Regression Baseline Update

Dominick Olivito (UCSD)

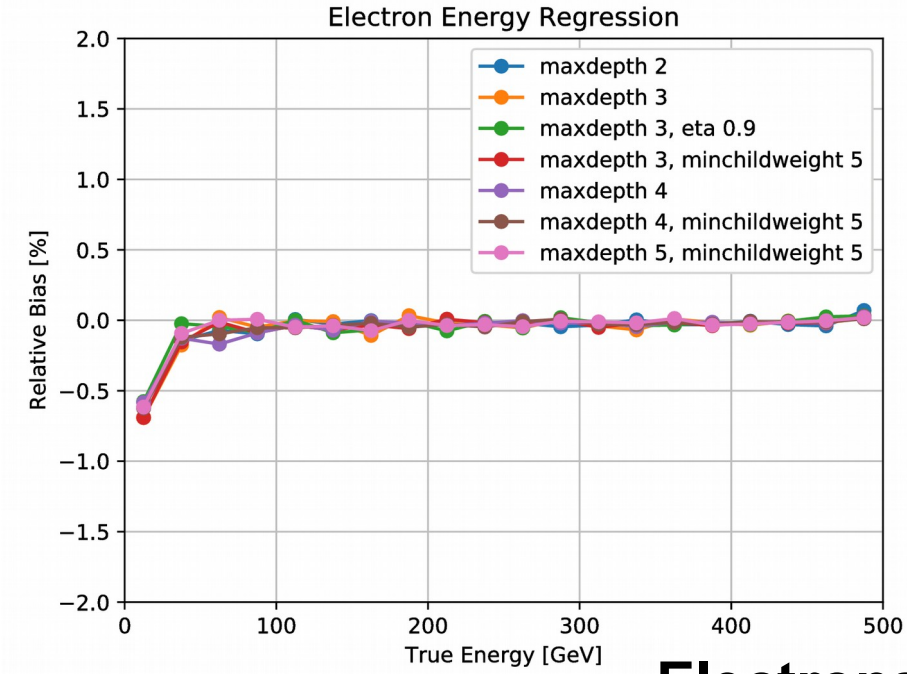
Overview

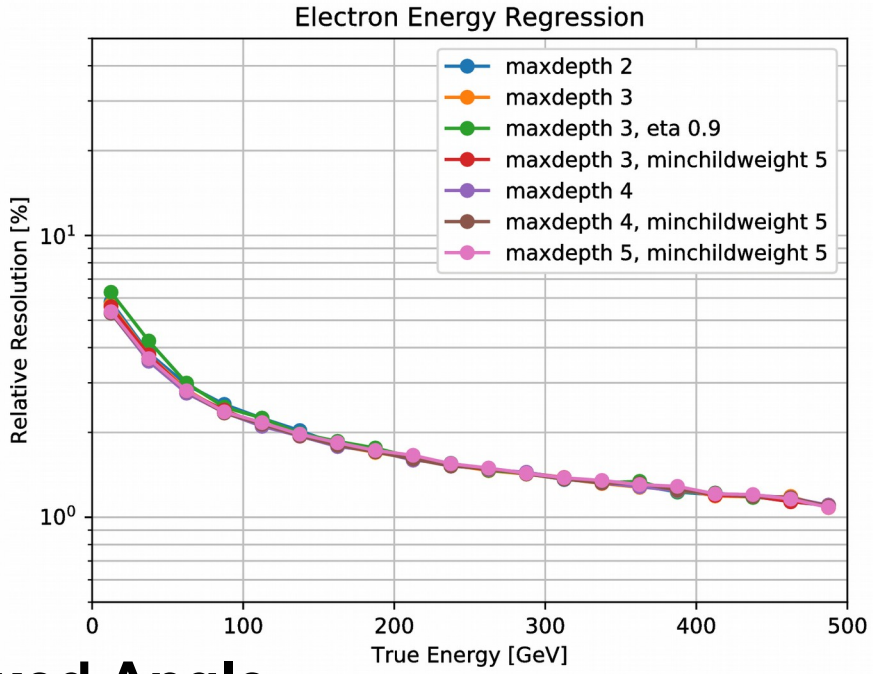
- Working on regression baseline using ECAL, HCAL energy sums and a few shower shapes

- Updates since last time:
  - Tried hyperparameter scan in xgboost, minimal changes
  - Re-discovered that Decision Trees can't extrapolate
  - Train on one particle type, test on another
  - First look at variable angle samples

Hyperparameters

- Tried relevant hyperparameters in xgboost:
  - maxdepth: maximum depth of an individual tree [was using 3]
  - minchildweight: min number of events to split [default: 1]
  - eta: learning rate [default: 0.3]
- No significant differences

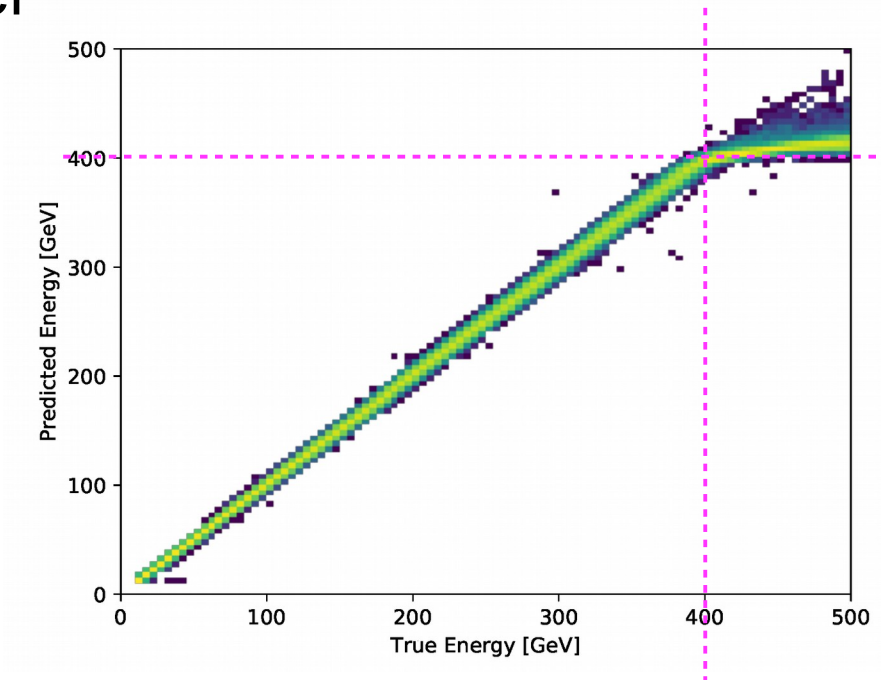




Electrons, Fixed Angle

Trees Can't Extrapolate

- Noticed that bias got slightly worse at 500 GeV, end of energy range
- Tried training with E < 400 GeV, predicting full energy range
- Re-discovered that Decision Trees can't extrapolate
  - In real experiment, wouldn't be able to directly use an algorithm like this
  - Could switch back to NN for a baseline, if people think this is a showstopper

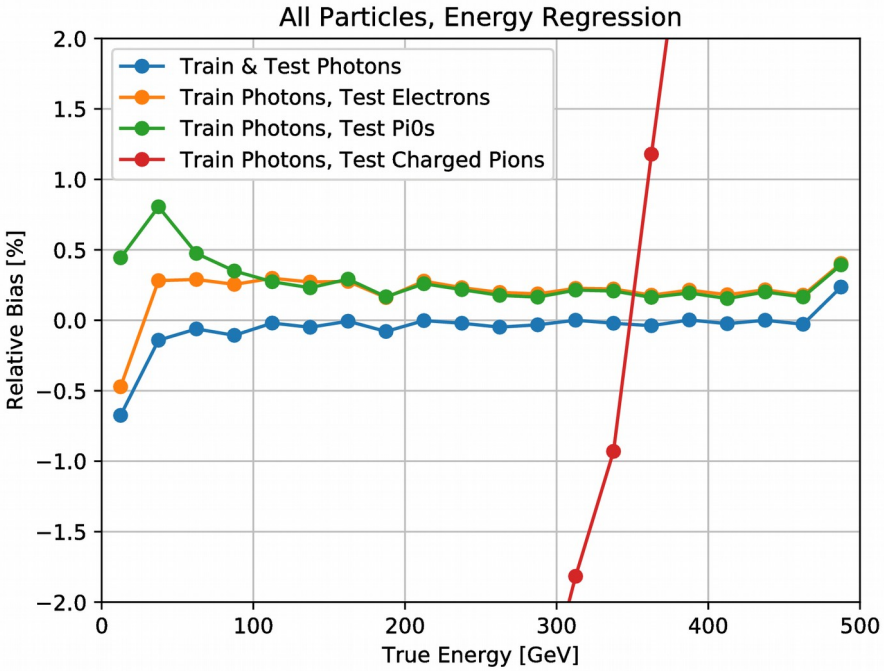


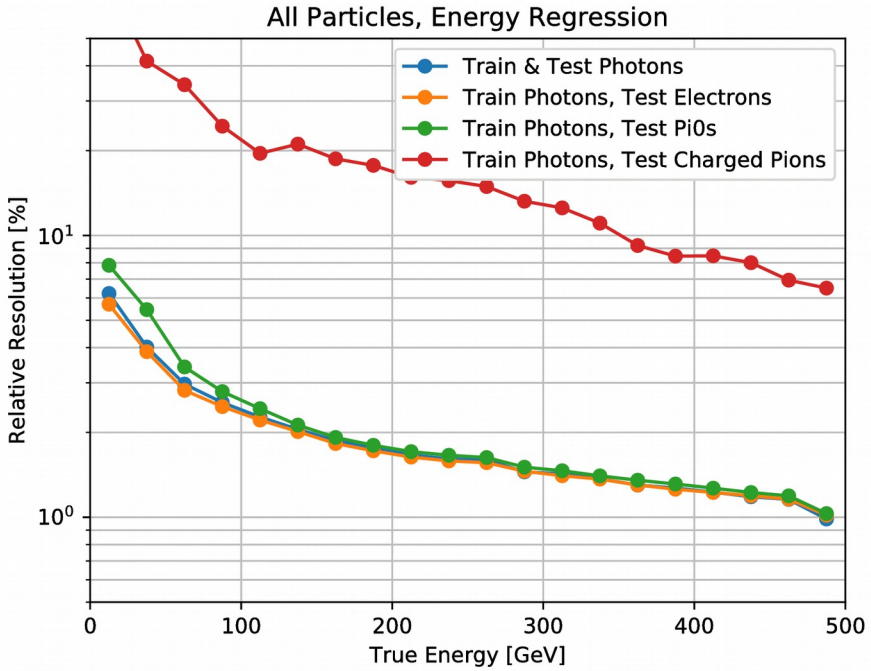
Xgboost never predicts values much higher than 400 GeV training cutoff

Electrons, Fixed Angle

Predicting Other Particles

- Train on photons (fixed angle), test on other particles
- Pretty good results for Electrons, Pi0s
  - Bias ~0.2% at high energy, resolution similar except at lowest energies for pi0
  - Compare to backup slide
- Charged pions much worse, as expected

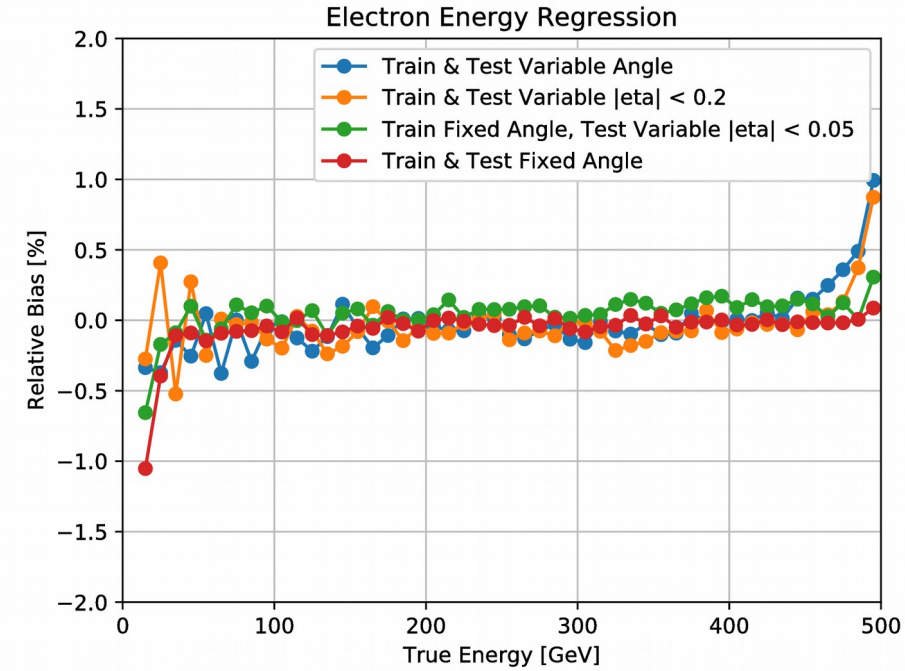


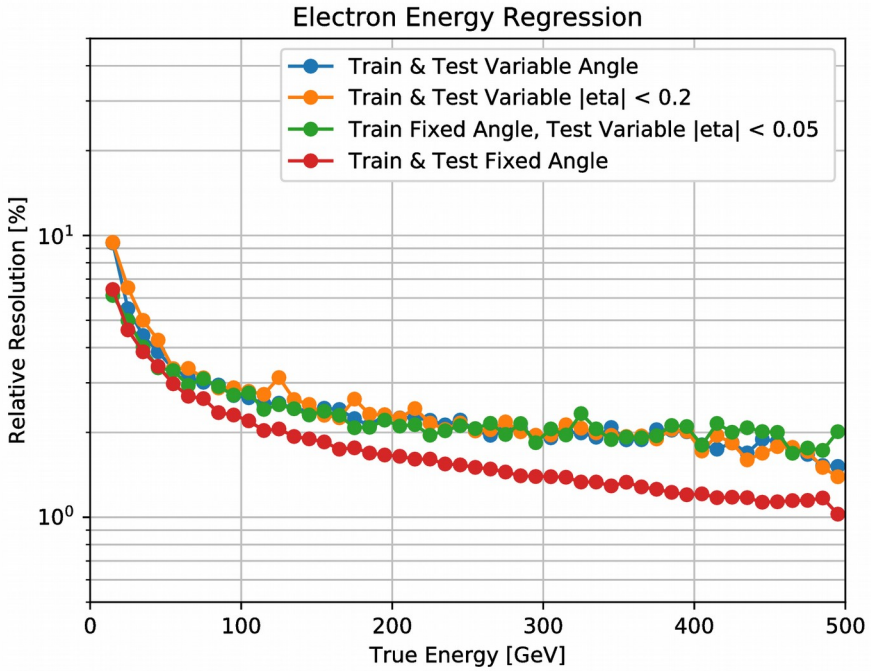


Train on Photons, Fixed Angle

Variable Angle Sample

- Training and testing on variable angle sample gives worse performance than fixed angle sample
  - Added eta to input variables, tried deeper trees, no improvement
- Training on fixed angle sample and evaluating on central events from variable angle sample, |eta| < 0.05, still gives worse performance
- Potential issue / differences with variable angle samples?

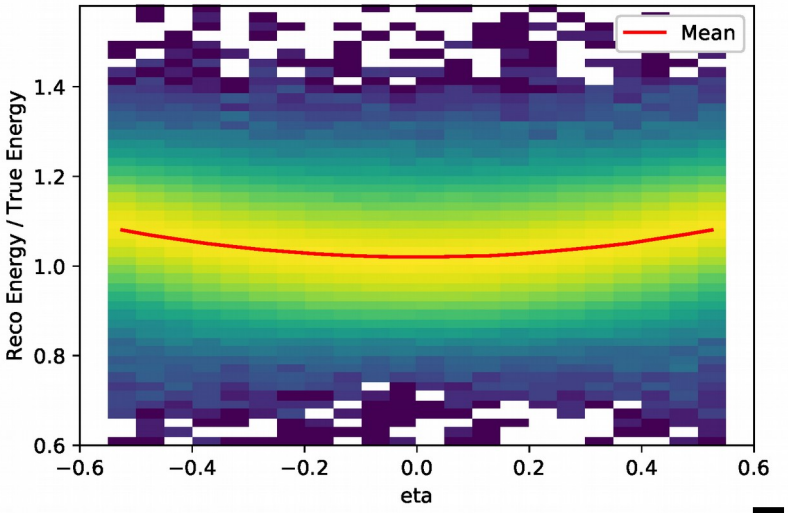


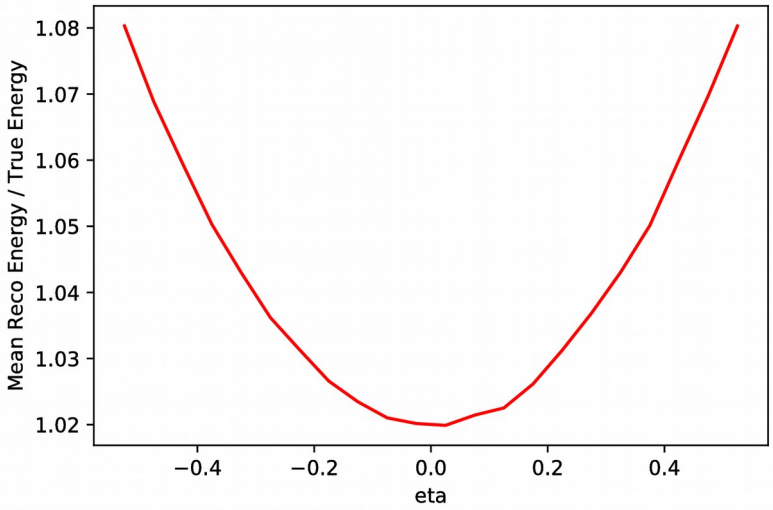


Electrons

Variable Angle Sample (2)

- Ratio of "raw" reco E / true E depends on eta
- Not sure if this is expected, or completely related to energy resolution issue
  - Doesn't explain why resolution is worse for |eta| < 0.05
- Is this from window selection in h5 step, or calo response?
  - Tried to find a .txt file on eos with eta info but didn't manage
  - Integrating over eta, looks like window selection could possibly account for a couple percent effect, not conclusive





Electrons

Summary

- Decision Tree can't extrapolate in energy – should we switch back to NN for feature baseline?
  - Or could try using xgboost to derive a “residual correction” on top of linear regression

- Variable angle sample shows worse resolution than fixed angle at high Energy, even for central objects |eta| < 0.05
  - Response depends on eta, but including eta doesn't improve regression
  - Not sure if this is related to window centering/size at all

Bonus Slides

Samples / Details

- Samples: new larger window samples, fixed angle, with features
  - On culture-plate at caltech:
    - /data/shared/LCDLargeWindow/fixedangle/*Escan/*.h5
    - /data/shared/LCDLargeWindow/varangle/*Escan/*.h5
  - Made slimmed versions with only features (no images):
    - /data/shared/LCDLargeWindow/fixedangle/*Escan/merged_featuresonly/
    - /data/shared/LCDLargeWindow/varangle/*Escan/merged_featuresonly/
  - ~800k events, 70% train, 30% test

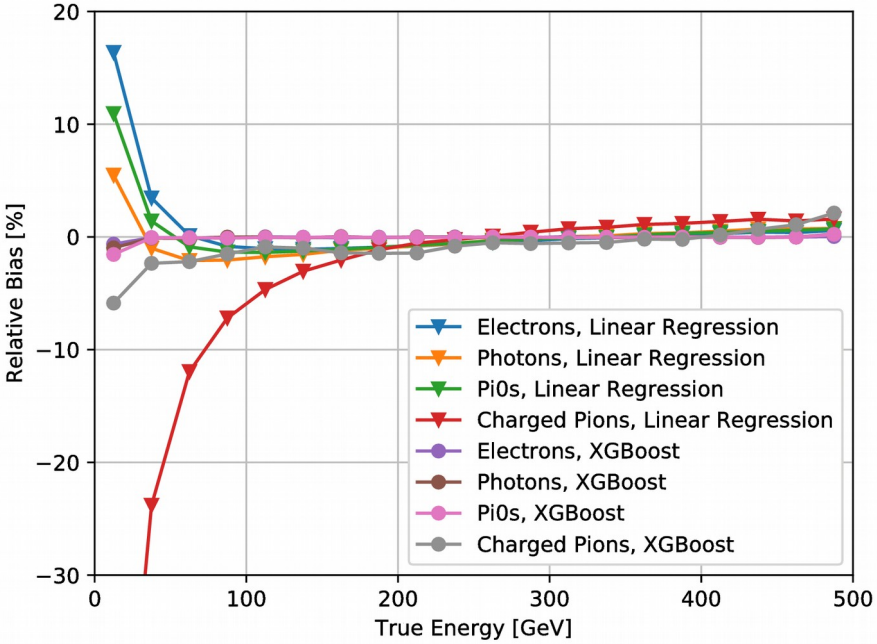
- Running XGBoost in python with:
  - maxdepth 3, up to 1000 rounds
  - Early stopping if test loss doesn't improve for 10 rounds

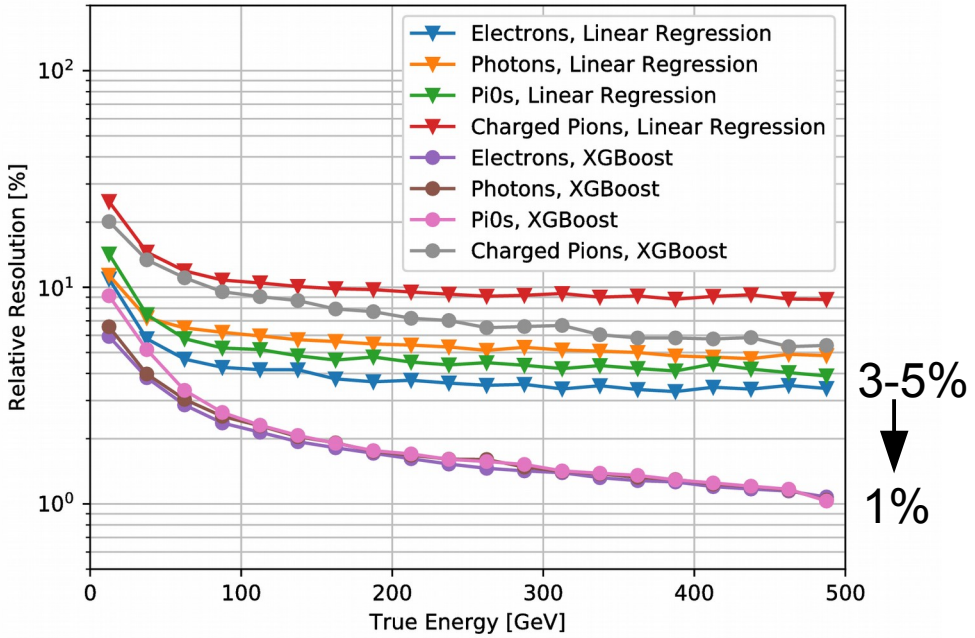
Mean Bias
mean(E_true – E_pred / E_true) * 100

Resolution
RMS(E_true – E_pred / E_true) *100

All Particle Types, Fixed Angle

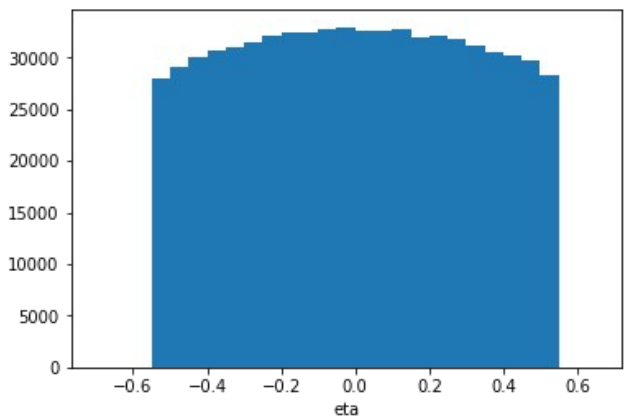
- Showing linear regression with ECAL / HCAL energies
- And XGBoost with energies + shower moments
- Good results for electrons, photons, pi0s
  - Similar resolution above 100 GeV
  - Resolution slightly worse for pi0 at lower energy
- Charged pions worse, for bias and for resolution
  - See next slide

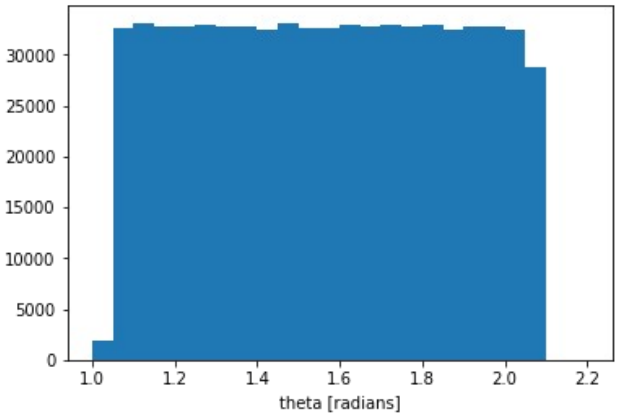


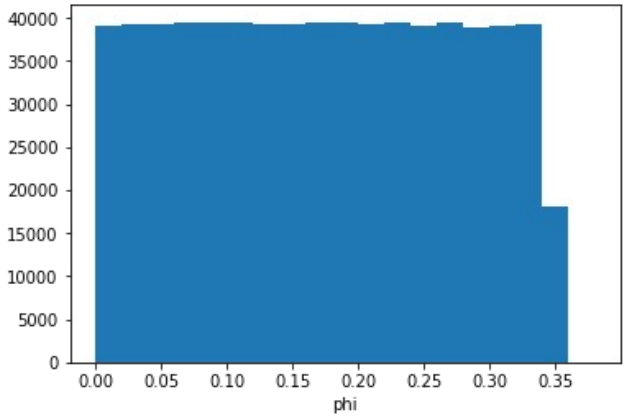


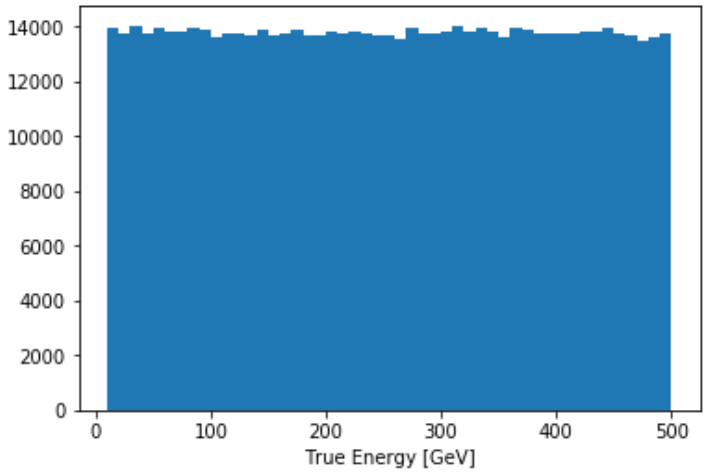
Variable Angle Electrons

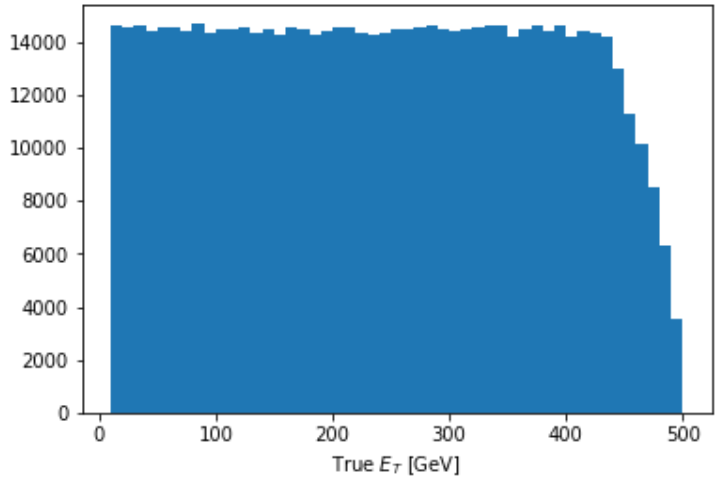
- Flat in theta, phi, energy





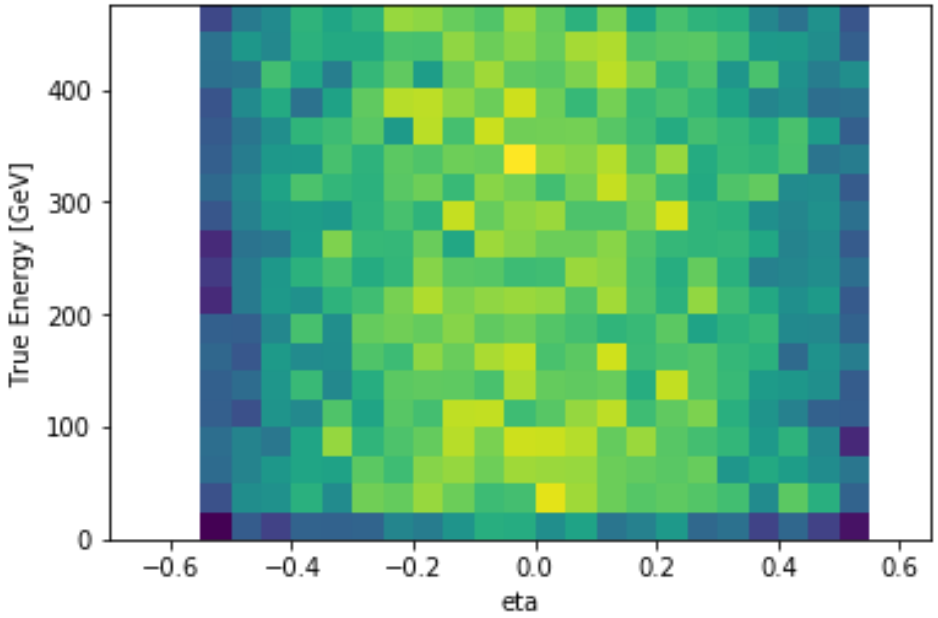


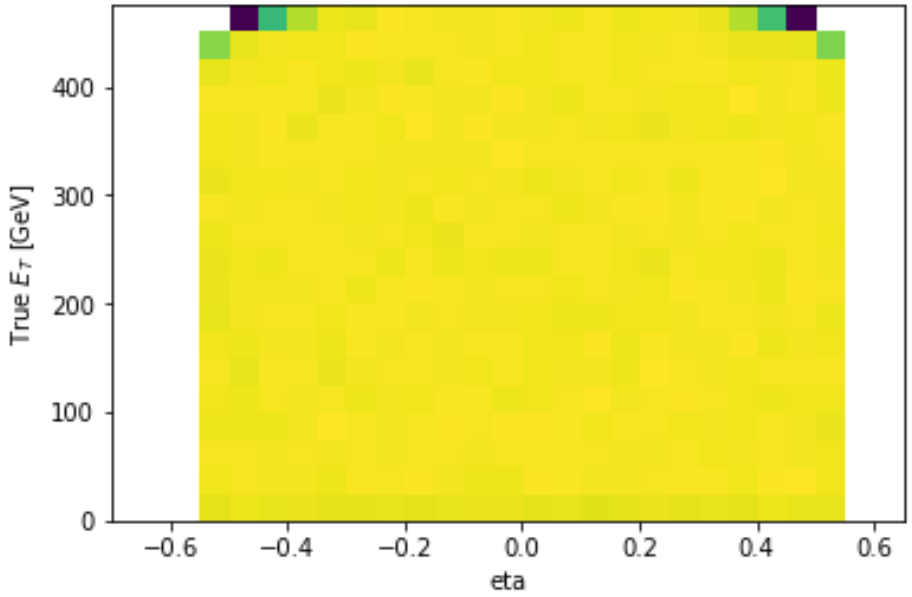




Variable Angle Electrons (2)

- ~Flat in 2d plane of ET, eta
  - Not flat in E, eta plane because of eta dependence





Log color scale