Higgs to 4leptons with EventViews

Stathes Paganis
(University of Sheffield)
with Rosy.Nikolaidou (Saclay)
Nektarios Benekos (Max Planck Institute)

PAT: Analysis Model Meeting, 1-Nov-05

Motivation

- Why we decided to use EventViews?
  - Need for a clean, transparent and modular analysis
  - We want to optimally collaborate with other groups:
    - Parallel development
    - Cross-check results, avoid confusion on how different people do their analyses
  - We want access to analysis tools and input from software experts
  - We want to have different 'views' of the Higgs to 4lepton:
    - Combine soft-muons with comb-muons etc
    - Combine soft-e with egamma-e and/or less isolated e
    - Study new electrons/muons which could be hybrid objects (they use track info in some cases and LAr/MuSpec info in other cases).
- Current Problem: see confusion in last Higgs WG with the cuts/approach used in 2/3 different analyses. Very hard to check what people are actually doing/using! We need a way for transparency in accord with recent Analysis guidelines (Fabiola et al).

Documentation

- Recent Note in Higgs to 4leptons: ATL-COM-PHYS-2005-043
- Talks in egamma+Higgs sessions October PhysWeek:
  http://agenda.cern.ch/fullAgenda.php?ida=a056348
  http://agenda.cern.ch/fullAgenda.php?ida=a056449
- EventView (Sep/05 Software Week) K.Cranmer, Amir Farbin:
  http://agenda.cern.ch/fullAgenda.php?ida=a045109
- EventView twiki documentation:
  https://uimon.cern.ch/twiki/bin/view/Atlas/EventView
- Egamma twiki documentation:
  https://uimon.cern.ch/twiki/bin/view/Atlas/ElectronGamma

Disclaimer

- We are still working with 10.0.1
- Still learning how to optimally use available tools
  - Maybe it would be nice to have eventView snipet examples on twiki?
- Don't use arrays yet (say for muons in eventView)
- Followed original instructions on twiki and nothing beyond this
- Counters per tool are vectors in the program which live in the looper (EV_Hto4l)
- We have written a muon/electron performance tool to study efficiencies at the single lepton level
  - Maybe these tools should go directly to the eventViewTools since they are general?

Datasets/People/Framework (1)

- Datasets-Running
  - Higgs Rome Samples no pile-up, SM rome sample with pile-up
  - Jobs running at lxplus and Sheffield (interactively)
  - JobOutput is a personalized ntuple (see discussion)
- Single Electrons
  - Rome.004021, 004022, 004027 with pile-up (thanks to M.Wielers)
- Z bosons
  - Rome.004201.recov10.lumi01.ZeeJimmy (SM group)
- Higgs
  - Rome.003045.recov10.H1_130_4l (no pile-up)
  - Rome.004321.recov10.AcerMC_Zbb_4l
- New Samples needed:
  - Higgs signal with (+cavern bgnd) pile-up, ttbar + Zbb (filtered)
  - Single Muon/Electron files

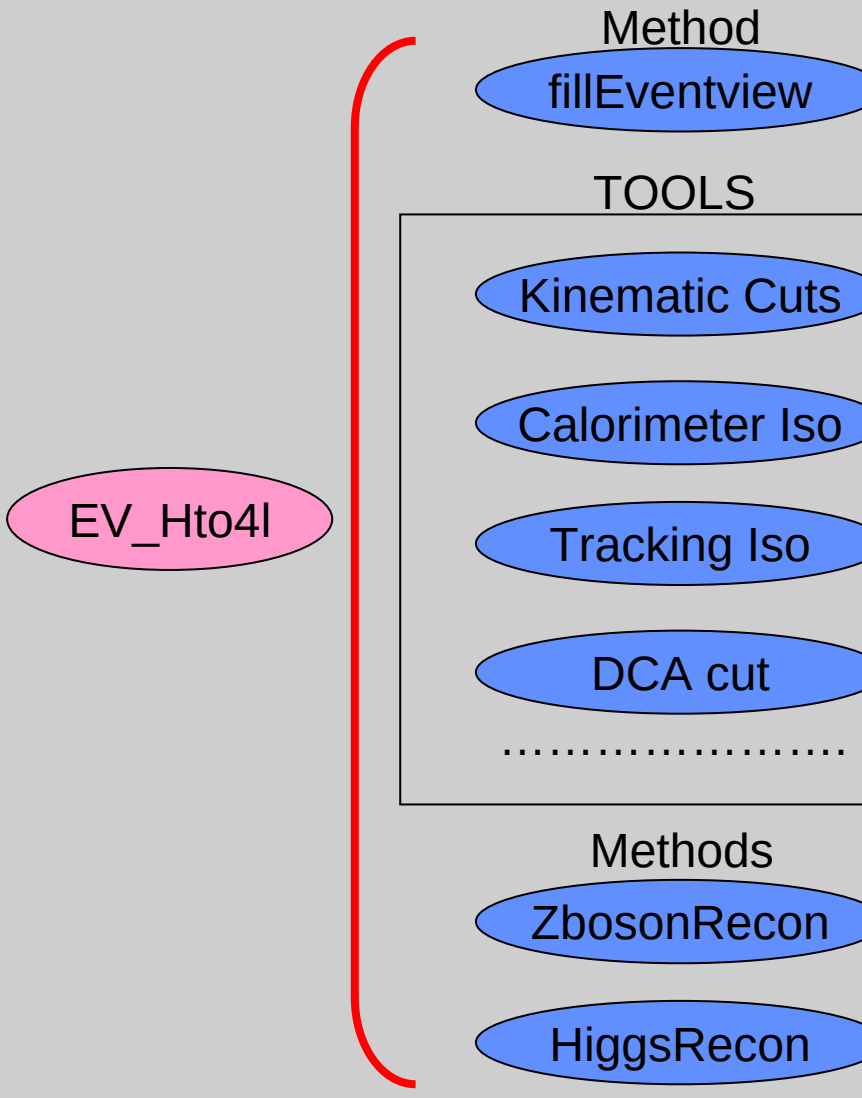
Datasets/People/Framework (2)

- Datasets-Running
  - Higgs Rome Samples no pile-up, SM rome sample with pile-up
  - Jobs running at lxplus and Sheffield (interactively)
  - JobOutput is a personalized ntuple (see discussion)
- People contributing (up to now)
  - Electron work: SP
  - Muon work: Rosy Nikolaidou, Nektarios Benekos
- Fraction of the analysis performed inside/outside framework
  - Analysis originally written on CBNT level outside framework
  - Cut factories were written in order to study alternative strategies
  - Currently these factories are transformed in eventViewTools
  - Refined analysis methods (e.g. constrained Z Mass fits) are still outside the framework working on CBNT output.

How we build the EventViews/UserData

- UserData design is under discussion
  - We store electron, muons, Zs, Higgs candidates (currently kinematic variables, cut flags, invariant masses)
- The goal is to be able to repeat the analysis at the ntuple level performing for example:
  - Z mass constraint fits
  - Extraction of Background under the peak
  - Significance calculations
- We may need input here from Amir et al; very little work has been done at the post-eventView ntuple level

EV_Hto4l: the modular pieces



Fills the eventView with user variables/objects including truth. These end-up in the output root-Tuple

The “cut” tools are just instantiations of one (simple) tool, inheriting from:
eventViewBaseTool
(like the usual particle inserter tools)

The reconstructed methods live in the EV_Hto4l and are also configurable from jobOptions like the cut tools

Z and Higgs boson building

- We have 2 methods for each Z and Higgs
  - One is using inferred objects
  - The other is simple and just fills the eventView with variables corresponding to the number of Zs, Hs, masses etc.
- Z->ee builder:
  - Loop over all electrons in eventView (they have passed all cuts)
  - Create Z candidates requiring that at least 2 Zs are found and one of them is on-shell (std TDR cuts depending on Higgs mass)
- H->ZZ builder:
  - Loop over all Zs making sure the same 4 electrons are not used by 2 Higgs candidates.
  - Keep all H candidates (usually one lepton-quadruplet survives)

EV_Hto4l : jobOptions(1)

```
#
# Setup Dlls for analysis and set algorithm options
#
theApp.Dlls      += ["AnalysisTools" , "TruthParticleAlgs"]
theApp.Dlls      += ["UserAnalysis"]
theApp.TopAlg += ["EV_Hto4l/higgs"]
higgs=Algorithm("higgs")
#
# Top Algorithm (I call it "higgs")
#
higgs.OutputLevel=ERROR

#
# The higgs analysis factory tools (order matters)
#
higgs.Hto4lTools=[
                "EV_Hto4lPreselectionTool/fourEMLArClusters",
                "EV_Hto4lPreselectionTool/fourTrkMatchClusters",
                "EV_Hto4lPreselectionTool/fourTrkQualClusters",
                "EV_Hto4lPreselectionTool/threeIsoEMLArClusters",
                "EV_Hto4lPreselectionTool/fourIsoEMLArClusters",
                "EV_Hto4lPreselectionTool/fourIsolatedTracks",
                "EV_Hto4lPreselectionTool/fourLowImpParamTrks"
                ]
```

Notice: there are 7 tools defined however all 7 are re-incarnations of the same “Preselection Tool” (in this case the tool is actually an eventViewBaseTool)

EV_Hto4l : jobOptions(2)

```
Hto4l.fourEMLArClusters.cutName            = "fourEMLArClusters"
Hto4l.fourEMLArClusters.doHto4e            = True
Hto4l.fourEMLArClusters.etaMax             = 2.5
Hto4l.fourEMLArClusters.ptLow              = 7.0*GeV
Hto4l.fourEMLArClusters.ptHigh             = 20.0*GeV
```

Kinematic cuts only
(TDR std cuts)

```
Hto4l.fourTrkMatchClusters.cutName         = "fourTrkMatchClusters"
Hto4l.fourTrkMatchClusters.doHto4e         = True
Hto4l.fourTrkMatchClusters.useTrackMatch   = True
```

Kinematic cuts
+ 4 egclusters/tracks

```
Hto4l.fourTrkQualClusters.cutName          = "fourTrkQualClusters"
Hto4l.fourTrkQualClusters.doHto4e          = True
Hto4l.fourTrkQualClusters.useTrackMatch    = True
Hto4l.fourTrkQualClusters.useTrkQualIsEM   = True
Hto4l.fourTrkQualClusters.useTRT           = False
```

Kinematic cuts
+ 4 egclusters/tracks
+ 4 isEM tracks (no TRT)

```
Hto4l.threeIsoEMLArClusters.cutName        = "threeIsoEMLArClusters"
Hto4l.threeIsoEMLArClusters.doHto4e        = True
Hto4l.threeIsoEMLArClusters.useTrackMatch = True
Hto4l.threeIsoEMLArClusters.useTrkQualIsEM= True
Hto4l.threeIsoEMLArClusters.useTRT         = False
Hto4l.threeIsoEMLArClusters.useCaloIsEM    = True
Hto4l.threeIsoEMLArClusters.nCaloNonIsoEle = 1 #num of Non-Iso elec (3+1)
```

Kinematic cuts
+ 4 egclusters/tracks
+ 4 isEM tracks (no TRT)
+ 3/4 isEM clusters

EV_Hto4l : jobOptions(3)

```
Hto4l.fourIsolatedTracks.cutName                = "fourIsolatedTracks"
Hto4l.fourIsolatedTracks.doHto4e                = True
Hto4l.fourIsolatedTracks.useTrackMatch          = True
Hto4l.fourIsolatedTracks.useTrkQualIsEM         = True
Hto4l.fourIsolatedTracks.useTRT                 = False
Hto4l.fourIsolatedTracks.useCaloIsEM            = True
Hto4l.fourIsolatedTracks.nCaloNonIsoEle         = 0
Hto4l.fourIsolatedTracks.useIsolation           = True
Hto4l.fourIsolatedTracks.isolationCone          = 0.2       # in deltaR
Hto4l.fourIsolatedTracks.absoluteIsolationCut   = 10.0*GeV
```

Kinematic cuts
+ 4 egclusters/tracks
+ 4 isEM tracks (no TRT)
+ 4 isEM clusters
+ 4 isolated tracks

```
Hto4l.fourLowImpParamTrks.cutName               = "fourLowImpParamTrks"
Hto4l.fourLowImpParamTrks.doHto4e               = True
Hto4l.fourLowImpParamTrks.useTrackMatch         = True
Hto4l.fourLowImpParamTrks.useTrkQualIsEM        = True
Hto4l.fourLowImpParamTrks.useTRT                = False
Hto4l.fourLowImpParamTrks.useCaloIsEM           = True
Hto4l.fourLowImpParamTrks.nCaloNonIsoEle        = 0
Hto4l.fourLowImpParamTrks.useIsolation          = True
Hto4l.fourLowImpParamTrks.isolationCone         = 0.2       # in deltaR
Hto4l.fourLowImpParamTrks.absoluteIsolationCut  = 10.0*GeV
Hto4l.fourLowImpParamTrks.useImpactParamCut     = True
Hto4l.fourLowImpParamTrks.impactParamSigmaCut= 5.0
```

Kinematic cuts
+ 4 egclusters/tracks
+ 4 isEM tracks (no TRT)
+ 4 isEM clusters
+ 4 isolated tracks
+ 4 small perigee tracks

Efficiency table S/B for default cuts

Higgs->4e

	Total	Passed	Efficiency	TotalEff
fourEMLArClusters	8076	734	0.0909	0.0909
fourTrkMatchClusters	734	734	1.0000	0.0909
fourTrkQualClusters	734	531	0.7234	0.0658
threeIsoEMLArClusters	531	435	0.8192	0.0539
fourIsoEMLArClusters	435	266	0.6115	0.0329
fourIsolatedTracks	266	250	0.9398	0.0310
fourLowImpParamTrks	250	223	0.8920	0.0276
ZeeReconstruction	223	223	1.0000	0.0276
HtoZZ4eReconstruction	223	165	0.7399	0.0204
Total				0.0204

Default Analysis (TDR-like)

Zbb->4l

	Total	Passed	Efficiency	TotalEff
fourEMLArClusters	94513	10745	0.1137	0.1137
fourTrkMatchClusters	10745	10745	1.0000	0.1137
fourTrkQualClusters	10745	4274	0.3978	0.0452
threeIsoEMLArClusters	4274	408	0.0955	0.0043
fourIsoEMLArClusters	408	53	0.1299	0.0006
fourIsolatedTracks	53	24	0.4528	0.0003
fourLowImpParamTrks	24	14	0.5833	0.0001
ZeeReconstruction	14	14	1.0000	0.0001
HtoZZ4eReconstruction	14	7	0.5000	0.0001
Total				0.0001

7/94500 Zbb events pass

EV_SingleLeptonPerformance(Tool)

Structure of Performance tool

```
#
# Top Algorithm (I call it "ePerf")
#
theApp.TopAlg += ["EV_SingleLeptonPerformance/ePerf"]
ePerf=Algorithm("ePerf")
ePerf.doHto4l         = False
ePerf.doSingleElec = True
ePerf.OutputLevel   = FATAL
#
# The performance tool
#
ePerf.PerformanceTool[
                "EV_SingleLeptonPerformanceTool/egammaOnly",
                "EV_SingleLeptonPerformanceTool/egCaloIsEM",
                "EV_SingleLeptonPerformanceTool/egTrackIsEM",
                "EV_SingleLeptonPerformanceTool/egTrackIso",
                "EV_SingleLeptonPerformanceTool/egCaloTrkIsEM",
                "EV_SingleLeptonPerformanceTool/egammaFull",
                ]
```

Notice: there are 6 tools defined however all 6 are re-incarnations of the same “Performance Tool” (in this case the tool is actually an eventViewBaseTool)

Multiple tool definition

```
toolSvc = Service( "ToolSvc" )

toolSvc.egammaOnly.cutName            = "egammaOnly"
toolSvc.egammaOnly.outputLevel        = FATAL

toolSvc.egCaloIsEM.cutName            = "egCaloIsEM"
toolSvc.egCaloIsEM.useCaloIsEM        = True
toolSvc.egCaloIsEM.outputLevel        = FATAL

toolSvc.egTrackIsEM.cutName           = "egTrackIsEM"
toolSvc.egTrackIsEM.useTrkQualIsEM    = True
toolSvc.egTrackIsEM.outputLevel       = FATAL

toolSvc.egTrackIso.cutName            = "egTrackIso"
toolSvc.egTrackIso.useIsolation       = True
toolSvc.egTrackIso.outputLevel        = FATAL

toolSvc.egCaloTrkIsEM.cutName         = "egCaloTrkIsEM"
toolSvc.egCaloTrkIsEM.useCaloIsEM     = True
toolSvc.egCaloTrkIsEM.useTrkQualIsEM  = True
toolSvc.egCaloTrkIsEM.outputLevel     = FATAL

toolSvc.egammaFull.cutName            = "egammaFull"
toolSvc.egammaFull.useCaloIsEM        = True
toolSvc.egammaFull.useTrkQualIsEM     = True
toolSvc.egammaFull.useIsolation       = True
toolSvc.egammaFull.outputLevel        = FATAL
```

Calorimeter Isolation

Track quality cuts

Track cone isolation

Calo Isolation+Track quality

All cuts combined

Single electrons pt=15GeV (pile-up: 10^{33})

	etamin	etamax	Total	Passed	Efficiency
egammaOnly	0.00	0.40	1468	1406	0.9578
egammaOnly	0.40	0.80	1529	1457	0.9529
egammaOnly	0.80	1.20	1514	1433	0.9465
egammaOnly	1.20	1.60	1495	1368	0.9151
egammaOnly	1.60	2.00	1609	1450	0.9012
egammaOnly	2.00	2.40	1494	1346	0.9009
egCaloIsEM	0.00	0.40	1468	1328	0.9046
egCaloIsEM	0.40	0.80	1529	1330	0.8698
egCaloIsEM	0.80	1.20	1514	1341	0.8857
egCaloIsEM	1.20	1.60	1495	1034	0.6916
egCaloIsEM	1.60	2.00	1609	1366	0.8490
egCaloIsEM	2.00	2.40	1494	1220	0.8166
egTrackIsEM	0.00	0.40	1468	1291	0.8794
egTrackIsEM	0.40	0.80	1529	1362	0.8908
egTrackIsEM	0.80	1.20	1514	1349	0.8910
egTrackIsEM	1.20	1.60	1495	1256	0.8401
egTrackIsEM	1.60	2.00	1609	1251	0.7775
egTrackIsEM	2.00	2.40	1494	1187	0.7945
egTrackIso	0.00	0.40	1468	1404	0.9564
egTrackIso	0.40	0.80	1529	1454	0.9509
egTrackIso	0.80	1.20	1514	1426	0.9419
egTrackIso	1.20	1.60	1495	1348	0.9017
egTrackIso	1.60	2.00	1609	1391	0.8645
egTrackIso	2.00	2.40	1494	1206	0.8072

Somewhat better than DC1:

	etamin	etamax	Total	Passed	Efficiency
egCaloTrkIsEM	0.00	0.40	1468	1224	0.8338
egCaloTrkIsEM	0.40	0.80	1529	1245	0.8143
egCaloTrkIsEM	0.80	1.20	1514	1262	0.8336
egCaloTrkIsEM	1.20	1.60	1495	956	0.6395
egCaloTrkIsEM	1.60	2.00	1609	1190	0.7396
egCaloTrkIsEM	2.00	2.40	1494	1084	0.7256

egammaFull	**0.00**	**0.40**	**1468**	**1224**	**0.8338**
egammaFull	**0.40**	**0.80**	**1529**	**1245**	**0.8143**
egammaFull	**0.80**	**1.20**	**1514**	**1262**	**0.8336**
egammaFull	**1.20**	**1.60**	**1495**	**947**	**0.6334**
egammaFull	**1.60**	**2.00**	**1609**	**1168**	**0.7259**
egammaFull	**2.00**	**2.40**	**1494**	**1017**	**0.6807**

Program text output

what would our ideal analysis environment be like?

- Together with the Phys. Analysis code we need performance tools, standardized so to avoid confusion on how efficiencies/rejections are defined.
- Cut tools and reconstruction methods should be fully modular, keep their own eff. counters/names
  - I guess this can be done completely in a python script
- Have the ability to produce a mini-ntuple:
  - With all events OR only events passing our cuts (using a flag)
  - Reaching a 100Hz speed on a single CPU.
  - Ntuple must be standardized per analysis with all necessary variables
  - With optimized truth info in the mini-ntuple
  - With the ability to read it back in athena

Problems, requests, comments, and plans

- 2Hz speed when we write out ntuple
- 50-100Hz when we don't write ntuples
- Need UserData variables as arrays not string indexed doubles (done already)
- Need some examples to make us aware of new PA Tools available
- Standard performance tools could be part of the eventViewTools and run (at least) as validation tools
- We need a Z mass constraint fitter (it is so widely used that it should be an athena PAT)

Plans (2)

- Code will go in “HiggsToFourLeptons” package
- Need to fully implement/test 2e2mu, 4mu
- Timescale ? It would be nice if the code is in place before the next Higgs meeting (mid-December)
- Cut optimization using eventView?
  - Amir and others made comments on a possible eventView setup which is targeting cut optimization (i.e. going beyond the std TDR analysis). For us this is the original goal for using eventView as a more efficient framework to tackle things like Higgs candidates with non-homogeneously classified leptons.