BAYESIAN NEURAL NETWORK APPLIED TO CHERENKOV EVENT RECONSTRUCTION

L. Perisse[1], A. Beauchêne[2], C. Ehrhardt[3], A. Ershova[2], E. Le Blévec[1], C. Quach[2], B. Quilain[1], A. Voulgari-Revof[3]

(1) ILANCE, CNRS - University of Tokyo, Japan (3) Master interns at ILANCE
(2) Ecole Polytechnique, IN2P3 - CNRS, Laboratoire Leprince-Ringuet, France

BAYESIAN NEURAL NETWORKS

Nextgen Cherenkov experiments need improved algorithms to reduce statistical and systematic uncertainties → Machine Learning

Artificial Neural Networks (ANN) are deterministic

- ✓ Faster than traditional algorithms once trained
- ✗ Black box algorithms
- ✗ Optimization solution depends on starting point

Bayesian Neural Networks (BNN) generalize uncertainty in NN

- ✓ Epistemic uncertainty (parameter uncertainty, dataset size)
- ✓ Robust to over-fitting
- ✗ BNN not yet investigated in Cherenkov low energy regression

BNN predict distributions with uncertainties related to their confidence about their own outputs

ANN
Language: PyTorch
Loss: depends on problem

↓

Pyro[1]: BNN library build on Pytorch

TyXe[2]: framework build on Pyro

↓

BNN
IID priors $\mathcal{N}(0,1)$
Loss: ELBO

Bayes' Theorem

Given inputs $X = \{x_1, \dots, x_N\}$ and outputs $Y = \{y_1, \dots, y_N\}$, we search parameters ω **likely to generate** $Y = f^{\omega}(X)$.
Priors $p(\omega)$ is modified according to the **likehood** $p(Y|X,\omega)$ to accustom parameter likeliness given observed data Y.
Posteriors over ω then obtained as: $p(\omega|X,Y) = \frac{p(Y|X,\omega)p(\omega)}{p(Y|X)}$

Stochastic Variational Inference (SVI)

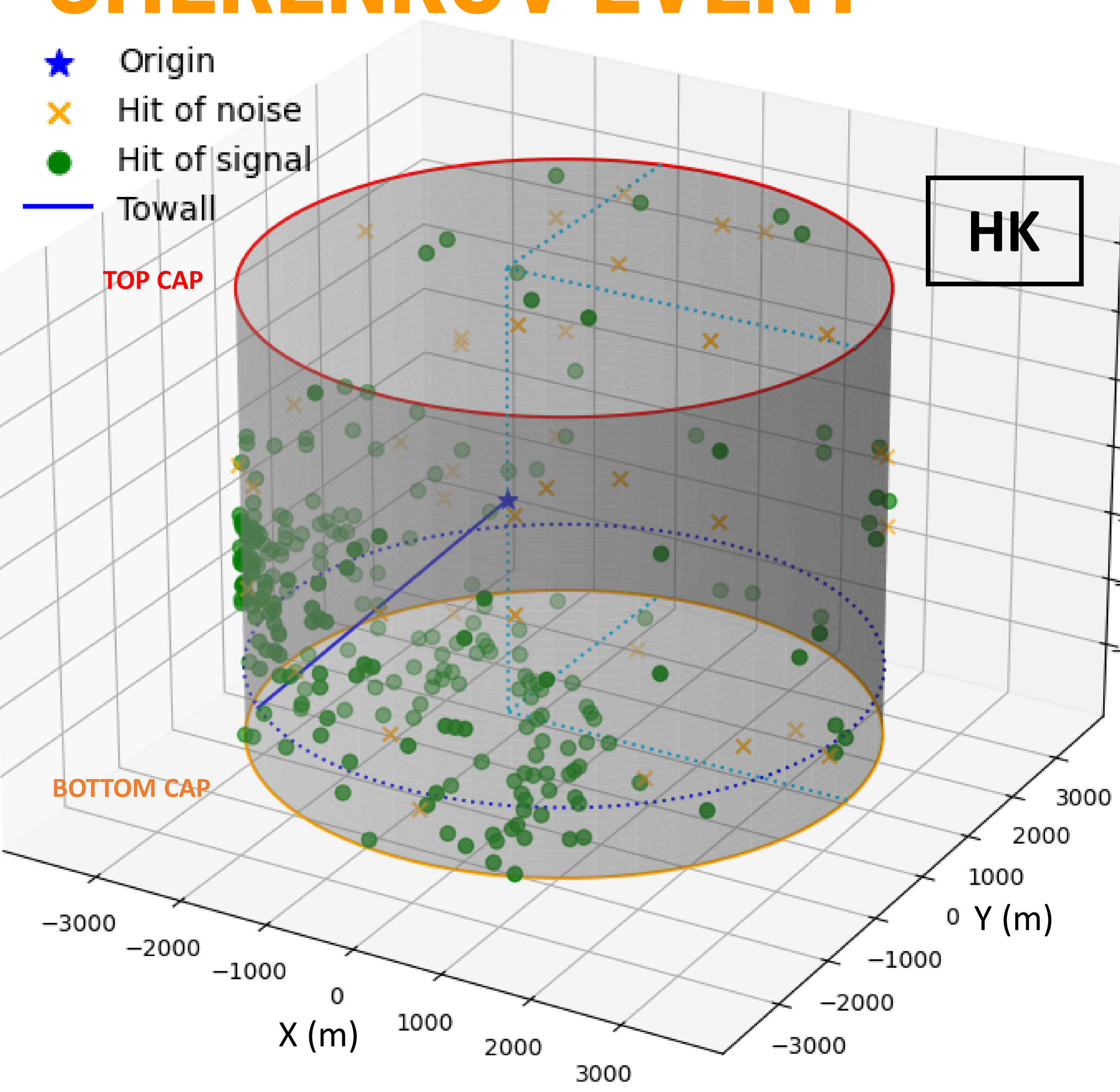
True posterior approximated by **variational dist.** $q_\theta(\omega)$ parametrized by θ.
Minimize Evidence Lower Bound (ELBO) loss = **Variational Inference**[3]:

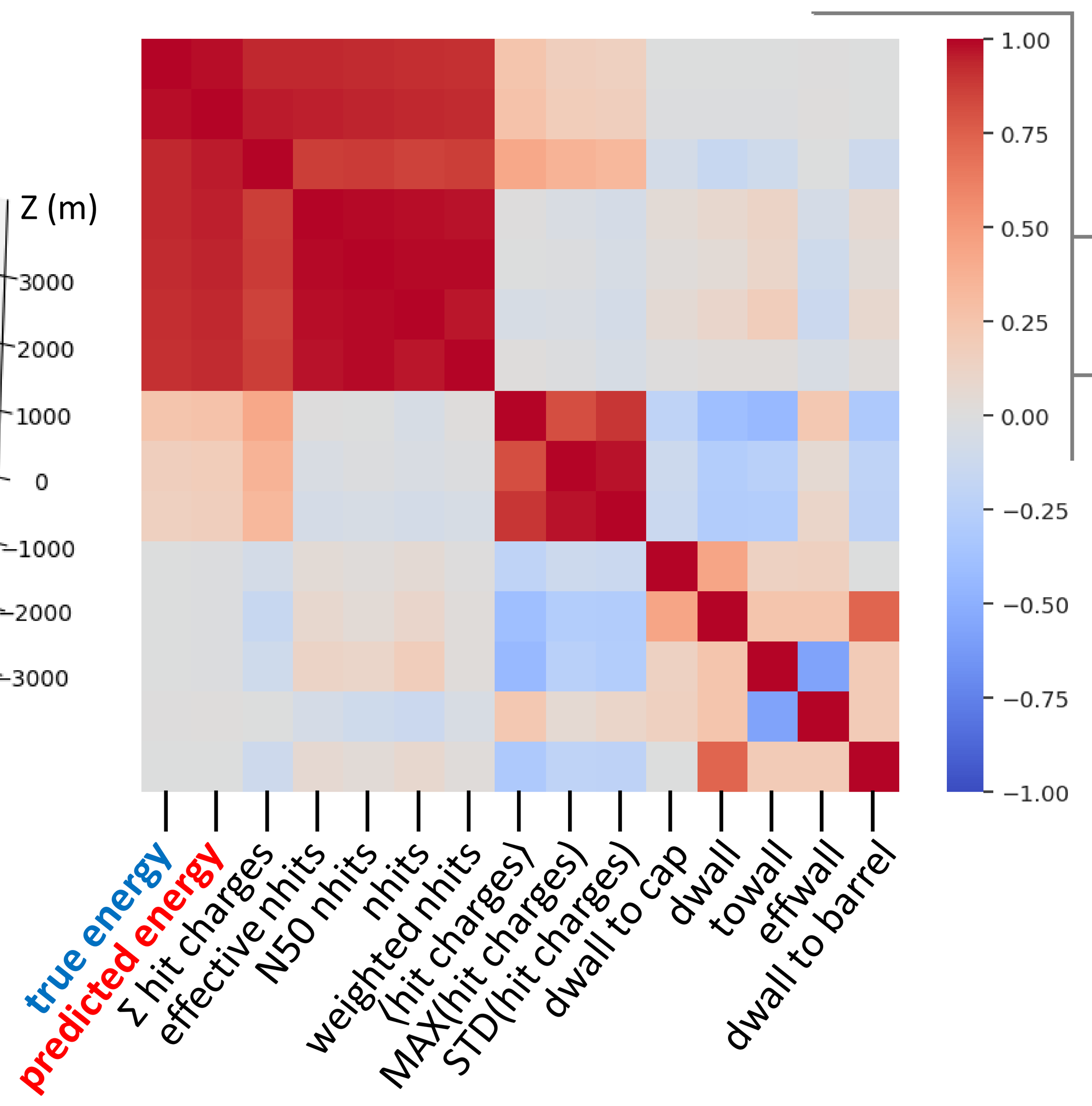
$$\mathrm{ELBO}(\theta) = \mathrm{KL}[q_\theta(\omega)||p(\omega)] - \int q_\theta(\omega)\log[p(Y|X,\omega)]d\omega$$

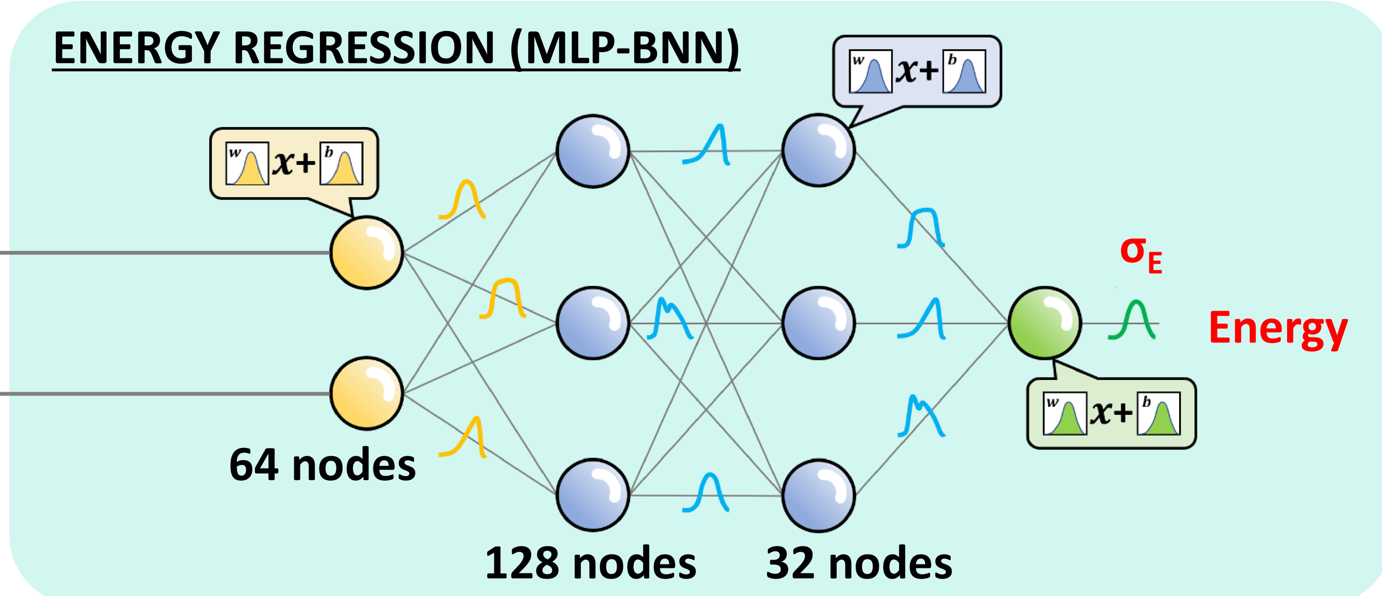
Kullback–Leibler divergence: Measure of dist. similarity[4]
Minimizing favors $q_\theta(\omega)$ to be close to prior

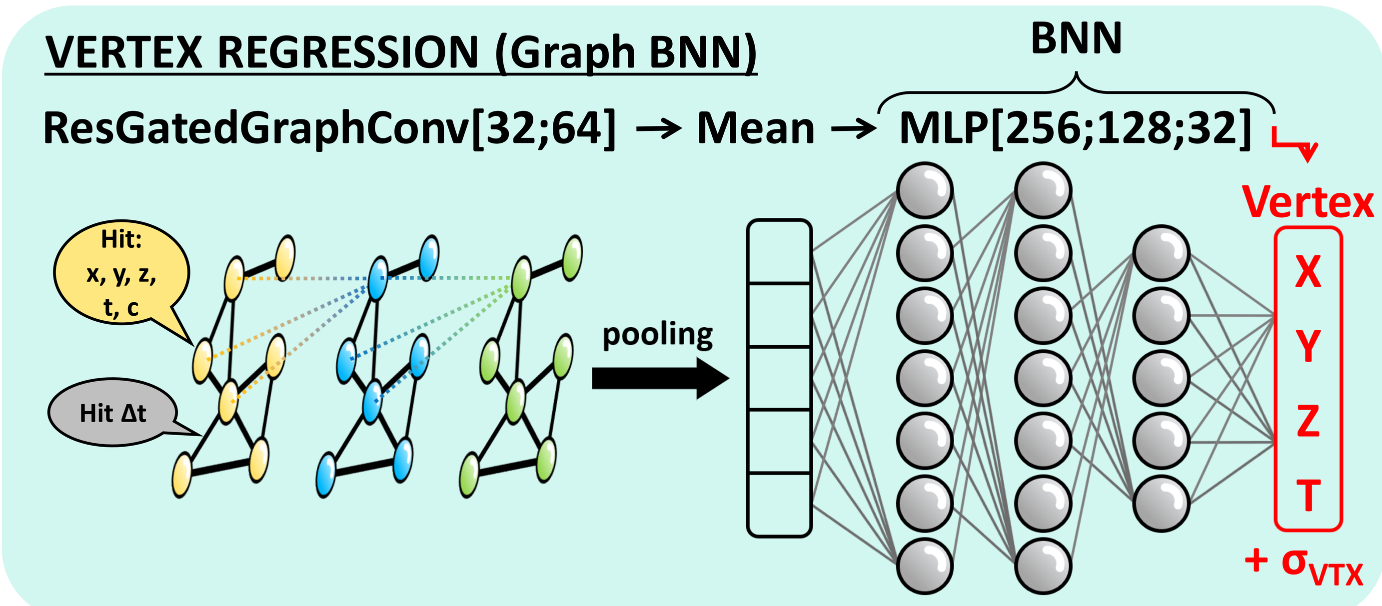
Expected log likelihood:
Maximizing favors $q_\theta(\omega)$ to explain data

CHERENKOV EVENT



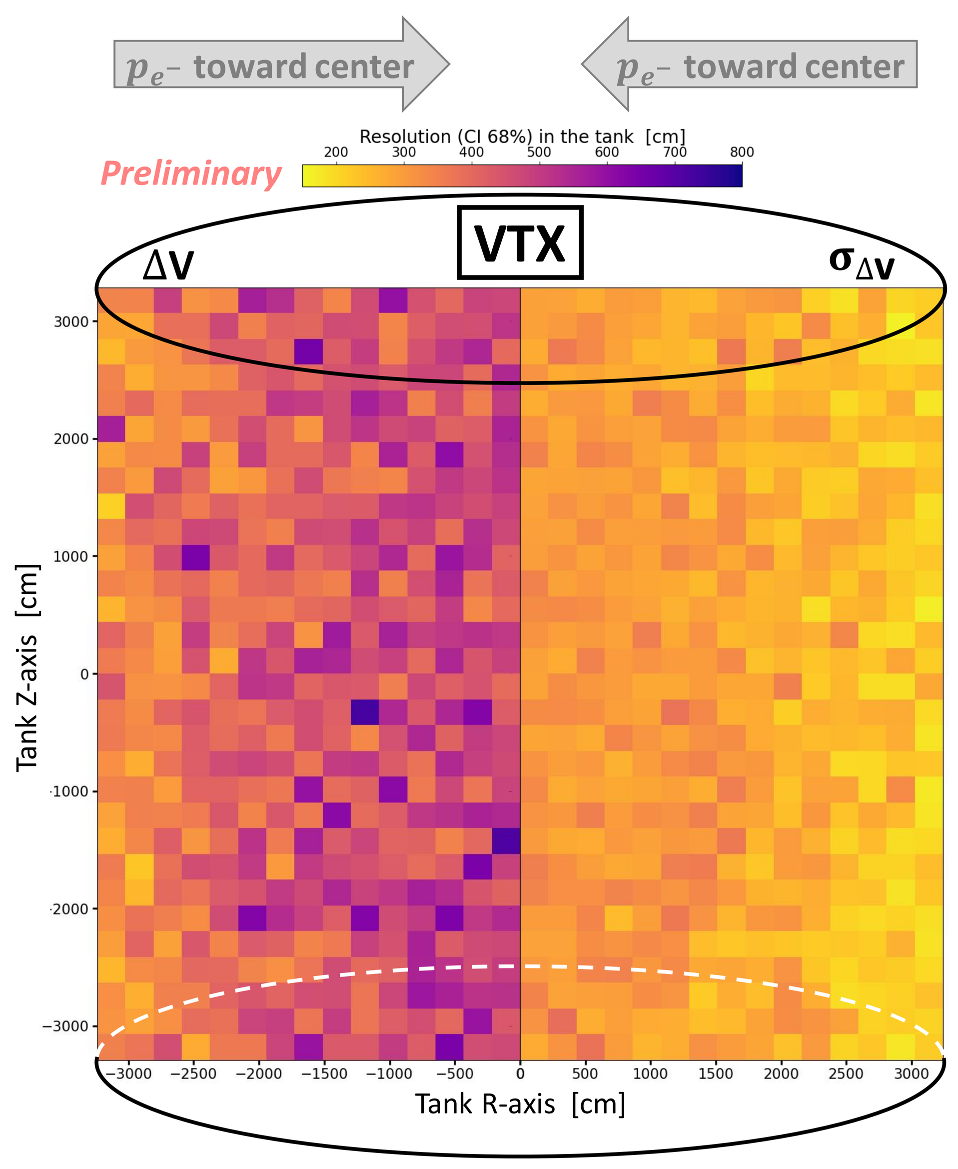


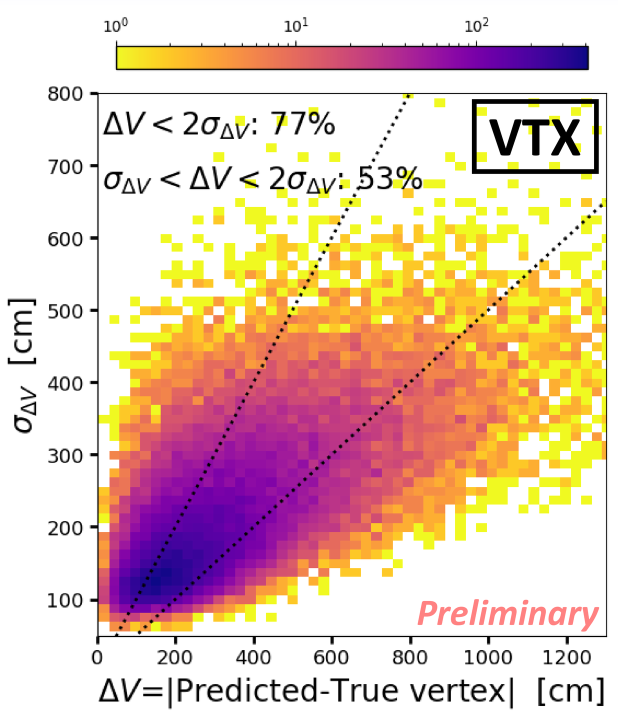


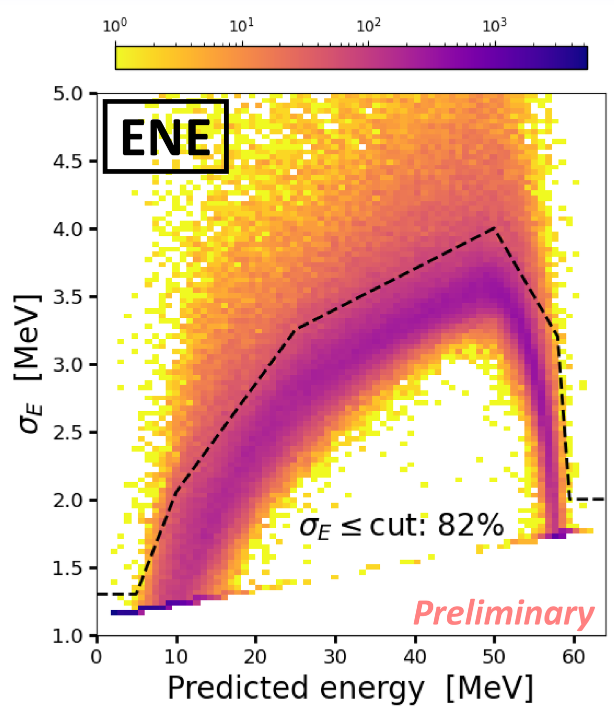


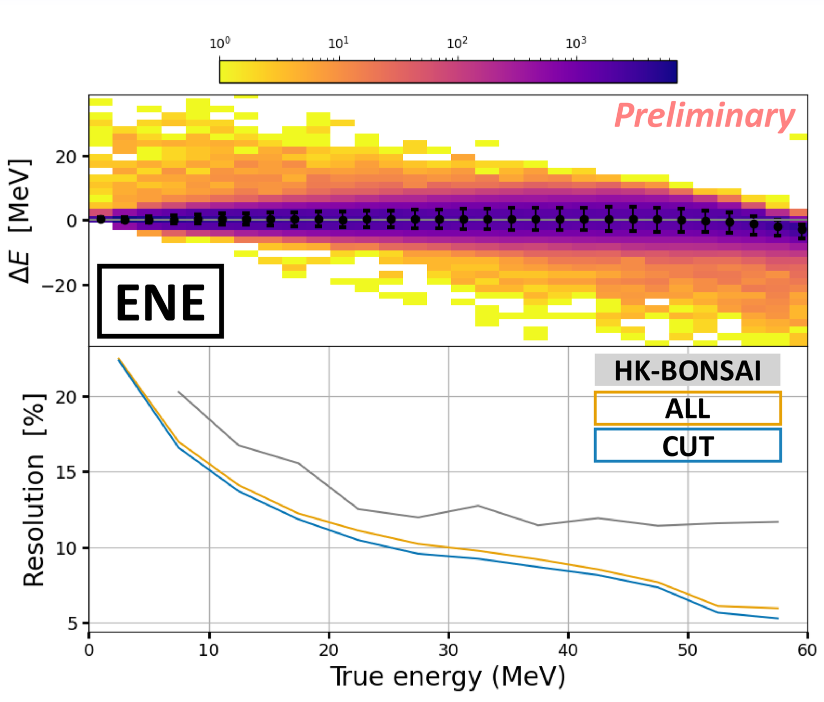
- Hyper-Kamiokande MC (WCSim)
- 500k e^- events (50% train, 50% test)
- Flat kinetic energy 0-60 MeV or fixed at 10 MeV
- Uniform vertex distribution + isotropic direction

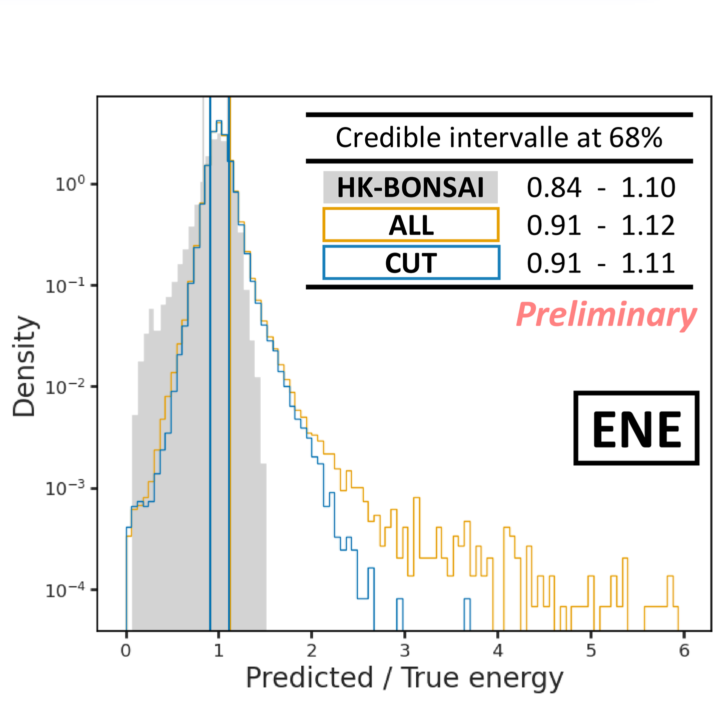
RESULTS











Resolution at 68% (all|cut)		ENERGY (@30 MeV)	VERTEX (@10 MeV) ΔV=$\|r-v\|$	Δt
μ	HK-BONSAI	3.6 MeV	0.7 m	-
	This work	3.0|2.7 MeV	3.1|2.5 m	8|2 ns
	Non-Bayesian version		*2.9 m*	*8 ns*
σ	This work	3.1 MeV	2.5 m	10 ns

- σ_E ~ fully correlated between events → energy spec. corr. matrix
- Regression time decreased compared to traditional algorithm
- Non-Bayesian version of our GNN shows similar regression time
- Vertex regression @500 MeV with similar non-Bayesian GNN ~2 m

VERTEX	$\sim 10^{-3}$ **sec/event/sample**
ENERGY	$\sim 10^{-6}$ **sec/event/sample**
HK-BONSAI	$\sim 10^{-1}$ **sec/event**

CONCLUSION & PERSPECTIVE

- ✓ Bayesian GNN-MLP works for **low energy event multidimensional regression**
- ✓ Relation between larger $\sigma_{\Delta V}$ and larger vertex difference → σ can be used as a discriminant
- ✗ σ_E not yet representative of model uncertainty ($\sigma_E \gg \Delta E$, dense location below cut)
- ✗ Vertex regression results too far from true vertex, but room for improvement
- › Identify key parameters impacting σ_E in our models + apply GNN to energy reconstruction
- › Optimization of models and hyperparameters (WatchMaL, W&B framework)

[1]: E. Bingham *et al.*, Journal of Machine Learning Research, 20(28):1–6 (2019)
[2]: H. Ritter & T. Karaletsos, Proceedings of Machine Learning and Systems (2022)
[3]: M. Jordan *et al.*, Machine learning, 37(2):183–233 (1999)
[4]: S. Kullback & R. A. Leibler, The annals of mathematical statistics, 22(1):79–86 (1951)