ttbar 500GeV benchmarking

Katsumasa Ikematsu (KEK)

ILD Meeting (12/Sep/'08 @Cambridge)

A benchmark process: e^+e^- -> ttbar

- The common LoI benchmarking
  - …The evaluation of the detector performance should be based on physics benchmarks, some of which will be the same for all LOIs based upon an agreed upon list and some which may be chosen to emphasize the particular strengths of the proposed detector… (from “Guideline for the definition of a Letter of Intent …”)
  - WWS Software panel in consultation with the detector concepts and the WWS Roadmap panel and starting from the Benchmark Panel Report has drafted a short list of processes
- e^+e^- -> ttbar: one of the common LoI benchmarking processes
  - CMS: 500GeV
  - Observables: σ_{tt}, m_{top}, A_{FB}
  - Decay mode: 6-jets mode (t->bW, W->qq’)
  - Comments from WWS: Test b-tagging and PFA in multi-jet events (m_{top}=175GeV)

Observables & analysis strategy

- σ_{tt} measurement
  - counting signal events => Based on hemisphere analysis (not necessary to full reconstruction of ttbar events)
  - Need realistic BG study w/ full SM background samples
- m_{top} measurement @500GeV (open top region)
  - 3-jets mass: top and anti-top full reconstruction w/ correct jet association => Double b-tagging is powerful tool

Current target!

- A_{FB} measurement w/ 6-jets mode
  - Need to know charge of a jet (at least one jet) => Vertex charge information (to reconstruct the vertex charge it is necessary to find all stable tracks from the B-decay chain: charged B-hadrons ~ 40% of the b-jets) Should be next step!

Common input: SLAC SM data samples

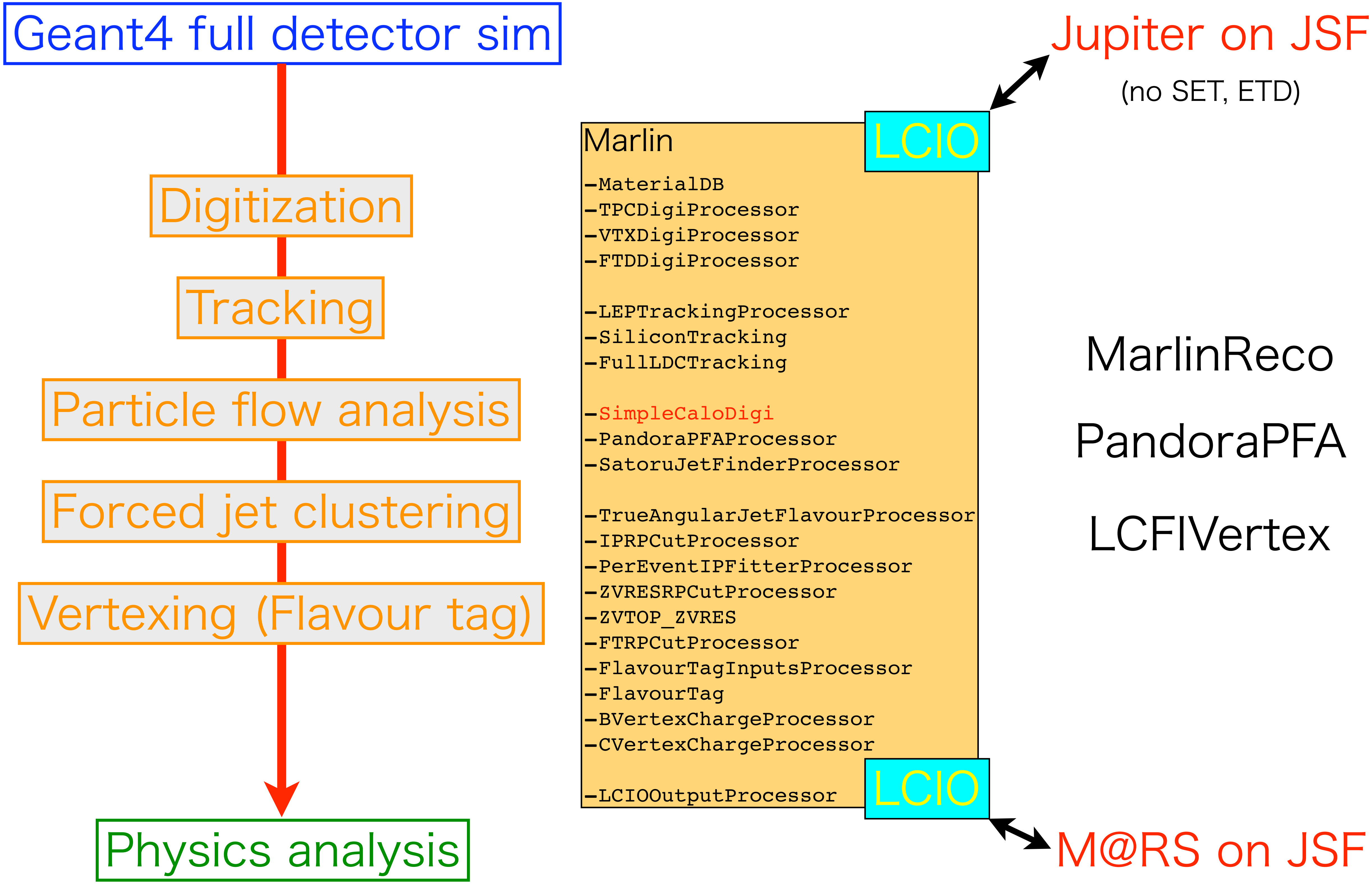
- All analyses used in the context of the detector optimization and LoI process need an inclusive sample of the SM Background
  - SLAC team produced a 2nd generation of a complete SM sample for 500fb^{-1} at all 4 polarization combinations (The WHIZARD Monte Carlo version 1.40 is used for parton generation)
  - Whizard input controls the properties of the colliding electron/positron beams. the spectrum (corresponds to the Guinea-Pig data) represents the default ILC design for E_{cm}=500GeV circa Aug. 2005, and includes both incoming LINAC energy spread and beamstrahlung
  - For each event in a derived StdHep file the variable IDRUPLH from the common block HEPEV4 is used to identify the process
  - PYTHIA 6.205 is used for final state QED/QCD parton showering and for the fragmentation of quarks and gluons. Parton showering is performed for all final state fermions with the exception of electrons (no top quarks appears in process_id)
- ttbar->6-jets samples are included in 6-fermion SM samples

SLAC SM data samples (cont’d)

- 6-fermion SM samples also contains no ttbar mediated events: (e^+e^- -> bb with γ->WW) and (e^+e^- -> WW with Z -> bb) => same signature for the ttbar events
- This work: 50fb^{-1} MC samples focusing on pre-selected bbuddu (17766) & bbcssc (17790) => different flavour components and low Xsec (limitation for CPUs & disks)

IDRUPLH	Process	Pol(e^-)	Pol(e^+)	Xsec (fb)
17765	bbuddu	-1.0	1.0	166.3
17766	bbuddu	1.0	-1.0	66.0
17769	bbudsc	-1.0	1.0	164.7
17770	bbudsc	1.0	-1.0	65.7
17785	bbcsdu	-1.0	1.0	164.7
17786	bbcsdu	1.0	-1.0	66.0
17789	bbcssc	-1.0	1.0	165.1
17790	bbcssc	1.0	-1.0	66.0

Analysis framework



Geant4 full detector sim

Digitization

Tracking

Particle flow analysis

Forced jet clustering

Vertexing (Flavour tag)

Physics analysis

Marlin

- MaterialDB
- TPCDigiProcessor
- VTXDigiProcessor
- FTDDigiProcessor

- LEPTrackingProcessor
- SiliconTracking
- FullLDCTracking

- SimpleCaloDigi
- PandoraPFAProcessor
- SatoruJetFinderProcessor

- TrueAngularJetFlavourProcessor
- IPRPCutProcessor
- PerEventIPFitterProcessor
- ZVRESRPCutProcessor
- ZVTOP_ZVRES
- FTRPCutProcessor
- FlavourTagInputsProcessor
- FlavourTag
- BVertexChargeProcessor
- CVertexChargeProcessor

- LCIOOutputProcessor

LCIO

LCIO

Jupiter on JSF

(no SET, ETD)

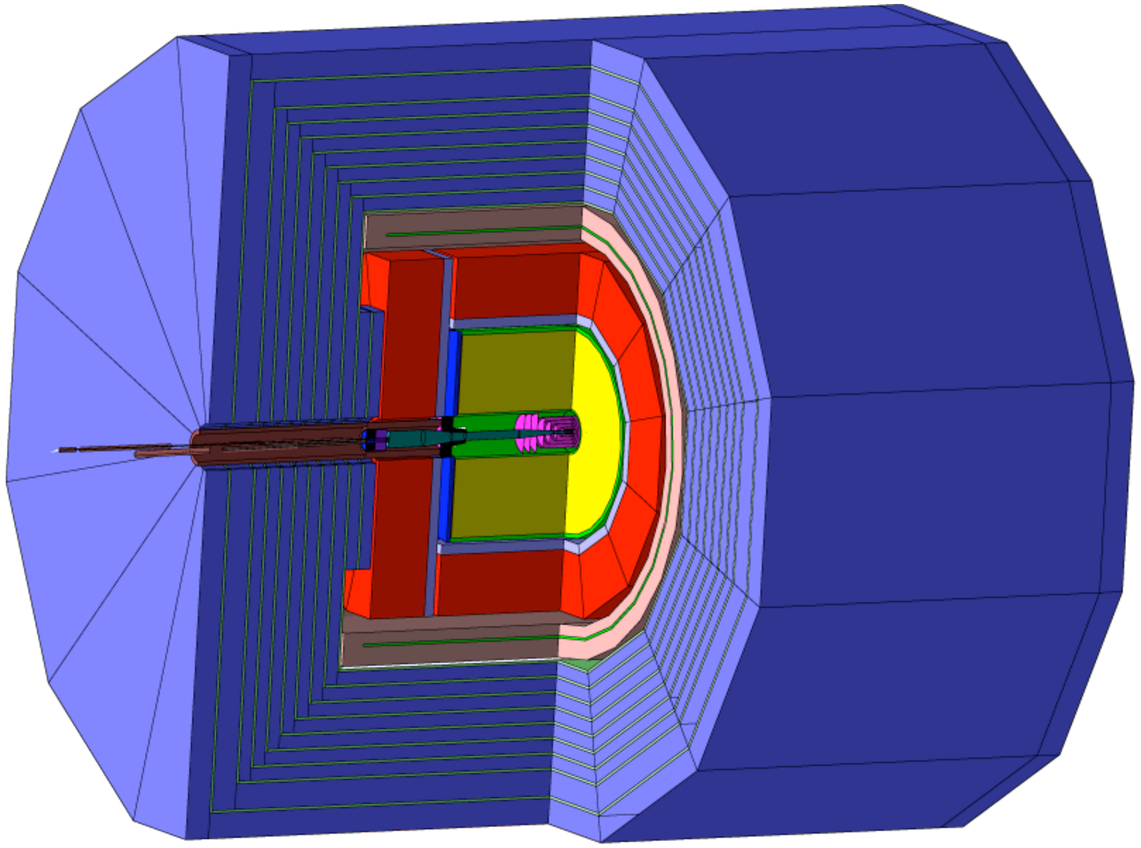
MarlinReco

PandoraPFA

LCFIVertex

M@RS on JSF

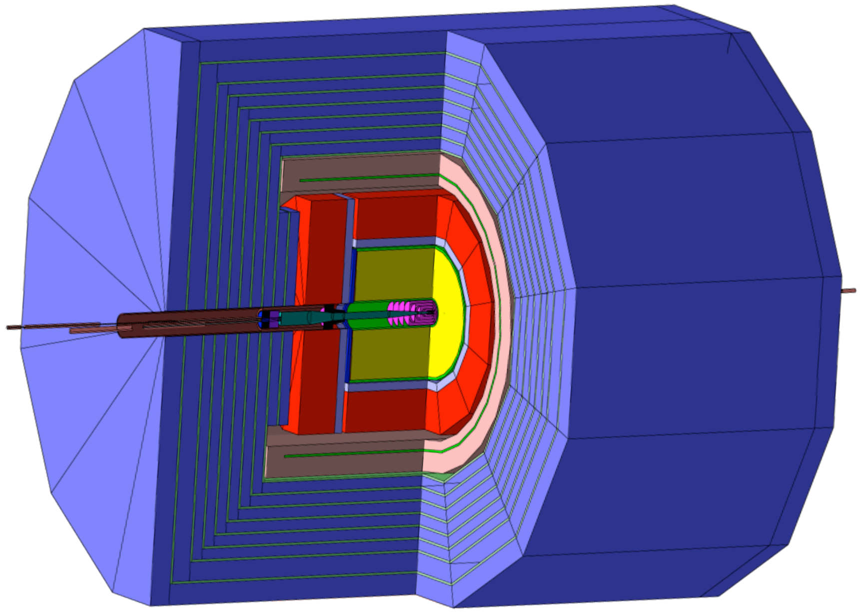
3 detector geometries: B&R



GLD

B = 3 Tesla

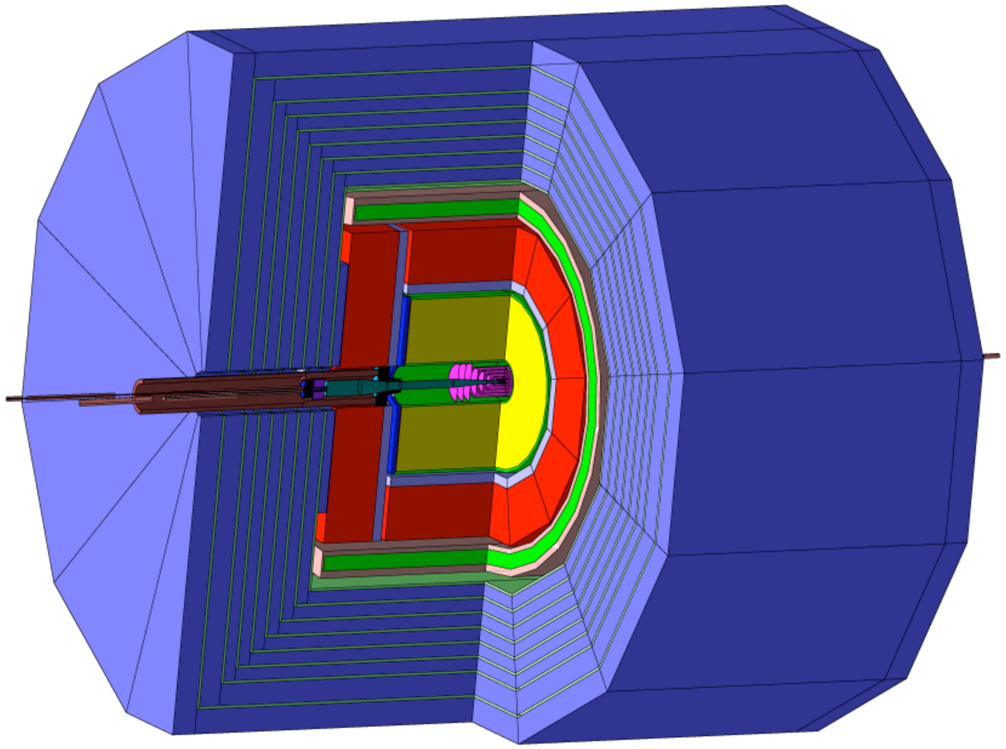
R_{min} (ECAL) = 210cm



J4LDC

B = 4 Tesla

R_{min} (ECAL) = 160cm



GLDPrim

B = 3.5 Tesla

R_{min} (ECAL) = 185cm

Run condition for Marlin processors (1)

- ilcsoft v01-04: same as LDC-DST mass production softwares
- Steering file: almost same as LDC-DST mass production steering file
  - `http://www-flc.desy.de/simulation/databasereco/recosteer.html`
  - Rec01-04_Slac_SM_LDCPrime_02Sc_ppr002 reconstruction (`stdreco_IN_Rec01-04.xml`)
- Difference between GLD and LDC side
  - VTXDigiProcessor

    no entries for SET in GLD

```
<!--R-Phi Resolution in SIT-->
<parameter name="PointResolutionRPhi_SIT" type="float">0.01 </parameter>
<!--Z Resolution in SIT-->
<parameter name="PointResolutionZ_SIT" type="float">0.01 </parameter>
```

  - VTXDigiProcessor

```
<!--Point Resolution in FTD-->
<parameter name="PointResolution" type="float">0.01 </parameter>
```

  - LEPTrackingProcessor

```
<parameter name="AlwaysRunCurlKiller" type="int">1</parameter>
<!--Cut for the number of hits allowed in one bin-->
<parameter name="MultiplicityCut" type="int">4 </parameter>
<parameter name="BinWidth" type="int">2</parameter>
<parameter name="BinHeight" type="int">2</parameter>
```

  - FullLDCTrackingProcessor

    delete SET & ETD related entries

Run condition for Marlin processors (2)

- Main difference between GLD and LDC comes from CaloDigiProcessors
  - LDC: LDCCaloDigi + SimpleMuonDigi + SimpleLCalDigi
  - GLD: SimpleCaloDigi

	ECALLayers	CalibrECAL	HCALLayers	CalibrHCAL
GLDapr08	33	24.56	46	27.44
GLDprim	33	24.33	42	27.09
J4LDC	33	23.954	37	28.196

  - PandoraPFAProcessor

    Tuned calibration factors for 3 different detector geometries (ECALMIPcalibration, HCALMIPcalibration, ECALThreshold, HCALThreshold, ECALEMMIPToGeV, ECALHadMIPToGeV, HCALEMMIPToGeV, HCALHadMIPToGeV)

```
<parameter name="ClusterFormationPadsECAL" type="float">2.5 </parameter>
<parameter name="SameLayerPadCutECAL" type="float">2.8 </parameter>
```

  - SatoruJetFinder (Only force 6-jets)

```
<parameter name="Mode" type="string" value="DurhamNJet"/>
<parameter name="NJetRequested" type="int" value="6"/>
```

  - Processors for LCFIVertex

    follow LDC-DST steering

ttbar -> 6jets reconstruction

I) Force 6-jets clustering

II) Confirm Max_cosθ_{jet} should be less than 0.99

III) Choose all the 15-possible pairs out of 6-jets => W_1 candidate

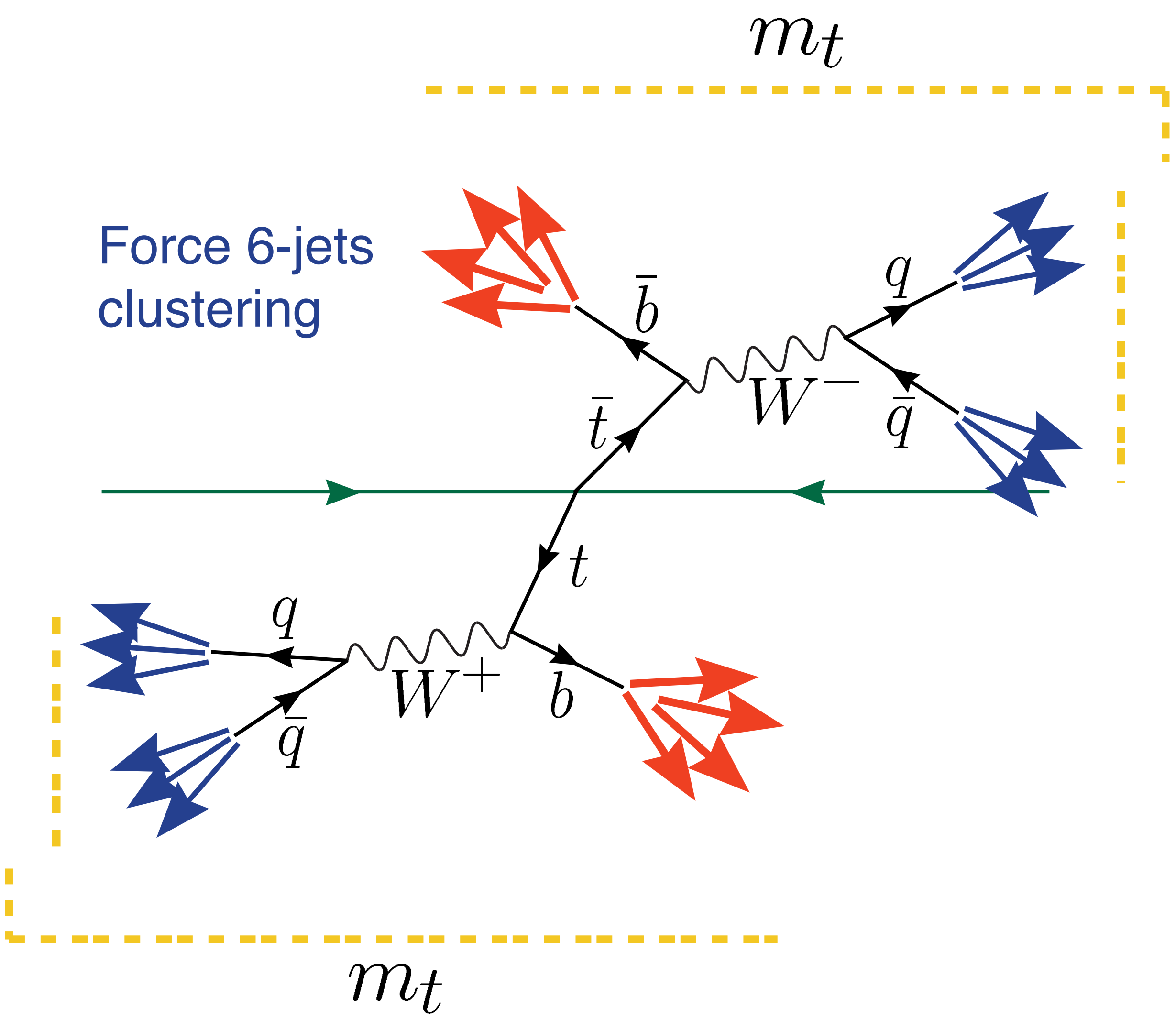
IV) Choose all the 6-possible paris out of remaining 4-jets => W_2 candidate

V) Remaining 2-jets should be b-jets: flavor tagging (charm/bottom tagging) is very important to eliminate both combinatorial and process BGs

VI) There are 2 possibilities to attach a b-jet to W_1 and W_2 candidates

VII) Store all solutions w/ $\chi^2 = (m_{w1} - m_w)^2/\sigma^2_{mw} + (m_{w2} - m_w)^2/\sigma^2_{mw} + (m_{t1} - m_t)^2/\sigma^2_{mt} + (m_{t2} - m_t)^2/\sigma^2_{mt}$

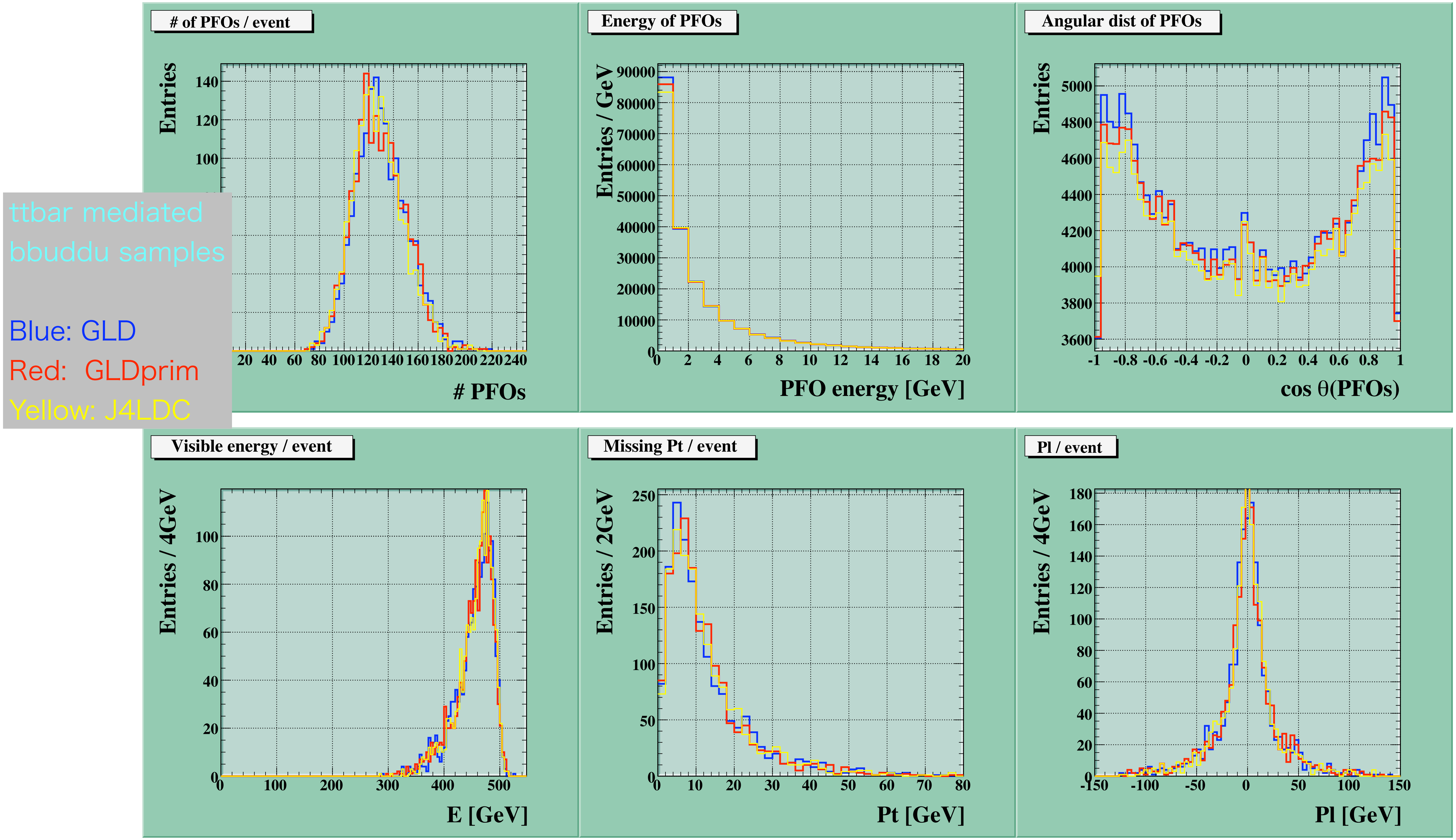
VIII) Sort solutions according to χ^2: choose the best solution



Force 6-jets clustering

Distribution w/ particle flow objects

- Before force 6-jets clustering is performed



ttbar mediated bbuddu samples

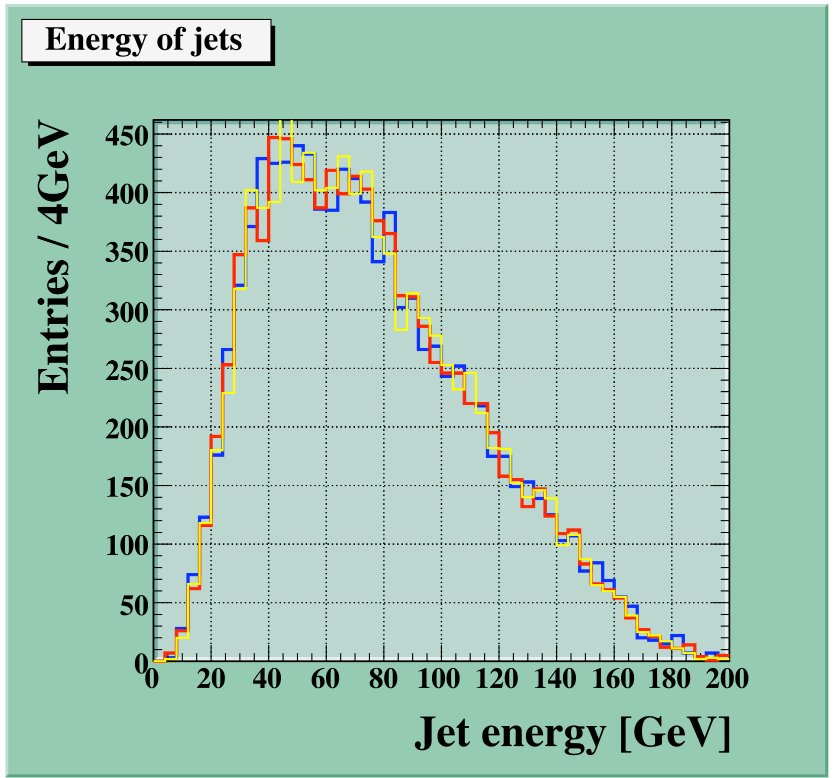
Blue: GLD
Red: GLDprim
Yellow: J4LDC

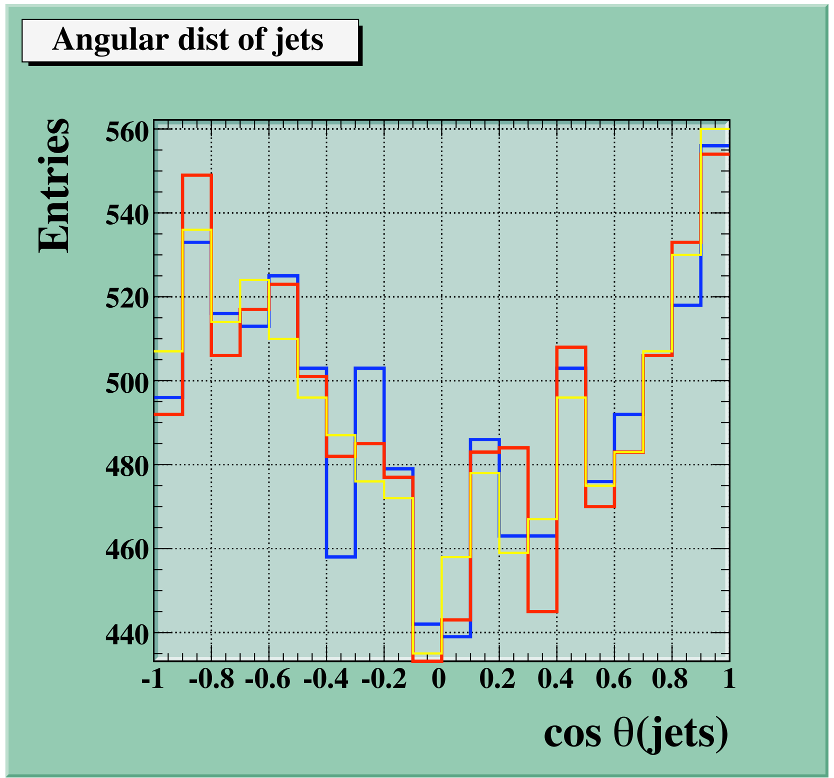
Distribution for jets

- After force 6-jets clustering is performed

ttbar mediated
bbuddu samples

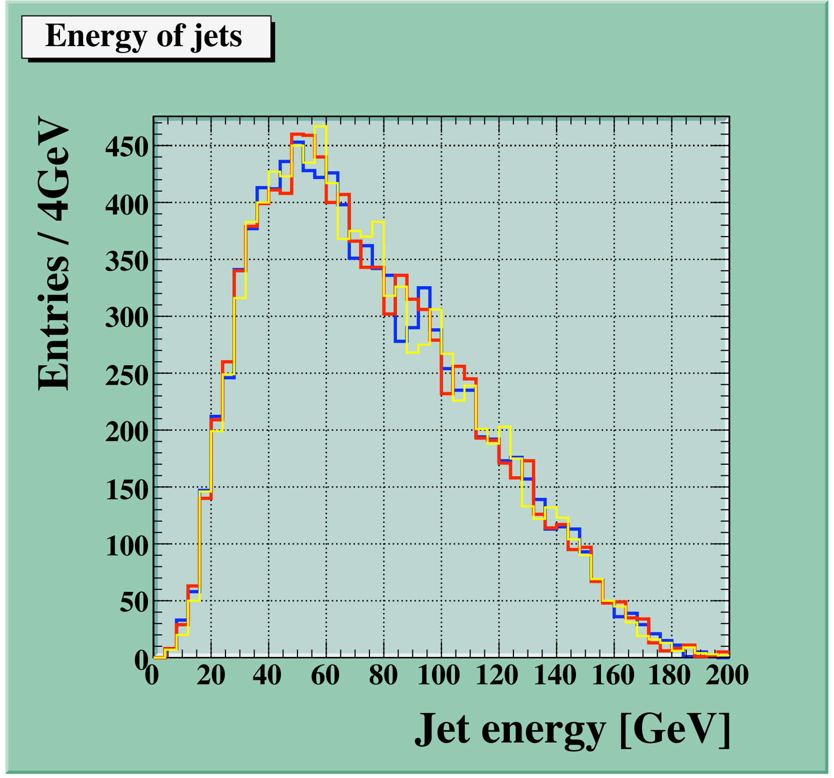
Blue: GLD
Red: GLDprim
Yellow: J4LDC

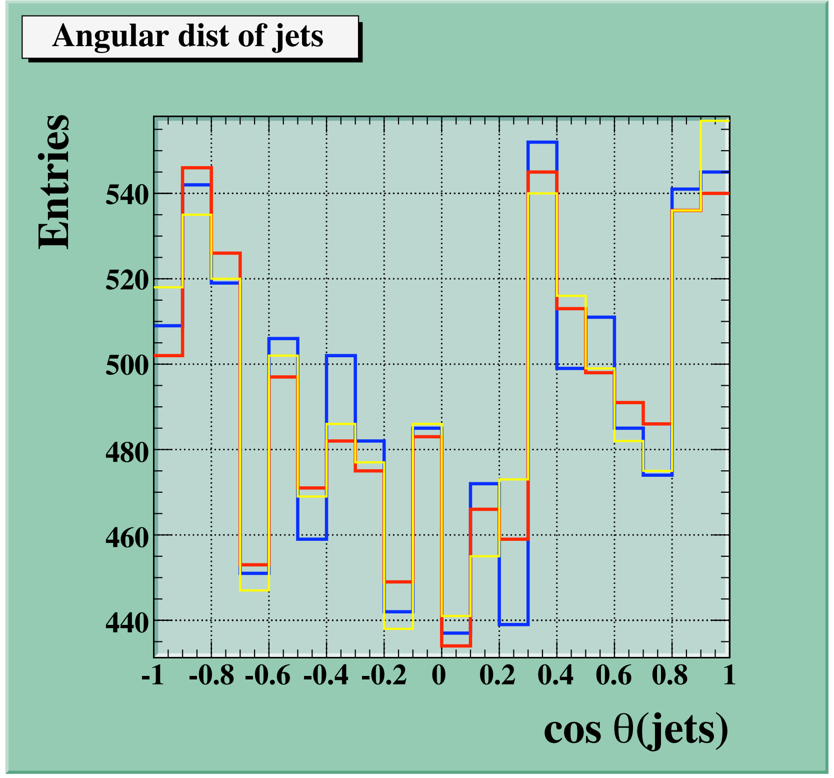




ttbar mediated
bbcssc samples

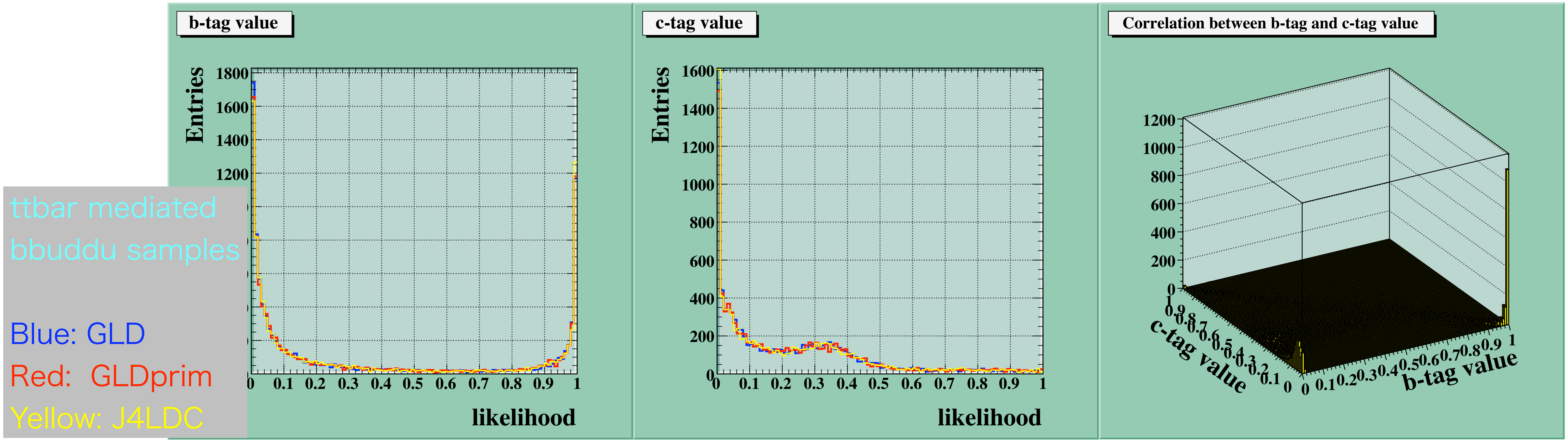
Blue: GLD
Red: GLDprim
Yellow: J4LDC





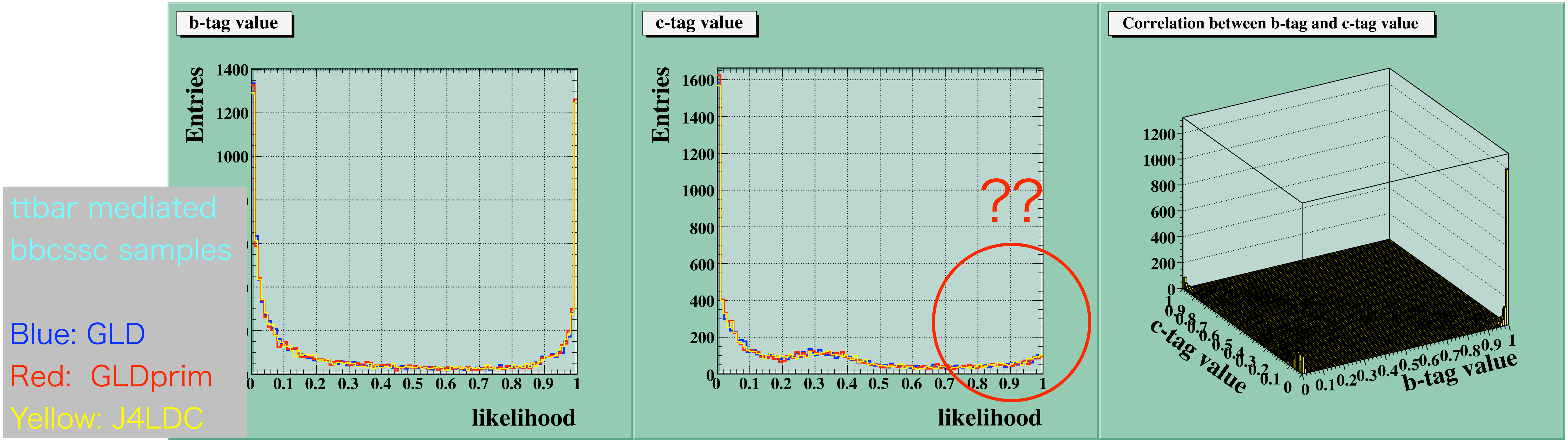
b-tag and c-tag value (LCFIVertex)

- After force 6-jets clustering is performed



ttbar mediated
bbuddu samples

Blue: GLD
Red: GLDprim
Yellow: J4LDC



ttbar mediated
bbcssc samples

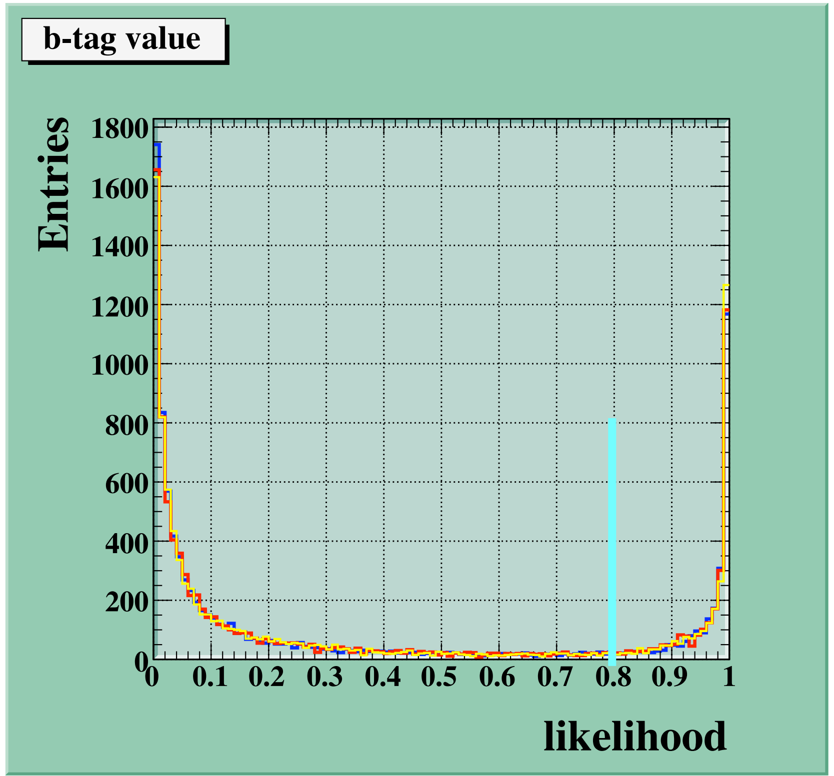
Blue: GLD
Red: GLDprim
Yellow: J4LDC

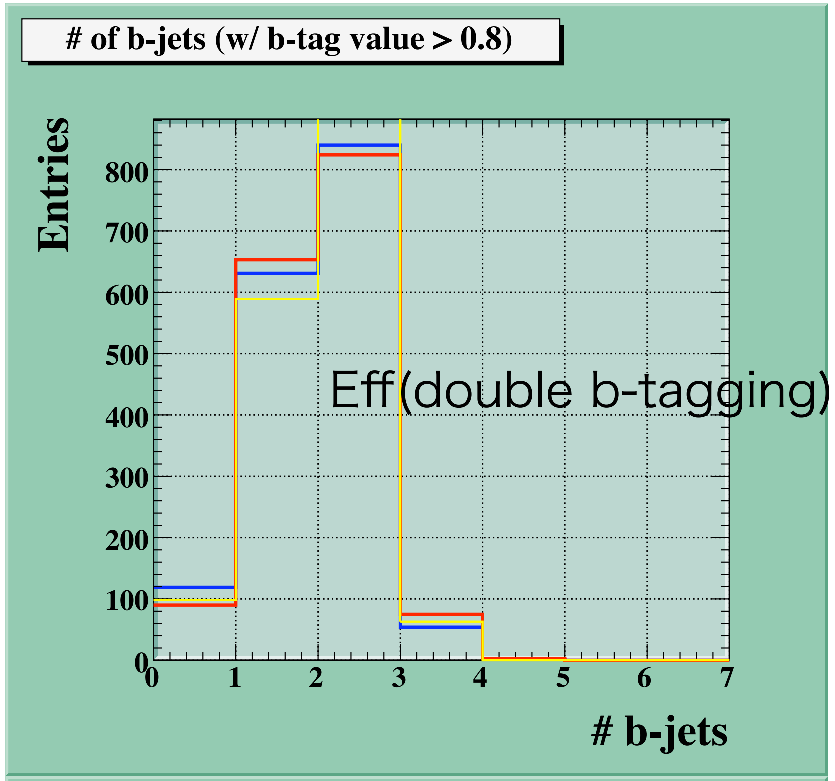
b-tagging using b-tag value

- Double b-tagging efficiency: worse than expected

ttbar mediated
bbuddu samples

Blue: GLD
Red: GLDprim
Yellow: J4LDC

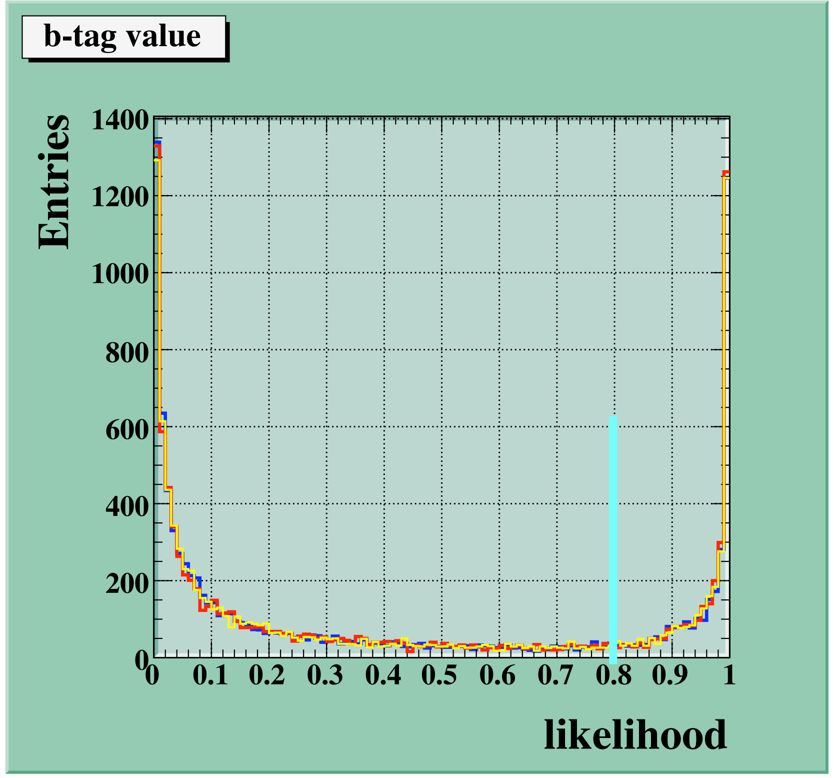


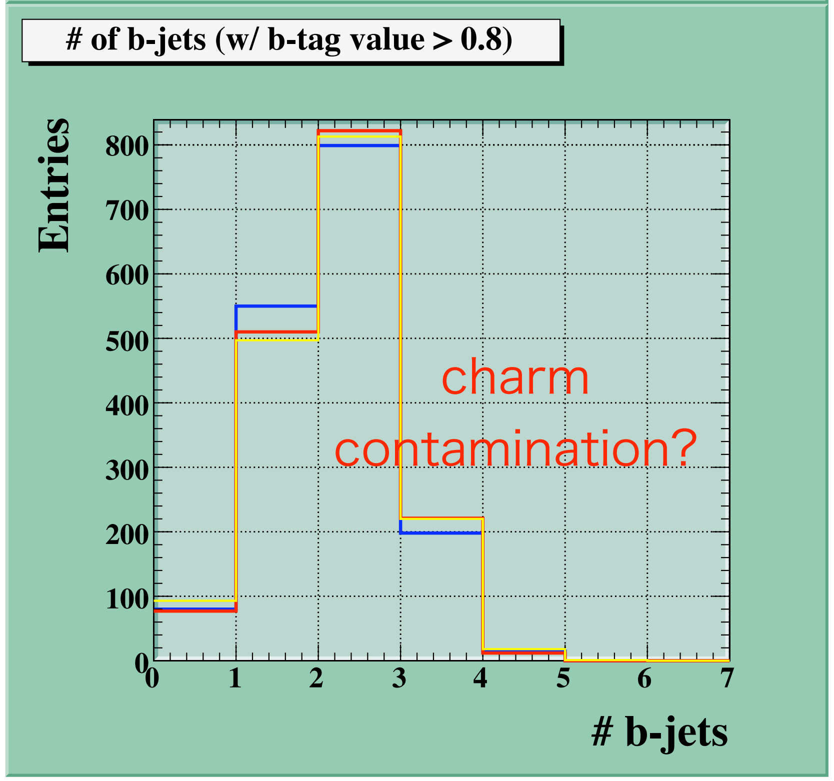


Eff(double b-tagging) ~ 50%

ttbar mediated
bbcssc samples

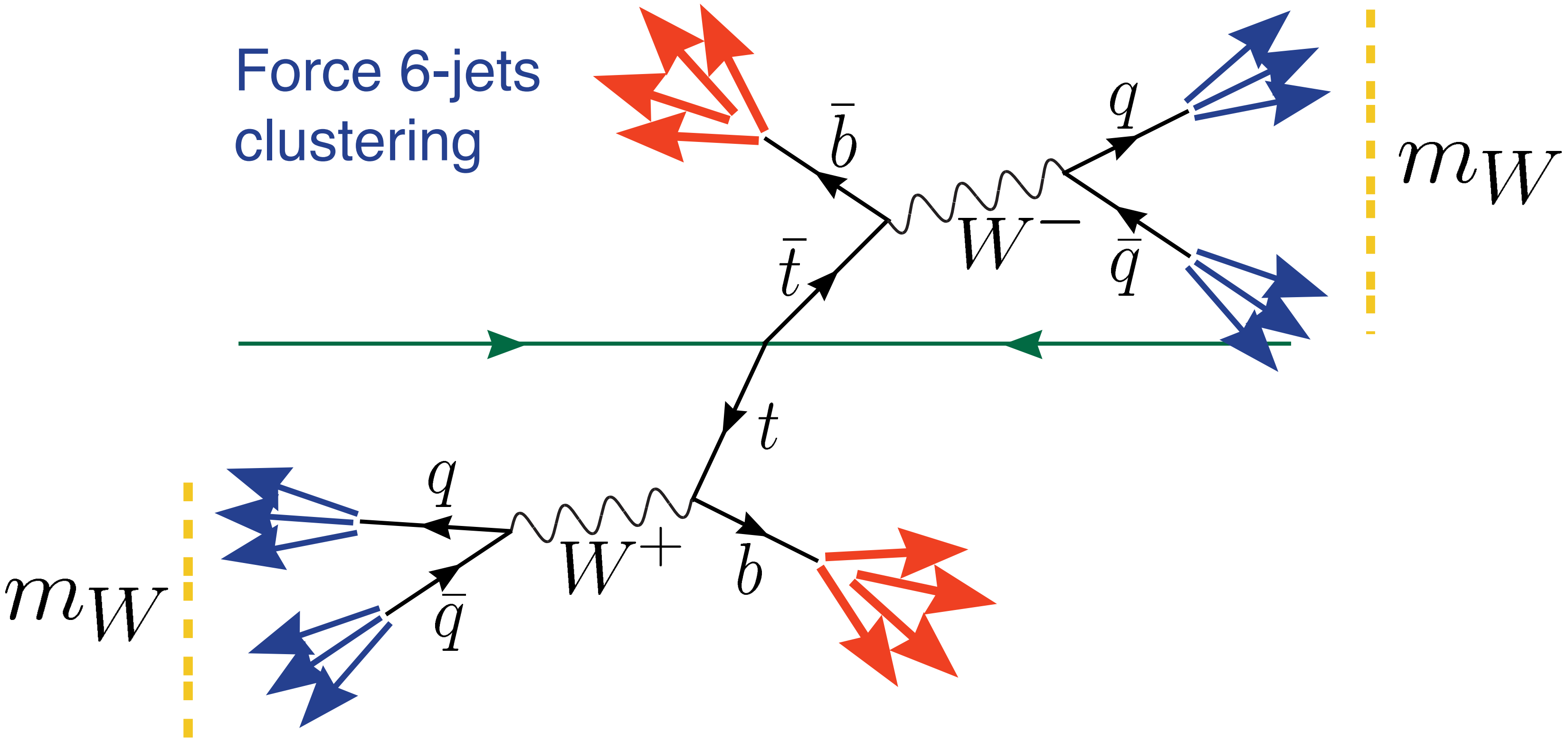
Blue: GLD
Red: GLDprim
Yellow: J4LDC

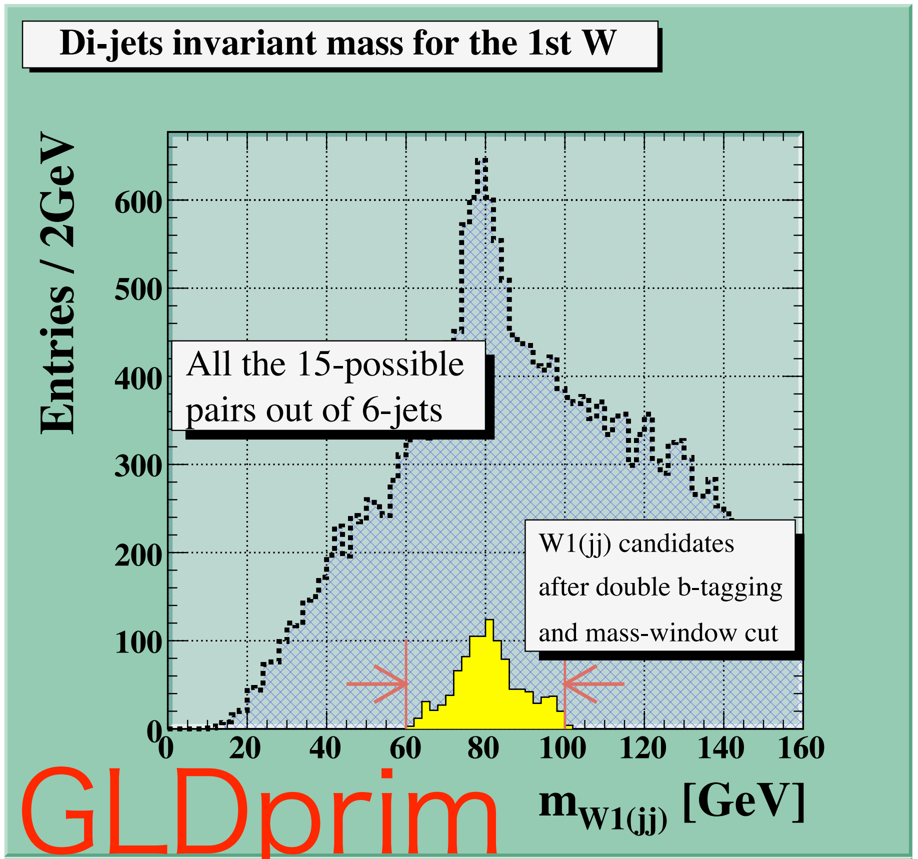


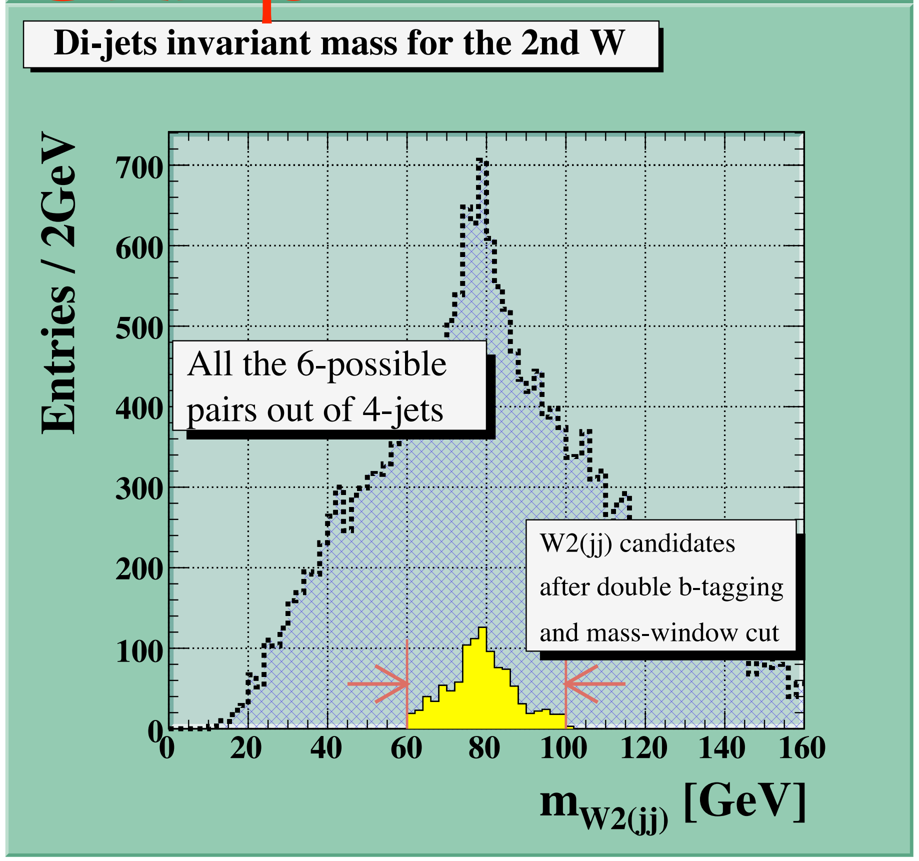


charm contamination?

W_1 & W_2 reconst w/ double b-tagging







I) Choose all the 15-possible pairs out of 6-jets => W_1 candidate

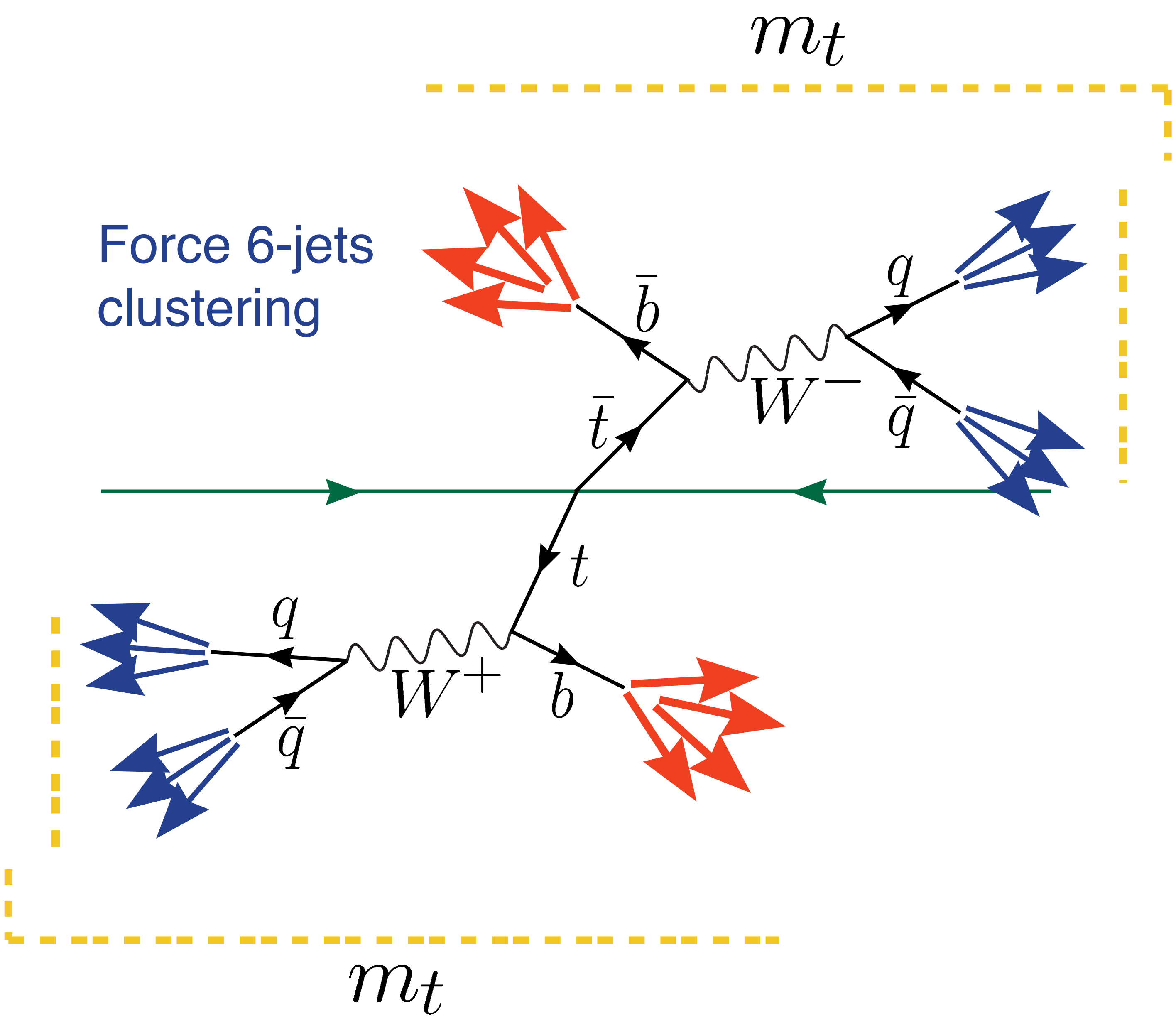
II) Choose all the 6-possible paris out of remaining 4-jets => W_2 candidate

III) Remaining 2-jets should be b-jets: flavor tagging (charm/bottom tagging) is very important to eliminate both combinatorial and process BGs

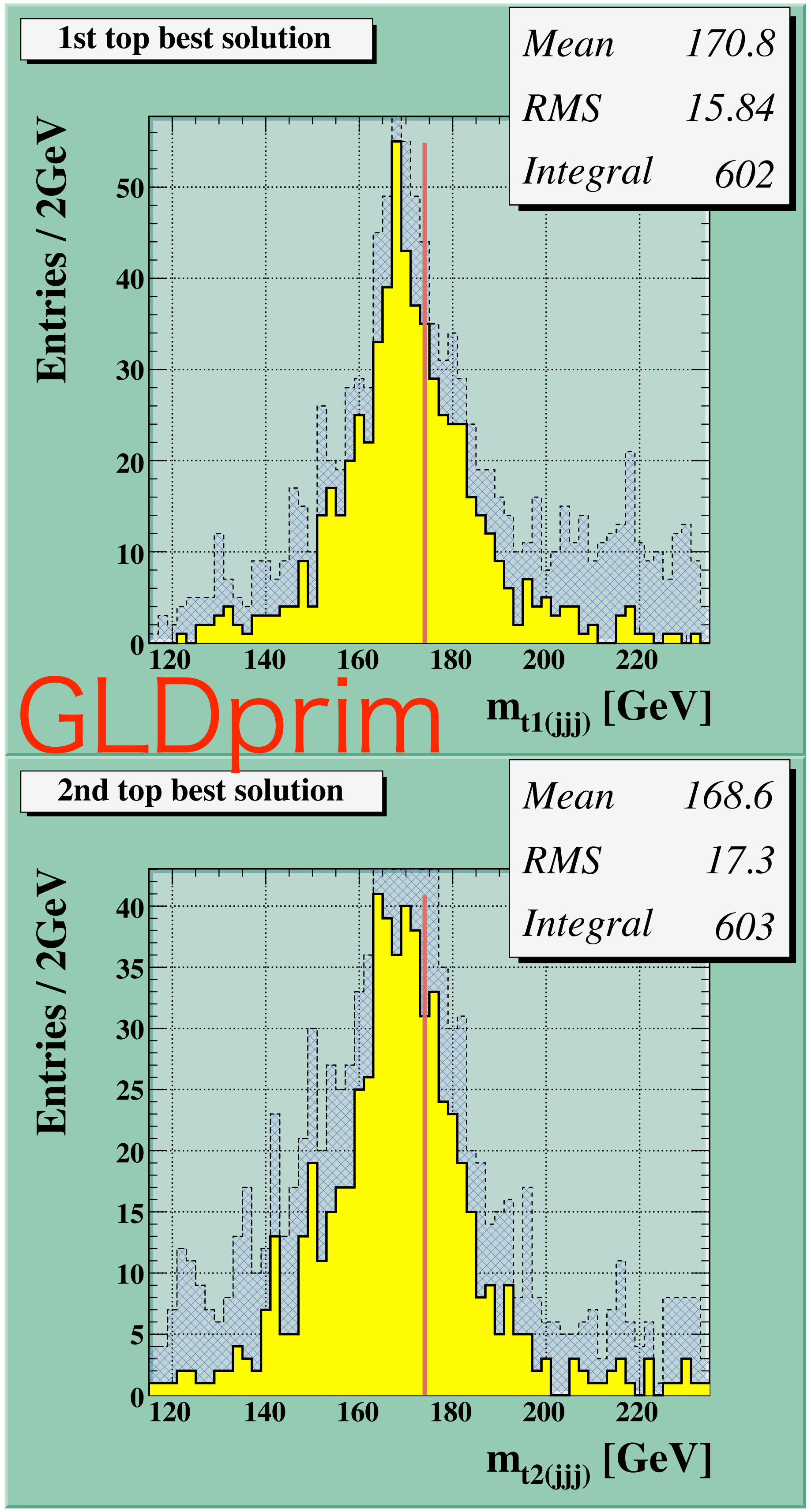
- Loose di-jet mass window cut (±20GeV)
- Double b-tagging (b-tag value > 0.8)

=> ~half of ttbar events were lost...

t_1 & t_2 reconst w/ double b-tagging



Force 6-jets clustering



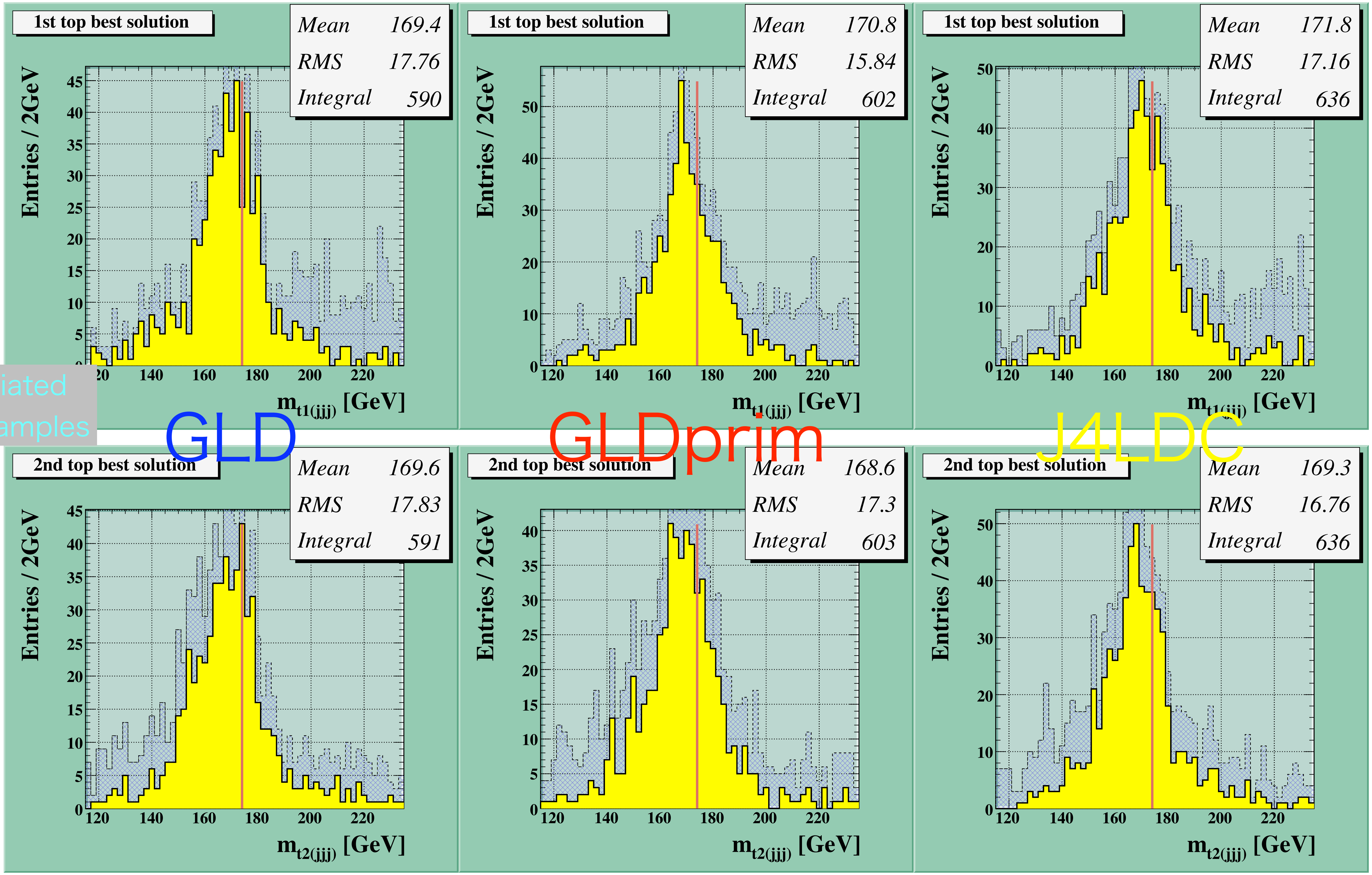
I) After double b-tagging, there are 2 possibilities to attach a b-jet to W_1 and W_2 candidates

II) Sort solutions according to χ^2: choose the best solution

$$\chi^2 = (m_{w1} - m_w)^2/\sigma^2_{mw} + (m_{w2} - m_w)^2/\sigma^2_{mw} + (m_{t1} - m_t)^2/\sigma^2_{mt} + (m_{t2} - m_t)^2/\sigma^2_{mt}$$

M_{t1} & M_{t2} for the best solution

- 3-jets mass resolution (t1 & t2 for the best solution)



ttbar mediated
bbuddu samples

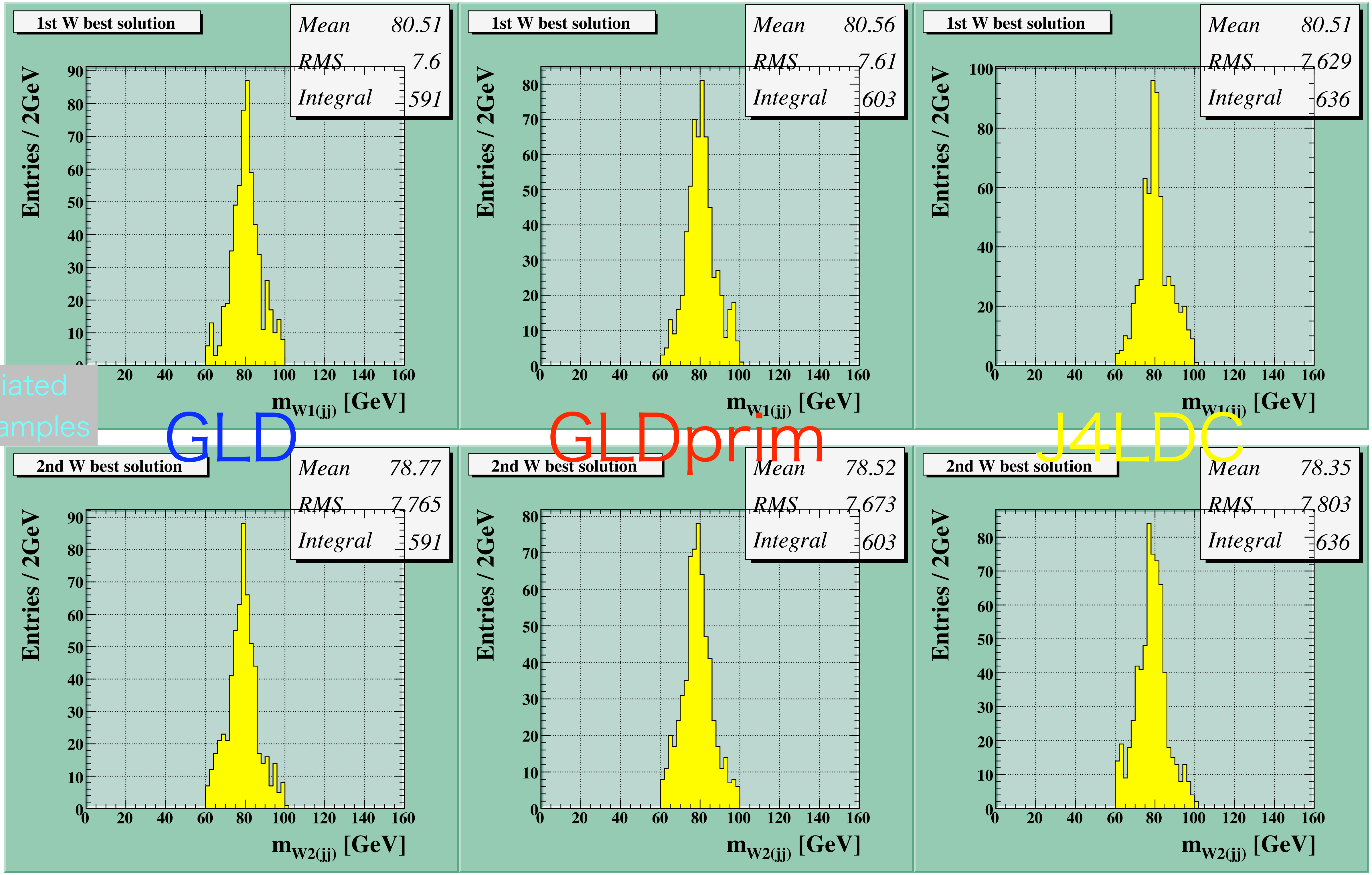
GLD

GLDprim

J4LDC

M_{W1} & M_{W2} for the best solution

- Di-jets mass resolution (W1 & W2 for the best solution)



ttbar mediated bbuddu samples

GLD

GLDprim

J4LDC

Summary / Plan

- Toward writing up the ILD-LoI:
  - e^+e^- -> ttbar @CMS 500GeV is one of the common LoI benchmarking process
    - 6-jets mode (t->bW, W->qq') should be analyzed to test b-tagging and PFA in multi-jet events
  - m_{top}(3jet) measurement for 3 different detector geometries
  - No significant differences for fully reconstructed ttbar events (m_{t1}, m_{t2}, m_{W1} & m_{W2})
  - LCFIVertex performance under 6-jets environment is worse than expected
    - Problem on NN training and/or forced jet clustering??
- Next step:
  - Deeper understanding of LCFIVertex under multi-jet environment (also Vertex charge measurement)
  - Check jet-parton association to solve our problem

Backup slides

GLD / GLDPrim / J4LDC

Name of Sheet of Properties		Properties					
Name of Sheet		glddec07_14m	gldapr08_14m	gldprim_v02	gldprim_v04	j4ldc_v02	j4ldc_v04
BeamPipe							
OuterRadius1		1.525	1.55	1.525	1.450	1.525	1.35
OuterRadius2		1.525	1.55	1.525	1.450	1.525	1.35
BeamPipeThickness		0.025	0.050	0.025	0.050	0.025	0.05
InnerRadius		1.5	1.5	1.5	1.4	1.5	1.3
Zmax J4.IR.ZEdges		5.0	8.0	5.0	8.0	5.0	8.0
Zmax J4IR.BeamPipeIP.HalfZLength		5.5	8.0	5.5	8.0	5.5	8
Zmax J4IR.BeamPipeMiddle.ZEdges		5.0	8.0	5.0	8.0	5.0	8.0
Vertex Detectors							
Number of Layers		6	6	6	6	6	6
Inner Radius		2.0	1.75	1.8	1.6	1.6	1.5
OuterRadius		5.0	6.0	4.8	6.0	4.6	6.0
VTX Thickness/Layers		0.005	0.0094	0.005	0.0094	0.005	0.0094
Silicon Radiation Length		9.36	9.36	9.36	9.36	9.36	9.36
VTX Thickness/Layer in Rad. Length	0.0010	0.000534188	0.001004274	0.000534188	0.001004274	0.00053419	0.0010043
Total VTX Thickness in Rad. Length		0.003205128	0.006025641	0.003205128	0.006025641	0.00320513	0.0060256
Half Z Length of 1st Layer		6.5	7.25	6.5	7.25	6.5	7.25
Outer Radius of 2nd Layer		2.205	1.9594	2.005	1.8094	1.805	1.7094
CosTheta Max of 2nd Layer	0.95	0.946994531	0.965365387	0.955572163	0.970240036	0.96353914	0.9733118
Half Z Length of 4th Layer		10	13.5	10	13.5	10	13.5
Outer Radius of 4th Layer		3.405	4.0094	3.205	3.9094	3.005	3.8594
CosTheta Max of 4th Layer	0.95	0.946628398	0.958616106	0.952285837	0.960535654	0.95769439	0.9614815
Half Z Length of 6th Layer		10	13.5	10	13.5	10	13.5
Spacial Resolution							
Intermediate Tracker Barrel							
Number of Layers		4	4	4	4	4	4
Inner Radius of 1st Layer		9.0	9.0	9.0	9.0	9.0	9.0
Inner Radius of Last Layer		30.0	30.0	30.0	30.0	29	29
Total Barrel IT Thickness (R.L.)		0.024	0.024	0.024	0.024	0.024	0.024
MaxCosTheta		0.9	0.9	0.9	0.9	0.9	0.9
HalfZ of 1st layer		18.58267444	18.58267444	18.58267444	18.58267444	18.5826744	18.582674
HalfZ of 2nd layer		33.03586568	33.03586568	33.03586568	33.03586568	33.0358657	33.035866
HalfZ of 3rd layer		47.48905691	47.48905691	47.48905691	47.48905691	47.4890569	47.489057
HalfZ of 4th layer		61.94224815	61.94224815	61.94224815	61.94224815	59.8775065	59.877507
J4IT.Barrel.HalfZ		18.5 33.0 47.5 62.0	18.5 33.0 47.5 62.0	18.5 33.0 47.5 62.0	18.6 33.0 47.5 62.0	18.5 33.0 47.5 57	18.5 33.0 47.5 57
Intermediate Tracker Endcap							
Number of layers		7	7	7	7	7	7
MinimumZ		15.5	15.5	15.5	15.5	15.5	15.5
Maximum Z		101.5	101.5	101.5	101.5	101.5	101.5
TPC							
TPC Region Inner Radius		39.5	39.5	39.5	39.5	30	30
Inner SupportTube Thickness		4.215	4.215	4	4	4	4
Drift Region Rmin		43.715	43.715	43.5	43.5	34	34
TPC Region Outer radius		206	206	180	180	158	158
Outer SupportTube Thickness		8.235	8.235	6	6	6	6
DriftRegion Rmax		197.765	197.765	174	174	152	152
Inner Support Tube Rad. Length		0.013000993	0.020864271	0.012337835	0.01980002	0.01233783	0.0198
Outer SupportTube Rad. Length		0.01700115	0.027283841	0.012386994	0.019878937	0.01238699	0.0198789
EndPlate and PadPlane Thickness		5	10	5	10	5	10
TPC Region HalfZ		260	260	235	235	216	216
TPC Drift Region HalfZ		255	250	230	225	211	206
EndPlate and PadPlane Thickness R.L.		0.0999972	0.147795858	0.0999972	0.147795858	0.0999972	0.1477959
# of Radial Sampling		200	256	170	217	150	196
Pad Radial Height		0.77025	0.601757813	0.767647059	0.601382488	0.78666667	0.6020408
Calorimeter							
ECAL Rmin		210	210	185	185	160	160
ECAL # of layers		33	33	33	33	33	33
ECAL Layer Thickness in cm		0.6	0.6	0.6	0.6	0.6	0.6
ECAL Layer Thickness in Rad. Length		0.860794846	0.860794846	0.860794846	0.860794846	0.86079485	0.8607948
ECAL Thickness		19.8	19.8	19.8	19.8	19.8	19.8
ECAL Total Rad. Length		28.40622991	28.40622991	28.40622991	28.40622991	28.4062299	28.40623
HCAL # of layers		46	46	41	42	37	37
HCAL Layer Thickness in cm		2.6	2.6	2.6	2.6	2.6	2.6
HCAL Total Thickness		119.6	119.6	106.6	109.2	96.2	96.2
HCAL Radius of Last Layer		349.4	349.4	311.4	314	276	276
Segment Half Angle (degree)		15	15	15	15	15	15
HCAL Rmax of Last Layer		361.7254974	361.7254974	322.3850026	325.0767206	285.736226	285.73623
ECAL layer Nucl. Int. Length		0.031924195	0.031924195	0.031924195	0.031924195	0.03192419	0.0319242
ECAL Total Nucl. Int. Length		1.053498424	1.053498424	1.053498424	1.053498424	1.05349842	1.0534984
HCAL layer Nucl. Int. Length		0.124730859	0.124730859	0.124730859	0.124730859	0.12473086	0.1247309
HCAL Total Nucl. Int. Length		5.73761952	5.73761952	5.113965224	5.238696083	4.61504179	4.6150418
Total Nucl. Int. Length		6.791117943	6.791117943	6.167463648	6.292194507	5.66854021	5.6685402
Barell CAL HalfZ		270	270	235	235	210	210
Endcap ECAL Front Z		280	280	245	245	220	220
Endcap ECAL LastZ		299.8	299.8	264.8	264.8	239.8	239.8
Endcal HCAL LastZ		419.4	419.4	371.4	374	336	336
Cryostat							
Cryostat Inner Radius		375	375	335	330	300	300
Coil-HCAL Distance of Barell		13.27450256	13.27450256	12.61499742	4.923279351	14.2637742	14.263774
Cryostat Outer Radius		440	440	400	385	375	375
Cryostat Thickness		65	65	65	55	75	75
Cryostat HalfZ		475	475	440	375	415	415
Solenoid Bfield		3	3	3.5	3.5	4	4
Coil Center Radius		400	400	360	355	325	325
Coil Thickness		5	5	5	30	5	5
Coil HalfZ		430	430	395	370	370	370
Return Yoke and Muon Detector							
Barrel Minimum Radius		445	445	405	405	380	380
Coil - Return Yoke Minimum Distance		15	15	10	35	10	10
Barrel Outer Radius		765	765	690	690	700	700
Barrel RY and MUD Total Thickness		315	315	280	280	280	280
Barrel RY Outer Radius		760	760	685	685	660	660
Barrel RY Maximum Radius		786.8098971	786.8098971	709.1641836	709.1641836	683.282279	683.28228
Endcap FrontEndcap FrontZ		425.1	425.1	390	375	365	365
Endcap FrontEndcap LastZ		485	485	450	400	425	425
Endcap FrontEndcap Nsuler Layers		2	2	2	1	2	2
Endcap LastZ		800	800	765	695	740	740
IR-FCAL							
FrontZ		230.0	230.0	230.0	230.0	230.0	230.0
Front Rmin		8.2	8.2	8.2	8.2	8.2	8.2
LastZ		284.99	284.99	284.99	284.99	284.99	284.99
Rmax		39.79	39.0	39.0	39.0	29.0	29.0
IR-BCAL							
CH2Mask FirstZ		405	405	405	405	405	405
CH2Mask LastZ		430	430	430	430	430	430
CH2Mask Outer Radius		20	20	20	20	20	20
BCAL FirstZ		430	430	430	430	430	430
BCAL LastZ		450	450	450	450	450	450
QC FirstZ		451.0	451.0	451.0	451.0	451.0	451.0

SLAC SM data samples: FAQ

- Whizard 1.40 has no gluon emission by default, leading to potentially incorrect multiplicity distributions.
- The WHIZARD version 1.40 that was used to generate this sample indeed did not include gluon emission. However gluon radiation was simulated using PYTHIA's parton showering algorithm. WHIZARD versions 1.50 and higher include gluon emission, and, starting with version 1.91, WHIZARD has its own parton showering code.
- Whizard 1.40 has an incorrect implementation of the CKM matrix. Only diagonal terms of the matrix are present (and = 1!), giving wrong W decays.
- Although true for the Whizard version 1.40 that was used to generate this data sample, it is extremely doubtful that this will have any effect on the current analyses. WHIZARD versions 1.51 and higher include the correct CKM matrix, and so future data samples will include the rarer W decays.
- This sample has generator level cuts a la SiD, providing a potential bias when used for ILD.
- There are, indeed, some kinematical cuts for processes with divergent cross-sections, which can be seen by looking at the whizard.in file as described above. However, the only kinematic cut that leads to a genuine loss of events is a 4 GeV minimum invariant mass cut on final state fermion-antifermion pairs.

M@RS

A DST analysis package

★M@RS = Modular Analysis with Root-based Subprograms

❖Aim

- provides a common framework for ILD-DST analysis
- same approach (= minimum user code modification) between Full simulator-Standard reconstruction (FullLDCTracking/PandoraPFA/LCFIVertex) and Quick simulator analyses
  - ➡make maximum use of the past resources!!
- use ROOT and OO features maximally = efficient and elegant reconstruction for complicated final states (e.g. ttbar -> 6jets) using LEDA/Anlib; ANL4DVector (Lockable TLorentzVector), ANLPair etc.

❖Dependencies

- ROOT, LEDA/Anlib and JSF-kern (Is it possible to distribute only JSF-kern??)

❖Impacts on our (ex-GLD side) physics benchmarking studies

- So far there was no interface between LCFIVertex and JSF => Flavor tagging (and Jet charge) is available
- Need MC truth of jet flavor (searchs the event for the leading hadron, and if this is a heavy flavor particle determines which of the jets in the event is closest in angle)
  - ➡ TrueAngularJetFlavourProcessor in LCFIVertex

❖Available to access all of the above information through MarsJet class

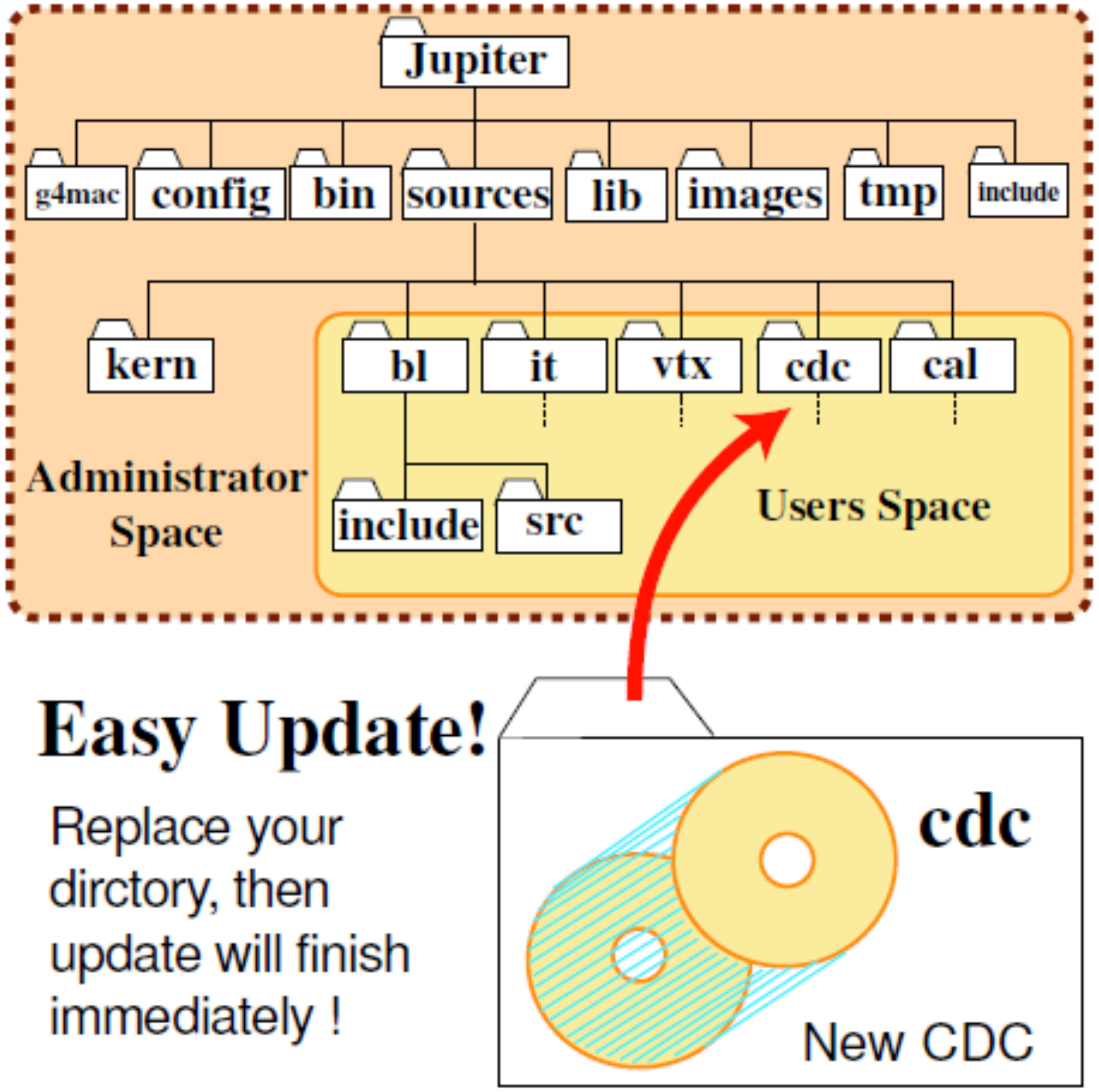
JSF: its features

- Framework: JSF = Root based application
  - All functions based on C++, compiled or through CINT
  - Provides common framework for event generations, detector simulations, analysis, and beam test data analysis
  - Unified framework for interactive and batch job: GUI, event display
  - Data are stored as root objects; root trees, ntuples, detector configuration in Jupiter run

- Release includes other tools; QuickSim, Event generators, beamstrahlung spectrum generator, etc.

Jupiter: its features (1)

- Currently w/ Geant4 9.1p1 Physics List: LCPysicsList (Default)
- Modular structure
  - easy installation of sub-detectors
- Geometry
  - Simple geometries are implemented (enough for the detector optimization)
  - parameters (size, material, etc) can be modified by an input ASCII file at run time
  - Parameters are saved as a ROOT object for use in Satellites later



Easy Update!

Replace your dirctory, then update will finish immediately !

Jupiter: its features (2)

- Input:
  - StdHep file (ASCII), HepEvt, CAIN, or any generators implemented in JSF
  - Interface to StdHep: Prepared as a JSFModule, using StdHep 5.06.01
- Output:
  - Exact Hits of each detectors (Smearing in Satellites)
  - Pre- and Post- Hits at before/after Calorimeter
    - Used to record true track information which enter CAL/FCAL/BCAL
  - Break points in tracking volume
  - Output in LCIO Format is through a JSFModule
- Run mode:
  - A standalone Geant4 application
  - JSF application to output a ROOT file