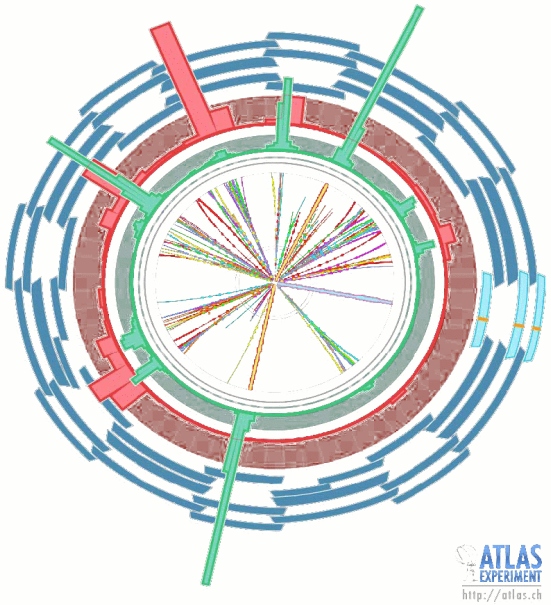
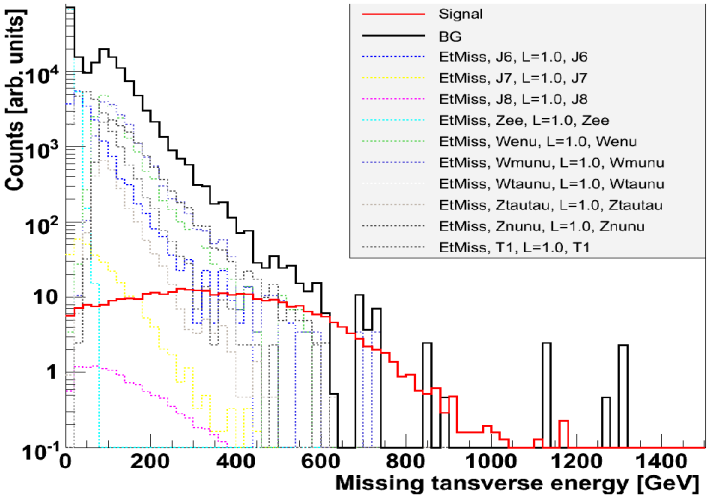
Atlas analysis using NorduGrid
A SUSY example



From here...

...to here...

...using this.

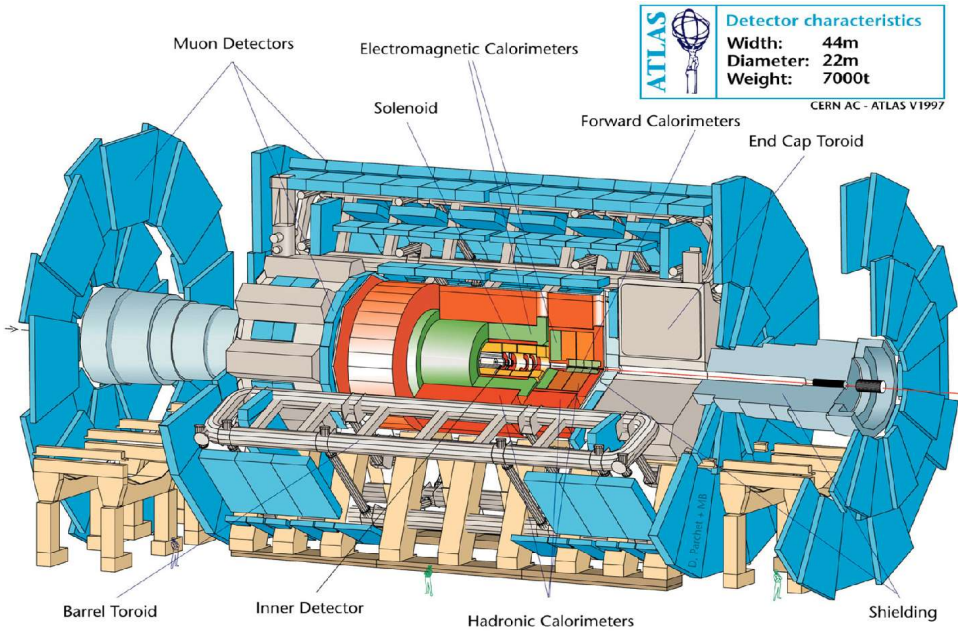


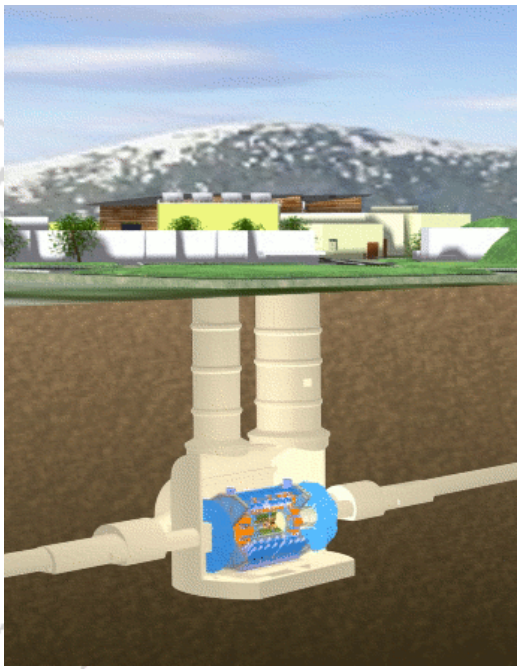
Bjørn H. Samset, University of Oslo

Outline

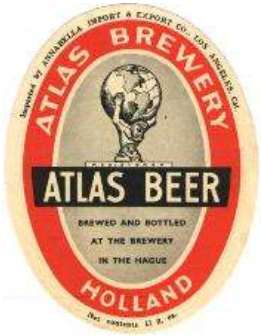
- Atlas distributed analysis on NG using Ganga
  - Atlas, Ganga and NorduGrid
  - Atlas computing model for users
  - **The new NG backend in Ganga**
  - Installing Ganga and running jobs on NG
- SUSY background study: A dijet, no-lepton channel
  - SUSY event topology for heavy gluinos
  - Collect relevant data from the grid
  - **Background study**
  - Relevance of this channel in mSUGRA
- Punchline
  - As a user, i.e. a physicist who wants to do analysis, I can access and process large datasets on the grid – even across grid boundaries. And it's easy, too.

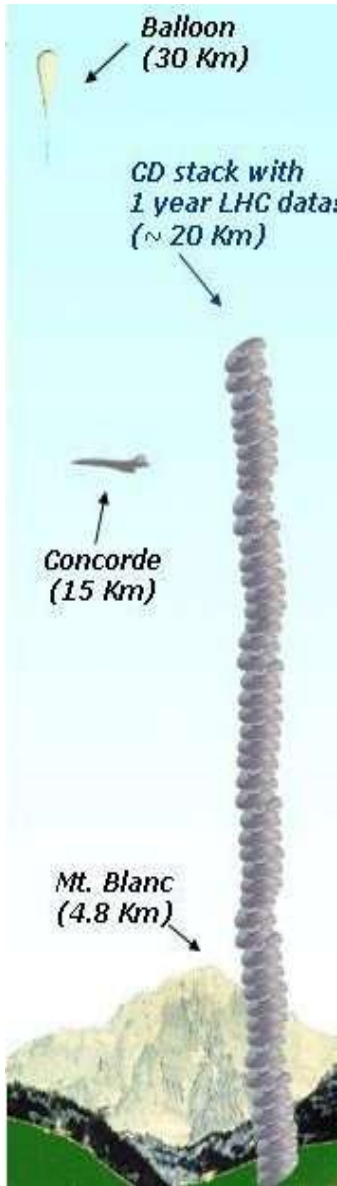
The Atlas experiment





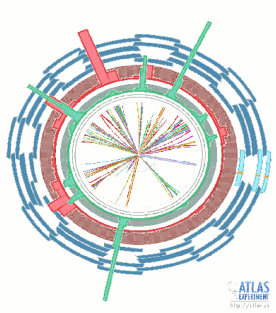
- Traditional «barrel»
  - Inner tracker (Si ++)
  - EM calorimeter
  - Hadronic calorimeter
  - Muon detectors
- Sees primarily
  - e, μ, τ
  - gammas
  - jets
- Will look for/study
  - higgs
  - SUSY
  - QCD
  - ...
- Will produce absurd amounts of data...

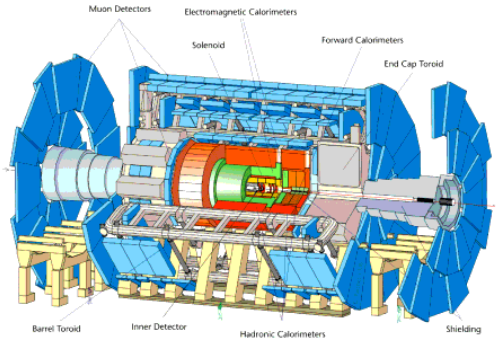




The Atlas computing model

Repeat after me: «Jobs go to data!»





Detector

RAW data

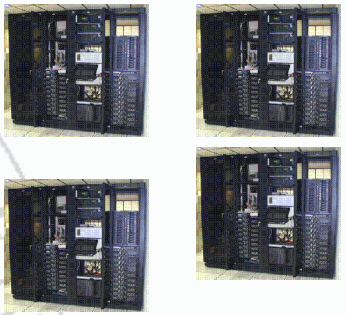
CERN Computer Centre + Tier 0

Reconstructed + RAW data

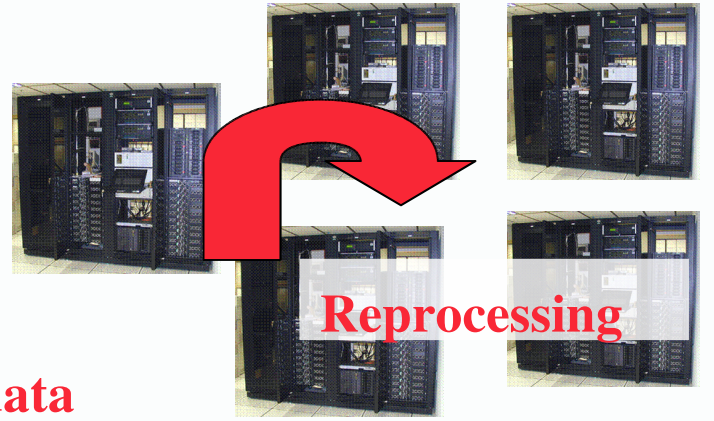
GRID

Small data products

Simulated data

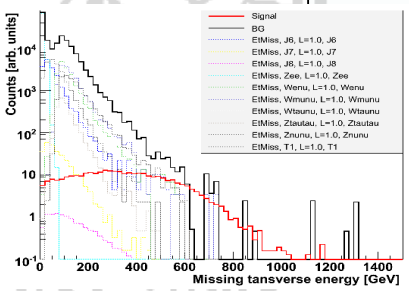


Tier 2 centres



Reprocessing

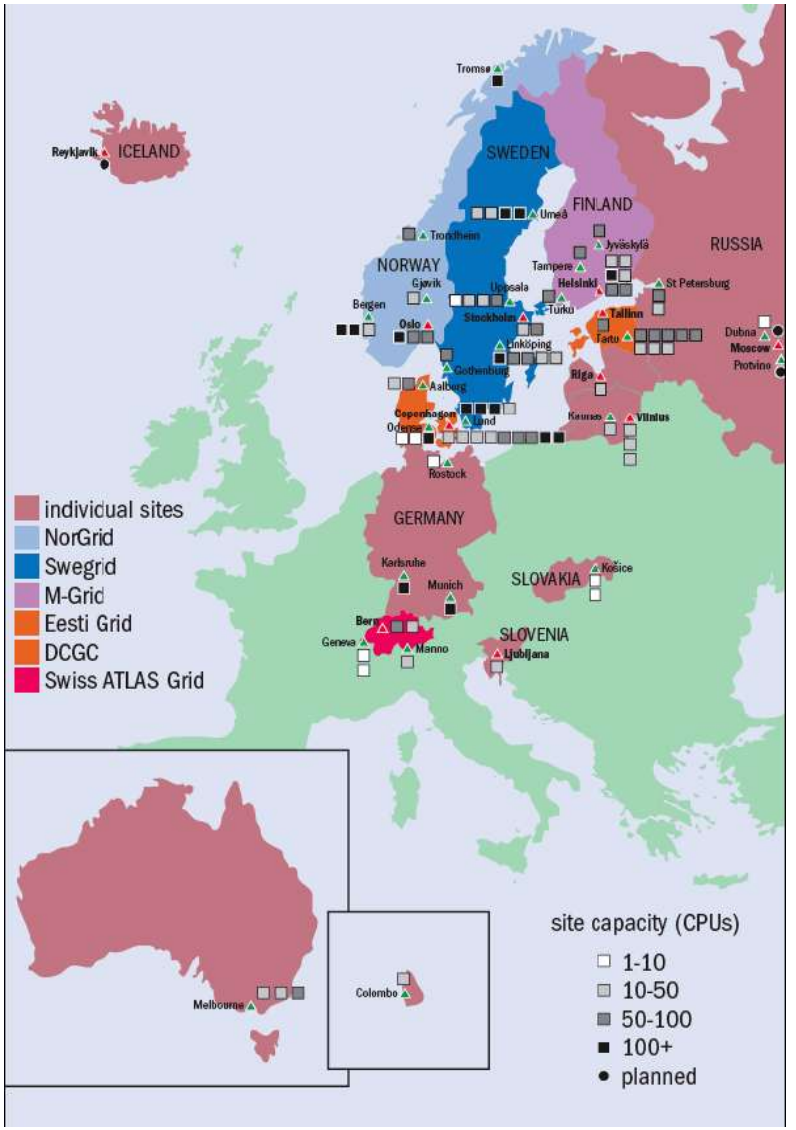
Tier 1 centres

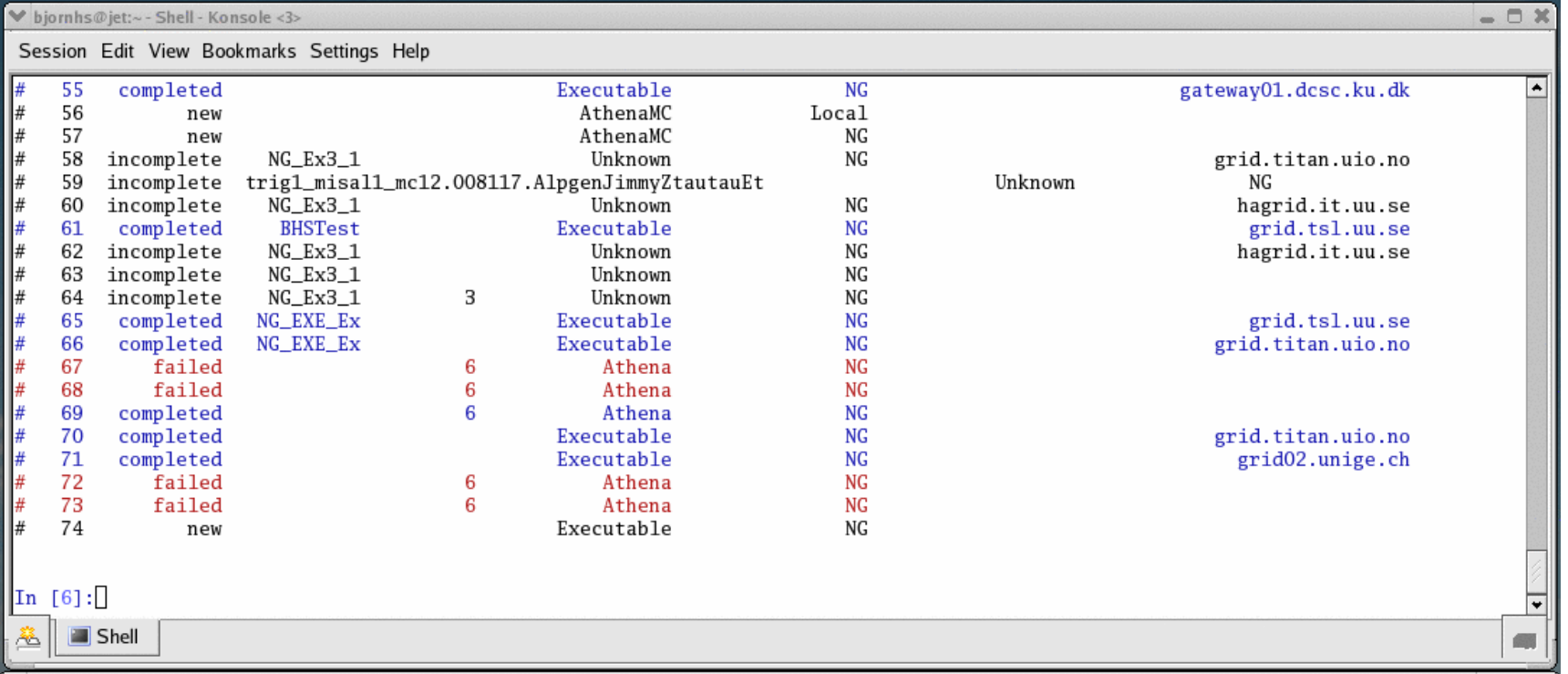


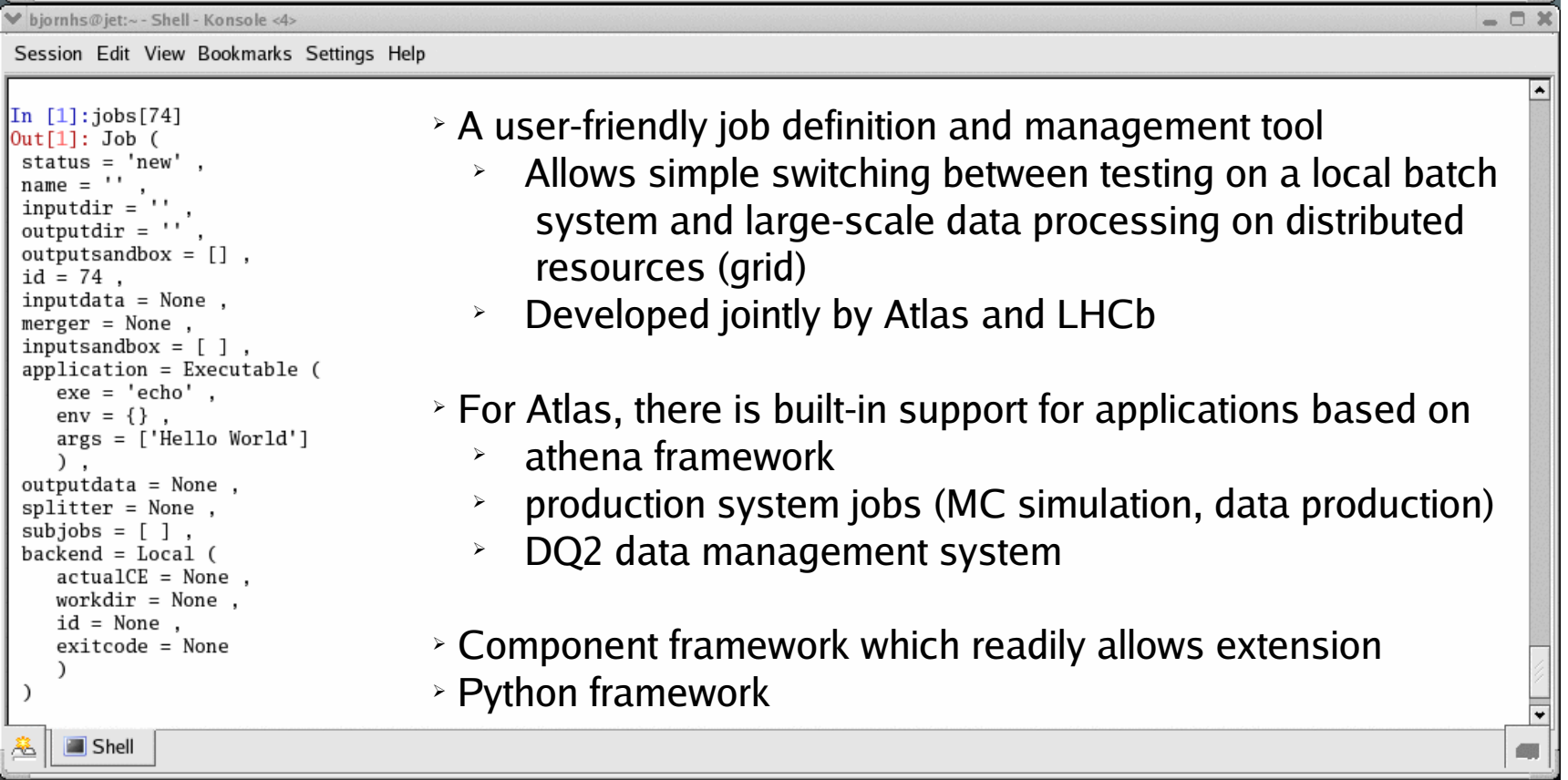
Slide borrowed from David Cameron

NorduGrid

- NorduGrid - A Grid Research and Development collaboration
- ARC - Advanced Resource Connector – the open source middleware coordinated by NorduGrid
- One of three grid flavours in Atlas
  - LCG
  - OSG
  - NG
- Data will be distributed to sites on all grids, but *each grid will only have a subset of the data*.
  - All data at all major sites in a very reduced form
  - More detailed data only at two sites
- As a user, I want my jobs to simply go to the data – preferably without me having to know where they are.
- We are moving towards that point.





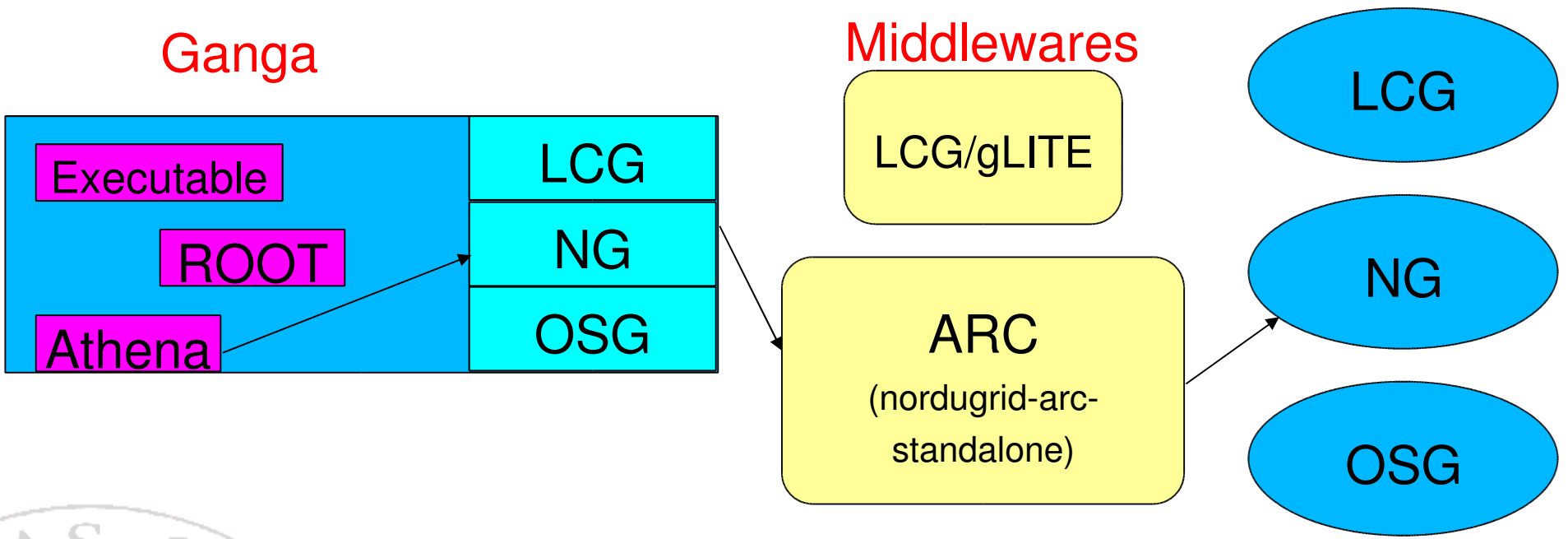


- A user-friendly job definition and management tool
  - Allows simple switching between testing on a local batch system and large-scale data processing on distributed resources (grid)
  - Developed jointly by Atlas and LHCb
- For Atlas, there is built-in support for applications based on
  - athena framework
  - production system jobs (MC simulation, data production)
  - DQ2 data management system
- Component framework which readily allows extension
- Python framework

http://www.cern.ch/ganga

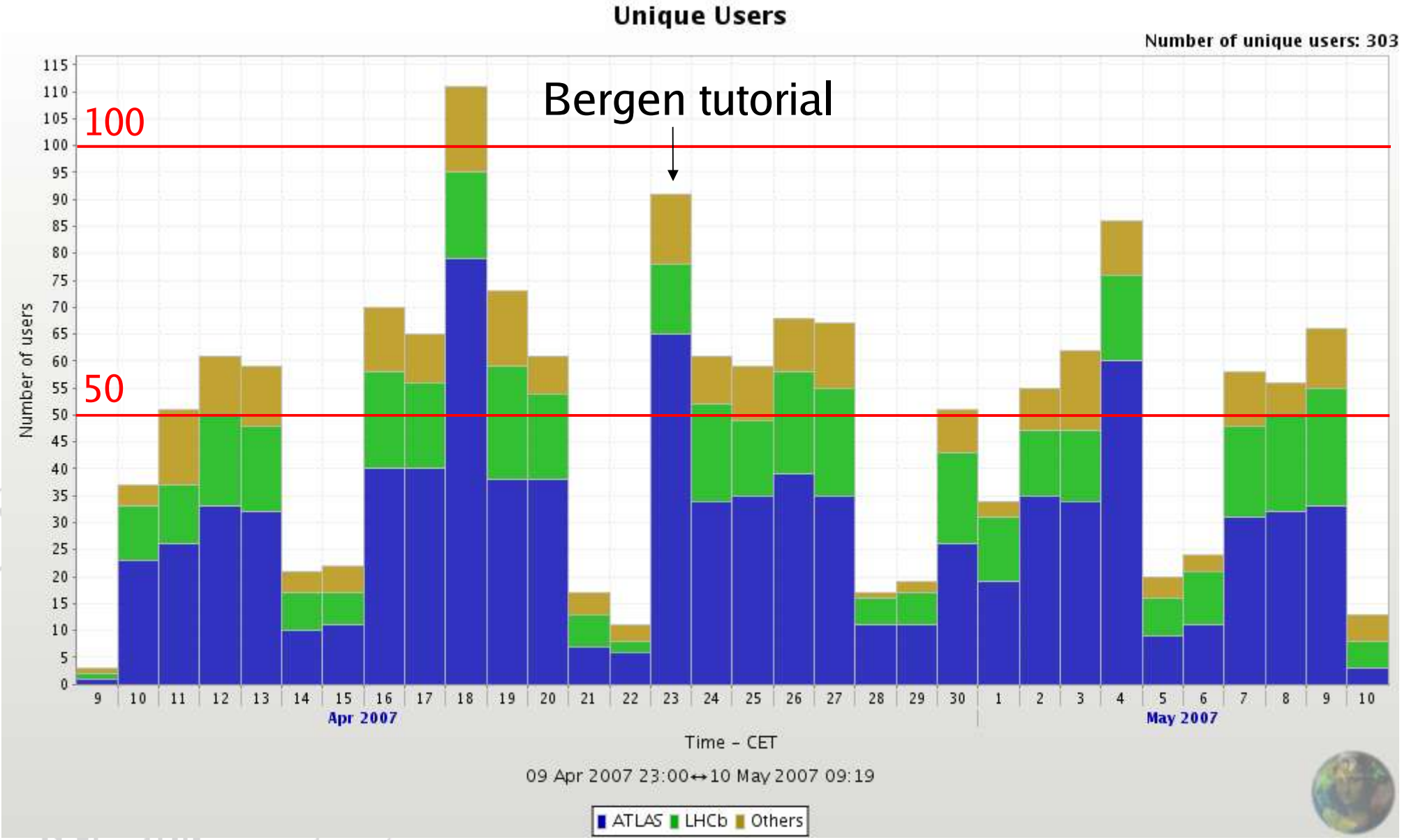
The GangaNG backend

Work done with Katarina Pajchel



- Links general purpouse Ganga to NorduGrid
  - Writes xrsl, i.e. the NG job language, for you
  - Talks to the middleware
- Ganga provides all NG middleware at install, you only need to provide your grid proxy

Ganga usage statistics



Example: Installing Ganga

If you want to install it locally:

- Get `ganga-install` from the ganga webpage:
  `wget http://cern.ch/ganga/download/ganga-install`
- Run the installer:
  `ganga-install --extern=GangaNG,GangaAtlas 4.3.1`
- Setup, after running ganga for the first time:
  - In `.gangarc`, under `[Configuration]`, specify
    `RUNTIME_PATH = GangaAtlas:GangaNG`
- That's it.
  - The middleware was downloaded in the install step
  - In the setup step, you activate the backend in Ganga

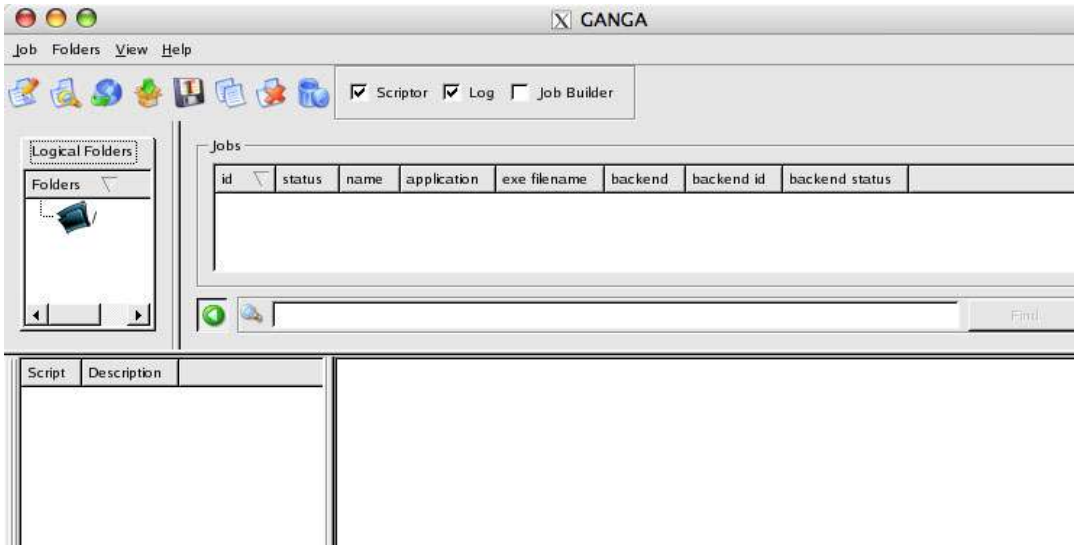
Tutorials:

- https://twiki.cern.ch/twiki/bin/view/Atlas/GangaTutorial43
- https://twiki.cern.ch/twiki/bin/view/Atlas/GangaNGTutorial430

Example: A Ganga job object

```
Out[6]: Job (
status = 'new' ,
name = '' ,
inputdir = '/home/scratch/bjornhs/gangadir/workspace//Local/74/input/' ,
outputdir = '/home/scratch/bjornhs/gangadir/workspace//Local/74/output/' ,
outputsandbox = [] ,
id = 74 ,
inputdata = None ,
merger = None ,
inputsandbox = [ ] ,
application = Executable (
   exe = 'echo' ,
   env = {} ,
   args = ['Hello World']
   ) ,
outputdata = None ,
splitter = None ,
subjobs = [ ] ,
backend = NG (
   status = '' ,
   check_availability = False ,
   actualCE = '' ,
   submit_options = [] ,
   middleware = 'ARC' ,
   CE = '' ,
   id = '' ,
   queue = '' ,
   reason = '' ,
   cputime = '' ,
   RejectCE = '' ,
   RLS = 'rls://atlasrls.nordugrid.org:39281' ,
   requirements = NGRequirements (
      runtimeenvironment = [] ,
      cputime = '30' ,
      other = [] ,
      memory = 500 ,
      disk = 500 ,
      walltime = '30'
      )
   )
)
```

- Python object
- Holds all information on
  - what the job is to do
  - where we want it to go
  - input/output
- status - running, completed, failed...
- Can be re-used, modified etc.
- Can also be edited via a GUI



Let's use it: Supersymmetry

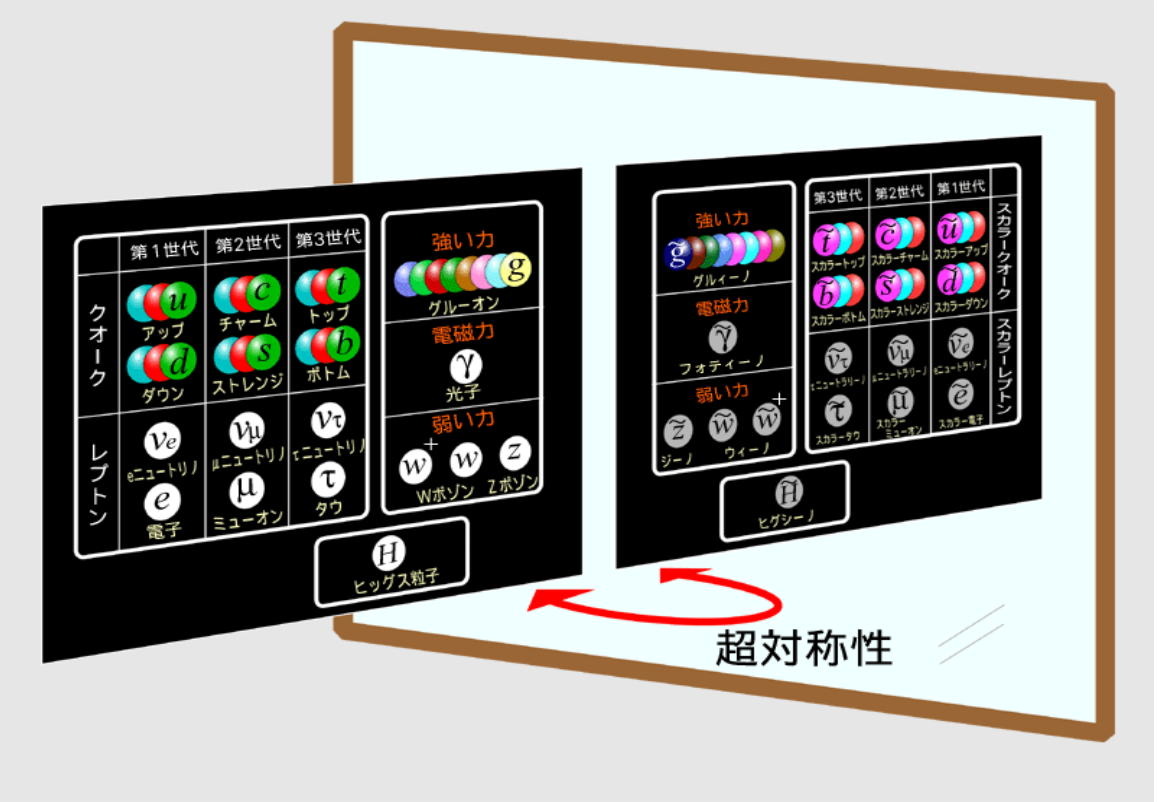


Image: KEK

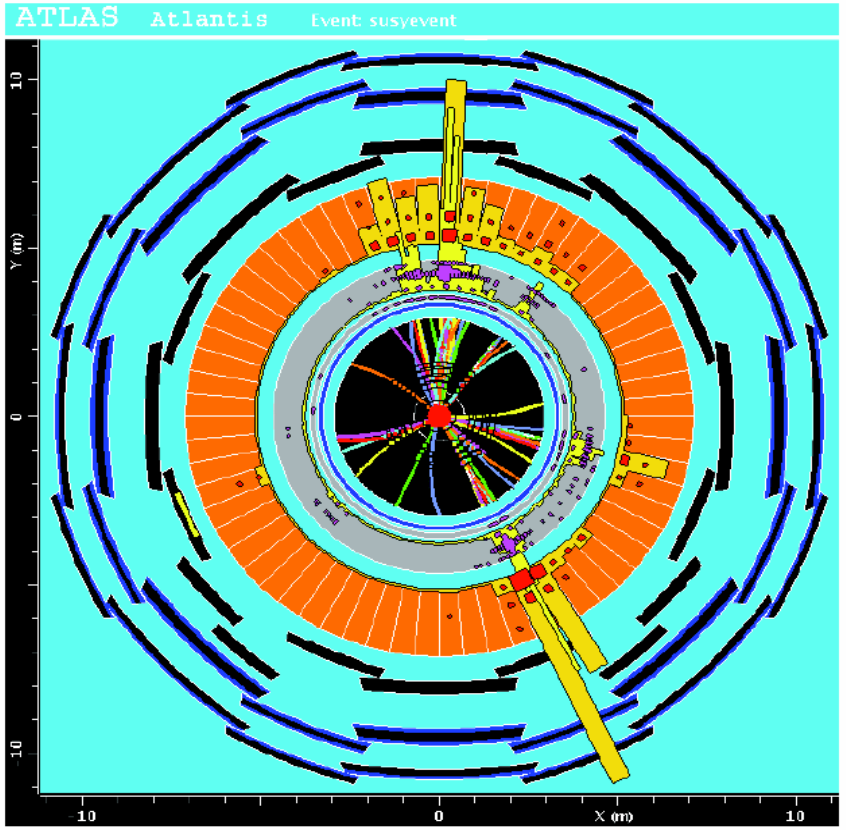
- A new «sparticle» for every particle (at least)
- Bosons->fermions and vice versa
- Adresses hierarchy problem
- Unifies force couplings at high energy scale
- Provides dark matter candidate
- etc...

...but is it real?

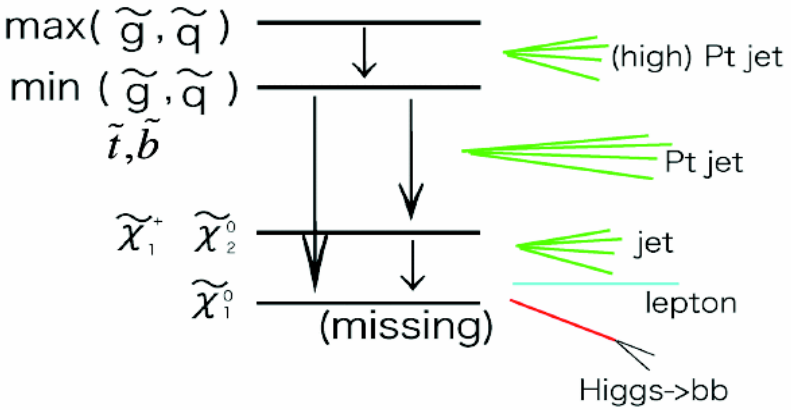
To find out we must first know what the SM predicts – i.e. study our expected backgrounds.

«Typical» SUSY event

"Typical" Events topology of SUSY signal is like this



Gluino/squark are produced first, then cascade decay is followed.



event topologies of SUSY

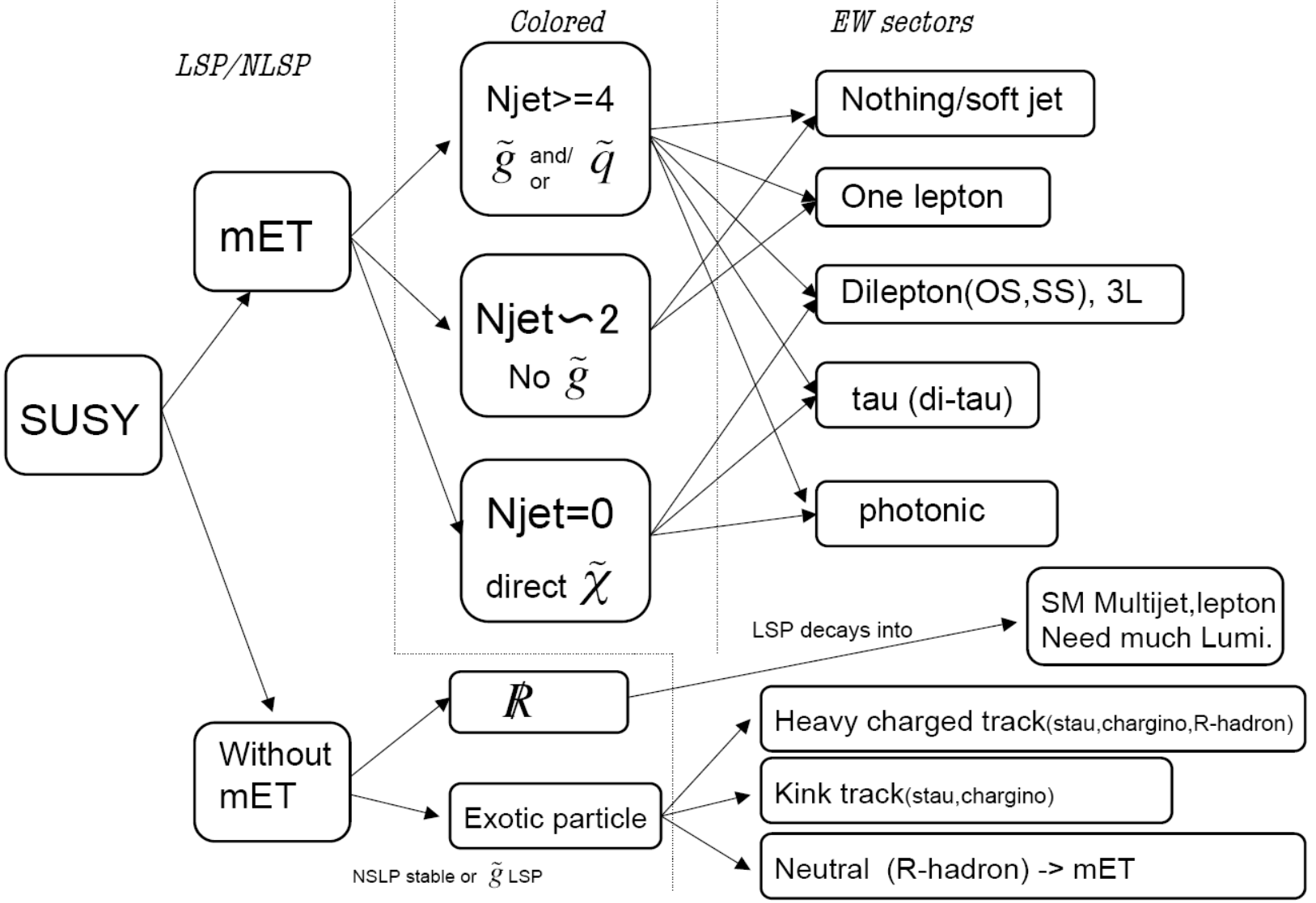
$$\not{E}_T + \text{High } P_T \text{ jets} + \begin{matrix}\text{multi} \\ \text{b-jets} \\ \tau\text{-jets}\end{matrix} \quad \text{leptons}$$

Especially
no or one lepton mode is promising for Discovery

Slide borrowed from Shoji Asai

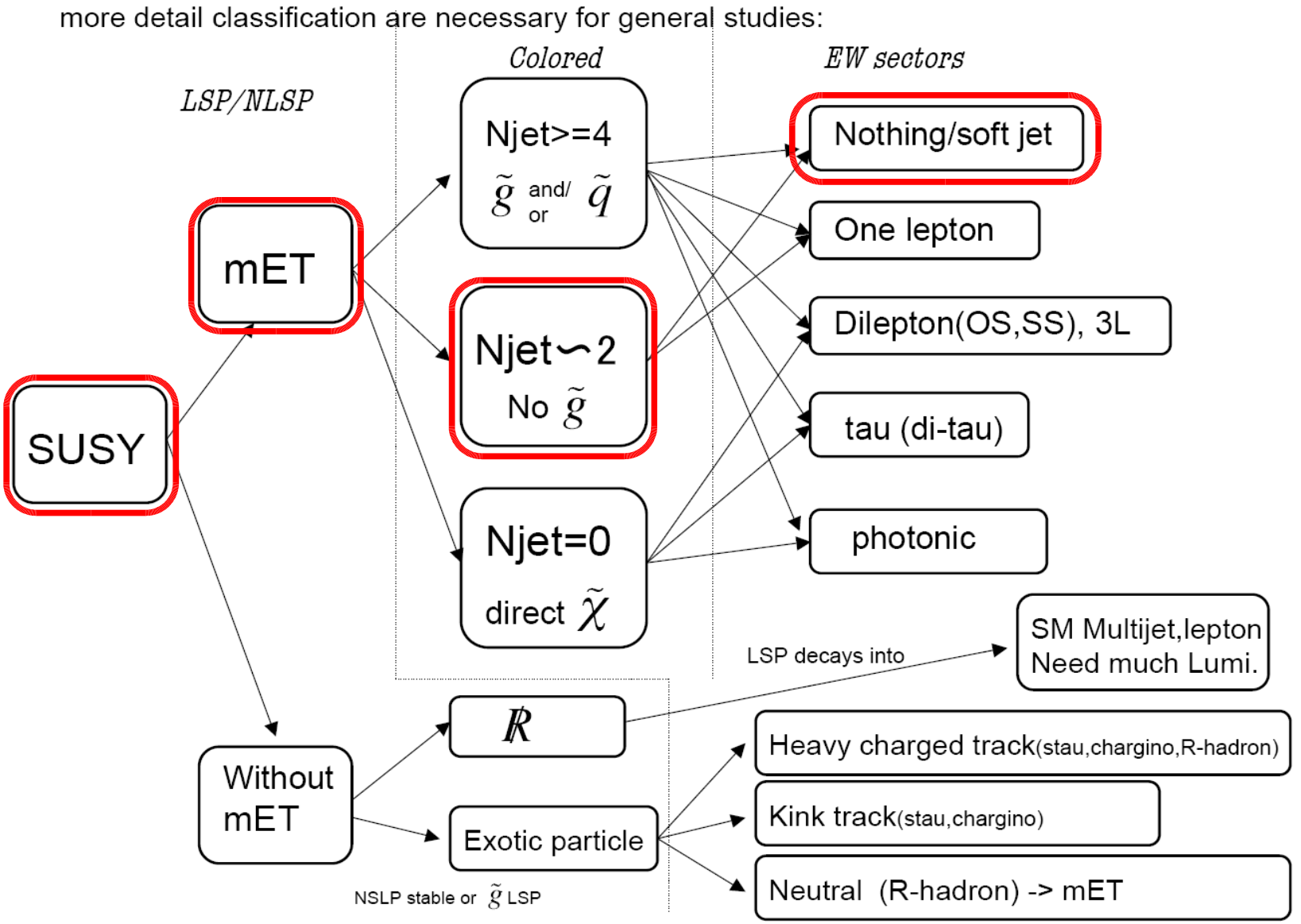
SUSY event topology

more detail classification are necessary for general studies:



Slide borrowed from Shoji Asai

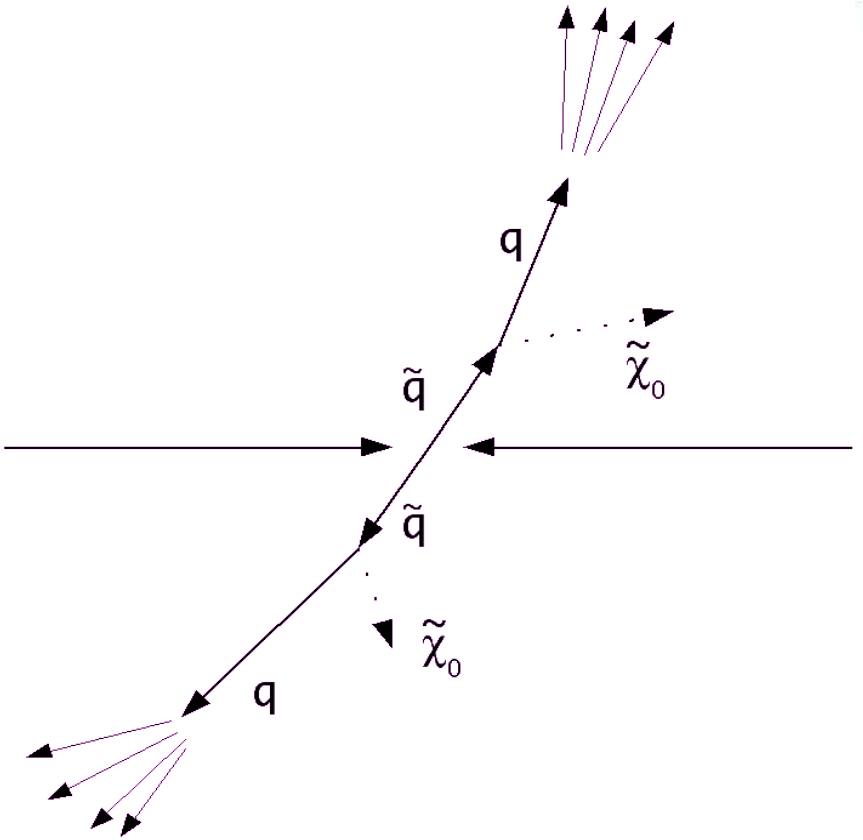
SUSY event topology



Slide borrowed from Shoji Asai

Two-jet, 0-lepton channel

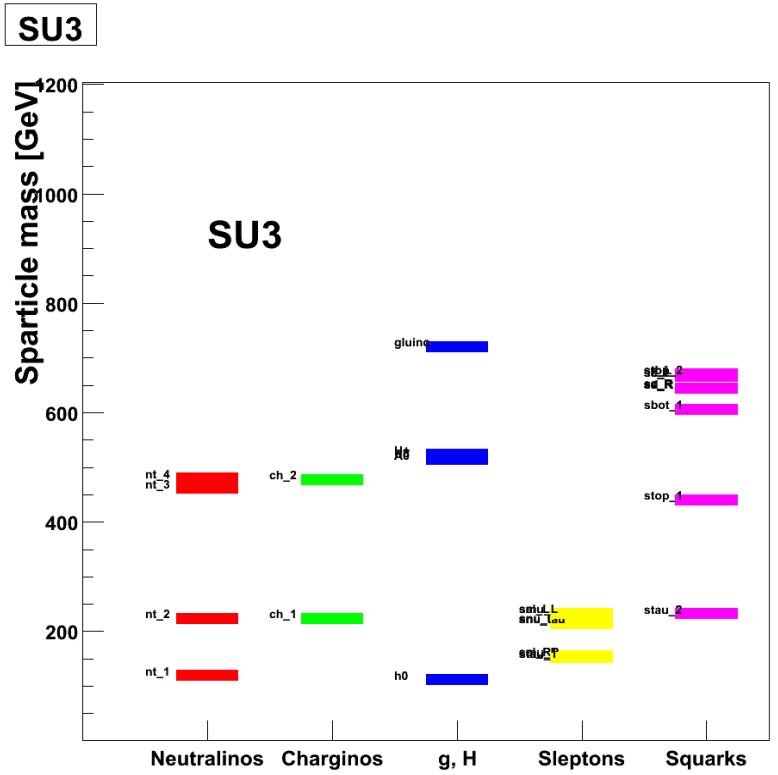
- Approximate topology:
- Most relevant if gluino is heavy, i.e. more likely to produce squark pair.
- Present but not dominatng in mSUGRA
  - $m_{squark} > m_{gluino}$ in most cases
- Still a relevant event topology



- Signals in this analysis
  - One realisation of mSUGRA (next slide)
- Backgrounds
  - QCD dijets
  - W/Z production with various decays
  - t-tbar events
  - ...

A «benchmark» mSUGRA point

- $m_0 = 100$
- $m_{1/2} = 300$
- $A_0 = -300$
- $\tan(\beta) = 6$
- $\text{sign}(\mu) = +$



$\sigma_{tot} = 19.3\text{pb}$

Production	$\sigma(fb)$	Production	$\sigma(fb)$
$f\bar{f} \to \tilde{l}_{L,R}\tilde{l}_{L,R}$	69.39	$qg \to \tilde{q}_{L,R}\tilde{\chi}^0$	345.60
$f\bar{f} \to \tilde{\tau}_{1,2}\tilde{\tau}_{1,2}$	35.30	$qg \to \tilde{q}_L\tilde{\chi}^{\pm}$	419.02
$q\bar{q} \to \tilde{l}_L\tilde{\nu}_l$	96.59	$qg \to \tilde{q}_{L,R}\tilde{g}$	8655.00
$q\bar{q} \to \tilde{\tau}_{1,2}\tilde{\nu}_\tau$	45.10	$f\bar{f} \to \tilde{t}_{1,2}\tilde{t}_{1,2}$	99.28
$f\bar{f} \to \tilde{\nu}_l\tilde{\nu}_l$	28.79	$gg \to \tilde{t}_{1,2}\tilde{t}_{1,2}$	391.20
$f\bar{f} \to \tilde{\nu}_\tau\tilde{\nu}_\tau$	13.93	$q\bar{q} \to \tilde{q}_{L,R}\tilde{q}_{L,R}$	4980.00
$f\bar{f} \to \tilde{\chi}^0\tilde{\chi}^0$	30.61	$f\bar{f} \to \tilde{q}_{L,R}\tilde{q}_{L,R}$	149.12
$f\bar{f} \to \tilde{\chi}^{\pm}\tilde{\chi}^{\pm}$	296.51	$gg \to \tilde{q}_{L,R}\tilde{q}_{L,R}$	851.45
$q\bar{q} \to \tilde{\chi}^0\tilde{\chi}^{\pm}$	549.70	$b\bar{q} \to \tilde{b}_{1,2}\tilde{q}_{L,R}$	165.86
$q\bar{q} \to \tilde{\chi}^0\tilde{g}$	115.30	$f\bar{f} \to \tilde{b}_{1,2}\tilde{b}_{1,2}$	60.40
$q\bar{q} \to \tilde{\chi}^{\pm}\tilde{g}$	143.18	$b\bar{b} \to \tilde{b}_{1,2}\tilde{b}_{1,2}$	2.42
$q\bar{q} \to \tilde{g}\tilde{g}$	98.91	$bg \to \tilde{b}_{1,2}\tilde{g}$	151.06
$gg \to \tilde{g}\tilde{g}$	1385	$f\bar{f} \to H/AH$	2.43

Table by Christina Potter, using ISAJET/Pythia

- All sparticles accessible at LHC
- No «special» degeneracies

Submit jobs to the grid...

...for each dataset and background

```
Out[4]: Statistics: 75  jobs

--------------

#   id      status        name   subjobs      application          backend    backend.actualCE

#   69   completed                    6           Athena               NG

In [5]:j = jobs[69].copy()

In [6]:j.backend=LCG()

In [7]:j.inputdata.dataset='some_dataset_only_available_on_LGC'

In [8]:j.submit()

Submitting job 75 to LCG backend.
```


- Easy to submit a new dataset if you want to look at a new signal or BG
- Just copy the old job object, select the new set name, submit
- New set not on your grid? Just switch the backend to another grid and resubmit.
- Jobs go to data.

Getting data from other users

- Other users also look at the same datasets for their analyses
- Common tools, in this case «SUSYView», are used as far as possible in the analysis chain
- Ntuples registered on the grid in the same way as datasets
- I can access these as easily as my own.
- Jobs go to data - until the point where the data files are small enough to fit on

Thanks to several collaborators:

Signal Samples

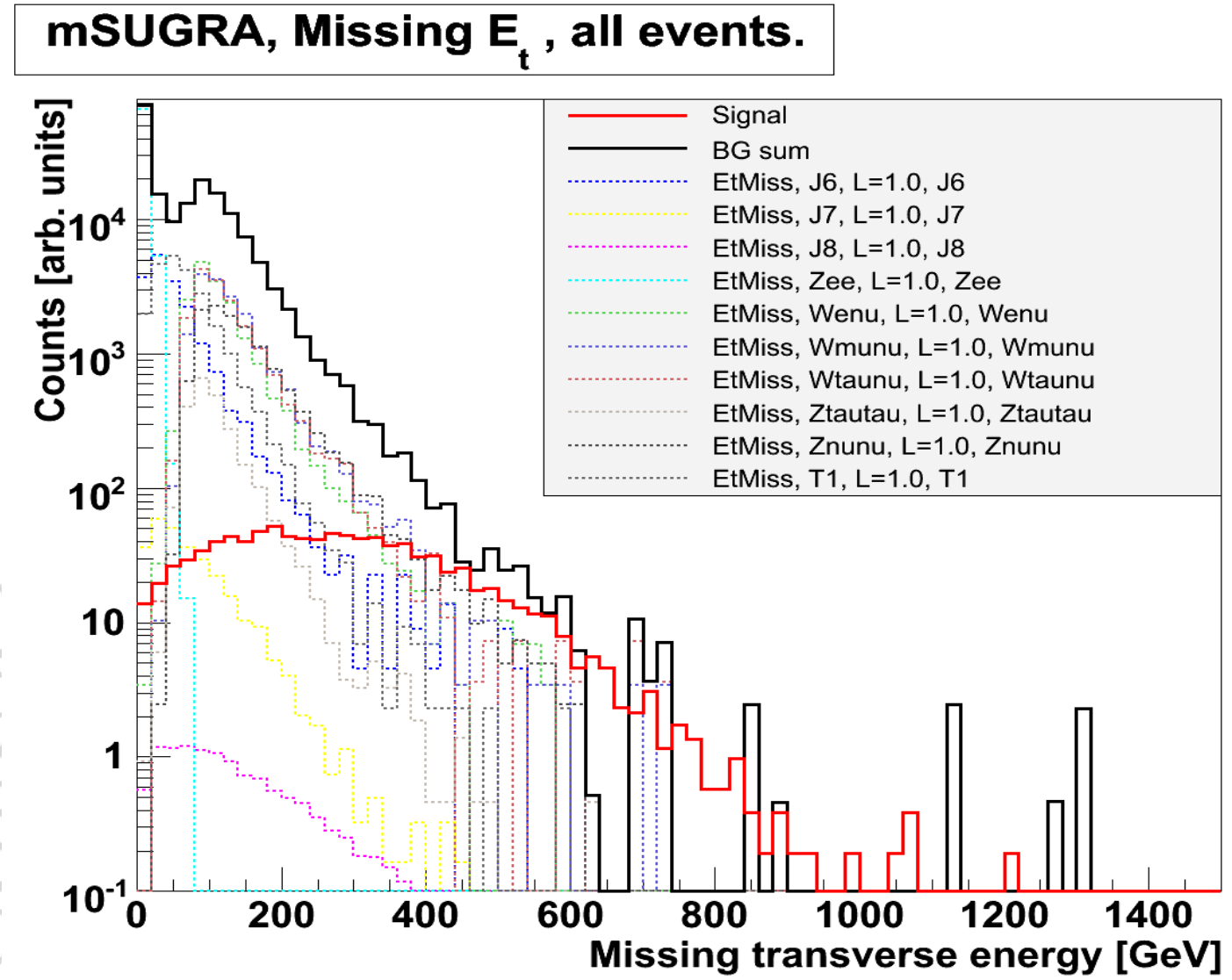
ntuples available as user datasets
user.MarijaMilosavljevic.SU1_jimmy_susy.HPTV.v12000601_tid006491
user.MarijaMilosavljevic.SU2_jimmy_susy.HPTV.v12000601_tid005862
user.JelenaKrstic.trig1_misal1_csc11.005403.SU3_jimmy_susy.recon.HPTV.v12000
user.MarijaMilosavljevic.006400.SU4_jimmy_susy.HPTV.v12000601
user.MarijaMilosavljevic.SU6_jimmy_susy.HPTV.v12000601_tid006098
user.MarijaMilosavljevic.SU8_jimmy_susy1.HPTV.v12000601_tid006099

Background Samples

ntuples available as user datasets
user.MarijaMilosavljevic.005200.T1_McAtNlo_Jimmy.HPTV.v12000601_tid005997
user.MarijaMilosavljevic.pythia_Wenu_qg_ckin80_Nj2.HPTV.v12000601_tid006105
user.MarijaMilosavljevic.pythia_Wmunu_qg_ckin80_Nj2.HPTV.v12000601_tid00610
user.MarijaMilosavljevic.pythia_Wtaunu_qg_ckin80_Nj2.HPTV.v12000601_tid00610
user.MarijaMilosavljevic.pythia_Znunu_qg_ckin80_Nj2.HPTV.v12000601_tid006108
user.MarijaMilosavljevic.pythia_Ztautau_qg_ckin80_Nj2.HPTV.v12000601_tid00670
user.MarijaMilosavljevic.pythia_Zee_qg_ckin80_Nj2.HPTV.v12000601_tid006109
user.MarijaMilosavljevic.AlpgenJimmyWenuNp2.HPTV.v12000601_tid006524
user.MarijaMilosavljevic.AlpgenJimmyWenuNp3.HPTV.v12000601_tid006525
user.MarijaMilosavljevic.AlpgenJimmyWenuNp4.HPTV.v12000601_tid006526

BG study 1: All events

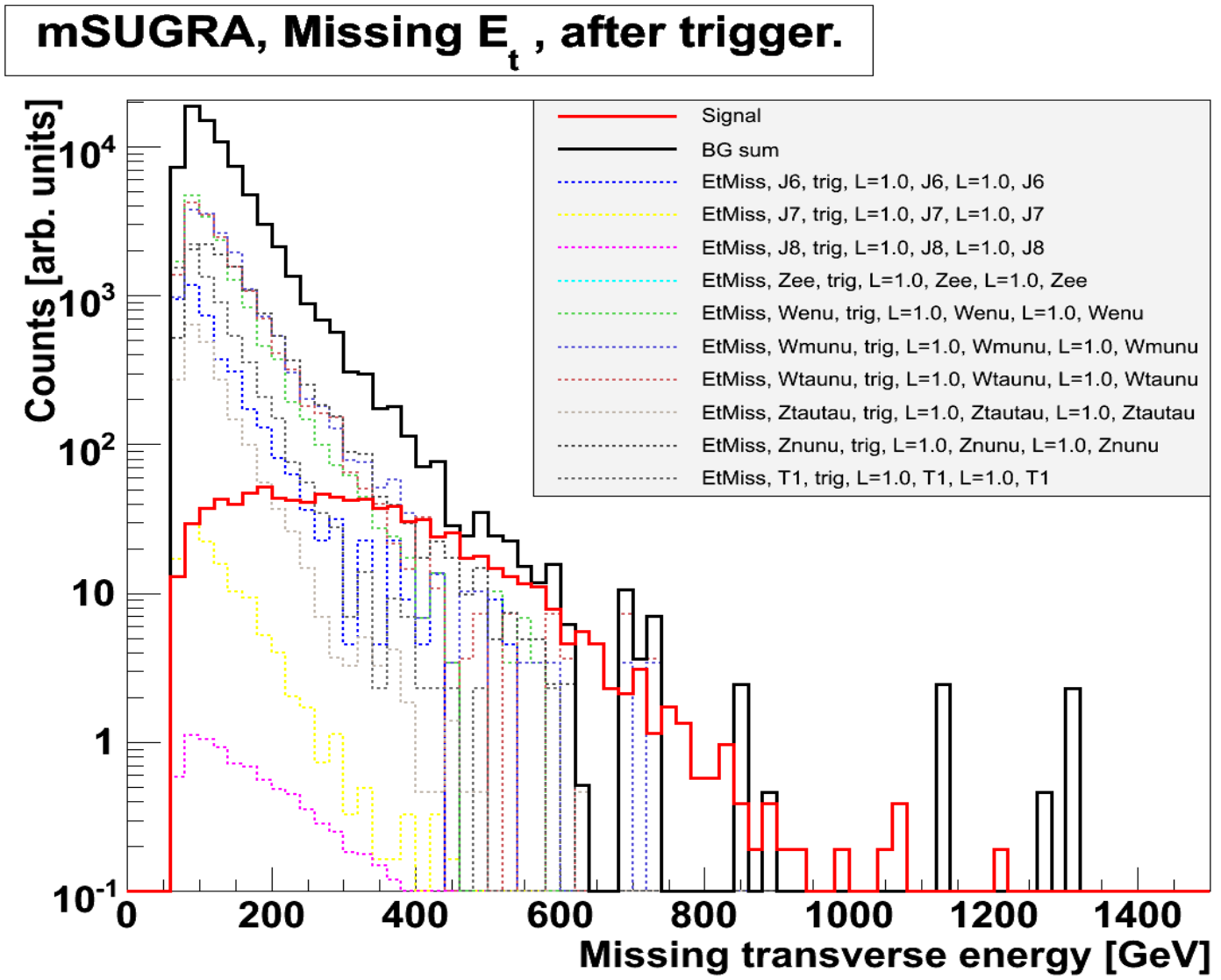
- Plot the missing transverse energy for all events.
- Sum up all background contributions.
- Overlay with SUSY signal.



BG study 2: Trigger

- Not all events reach the data stream.
- Require that the (simulated) trigger actually fired.

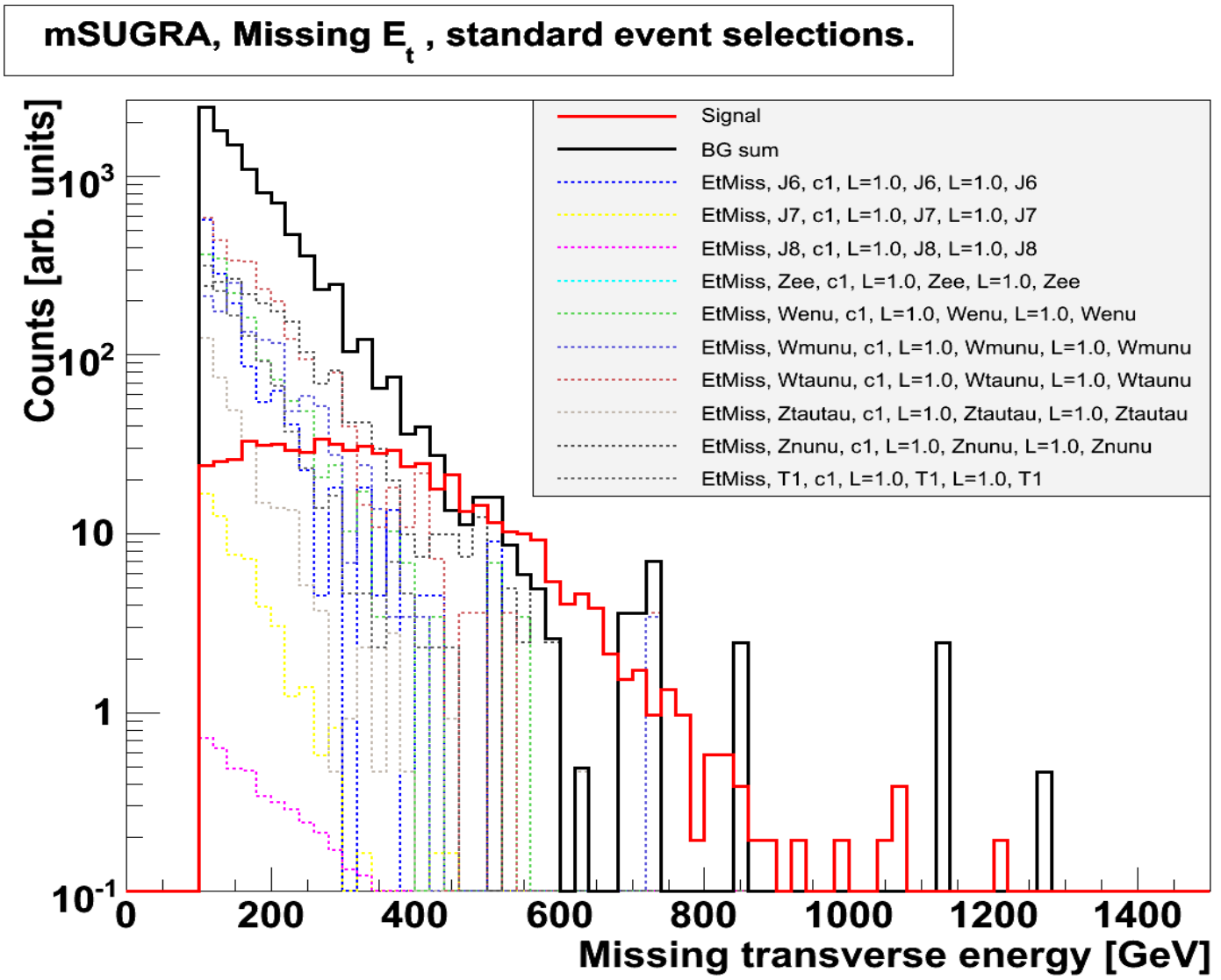
`EtMiss>70 && JetPt>70`



BG study 3: Event selection

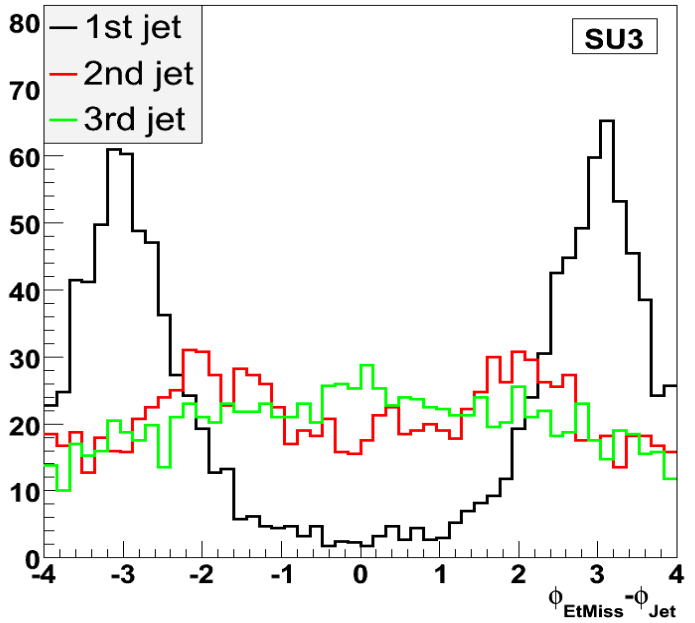
- Select events that are consistent with our chosen SUSY topology.
- Missing Et, number of jets, number of leptons

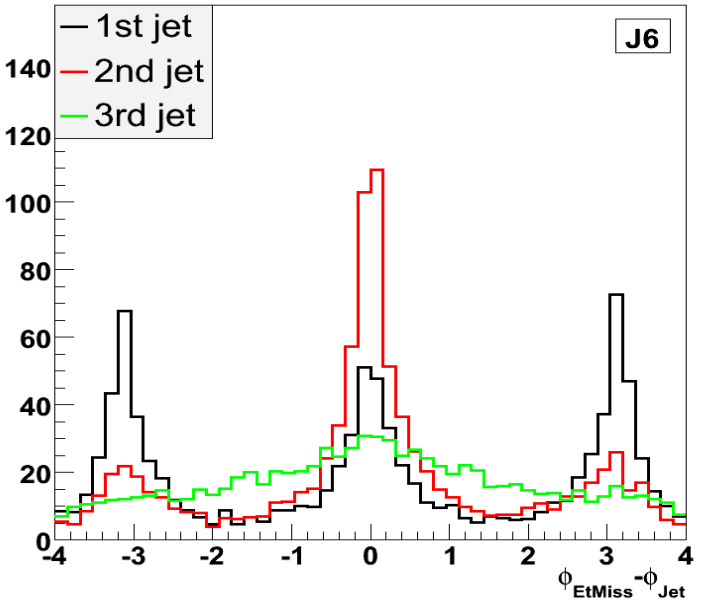
EtMiss>100GeV, pt_{Jet1}>100GeV/c, pt_{Jet2}>80GeV/c, pt_{Lep1}<20GeV/c



BG study 4: Jet/EtMiss corr.

- Make additional cuts that reduce BG but not signal.
- E.g. correlate the phi direction of the total missing Et vector with the jet directions.

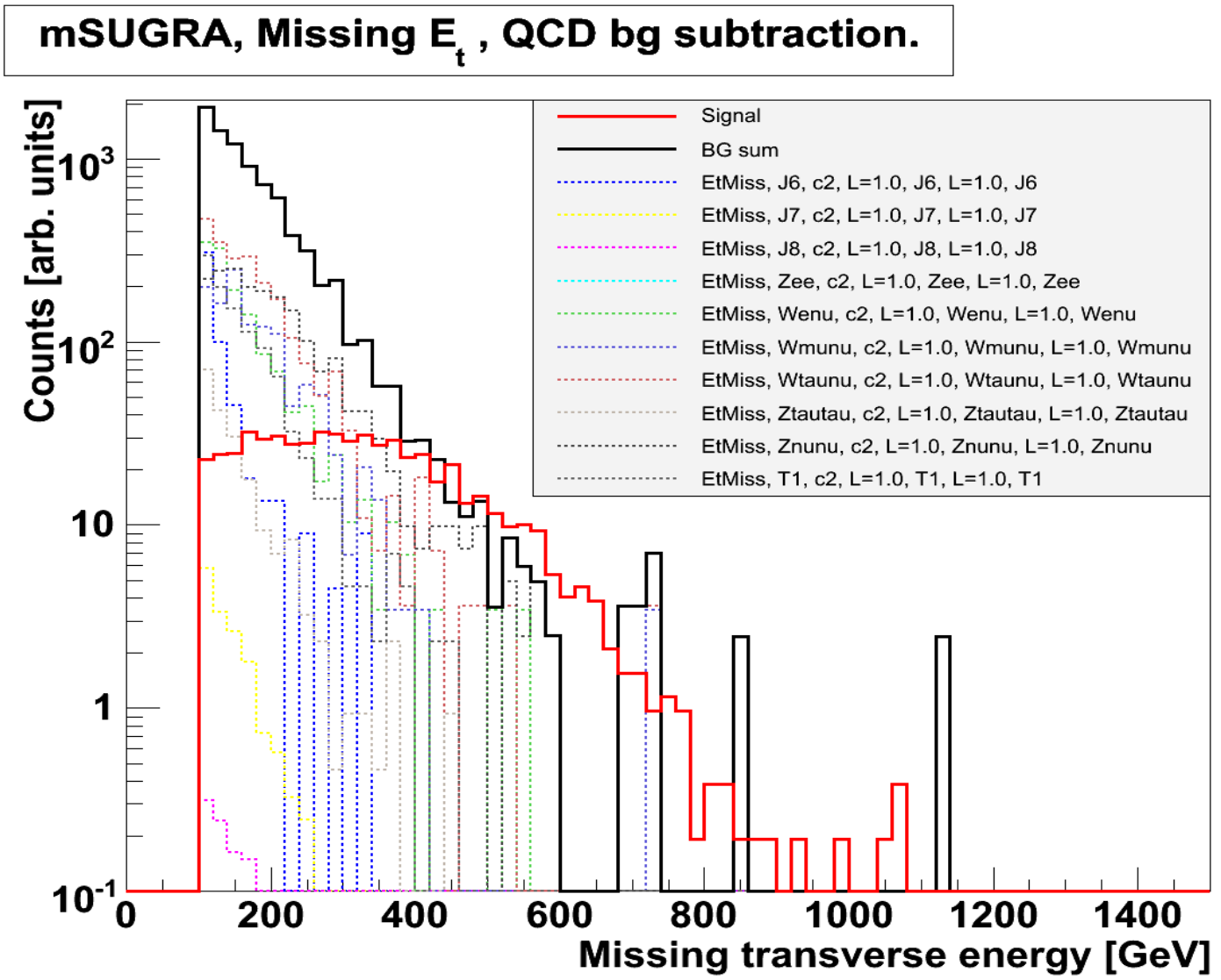




BG study 4: Jet/EtMiss corr.

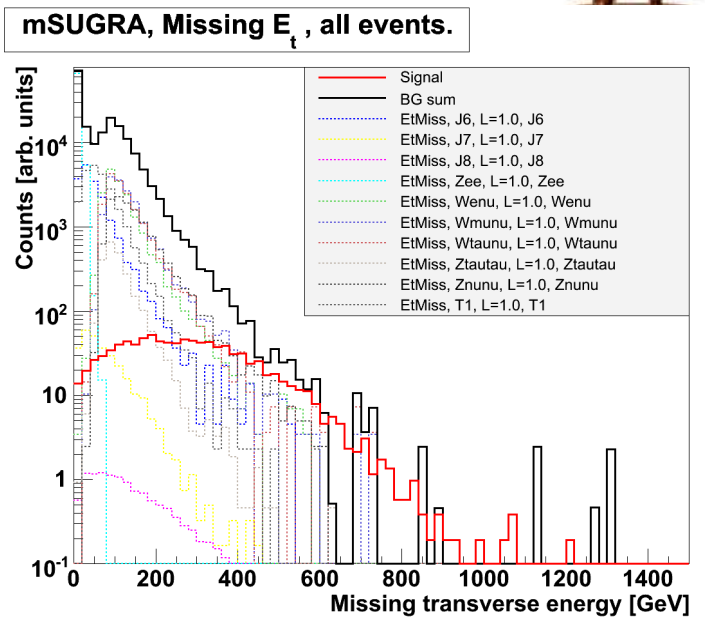
- Make additional cuts that reduce BG but not signal.
- E.g. correlate the phi direction of the total missing Et vector with the jet directions.

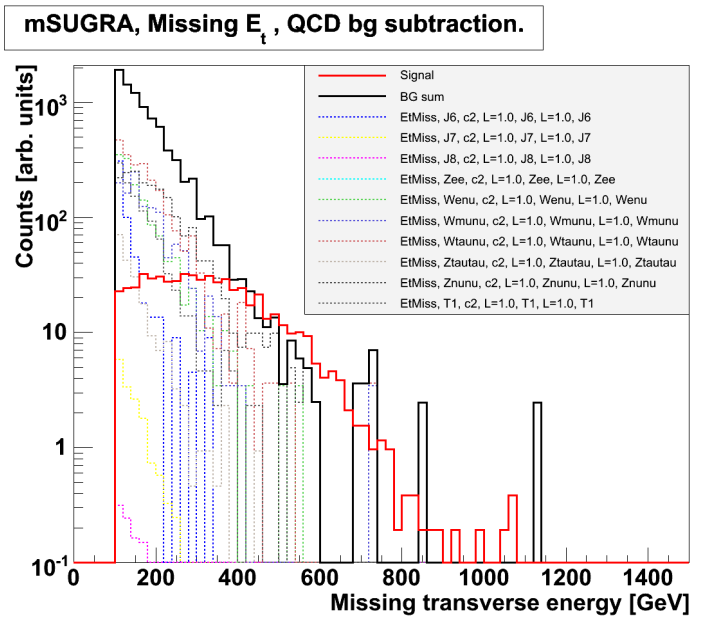
$\texttt{EtMissPhi-Phi}_{\texttt{Jet2}}\texttt{<0.2}$



Physics conclusions

- A dijet, lepton-free SUSY signal is present, even in mSUGRA where gluinos are heavier than squarks.
- However, not dominant above backgrounds, even after event selection and a first round of kinematic cuts.
- If we see an excess of missing transverse energy in this channel, should take note and look outside (this particular realisation of) mSUGRA for explanations.
- (More to be done here, of course...)





Grid conclusions

- NorduGrid is (very nearly) ready for full-scale distributed analysis use by any Atlas user
- Anyone in the Atlas Virtual Organisation can use NorduGrid resources and access data on NG storage elements, through Ganga
- Interplay with other grids is improving
- Punchline
  - As a user, i.e. a physicist who wants to do analysis, I can access and process large datasets on the grid – even across grid boundaries. And it's easy, too.