Short-tracks investigation continued

A. Rakitin
Lancaster University

September 27, 2007
Tralgo Meeting

http://www-d0.fnal.gov/~rakitin/d0_private/tex/2007.Sep.27.Tralgo/tr.pdf

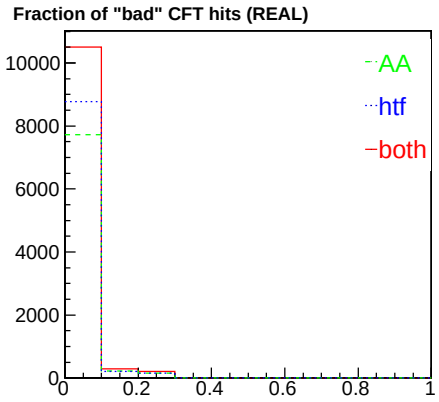
Disclaimer

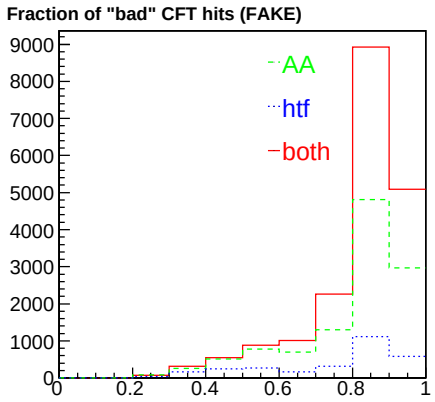
- This talk is based on my last talk on Tralgo meeting on September 6
- Some plots and numbers are old, some are new

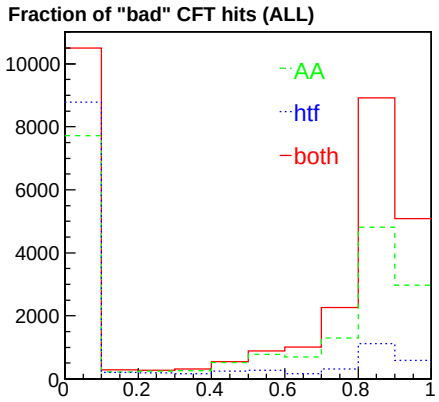
Outline

- I used one of Mike Hildreth's high-luminosity MC files
  /prj_root/1403/www_algo/mikeh/qcd_minbias_smtmod/dest/d0reco/
  d0reco_p20.08.02_NumEv-0_poiss15_mcp17_00015_07103224449
  to look at AA and HTF tracks separately
- Created three TMB files (250 events each):
  - Reco'ed with AA only
  - Reco'ed with HTF only
  - Reco'ed with HTF+AA (standard reconstruction)
- For each track looked at MC information for SMT and CFT hits
- Determined MC particle which made the track as the particle which left the majority of hits
- Identified "bad" hits coming from other particles (of from nowhere)
- Called tracks with > 25% bad CFT hits - "fake" (this is a bit different from my last talk definition: >= 8 bad CFT tracks)
- All the other tracks are real
- Tracks with zero denominator (i.e. numCFThits==0, i.e. SMT-only tracks) are real too
- Plotted a few distributions for fake/real tracks reco'ed with AA/HTF

Fraction of “bad” CFT hits

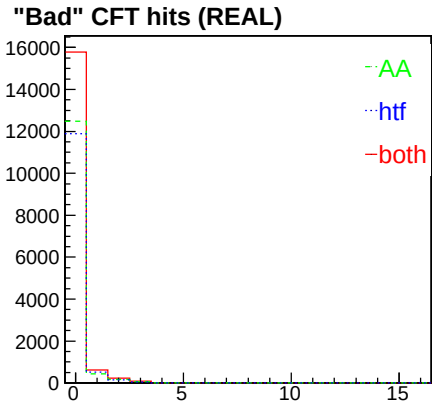


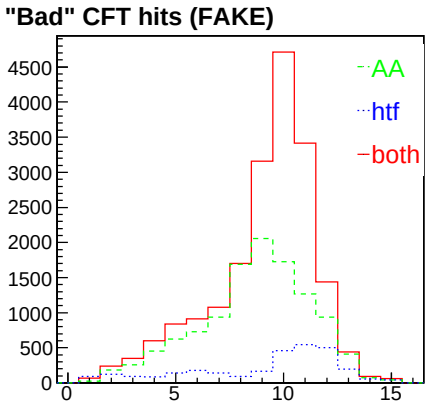


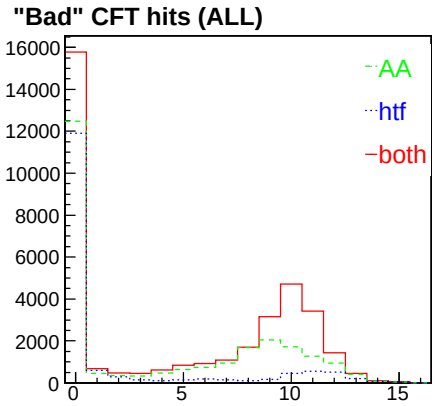


- In this plot we see a problem - many tracks with high fraction of “bad” CFT hits
- Both AA and HTF create such tracks
- Based on this plot I cut at > 25% “bad” CFT hits to distinguish real and fake tracks
- The threshold seems to be the same for AA, HTF and their combination

Number of “bad” CFT hits



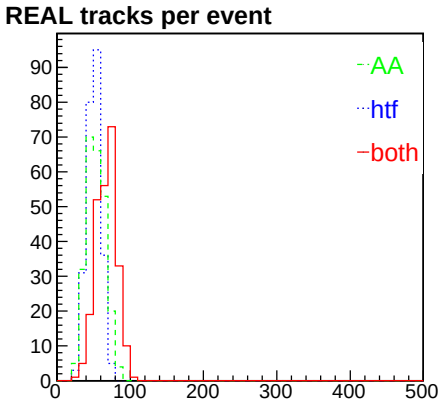


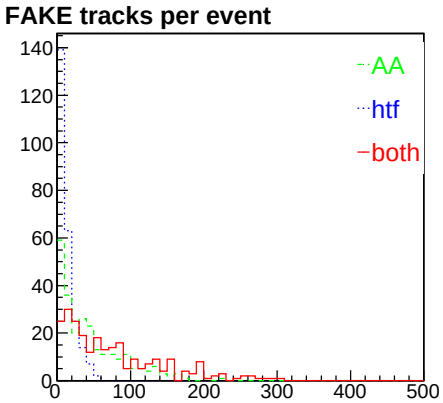


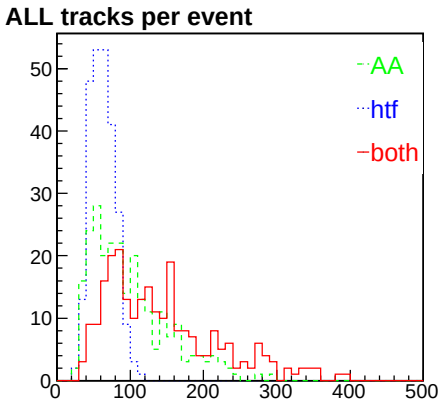
- The cut on *fraction* of “bad” CFT hits separates tracks with high *number* of “bad” CFT hits quite well

Number of tracks

- Using this defi nition of fake/real tracks plot the number of tracks per event







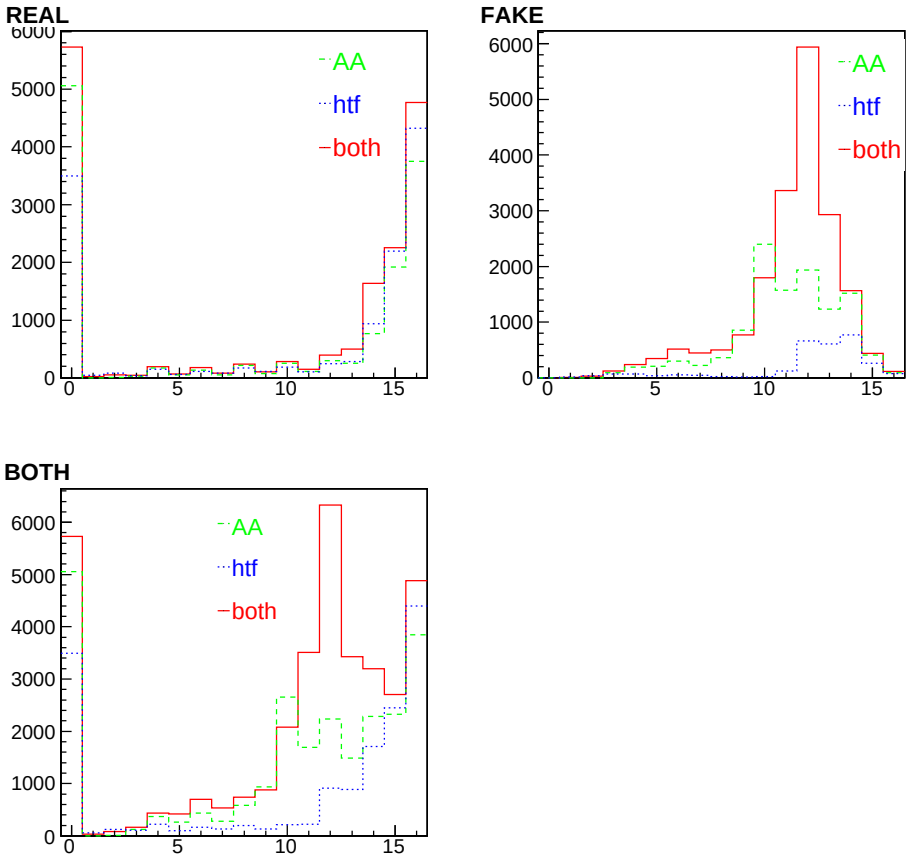
- AA produces more tracks than HTF
- AA produces more fake tracks than HTF

Number of tracks

Algorithm	Real tracks	Fake tracks	All tracks	Real tracks ($\|\eta\| < 1$)	Fake tracks ($\|\eta\| < 1$)	All tracks ($\|\eta\| < 1$)
AA	13158	11432	24590	4557	5237	9794
HTF	12622	2921	15543	4762	1144	5906
AA+HTF	16717	19129	35846	6013	7462	13475

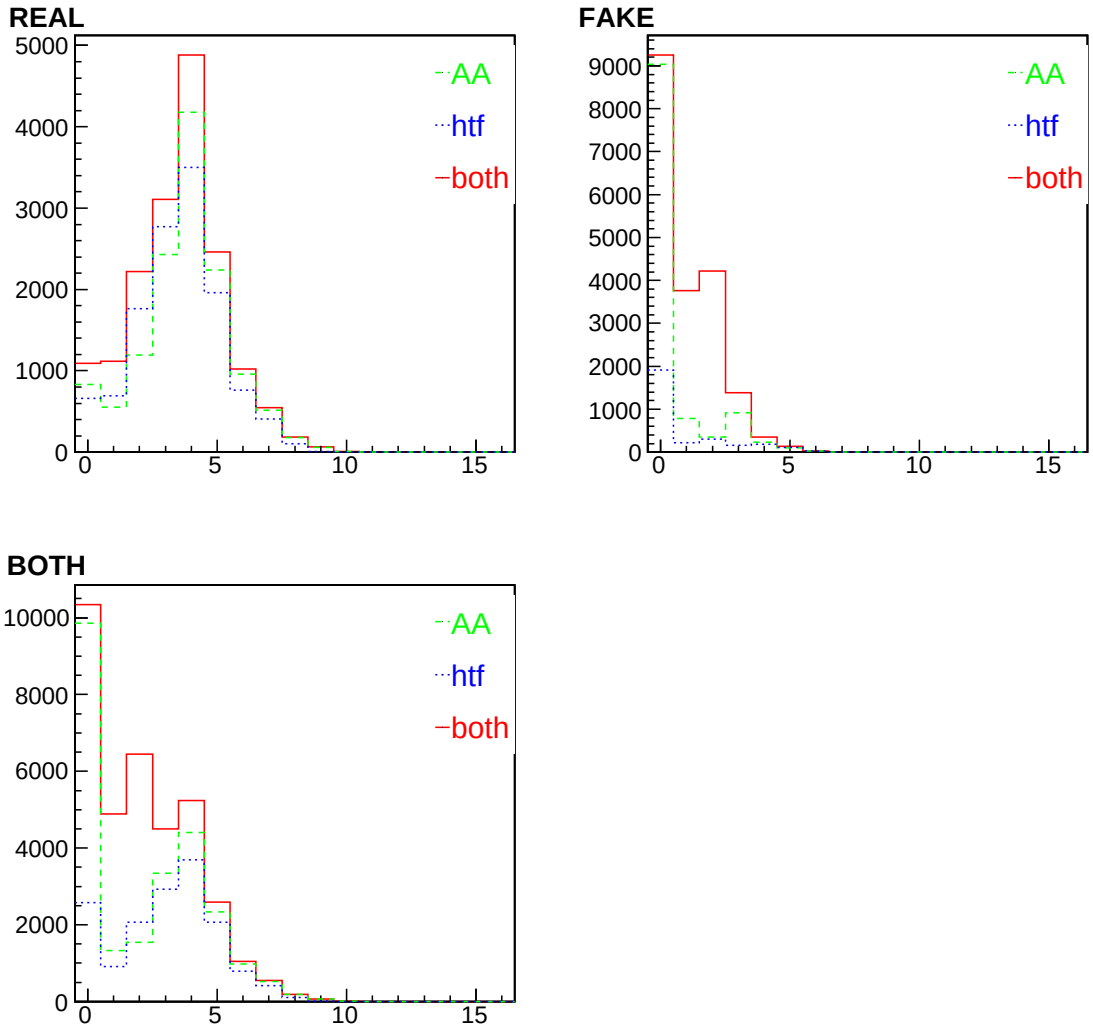
- Both algorithms produce approximately equal amount of tracks
- Both algorithms produce fake tracks
- Number of fake tracks produced by AA+HTF is higher than number of fake tracks produced by AA and HTF separately
- The percentage of fake tracks is very high (53%)
- AA produces factor of 4 more fake tracks than HTF
- Looking at only central region ($|\eta| < 1$) does not change the proportions

CFT hits:



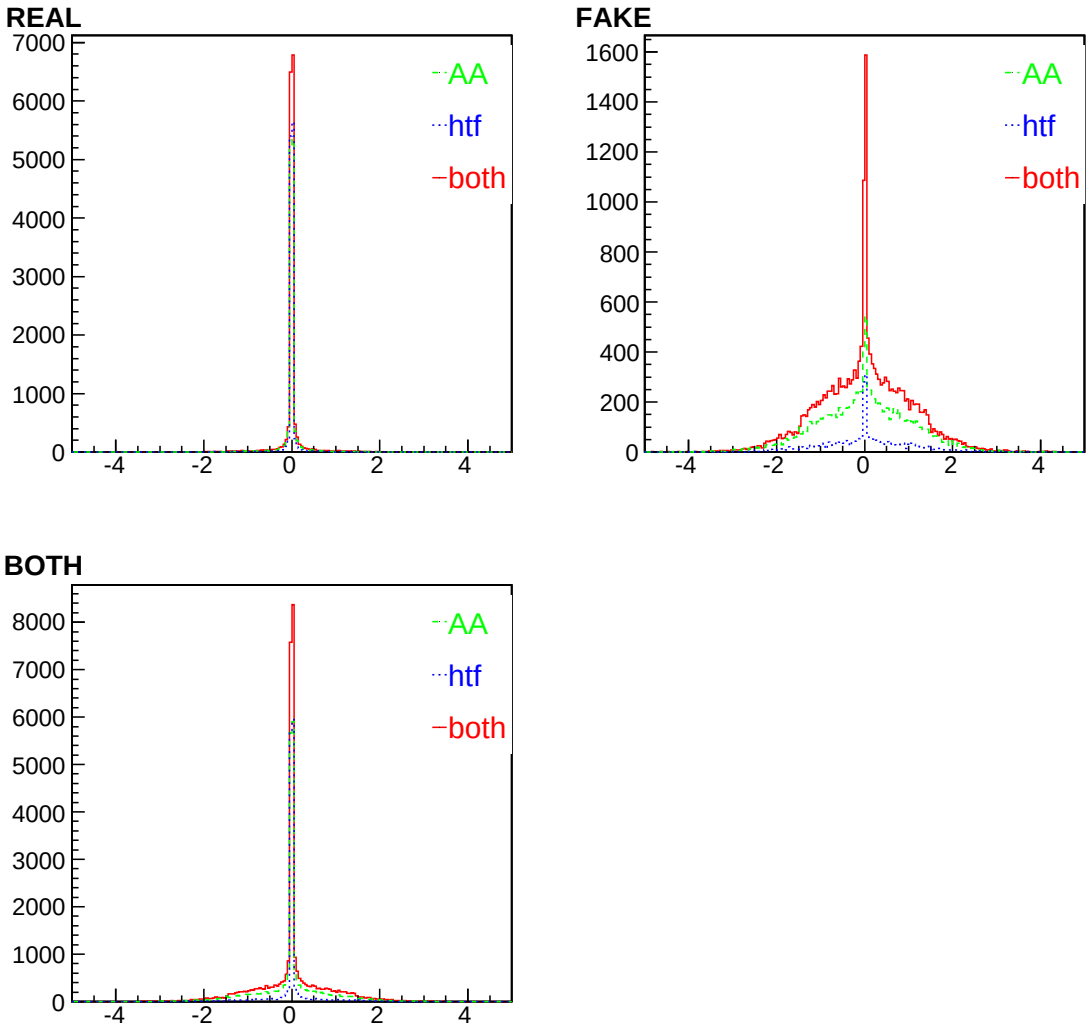
- Fake tracks have >= 10 and < 14 CFT hits
- Real tracks have >= 14 CFT hits
- So, one can separate them by imposing cut numCFThits>= 14
- Well-known spike at 12 comes from fake tracks
- **Neither algorithm alone is responsible for this spike**

SMT hits:



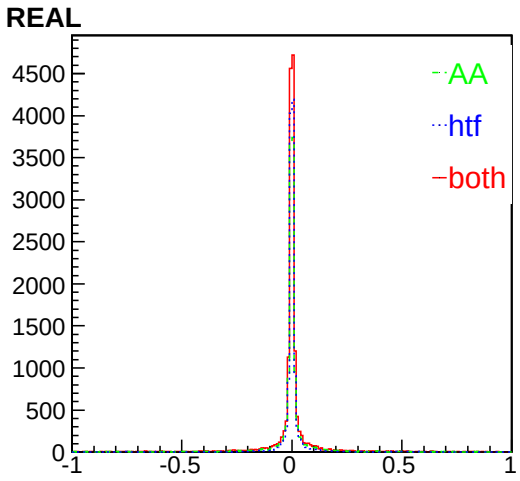
- Fake tracks have < 3 SMT hits (i.e. mostly CFT-only)

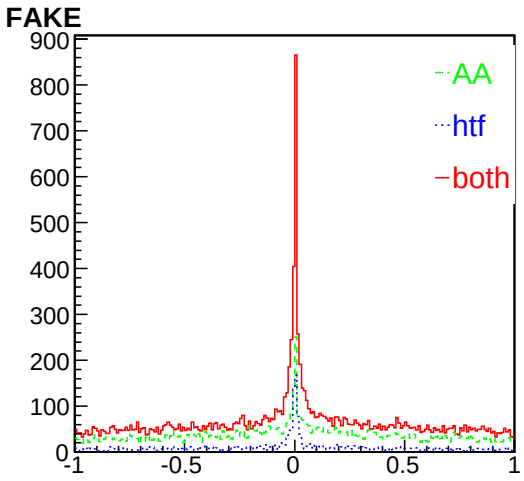
Impact parameter distribution:

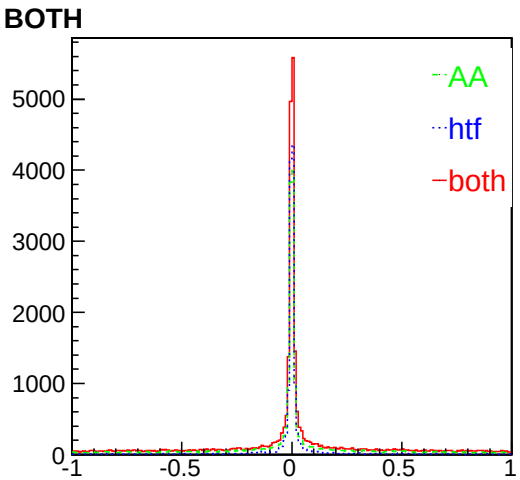


- Fakes have much broader impact parameter (d_0) distribution
- So, one can separate fake and real tracks by imposing corresponding cut

Impact parameter distribution:

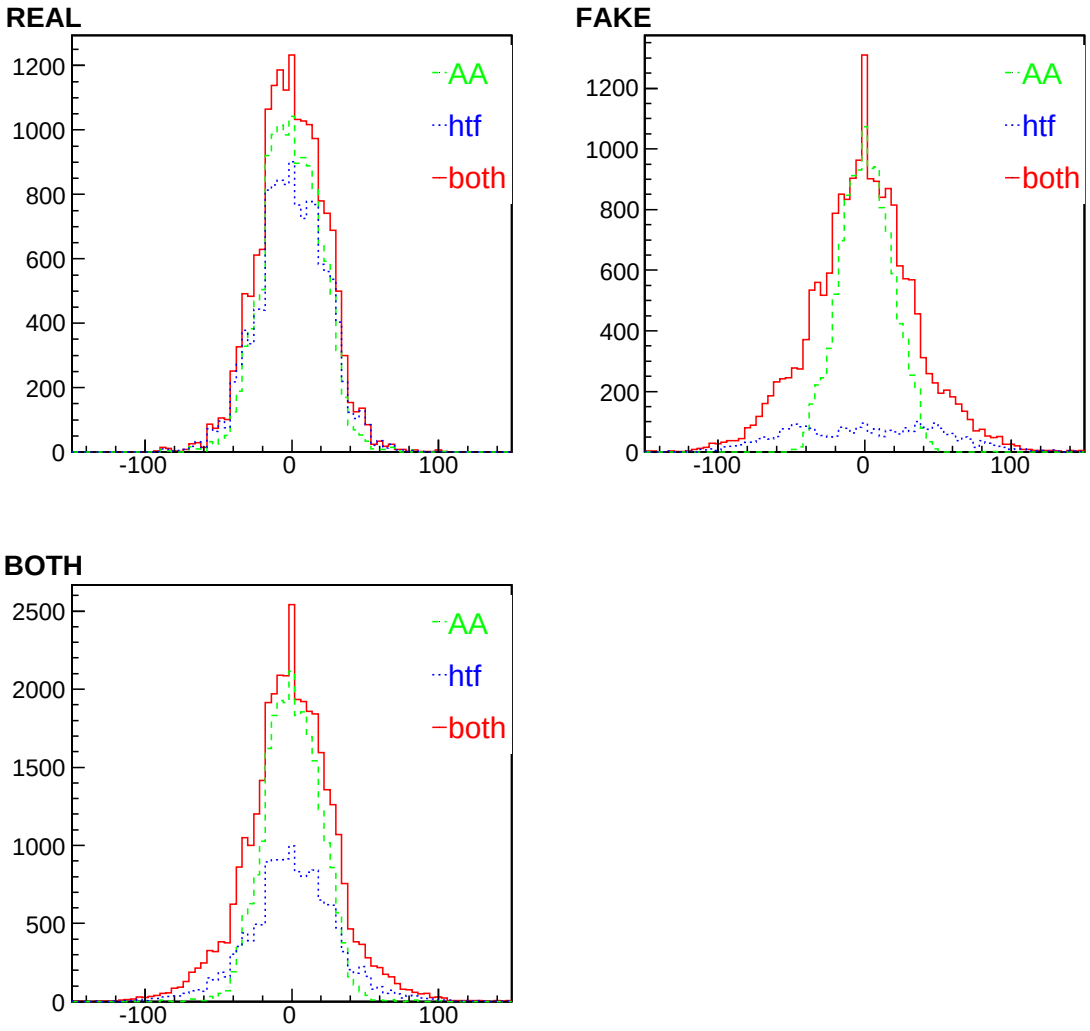






- Same plots zoomed in
- Let's make a cut on impact parameter d_0 e.g $|d_0| < 0.1$ (not really optimized)

z_0 distribution:



- Fakes have much broader z_0 distribution because of HTF part
- Explanation - see next page

AA and HTF interplay

According to Guennadi:

- AA alone finds only primary vertices inside SMT acceptance
- AA together with HTF uses primary vertices found by HTF outside SMT acceptance as well
- This is why combination of two algorithms finds more tracks than individual algorithms
- This is why z_0 distribution looks this way

My suggestion:

- Reduce the number of fake tracks by cutting on z_0 as well...

Number of tracks

- So, I impose cut on numCFThits>= 14 and impact parameter $|d_0| < 0.1$ saving cut on z_0 for future investigation

	Algorithm	Real tracks	Fake tracks	All tracks	Real tracks ($\|\eta\| < 1$)	Fake tracks ($\|\eta\| < 1$)	All tracks ($\|\eta\| < 1$)
Before cuts	AA	13158	11432	24590	4557	5237	9794
	HTF	12622	2921	15543	4762	1144	5906
	AA+HTF	16717	19129	35846	6013	7462	13475
After cuts	AA	9871	248	10119	3663	149	3812
	HTF	9905	147	10052	4002	71	4073
	AA+HTF	12201	241	12442	4611	129	4740

- Cuts diminish the number of real tracks by 22-27%
- Number of fake tracks almost vanishes
- May be we need to optimize the cuts better...

Conclusion:

Observed facts:

- I defi ne “fake” tracks as tracks having > 25% bad CFT hits
- There are 53% of such tracks
- AA produces factor of 4 more fake tracks than HTF, both in central ($|\eta| < 1$) and forward ($1 < |\eta| < 2$)
- Well-known spike at 12 CFT hits comes from fake tracks, but neither algorithm is solely responsible for it
- Fake tracks are mostly CFT-only tracks
- It is possible to distinguish between fake and real tracks using the number of CFT hits (>= 14)
- Fake tracks have much broader impact parameter distribution than real tracks
- It is possible to distinguish between fake and real tracks by imposing cut on impact parameter ($|d_0| < 0.1$)
- Application of both cuts reduce the fraction of fake tracks from 53% to 2% while keeping 73% of real tracks
- May be we need to optimize the cuts better, including cut on z_0