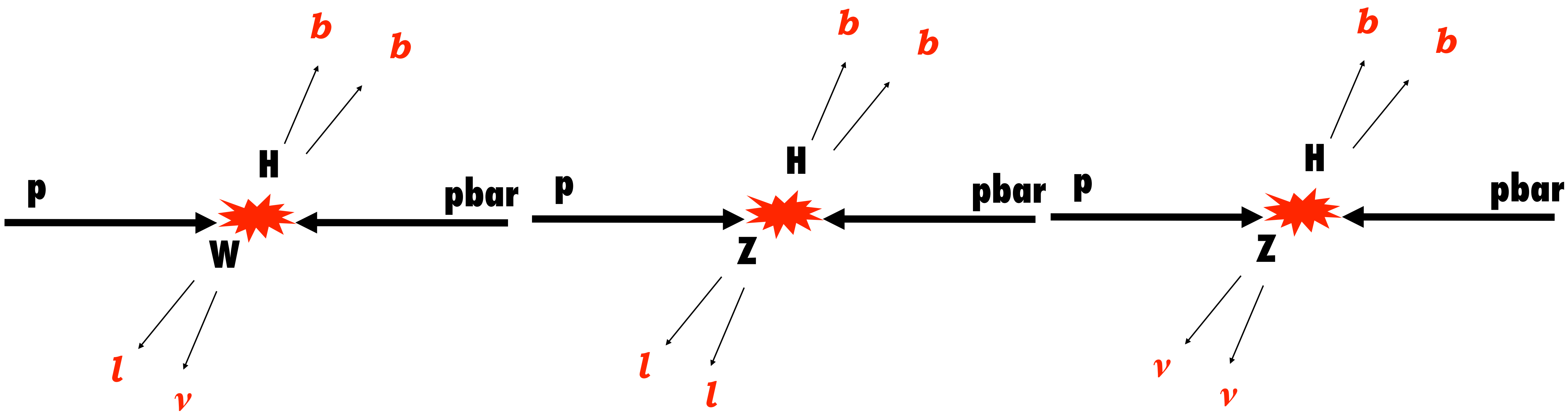
Methods and algorithms for diboson production in hadronic final states

Vadim Rusu

Outline

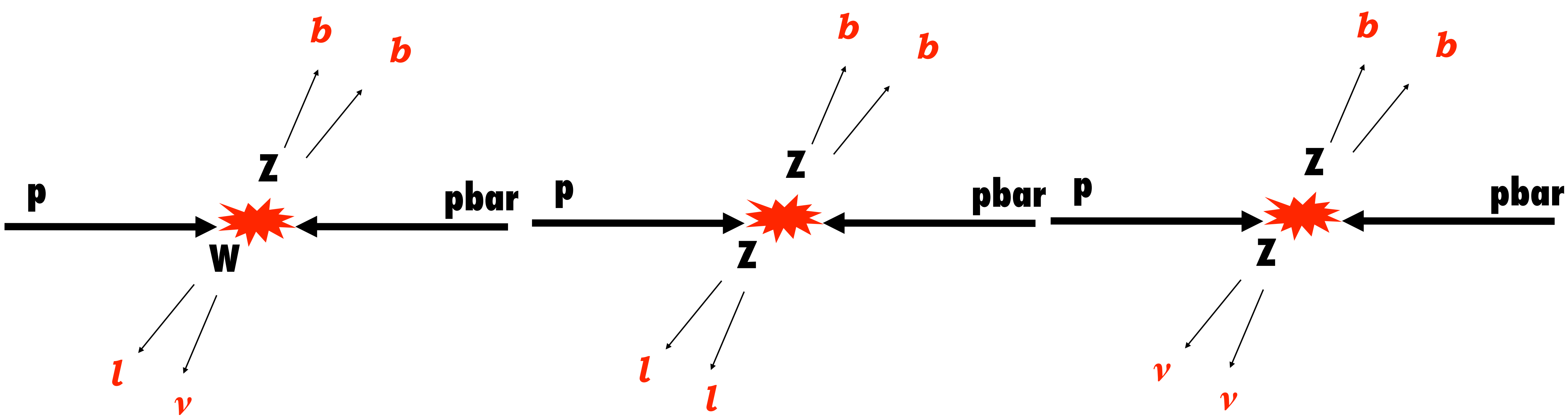
- High P_T physics with jets
- Issues
- Quark-gluon separation
- b tagging
- Conclusions

Associated Higgs



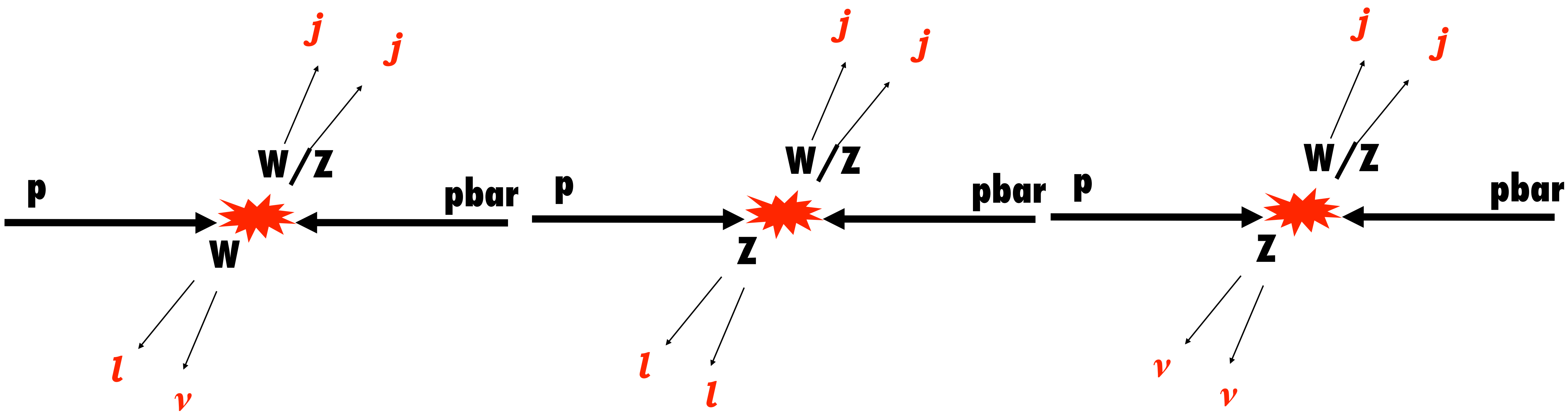
- The most promising channel for low Higgs
  - at least at the Tevatron

Dibosons



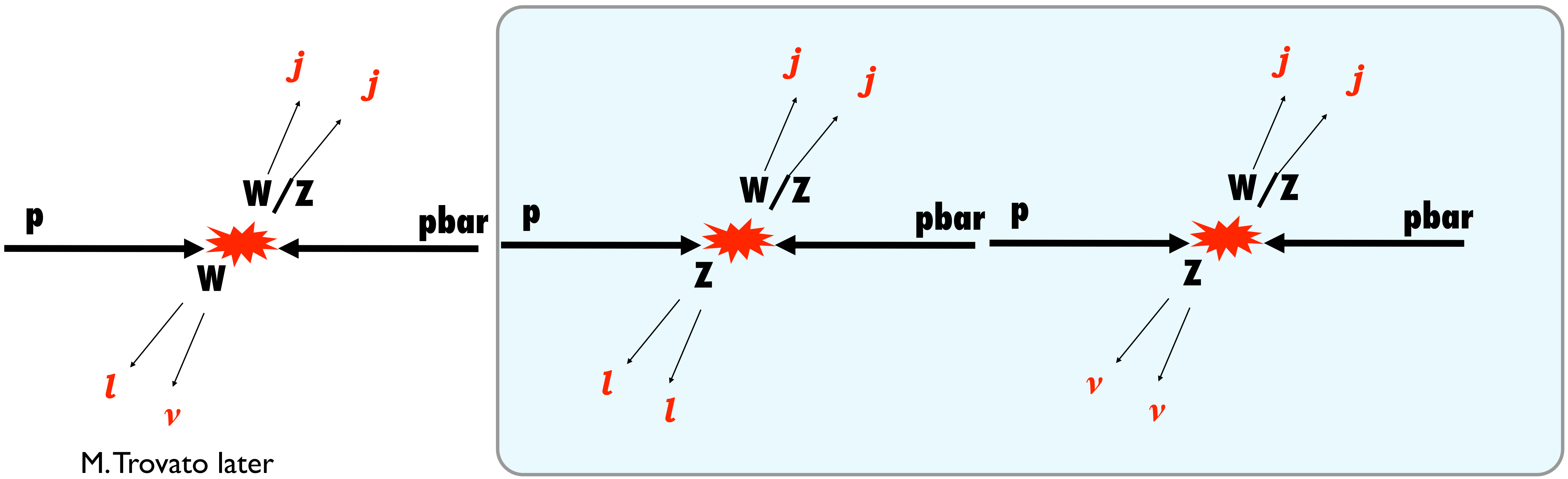
- As close as possible to the Higgs
  - calibration of tools
- The only chance to observe Z→bb

Dibosons



- One of the few (if not the only) standard candle in jet spectroscopy
- W-Z indistinguishable because of energy resolution
- Anomalous Triple Gauge Coupling study ground

Dibosons

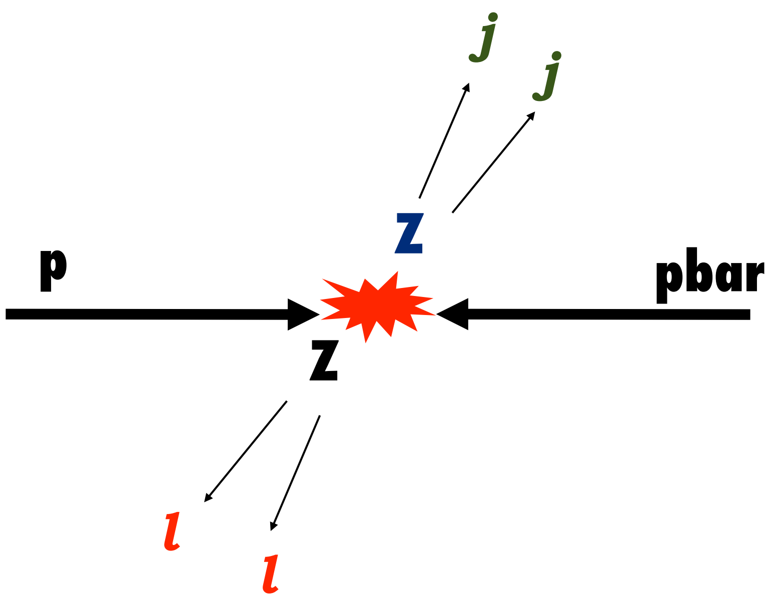


M.Trovato later

- One of the few (if not the only) standard candle in jet spectroscopy
- W-Z indistinguishable because of energy resolution
- Anomalous Triple Gauge Coupling study ground

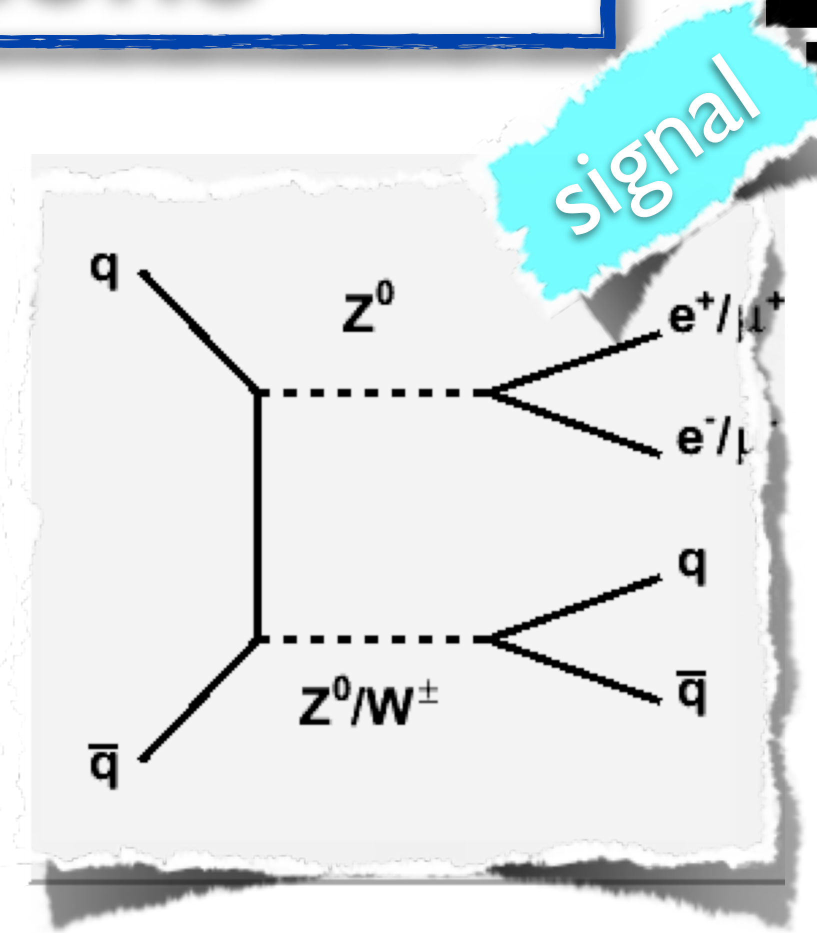
W(Z)+Z→lljj in a slide

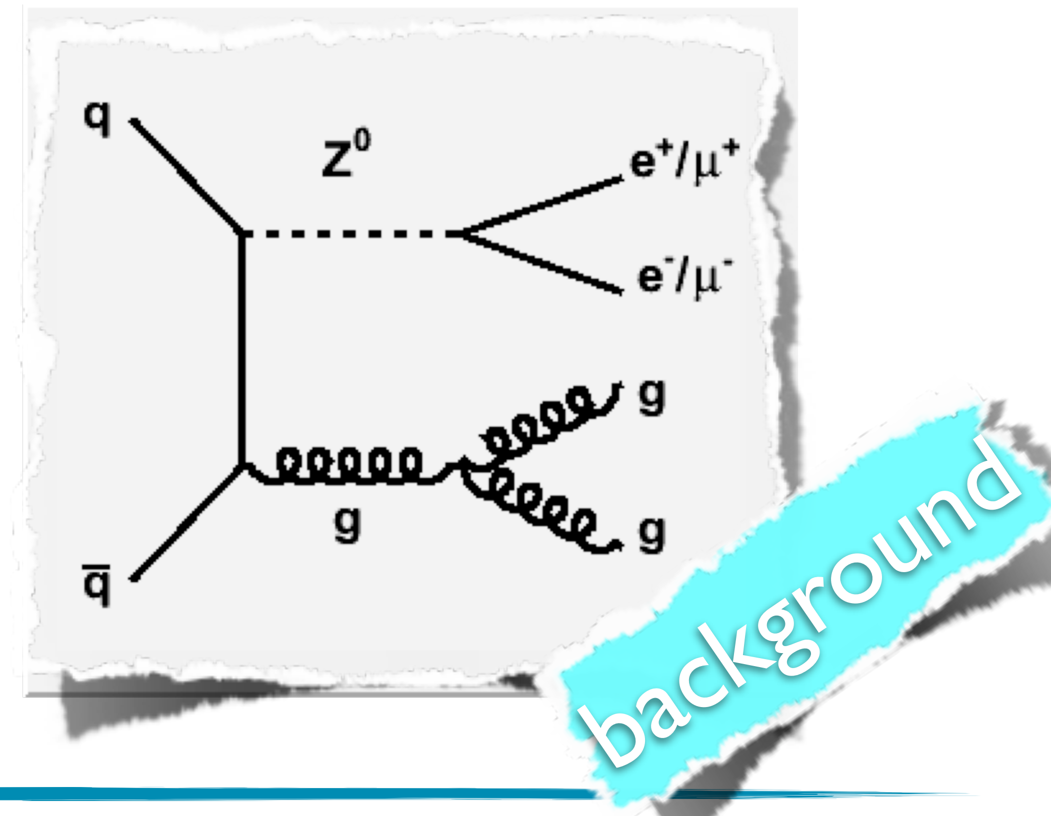
- Precursor for the cleanest VH channel
  - clean but hard
  - low xsection x BR is a killer
- Typical selection:
  - Two leptons above 20GeV
  - Two jets above 20 GeV
  - Low MET



Quarks vs gluons

- Large Z+jets background
  - lots of gluons
- Signal mostly quarks
- Separating would improve S/B a lot
- Several attempts in the past by using shape of jets





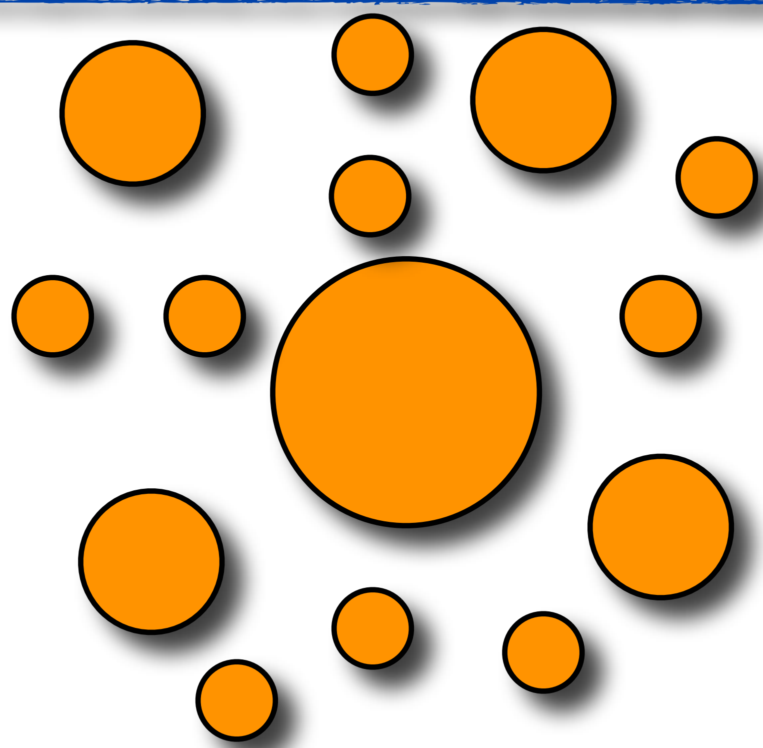
Separating quarks and gluons

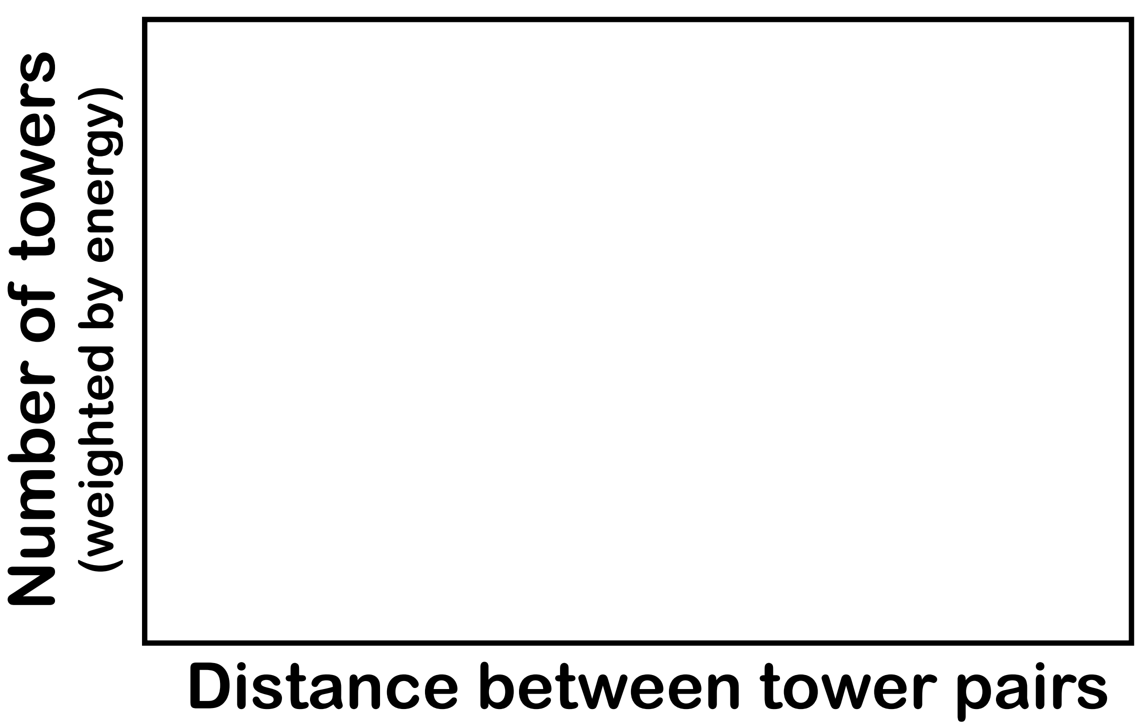
Developed by Wesley Ketchum, PhD student at UChicago

- To distinguish between quarks and gluons, look at distribution of ΔR between all pairs of towers in jet
  - Weight tower pairs by energy relevance ($E_i \times E_j \ / \ E_{tot}^2$)
  - Compare to "typical" quark and gluon jet templates by computing a likelihood (or at least likelihood-like quantity)
  - Templates from ALPGEN $Z \rightarrow ll$ + 2 parton MC
  - Construct Likelihood Ratio: LR = L (jet from quark) / L (jet from gluon)

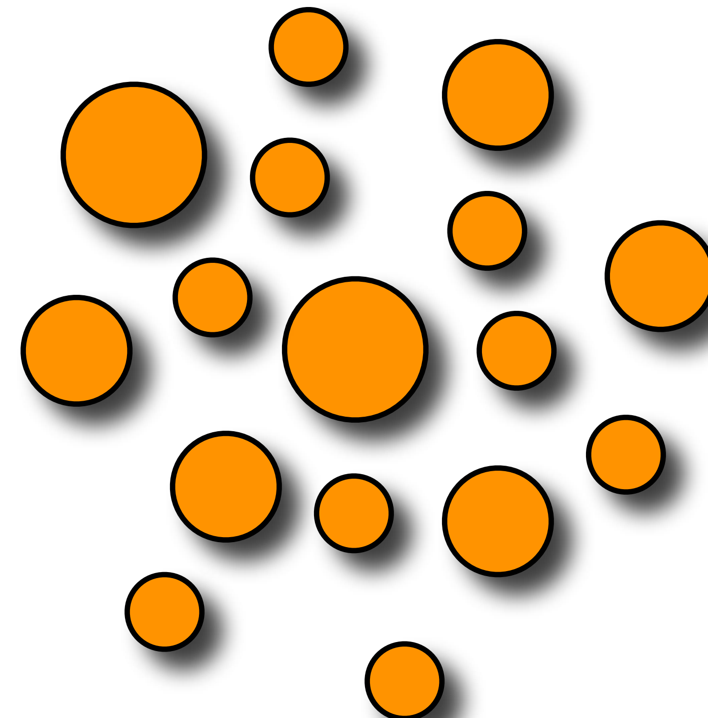
Just like two point correlation function in astrophysical observations

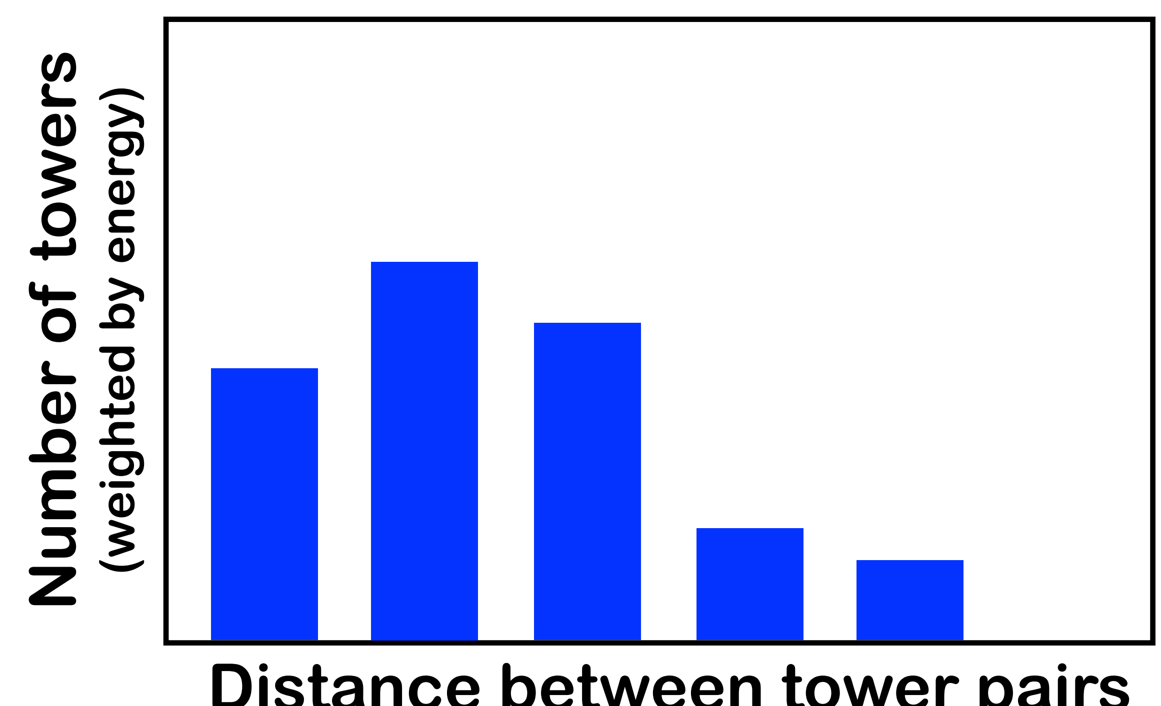
(Light) quarks



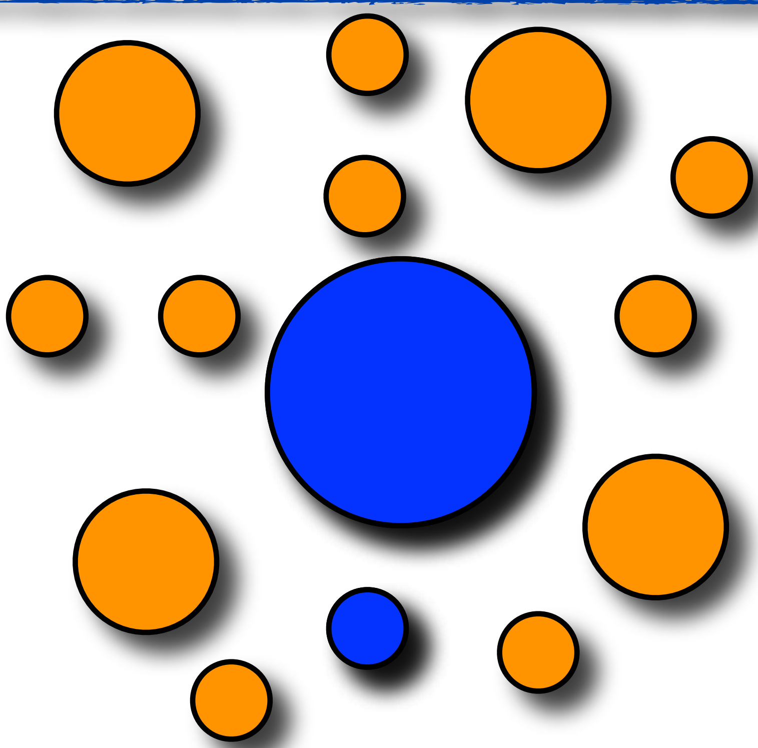


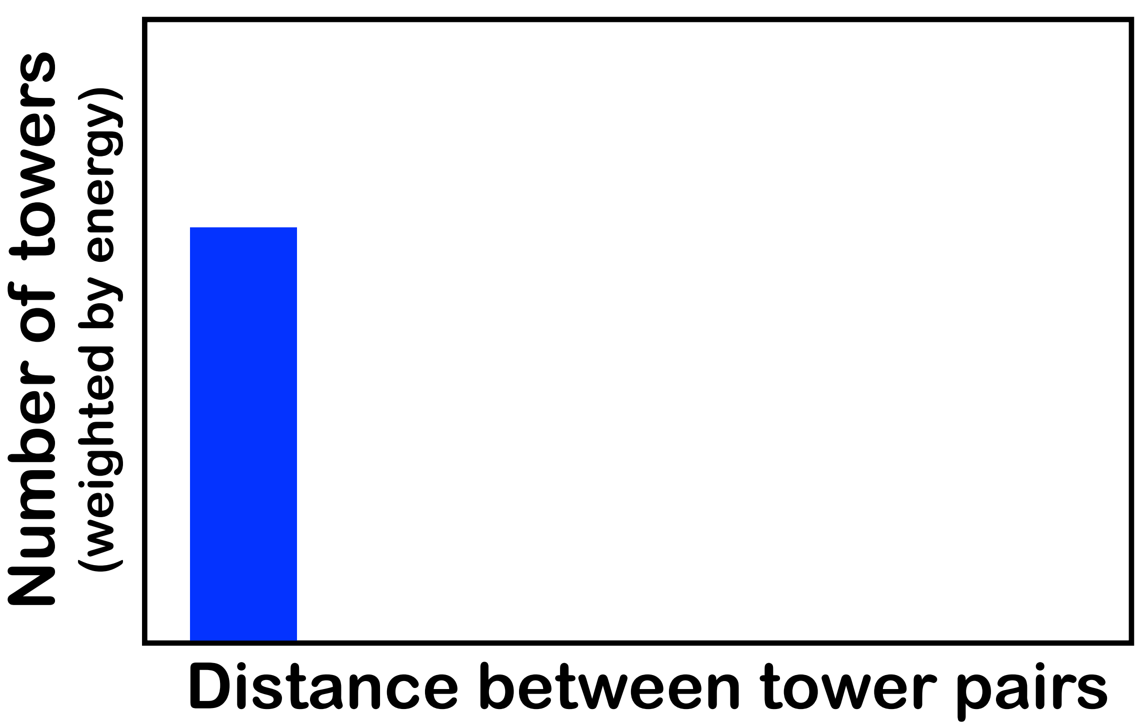
gluons



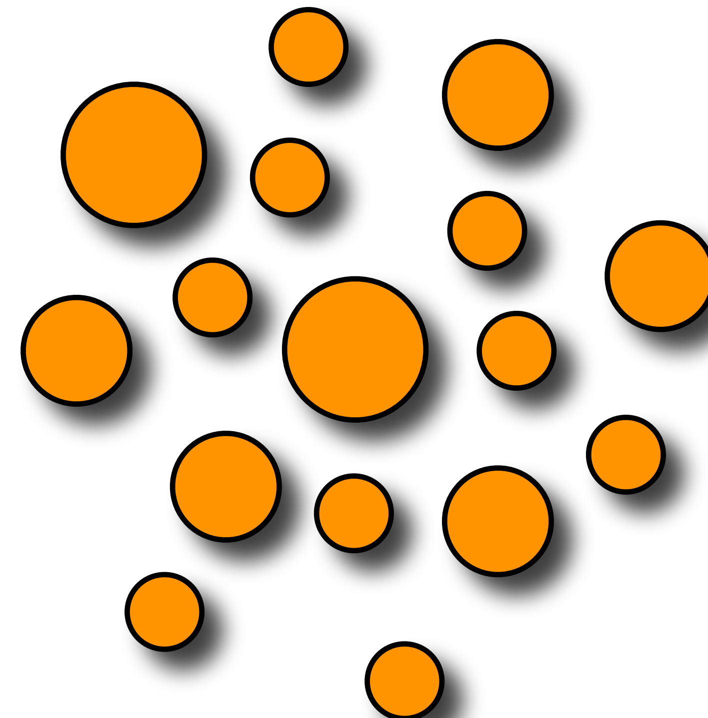


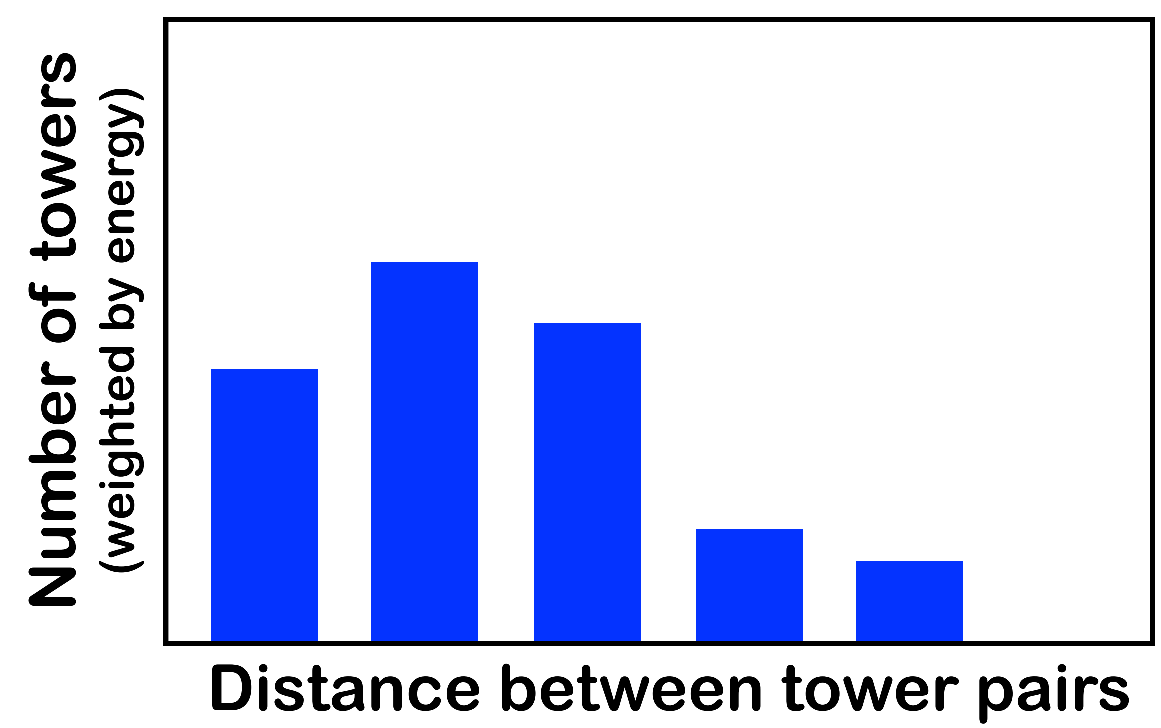
(Light) quarks



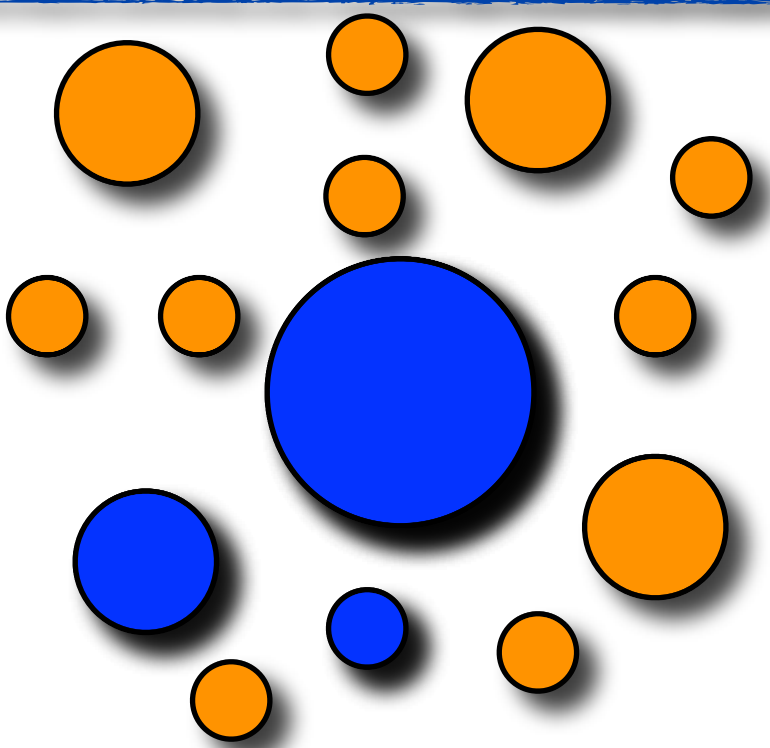


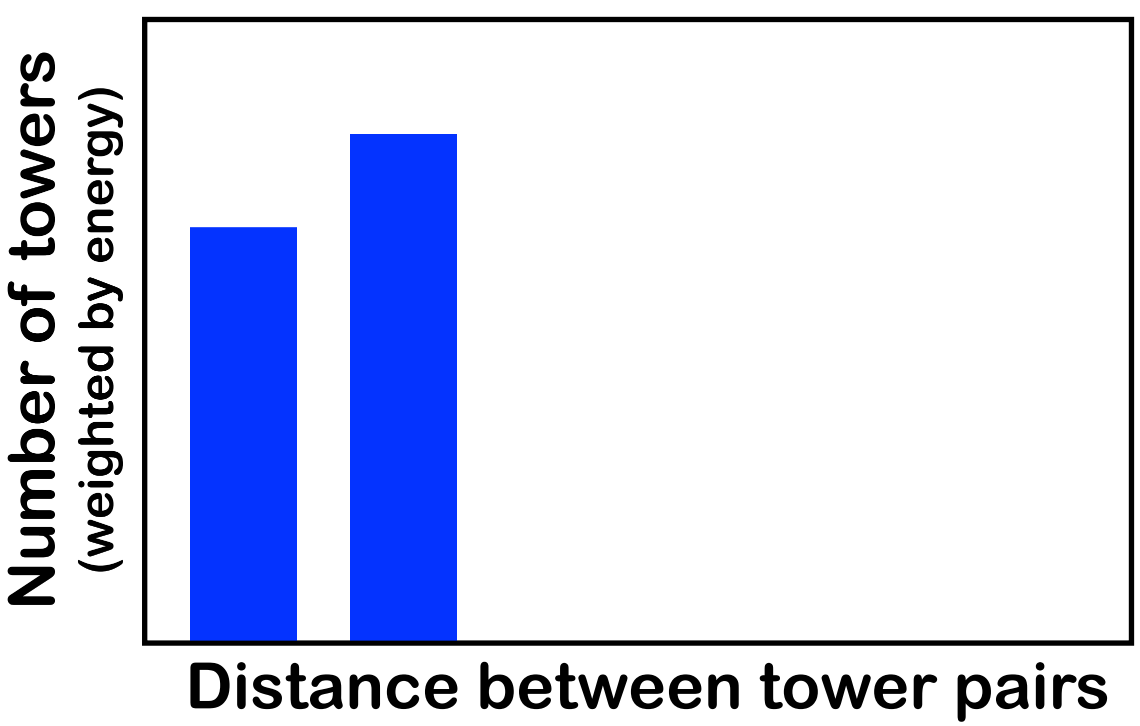
gluons



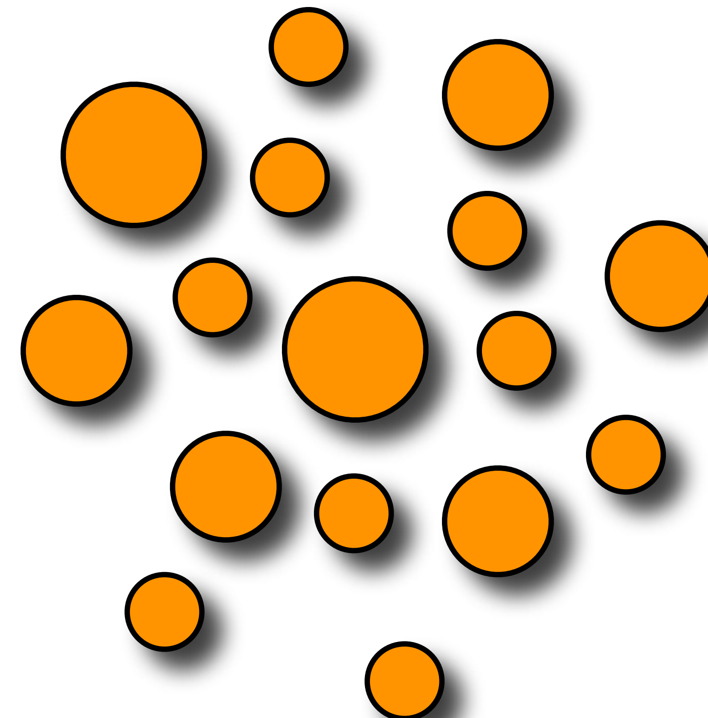


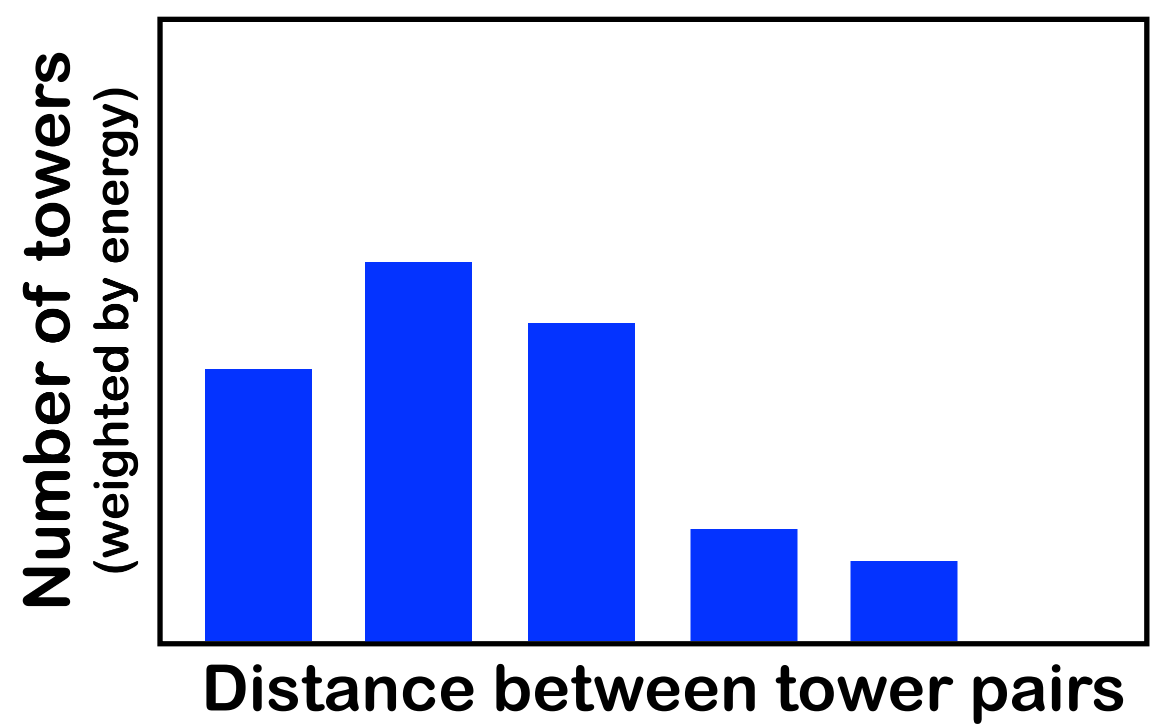
(Light) quarks



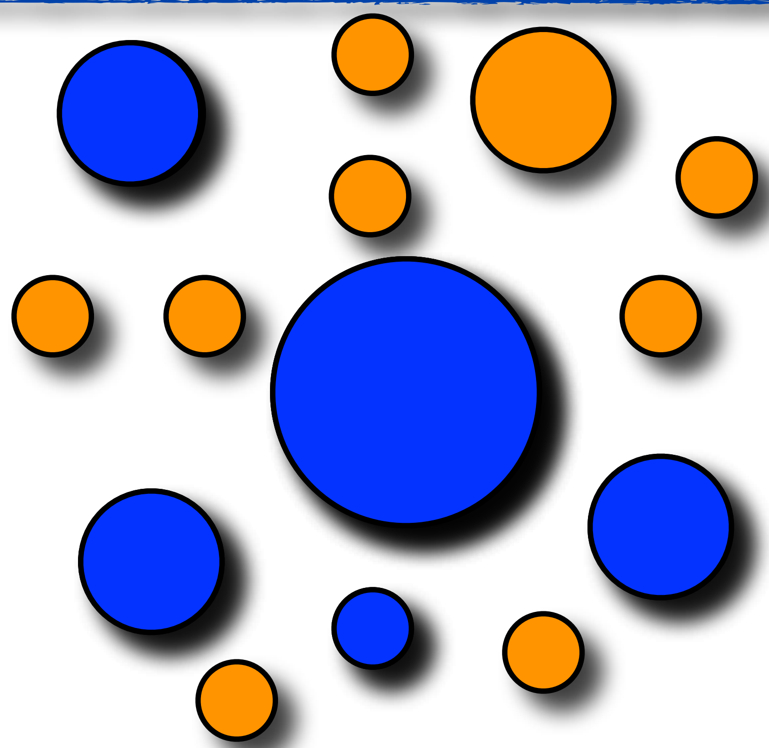


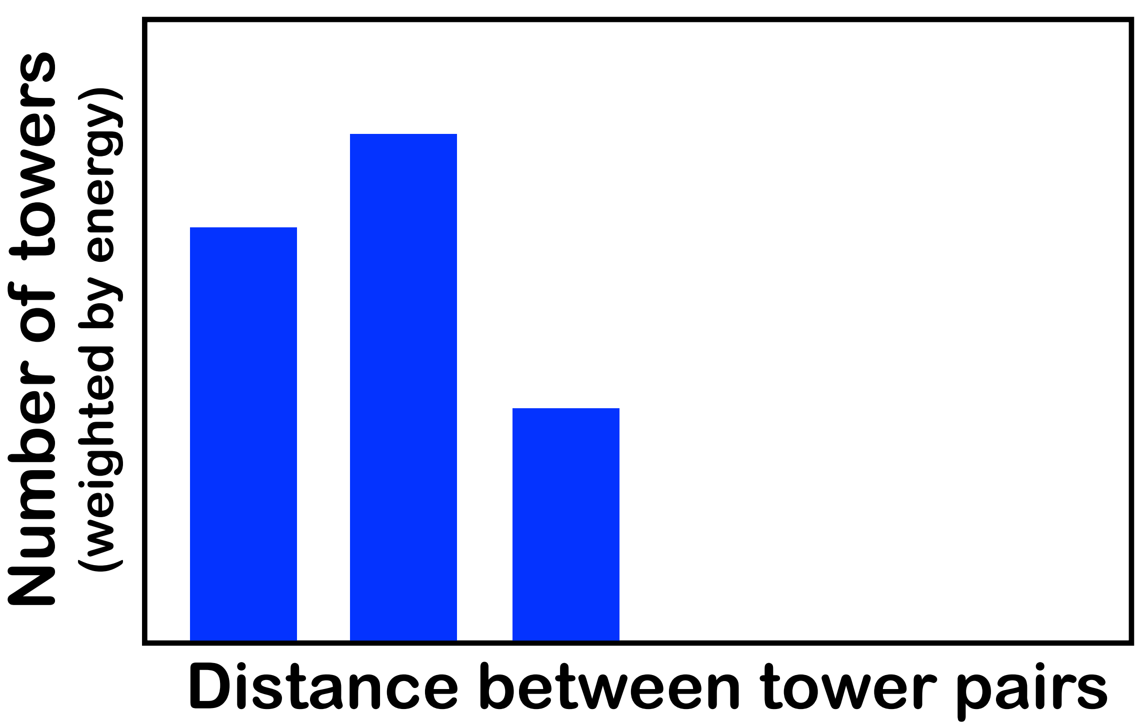
gluons



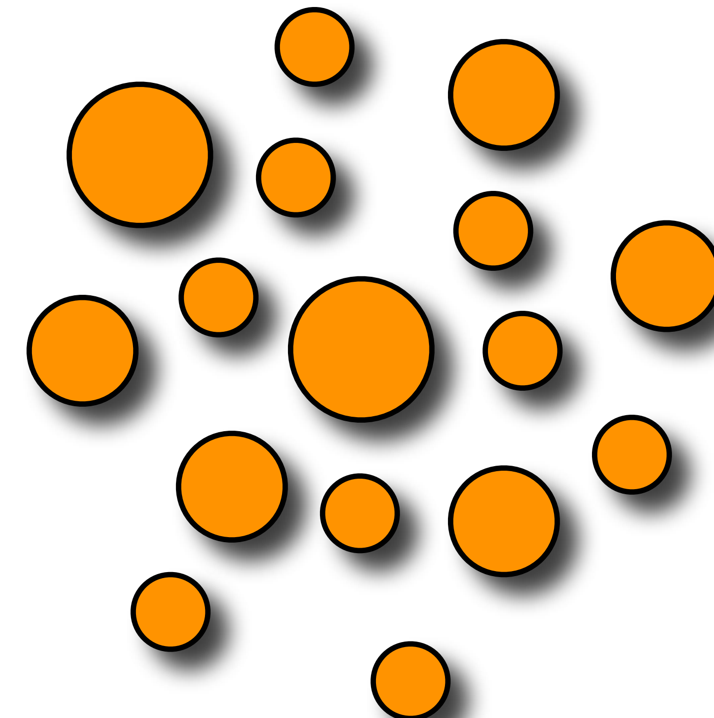


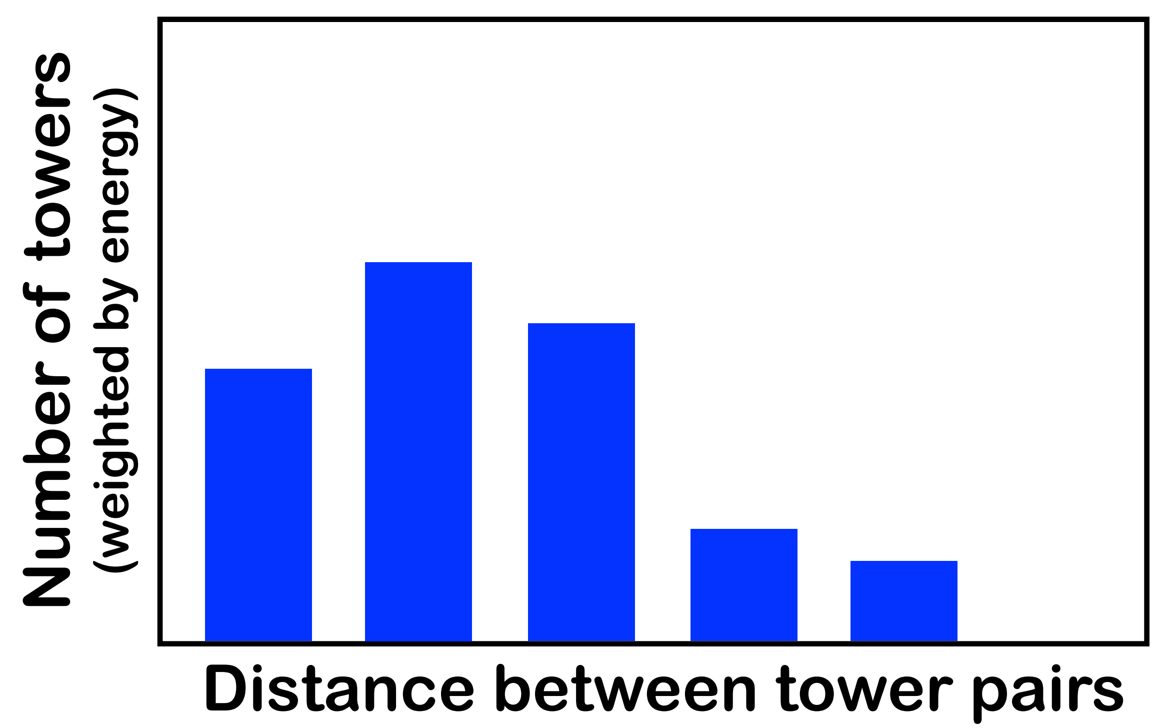
(Light) quarks



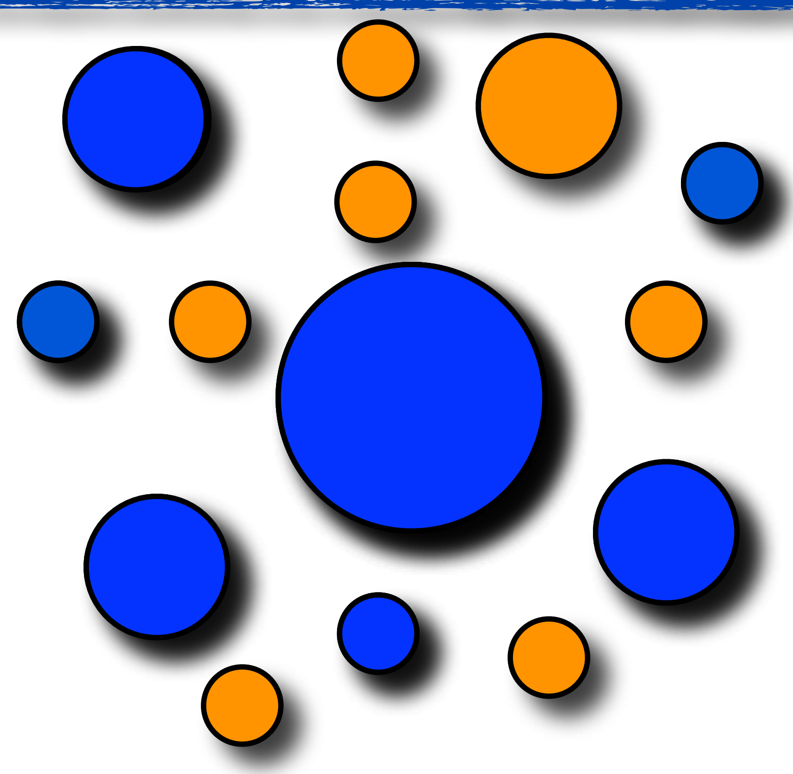


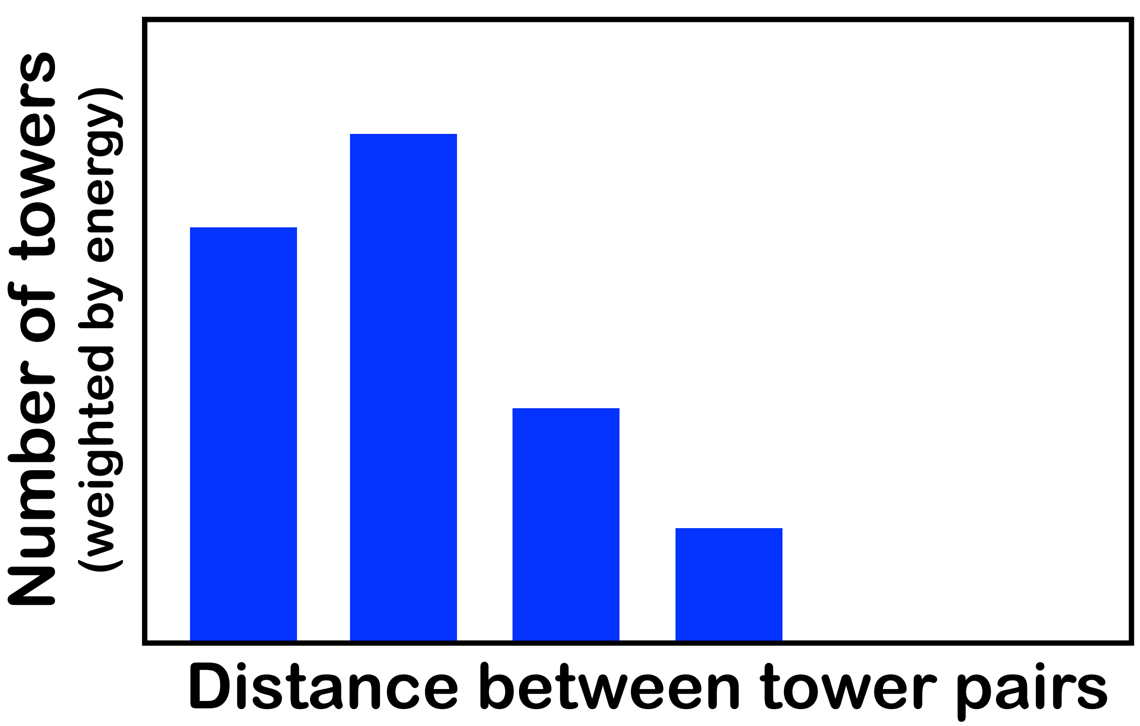
gluons



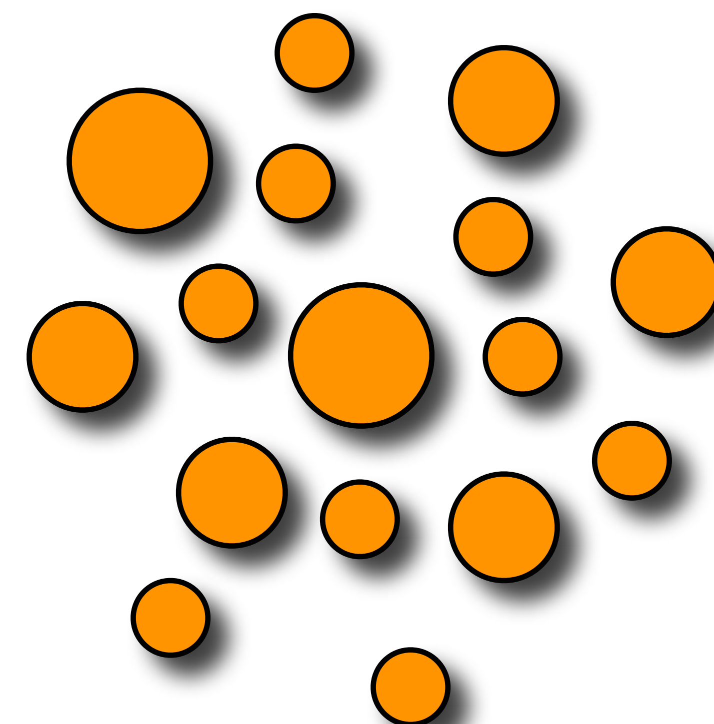


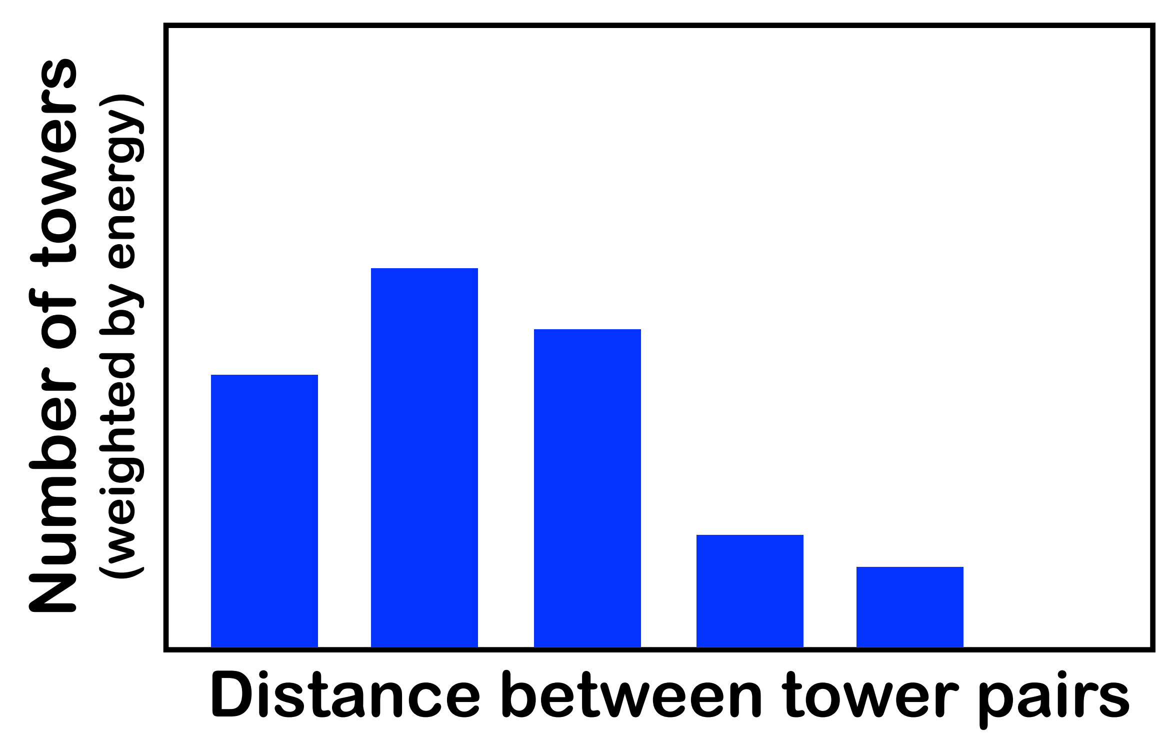
(Light) quarks





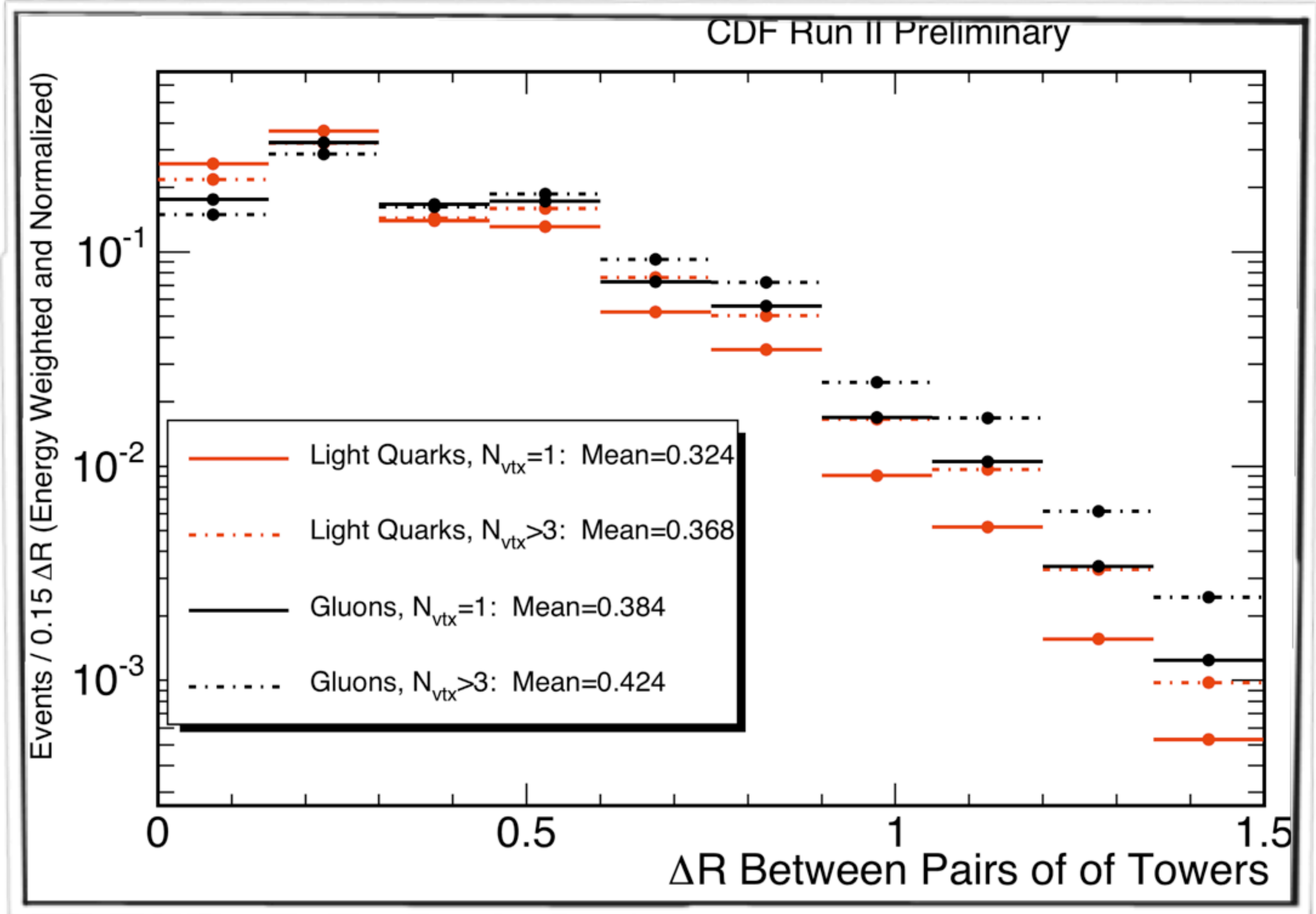
gluons





What do we see?

- The quark and gluon ΔR are very similar
- Remember the power of large numbers
  - There are ~20 towers in a jet
  - N(N-1) large

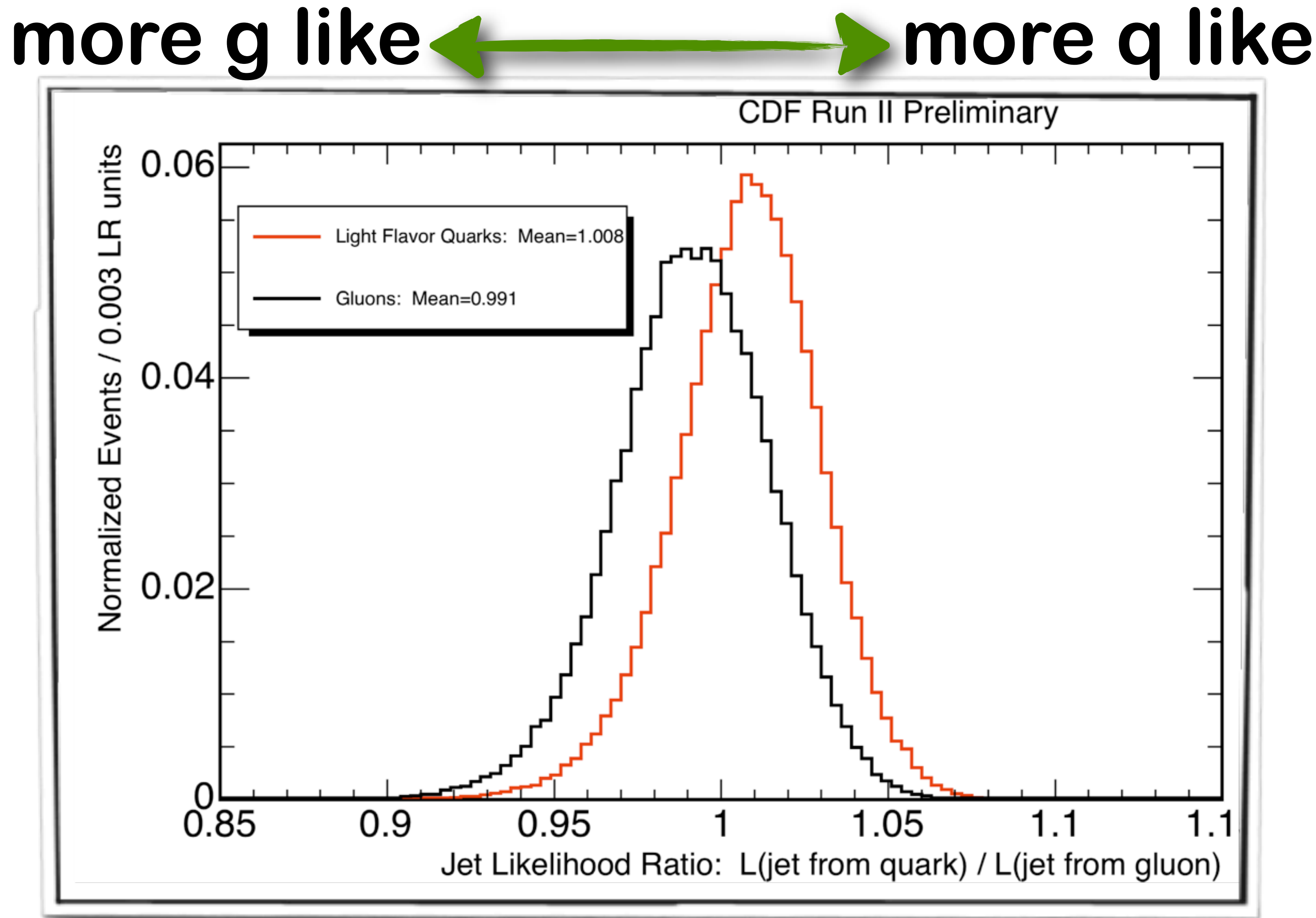


Use different templates for different number of vertices:

More interaction vertices ➜ more energy distributed throughout calorimeter ➜ jets look more spread out ➜ jets look more gluon-like

The likelihood ratio (LLR)

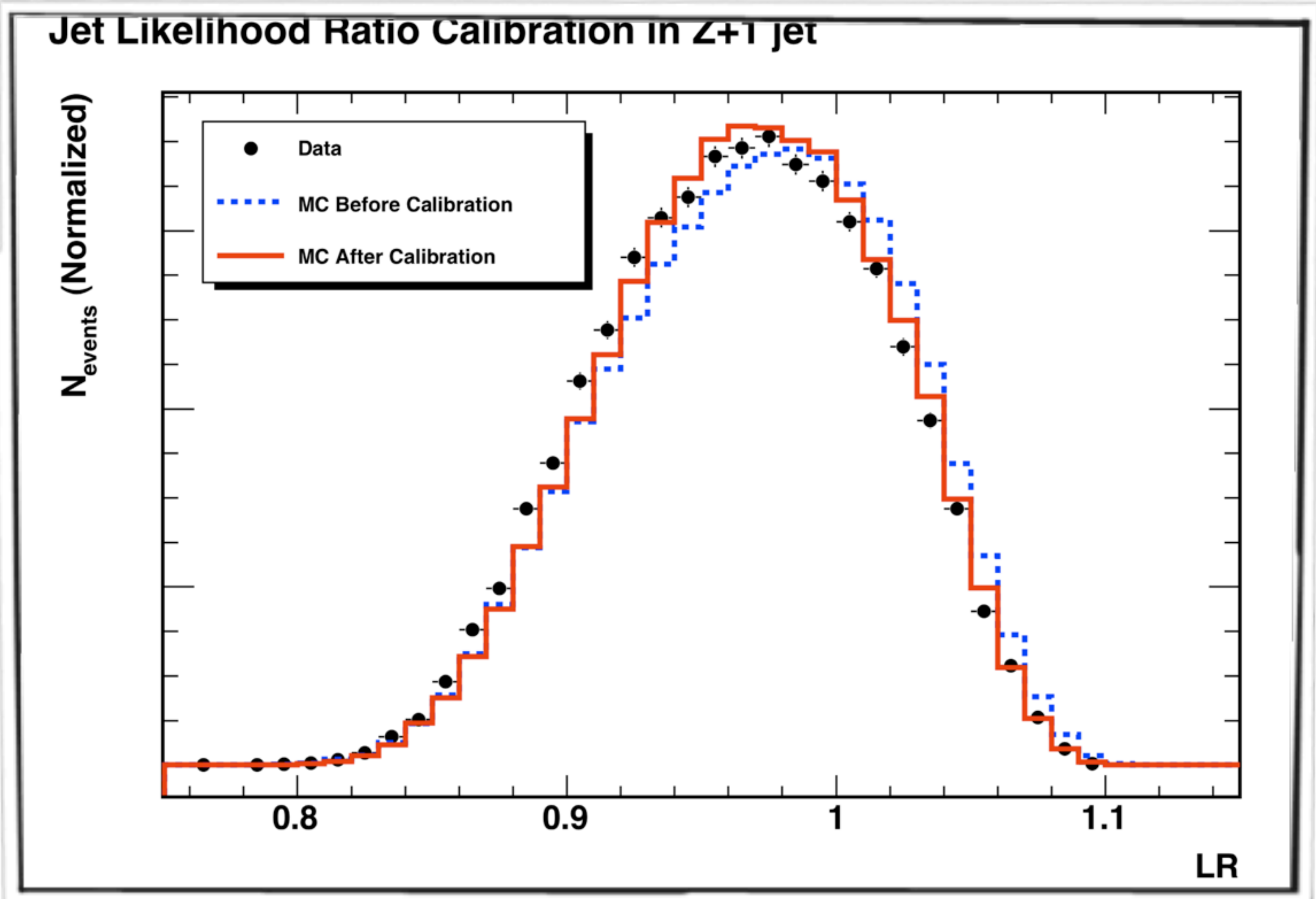
- Quarks and gluons look different enough
- Separate on a statistical basis



Not enough for a straight cut but useful in a multivariate analysis

Check in Z+1 jet

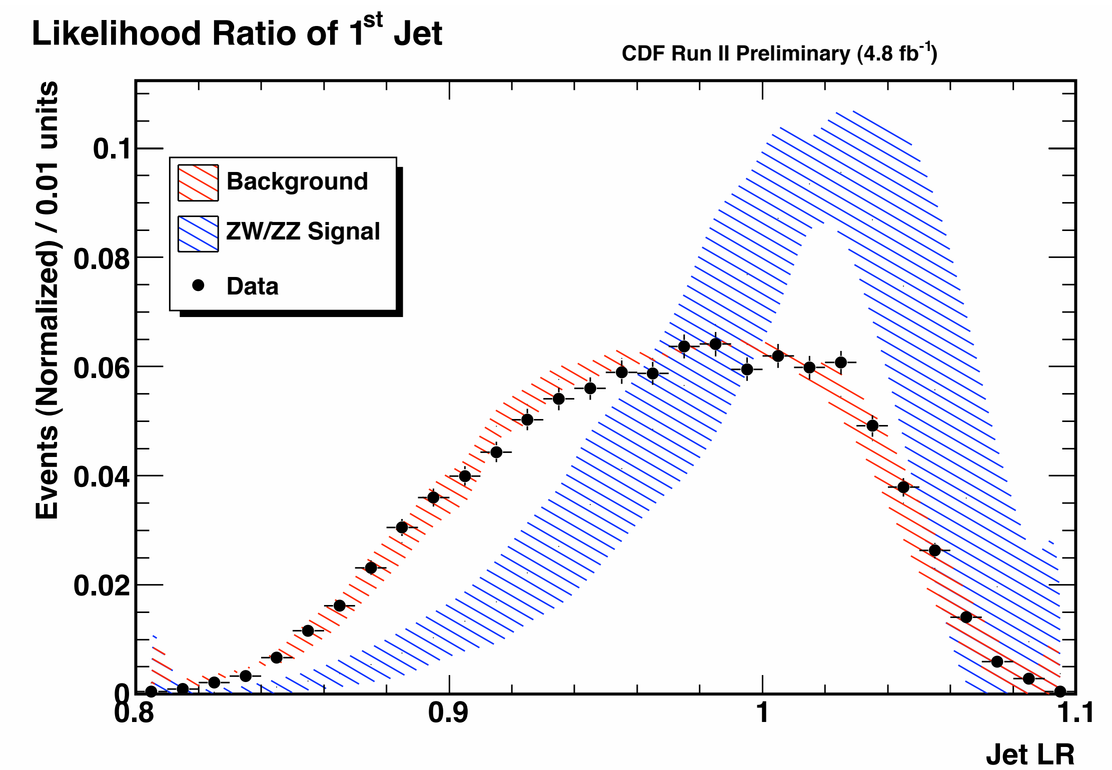
- Still not the “signal” region
  - Large stats sample
  - Check how the calibration worked
- A bit off but much better agreement
  - This final difference can be used to asses systematics

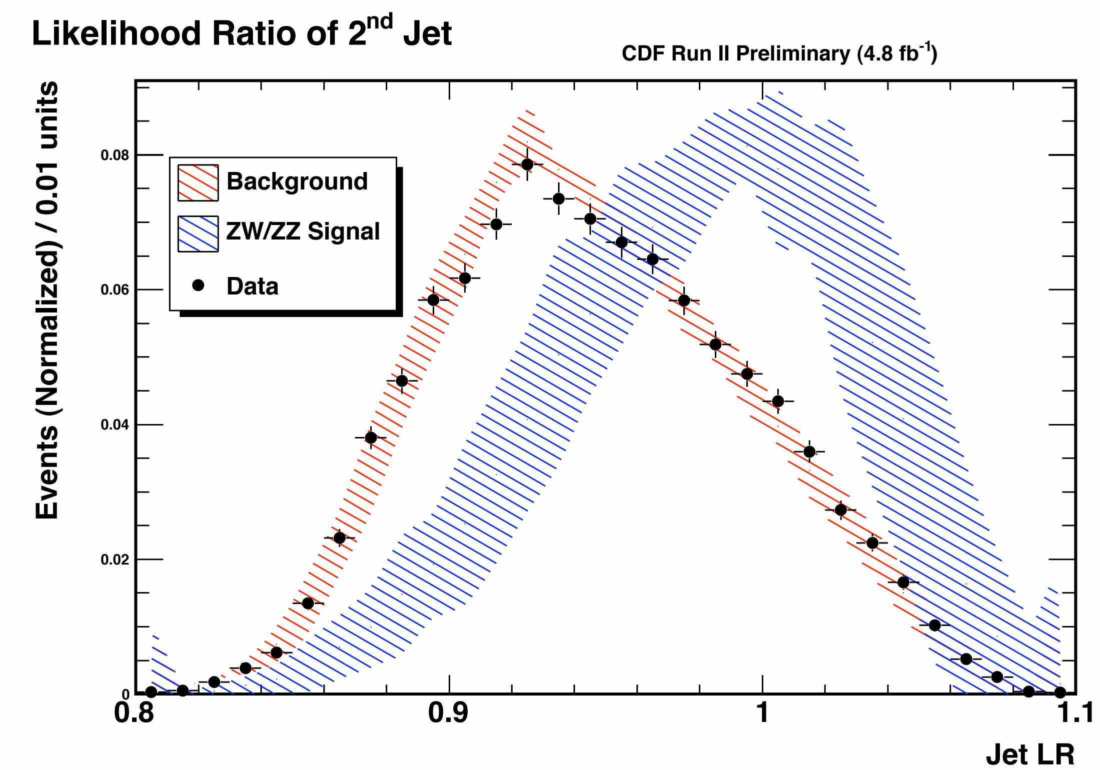


We now have a q/g separator which MC can described quite well

Finally, the Z+2 jets

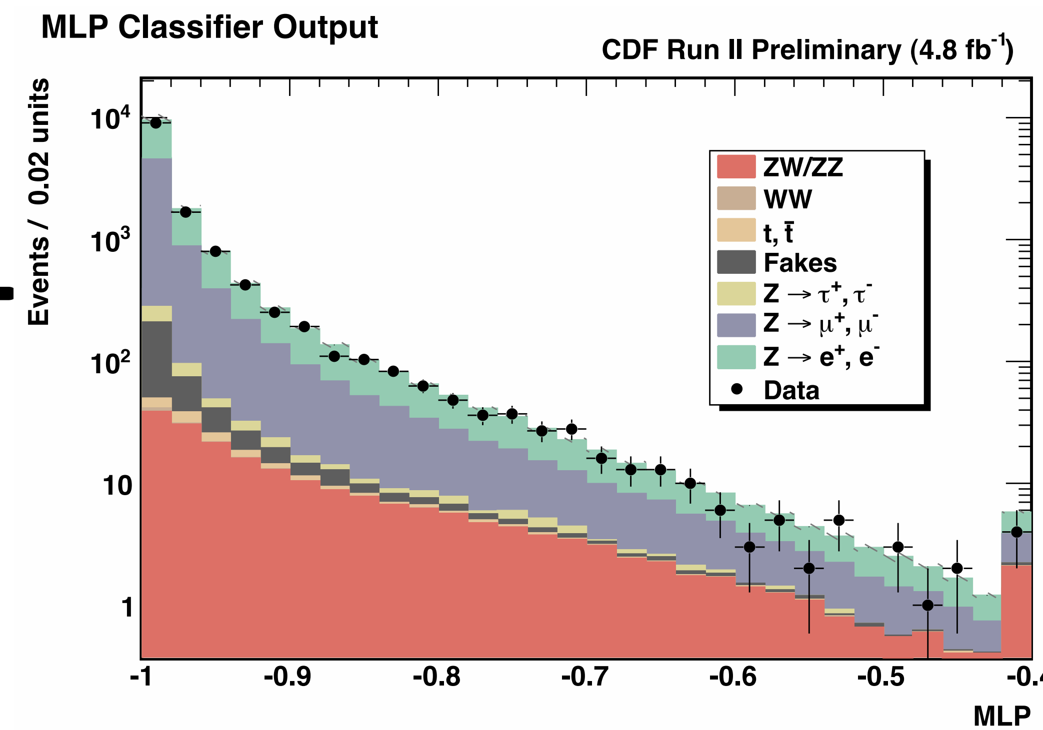
- This is our final sample of interest
  - Significant separation between quark and gluon jets
  - Can be used on a multivariate analysis
  - Double gain, because we can use it on both jets
  - Systematic band derived from Z+1jet
- Ways to improve?

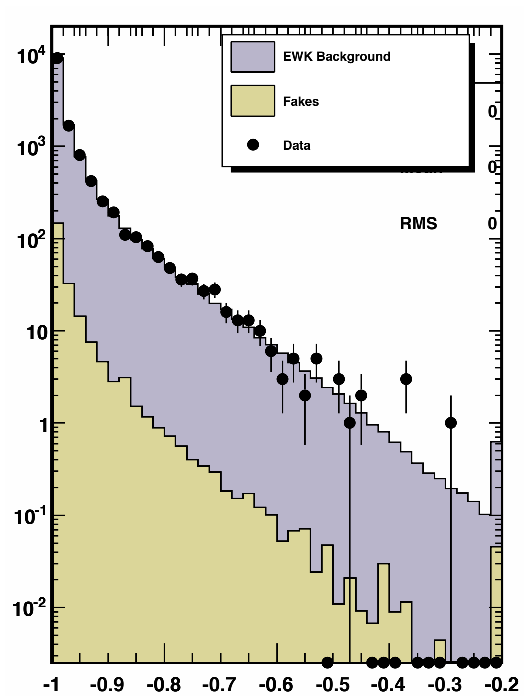


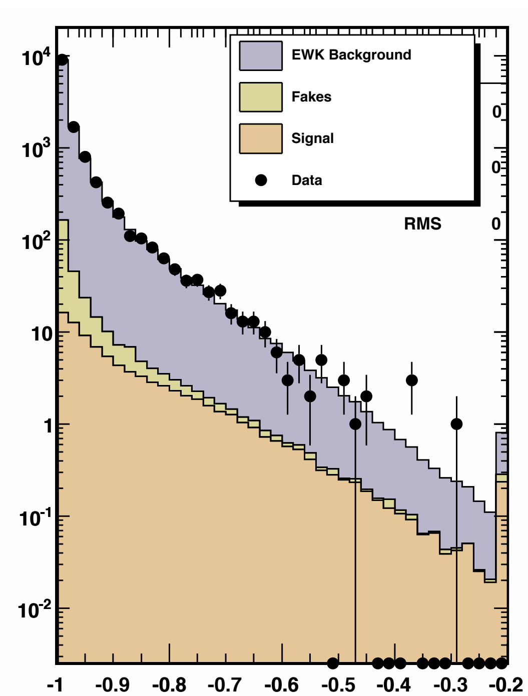


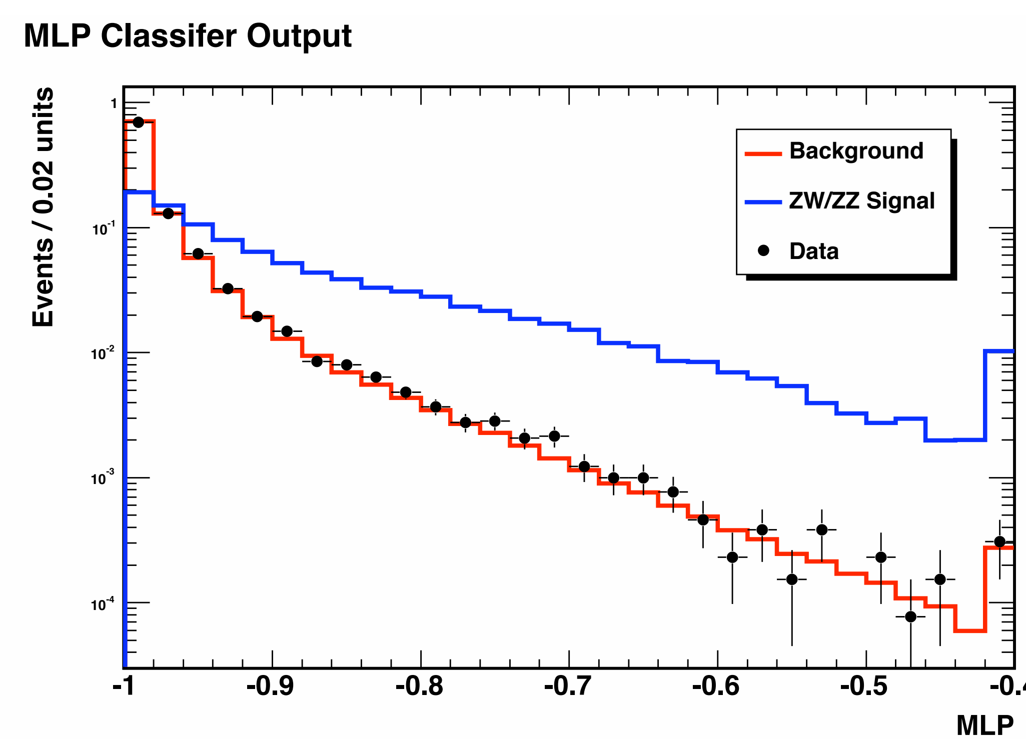
Final result

- σ_{WZ+ZZ} < 18pb at 95% CL
  - 2.4 × SM



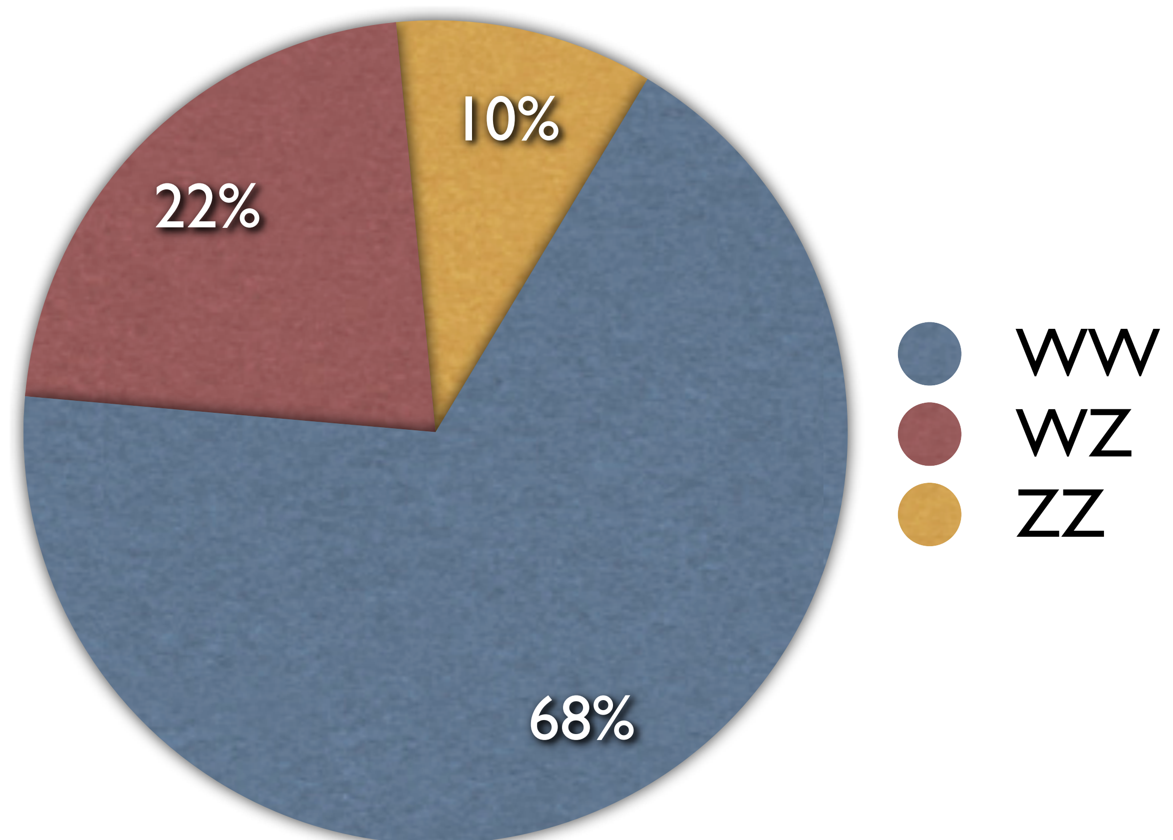






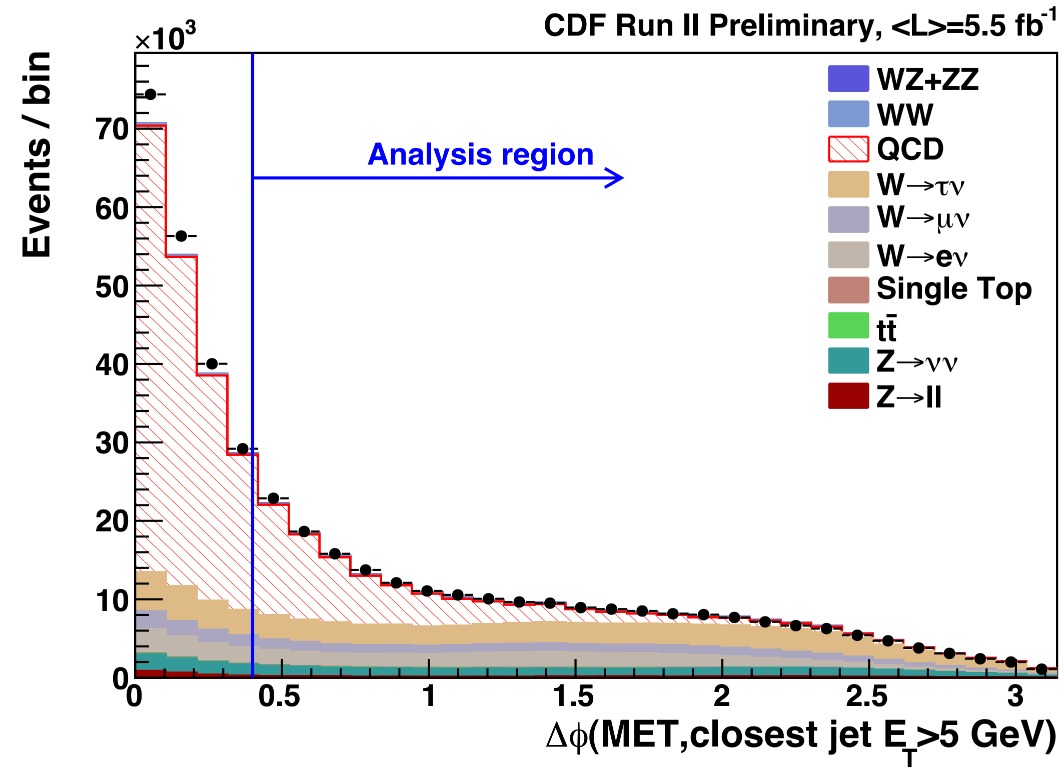
MET+jets

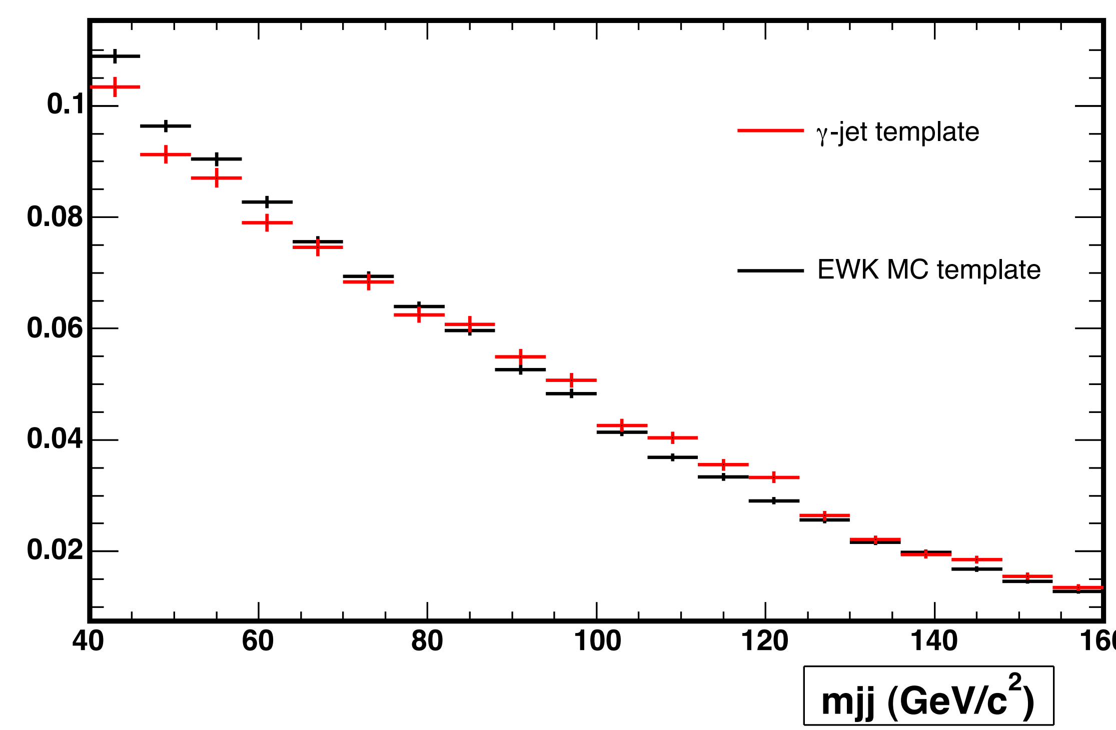
- The highest acceptance
  - it pays
- First observation of dibosons in hadronic final states
- Selection:
  - Jets ET above 25
  - MET above 60
  - +++ lots of cleaning cuts
- No btagging → WW dominates



Two big backgrounds

- QCD - multijet production
  - MET comes from mismeasurement
  - or resolution effects
- EWK - W+jets, Z+jets
  - Rely on MC model
    - ALPGEN, pythia, MadGraph, Sherpa, etc
    - none of them reproduce the CDF data completely
  - γ+jets similar production, use for systematics

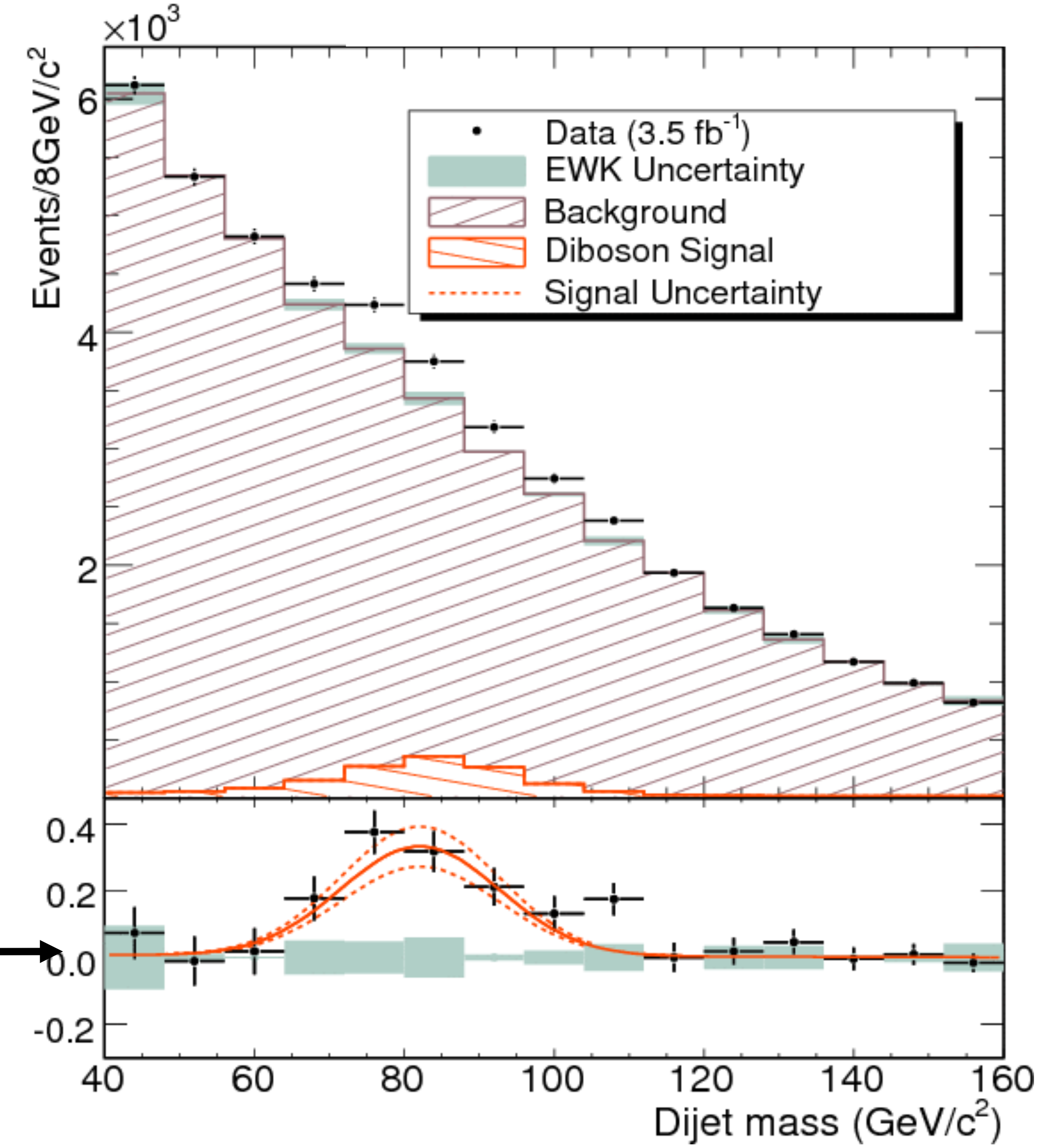




Results

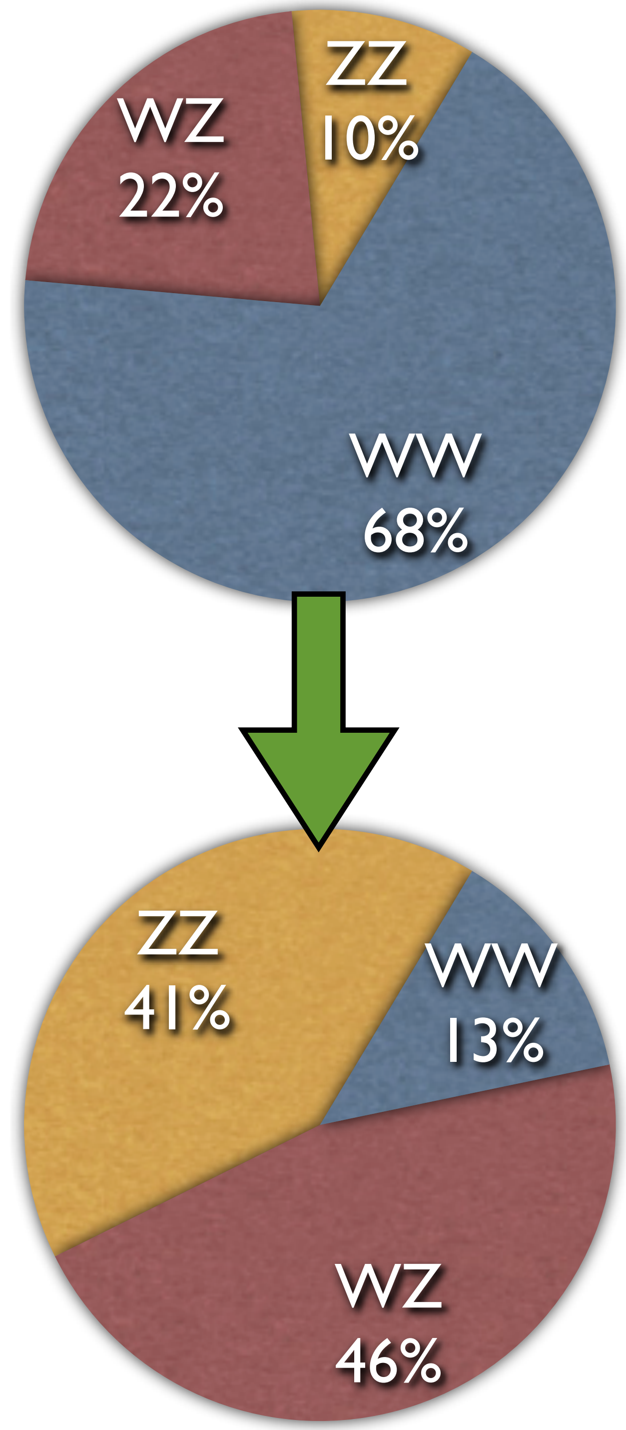
- Fit result:
  - 1516 ±239 (stat) ± 144(sys)
  - Expected 1398 ± 243

Data - background



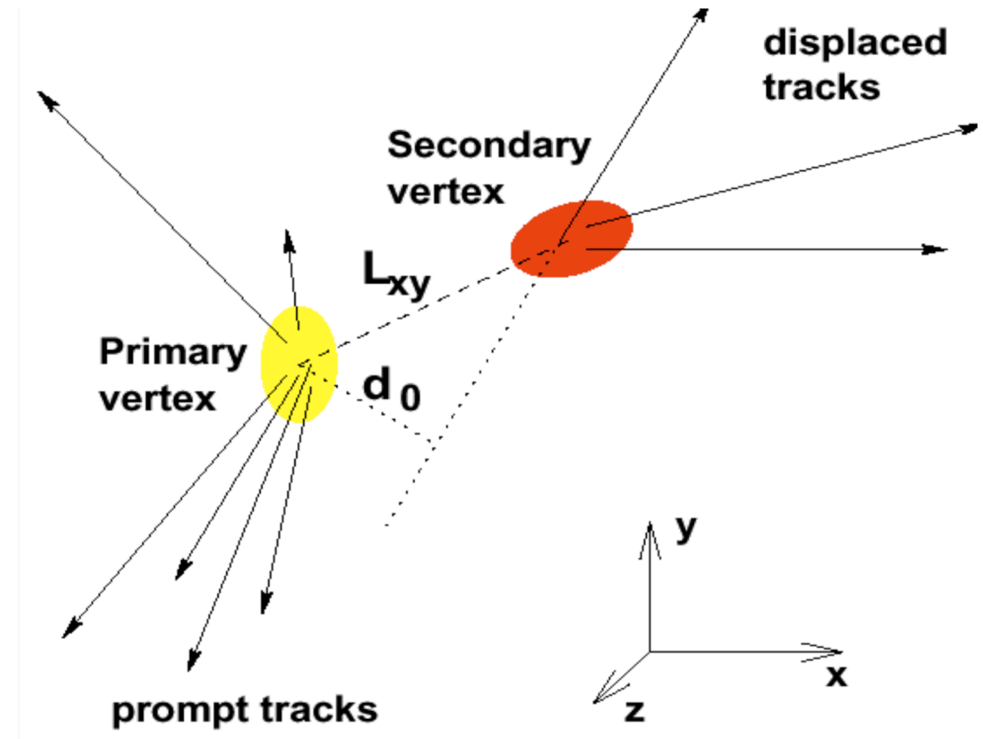
b quarks

- Why is it important?
  - H→bb
  - top physics BR(t→Wb) = 100%
  - New Physics most likely in 3'rd gen.
    - sbottom, stop decays into b+X
- Not something new
  - LEP, HERA, Run 1
- Precision vertexing
  - Silicon vertex detector



Signature of b-jets

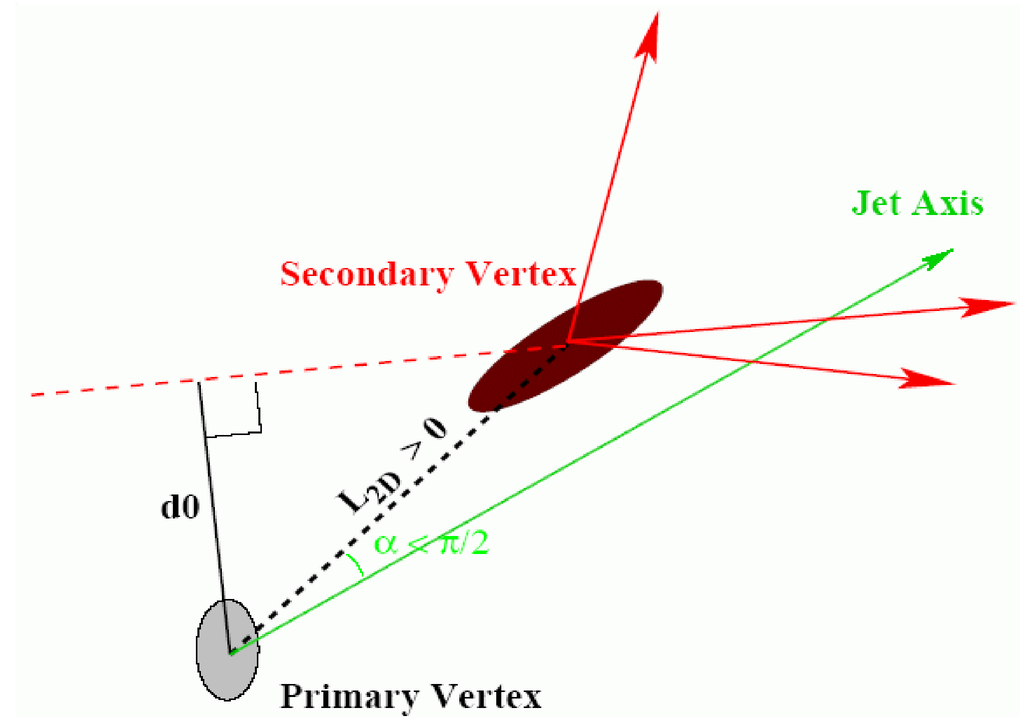
- b - long lived (!)
  - 1.5ps → cτ = 450micron
- b heavy (!)
  - 5.3GeV/c^2
- hard fragmentation
  - B hadron retains 70% of b- quark momentum



- High PT tracks
- Secondary vertices
  - Large impact parameters

A new btagger

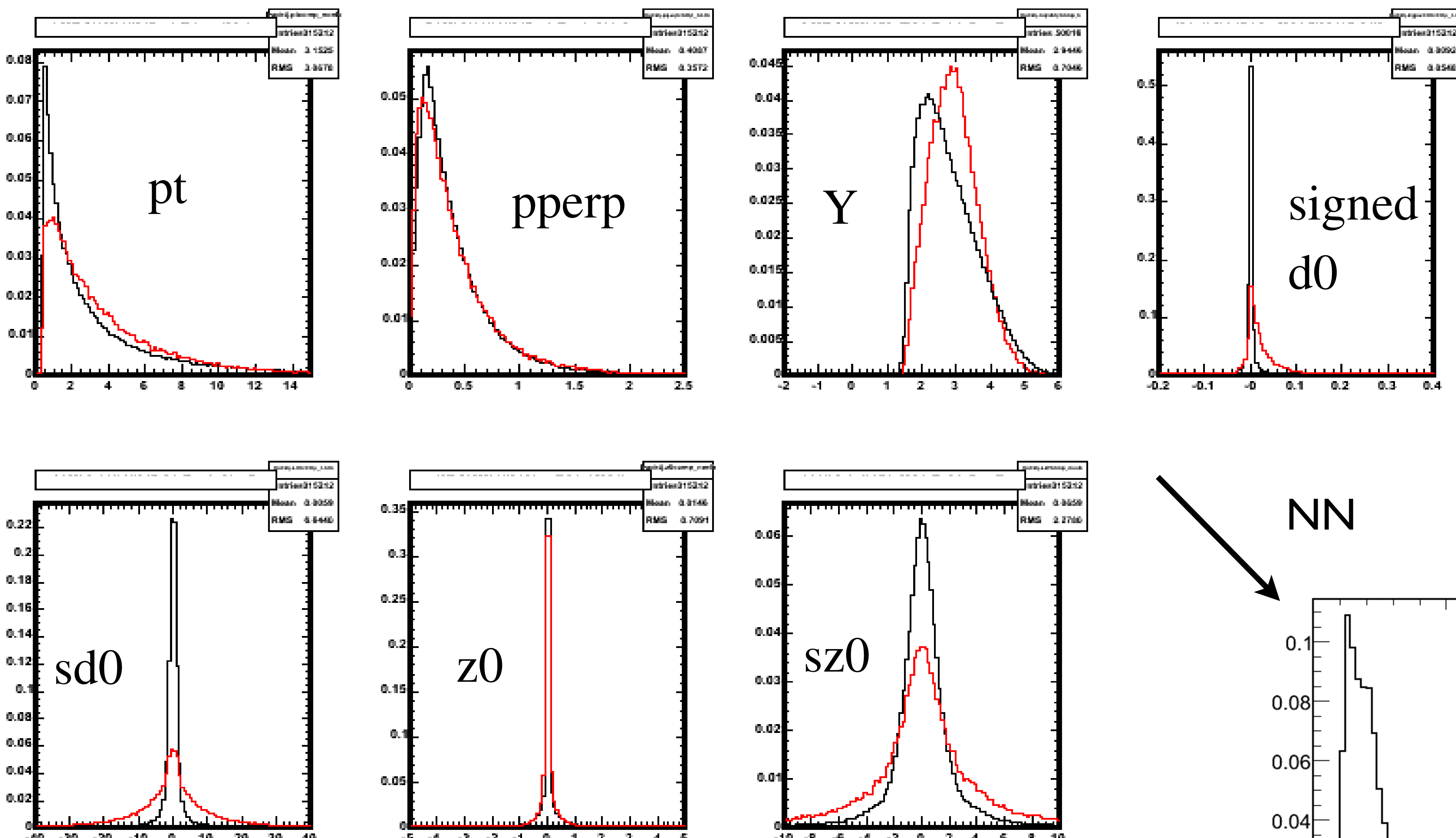
- First rank tracks
  - track displacement
    - signed impact parameter (d0) and significance for impact parameter (sd0)
    - z position wrt primary vertex and its significance
  - high B momentum
    - Track Pt
    - Momentum transverse to jet axis (pperp)
    - Rapidity wrt jet axis (Y)

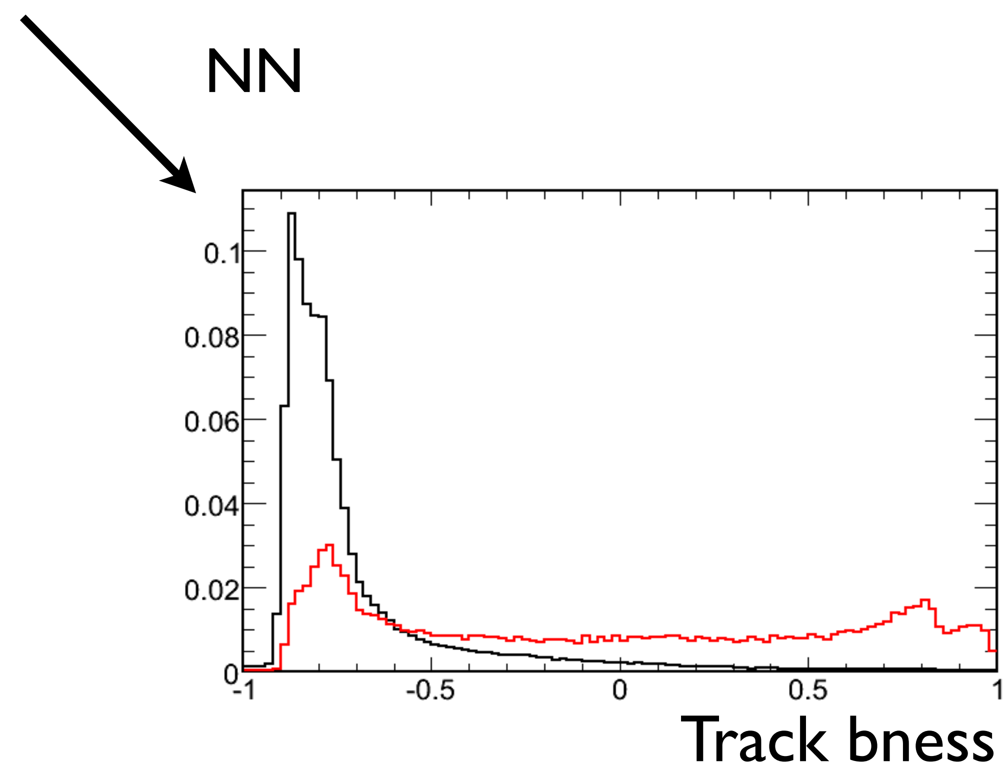


developed by John Freeman (Fermilab postdoc) and by Stephen Poprocki (Cornell PhD student)

Track bness inputs

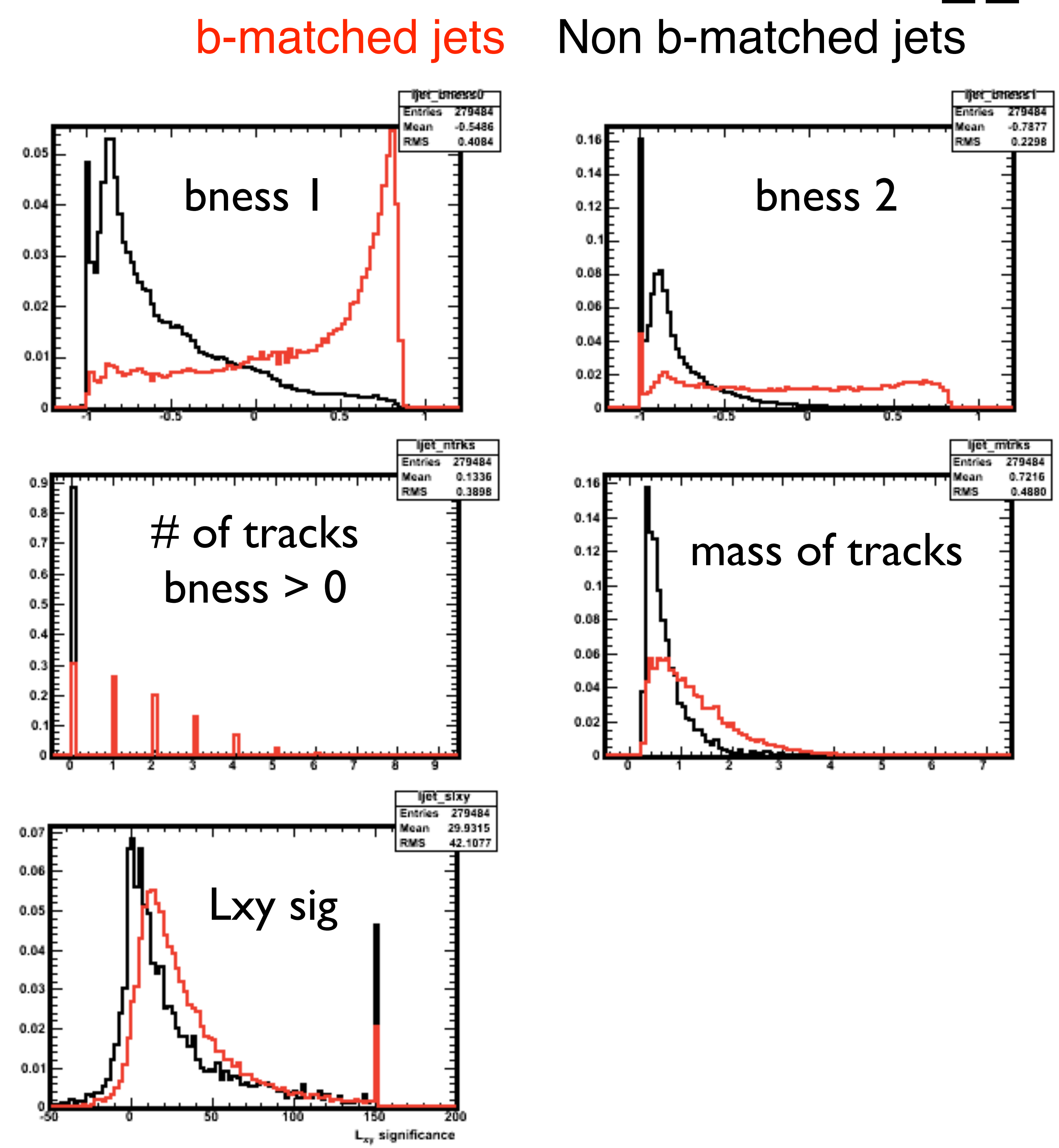
B daughters   Non-B daughters





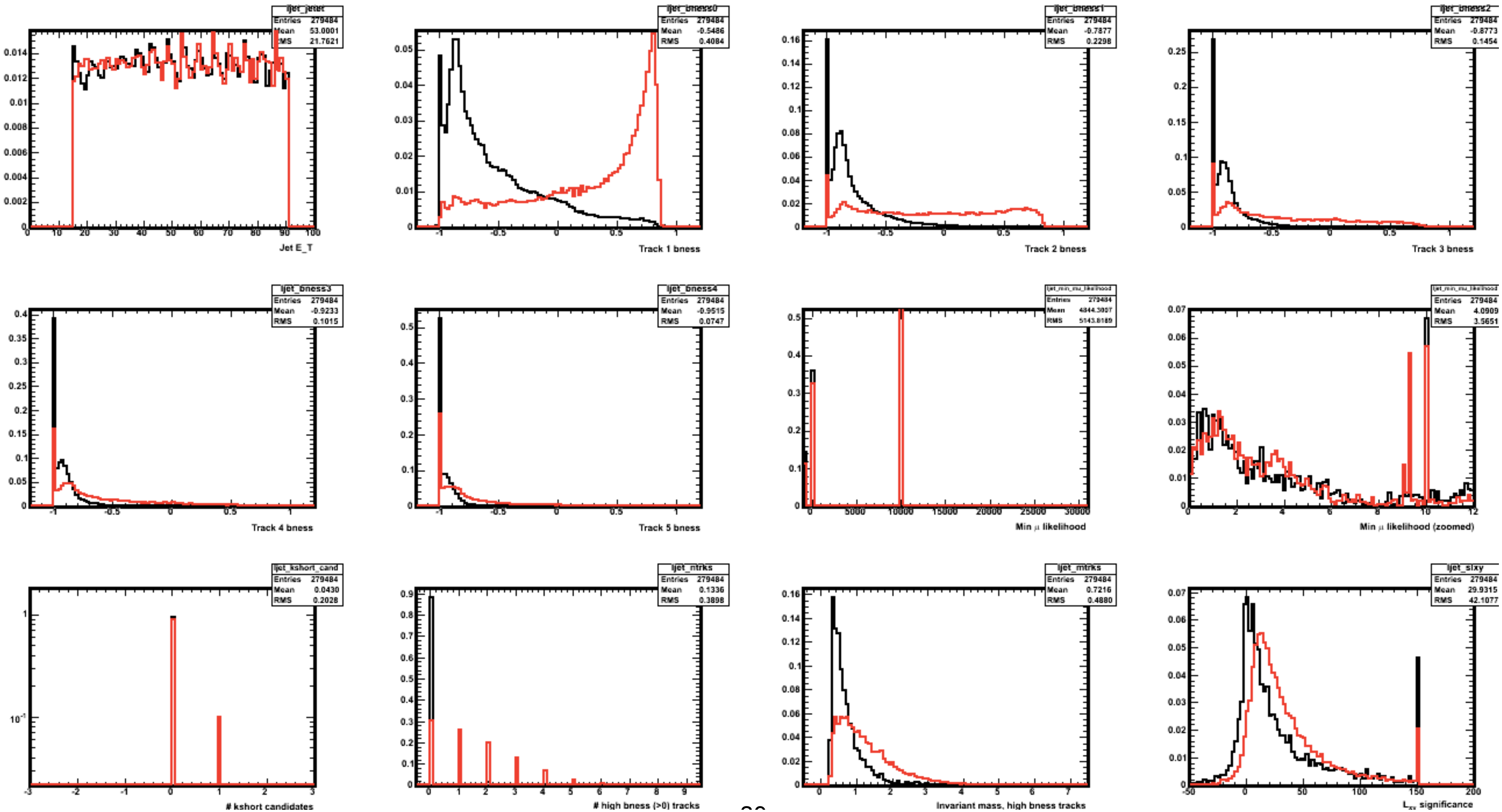
Jet bness

- Top 5 tracks bness (bness1, etc.)
- # of tracks with bness > 0
  - inv. mass of those
- Lxy significance
- Min muon likelihood form SLT
- # of KS candidates



All jet bness inputs

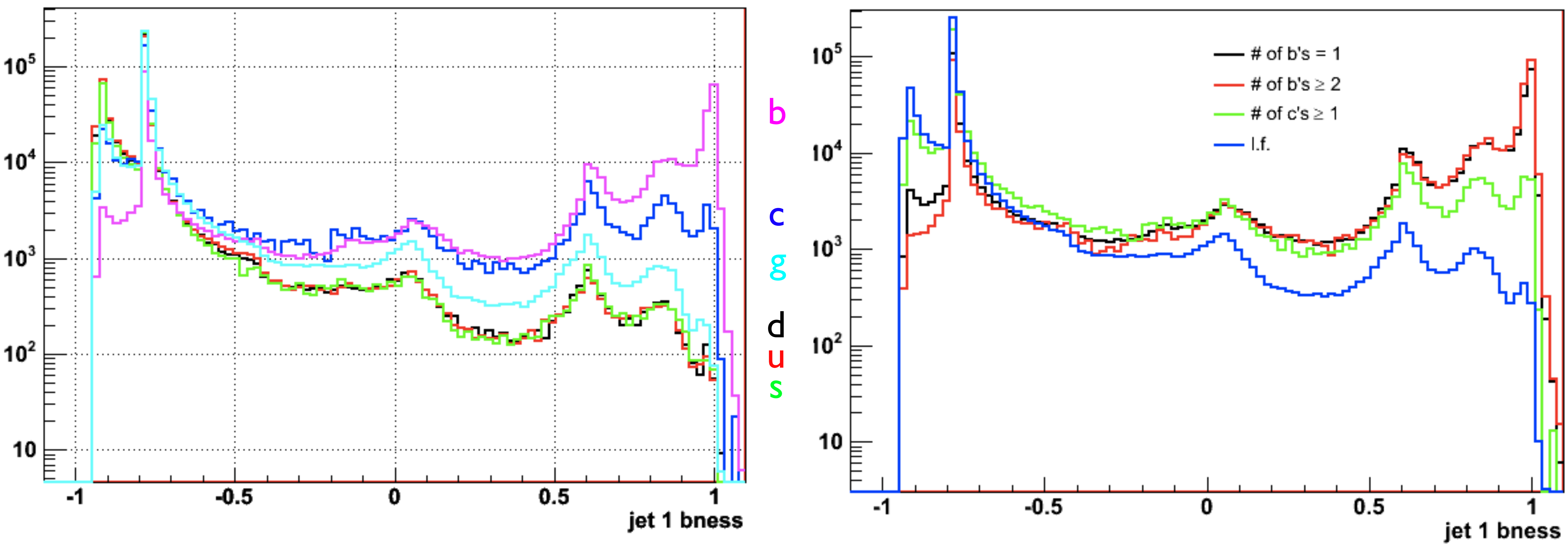
B daughters   Non-B daughters



Jet bness output

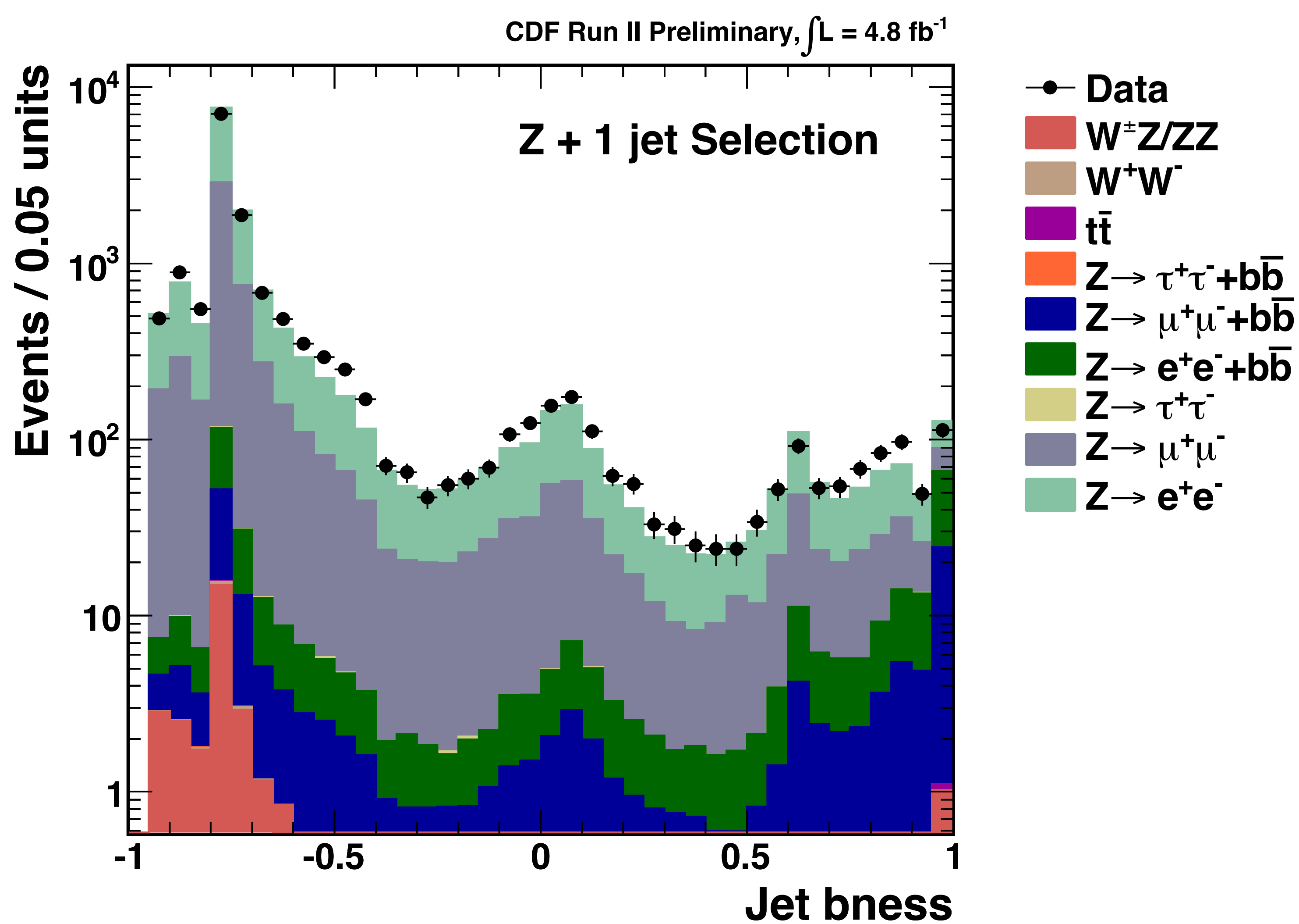
- First try: dijet MC
- Jets matched to just 1 parton (DR<0.4)
- Normalized to equal area

- Including jets matched to multiple partons
- If there is at least 1b does not seem to matter if other partons are also matched



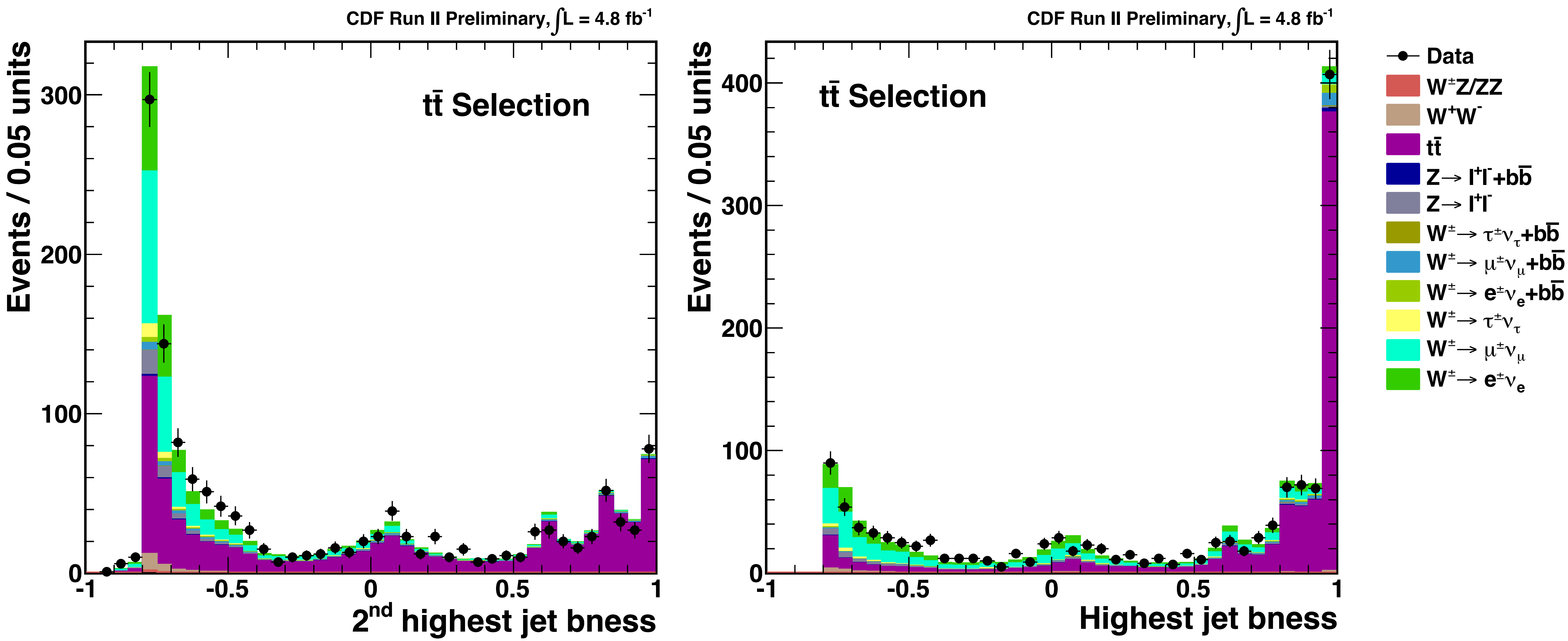
How about data

- Again, all this nice MC have to get a reality check from the data (agh!)
- Z+1jet mostly non-b jets
- used to measure the mistag rates

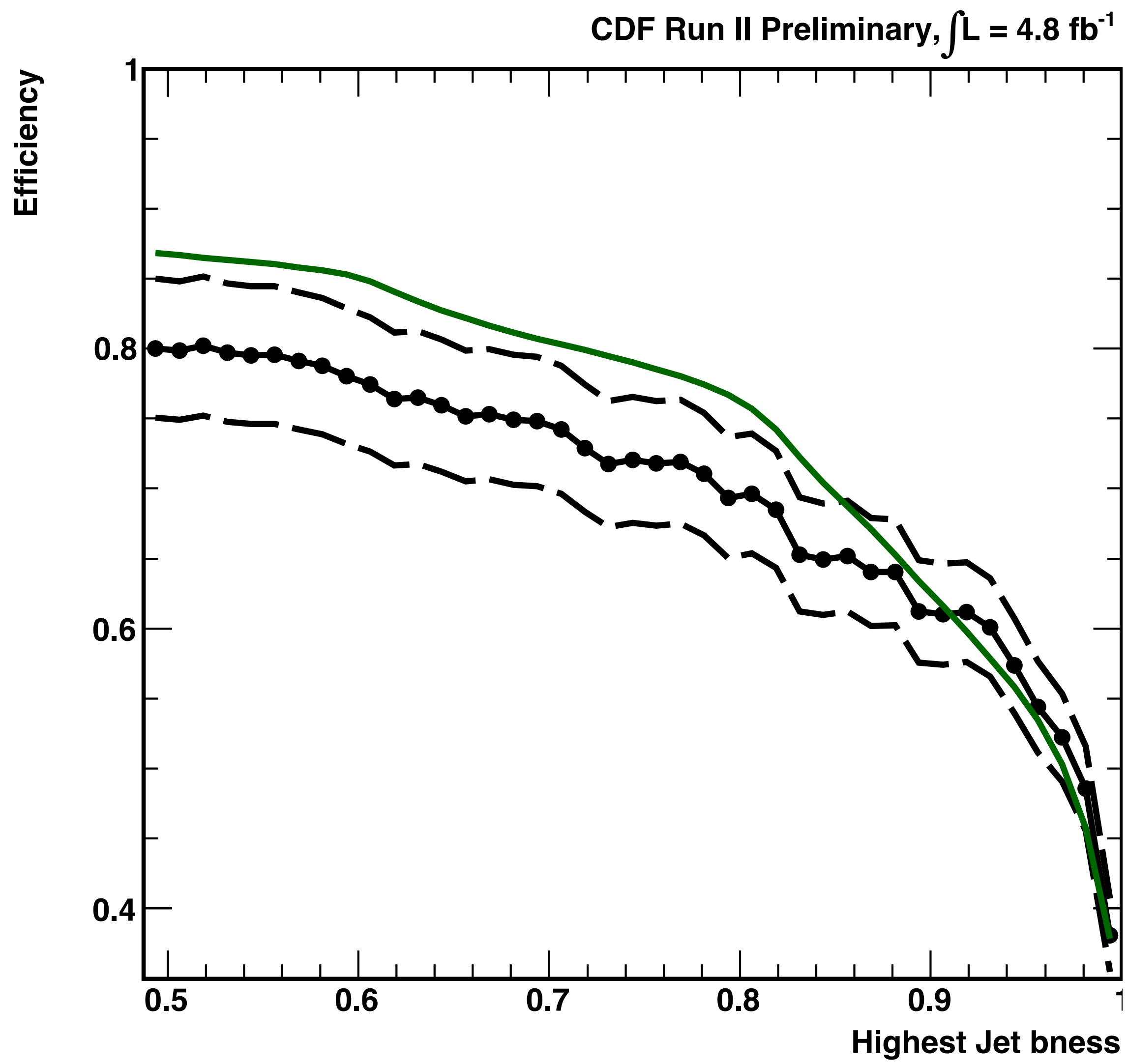


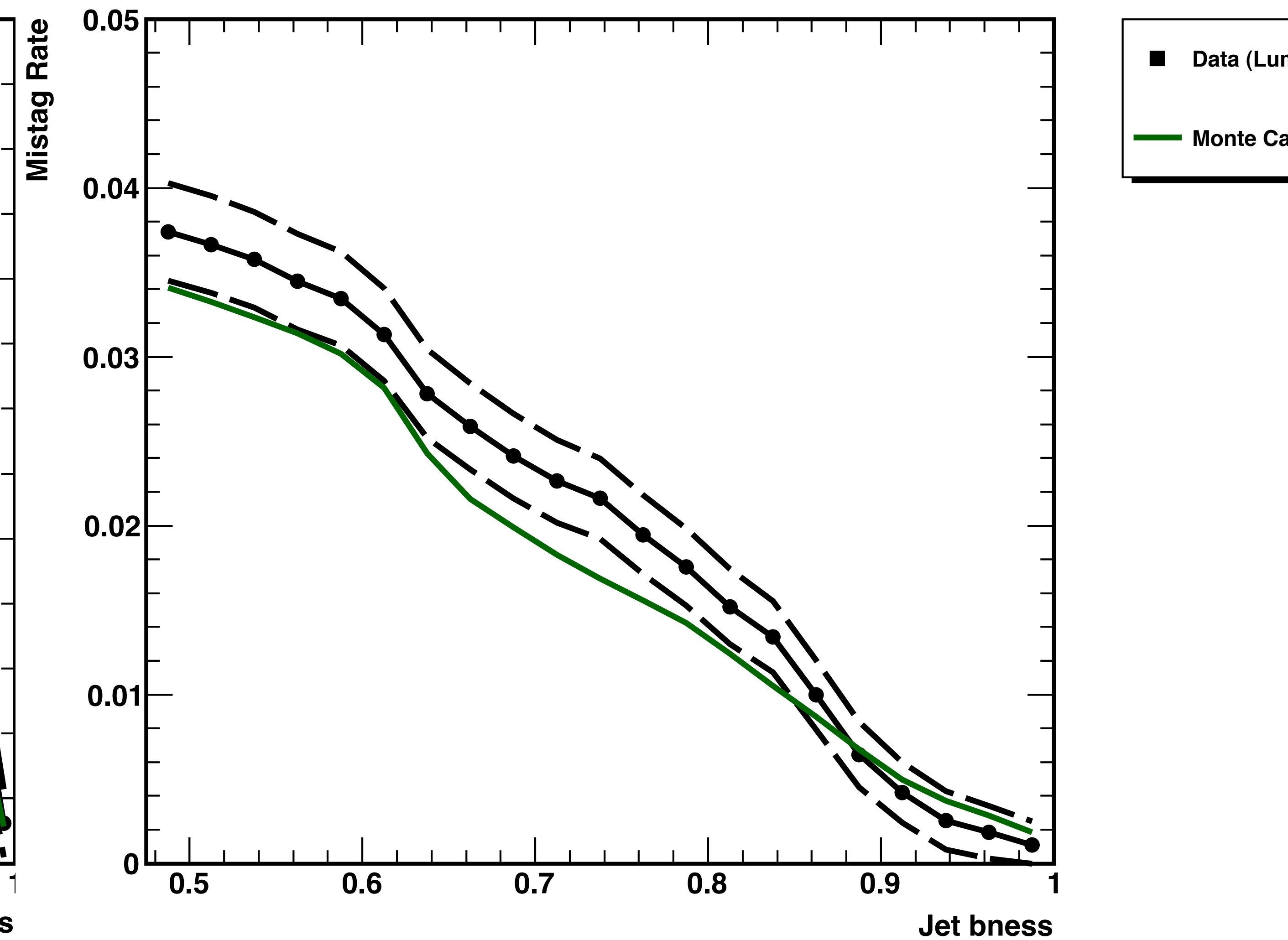
ttbar selection

- Enhanced in b
- lepton+jets ttbar with bkg. suppressing cuts
- use to measure the efficiency

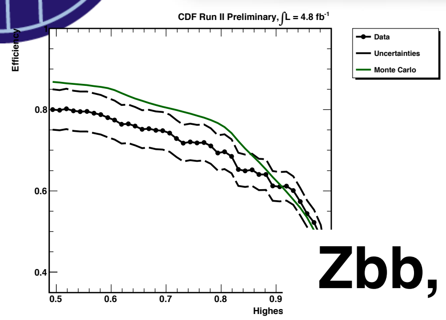


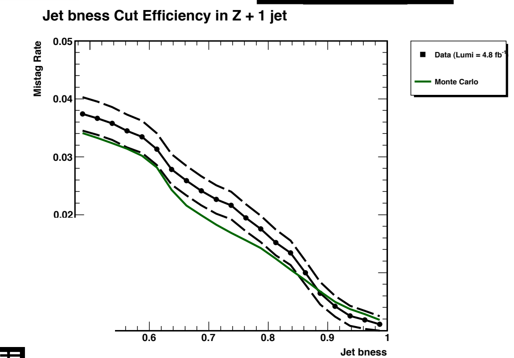
Efficiencies and mistags



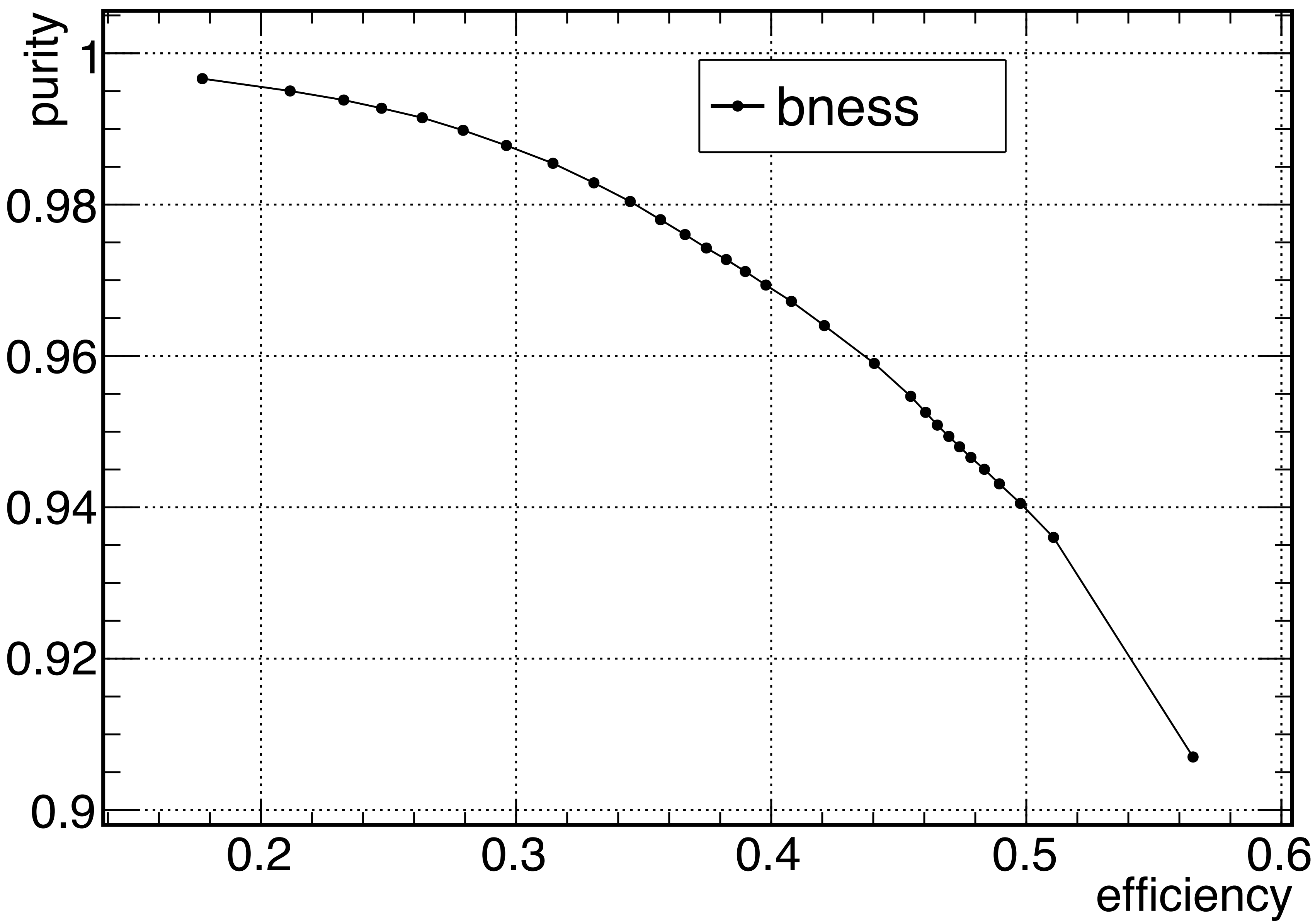


Efficiencies and mistags





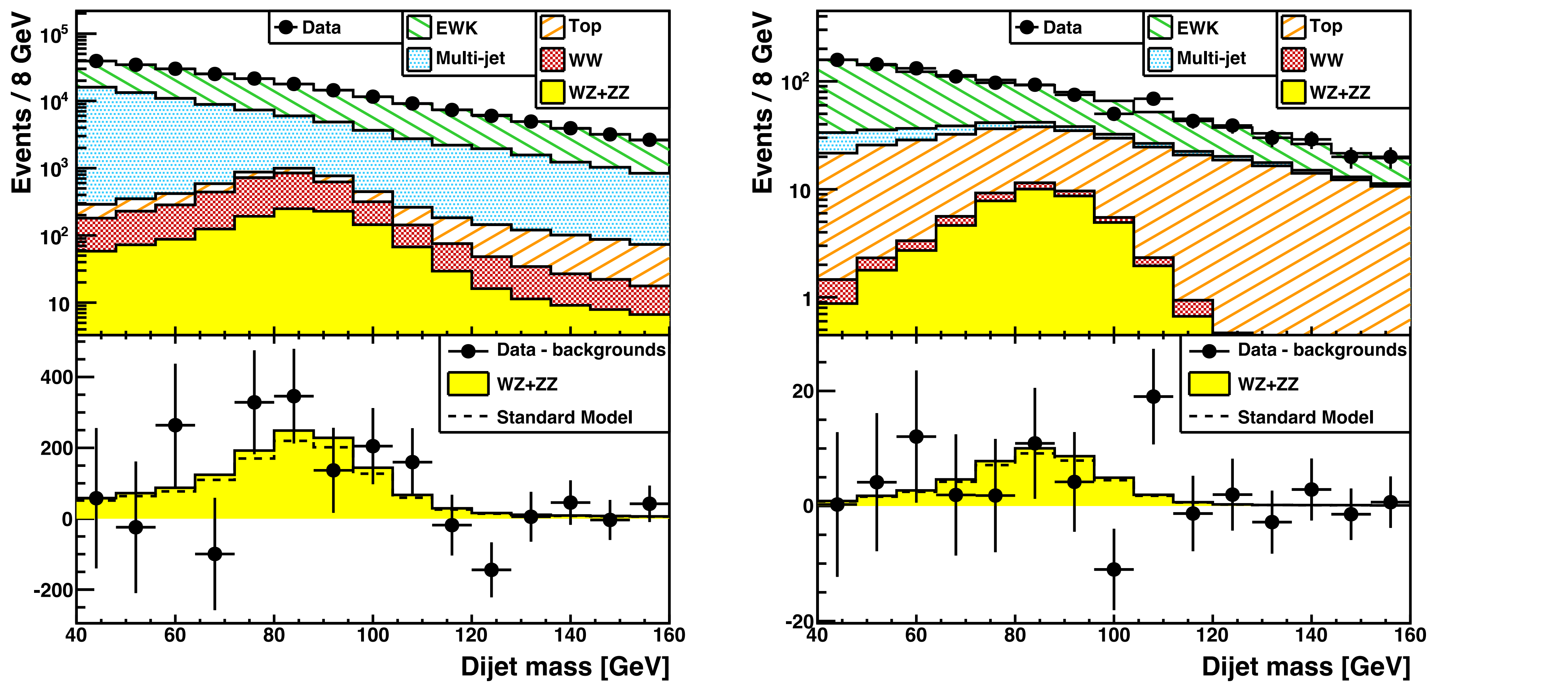
Zbb, Zqq MC



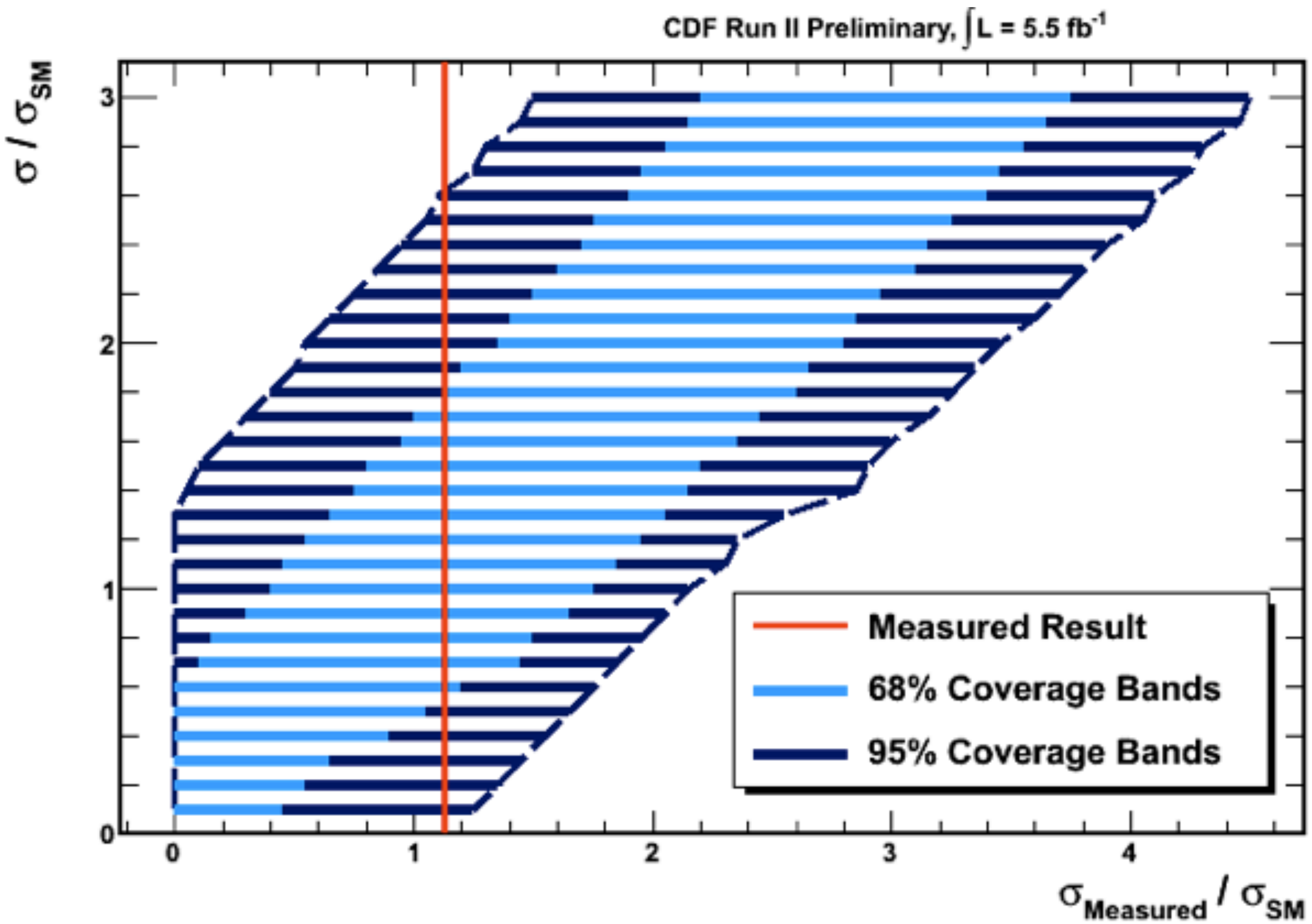
Results

- search for WZ+ZZ
  - two channels differentiated by bness
- WW constrained to theoretical cross section
- This is as close to the Higgs as it gets

Process(es)	Fit N_{events} (no-tag)	Fit N_{events} (2-tag)
EWK	$149100\ ^{+5700}_{-5200}$	$687\ ^{+52}_{-51}$
$t\bar{t}$ and single t	1700 ± 150	$312\ ^{+36}_{-42}$
Multi-jet	$76600\ ^{+5000}_{-5400}$	59.4 ± 8.0
WW	2700 ± 210	$8.3\ ^{+2.0}_{-2.2}$
WZ/ZZ	$1320\ ^{+790}_{-780}$	45 ± 26



A measurement of significance

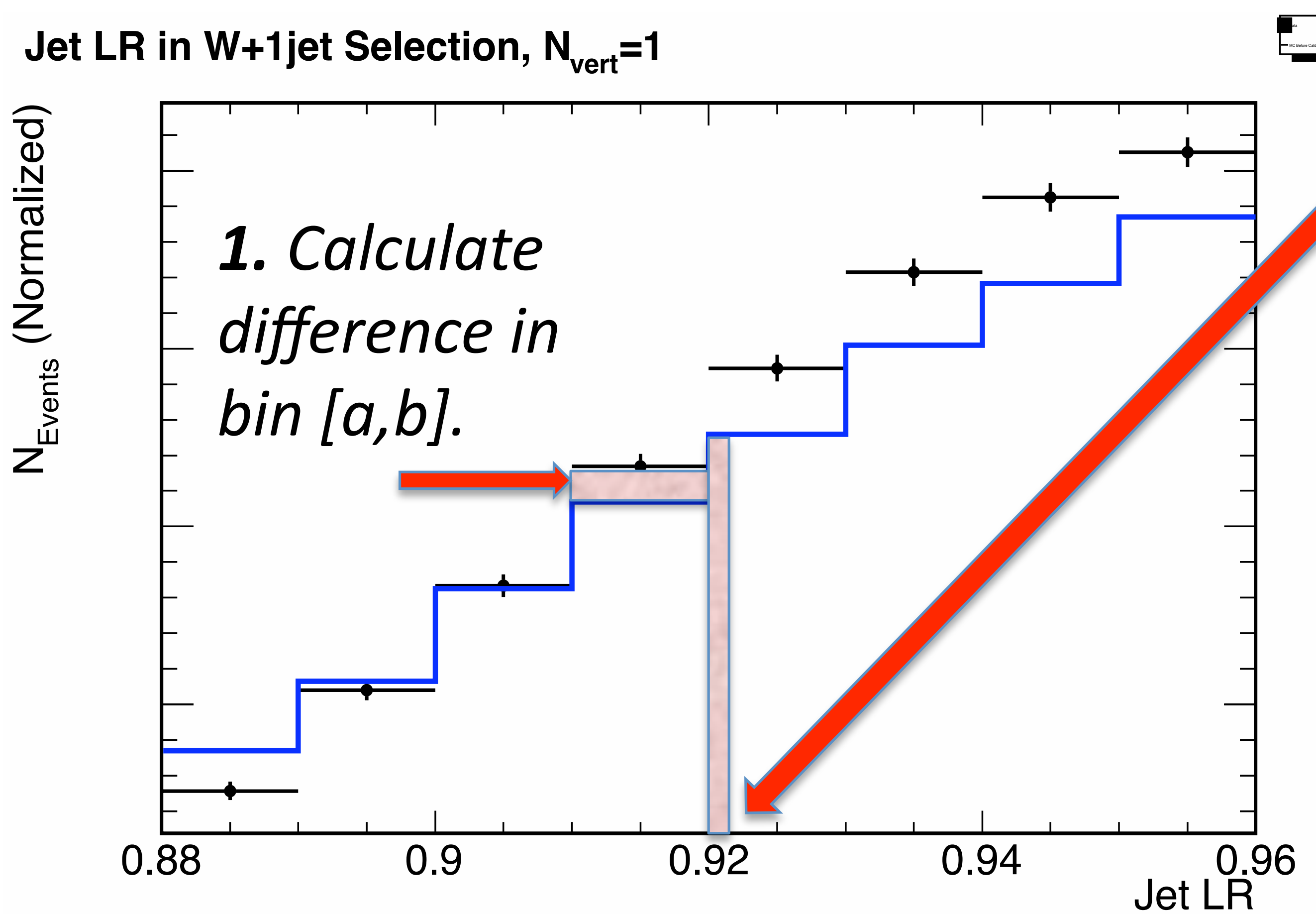


- Almost 2σ measurement

Conclusions

- To do any spectroscopy we need to prove first that we can identify the standard candles
- This is highly nontrivial because of the large backgrounds
- Any tools that we developed have to be benchmarked on them
- Dibosons in hadronic final states have been unambiguously demonstrated at CDF

- **Rescale LR in MC to match data in discrete bins**



1. *Calculate difference in bin [a,b].*

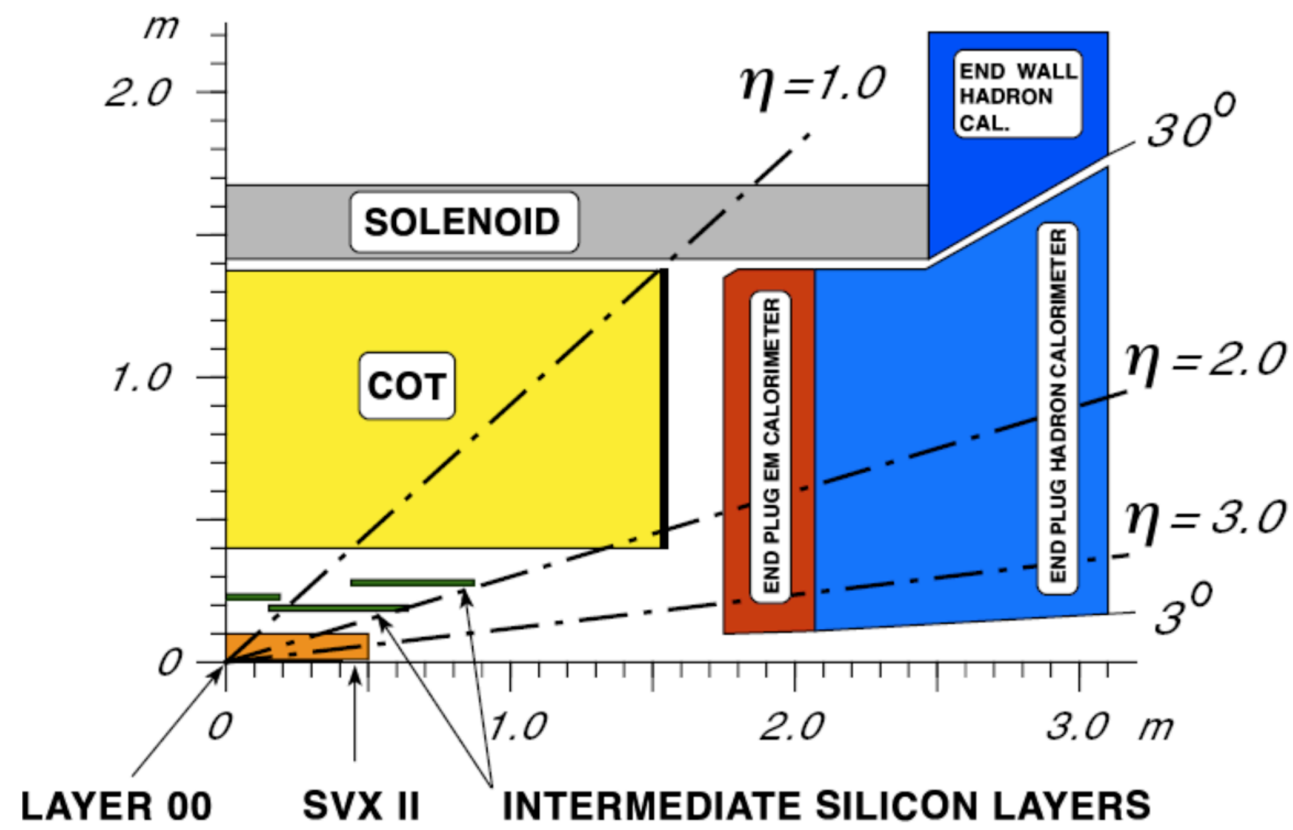
2. *Calculate bound in next bin to match that area (**b'**)*

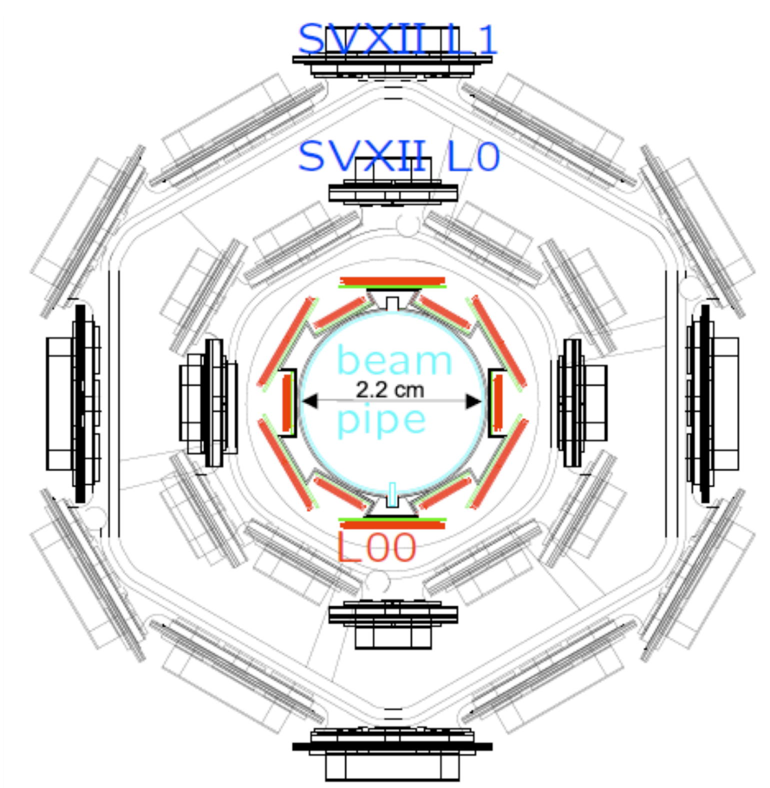
3. *Rescale value of LR in MC*

$$LR \to a + (LR - a) \times \frac{(b - a)}{(b' - a)}.$$

CDF tracking

- COT: 96 layers drift chamber
  - 4 axial 4 stereo superlayers
  - 1.4 T field
- Silicon tracker:
  - 95 cm long L00
  - 3 SVX II layers each 29cm long

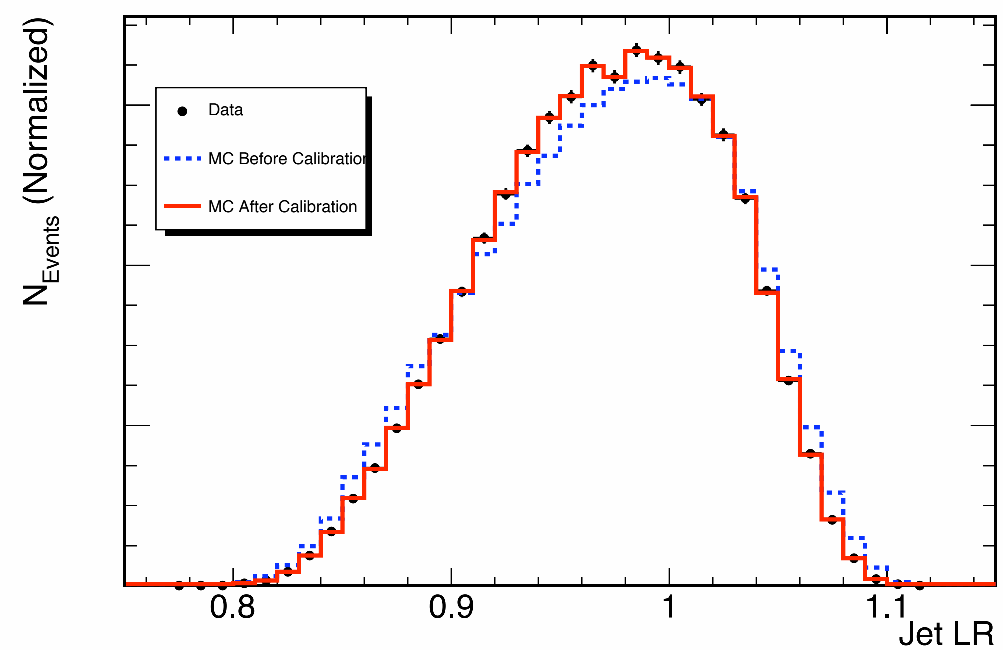




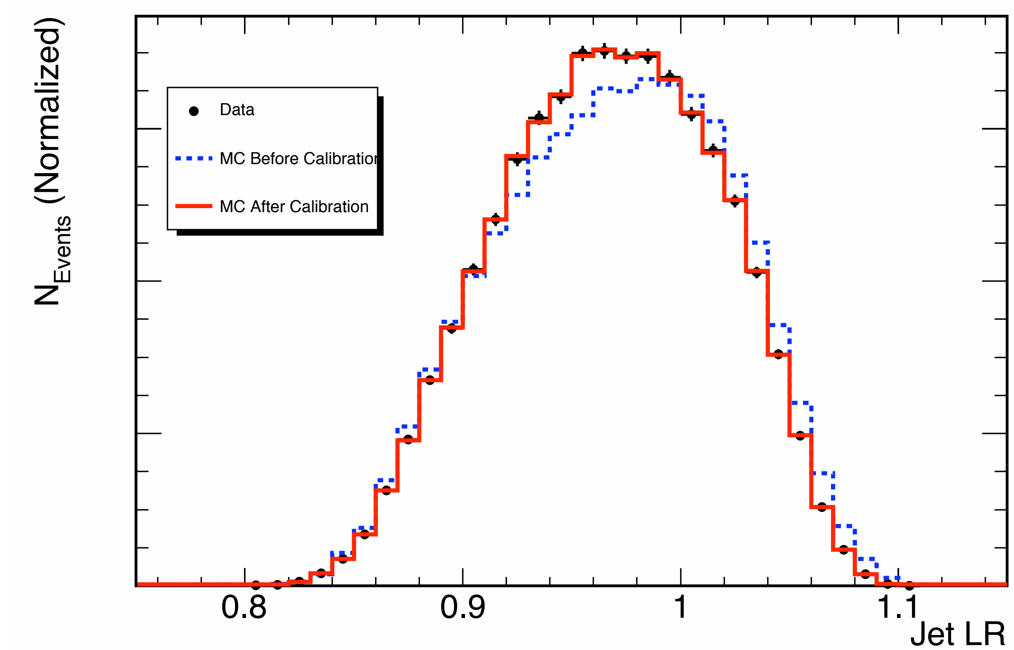
- L00 - single sided r=1.2cm
- SVX II 5 double sided layers
  - L0 rφ-rz, r=2.5cm
  - L1 rφ-rz, r=4.5cm
  - L2 rφ-1.2° stereo, r=6.5cm
  - L3 rφ-rz, r=8.5cm
  - L4 rφ-1.2° stereo, r=10.5cm
- Total sensor area 6m
- 720k electronic channels

...and vs number of vertices

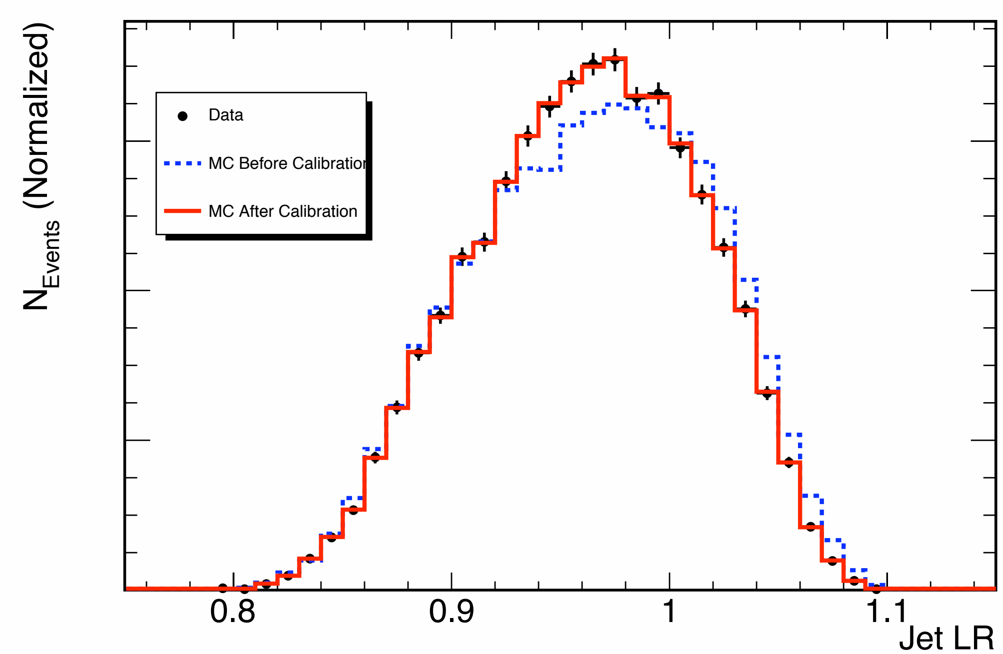
Jet LR in W+1jet Selection, N_{vert}=1



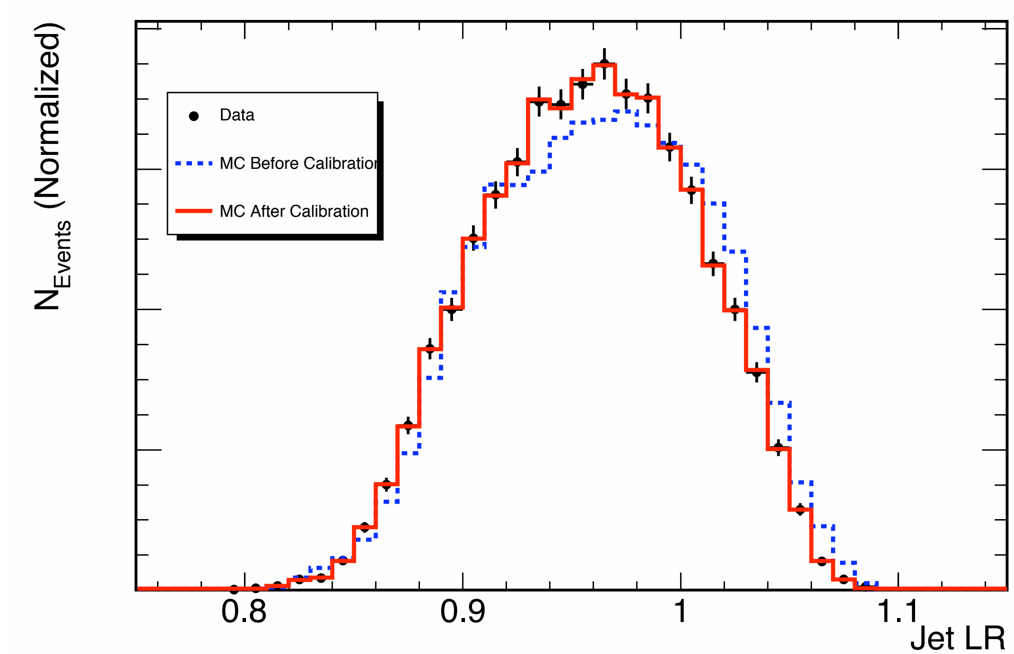
Jet LR in W+1jet Selection, N_{vert}=2



Jet LR in W+1jet Selection, N_{vert}=3

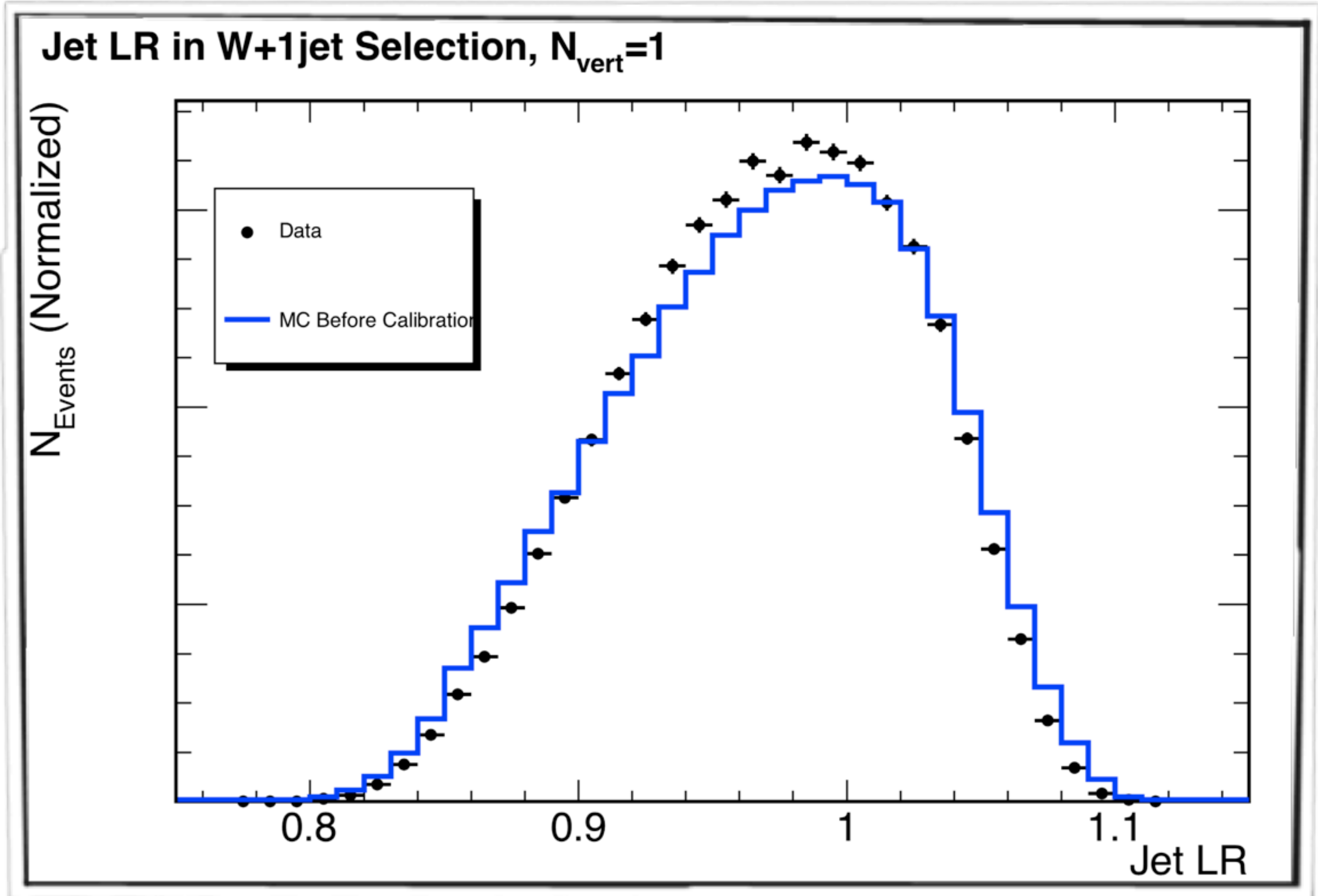


Jet LR in W+1jet Selection, $N_{vert} \geq 4$



How about data?

- Use W+jets because large stats
  - also, similar to what we are in the end looking for
- The shape looks quite good given the level of accuracy the simulation has to get right
- Still data systematically higher (more quark like)
- Strong assumption: the quark and gluon ratio is correctly described in MC

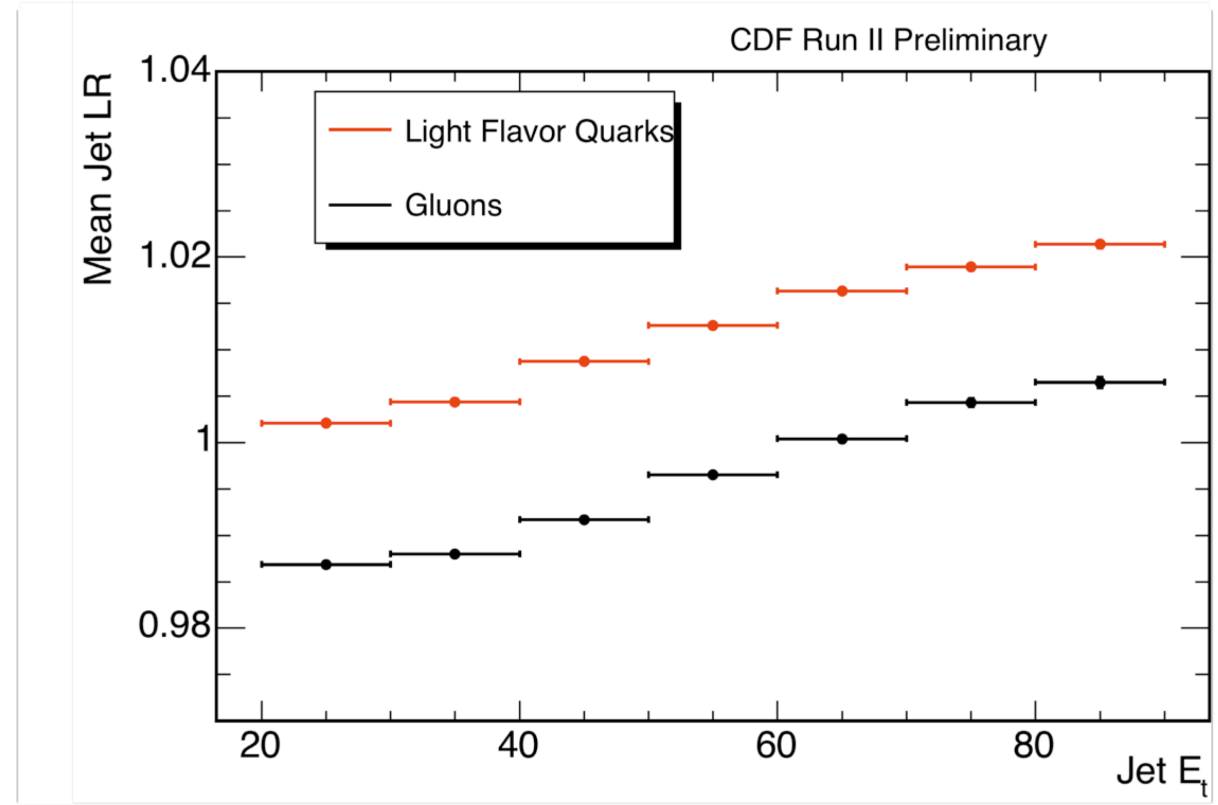


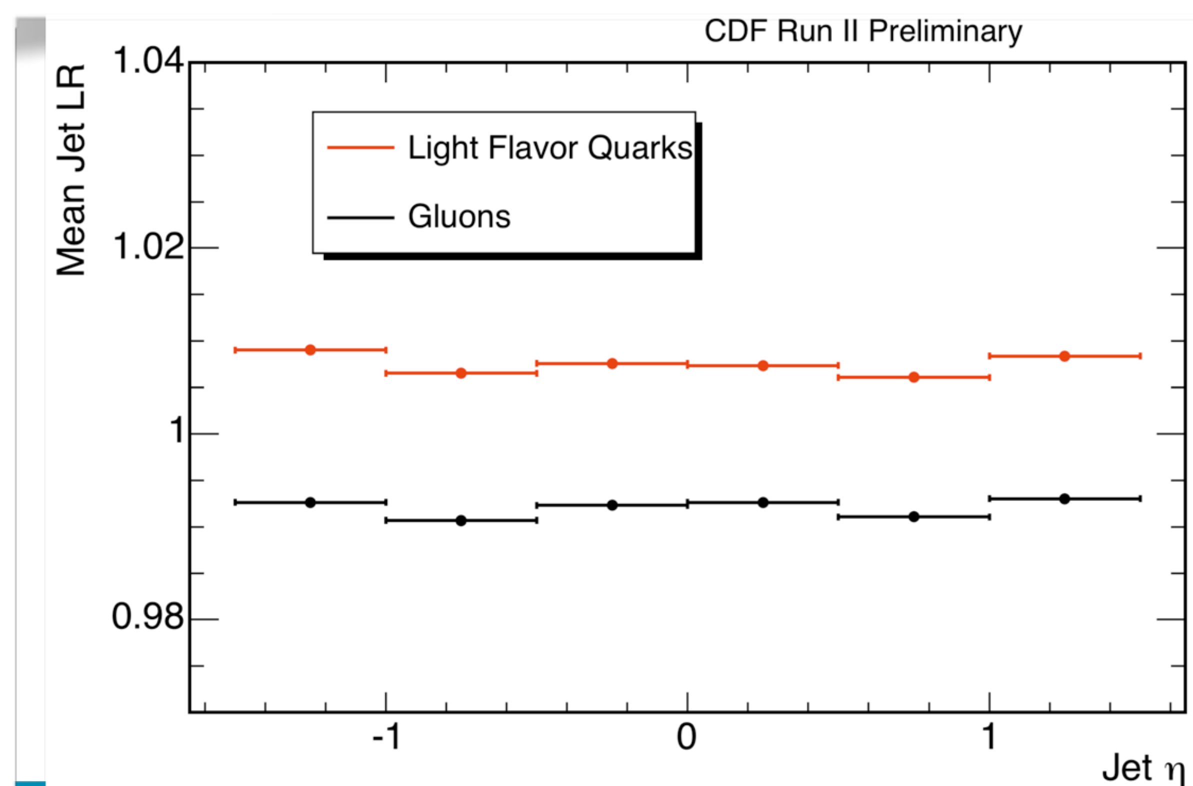
Assume all differences are due to calorimeter simulation

Calibrate based on the W+jets data

Correlations
(Is It Really True?)

- **Jet LR and E_T show strong correlation**
  - Expect jets with greater ET to be more collimated → look mre like quarks → have higher LR
  - Still see discrimination in similar ET jets
- **Jet LR looks uncorrelated with jet η**
  - Small differences possibly from different calorimeter structure inputs





Other taggers

- Soft Lepton Tagger
- Count displaced tracks
  - 3 2 sigma or 2 3 sigma
  - signed IPs against jet direction
- Jet probability
  - Joint probability for all tracks to come from a primary vertex
  - Track resolution derived from negative IP tracks
  - Use only positive IP tracks in probability calculation
- Multivariate taggers
  - CDF NN tagger
  - D0 NN tagger

Secondary vertex tagger

- Identify secondary vertices
  - Impact parameter significant
  - Fit displaced tracks (>Pt) to a common vertex
  - Prune tracks and cut on fit chi2
  - Cut on Lxy significance