

Generating Accurate Showers in Highly Granular Calorimeters Using Normalizing Flows

Thorsten Buss, Sascha Diefenbacher, Engin Eren, Frank Gaede,
Gregor Kasieczka, Claudius Krause, Imahn Shekhzadeh, David Shih

thorsten.buss@uni-hamburg.de

Institut für Experimentalphysik
Universität Hamburg

March 23, 2023



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE



Detector Simulation

Detector Simulation

International Large Detector

Architecture

Upscaling

Preliminary Results

Summary & Outlook

- monte carlo (MC) necessary to compare theory and measurements
- detector simulation most expensive part of simulation chain
- computational requirements expected to exceed available resources soon
 - need for speeding up detector simulation
- generative neuronal networks learn distributions and can sample from them
- work flow:
 - simulate small amounts of data using slow monte carlo
 - train generative model on these data
 - draw large amounts of data from fast ML models

International Large Detector (ILD)

Detector Simulation

International Large Detector

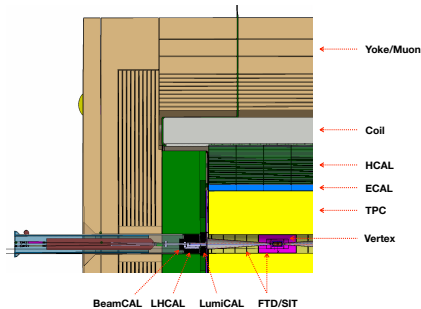
Architecture

Upscaling

Preliminary Results

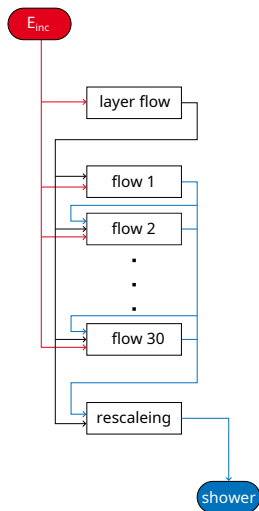
Summary & Outlook

- proposed detector for the ILC
- has two sampling calorimeters
 - electromagnetic calorimeter
 - 30 layers, 5mm x 5mm cells
 - hadronic calorimeter
 - 48 layers, 30mm x 30mm cells
- MC simulation especially slow
 - speed up calorimeter simulation



- based on CaloFlow² and L2L Flow³
- one layer flow
 - learns distribution of layer energies
 - conditioned on incident energy
- 30 multiple flows
 - learn shower shape in layer
 - conditioned on
 - incident energy
 - layer energy
 - previous layers
 - summary network reduces cardinality
- generation
 - sample layer energies using layer flow
 - sample shower shape using multiple flows
 - rescale voxel energies

Architecture

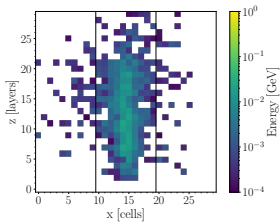
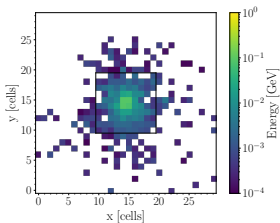


²Claudius Krause and David Shih. *CaloFlow: Fast and Accurate Generation of Calorimeter Showers with Normalizing Flows*. 2021. arXiv: 2106.05285.

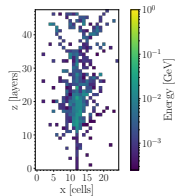
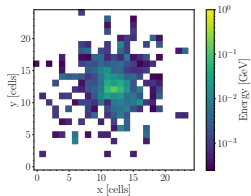
³Sascha Diefenbacher et al. *L2LFlows: Generating High-Fidelity 3D Calorimeter Images*. 2023. arXiv: 2302.11594.

Upscaling

- scaling up to 30x30x30 voxels
- training gets unstable
 - apply gradient clipping
- improve training using a LR scheduler
- features in energy spectrum are smeared out
 - apply element with function to get them back

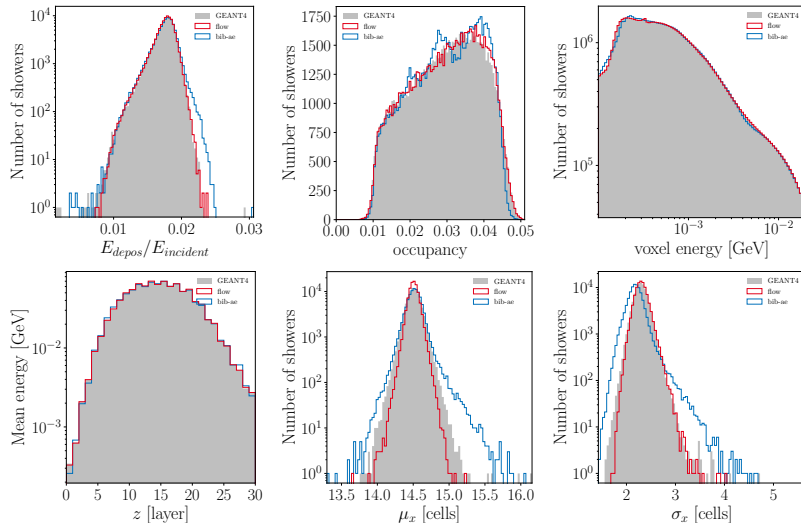


photon shower



pion shower

Preliminary Results



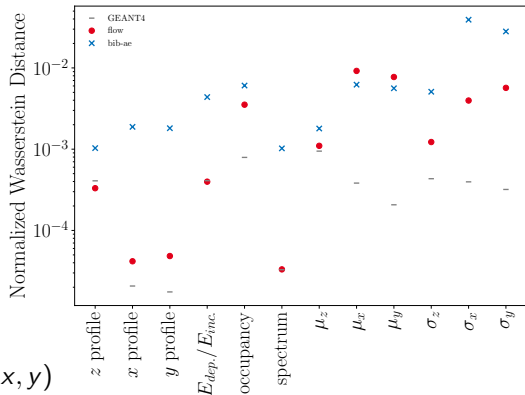
High Level Classifier:

	AUC	Acc	JSD
flow	0.71	0.65	0.11
bib-ae	0.90	0.81	0.43

$$W(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu, \nu)} \int d(x, y) d\gamma(x, y)$$

$$JSD(P||Q) := \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M)$$

Metrics



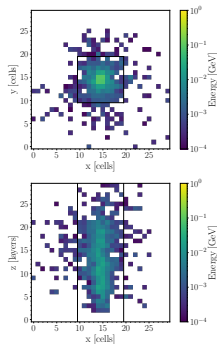
Summary & Outlook

Summary:

- ML generators fast alternative to MC simulations
- the ILD has highly granular calorimeter
 - hard to learn
- flows can generate highly accurate showers

Outlook:

- generate hadron showers
- combine electromagnetic and hadronic calorimeter
- train inverse autoregressive flows to speed up generation
- improve architecture to save weights, computing time



Normalizing Flows

- diffeomorphism between physics space and latent space
- transform physics space distribution into a simple prior distribution
- change of variables formula allows for physics space density estimation
- training: minimize negative log-likelihood
- generation: sample from latent distribution and apply inverse of flow

$$p(x) = q(f(x)) |J(x)|$$

$$\mathcal{L} = -\log q(f(x)) - \log|J(x)|$$

