Optimizing Reconstruction and Error Estimation of IceCube Events Using Graph Neural Networks

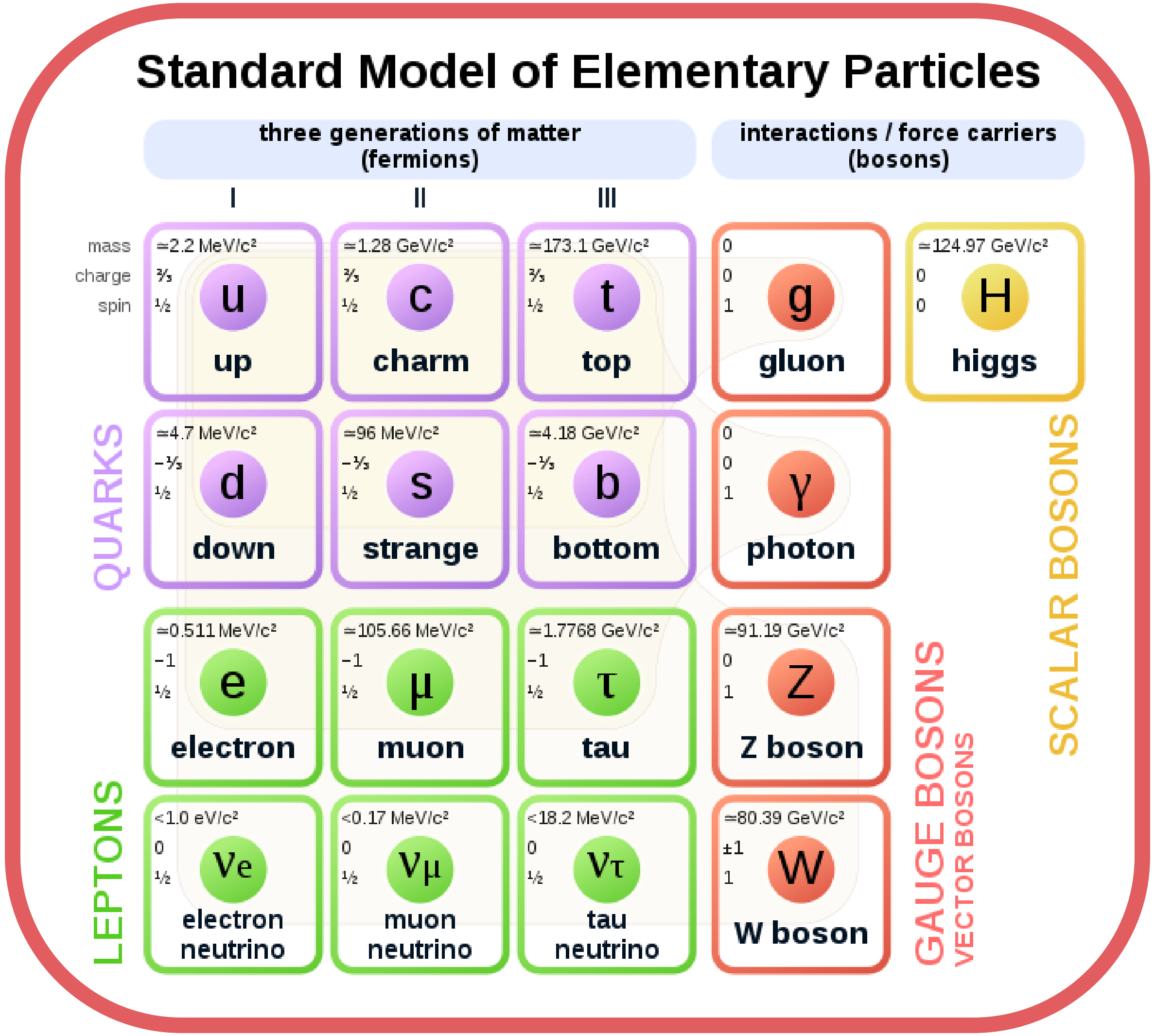
Bachelor Defence

Christian Kragh Jespersen
Jakob Hallundbæk Schauser
Johann Bock Severin
Jonas Vinther

KØBENHAVNS UNIVERSITET

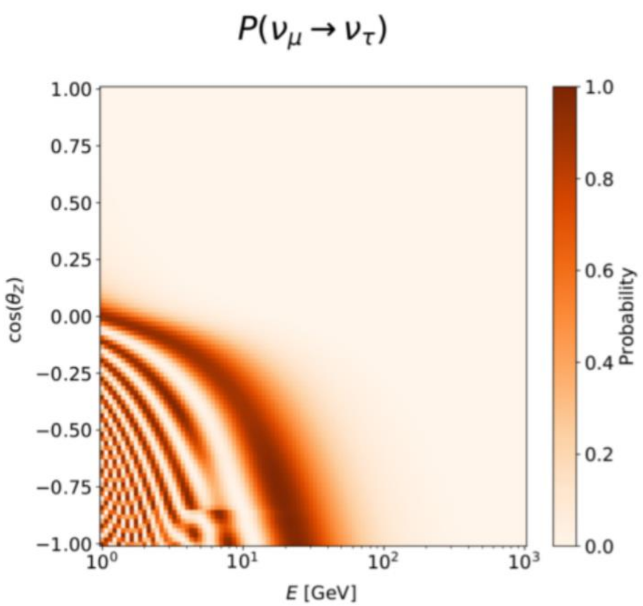
Outline

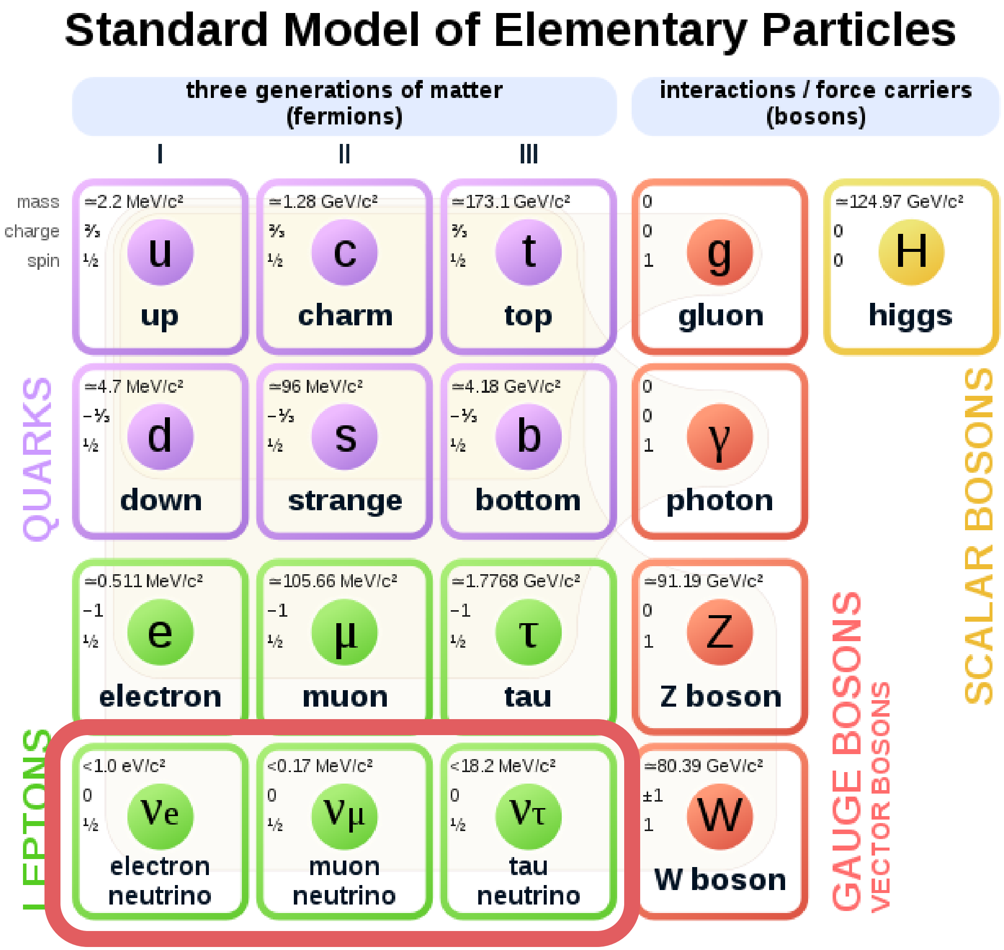
The Standard Model



Neutrinos

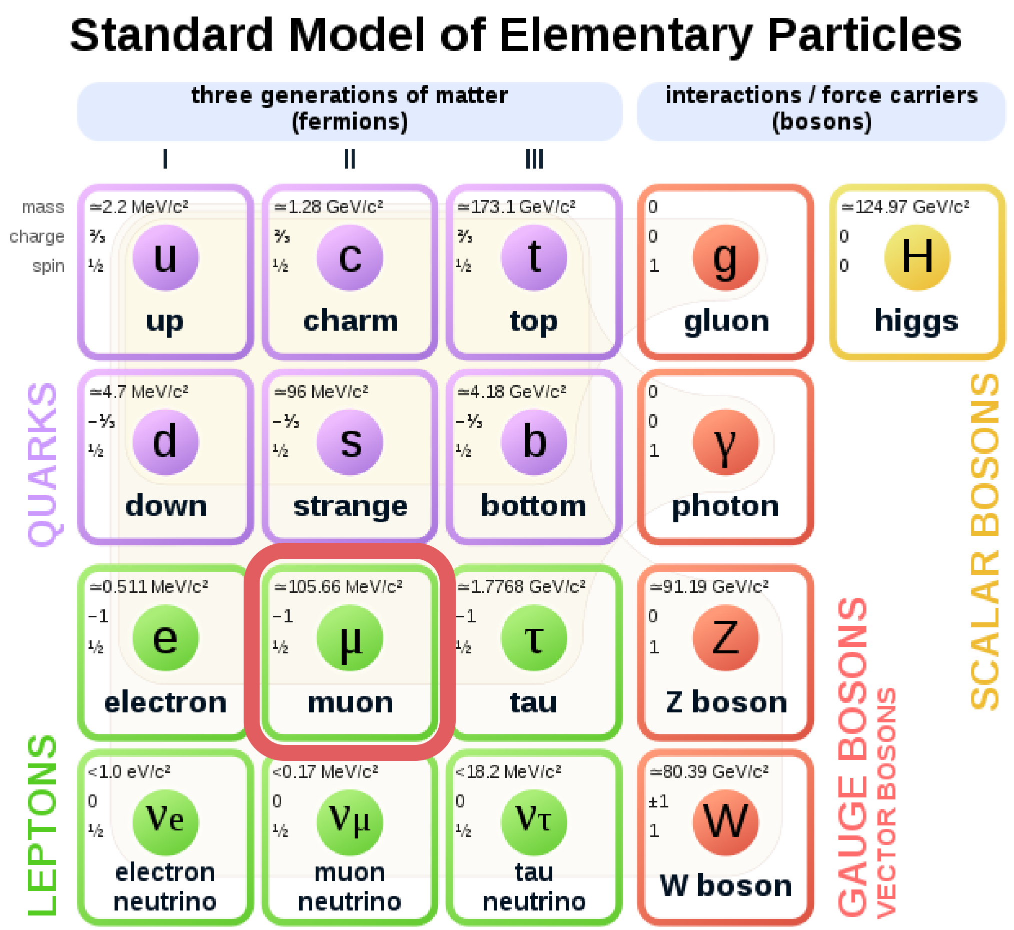
- Only weak-force/gravity
- Unknown masses (lightweight)
- Flavour oscillation



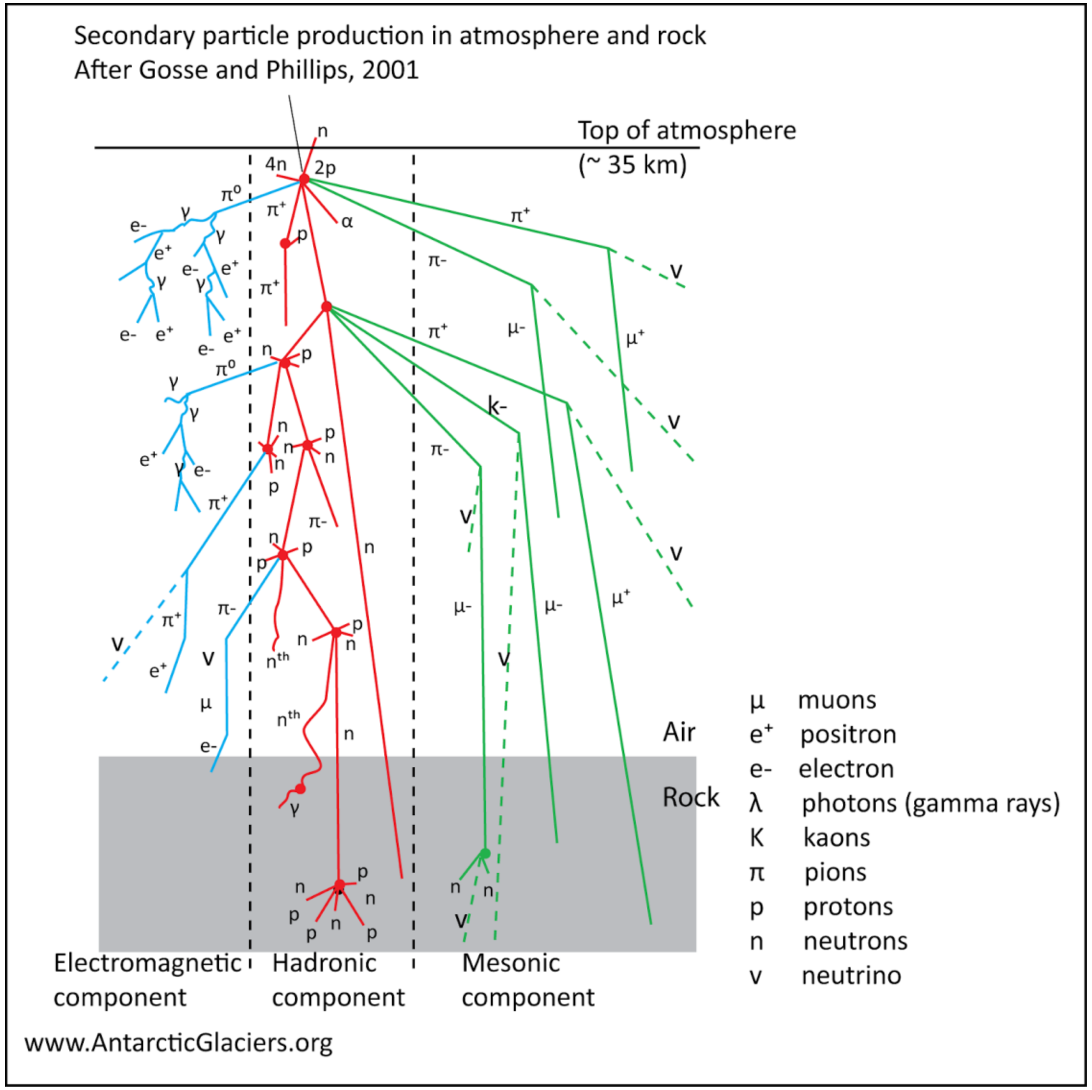


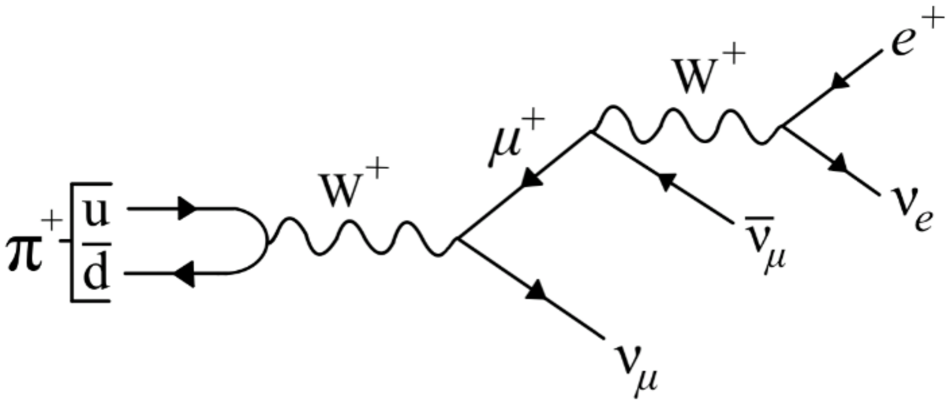
Muons

- Forces: EM-weak-gravity
- Electron with ~200x the mass
- Lifetime: 2.2×10^{-6} s



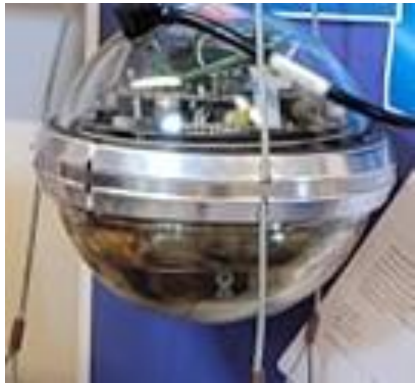
Where Do the Particles Come From?

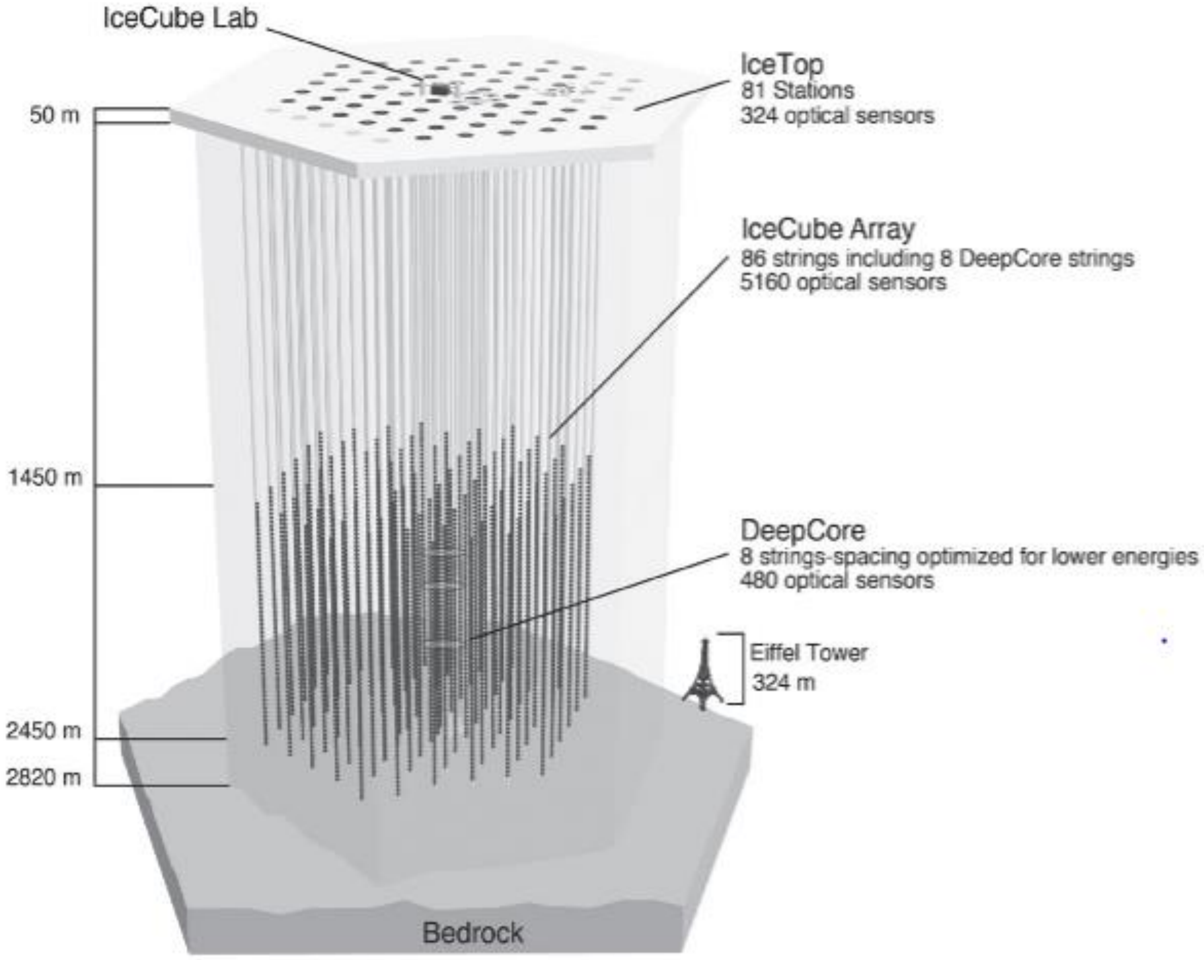




IceCube Neutrino Observatory

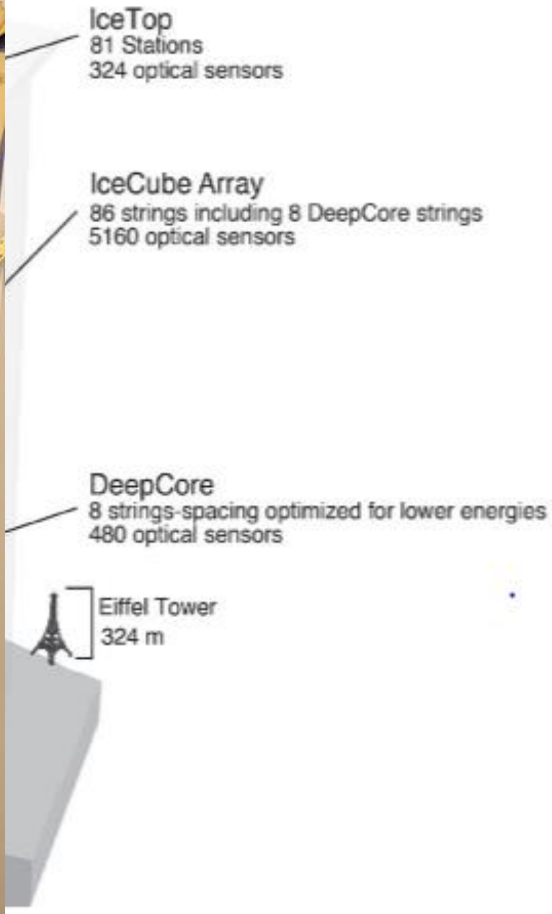
- As much mass as possible
- Digital Optical Module (DOM)
- 86 + 8 strings
- Clear ice
- Avoid hadronic/ electromagnetic biproducts





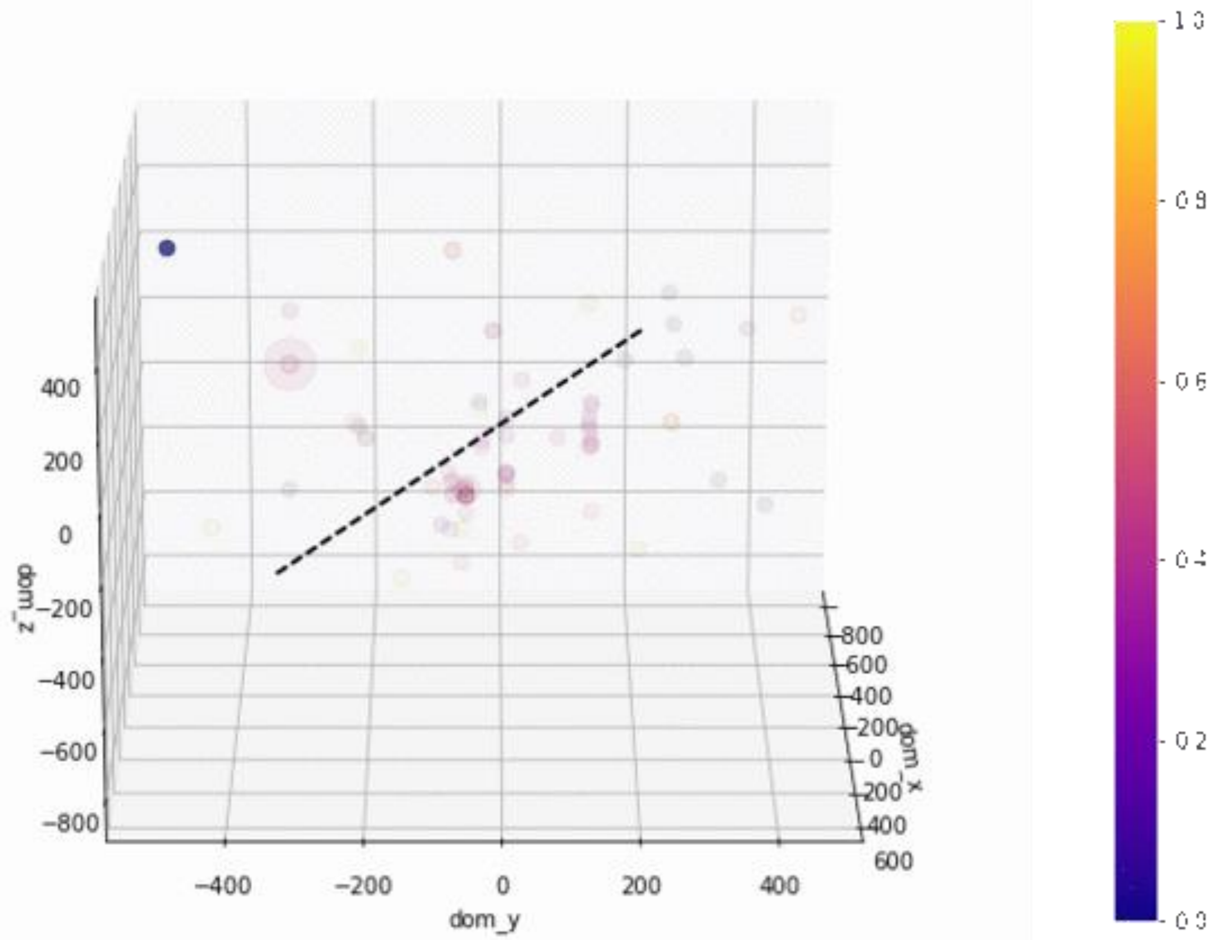
IceCube Neutrino Observatory

- As muc
- Digital
- 86 + 8
- Clear ic
- Avoid h
  electro



The Data

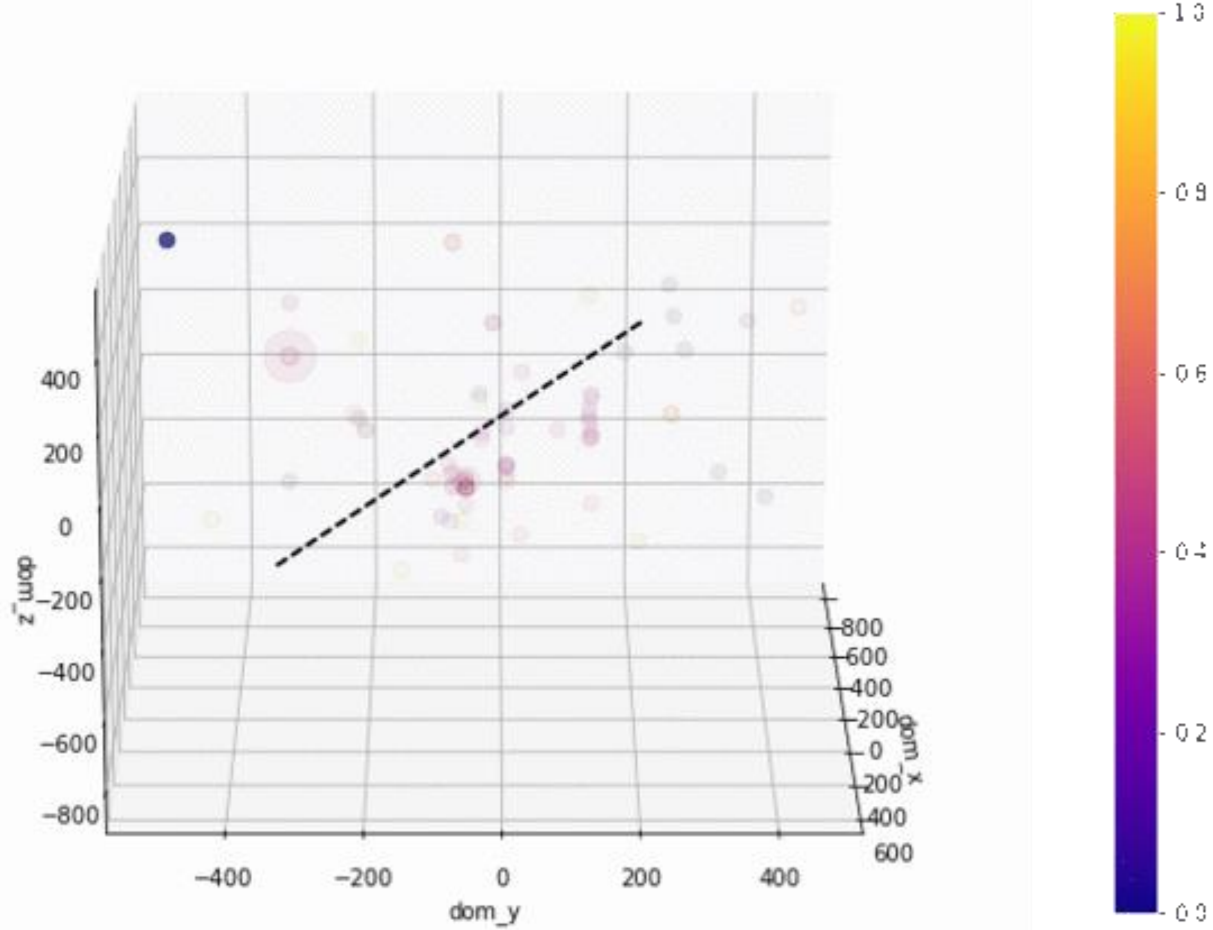
- What We See
  - Position
  - Time
  - Charge
  - Precision
- What We Want
  - Zenith
  - Azimuth
  - Energy
  - (Classification)
  - (Stopped)



Why New Algorithms?

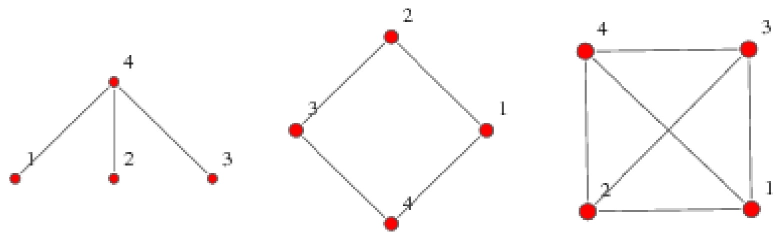
- Retro
  - Slow
  - High-energy
  - Unflexible

- Transition to Machine Learning
  - Fast
  - Flexible

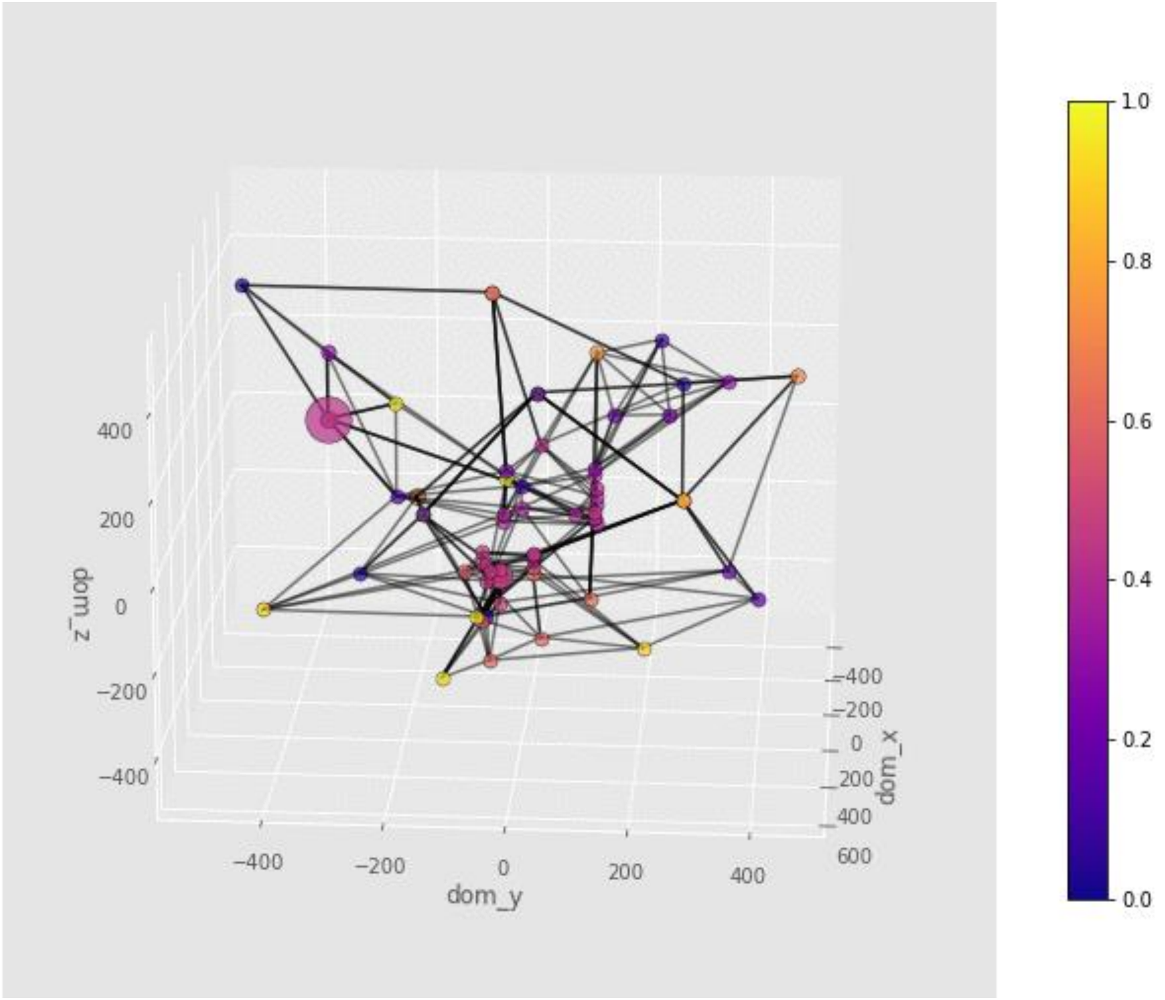


How About Graphs?

- Nodes and edges
- Represented in Linear Algebra:
    - Feature Matrix, $\boldsymbol{X}$
    - Adjacency Matrix, $\boldsymbol{A}$
    - Edge-features, $\boldsymbol{E}$



$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

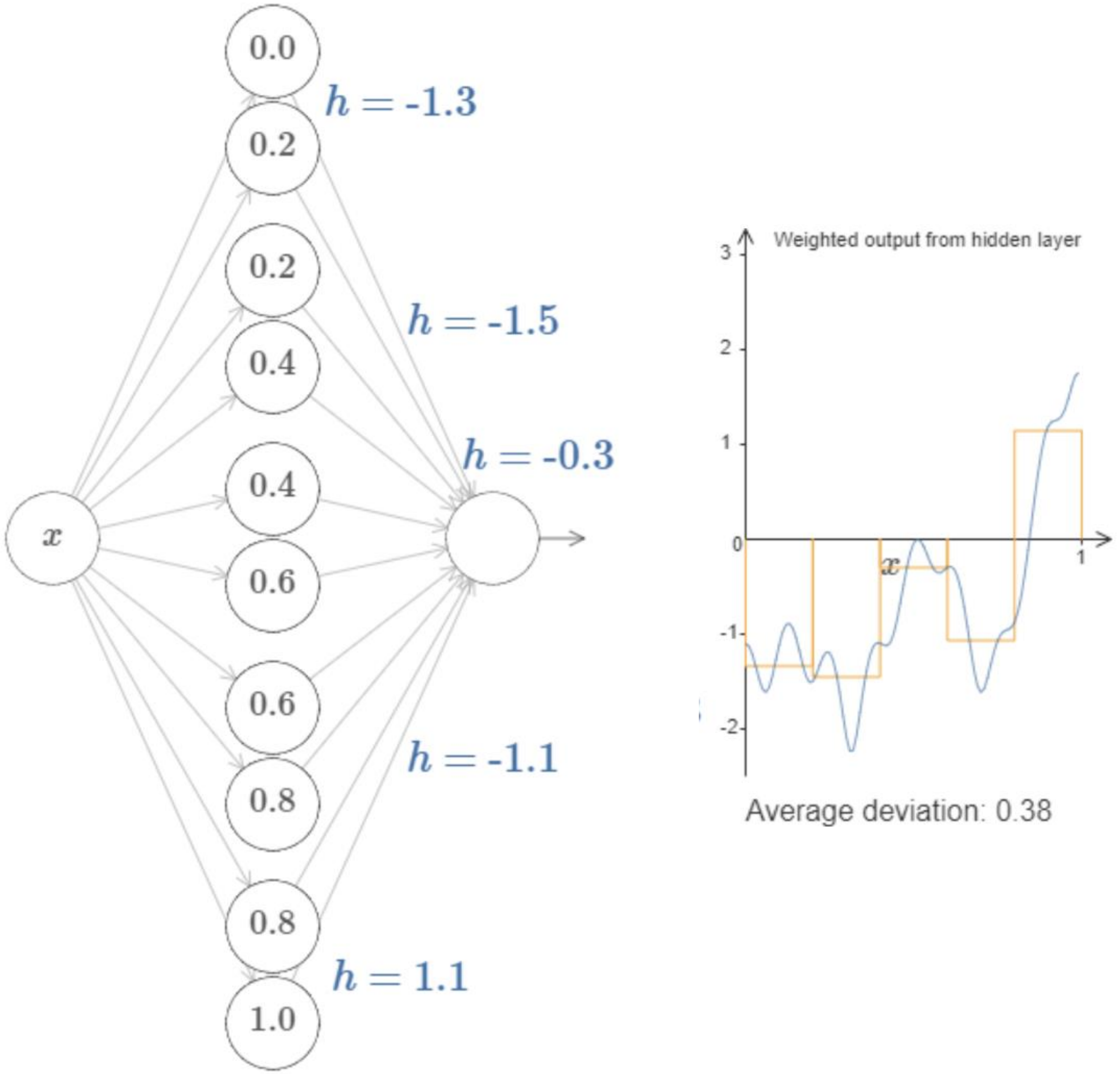


Neural Networks

- Multi-Layer Perceptron (MLP)

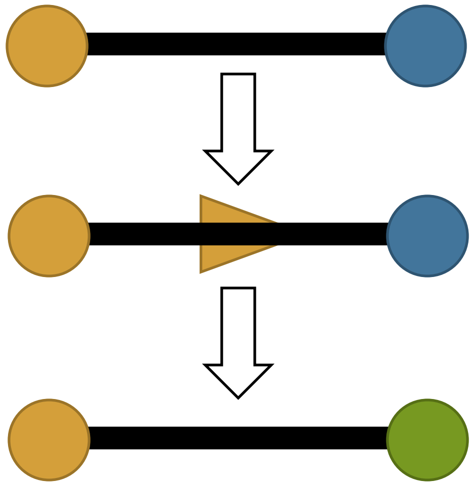
$$f(x) = \phi^N \left(W^N \phi^{N-1} \left(W^{N-1} \cdots \phi^1 \left(W^1 x \right) \cdots \right) \right)$$

- Universal Approximation Theorem



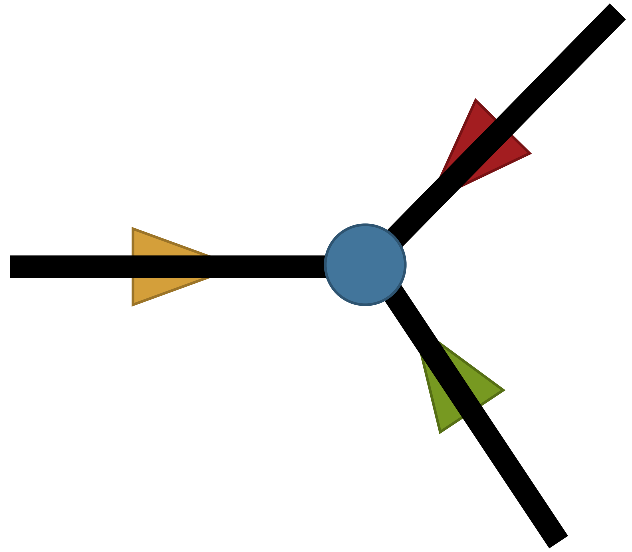
Graph Neural Networks

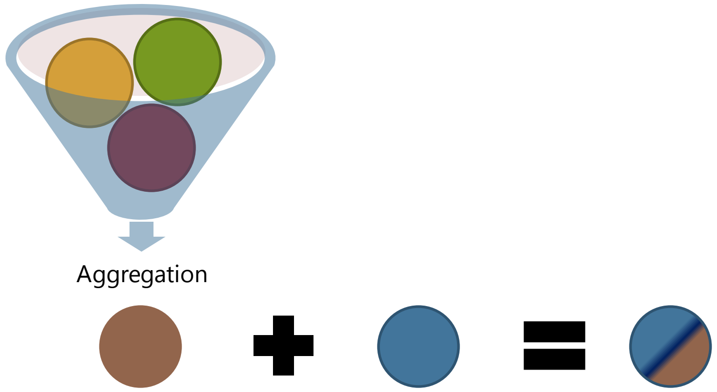
- **Message passing method**
- Aggregation
- Convolutions
- Pooling



Graph Neural Networks

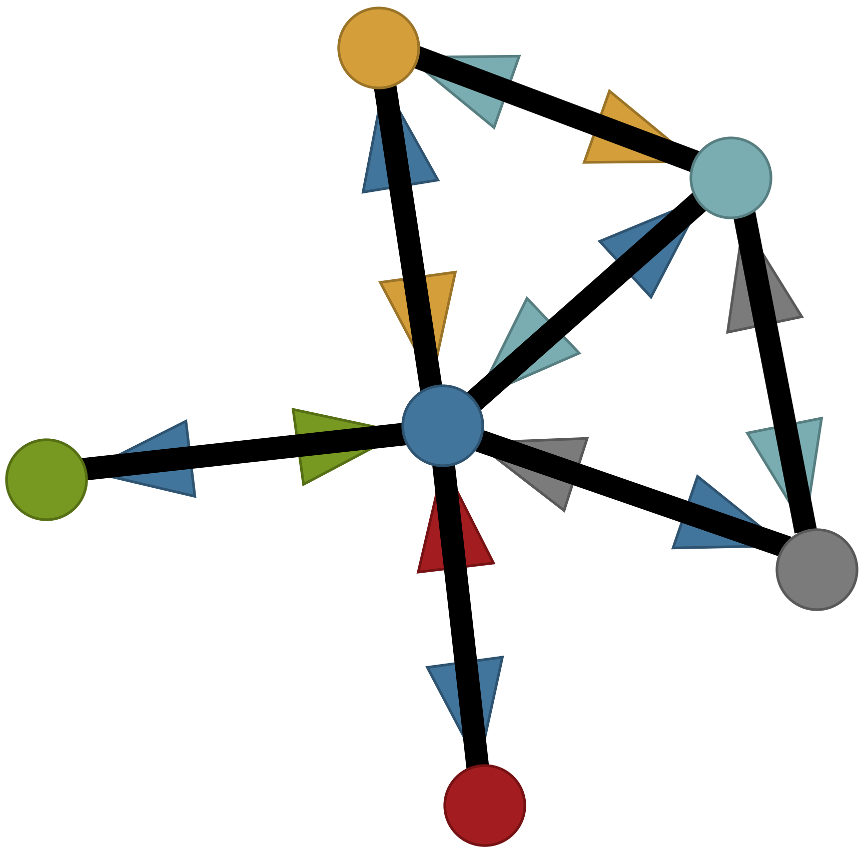
- Message passing method
- **Aggregation**
- Convolutions
- Pooling





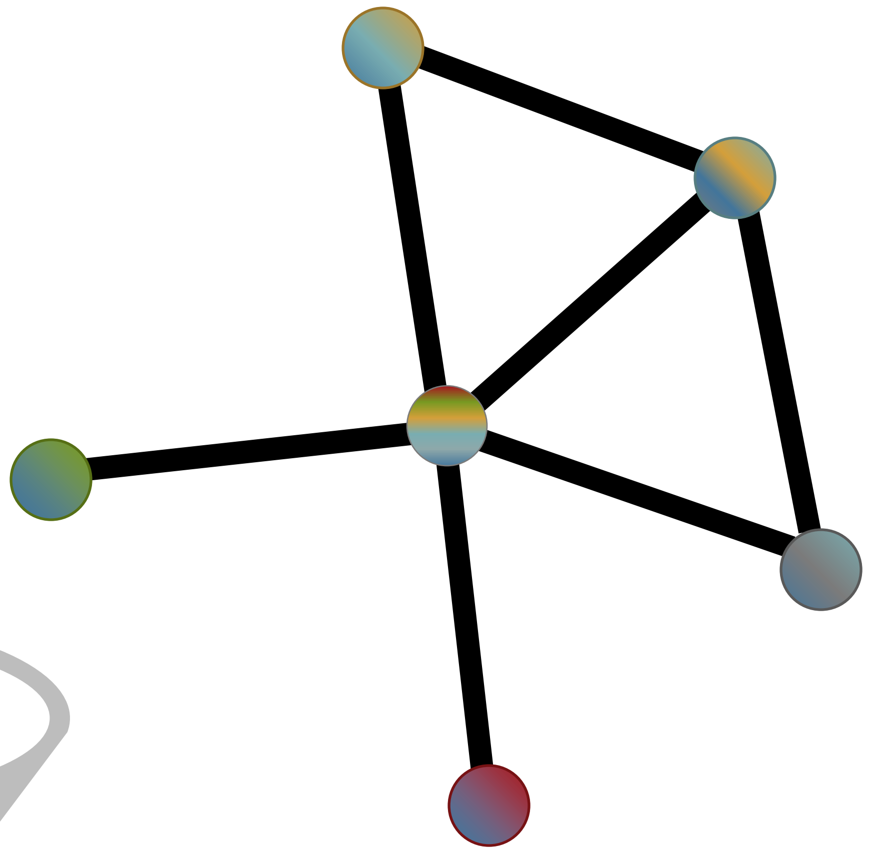
Graph Neural Networks

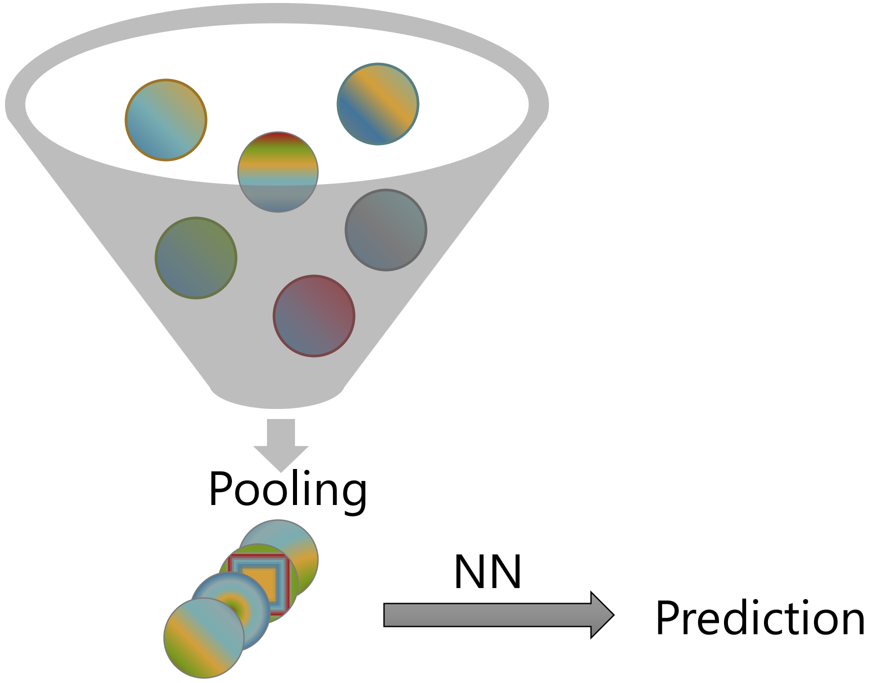
- Message passing method
- Aggregation
- **Convolutions**
- Pooling



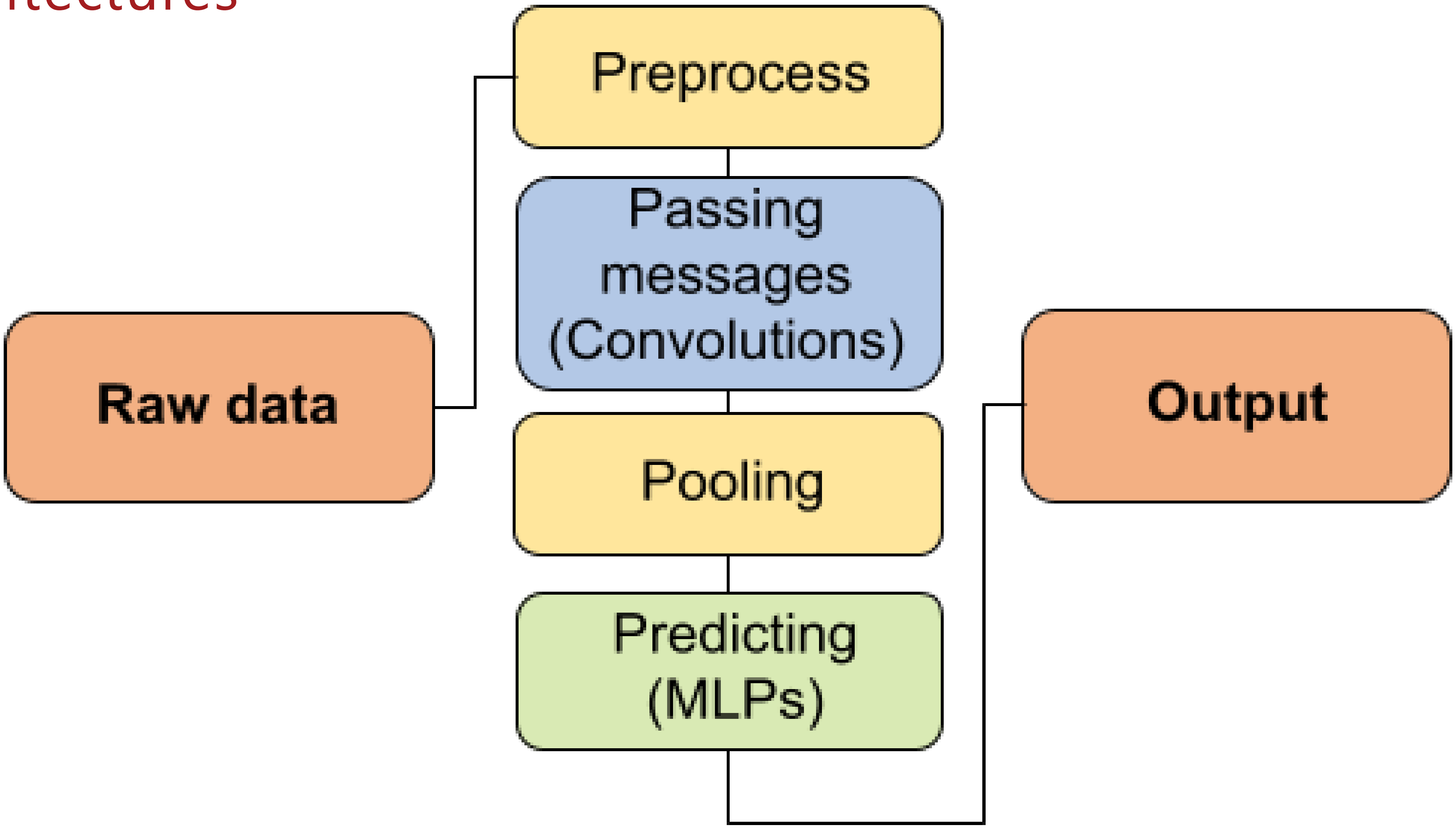
Graph Neural Networks

- Message passing method
- Aggregation
- Convolutions
- **Pooling**



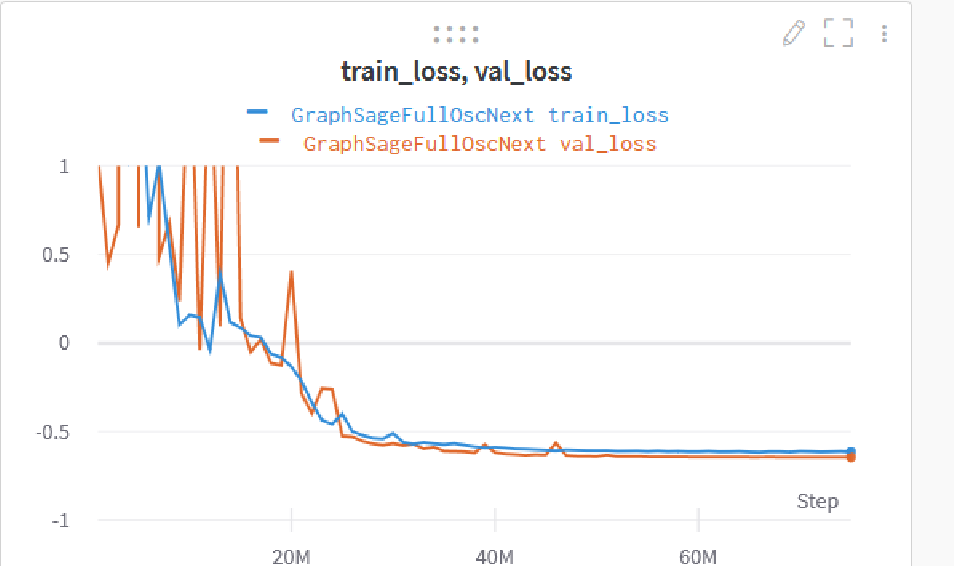


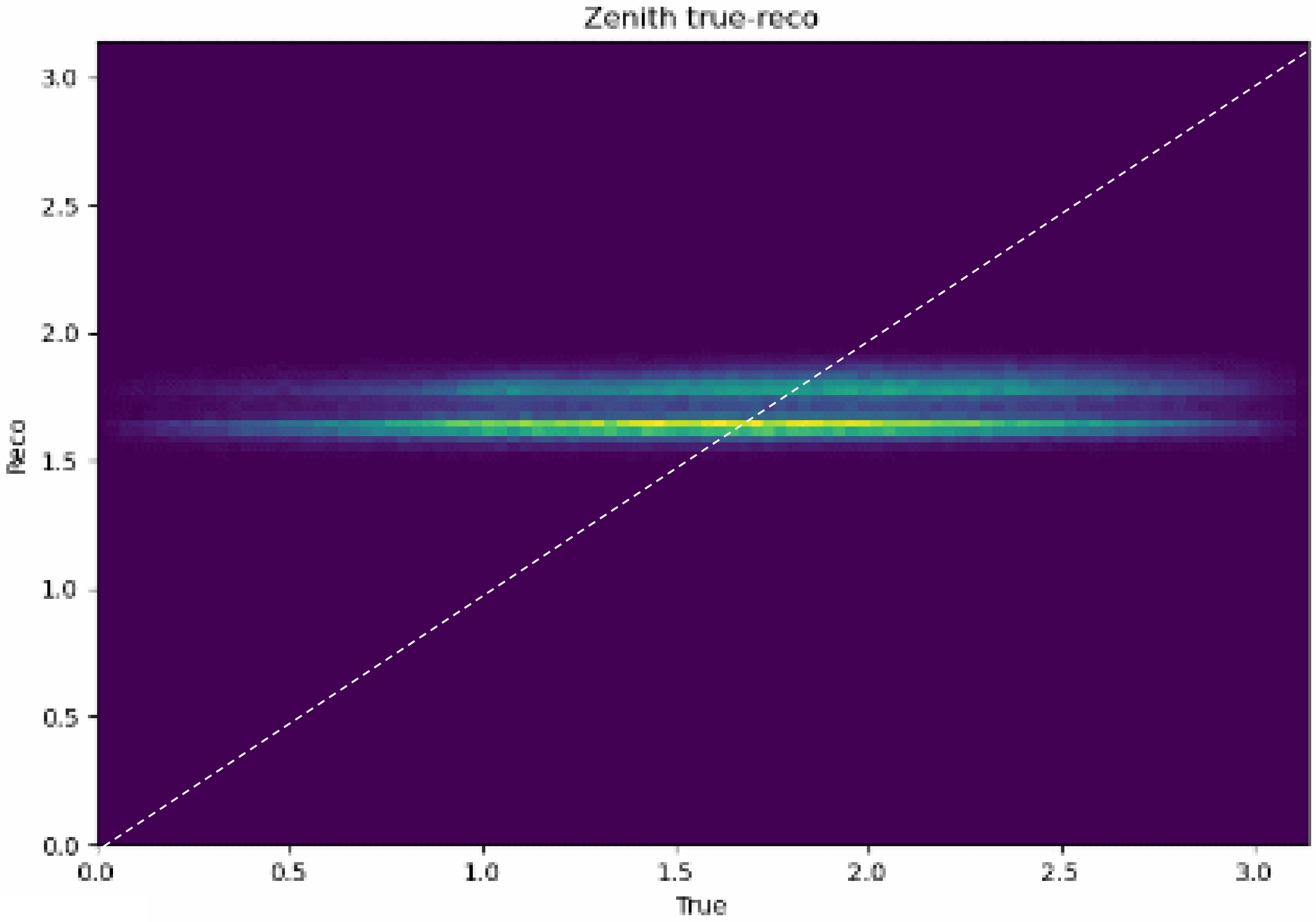
Architectures



Training the Network

- Loss functions
- Gradient descent



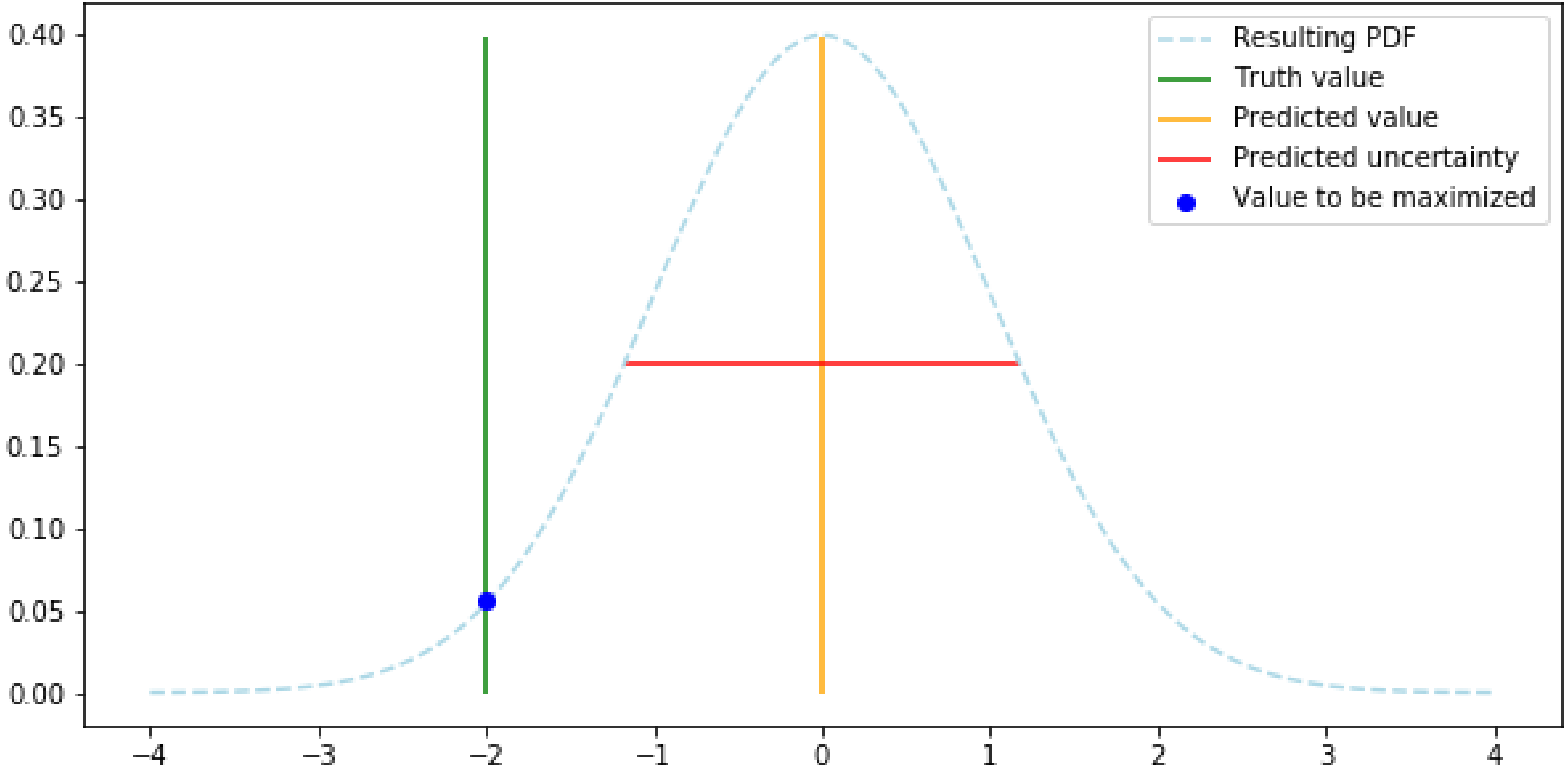


Results

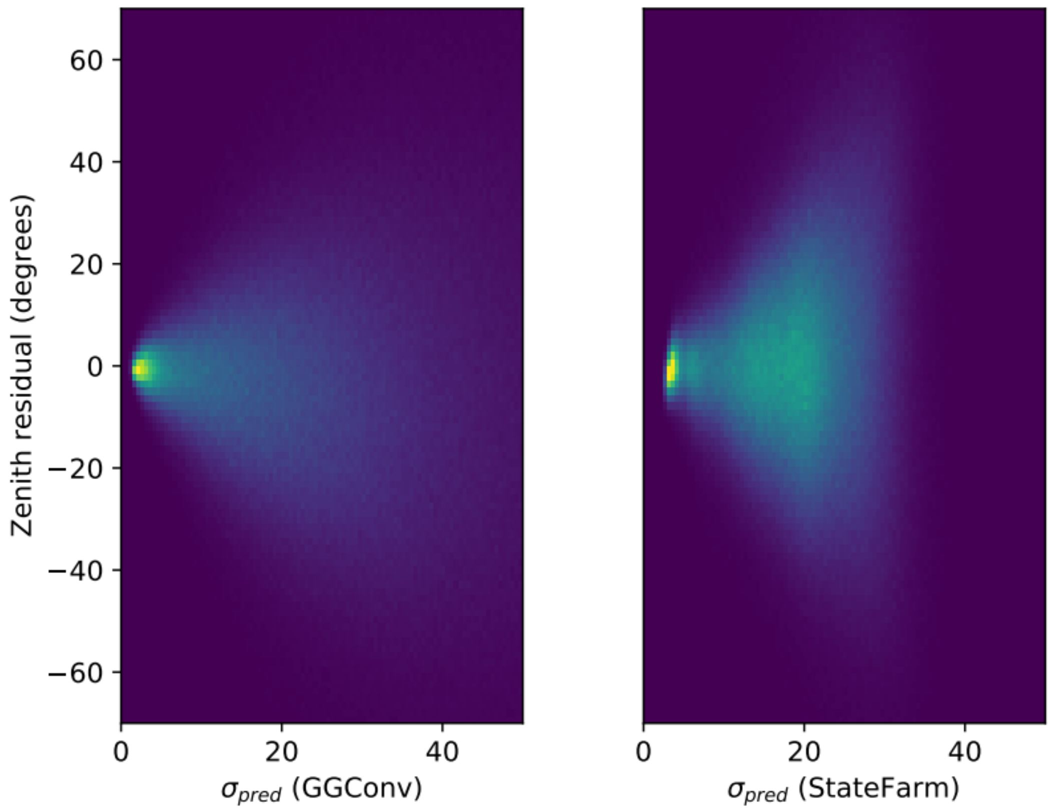
	Loss ($\theta, \phi / x, y, z$)	Loss ($log(E)$)	θ	ϕ	Ω	$log(E)$	Speed	Params (10^3)
StateFarm	2xPvM	MAE	**11.92 °**	31.23 °	33.37 °	**0.1487**	$9 \cdot 10^3$	643
GGConv	vG	Normal	15.94 °	**29.65 °**	**32.84 °**	0.2105	$3 \cdot 10^4$	176
AntHill	2xPvM/SvM	Normal	14.20 °	33.53 °	36.37 °	0.2357	$1 \cdot 10^4$	2,218
LifeGuard	2xPvM+σ	MAE	15.91°	-	-	0.2634	$2 \cdot 10^4$	2.2
Retro	-	-	15.00 °	37.93 °	40.56 °	0.2390	$O(10^{-2})$	-
σ	-	-	0.01 °	0.05 °	0.04 °	0.0003	-	-

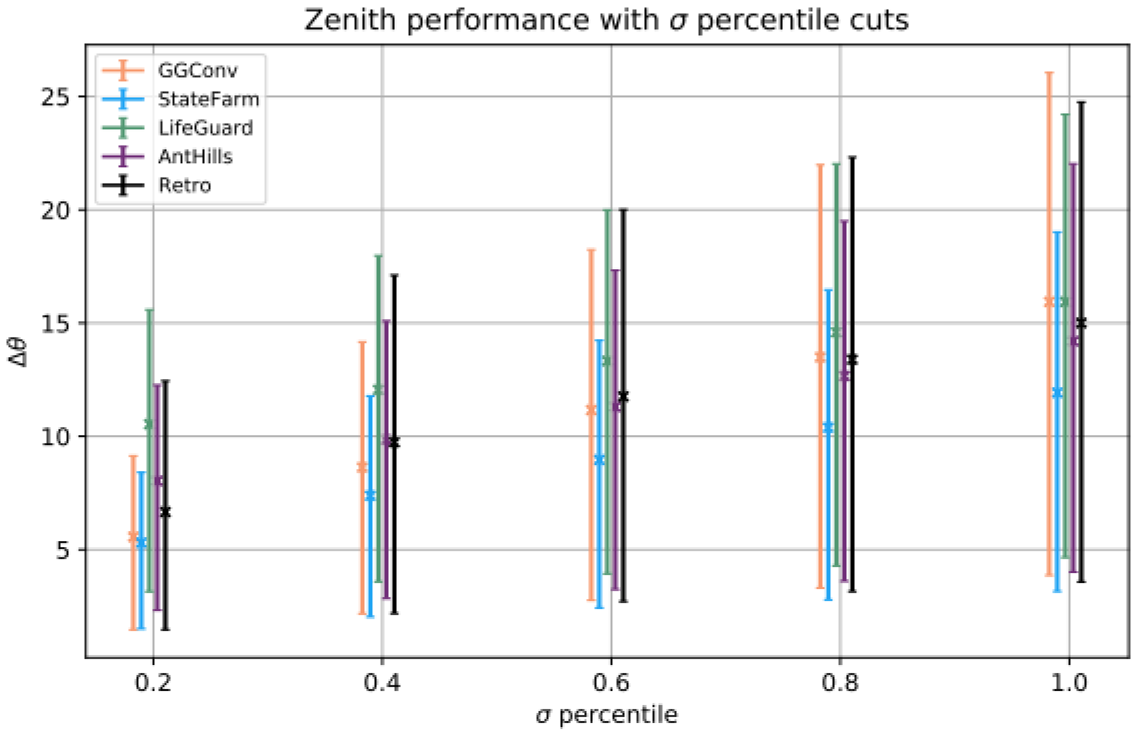
	Best GNN	Retro
$\log_{10}(E)$	0.15	0.24
θ (deg)	11.9	15.0
ϕ (deg)	29.6	38.0
Ω (deg)	32.8	40.6

Probabilistic Loss Functions

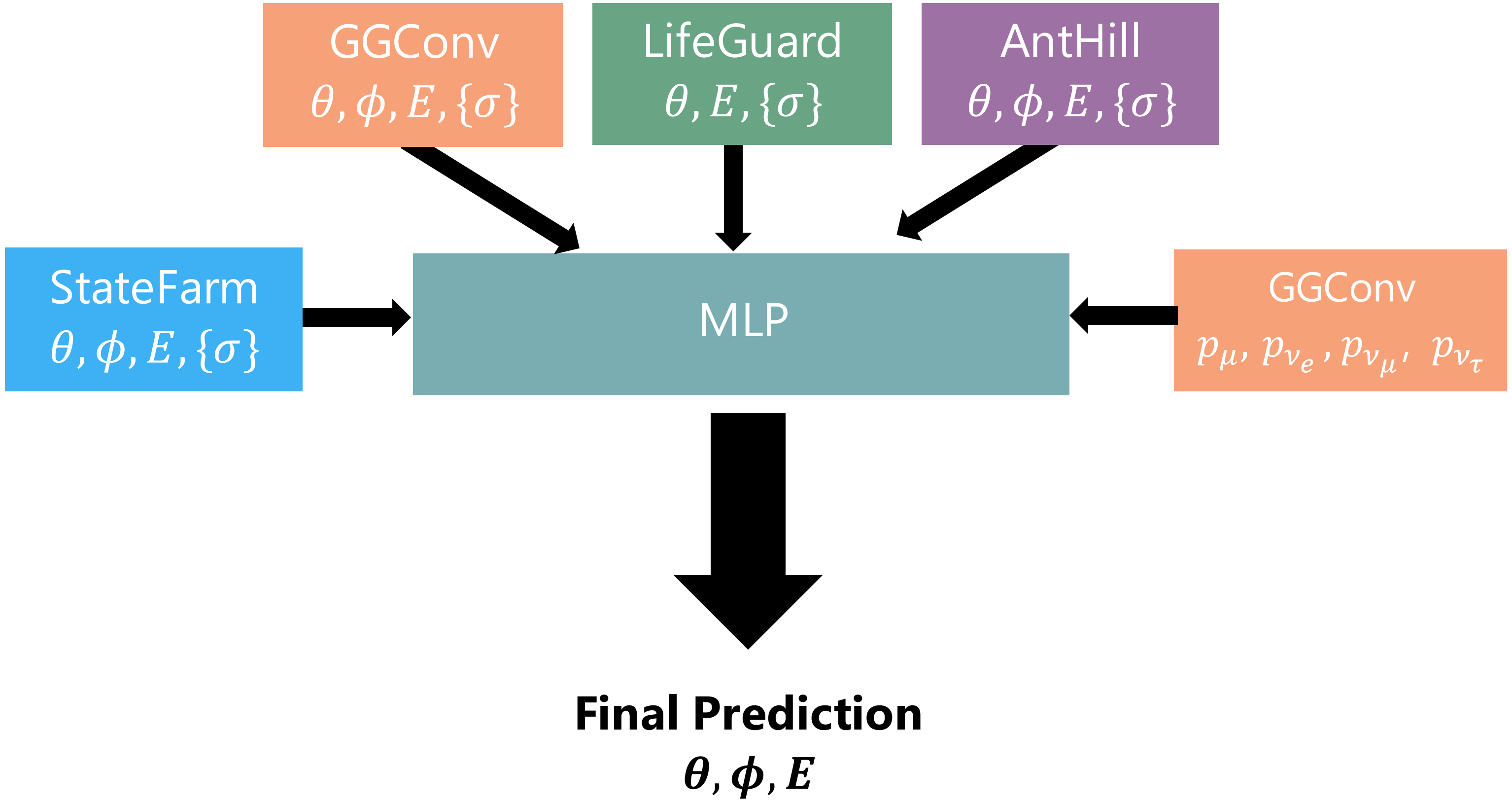


Uncertainty estimation

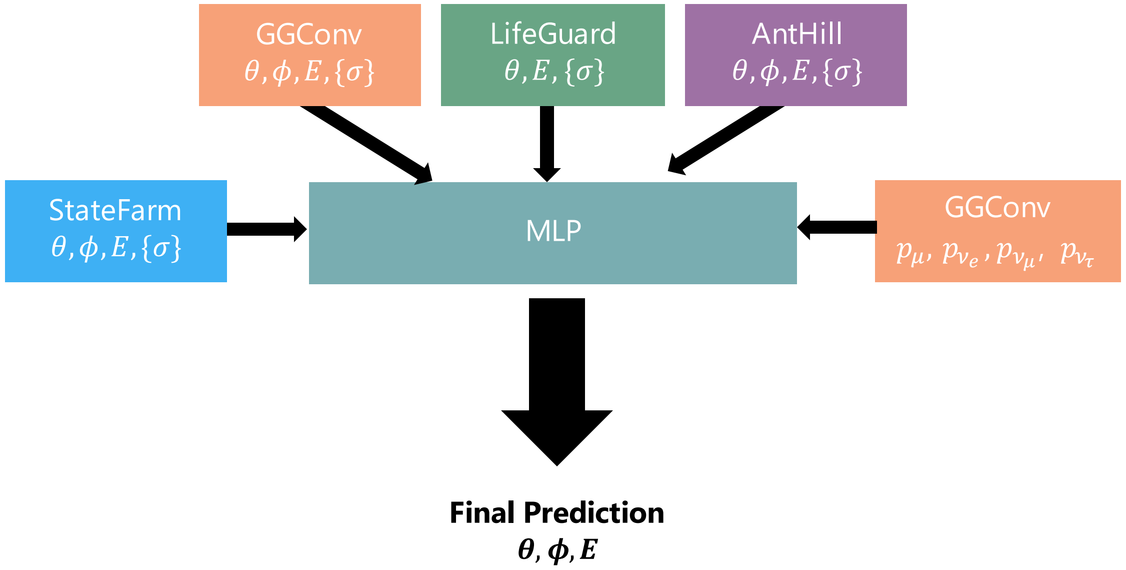




Ensemble

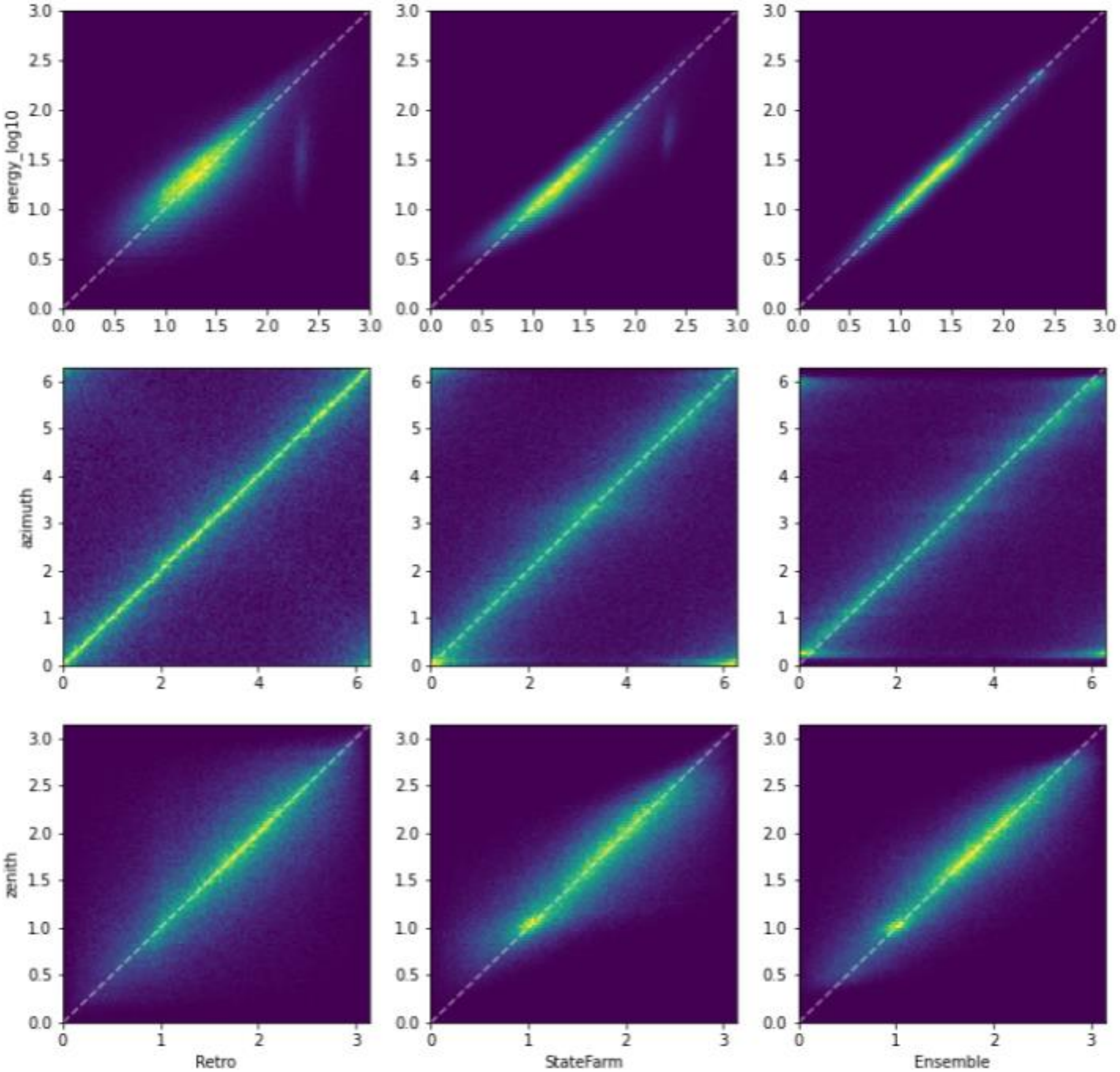

GGConv
$\theta, \phi, E, \{\sigma\}$
LifeGuard
$\theta, E, \{\sigma\}$
AntHill
$\theta, \phi, E, \{\sigma\}$
StateFarm
$\theta, \phi, E, \{\sigma\}$
MLP
GGConv
$p_{\mu}, p_{\nu_e}, p_{\nu_\mu}, p_{\nu_\tau}$
Final Prediction
$\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{E}$

Ensemble



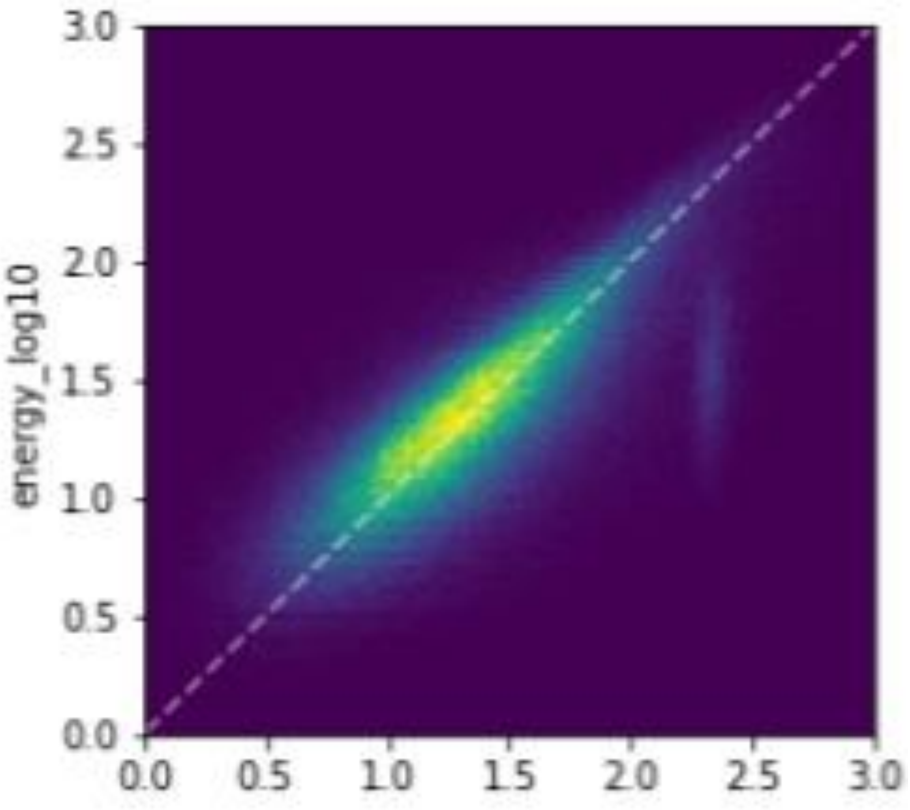
	Ensemble	Best GNN	Retro
$\log_{10}(E)$	**0.09**	0.15	0.24
θ (deg)	**10.8**	11.9	15.0
ϕ (deg)	29.8	29.6	38.0
Ω (deg)	**30.9**	32.8	40.6

Ensemble

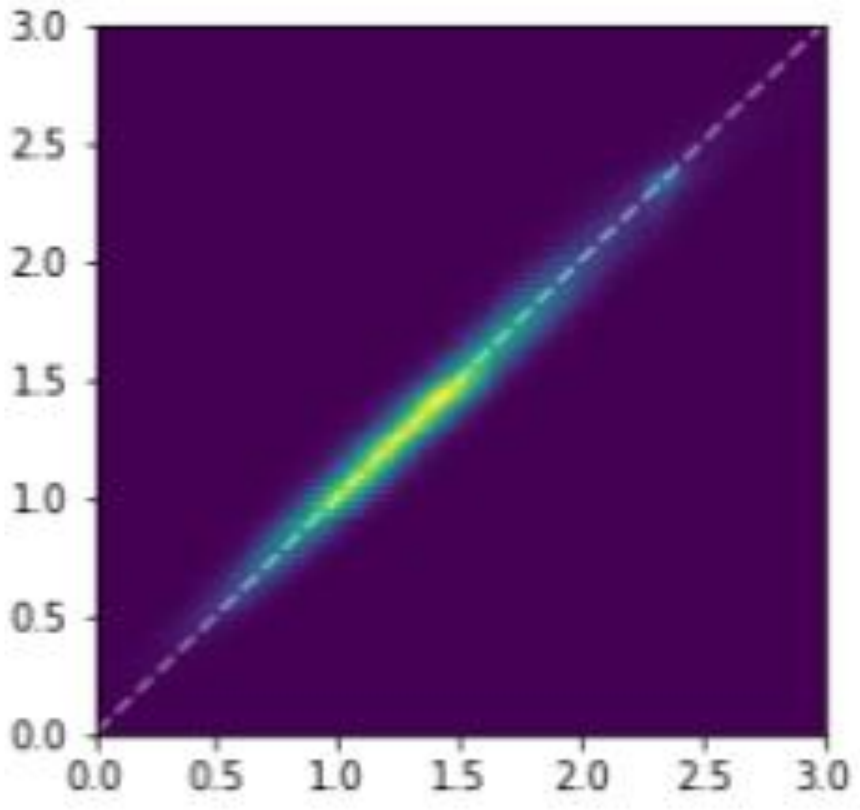


Ensemble – Energy Predictions

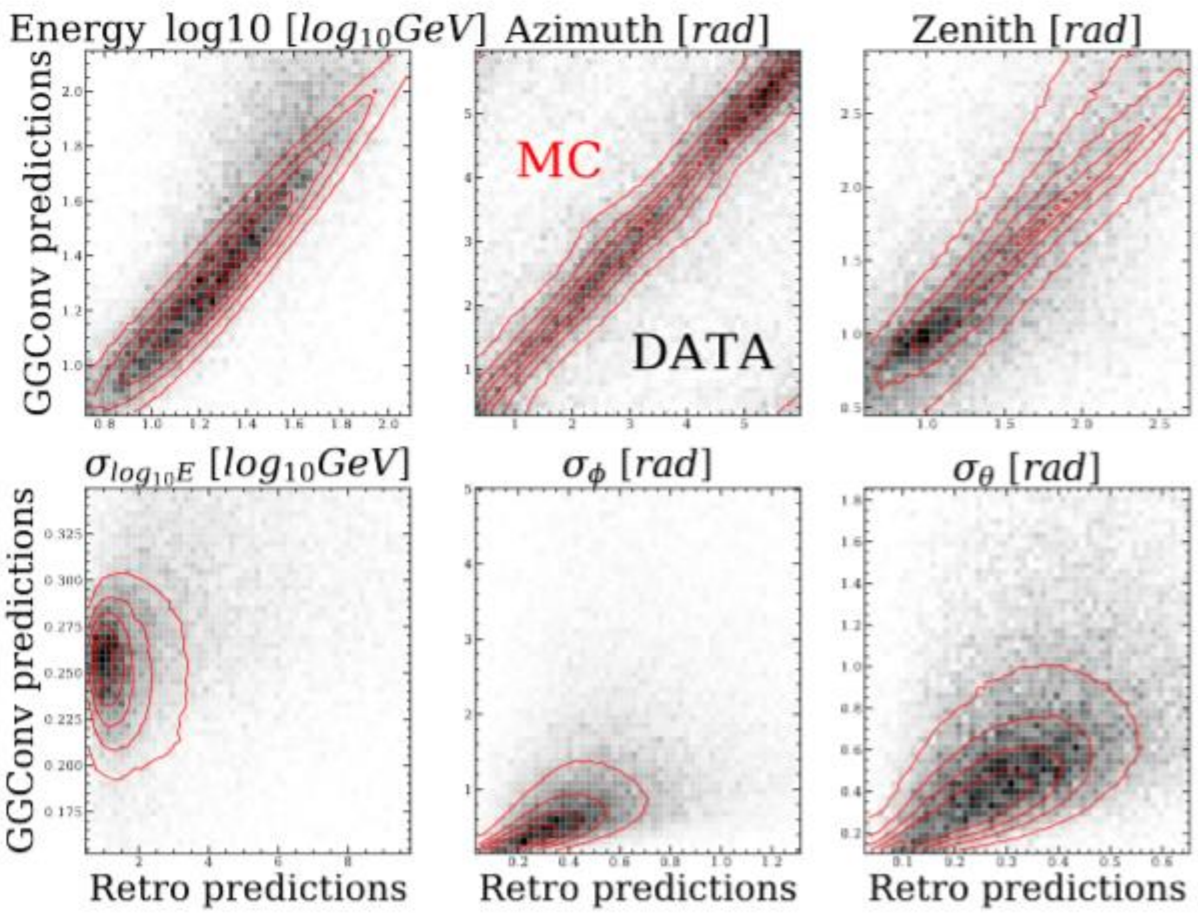
Retro



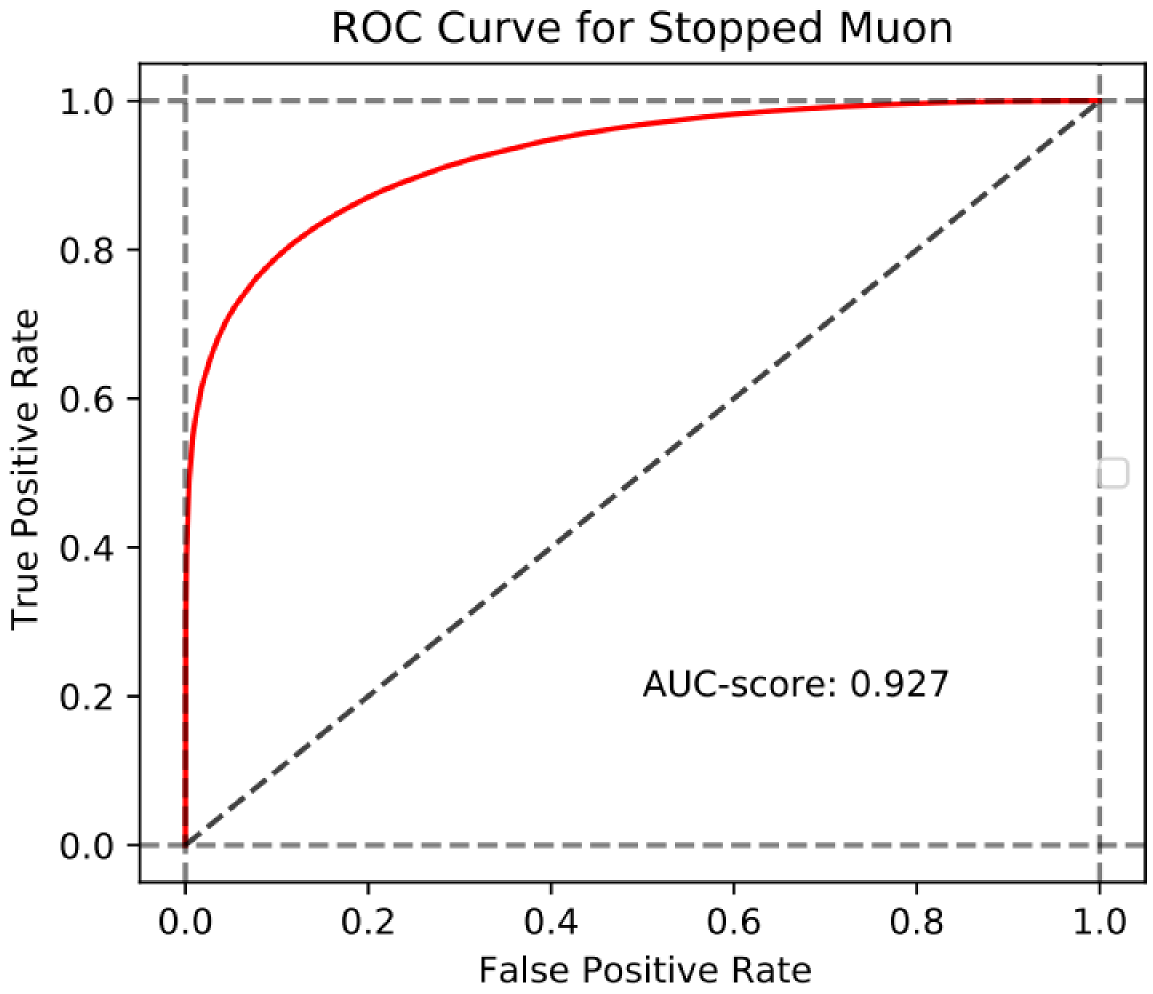
Ensemble



From simulation to observation



Muons and the Moon



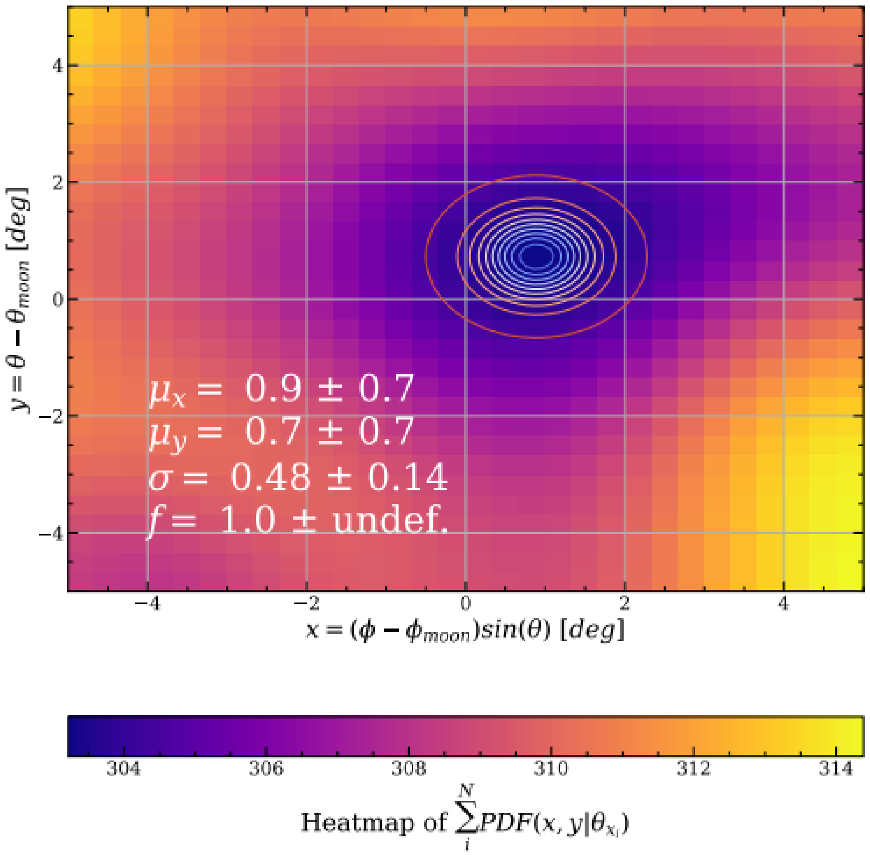
Reconstruction of MuonGun

	θ	ϕ	Ω	$log_{10}(E)$
StateFarm	**1.72 °**	**5.49 °**	**4.09 °**	**0.1340**
GGConv	2.77 °	7.52 °	5.77 °	0.1376
AntHill	1.88 °	5.57 °	4.19 °	-
σ	0.01°	0.02 °	0.01 °	0.0007

Moon

- Calibration tool for real observations

- Uncertain Unbinned Maximum Likelihood Estimation

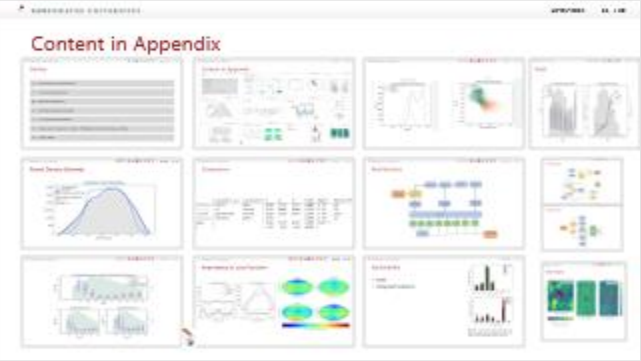


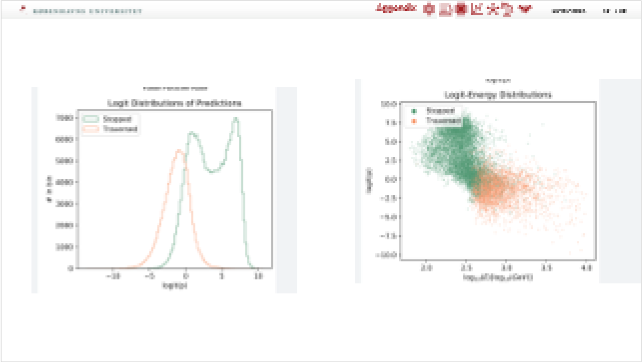
Further work

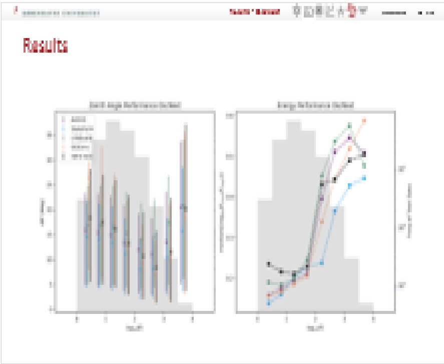
- More Work on Ensembles
- Improved Explainability
- Better Implementation/Closer to the I3 Files
- GNNs for Cleaning

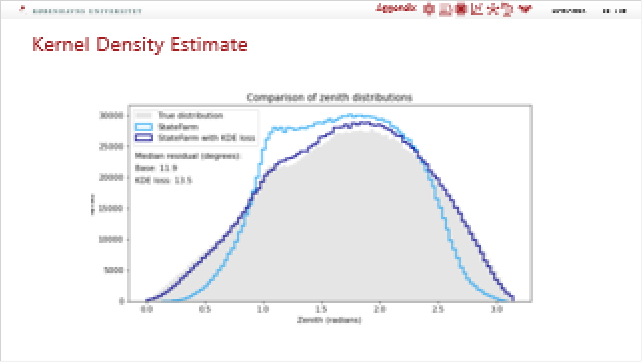
Appendix

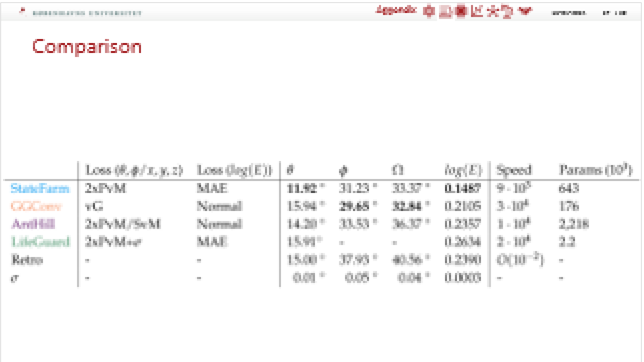
Content in Appendix



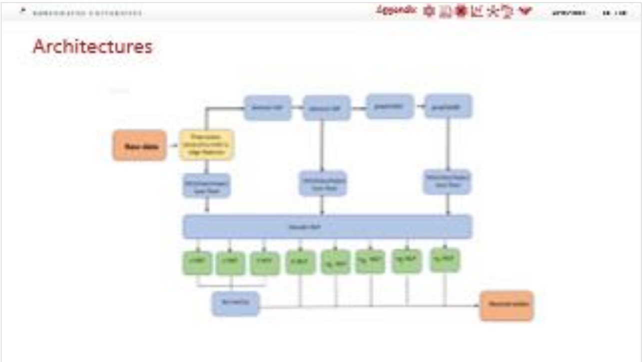


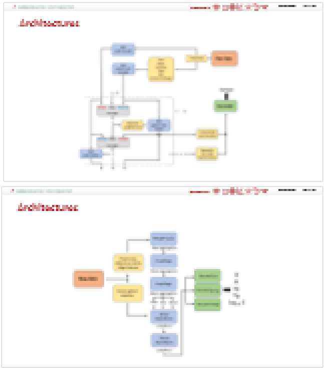


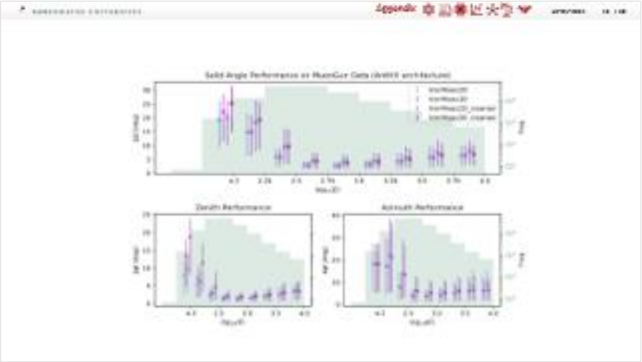


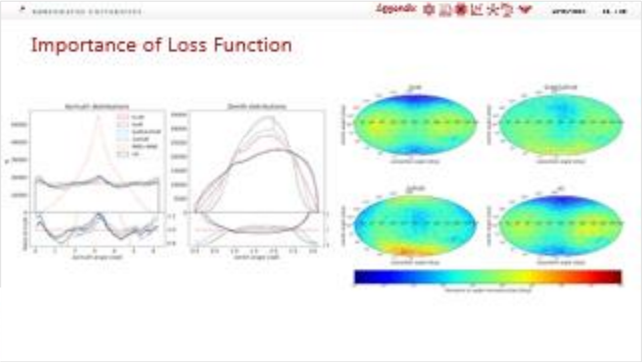

Comparison

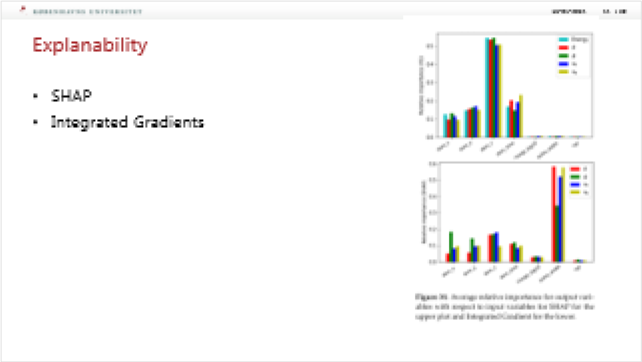
	Loss ($\theta, \phi / x, y, z$)	Loss ($log(E)$)	θ	ϕ	Ω	$log(E)$	Speed	Params (10^3)
StateFarm	2xPvM	MAE	**11.92°**	31.23°	33.37°	**0.1487**	$9 \cdot 10^5$	643
GGConv	vG	Normal	15.94°	**29.65°**	**32.84°**	0.2105	$3 \cdot 10^4$	176
AntHill	2xPvM/SvM	Normal	14.20°	33.53°	36.37°	0.2357	$1 \cdot 10^4$	2,218
LifeGuard	2xPvM+σ	MAE	15.91°	-	-	0.2634	$2 \cdot 10^4$	2.2
Retro	-	-	15.00°	37.93°	40.56°	0.2390	$O(10^{-2})$	-
σ	-	-	0.01°	0.05°	0.04°	0.0003	-	-

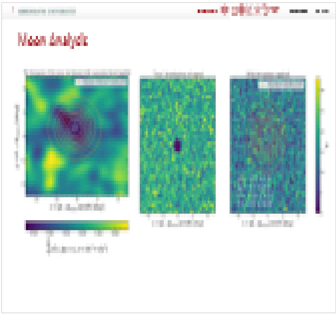




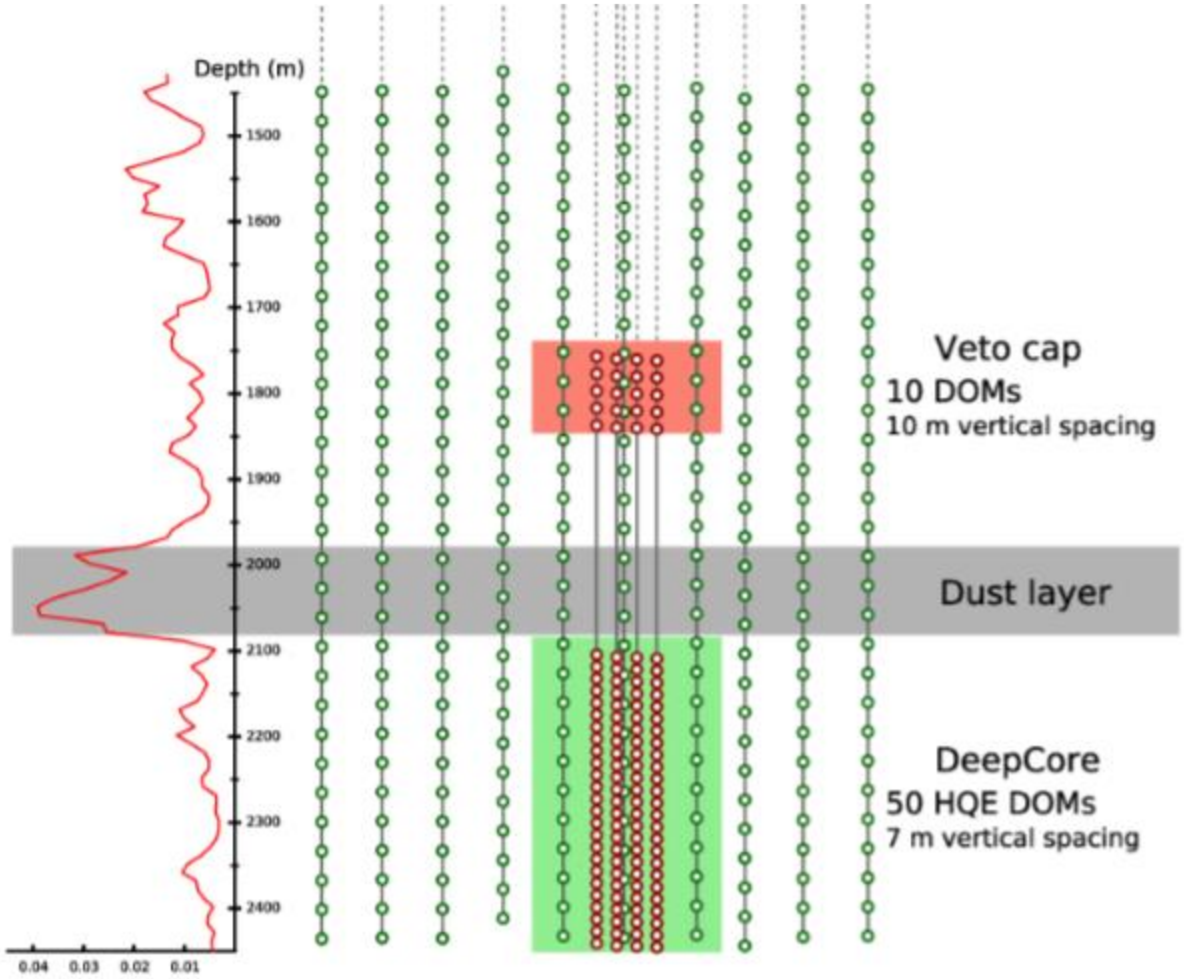




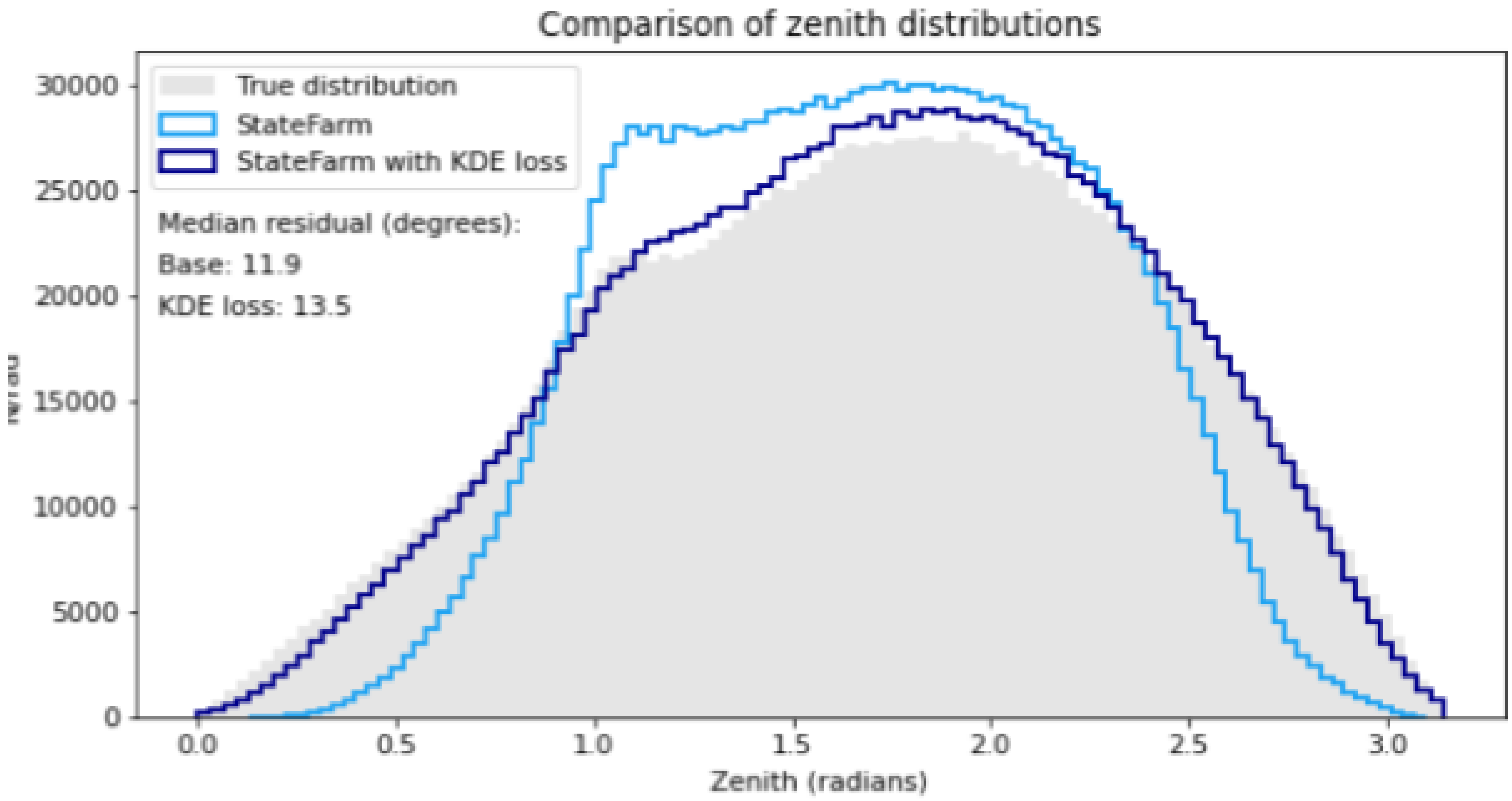




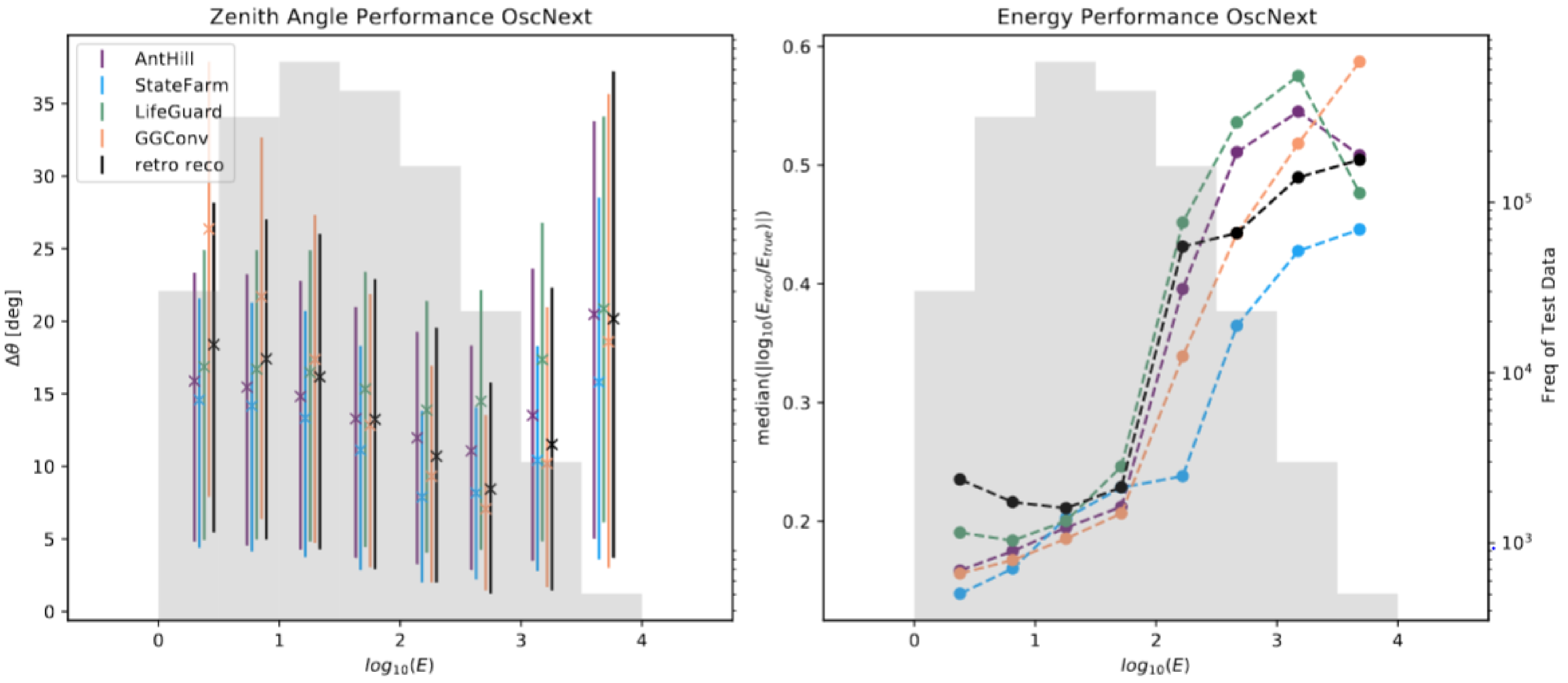
Sensors of IceCube: DOMs (Digital Optical Module)



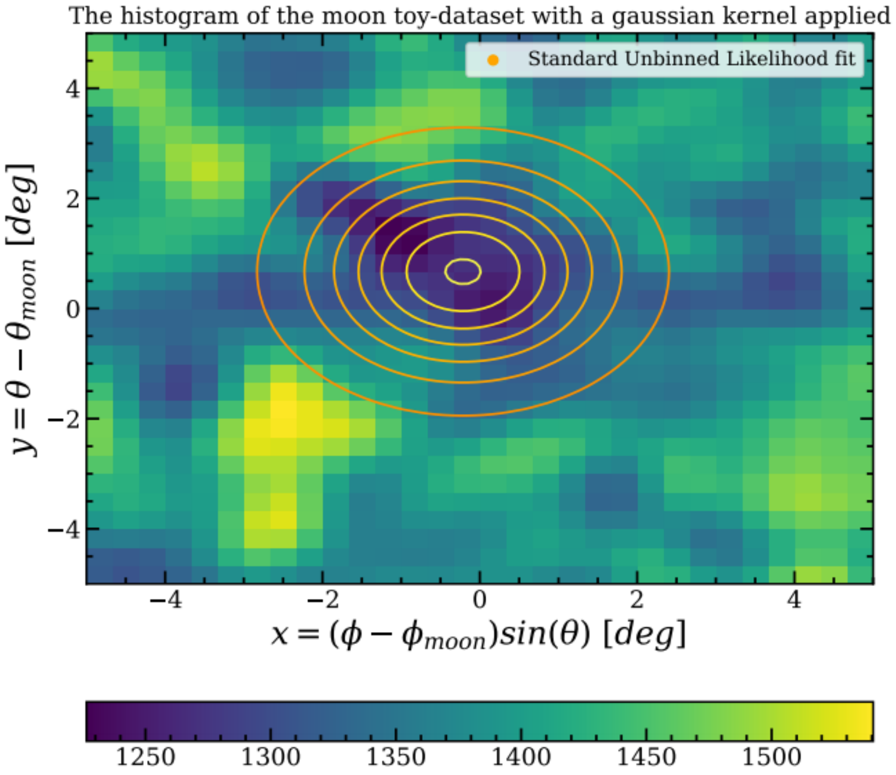
Kernel Density Estimate



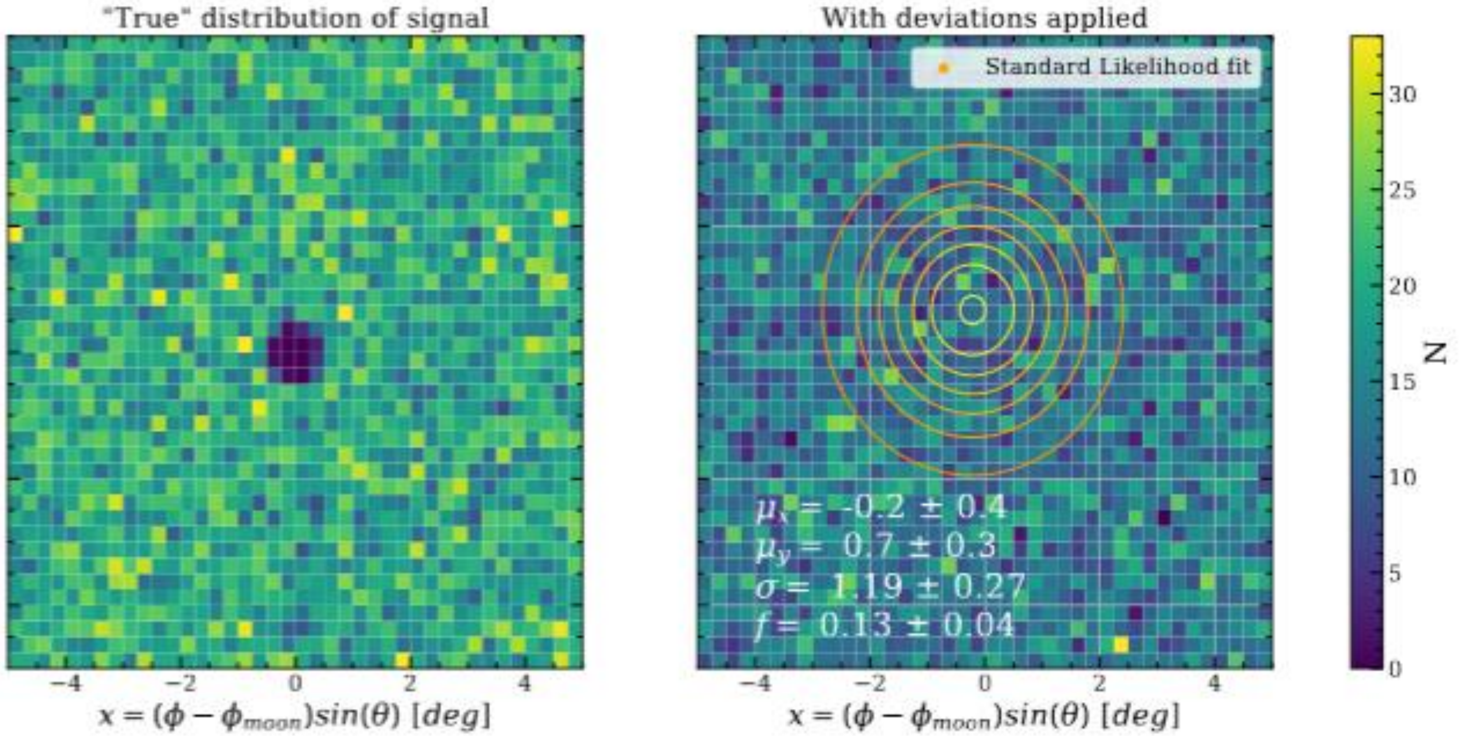
Results


Zenith Angle Performance OscNext
AntHill
StateFarm
LifeGuard
GGConv
retro reco
Δθ [deg]
$log_{10}(E)$
Energy Performance OscNext
median(|$log_{10}(E_{reco}/E_{true})$|)
Freq of Test Data
$log_{10}(E)$

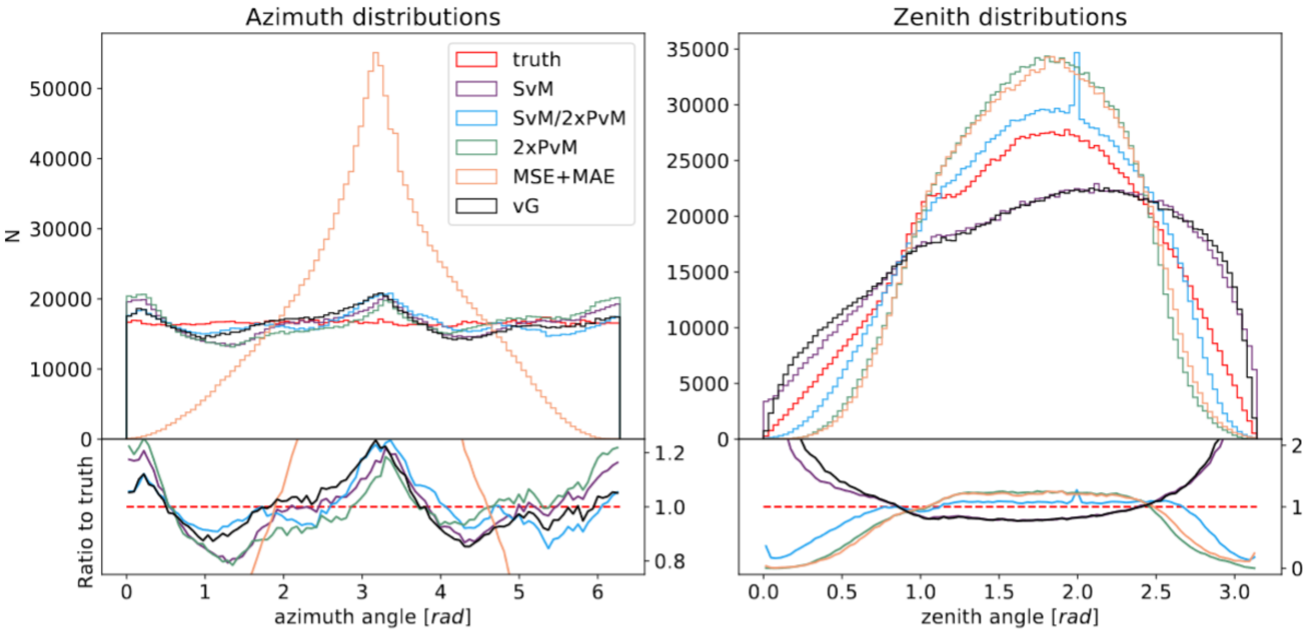
Moon Analysis

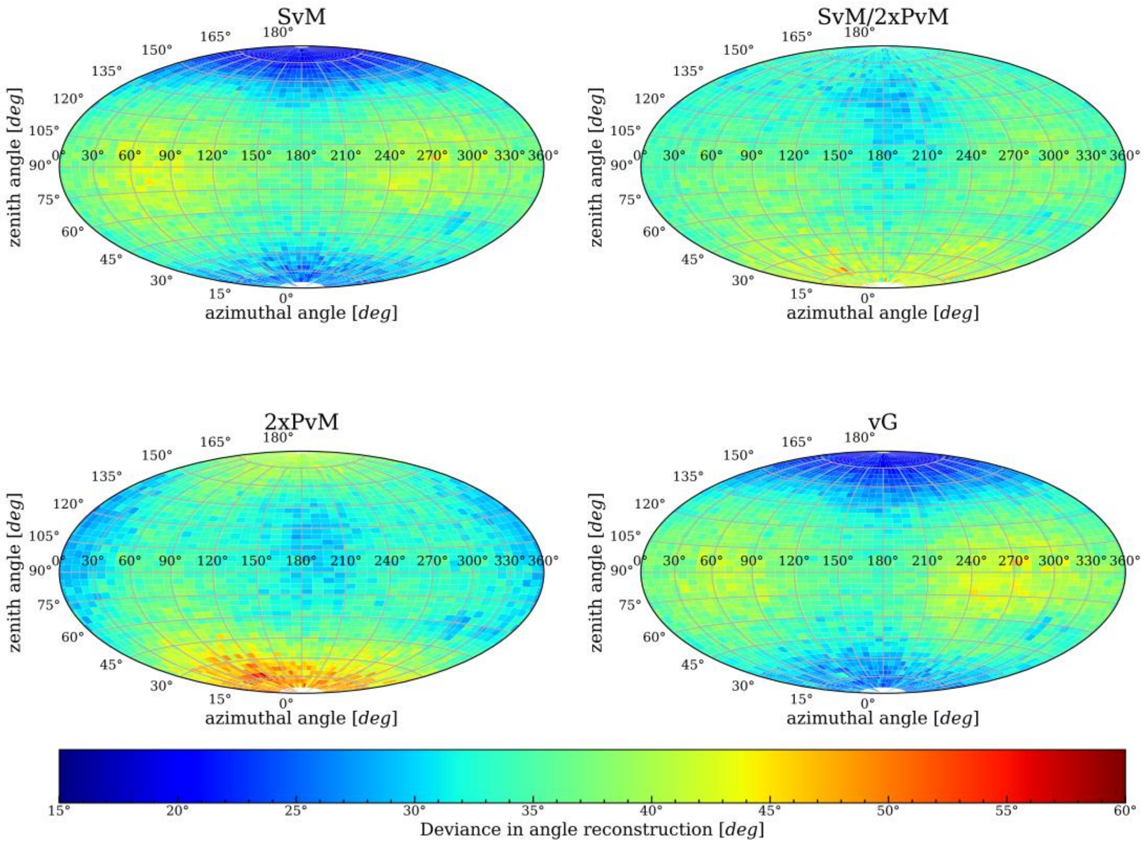


$$\sum_i^N \mathcal{N}(x, y|\mu = x_i, \sigma = dx^2 + dy^2)$$



Importance of Loss Function

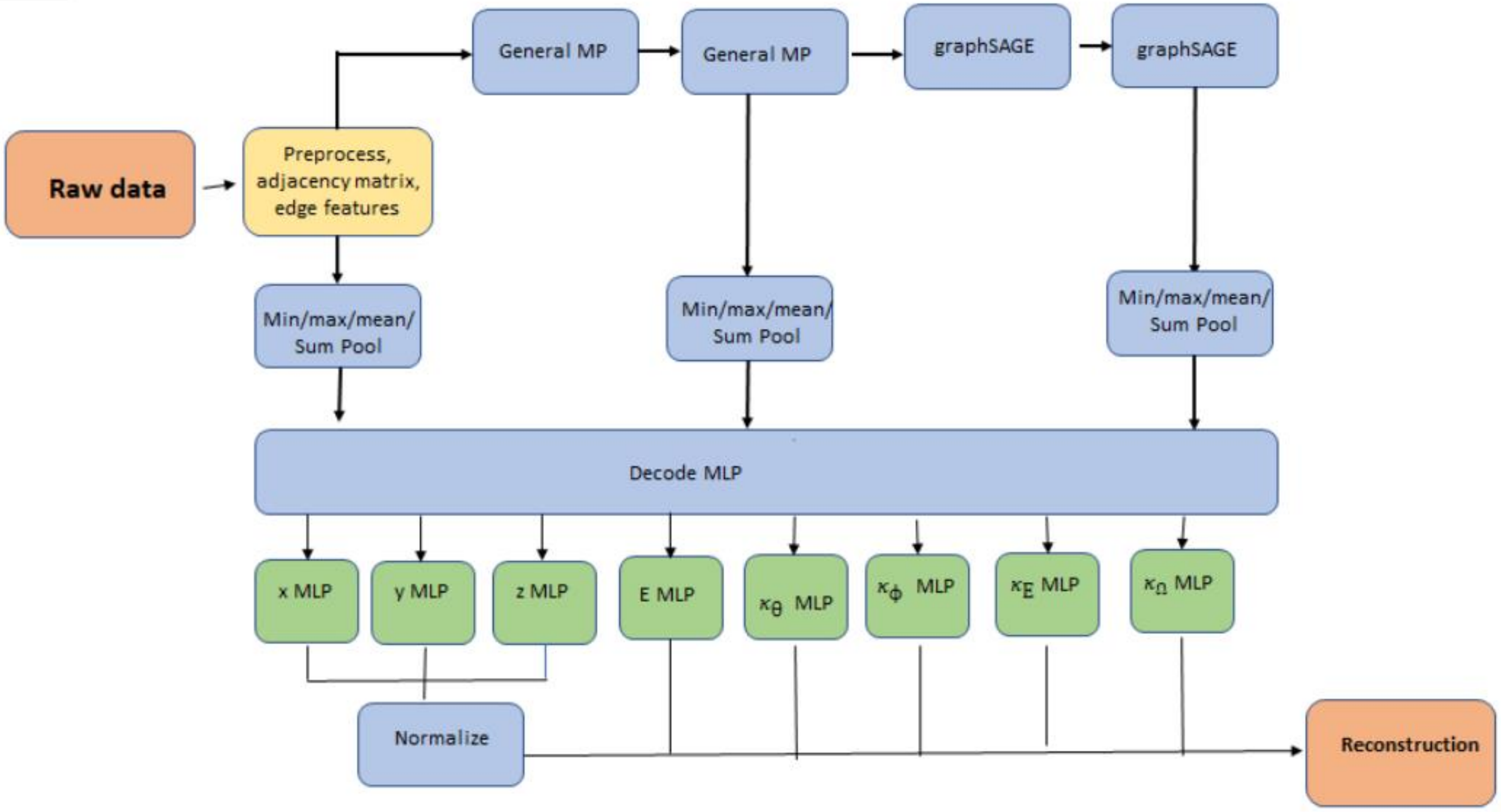




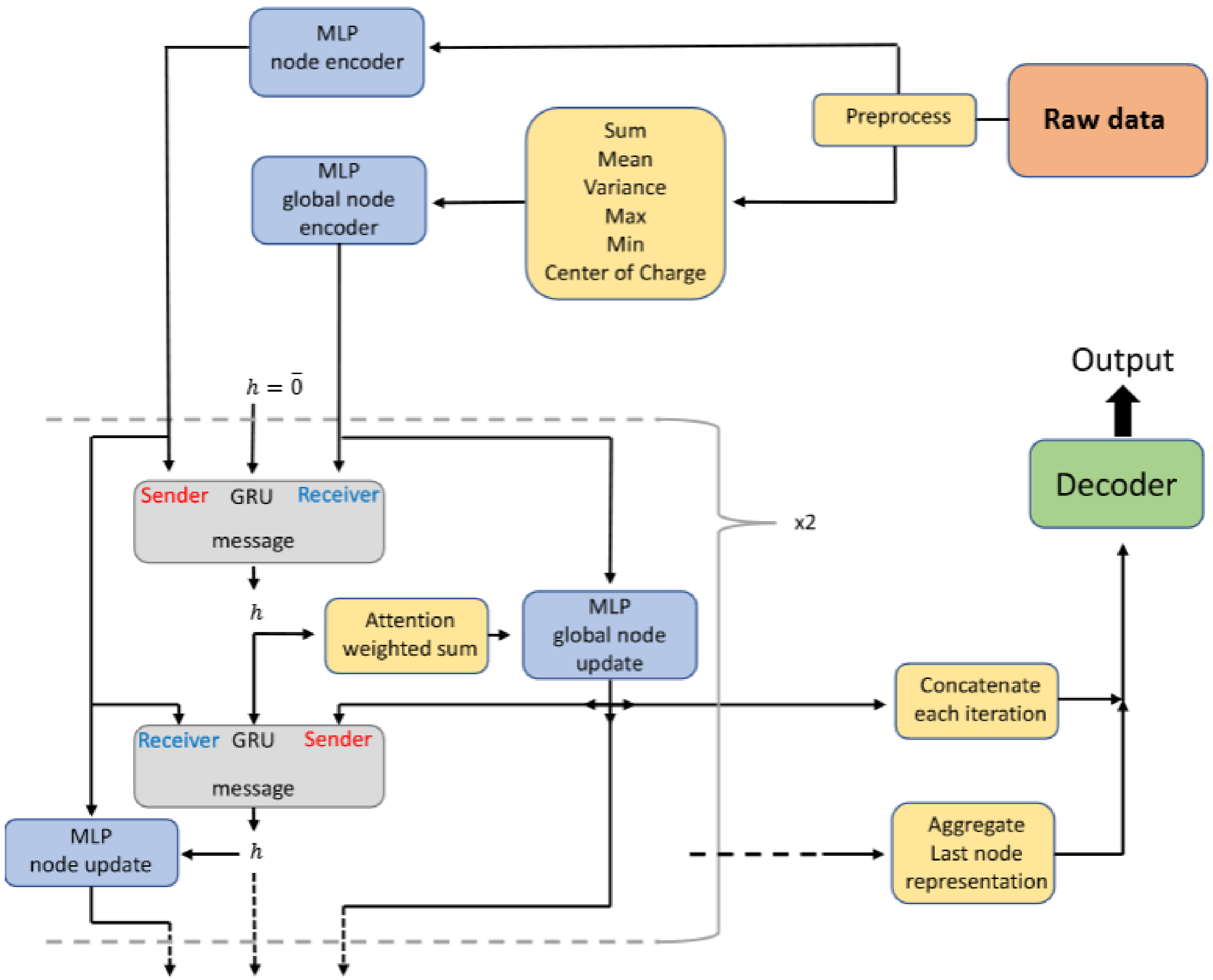
Comparison

	Loss $(\theta, \phi / x, y, z)$	Loss $(log(E))$	θ	ϕ	Ω	$log(E)$	Speed	Params (10^3)
StateFarm	2xPvM	MAE	**11.92** °	31.23 °	33.37 °	**0.1487**	$9 \cdot 10^3$	643
GGConv	vG	Normal	15.94 °	**29.65** °	**32.84** °	0.2105	$3 \cdot 10^4$	176
AntHill	2xPvM/SvM	Normal	14.20 °	33.53 °	36.37 °	0.2357	$1 \cdot 10^4$	2,218
LifeGuard	2xPvM+σ	MAE	15.91°	-	-	0.2634	$2 \cdot 10^4$	2.2
Retro	-	-	15.00 °	37.93 °	40.56 °	0.2390	$O(10^{-2})$	-
σ	-	-	0.01 °	0.05 °	0.04 °	0.0003	-	-

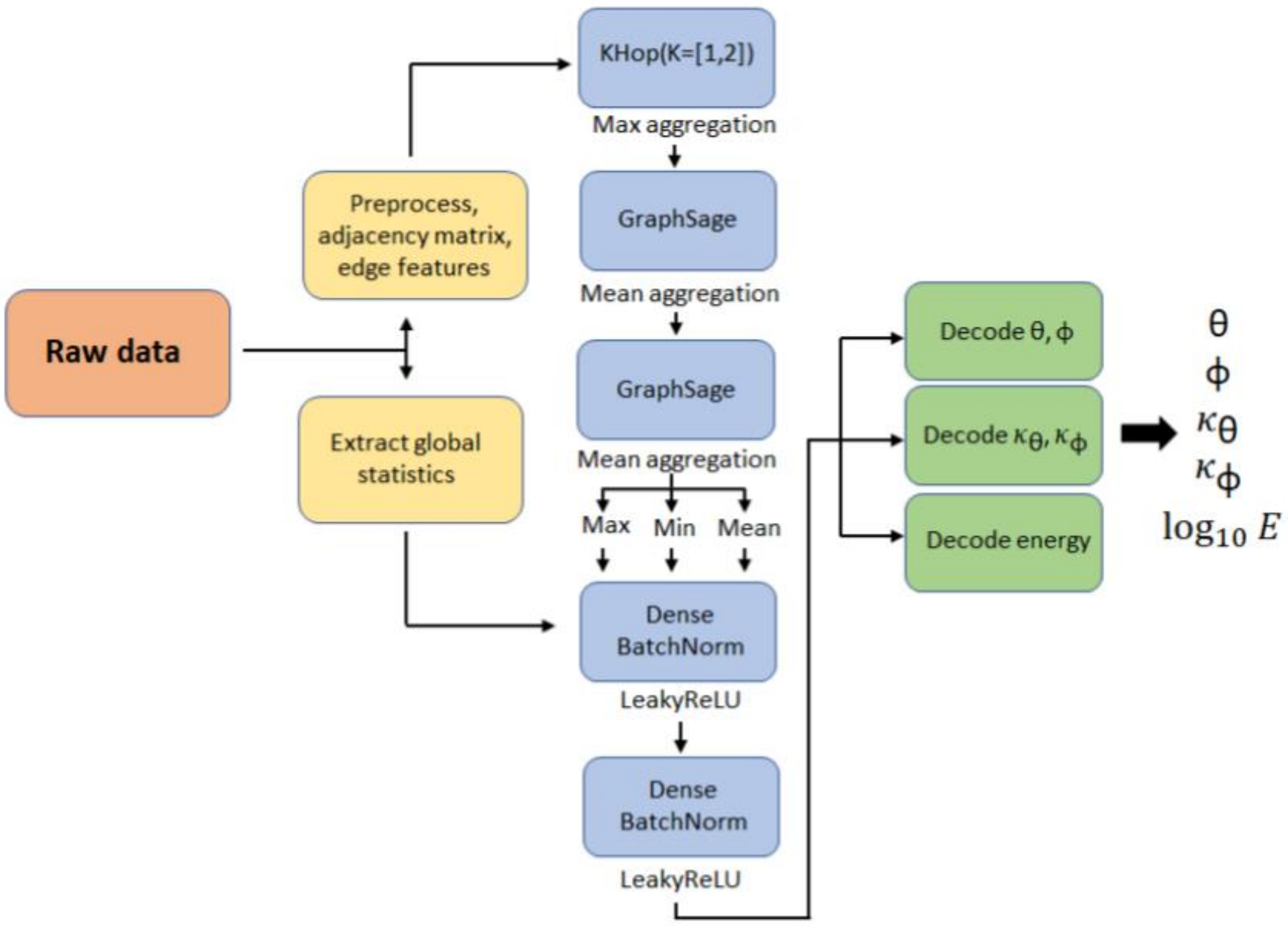
Architectures



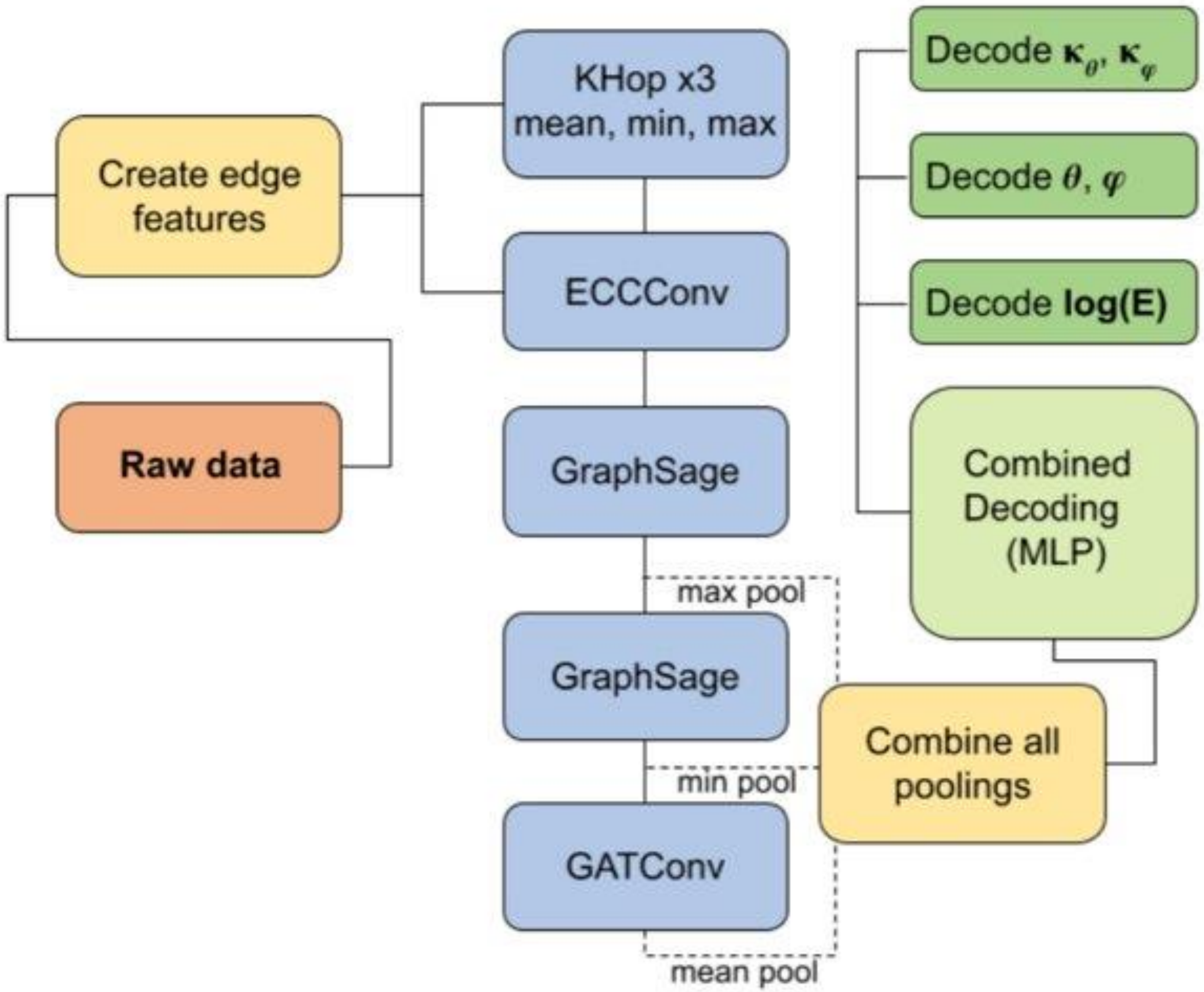
Architectures

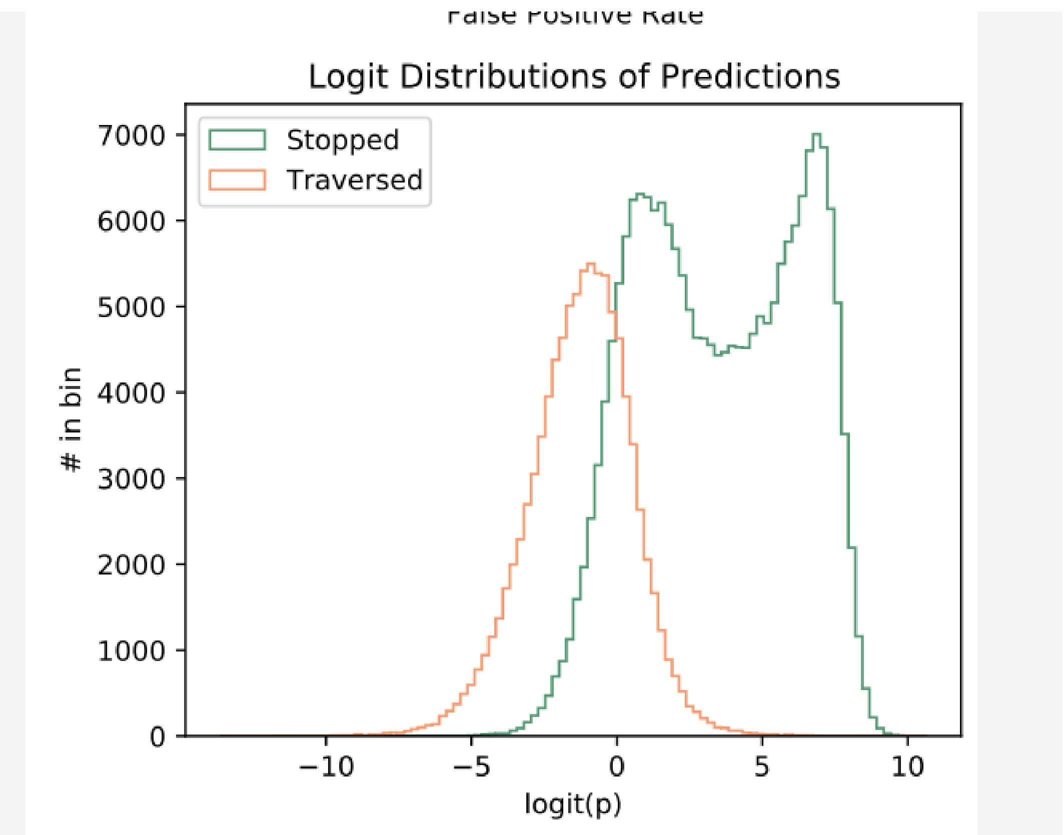


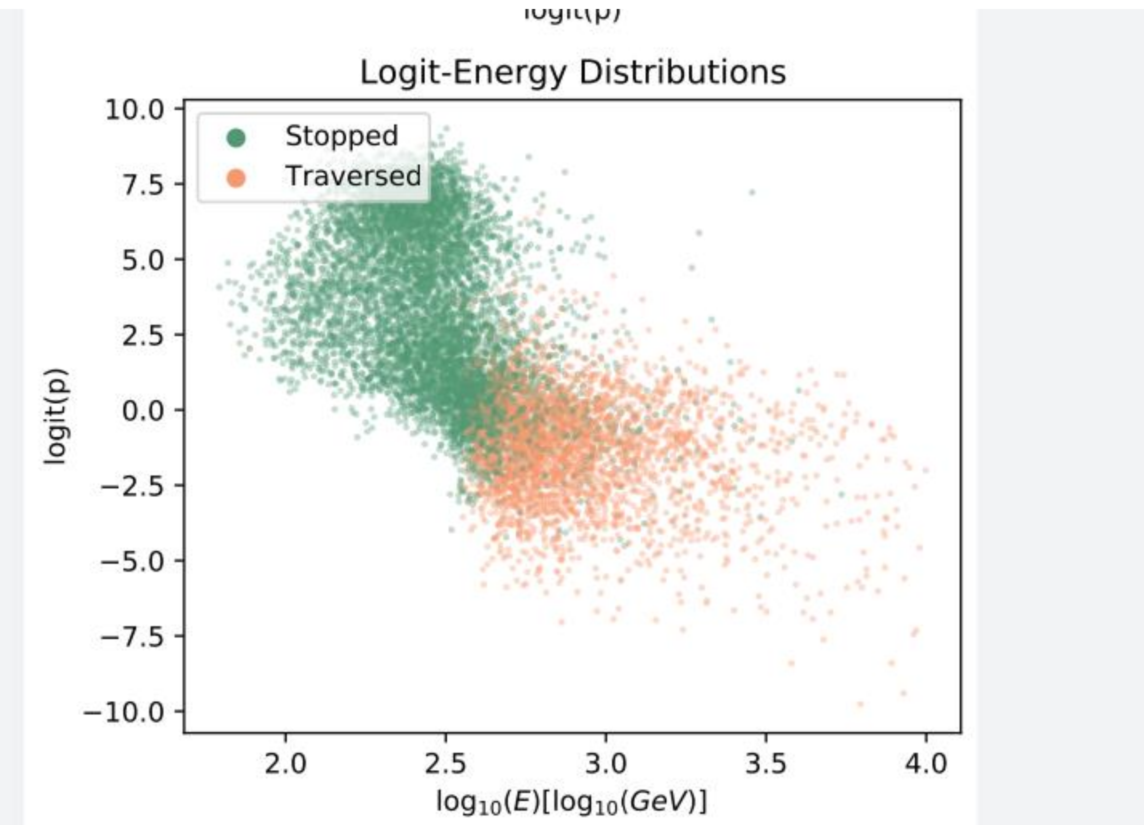
Architectures

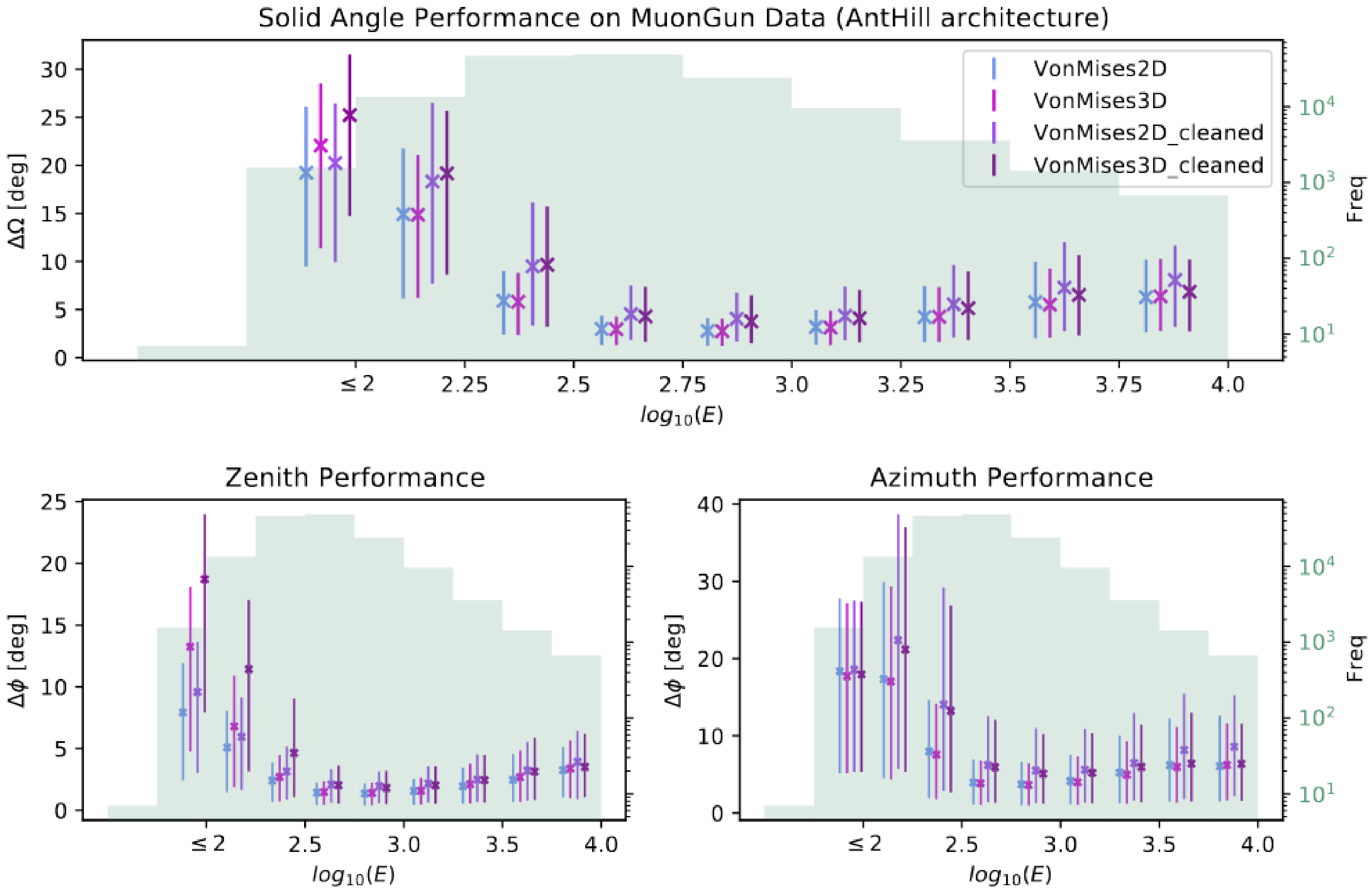


Architectures




Logit Distributions of Predictions
Stopped
Traversed
# in bin
logit(p)
7000
6000
5000
4000
3000
2000
1000
0
−10
−5
0
5
10


Logit-Energy Distributions
Stopped
Traversed
logit(p)
$\log_{10}(E)[\log_{10}(GeV)]$
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
2.0
2.5
3.0
3.5
4.0


Solid Angle Performance on MuonGun Data (AntHill architecture)
VonMises2D
VonMises3D
VonMises2D_cleaned
VonMises3D_cleaned
ΔΩ [deg]
Freq
log10(E)
Zenith Performance
Δϕ [deg]
log10(E)
Azimuth Performance
Δϕ [deg]
Freq
log10(E)

Explanability

- SHAP
- Integrated Gradients

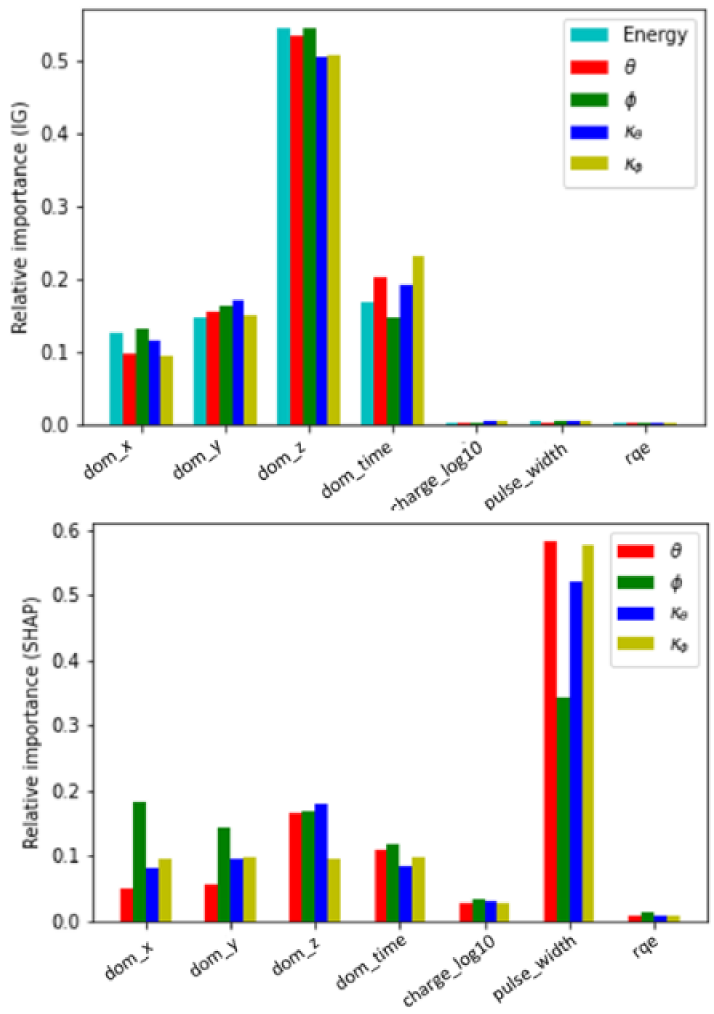


Figure 31: Average relative importance for output variables with respect to input variables for SHAP for the upper plot and Integrated Gradient for the lower.