



VAE Results

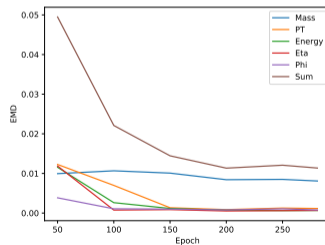
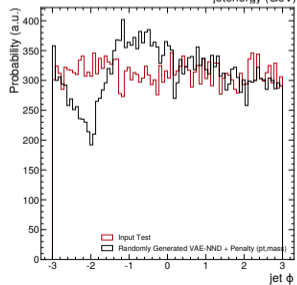
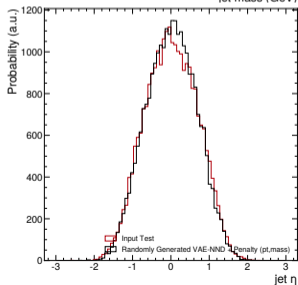
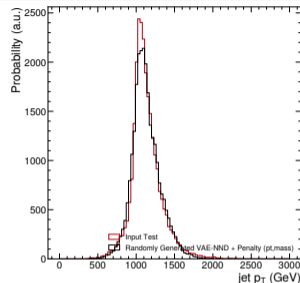
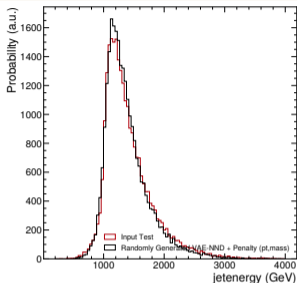
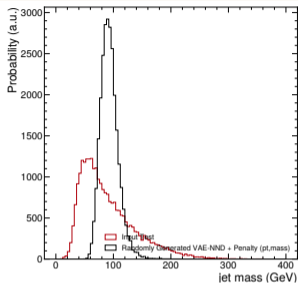
Breno Orzari

Sprace

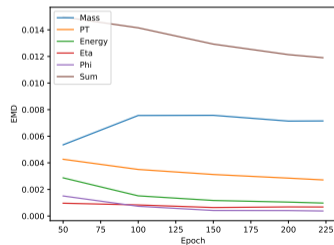
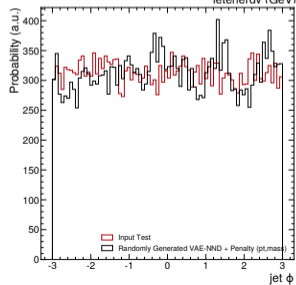
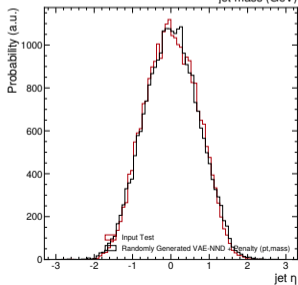
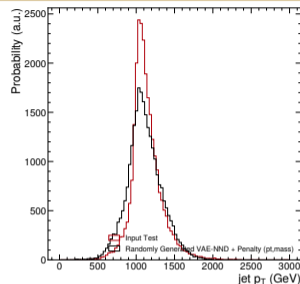
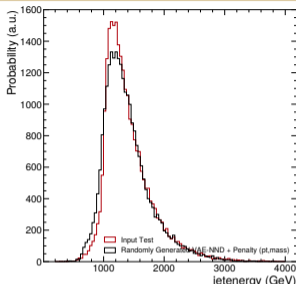
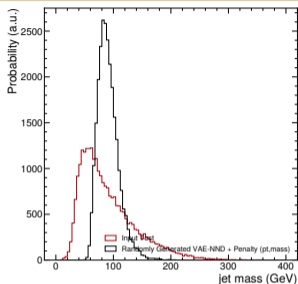
VAE with jets

- Performed in the particles features a per feature normalization from 0.0 to 1.0
 - $p_i^N = \frac{p_i - \min(p_i)}{\max(p_i) - \min(p_i)}$
- Test inverting normalization in loss function
 - $p_i = p_i^N \times (\max(p_i) - \min(p_i)) + \min(p_i)$
 - With early stopping:
 - $\beta=10000$
 - $\beta=5000$
 - $\beta=0.9998$; $\gamma_{mass}=10.0$
 - Masking as fourth feature
 - 1500 epochs:
 - $\beta=5000$
 - $\beta=5000$; only jet mass in loss function
 - $\beta=0.9998$; $\gamma_{mass}=10.0$
 - Masking as fourth feature
 - 3000 epochs:
 - Masking as fourth feature

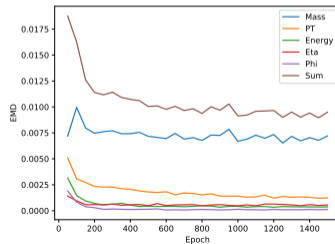
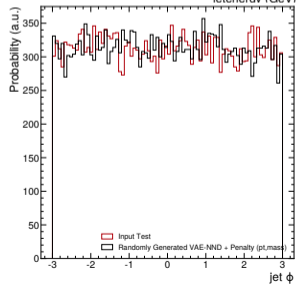
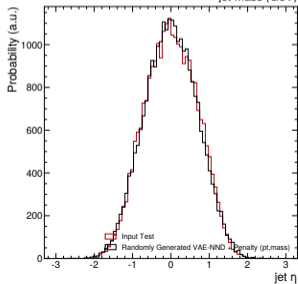
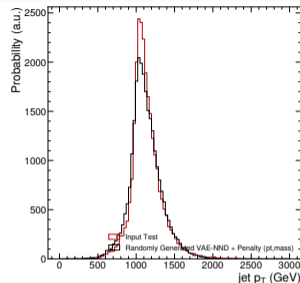
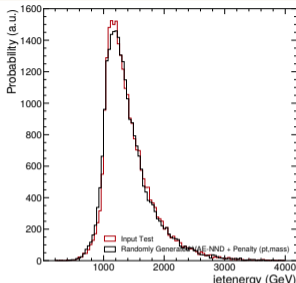
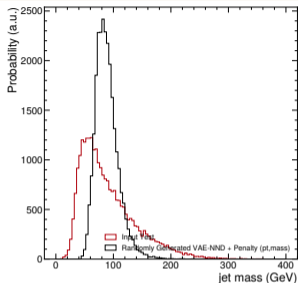
With early stopping: $\beta=10000$



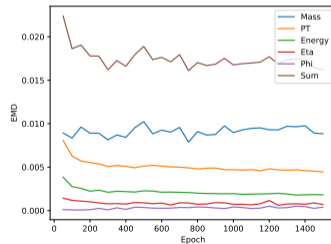
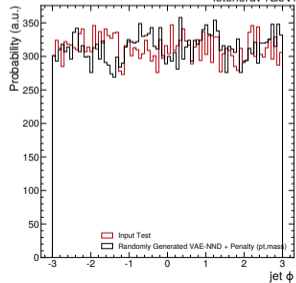
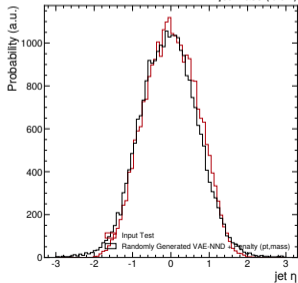
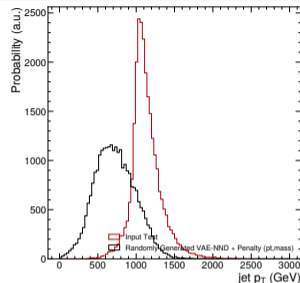
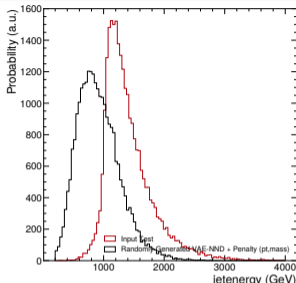
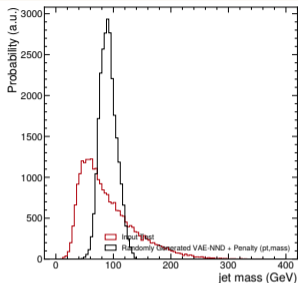
With early stopping: $\beta=5000$



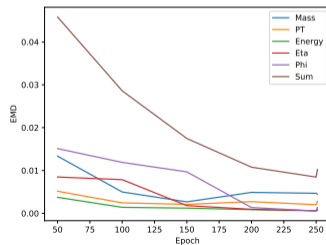
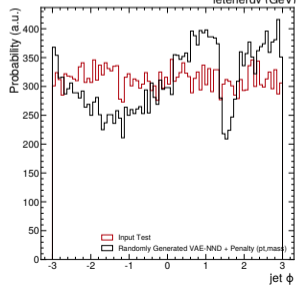
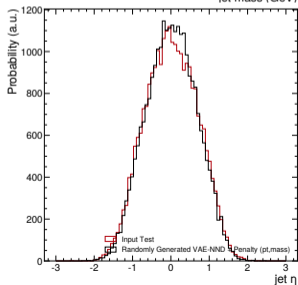
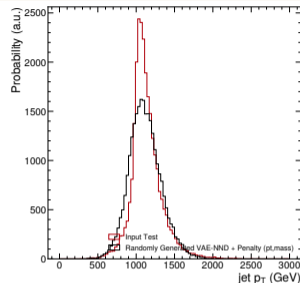
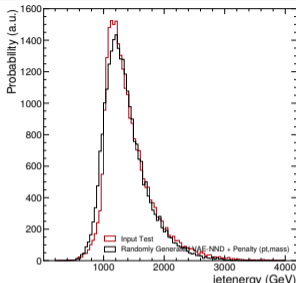
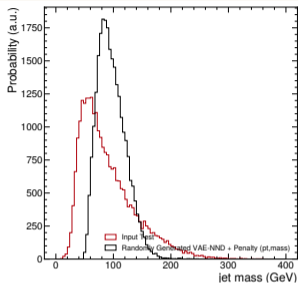
With 1500 epochs: $\beta=5000$



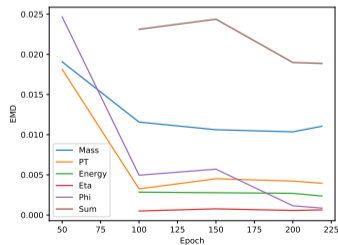
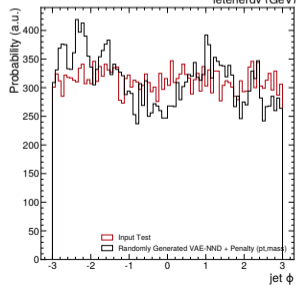
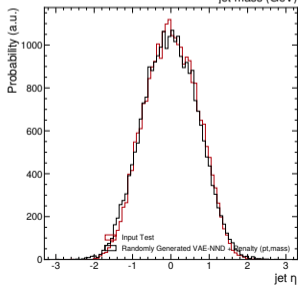
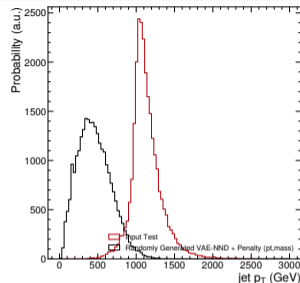
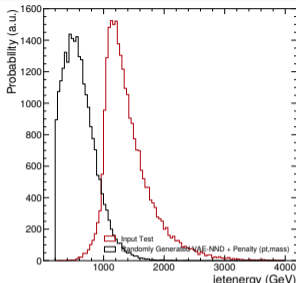
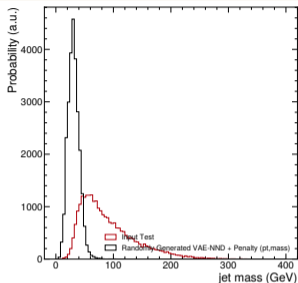
With 1500 epochs: $\beta=5000$; only jet mass in loss function



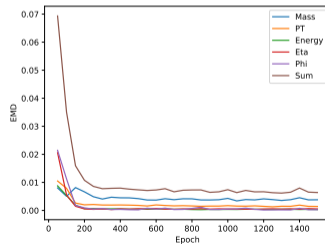
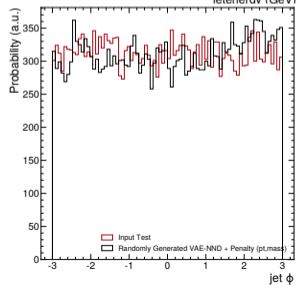
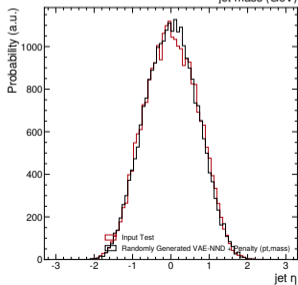
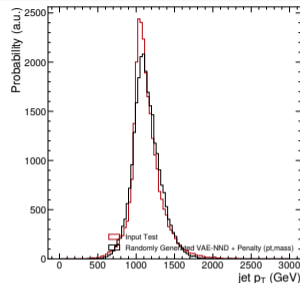
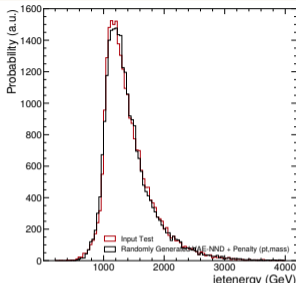
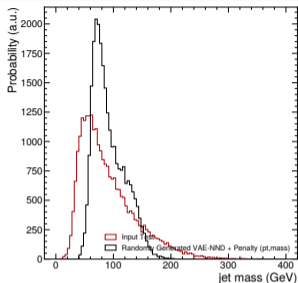
With early stopping: $\beta=0.9998$; $\gamma_{mass}=10.0$



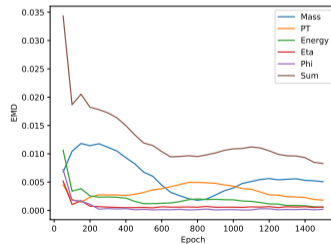
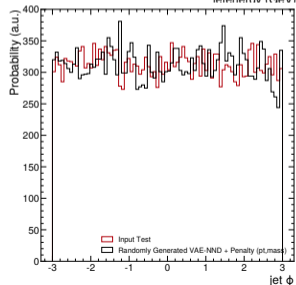
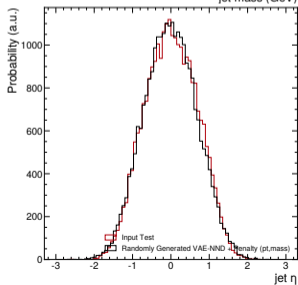
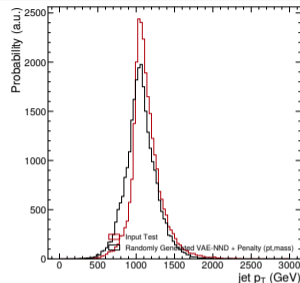
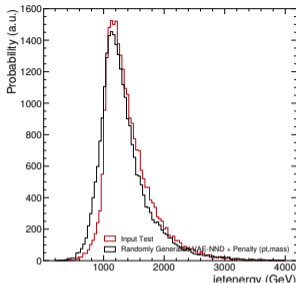
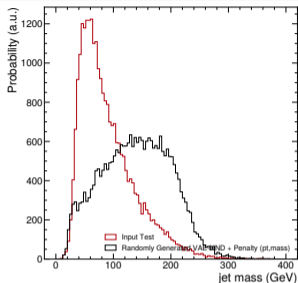
With early stopping: masking as fourth feature



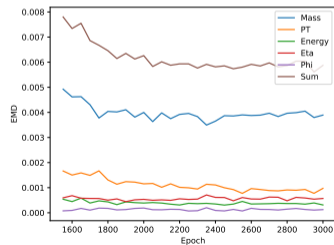
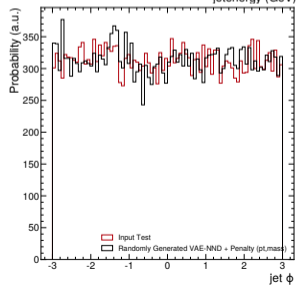
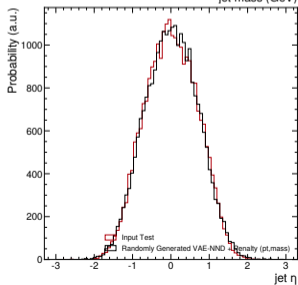
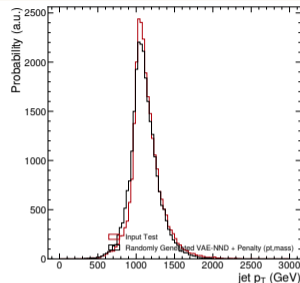
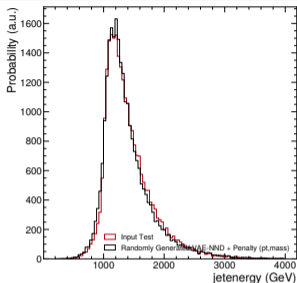
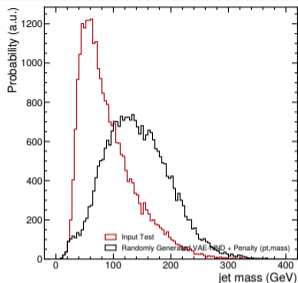
1500 epochs: $\beta=0.9998$; $\gamma_{mass}=10.0$



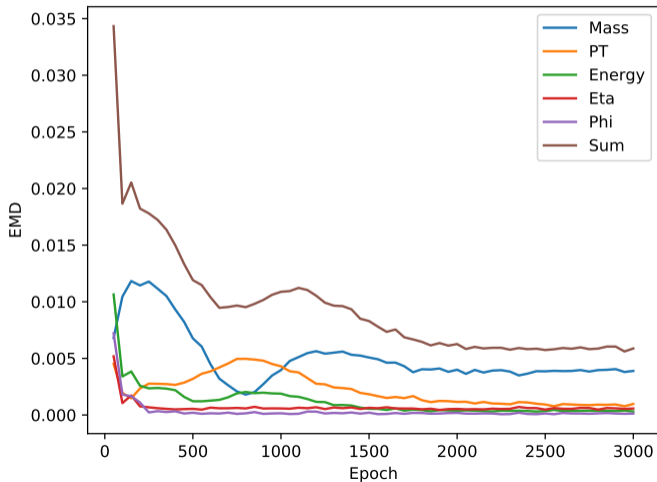
1500 epochs: masking as fourth feature



3000 epochs: masking as fourth feature



3000 epochs: masking as fourth feature (EMD graph from 0 to 3000 epochs)



Next steps

- Check what is wrong with mass in generation
 - Maybe adding other jet features to the loss might help (?)
- Implement masking feature in network
 - The contribution of the "masking feature" to the loss might be much smaller than the rest (add a weight to it (?))
- Distinct techniques to try in the near future
 - Test normalizing flows alone (VAE + flows)
 - With the best VAE model saved, start testing the GAN approach to train the normalizing flows

Backup

$$L = (1 - \beta)L_{\text{reco}} + \beta D_{KL} \quad (1)$$

where

$$L_{\text{reco}} = NND(\text{particle features}) + \gamma_{p_T} * MSE(\text{jet } p_T) + \gamma_{\text{mass}} * MSE(\text{jet mass}) \quad (2)$$

KL divergence per jet feature: 1500 epochs (continue in next slides)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	8.2605e-01	2.0940e-02	7.5556e-03	4.7695e-03	1.0276e-02	8.6959e-01
100	4.5138e-01	9.4429e-03	4.9566e-03	5.0757e-03	2.0010e-03	4.7286e-01
150	3.2272e-01	6.8696e-03	3.2067e-03	4.7887e-03	1.2921e-03	3.3887e-01
200	3.7080e-01	3.9139e-03	3.3406e-03	6.0557e-03	6.5329e-04	3.8476e-01
250	2.3222e-01	6.3299e-03	4.6964e-03	2.1448e-03	3.8649e-04	2.4578e-01
300	2.7877e-01	8.4667e-03	7.5722e-03	3.1152e-03	1.2716e-03	2.9920e-01
350	1.7044e-01	4.7971e-03	3.8813e-03	1.4014e-03	5.7183e-04	1.8109e-01
400	8.4982e-02	6.6242e-03	6.0587e-03	1.6277e-03	3.8991e-04	9.9683e-02
450	7.9693e-02	5.8303e-03	5.5037e-03	7.2171e-04	5.7014e-04	9.2319e-02
500	7.6234e-02	4.1459e-03	5.2709e-03	2.4886e-03	4.4676e-04	8.8586e-02
550	9.2720e-02	5.4468e-03	4.2510e-03	2.8315e-03	9.0427e-04	1.0615e-01
600	6.1575e-02	6.1131e-03	2.8721e-03	1.8787e-03	3.7549e-04	7.2815e-02

KL divergence per jet feature: 1500 epochs (continue in next slide)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
650	1.8751e-01	6.6454e-03	5.1860e-03	2.9980e-03	1.4167e-03	2.0375e-01
700	8.0659e-02	4.4065e-03	3.1377e-03	1.7375e-03	3.7717e-04	9.0318e-02
750	7.3112e-02	5.4230e-03	4.5853e-03	2.0127e-03	7.8947e-04	8.5922e-02
800	9.0053e-02	5.4324e-03	4.5310e-03	1.9827e-03	3.5932e-04	1.0236e-01
850	6.0397e-02	5.3190e-03	3.7719e-03	1.9800e-03	4.8825e-04	7.1957e-02
900	7.8598e-02	4.4622e-03	3.0582e-03	2.6275e-03	6.2821e-04	8.9374e-02
950	7.0326e-02	3.0026e-03	3.6077e-03	3.3163e-03	5.0378e-04	8.0757e-02
1000	5.0762e-02	4.6015e-03	3.9353e-03	1.1393e-03	3.7752e-04	6.0815e-02
1050	7.1664e-02	4.1297e-03	2.4199e-03	1.0617e-03	5.7034e-04	7.9846e-02
1100	7.1252e-02	2.2578e-03	3.1200e-03	1.8093e-03	2.8116e-04	7.8720e-02
1150	7.7812e-02	5.7099e-03	1.9522e-03	1.0304e-03	5.7690e-04	8.7082e-02
1200	8.2688e-02	3.6579e-03	2.4943e-03	1.7522e-03	4.9098e-04	9.1083e-02

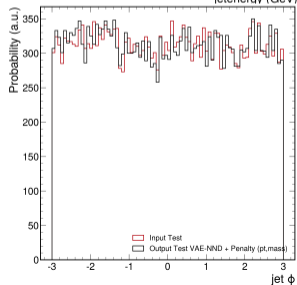
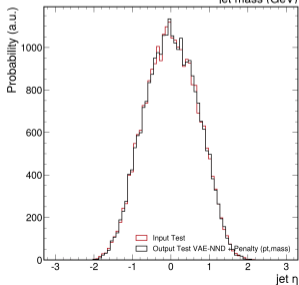
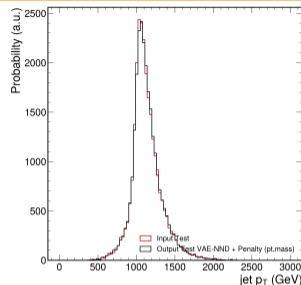
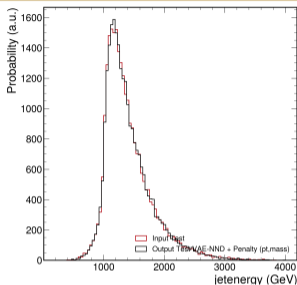
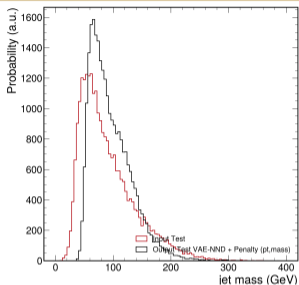
KL divergence per jet feature: 1500 epochs

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1250	9.7535e-02	4.0079e-03	4.2762e-03	1.7523e-03	9.2780e-04	1.0850e-01
1300	5.8758e-02	4.0791e-03	2.3912e-03	1.6562e-03	3.9935e-04	6.7284e-02
1350	5.6689e-02	3.2040e-03	2.9820e-03	1.7787e-03	4.9749e-04	6.5151e-02
1400	5.4620e-02	2.5694e-03	4.1161e-03	1.0080e-03	3.6984e-04	6.2683e-02
1450	4.7014e-02	3.7796e-03	3.0786e-03	1.0542e-03	4.3585e-04	5.5362e-02
1500	5.1686e-02	3.3552e-03	3.0053e-03	1.1447e-03	3.5584e-04	5.9547e-02

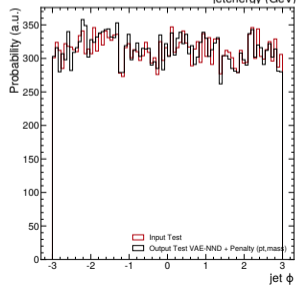
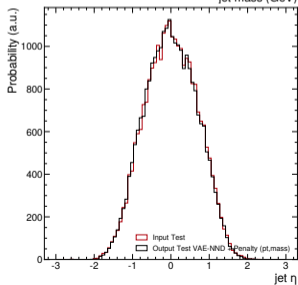
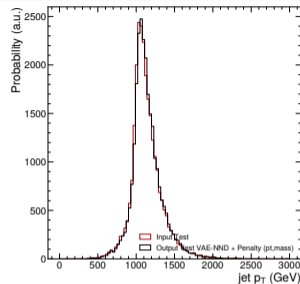
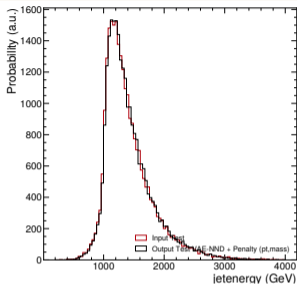
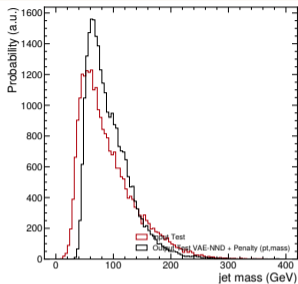
Reconstruction tests

- Performed in the particles features a per feature normalization from 0.0 to 1.0
 - $p_i^N = \frac{p_i - \min(p_i)}{\max(p_i) - \min(p_i)}$
- Test inverting normalization in loss function
 - $p_i = p_i^N \times (\max(p_i) - \min(p_i)) + \min(p_i)$
 - With early stopping:
 - No division; linear activation function
 - No division; sigmoid activation function
 - Per batch, per feature (jet or particle) division by standard deviation; linear activation function
 - Per batch, per feature (jet or particle) division by standard deviation; sigmoid activation function
 - No jet feature in reconstruction loss or only jet mass or jet p_T
 - Mask particles in loss function
 - 1500 epochs:
 - No division; sigmoid activation function
 - Distinct normalization: mean = 0.0 and std = 1.0

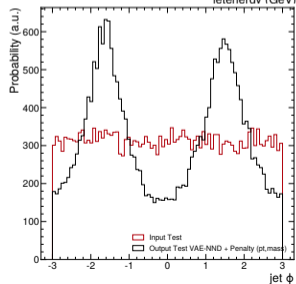
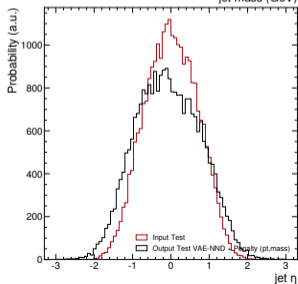
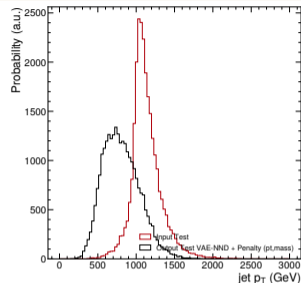
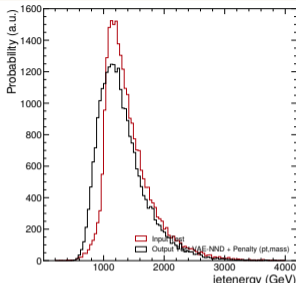
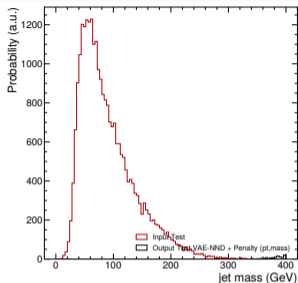
With early stopping: no division; linear activation



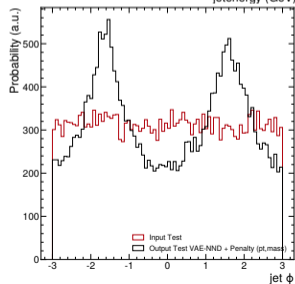
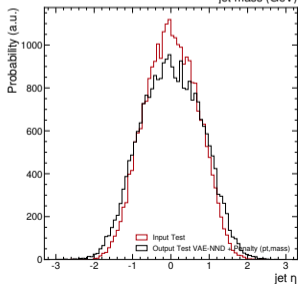
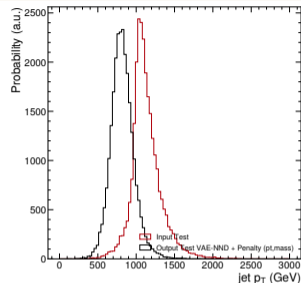
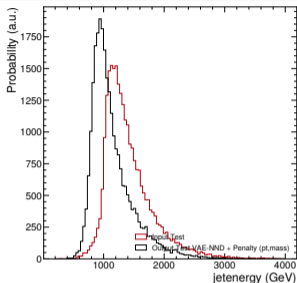
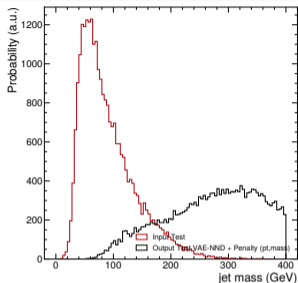
With early stopping: no division; sigmoid activation



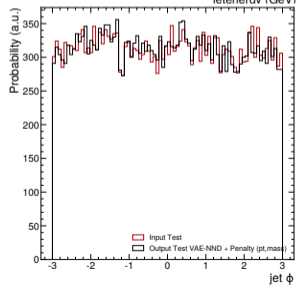
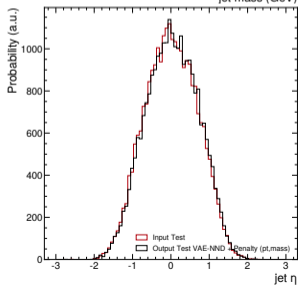
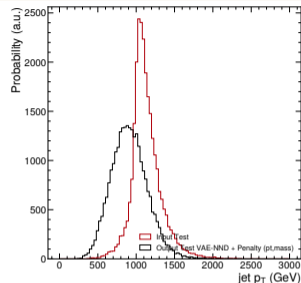
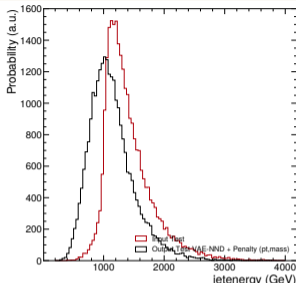
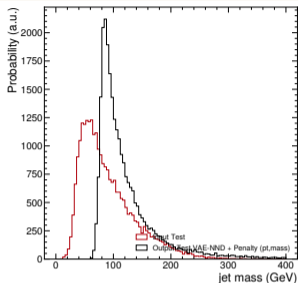
With early stopping: division by standard deviation; linear activation



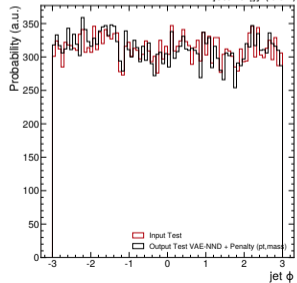
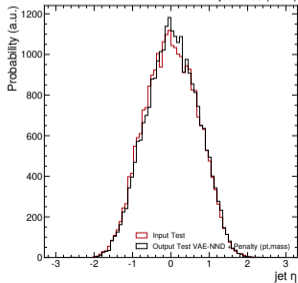
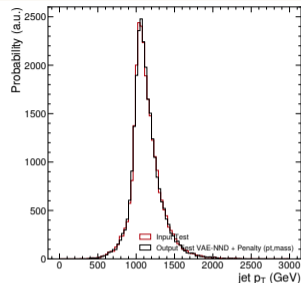
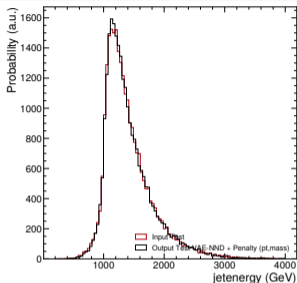
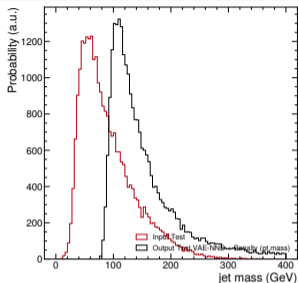
With early stopping: division by standard deviation; sigmoid activation



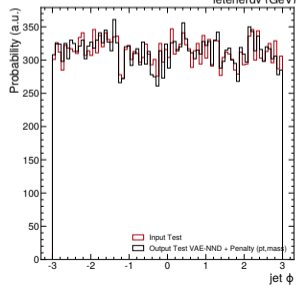
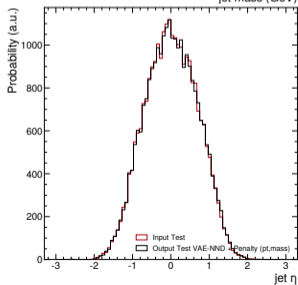
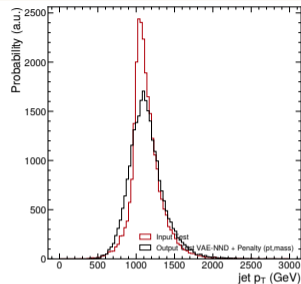
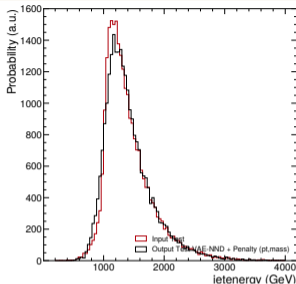
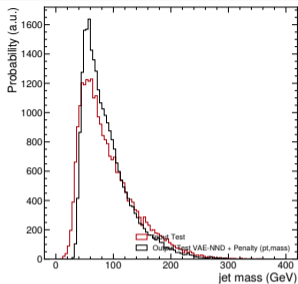
With early stopping: no jet features



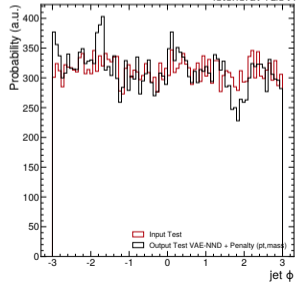
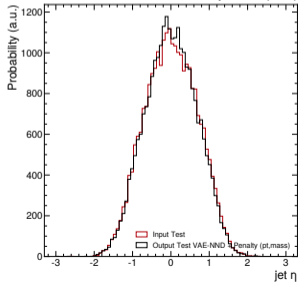
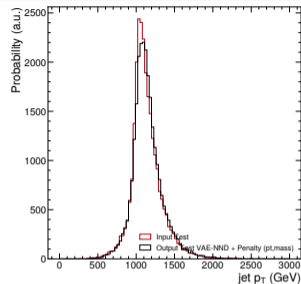
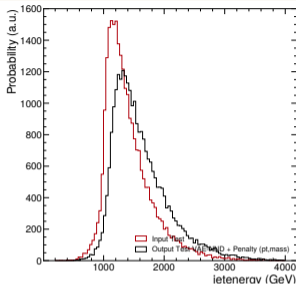
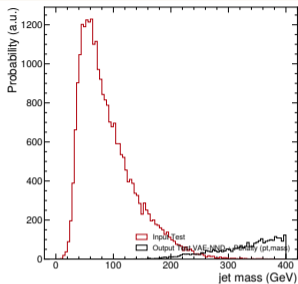
With early stopping: only jet p_T



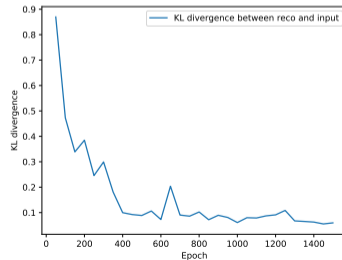
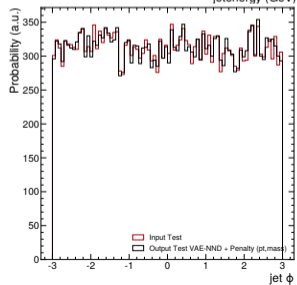
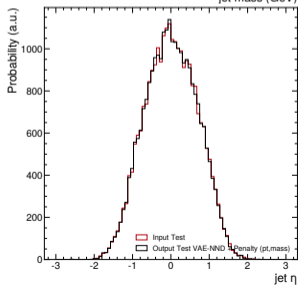
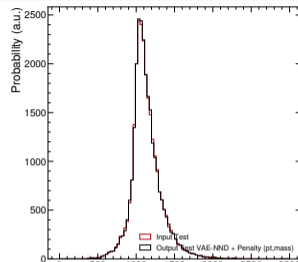
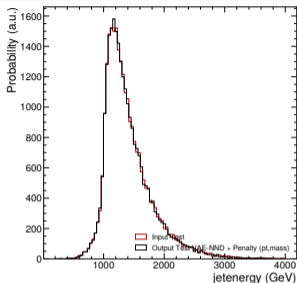
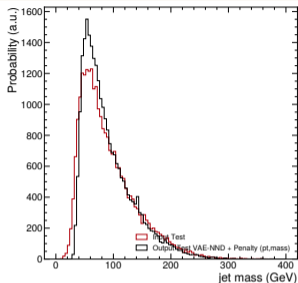
With early stopping: only jet mass



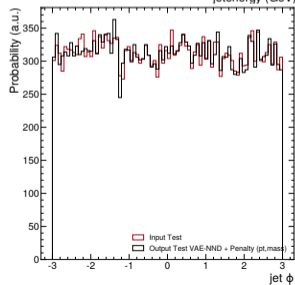
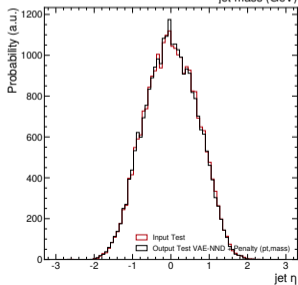
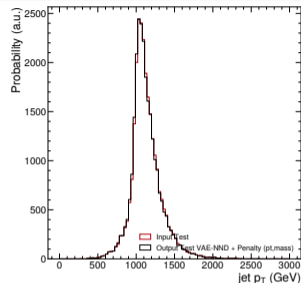
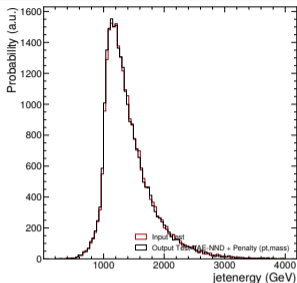
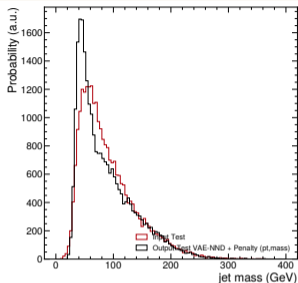
With early stopping: masking particles



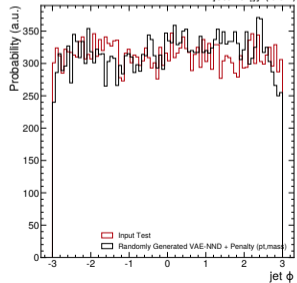
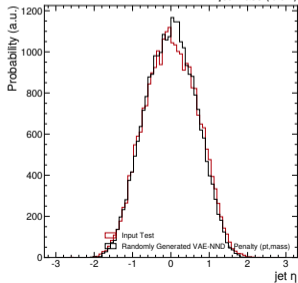
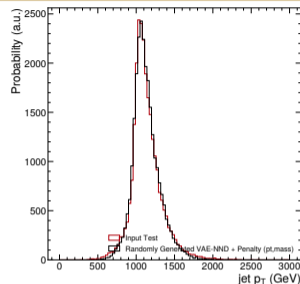
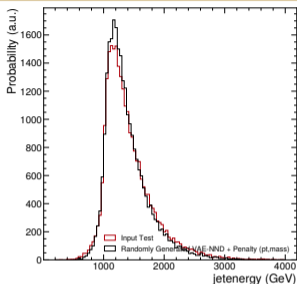
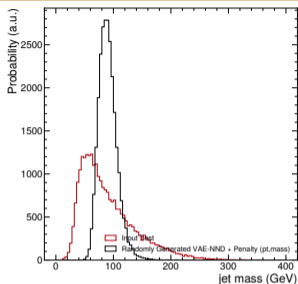
1500 epochs: no division; sigmoid activation



1500 epochs: distinct normalization



1500 epochs: no division; sigmoid activation (generation different beta)



No jet feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	1.7203e+00	8.1025e-01	5.0701e-01	1.2710e-02	2.5291e-03	3.0528e+00
1.0000e+02	1.9530e+00	7.8136e-01	4.6108e-01	2.3272e-03	1.6261e-03	3.1994e+00
1.5000e+02	8.2363e-01	7.6466e-01	4.9012e-01	3.5510e-03	1.0814e-03	2.0830e+00
1.6200e+02	6.9644e-01	8.7703e-01	5.8290e-01	4.7995e-03	1.1405e-03	2.1623e+00

KL divergence tables

Only jet p_T

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	5.6156e+00	8.7501e-03	2.2692e-02	1.3879e-02	1.0836e-01	5.7693e+00
1.0000e+02	2.3496e+00	7.1530e-03	8.2442e-03	4.9062e-03	2.3426e-02	2.3934e+00
1.5000e+02	1.8712e+00	9.7693e-03	6.8163e-03	4.2073e-03	5.8810e-03	1.8979e+00
2.0000e+02	1.3406e+00	5.8739e-03	5.3942e-03	2.4816e-03	2.6825e-03	1.3570e+00
2.5000e+02	1.2863e+00	1.3186e-02	6.4062e-03	3.6992e-03	1.9623e-03	1.3115e+00
2.6900e+02	1.1797e+00	5.8944e-03	5.1474e-03	3.5691e-03	1.4157e-03	1.1957e+00

Only jet mass

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	7.2973e-01	7.5553e-01	2.3685e-01	2.2178e-03	2.2108e-03	1.7265e+00
1.0000e+02	4.7350e-01	5.0531e-01	1.5473e-01	2.9448e-03	9.1262e-04	1.1374e+00
1.5000e+02	9.7626e-02	3.6127e-01	1.0979e-01	3.6480e-03	6.4721e-04	5.7299e-01
2.0000e+02	1.9201e-01	2.1016e-01	6.4102e-02	5.7306e-04	9.6638e-04	4.6781e-01
2.5000e+02	1.1031e-01	8.0245e-02	1.6744e-02	1.1445e-03	6.1319e-04	2.0906e-01
2.8500e+02	7.0034e-02	7.1115e-02	1.7281e-02	1.8337e-03	8.4188e-04	1.6111e-01

Masking particles

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	1.1208e+01	3.9599e-02	1.9996e-01	7.9067e-03	8.5002e-03	1.1464e+01
1.0000e+02	1.1155e+01	2.4435e-02	1.7593e-01	7.9958e-03	4.1014e-03	1.1367e+01
1.5000e+02	1.1072e+01	1.7085e-02	1.5861e-01	7.7367e-03	3.0032e-03	1.1259e+01
2.0000e+02	1.1128e+01	1.5187e-02	1.5099e-01	9.4952e-03	2.9402e-03	1.1307e+01
2.5000e+02	1.1185e+01	1.0042e-02	1.4501e-01	6.5557e-03	3.9695e-03	1.1351e+01

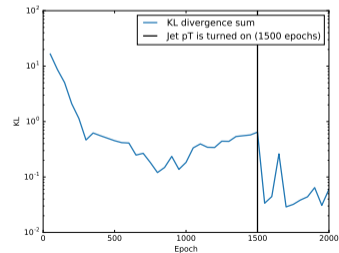
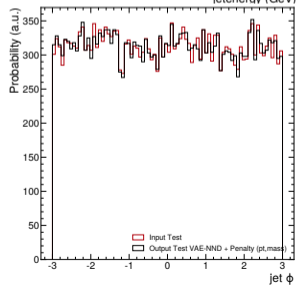
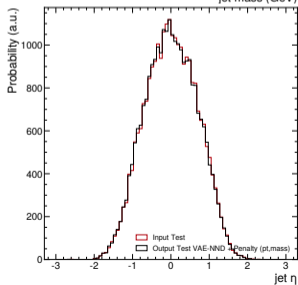
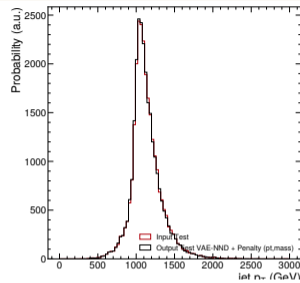
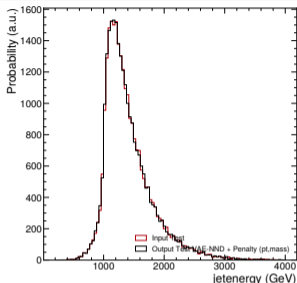
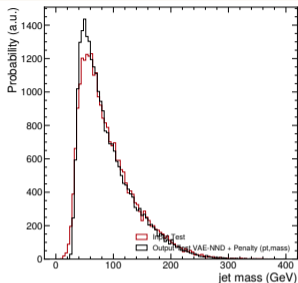
KL divergence tables: only jet mass (1500 epochs)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	8.3429e-01	1.9618e+00	5.5069e-01	1.9748e-03	1.6690e-03	3.3504e+00
1.0000e+02	2.5774e-01	1.0858e+00	3.2862e-01	2.2308e-03	1.3312e-03	1.6757e+00
1.5000e+02	1.6704e-01	6.7980e-01	2.3668e-01	8.9592e-04	8.6976e-04	1.0853e+00
2.0000e+02	1.5310e-01	4.5228e-01	1.6439e-01	2.2594e-03	7.4092e-04	7.7277e-01
2.5000e+02	1.3115e-01	1.8167e-01	6.8209e-02	2.1015e-03	7.0949e-04	3.8385e-01
3.0000e+02	1.0506e-01	8.9207e-02	2.3942e-02	2.0050e-03	6.8123e-04	2.2090e-01
3.5000e+02	1.1757e-01	6.0216e-02	1.0326e-02	1.1650e-03	4.9630e-04	1.8978e-01
4.0000e+02	1.0914e-01	5.7485e-02	1.1507e-02	2.8360e-03	9.9499e-04	1.8196e-01
4.5000e+02	1.0584e-01	7.6392e-02	1.7920e-02	1.5105e-03	4.0530e-04	2.0207e-01
5.0000e+02	7.2738e-02	8.0016e-02	2.4940e-02	6.3612e-04	6.4232e-04	1.7897e-01
5.5000e+02	7.2375e-02	1.0301e-01	3.5000e-02	2.6056e-03	4.8390e-04	2.1347e-01
6.0000e+02	6.0272e-02	1.5636e-01	6.6520e-02	1.8379e-03	6.9209e-04	2.8569e-01
6.5000e+02	1.4106e-01	1.3385e-01	5.6045e-02	2.3172e-03	8.8427e-04	3.3416e-01
7.0000e+02	3.1441e-02	2.5606e-01	1.2172e-01	1.7073e-03	3.7094e-04	4.1130e-01

KL divergence tables: only jet mass (1500 epochs)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
7.5000e+02	4.7349e-02	2.5590e-01	1.2183e-01	3.9613e-04	6.4038e-04	4.2612e-01
8.0000e+02	3.2933e-01	3.4611e-01	1.7438e-01	5.1130e-03	9.0016e-04	8.5584e-01
8.5000e+02	8.4957e-02	2.8877e-01	1.4262e-01	2.3523e-03	8.1481e-04	5.1951e-01
9.0000e+02	2.5542e-02	3.5822e-01	1.8228e-01	2.6595e-04	2.7613e-04	5.6658e-01
9.5000e+02	3.2384e-02	4.4881e-01	2.4212e-01	8.7439e-04	3.1379e-04	7.2450e-01
1.0000e+03	2.3604e-02	4.5254e-01	2.4052e-01	4.0877e-04	4.0708e-04	7.1748e-01
1.0500e+03	2.1989e-02	5.4815e-01	3.0093e-01	9.1641e-04	3.9321e-04	8.7238e-01
1.1000e+03	3.1847e-02	5.8405e-01	3.2053e-01	1.5835e-03	4.6906e-04	9.3848e-01
1.1500e+03	1.8269e-02	5.3131e-01	2.8774e-01	8.6460e-04	3.3941e-04	8.3852e-01
1.2000e+03	3.1049e-02	5.2041e-01	2.8332e-01	1.6177e-03	4.2153e-04	8.3682e-01
1.2500e+03	2.3144e-02	5.8097e-01	3.2106e-01	9.6040e-04	4.3700e-04	9.2658e-01
1.3000e+03	1.4899e-02	5.9190e-01	3.2388e-01	9.0138e-04	3.3964e-04	9.3192e-01
1.3500e+03	1.5377e-02	6.5026e-01	3.6592e-01	8.3121e-04	4.4292e-04	1.0328e+00
1.4000e+03	2.8221e-02	4.4985e-01	2.3394e-01	1.6590e-03	4.8325e-04	7.1415e-01

1500 epochs: turning jet p_T on



KL divergence tables: turning jet p_T on (1500 epochs)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
5.0000e+01	5.8105e-01	9.0074e+00	6.8920e+00	1.7135e-02	9.8977e-03	1.6507e+00
1.0000e+02	1.0443e+00	4.3698e+00	3.2529e+00	1.6233e-02	3.3772e-03	8.6865e+00
1.5000e+02	7.0849e-01	2.4565e+00	1.8805e+00	2.0503e-03	1.6588e-03	5.0492e+00
2.0000e+02	5.4536e-01	9.3513e-01	6.1444e-01	2.2097e-03	9.9026e-04	2.0981e+00
2.5000e+02	6.0268e-01	3.6640e-01	1.6600e-01	4.4356e-03	1.2600e-03	1.1408e+00
3.0000e+02	2.2637e-01	1.8683e-01	4.9413e-02	1.7924e-03	9.7808e-04	4.6538e-01
3.5000e+02	1.6429e-01	3.2655e-01	1.2821e-01	2.1727e-03	7.6609e-04	6.2198e-01
4.0000e+02	1.0693e-01	3.1789e-01	1.2889e-01	2.0168e-03	6.4110e-04	5.5637e-01
4.5000e+02	8.6843e-02	2.8549e-01	1.2528e-01	3.1931e-03	5.0378e-04	5.0131e-01
5.0000e+02	1.3440e-01	2.1992e-01	9.4820e-02	1.7476e-03	6.7136e-04	4.5156e-01
5.5000e+02	1.3127e-01	1.9580e-01	8.7220e-02	1.8269e-03	6.3895e-04	4.1676e-01
6.0000e+02	1.3436e-01	1.8852e-01	8.6361e-02	3.3759e-04	5.0546e-04	4.1008e-01
6.5000e+02	7.3624e-02	1.2029e-01	5.1202e-02	2.6113e-03	6.0248e-04	2.4833e-01
7.0000e+02	9.2705e-02	1.2090e-01	5.2413e-02	9.1404e-04	5.8713e-04	2.6752e-01

KL divergence tables: turning jet p_T on (1500 epochs)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
7.5000e+02	8.7852e-02	6.9756e-02	2.3157e-02	1.1279e-03	8.2440e-04	1.8272e-01
8.0000e+02	3.9082e-02	6.1365e-02	1.8175e-02	1.0056e-03	5.7011e-04	1.2020e-01
8.5000e+02	7.7526e-02	5.6449e-02	1.2810e-02	1.2903e-03	5.5646e-04	1.4863e-01
9.0000e+02	1.4934e-01	6.4810e-02	1.8602e-02	1.7107e-03	5.9781e-04	2.3506e-01
9.5000e+02	4.1414e-02	7.3519e-02	2.1200e-02	4.5477e-04	7.5774e-04	1.3735e-01
1.0000e+03	5.2687e-02	9.6439e-02	3.3302e-02	4.4463e-04	5.5535e-04	1.8343e-01
1.0500e+03	1.3525e-01	1.4037e-01	5.7723e-02	1.0046e-03	5.0116e-04	3.3485e-01
1.1000e+03	9.2158e-02	2.0662e-01	9.3224e-02	2.4395e-03	7.5486e-04	3.9519e-01
1.1500e+03	4.3587e-02	2.0395e-01	9.4498e-02	1.1500e-03	4.4625e-04	3.4363e-01
1.2000e+03	3.0271e-02	2.0870e-01	1.0040e-01	9.4153e-04	4.4434e-04	3.4075e-01
1.2500e+03	3.3285e-02	2.7219e-01	1.3456e-01	1.5952e-03	4.7638e-04	4.4211e-01
1.3000e+03	4.7517e-02	2.6081e-01	1.2906e-01	1.5885e-03	4.2980e-04	4.3941e-01
1.3500e+03	4.3789e-02	3.2462e-01	1.6628e-01	2.3311e-03	5.6190e-04	5.3758e-01
1.4000e+03	3.8437e-02	3.4000e-01	1.7628e-01	6.5268e-04	3.8760e-04	5.5576e-01

KL divergence tables: turning jet p_T on (1500 epochs)

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1.4500e+03	2.7388e-02	3.5773e-01	1.8652e-01	1.6204e-03	2.7165e-04	5.7354e-01
1.5000e+03	3.6458e-02	3.9415e-01	2.1079e-01	1.8175e-03	3.2926e-04	6.4355e-01
1.5500e+03	2.3336e-02	5.4626e-03	3.1835e-03	1.0003e-03	7.0038e-04	3.3683e-02
1.6000e+03	3.5534e-02	4.4079e-03	2.9660e-03	1.0540e-03	3.3467e-04	4.4297e-02
1.6500e+03	2.3763e-01	7.1117e-03	7.7155e-03	7.2367e-03	1.6315e-03	2.6132e-01
1.7000e+03	2.0747e-02	3.5943e-03	3.1796e-03	9.9660e-04	3.6448e-04	2.8882e-02
1.7500e+03	2.4102e-02	4.2656e-03	2.5554e-03	8.3674e-04	2.9781e-04	3.2058e-02
1.8000e+03	2.8158e-02	5.7598e-03	3.0466e-03	9.2619e-04	4.8386e-04	3.8375e-02
1.8500e+03	3.3551e-02	4.9411e-03	4.0169e-03	9.7525e-04	4.1273e-04	4.3897e-02
1.9000e+03	5.4786e-02	3.9689e-03	3.3109e-03	1.0712e-03	1.1253e-03	6.4263e-02
1.9500e+03	1.9970e-02	3.7597e-03	5.8563e-03	8.5153e-04	3.5092e-04	3.0788e-02
2.0000e+03	5.0546e-02	3.8820e-03	3.7668e-03	1.3172e-03	4.8323e-04	5.9996e-02

Generation tests

Early stopping; $\beta=10000$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0099	0.0123	0.0116	0.0119	0.0038	0.0495
100	0.0107	0.0070	0.0026	0.0008	0.0011	0.0221
150	0.0101	0.0013	0.0012	0.0009	0.0010	0.0144
200	0.0084	0.0009	0.0007	0.0005	0.0008	0.0113
250	0.0085	0.0013	0.0006	0.0006	0.0011	0.0121
281	0.0081	0.0012	0.0007	0.0007	0.0008	0.0114

Early stopping; $\beta=5000$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0054	0.0043	0.0029	0.0010	0.0015	0.0150
100	0.0076	0.0035	0.0015	0.0008	0.0007	0.0142
150	0.0076	0.0031	0.0012	0.0006	0.0004	0.0129
200	0.0071	0.0029	0.0010	0.0007	0.0004	0.0121
223	0.0072	0.0027	0.0010	0.0007	0.0004	0.0119

1500 epochs; $\beta=5000$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0072	0.0051	0.0032	0.0014	0.0019	0.0188
100	0.0100	0.0031	0.0015	0.0010	0.0008	0.0163
150	0.0080	0.0027	0.0009	0.0006	0.0004	0.0126
200	0.0075	0.0023	0.0007	0.0006	0.0003	0.0114
250	0.0076	0.0023	0.0006	0.0006	0.0001	0.0112
300	0.0077	0.0023	0.0006	0.0006	0.0002	0.0114
350	0.0074	0.0021	0.0007	0.0005	0.0001	0.0109
400	0.0074	0.0021	0.0005	0.0006	0.0001	0.0107
450	0.0076	0.0019	0.0004	0.0006	0.0001	0.0106
500	0.0072	0.0018	0.0004	0.0005	0.0001	0.0100
550	0.0071	0.0018	0.0004	0.0007	0.0002	0.0101
600	0.0069	0.0018	0.0004	0.0005	0.0001	0.0098

1500 epochs; $\beta=5000$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
650	0.0074	0.0015	0.0004	0.0006	0.0001	0.0101
700	0.0069	0.0017	0.0004	0.0006	0.0001	0.0097
750	0.0070	0.0017	0.0004	0.0006	0.0001	0.0098
800	0.0068	0.0015	0.0005	0.0005	0.0001	0.0094
850	0.0073	0.0016	0.0004	0.0005	0.0001	0.0100
900	0.0073	0.0014	0.0003	0.0006	0.0001	0.0097
950	0.0078	0.0014	0.0004	0.0005	0.0001	0.0103
1000	0.0067	0.0014	0.0004	0.0005	0.0001	0.0091
1050	0.0069	0.0013	0.0003	0.0006	0.0001	0.0092
1100	0.0073	0.0013	0.0004	0.0005	0.0001	0.0096
1150	0.0070	0.0015	0.0004	0.0006	0.0001	0.0096
1200	0.0073	0.0012	0.0003	0.0006	0.0001	0.0097

1500 epochs; $\beta=5000$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1250	0.0065	0.0014	0.0004	0.0006	0.0001	0.0090
1300	0.0072	0.0013	0.0004	0.0006	0.0001	0.0095
1350	0.0067	0.0013	0.0004	0.0005	0.0001	0.0090
1400	0.0070	0.0013	0.0004	0.0006	0.0001	0.0094
1450	0.0068	0.0012	0.0004	0.0005	0.0001	0.0090
1500	0.0072	0.0012	0.0004	0.0006	0.0001	0.0095

1500 epochs; $\beta=5000$; only jet mass in loss function

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0089	0.0081	0.0039	0.0014	0.0001	0.0224
100	0.0083	0.0063	0.0028	0.0012	0.0001	0.0186
150	0.0096	0.0057	0.0026	0.0011	0.0001	0.0191
200	0.0089	0.0055	0.0022	0.0010	0.0001	0.0178
250	0.0089	0.0054	0.0024	0.0009	0.0003	0.0178
300	0.0081	0.0051	0.0021	0.0008	0.0001	0.0162
350	0.0087	0.0052	0.0022	0.0008	0.0003	0.0173
400	0.0084	0.0051	0.0022	0.0007	0.0001	0.0166
450	0.0095	0.0049	0.0021	0.0009	0.0004	0.0179
500	0.0102	0.0051	0.0023	0.0009	0.0004	0.0189
550	0.0088	0.0052	0.0022	0.0008	0.0003	0.0174
600	0.0093	0.0051	0.0021	0.0009	0.0003	0.0176

1500 epochs; $\beta=5000$; only jet mass in loss function

Epoch	Mass	p_T	Energy	η	ϕ	Sum
650	0.0090	0.0050	0.0021	0.0006	0.0003	0.0171
700	0.0096	0.0050	0.0021	0.0009	0.0004	0.0179
750	0.0079	0.0049	0.0021	0.0009	0.0003	0.0161
800	0.0091	0.0048	0.0020	0.0007	0.0004	0.0170
850	0.0087	0.0049	0.0020	0.0008	0.0004	0.0167
900	0.0088	0.0049	0.0020	0.0009	0.0003	0.0168
950	0.0097	0.0047	0.0019	0.0009	0.0002	0.0175
1000	0.0090	0.0047	0.0019	0.0007	0.0004	0.0168
1050	0.0093	0.0047	0.0019	0.0007	0.0004	0.0169
1100	0.0095	0.0047	0.0019	0.0007	0.0002	0.0170
1150	0.0095	0.0046	0.0019	0.0008	0.0003	0.0171
1200	0.0093	0.0048	0.0019	0.0012	0.0005	0.0177

1500 epochs; $\beta=5000$; only jet mass in loss function

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1250	0.0093	0.0047	0.0020	0.0006	0.0003	0.0169
1300	0.0097	0.0047	0.0019	0.0008	0.0003	0.0174
1350	0.0096	0.0047	0.0018	0.0008	0.0005	0.0174
1400	0.0098	0.0046	0.0018	0.0007	0.0005	0.0174
1450	0.0089	0.0045	0.0019	0.0009	0.0003	0.0165
1500	0.0088	0.0044	0.0018	0.0007	0.0004	0.0162

Early stopping; $\beta=0.9998$; $\gamma_{mass}=10.0$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0133	0.0052	0.0037	0.0085	0.0151	0.0458
100	0.0050	0.0024	0.0014	0.0079	0.0119	0.0286
150	0.0027	0.0021	0.0012	0.0018	0.0097	0.0175
200	0.0049	0.0027	0.0009	0.0009	0.0014	0.0108
250	0.0047	0.0020	0.0006	0.0006	0.0005	0.0085
251	0.0045	0.0028	0.0008	0.0008	0.0014	0.0102

1500 epochs; $\beta=0.9998$; $\gamma_{mass}=10.0$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0080	0.0105	0.0088	0.0206	0.0214	0.0693
100	0.0051	0.0079	0.0053	0.0055	0.0113	0.0351
150	0.0082	0.0026	0.0019	0.0015	0.0018	0.0159
200	0.0066	0.0020	0.0006	0.0006	0.0009	0.0107
250	0.0049	0.0021	0.0007	0.0004	0.0005	0.0086
300	0.0041	0.0019	0.0005	0.0005	0.0007	0.0078
350	0.0047	0.0019	0.0006	0.0003	0.0003	0.0079
400	0.0045	0.0020	0.0006	0.0006	0.0004	0.0080
450	0.0045	0.0019	0.0005	0.0005	0.0003	0.0076
500	0.0042	0.0018	0.0005	0.0006	0.0002	0.0073
550	0.0037	0.0016	0.0004	0.0006	0.0008	0.0071
600	0.0037	0.0020	0.0005	0.0007	0.0005	0.0073

1500 epochs; $\beta=0.9998$; $\gamma_{mass}=10.0$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
650	0.0042	0.0018	0.0005	0.0005	0.0008	0.0078
700	0.0038	0.0016	0.0004	0.0004	0.0003	0.0067
750	0.0042	0.0017	0.0004	0.0005	0.0004	0.0072
800	0.0041	0.0016	0.0003	0.0005	0.0007	0.0073
850	0.0037	0.0015	0.0003	0.0006	0.0012	0.0073
900	0.0037	0.0016	0.0004	0.0005	0.0003	0.0065
950	0.0038	0.0016	0.0003	0.0006	0.0003	0.0066
1000	0.0043	0.0017	0.0004	0.0005	0.0005	0.0074
1050	0.0034	0.0016	0.0004	0.0005	0.0006	0.0064
1100	0.0039	0.0015	0.0004	0.0005	0.0008	0.0071
1150	0.0038	0.0016	0.0004	0.0004	0.0004	0.0066
1200	0.0041	0.0015	0.0004	0.0005	0.0002	0.0067

1500 epochs; $\beta=0.9998$; $\gamma_{mass}=10.0$

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1250	0.0039	0.0013	0.0004	0.0006	0.0002	0.0063
1300	0.0036	0.0015	0.0004	0.0005	0.0002	0.0062
1350	0.0038	0.0015	0.0004	0.0006	0.0002	0.0065
1400	0.0045	0.0019	0.0003	0.0005	0.0008	0.0080
1450	0.0038	0.0015	0.0004	0.0006	0.0003	0.0065
1500	0.0038	0.0014	0.0003	0.0006	0.0003	0.0064

Early stopping; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0191	0.0181	nan	nan	0.0247	nan
100	0.0116	0.0033	0.0028	0.0005	0.0050	0.0231
150	0.0106	0.0045	0.0028	0.0008	0.0057	0.0244
200	0.0104	0.0042	0.0027	0.0006	0.0012	0.0190
219	0.0110	0.0040	0.0024	0.0006	0.0009	0.0189

3000 epochs; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
50	0.0068	0.0045	0.0106	0.0052	0.0072	0.0343
100	0.0105	0.0020	0.0034	0.0011	0.0018	0.0187
150	0.0118	0.0015	0.0038	0.0017	0.0017	0.0205
200	0.0114	0.0023	0.0026	0.0008	0.0011	0.0182
250	0.0118	0.0028	0.0024	0.0007	0.0002	0.0178
300	0.0112	0.0028	0.0024	0.0006	0.0003	0.0172
350	0.0105	0.0027	0.0023	0.0005	0.0003	0.0164
400	0.0093	0.0027	0.0022	0.0005	0.0003	0.0150
450	0.0082	0.0029	0.0016	0.0005	0.0001	0.0133
500	0.0068	0.0032	0.0012	0.0005	0.0002	0.0119
550	0.0060	0.0036	0.0012	0.0005	0.0001	0.0115
600	0.0045	0.0039	0.0013	0.0006	0.0002	0.0104

3000 epochs; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
650	0.0032	0.0042	0.0013	0.0006	0.0001	0.0095
700	0.0026	0.0045	0.0016	0.0006	0.0002	0.0095
750	0.0021	0.0050	0.0019	0.0006	0.0001	0.0097
800	0.0018	0.0050	0.0020	0.0006	0.0001	0.0095
850	0.0020	0.0049	0.0019	0.0007	0.0003	0.0098
900	0.0027	0.0048	0.0020	0.0006	0.0001	0.0102
950	0.0035	0.0045	0.0019	0.0006	0.0001	0.0106
1000	0.0040	0.0043	0.0019	0.0006	0.0002	0.0109
1050	0.0048	0.0039	0.0017	0.0006	0.0001	0.0109
1100	0.0051	0.0038	0.0016	0.0006	0.0001	0.0112
1150	0.0055	0.0033	0.0014	0.0006	0.0003	0.0110
1200	0.0056	0.0028	0.0012	0.0007	0.0003	0.0105

3000 epochs; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1250	0.0054	0.0027	0.0011	0.0005	0.0001	0.0099
1300	0.0055	0.0024	0.0009	0.0007	0.0002	0.0097
1350	0.0056	0.0024	0.0009	0.0006	0.0002	0.0096
1400	0.0053	0.0023	0.0008	0.0007	0.0002	0.0093
1450	0.0052	0.0020	0.0007	0.0005	0.0001	0.0085
1500	0.0051	0.0018	0.0006	0.0006	0.0002	0.0083
1550	0.0049	0.0017	0.0005	0.0006	0.0001	0.0078
1600	0.0046	0.0015	0.0004	0.0007	0.0001	0.0073
1650	0.0046	0.0016	0.0006	0.0006	0.0002	0.0076
1700	0.0043	0.0015	0.0004	0.0006	0.0001	0.0069
1750	0.0038	0.0017	0.0005	0.0006	0.0002	0.0067
1800	0.0040	0.0013	0.0004	0.0005	0.0002	0.0065

3000 epochs; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
1850	0.0040	0.0011	0.0003	0.0006	0.0001	0.0061
1900	0.0041	0.0012	0.0004	0.0004	0.0001	0.0063
1950	0.0038	0.0012	0.0004	0.0005	0.0002	0.0061
2000	0.0040	0.0012	0.0004	0.0005	0.0002	0.0063
2050	0.0036	0.0012	0.0004	0.0005	0.0001	0.0058
2100	0.0040	0.0010	0.0004	0.0005	0.0001	0.0060
2150	0.0037	0.0012	0.0003	0.0005	0.0001	0.0059
2200	0.0039	0.0010	0.0003	0.0006	0.0001	0.0059
2250	0.0040	0.0010	0.0004	0.0005	0.0001	0.0059
2300	0.0038	0.0009	0.0004	0.0005	0.0001	0.0058
2350	0.0035	0.0011	0.0004	0.0007	0.0002	0.0059
2400	0.0037	0.0011	0.0003	0.0006	0.0001	0.0058

3000 epochs; masking as fourth feature

Epoch	Mass	p_T	Energy	η	ϕ	Sum
2450	0.0039	0.0010	0.0003	0.0006	0.0001	0.0059
2500	0.0039	0.0009	0.0003	0.0005	0.0001	0.0057
2550	0.0039	0.0008	0.0005	0.0006	0.0001	0.0058
2600	0.0039	0.0010	0.0003	0.0005	0.0002	0.0059
2650	0.0039	0.0009	0.0004	0.0005	0.0001	0.0058
2700	0.0040	0.0009	0.0004	0.0006	0.0001	0.0060
2750	0.0038	0.0009	0.0004	0.0006	0.0001	0.0058
2800	0.0040	0.0009	0.0004	0.0005	0.0001	0.0059
2850	0.0040	0.0009	0.0004	0.0006	0.0002	0.0060
2900	0.0040	0.0009	0.0003	0.0006	0.0001	0.0060
2950	0.0038	0.0008	0.0004	0.0005	0.0001	0.0056
3000	0.0039	0.0010	0.0003	0.0006	0.0001	0.0059