Run 3 Electron ID for HZZ analysis

Andro Petković for CJLST group

Training setup

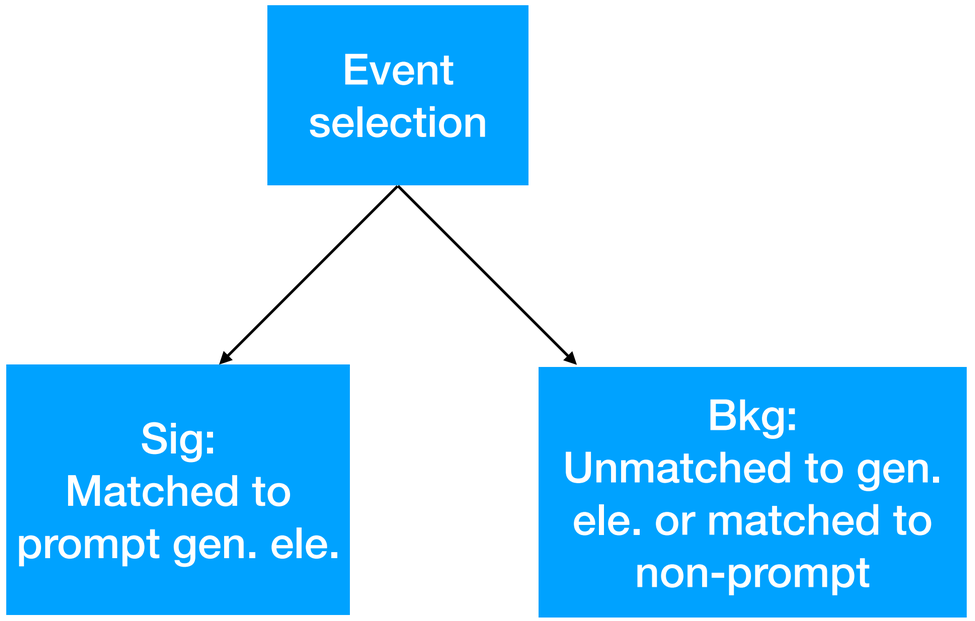
BDT input

→ **Run2**

For Run3 we switched from particle to cluster-based isolation variables

ele_ecalPFClusterIso
ele_hcalPFClusterIso
ele_dR03TkSumPt

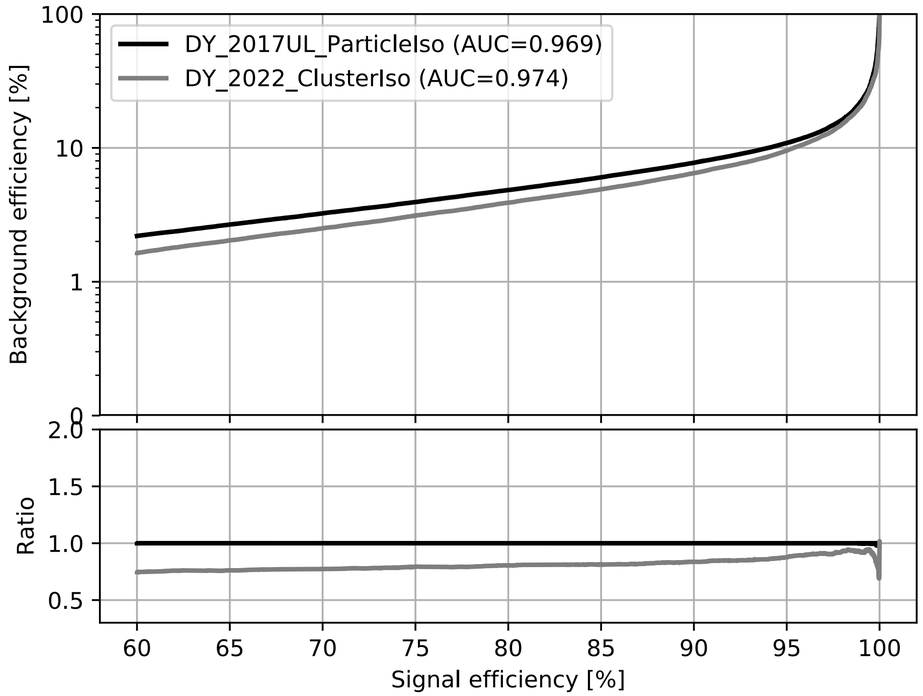
- **Use DY MC to target HZZ analysis**
- **Split training into 6 categories based on electron p_T and η**



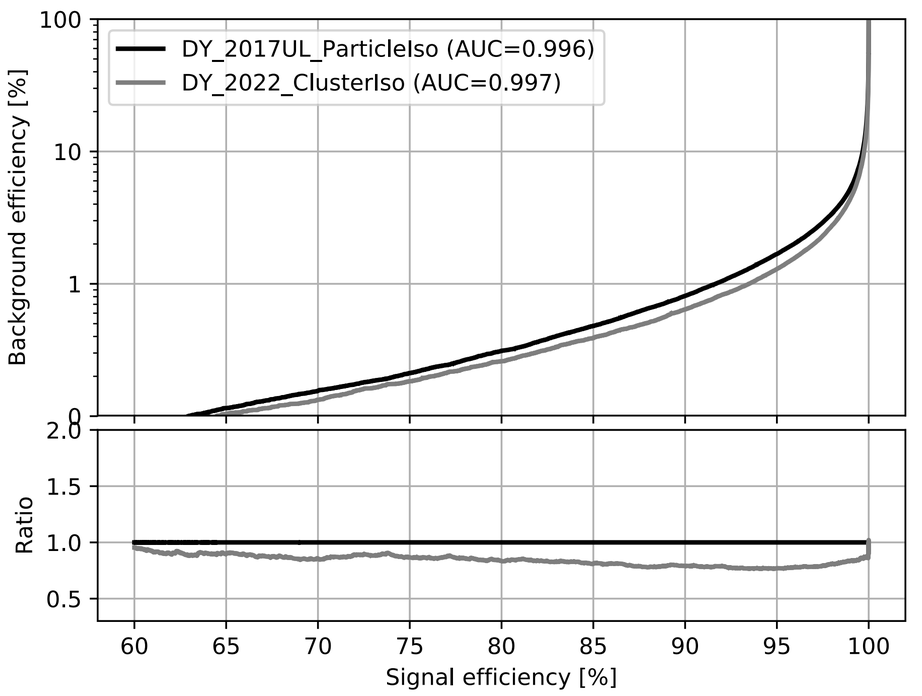
Training workflow and datasets

- **Reuse training setup from Run 2 (xgbo+bayesian optimization)**
- **Compare 2017UL vs 2022 results to reproduce previous results and check if machinery works (check Ana's presentation here)**
- **Compare Run2 vs Run3 training variable distributions**
- **2017UL: /DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/ RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9_ext1-v1**
- **2022 - 3 datasets merged: /DYJetsToLL_M-50_TuneCP5_13p6TeV-madgraphMLM-pythia8/ Run3Winter22MiniAOD-122X_mcRun3_2021_realistic_***

ROC curve comparison, inner barrel

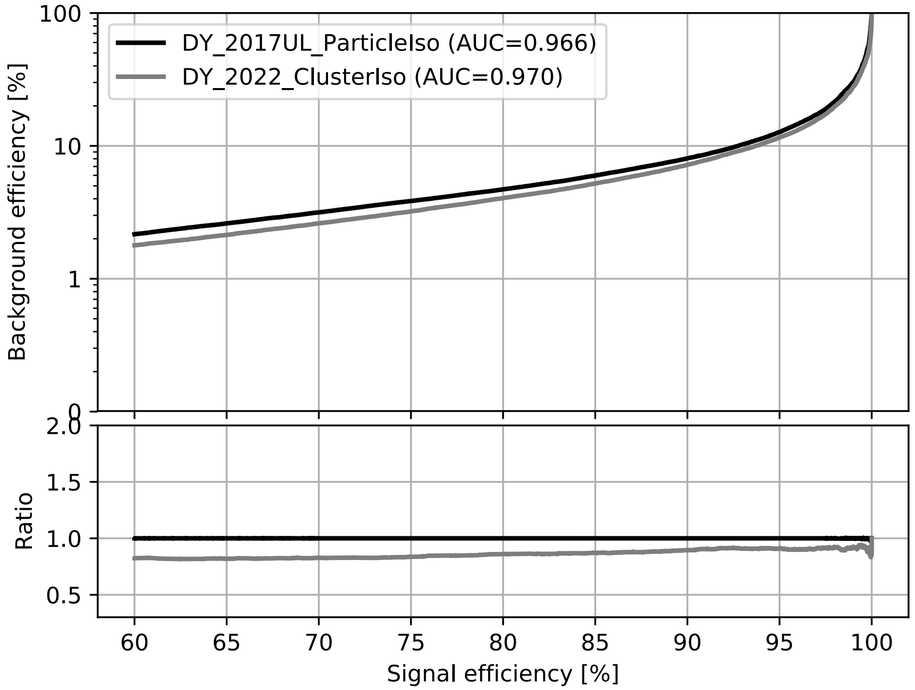


p_T < 10 GeV

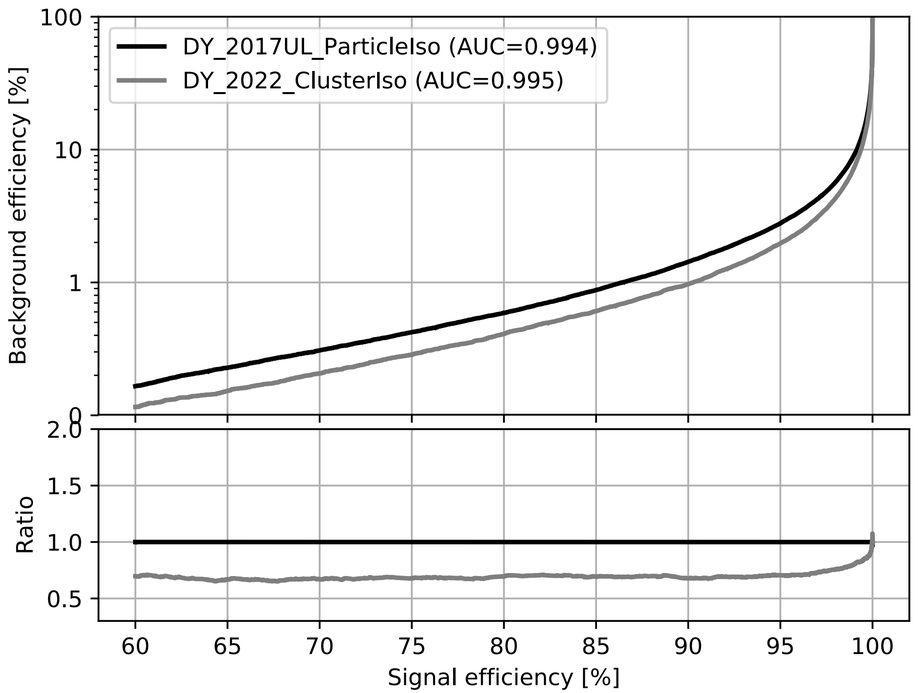


p_T > 10 GeV

ROC curve comparison, outter barrel

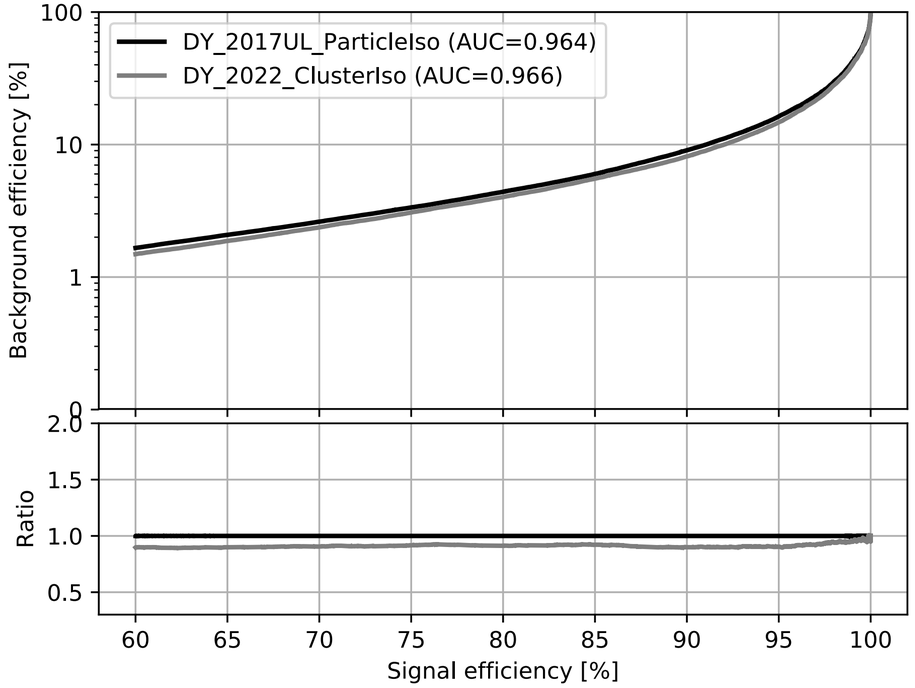


p_T < 10 GeV

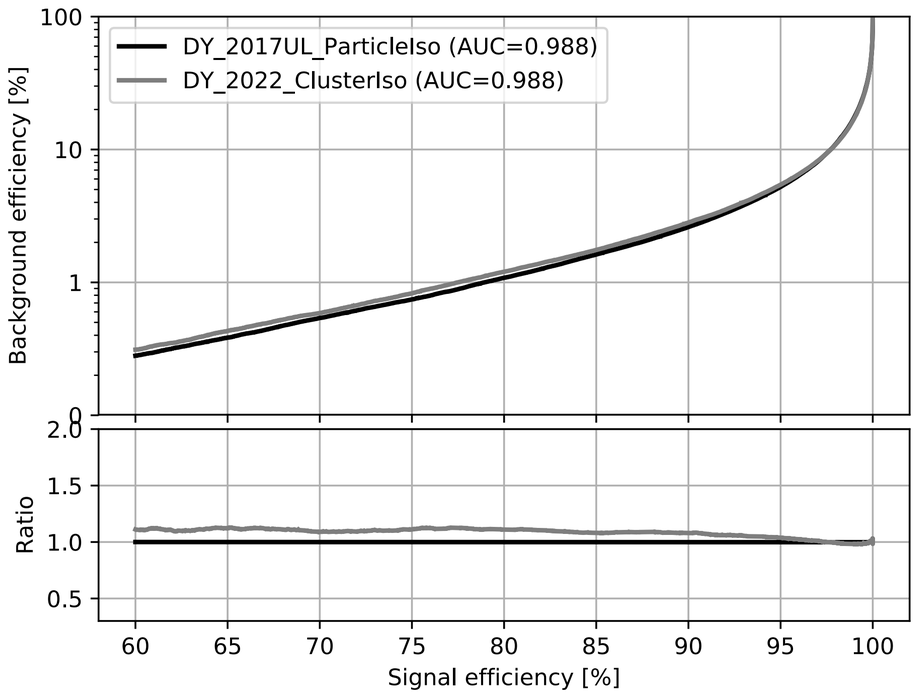


p_T > 10 GeV

ROC curve comparison, endcap



p_T < 10 GeV



p_T > 10 GeV

2017UL vs 2022 for target signal efficiency

Region	Sig. eff. target	2017UL bkg. eff.	2022 bkg. eff. particle iso	2022 bkg. eff. cluster iso
EB1_5	81.64 %	5.18%	5.89%	4.2%
EB1_10	97.44%	2.87%	2.84%	2.27%
EB2_5	80.31%	4.77%	5.63%	4.11%
EB2_10	96.68%	3.92%	3.48%	2.83%
EE_5	74.38%	3.25%	3.99%	2.97%
EE_10	96.62%	7.36%	8.46%	7.46%

EB1_5:
$5\ \textbf{GeV} < p_T < 10\ \textbf{Gev}, |\eta| < 0.8$

EB1_10:
$p_T > 10\ \textbf{Gev}, |\eta| < 0.8$

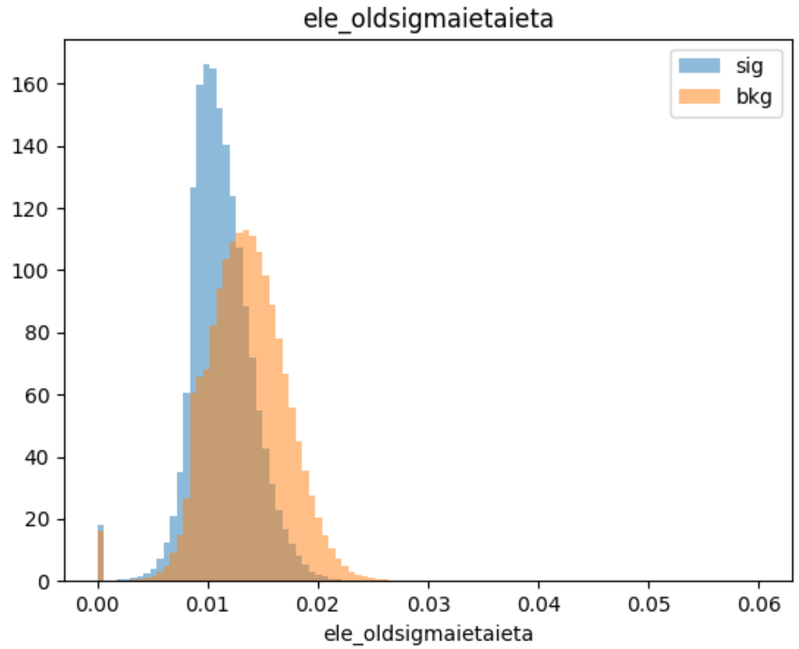
EB2_5:
$5\ \textbf{Gev} < p_T < 10\ \textbf{Gev},\ 0.8 < |\eta| < 1.479$

EB2_10:
$p_T > 10\ \textbf{Gev},\ 0.8 < |\eta| < 1.479$

EE_5:
$5\ \textbf{GeV} < p_T < 10\ \textbf{Gev}, |\eta| > 1.479$

EE_10:
$p_T > 10\ \textbf{Gev},\ |\eta| > 1.479$

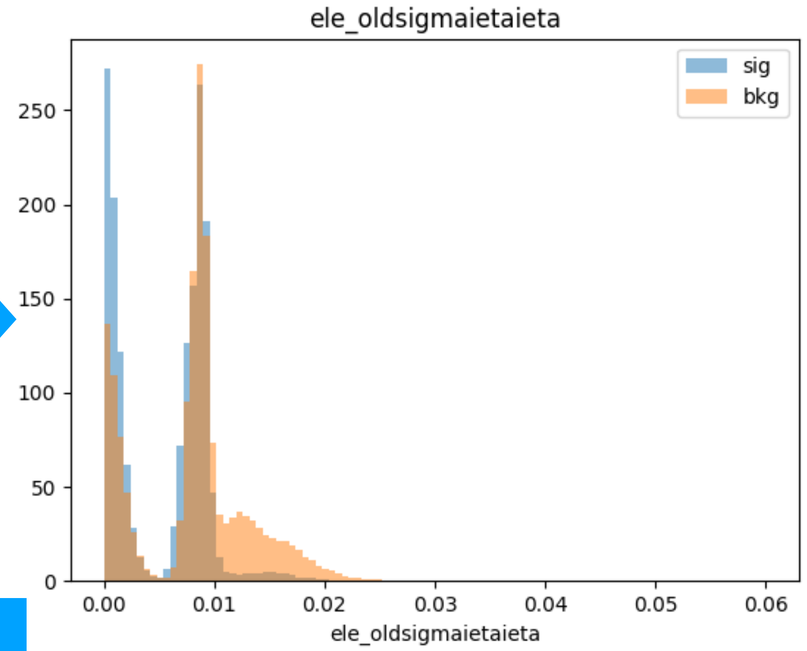
Training variable distributions



2017 UL, inner barrel, p_T < 10 GeV

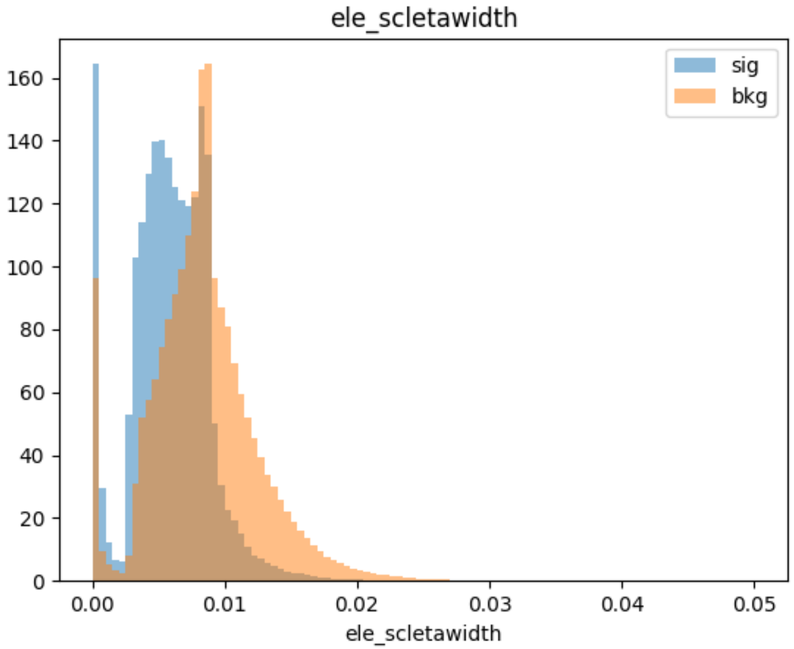
Crystal thresholds raised

Thresholds increased to cope with higher pileup -> more crystals have energy set to 0, check the details of redefinition here

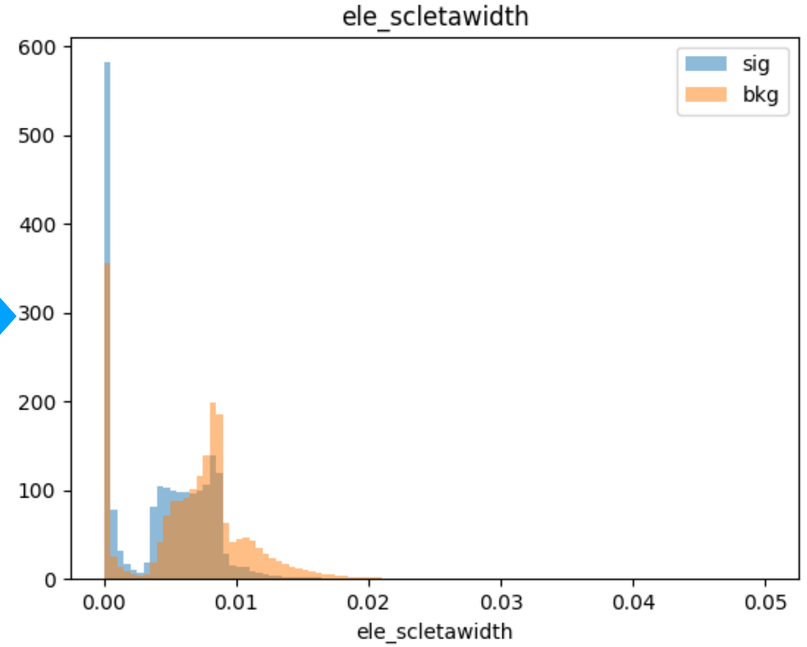


2022, inner barrel, p_T <10 GeV

Training variable distributions



Crystal thresholds raised



2017 UL, inner barrel, p_T < 10 GeV

2022, inner barrel, p_T < 10 GeV

Summary and outlook

- Shower shape distributions affected by raised thresholds
- Performance boost when switched from particle to cluster-based isolation variables
- Check distributions of all 2017 UL and 2022 training variables
- Add this BDT (version trained on cluster isolation variables) to next nanoAOD production
- Use in analysis targeting Run3 results