Elias Coniavitis
Martin Flechl

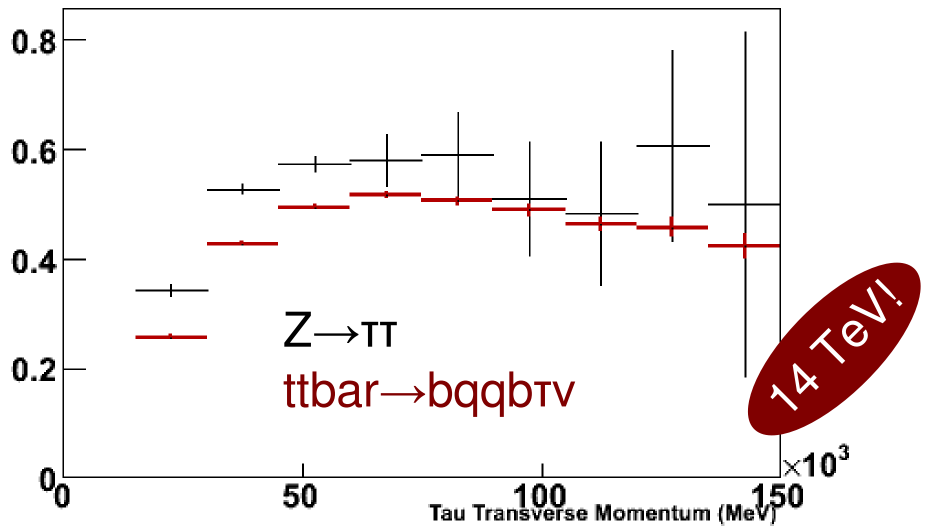
UPPSALA UNIVERSITET

Taus in Dense Environments

ATLAS Workshop on Tau-lepton Physics,
Copenhagen,
April 16th, 2009

Introduction

- τ-ID performance drops in high-multiplicity environments, such as ttbar→bqqbτν events
  - Important for new physics searches involving τ:s, such as H^+

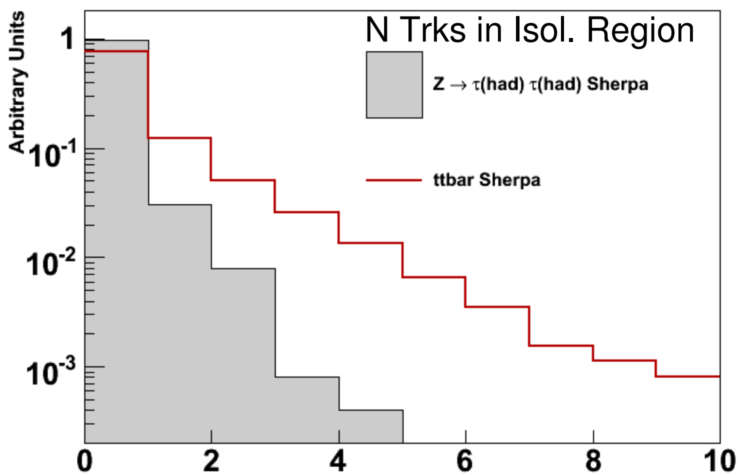


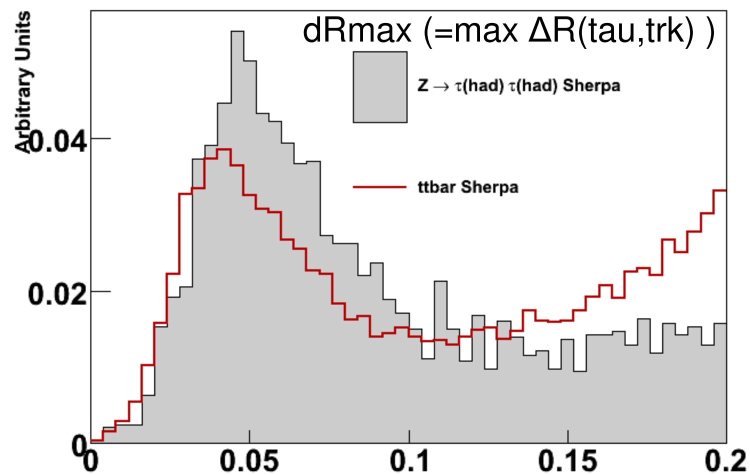
- Possibility of having a LLH optimized for high-multiplicity environments?

Distributions of variables

UPPSALA UNIVERSITET

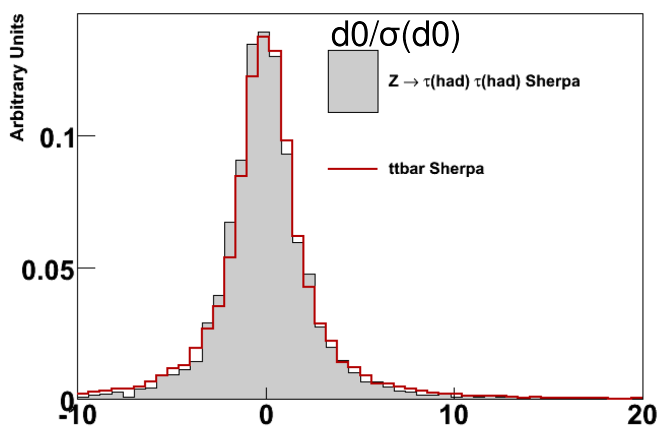
- Compare Z→ττ and ttbar→bqqbτν events:
  Some of the variables entering the LLH have different distributions

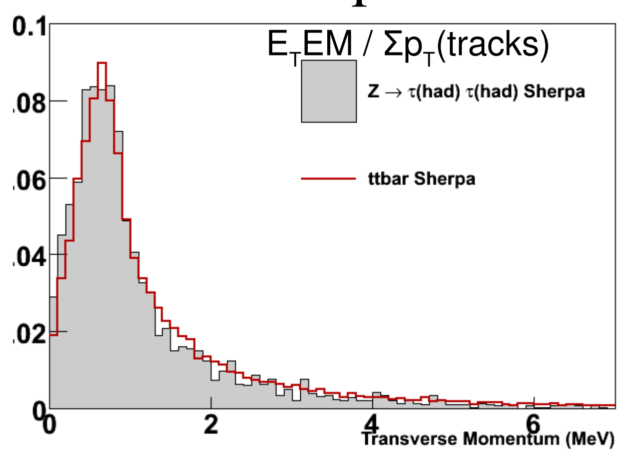


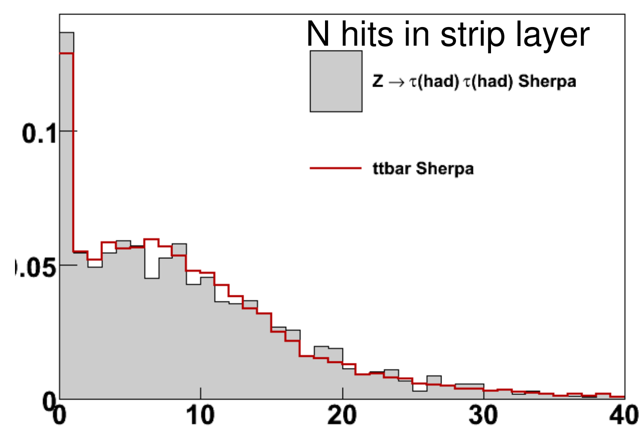


14 TeV!

- However, most of the variables are quite similar







(Example variables for 3 prong τ:s with $30<p_T<60$ GeV @ 14 TeV)

ttbar PDFs

UPPSALA UNIVERSITET

- Default 2008 Likelihood PDFs generated with $Z\rightarrow\tau\tau$ and $A\rightarrow\tau\tau$, i.e. 'cleaner' environment compared to ttbar

➔ Produce PDFs directly from ttbar events instead, and see if that improves situation

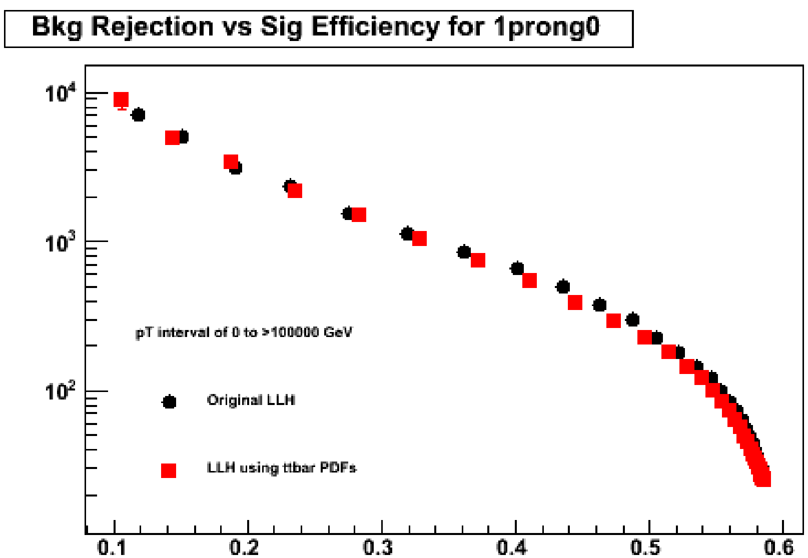
ttbar PDFs

- Results from new, 10 TeV samples for the rest of this talk
- Samples used for PDF production:
  - Taus: 40k ttbar→bqqbτν (privately produced, Sherpa, reco 14.2.25.8)
  - Jets: 350k 105009-105015 (central production, reco 14.2.25.8)
- Tested performance using (different from PDF production):
  - Taus: 40k ttbar→bqqbτν (privately produced, Sherpa, reco 14.2.25.8)
  - Jets: 1.3M 105009-105013 (central production, reco 14.2.25.8)
- Used ARA code provided by Artur Kalinowski; running in 14.5.0
- Currently only considering overlap candidates (calo + track seeded candidates)

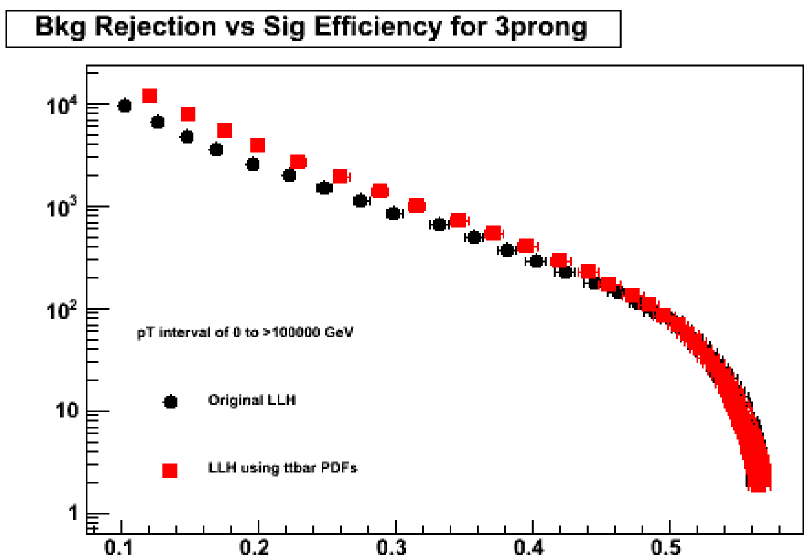
PDFs from ttbar samples

UPPSALA UNIVERSITET

- Producing PDFs from ttbar events does not improve situation much



1 prong



3 prong

(a subset of the QCD “test” sample was used for these plots, for technical reasons)

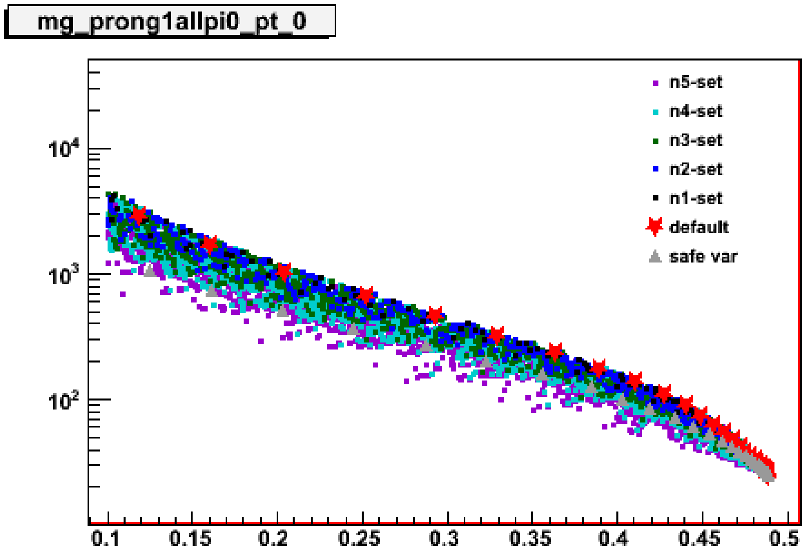
ttbar PDFs

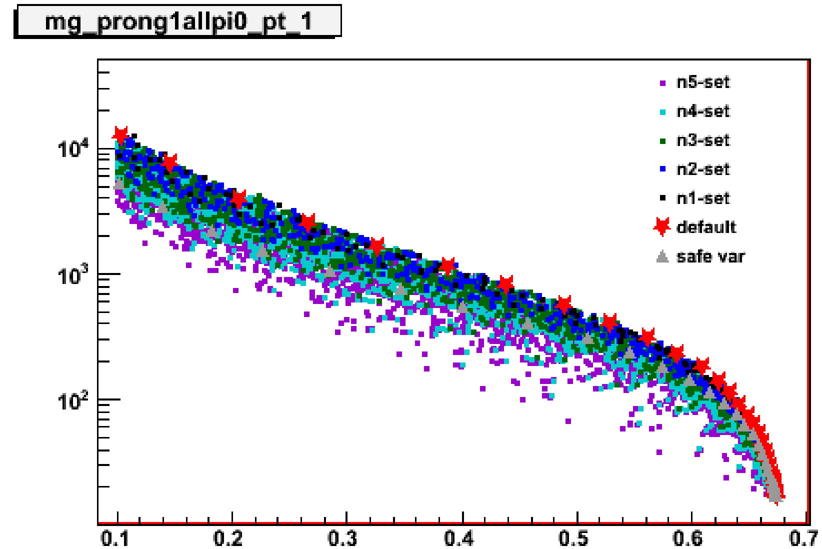
UPPSALA UNIVERSITET

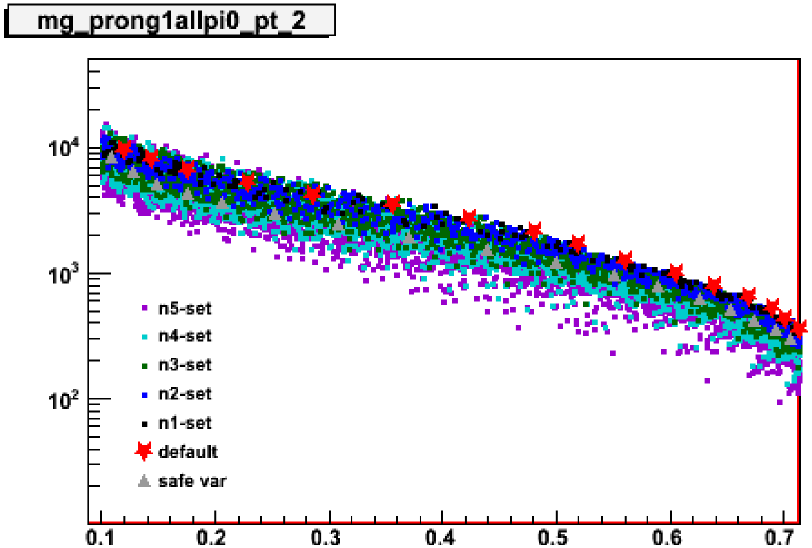
- Relative importance of variables entering this likelihood?
  - Reach similar performance with fewer variables? (→ higher robustness, lower systematics)
- Re-run likelihood from ttbar PDFs, removing one or more variables and compare
  - Define nX-set: Set of all input variable combinations where X variables have been removed from the likelihood
    - e.g. for four variables 0,1,2,3 this would be (1,2),(1,3),(1,4),(2,3),(2,4),(3,4)
  - In each set find variable which makes the least/most difference when removed

UPPSALA UNIVERSITET

1 prong: Efficiency vs Rejection







nX-set: set of all input variable combinations in which x variables from the standard tau likelihood have been removed

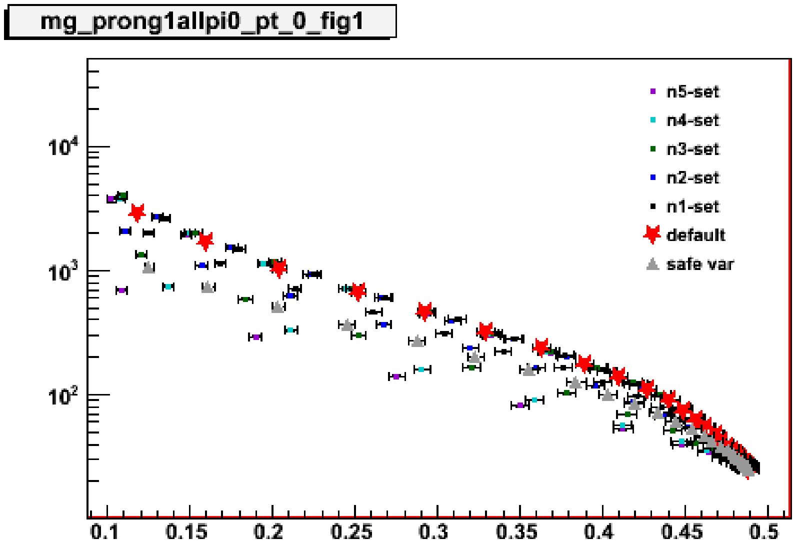
Default: Variables used in Likelihood2008 (but with ttbar PDFs here)

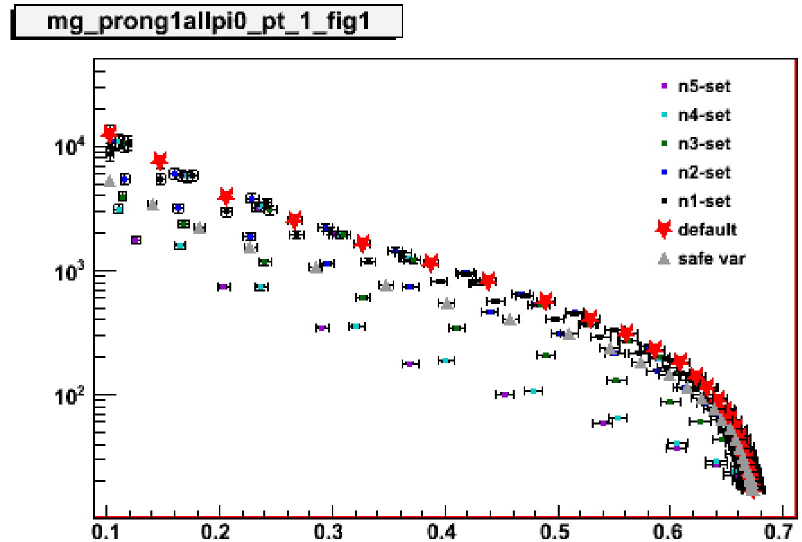
Safe vars: A likelihood constructed using only the 'safe variables'

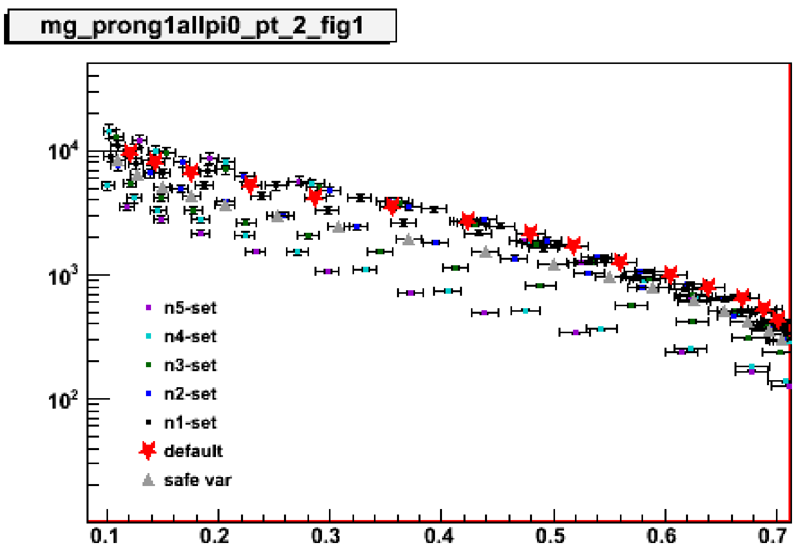
Points are projected on top of each other in the order given in the legend

1 prong: Best/worst performing variable-combinations

UPPSALA UNIVERSITET







Performance definition:

- take rejection from 3 pt-bins for 5 given efficiencies (0.2, 0.25, 0.3, 0.35, 0.4)
- normalize values by rejection of default set
- add all 15 values
- highest number → best-performing combination, lowest number → worst-performing combination (within a given nX-set)

1 prong: Best/Worst performing variables

UPPSALA UNIVERSITET

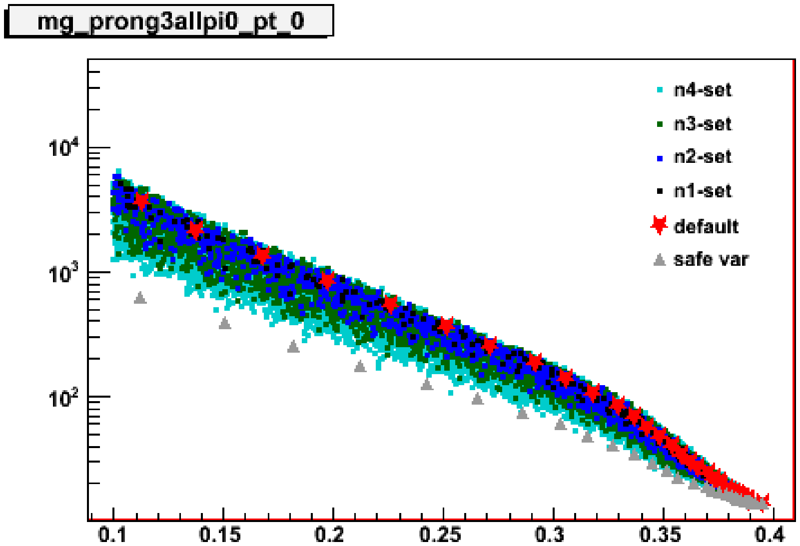
- Best-performing variable-combination → Identify Least Contributing Variables
- Worst-performing variable-combination → Identify Most Contributing Variables

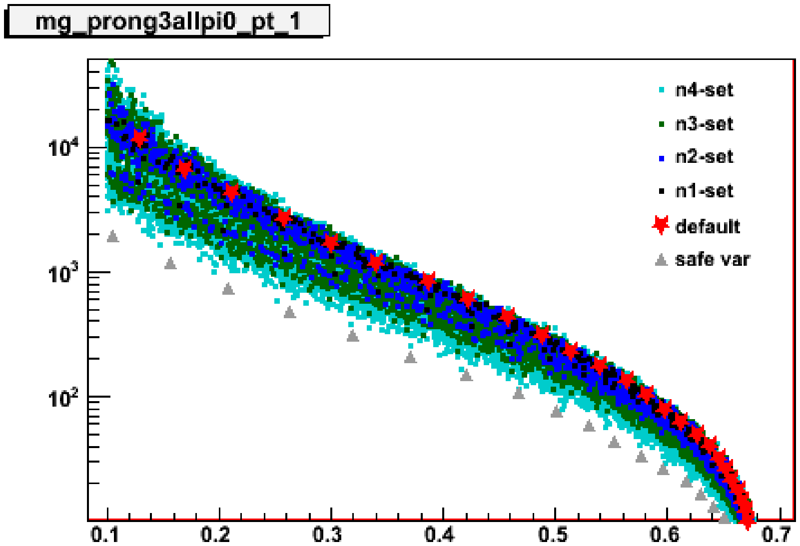
Best	Worst
nAssocTracksIsol	etEM2etTracks
stripWidth2	numStrip
emRadius	etHad2etTracks
isoFrac	MvisEflow
SignD0Trk3p	z0SinThetaSig

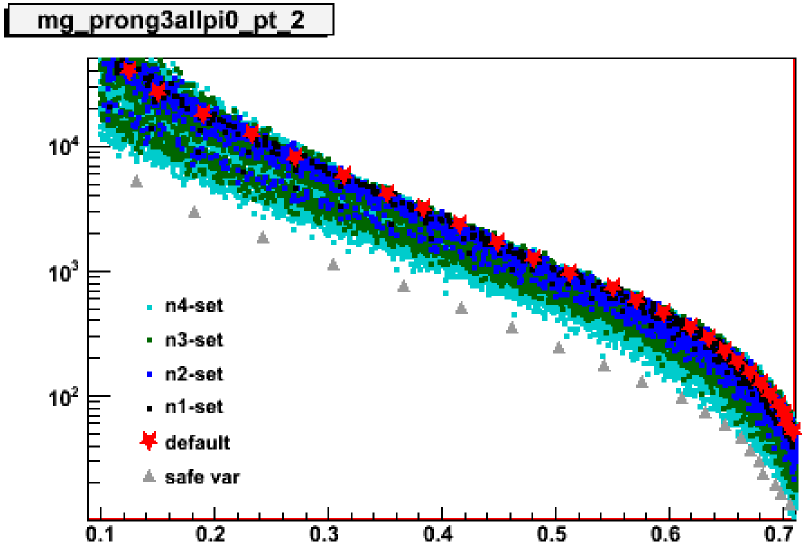
(these are cumulative)

Best variables: Variables removed from worst-performing n5-combination. Top 4 were also the ones removed from worst n4, same with top 3 for worst n3, etc...
Worst variables: Similarly but with the best-performing combinations.

3 prong







Best	Worst
trFlightPathSig	emRadius
nAssocTracksIsol	etEM2etTracks
etTracks2et	rWidth2Trk3P
etEM2etTracks	isoFrac

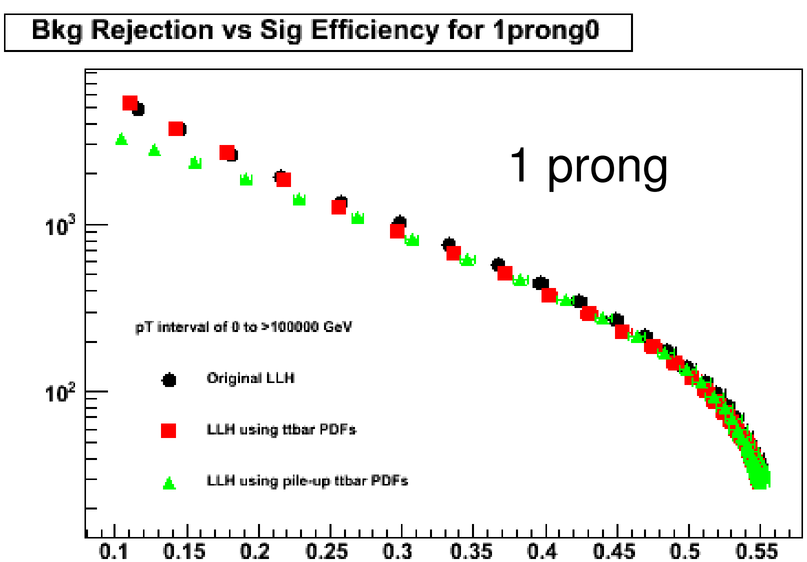
(cumulative)

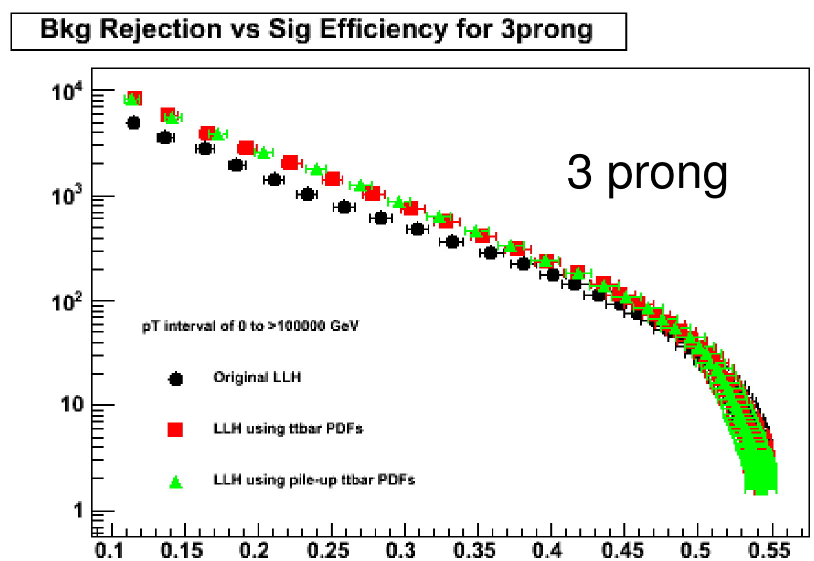
Pile-up

- Pile-up: L=10^{31}, 450 ns
- Produced PDFs from pile-up samples (both tau and jets)
  - Tau: 40k ttbar$\rightarrow$bqqb$\tau\nu$ (privately produced, Sherpa, reco 14.2.25.8)
  - Jets: 120k 105010, 105012, 105013 (centrally produced, reco 14.2.25.8)
- Tested on:
  - Tau: 40k ttbar$\rightarrow$bqqb$\tau\nu$ (privately produced, Sherpa, reco 14.2.25.8)
  - Jets: 1.1M 105010, 105012, 105013 (centrally produced, reco 14.2.25.8)

Pile-up

- Also for pile-up the improvement is very marginal, if at all
  - Difference at low efficiencies for 1 prong not really understood

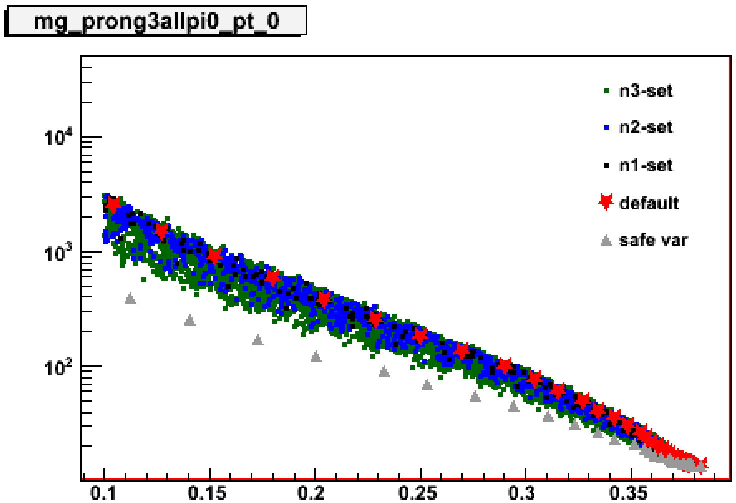


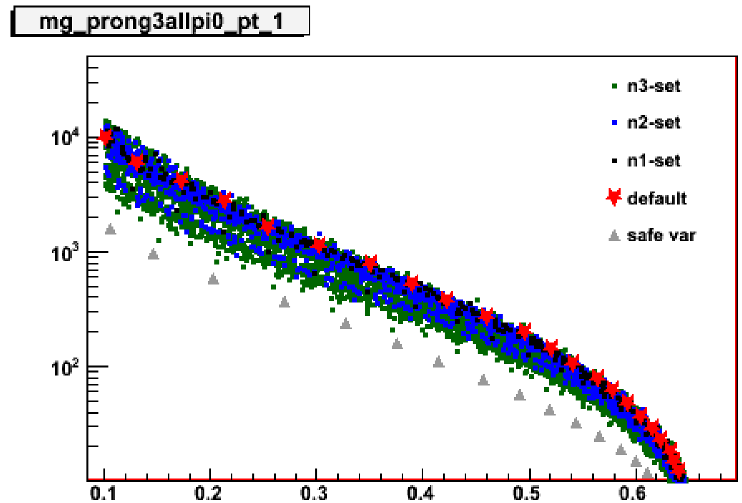


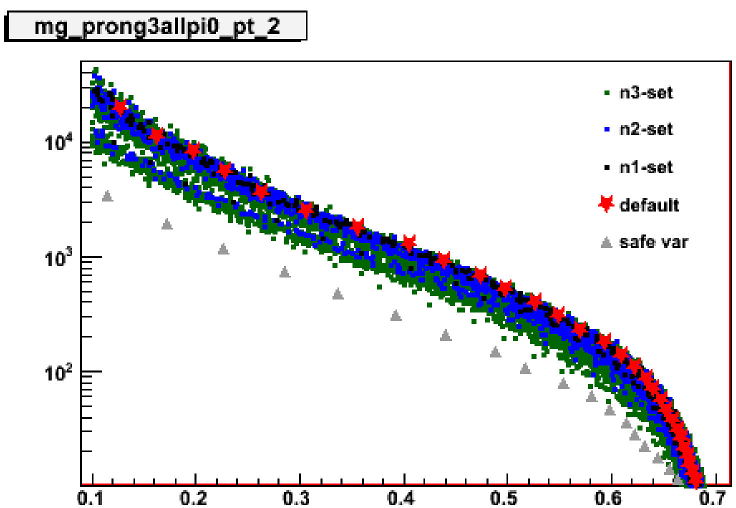
- Repeat study of variable contribution for pile-up case

(a subset of the "test" sample was used for these plots, for technical reasons)

Pile-up







1p		3p	
Best	Worst	Best	Worst
nAssocTracksIsol	etEM2etTracks	trFlightPathSig	emRadius
isoFrac	mVisEflow	etTracks2et	dRmax
emRadius	numStrip	etEM2etTracks	etEM2etTracks

(cumulative)

Summary & Outlook

UPPSALA UNIVERSITET

- Studying performance of tau identification in busy environments (ttbar)
- Produced PDFs from ttbar samples → very marginal improvement
- Have studied which variables contribute most/least
  - Can get very similar performance to default configuration with fewer variables
- Next step: Add new variables and study effects
  - Machinery for that already in place!
- First look at pile-up: situation seems quite similar, but need to look closer.

Backup

Definitions

UPPSALA UNIVERSITET

- Efficiency:
  - (# truth-matched reco taus passing cut) / (# true taus in kinematic region with corresponding # prongs)
    - Truth-matching: within $\Delta R<0.2$, with matching # prongs (1, >1) & # $\pi 0$:s (0, >0)
- Rejection:
  - $(1-\varepsilon_{jets}) / \varepsilon_{jets}$
    - ε_{jets}= (#reco taus matched to true jet passing cut)/(# true jets in kinematic region)

UPPSALA UNIVERSITET

Best/worst variables comparison

No Pile-up

1 prong

Best	Worst
nAssocTracksIsol	etEM2etTracks
stripWidth2	numStrip
emRadius	etHad2etTracks

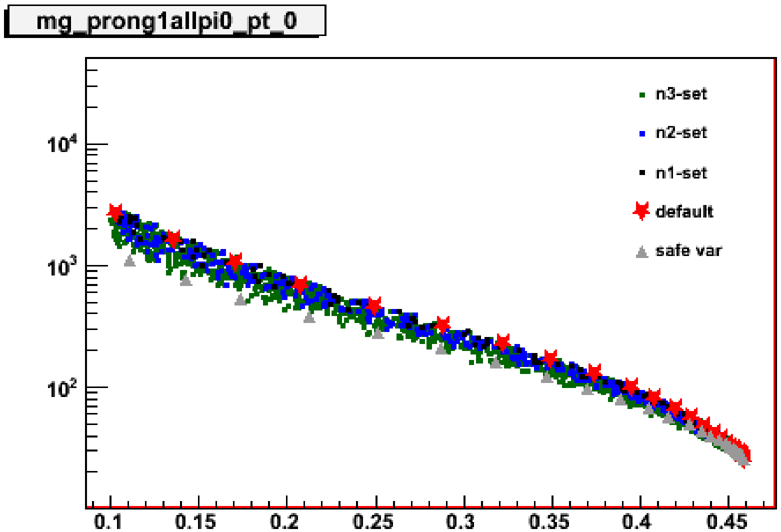
3 prong

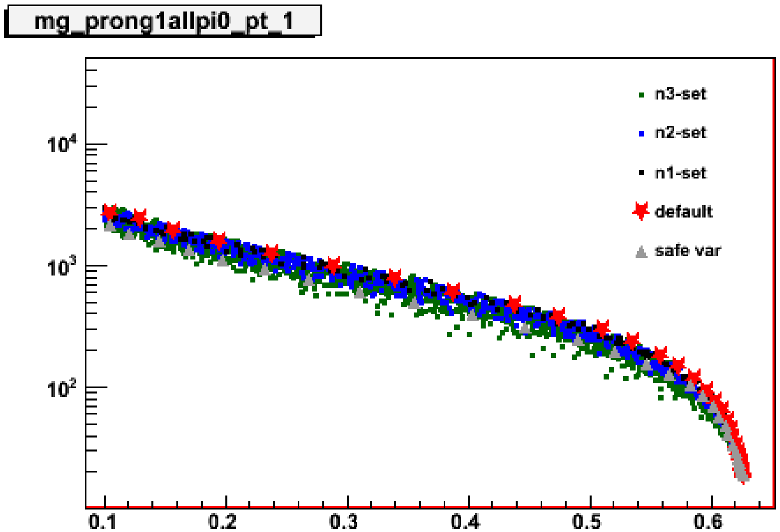
Best	Worst
trFlightPathSig	emRadius
nAssocTracksIsol	etEM2etTracks
etTracks2et	rWidth2Trk3P

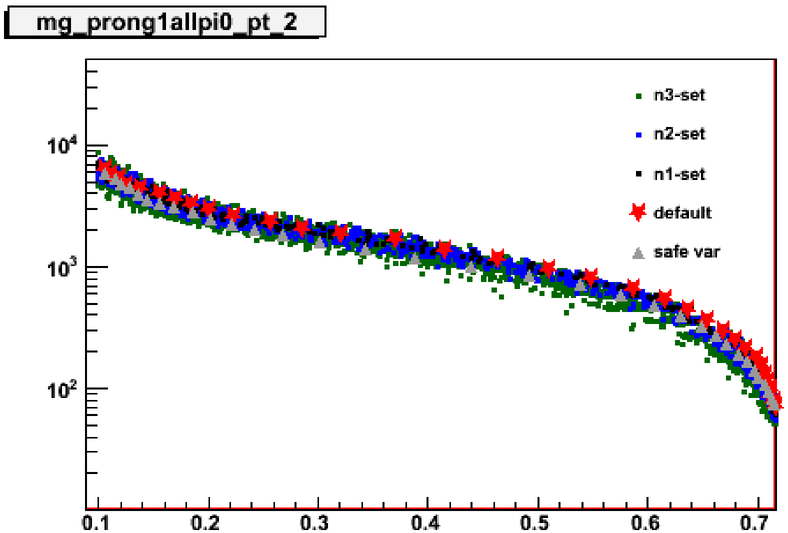
Pile-up

1p		3p	
Best	Worst	Best	Worst
nAssocTracksIsol	etEM2etTracks	trFlightPathSig	emRadius
isoFrac	mVisEflow	etTracks2et	dRmax
emRadius	numStrip	etEM2etTracks	etEM2etTracks

Pile-up 1 prong







UPPSALA UNIVERSITET

Variables

var	name	1p	3p
0	emRadius	X	X
1	isoFrac	X	X
2	stripWidth2	X	X
3	numStrip	X	X
4	etHad2etTracks	X	X
5	etEM2etTracks	X	X
6	etTracks2et	X	X
7	etStrip2et		
8	etRatio		
9	dRmin		X
10	dRmax		X
11	rWidth2Trk3P		X
12	massTrk3P		X
13	nAssocTracksIsol	X	X
14	MVisEflow	X	X
15	z0SinThetaSig	X	
16	SignD0Trk3P	X	X
17	trFlightPathSig		X

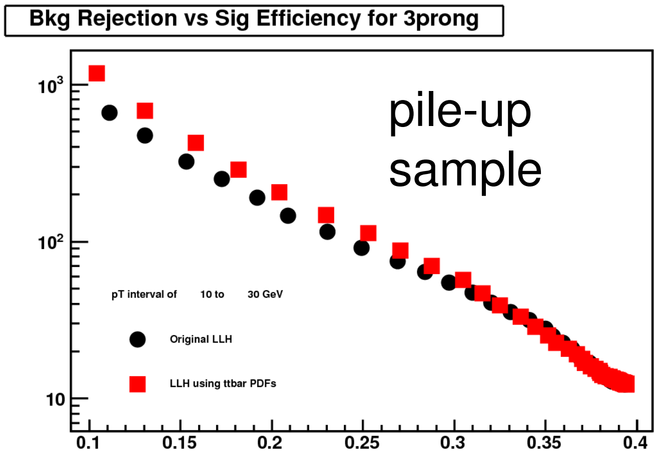
Safe Variables

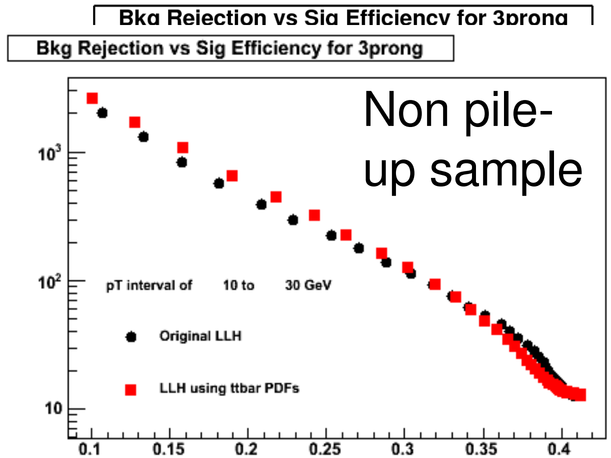
		calo		track		sum	
var	**name**	**1p**	**3p**	**1p**	**3p**	**1p**	**3p**
0	emRadius	S	S			S	S
1	isoFrac	S	S			S	S
2	stripWidth2	S	S			S	S
3	numStrip						
4	etHad2etTracks			S	S	S	S
5	etEM2etTracks			S	S	S	S
6	etTracks2et			S	S	S	S
7	etStrip2et						
8	etRatio						
9	dRmin						
10	dRmax						
11	rWidth2Trk3P				S		S
12	massTrk3P						
13	nAssocTracksIsol						
14	MVisEflow						
15	z0SinThetaSig						
16	SignD0Trk3P						
17	trFlightPathSig						
X	etEM2Et			S	S	S	S
Y	etOverPtLeadTrk			S	S	S	S

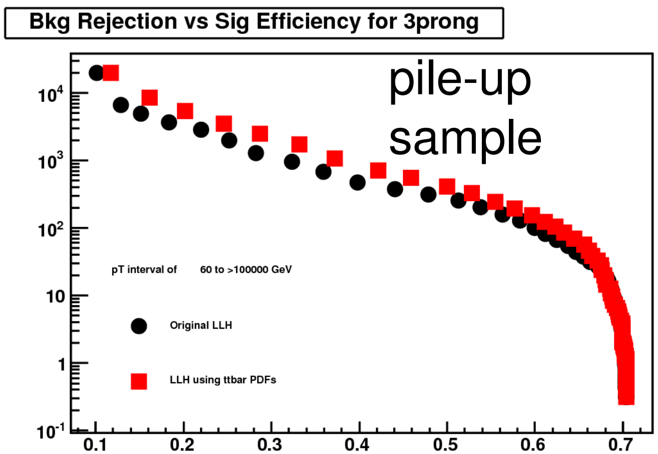
14 TeV samples used

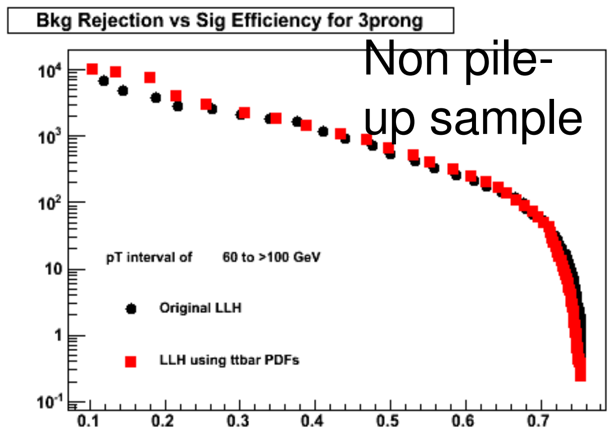
- Z: 25k, privately produced, Sherpa, 14.2.10.1+14.2.20.3
- ttbar: privately produced, Sherpa, 14.2.10.1+14.2.20.3
  - 50 k for PDF production
  - 50 k for testing
- QCD: 120k, central production, 5012 & 5013
- For more details on the 14 TeV plots, please look at previous talks at Tau WG meetings (2/10/08, 18/11/08 & 3/3/09)

Pile-up at 14 TeV









TauRec performance for
- default tauRec2008 pdfs
- ttbar→tau+jets pdfs

(in the progess of producing pile-up pdfs...)

NB: comparison non-trivial: For pile-up, the computed rejection combines jets from the hard dijet event and from minimum bias events

(could be “easier” to reject the latter, due to different quark/gluon content, pt spectrum, rapidity, ...)