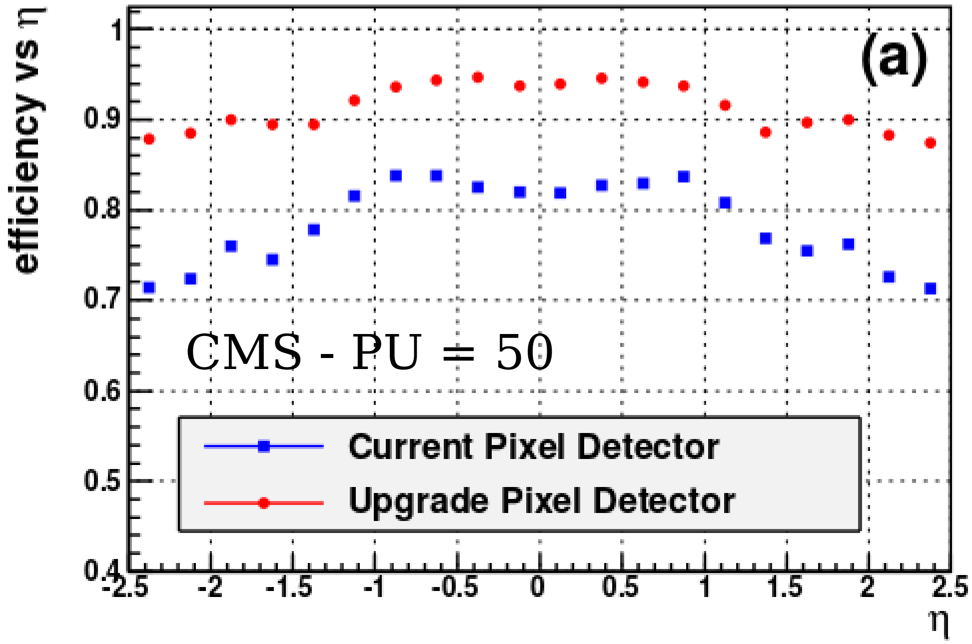
1. "Snowmass Energy Frontier Simulations for Hadron Colliders", A. Avetisyan et. al. arXiv:1307.XXX, July 2013

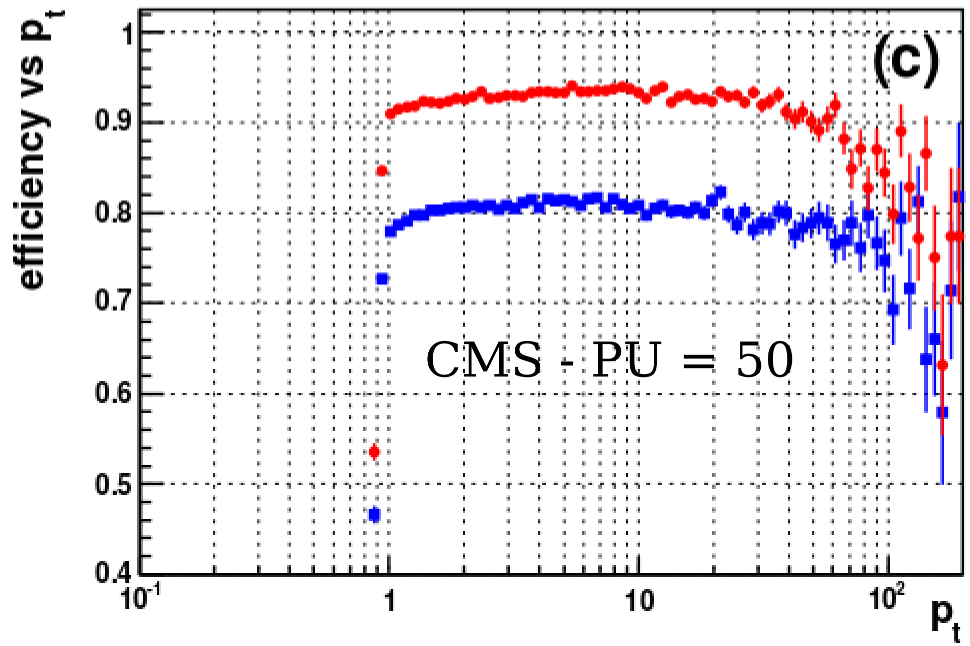
2. "Standard Model Background Generation for Snowmass using Madgraph", A. Avetisyan et. al. arXiv:1307.XXX, July 2013

Backup slides

Tracking performance and expectations

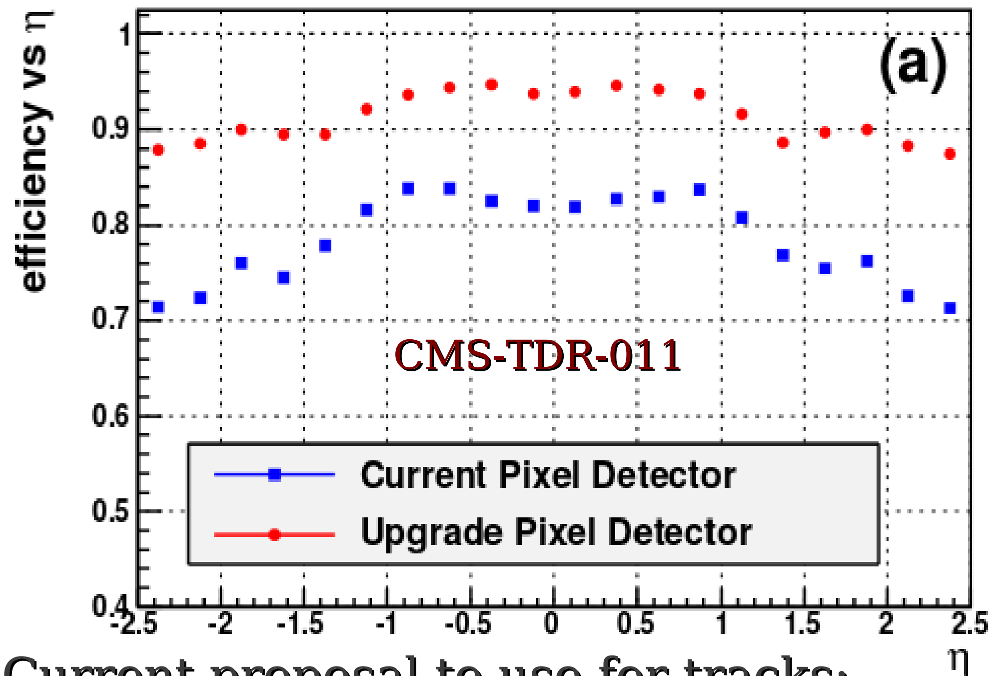
CMS-TDR-011 : http://cds.cern.ch/record/1481838/files/CMS-TDR-011.pdf

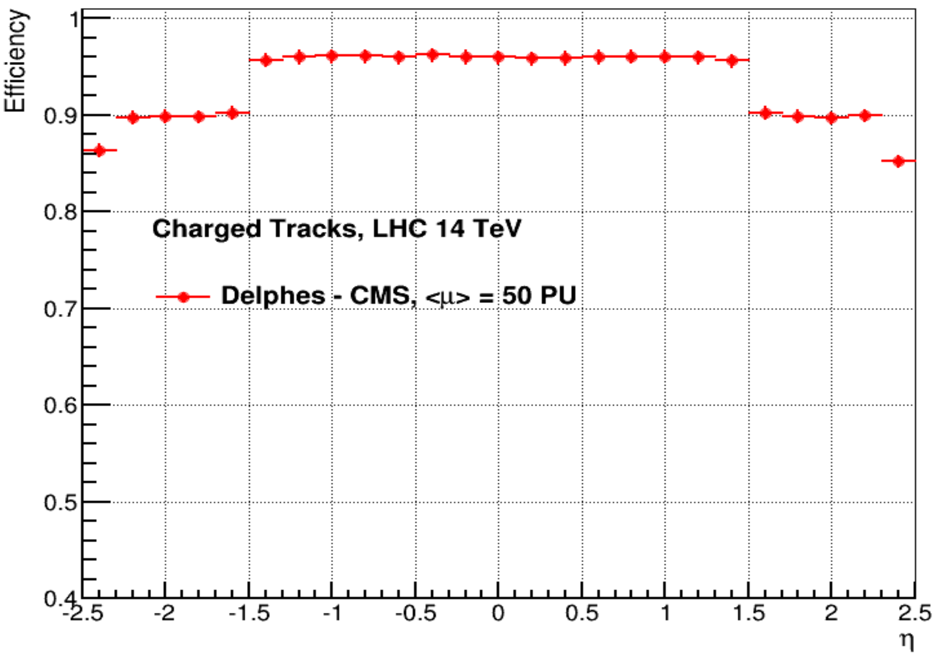




Sample and Conditions		Tracking Efficiency (%)		Track Fake Rate (%)	
Sample	PU/DL/Cuts	Current	Upgrade	Current	Upgrade
Muon	0/No/Cleanup	97.4	98.1	0.0	0.0
Muon	0/Yes/Cleanup	93.9	97.9	0.0	0.0
Muon	50/No/Cleanup	90.1	94.9	0.22	0.17
Muon	50/Yes/Cleanup	81.5	94.4	0.23	0.17

Tracking performance and expectations



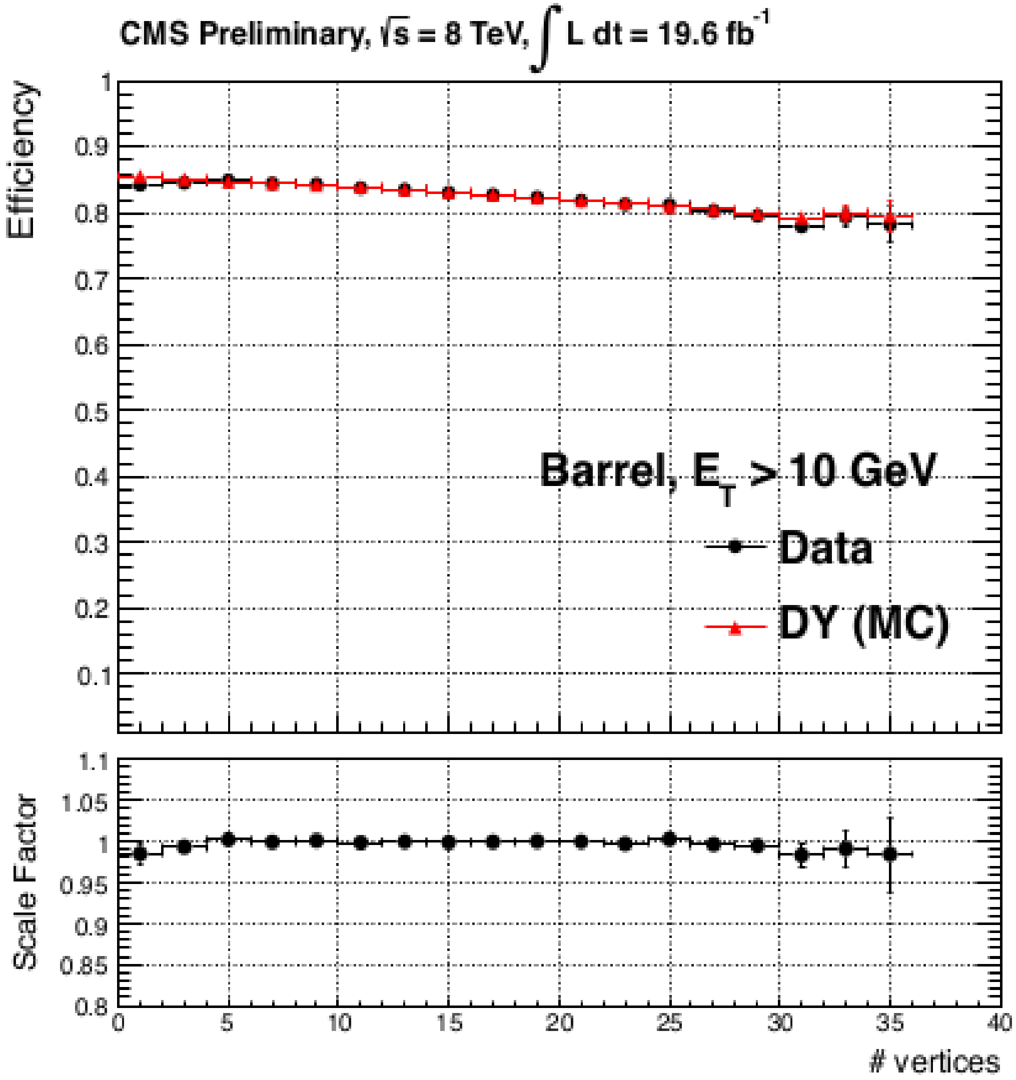


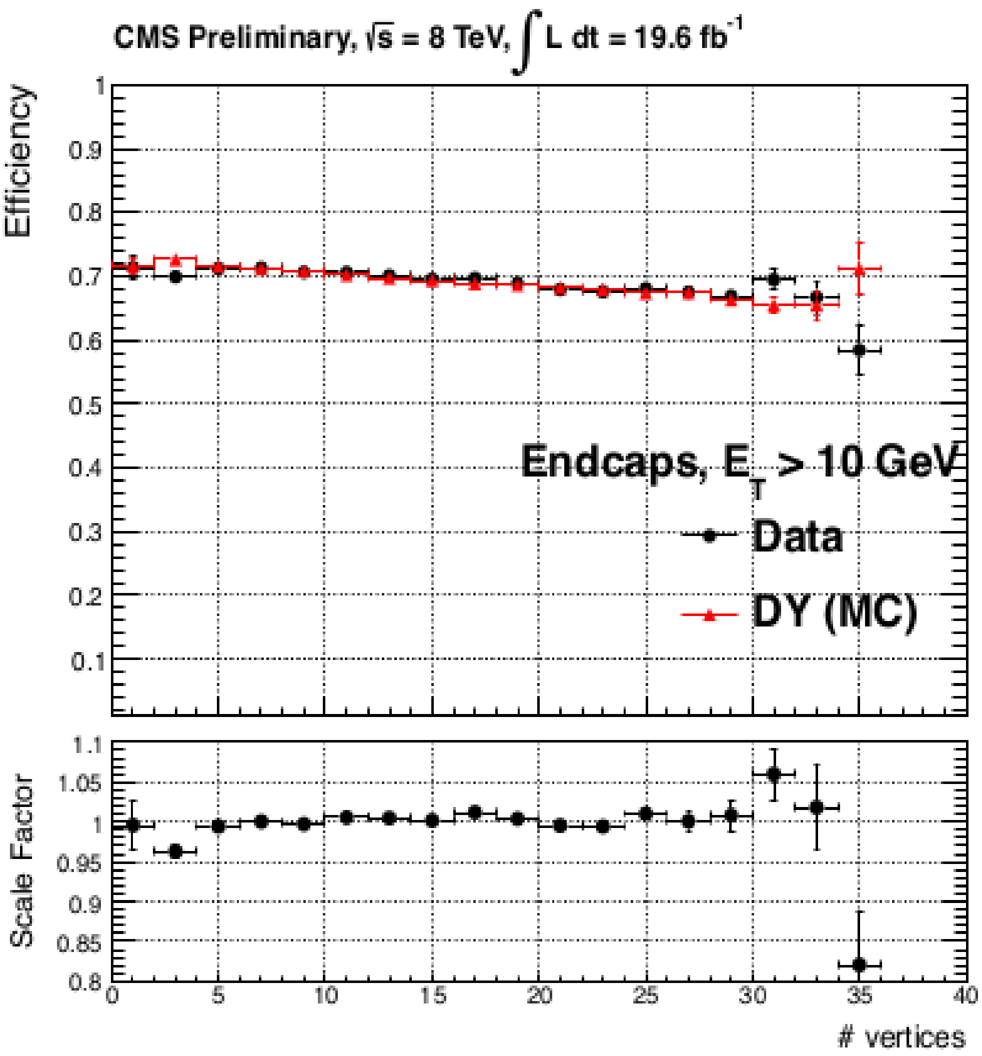
Current proposal to use for tracks:

attribute	ATLAS	CMS	Snowmass
Minimum track p_T to reach calorimeter	0.5 GeV	0.7 GeV	0.6 GeV
Tracking efficiency (DELPHES2)	97%	95%	96%
Muon Efficiency (DELPHES3)	95% / 85%	95% / ~~85%~~	95% / ~~85%~~
Muon Efficiency (upgrade)	95% / 85%	95% / ~~85%~~	95% / ~~85%~~
Electron & Pion Efficiency (upgrade)	95% / 85%	95% / ~~85%~~	95% / ~~85%~~
Momentum resolution @100 GeV (upgrade)	2%	1.5%	1.5%

Electrons (CMS Full Simulation and data)

<u>Electrons (cut based) with medium working point</u> (CMS DP -2013/003)

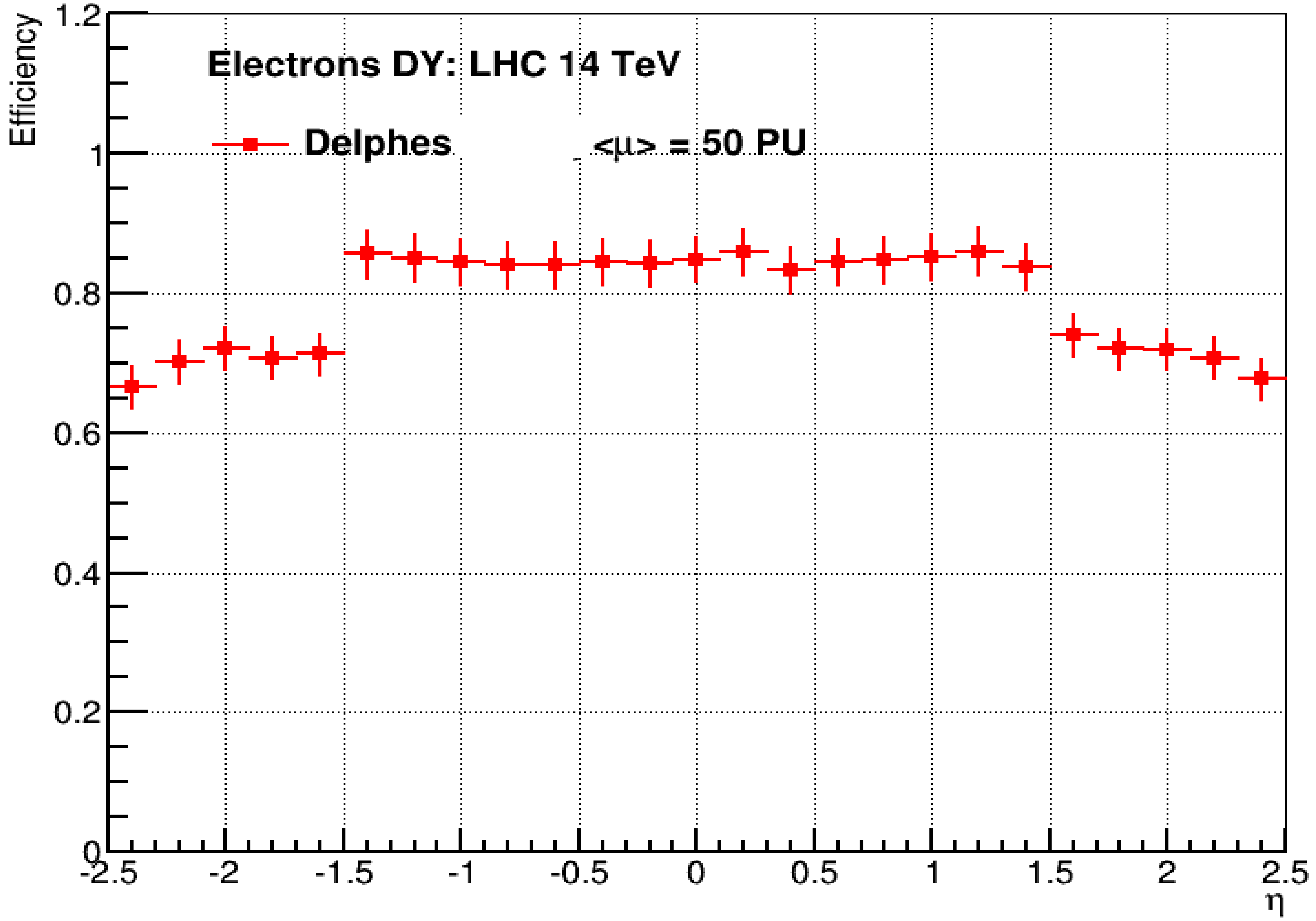




Expected efficiency with 50 PU: ~80% (barrel) and 70% (endcap)

Electrons (Snowmass detector with parameterization)

Electrons (cut based) with medium working point (with 50 PU events)



Simulation agrees with the expectations with 50 PU after subtraction.

We use_this for the combined Snowmass LHC detector