Glow

December 26, 2022

[1]:
```
%cd ../../../
```

```
/home/jan/FMF/masters
```

[2]:
```
saved = "ml_hep_sim/notebooks/article_notebooks/saved/"
```

[3]:
```
from ml_hep_sim.notebooks.article_notebooks.test_runs import *
from ml_hep_sim.pipeline.pipes import *
from ml_hep_sim.pipeline.blocks import *

from ml_hep_sim.plotting.style import style_setup
from ml_hep_sim.stats.stat_plots import two_sample_plot

from ml_hep_sim.data_utils.higgs.process_higgs_dataset import LATEX_COLNAMES,␣
↪LOG_BIN_RANGES

from scipy.optimize import curve_fit
import matplotlib.pyplot as plt
from tqdm import tqdm
import copy

style_setup(seaborn_pallete=True)
```

[4]:
```
num_flows = np.arange(4, 32, 2)
```

[5]:
```
pipelines = run_glow_pipeline(train=False, gen=False, test=False)
```

```
  0%|
| 0/14 [00:00<?, ?it/s]WARNING:root:Number of composed and loaded pipes did not
match! Loading anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
 43%|

| 6/14 [00:00<00:00, 57.20it/s]WARNING:root:Number of composed and loaded pipes
did not match! Loading anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
 86%|


| 12/14 [00:00<00:00, 58.43it/s]WARNING:root:Number of composed and loaded pipes
did not match! Loading anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
100%|


           | 14/14 [00:00<00:00, 45.06it/s]
```

[6]:
```
results = []

N = 20 # batch size (i.e. 1024 * N data points)

for pipe in tqdm(pipelines):
    x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipe[0].pipes

    x1 = ModelLoaderBlock()(x_ConfigBuilderBlock, x_ModelTrainerBlock)
    x2 = DatasetBuilderBlock()(x_ConfigBuilderBlock)
    x3 = CouplingModelTestingBlock(N, loss_cutoff=20)(x2, x1,␣
 ↪x_ConfigBuilderBlock)

    test_pipe = Pipeline()
    test_pipe.compose(x1, x2, x3)
    test_pipe.fit()
    results.append(test_pipe)
```

```
  0%|
| 0/14 [00:00<?, ?it/s]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7cd4864ca0>!
```

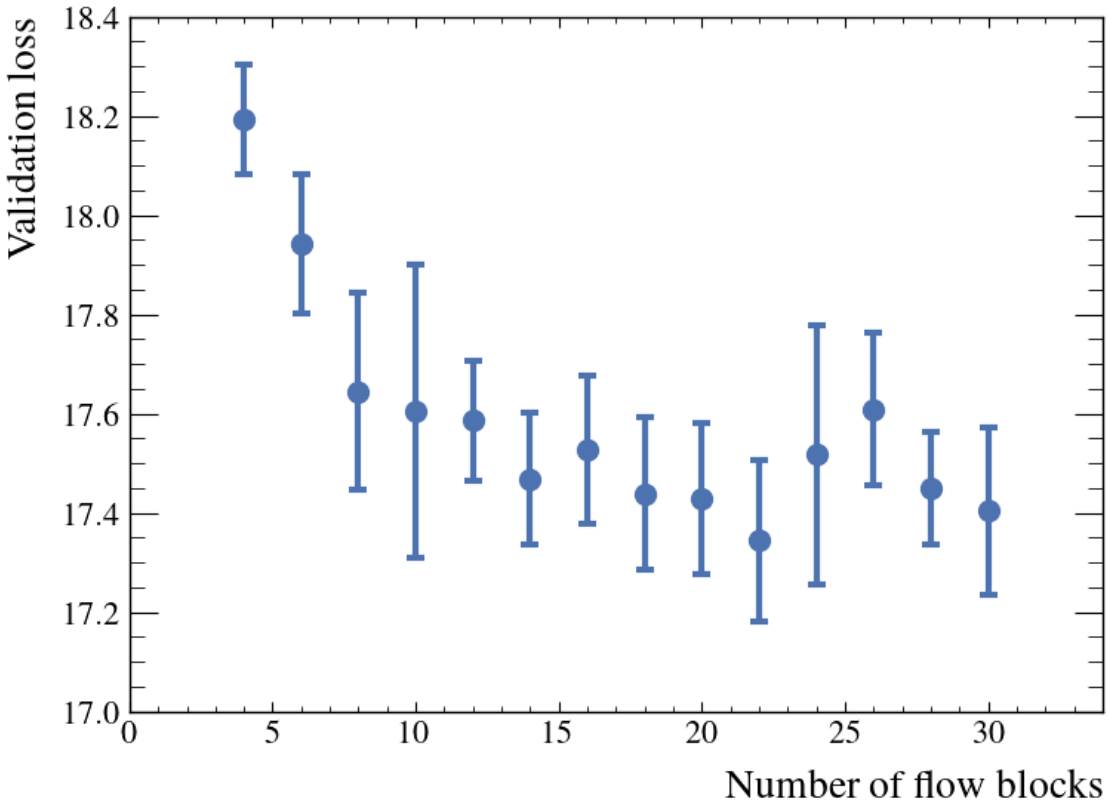
```
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cd4864c10>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7cd4864bb0>!
  7%|
| 1/14 [00:02<00:34,  2.69s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e586c6760>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e586bf220>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e58649c10>!
 14%|
| 2/14 [00:05<00:32,  2.68s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4ffaf460>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4ffaf040>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4ffafa30>!
 21%|
| 3/14 [00:08<00:29,  2.71s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4ffa1b50>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4feae370>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4feae040>!
 29%|
| 4/14 [00:11<00:28,  2.84s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4fd58490>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4fd584f0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4fd58460>!
 36%|

| 5/14 [00:14<00:26,  2.92s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4fdbd4c0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4fbcdb50>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4fbcdd30>!
 43%|

| 6/14 [00:17<00:24,  3.02s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4fa59fa0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4fa598e0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4fa59f40>!
```

```
 50%|

| 7/14 [00:20<00:21,  3.12s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4f1f5f70>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4f1f5e20>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4f1f5f10>!
 57%|

| 8/14 [00:24<00:19,  3.18s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e4fa45430>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e397eb0d0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4fc20fa0>!
 64%|


| 9/14 [00:27<00:16,  3.20s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e3957af10>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e3957ad60>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e3957aeb0>!
 71%|


| 10/14 [00:30<00:13,  3.28s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e3957af70>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4fc0f370>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4fc0f190>!
 79%|


| 11/14 [00:34<00:10,  3.40s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e39408f40>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e39408fa0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e39408f70>!
 86%|


| 12/14 [00:38<00:06,  3.47s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e39408ca0>!
```

```
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e39781e50>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e39781e20>!
 93%|


                    | 13/14 [00:41<00:03,  3.50s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f7e586ddfd0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e4fbc32e0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e4ff444c0>!
100%|


           | 14/14 [00:45<00:00,  3.24s/it]
```

[7]:
```
m_lst, s_lst = [], [] # mean and std

for r in results:
    m, s, _ = r.pipes[-1].results
    m_lst.append(m)
    s_lst.append(s)
```

[8]:
```
plt.scatter(num_flows, m_lst, s=100)
plt.errorbar(num_flows, m_lst, yerr=s_lst, capsize=4, ls="none", lw=3,␣
 ↪capthick=3)
plt.xlim([0, 34])
plt.xlabel("Number of flow blocks")
plt.ylabel("Validation loss")
plt.savefig(saved + "loss_vs_num_flows_glow.pdf")
plt.tight_layout()
```

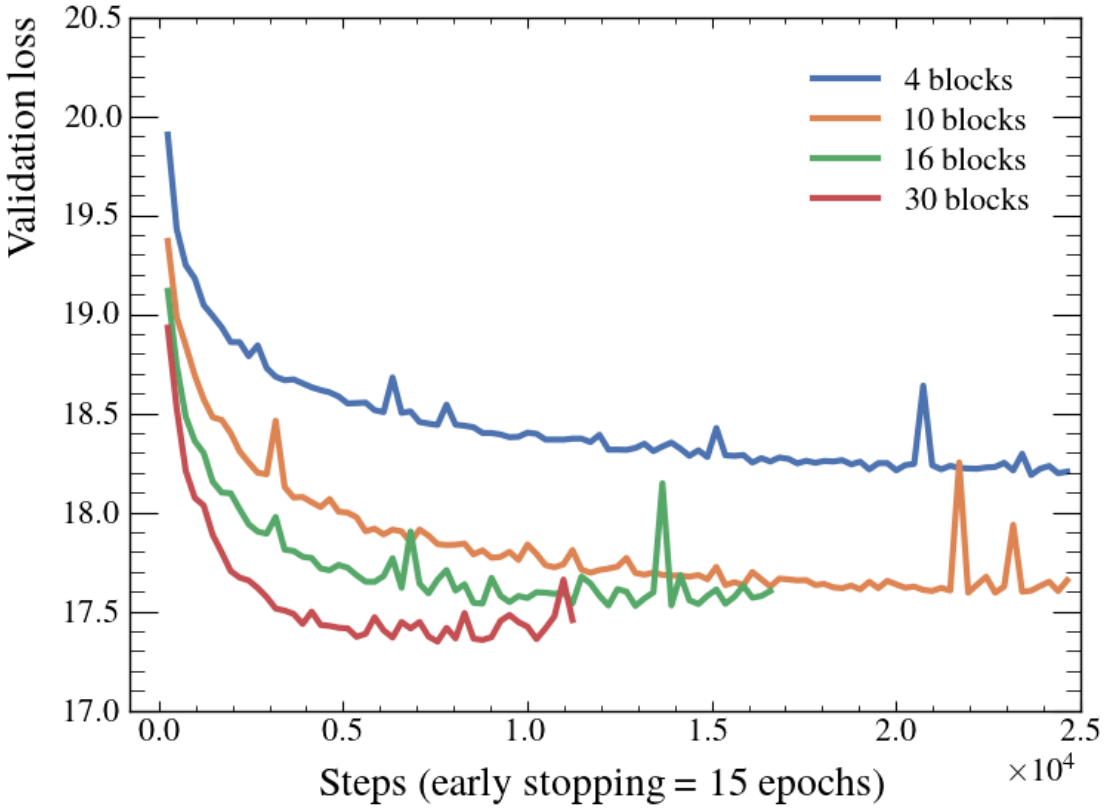


[9]:
```
t = pipe[1].pipes[0].metrics[0]["timestamp"]
```

[10]:
```
val_losses, steps, times = [], [], []

for pipe in pipelines:
    val_loss = pipe[1].pipes[0].metrics[-1]["val_loss"]
    step = pipe[1].pipes[0].metrics[-2]["step"]
    t = pipe[1].pipes[0].metrics[0]["timestamp"].to_numpy()
    times.append(t[-1] - t[0])
    val_losses.append(val_loss)
    steps.append(step)
```

[11]:
```
plt.plot(steps[0], val_losses[0], lw=3)
plt.plot(steps[3], val_losses[3], lw=3)
plt.plot(steps[6], val_losses[6], lw=3)
plt.plot(steps[-1], val_losses[-1], lw=3)
plt.legend(["4 blocks", "10 blocks", "16 blocks", "30 blocks"])
plt.xlim([-800, 2.5*10**4])
plt.xlabel("Steps (early stopping = 15 epochs)", loc="center")
plt.ylabel("Validation loss")
plt.savefig(saved + "val_loss_vs_steps_glow.pdf")
```

```
plt.tight_layout()
```



[13]:
```
device = "cuda"

r = 15 # repeats
s = 30 # scaling plot points
N = 10**5

res_lst = []
class_res = []

for _ in range(r):
    x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipelines[3][0].pipes
    x_ConfigBuilderBlock.config["datasets"]["data_params"]["subset_n"] =␣
 ↪[250000, 100000, N]

    x1 = ModelLoaderBlock(device=device)(x_ConfigBuilderBlock,␣
 ↪x_ModelTrainerBlock)

    x2 = DataGeneratorBlock(N, model_type="flow", chunks=10, device=device)(x1)
```

```
   x3 = GeneratedDataVerifierBlock(save_data=False, device=device,␣
↪rescale_data=False)(x1, x2)

   x4 = DatasetBuilderBlock()(x_ConfigBuilderBlock)
   x5 = ReferenceDataLoaderBlock(rescale_reference="logit_normal",␣
↪device=device)(x4)

   class_run_name = "Higgs_resnet_classifier_train_pipeline"
   class_train_pipeline = Pipeline(pipeline_name=class_run_name,␣
↪pipeline_path="ml_pipeline/")
   class_train_pipeline.load()

   x6 = ModelLoaderBlock(device=device)(class_train_pipeline.pipes[0],␣
↪class_train_pipeline.pipes[-1])
   x7 = ClassifierRunnerBlock(save_data=False, device=device)(x5, x6)
   x8 = ClassifierRunnerBlock(save_data=False, device=device)(x3, x6)

   class_res.append(x7.results)

   x9 = ScalingTestBlock(1000, N, s)(x7, x8)

   scaling_pipe = Pipeline()
   scaling_pipe.compose(x1, x2, x3, x4, x5, x6, x7, x8, x9)
   scaling_pipe.fit()

   res = scaling_pipe.pipes[-1].results

   res_lst.append(res)
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e31840790>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e31840160>!
100%|


           | 10/10 [00:00<00:00, 74.53it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e318405e0>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
```

```
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e318404f0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e31840c40>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e586a6d00>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cfa30a0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cfa30d0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2cfa3160>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfa33a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2cbf39d0>!
100%|


           | 10/10 [00:00<00:00, 82.08it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2cbf3940>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cc28ac0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cc28eb0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cc2a280>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cc2a340>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cc2a400>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2cc2a490>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cc6ac10>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2ca7d610>!
```

```
100%|

           | 10/10 [00:00<00:00, 77.53it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2ca7d940>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2ca94d90>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2ca96190>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2ca96550>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2ca96610>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2ca966d0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2ca96760>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e40d81f40>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e40d81340>!
100%|

           | 10/10 [00:00<00:00, 80.14it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2cad9c70>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cad9070>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cad90d0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cad9610>!
```

```
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cad9ee0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cadf3a0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2cadf5e0>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e3be38040>!
100%|


           | 10/10 [00:00<00:00, 81.25it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e3be380a0>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e318c0a90>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e318c0c70>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e31927040>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e31927280>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e31927160>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e31927490>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfc2d30>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e3beea730>!
100%|


           | 10/10 [00:00<00:00, 77.47it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e3bf2edf0>!
```

```
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e3bf5d1f0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e3bf5db80>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e3bf5d310>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e3bf5d250>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e3bf5d730>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e3bf5d6d0>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e3bfb4d60>!
100%|


           | 10/10 [00:00<00:00, 80.13it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e3bfb49a0>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cbceb80>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cbcef70>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cbd6340>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cbd6400>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cbd64c0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2cbd6550>!
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfc2d30>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2c8ec700>!
100%|


           | 10/10 [00:00<00:00, 78.92it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2c8eca30>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2c876d90>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2c877190>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c8def40>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2c8def70>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2c8dedf0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2c8ded60>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7ceab81670>!
100%|


           | 10/10 [00:00<00:00, 80.21it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7ceab81fd0>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
```

```
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7ceab8aac0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7ceab8a8e0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7ceab8b7f0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7ceab8b8b0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7ceab8b970>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7ceab8ba00>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfc2d30>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7ceaa9e130>!
100%|


           | 10/10 [00:00<00:00, 82.89it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e58103a30>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7ceaa411c0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7ceaa415b0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7ceaa41940>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7ceaa41a00>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7ceaa41ac0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7ceaa41b50>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7cea911c10>!
```

```
100%|

           | 10/10 [00:00<00:00, 82.74it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7cea911f40>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cea8a32e0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7cea8a36d0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cea8a3a60>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea8a3b20>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea8a3be0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7cea8a3c70>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfc2d30>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7cea7dbe20>!
100%|

           | 10/10 [00:00<00:00, 82.70it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7cea7dbf40>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cea7937f0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7cea793be0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cea796040>!
```

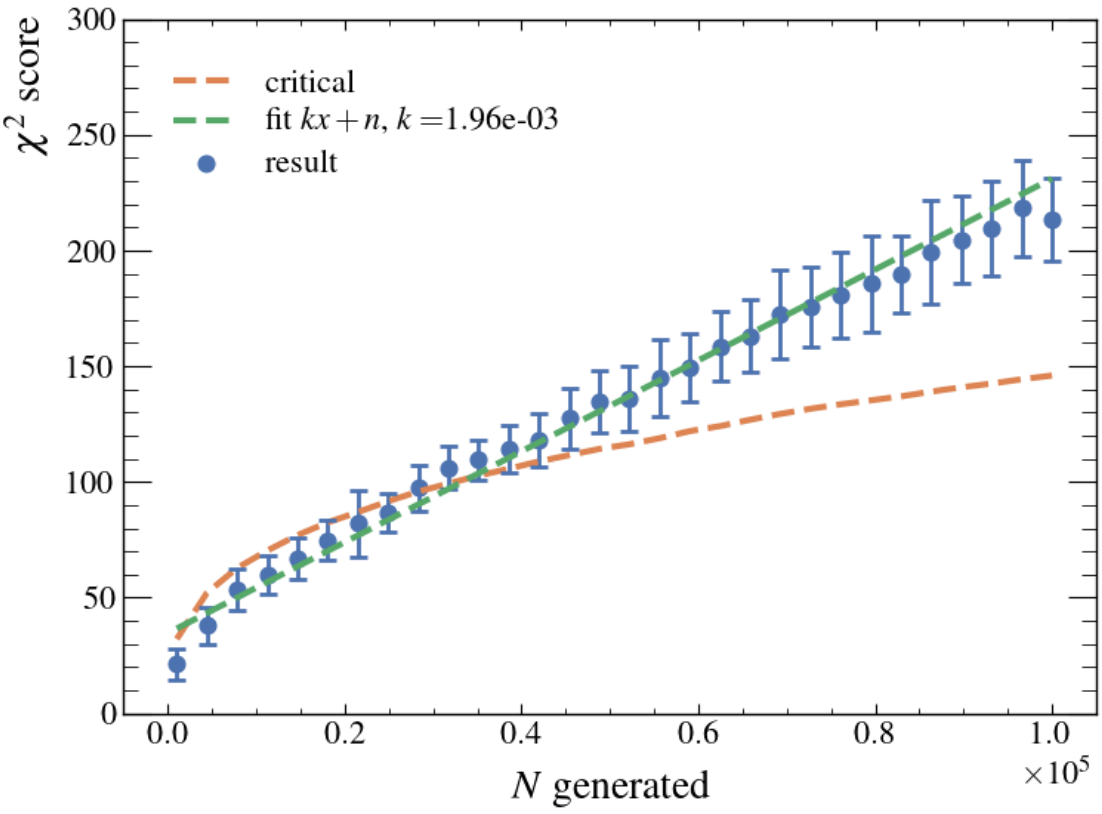
```
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea796070>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea796130>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7cea7961c0>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7cea632a30>!
100%|


           | 10/10 [00:00<00:00, 82.82it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7cea632d60>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cea64a4c0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7cea64a4f0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cea64a880>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea64a940>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea64aa00>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7cea64aa90>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cfc2d30>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7cea562c40>!
100%|


           | 10/10 [00:00<00:00, 83.01it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7cea562f70>!
```

```
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cea4f9310>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7cea4f9700>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cea4f9a90>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea4f9b50>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cea4f9c10>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7cea4f9ca0>!
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c9550a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7cc83e0e50>!
100%|


           | 10/10 [00:00<00:00, 83.04it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7cc83c5b80>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7cc8379550>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7cc8379910>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cc8379ca0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cc8379d60>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cc8379e20>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7cc8379eb0>!
```

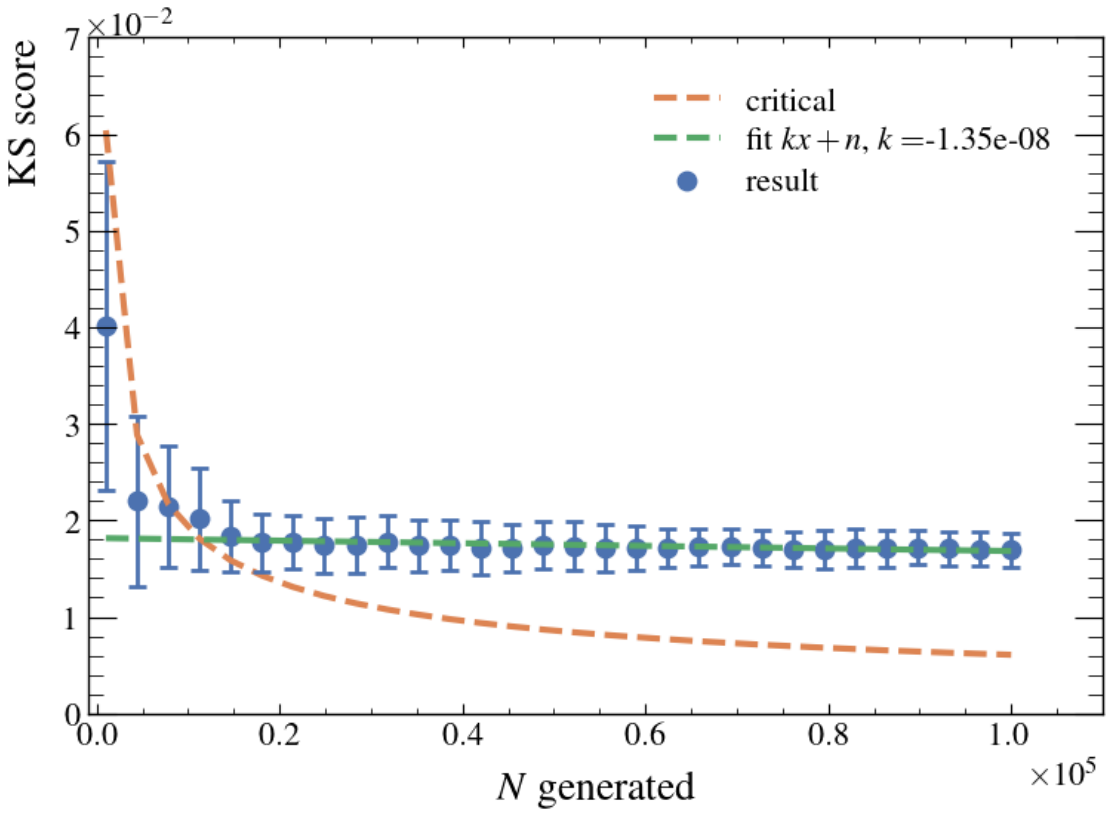
[14]:
```
chi2_m = np.zeros((r, s))
ks_m = np.zeros((r, s))
chi2_m_crit = np.zeros((r, s))
ks_m_crit = np.zeros((r, s))

for ri in range(r):
    for si in range(s):
        chi2, ks = res_lst[ri][si]

        chi2_m[ri, si] = chi2["chi2"].to_numpy()[0]
        ks_m[ri, si] = ks["ks"].to_numpy()[0]
        chi2_m_crit[ri, si] = chi2["crit"].to_numpy()[0]
        ks_m_crit [ri, si] = ks["crit"].to_numpy()[0]
```

[15]:
```
N_range = x9.N_range
```

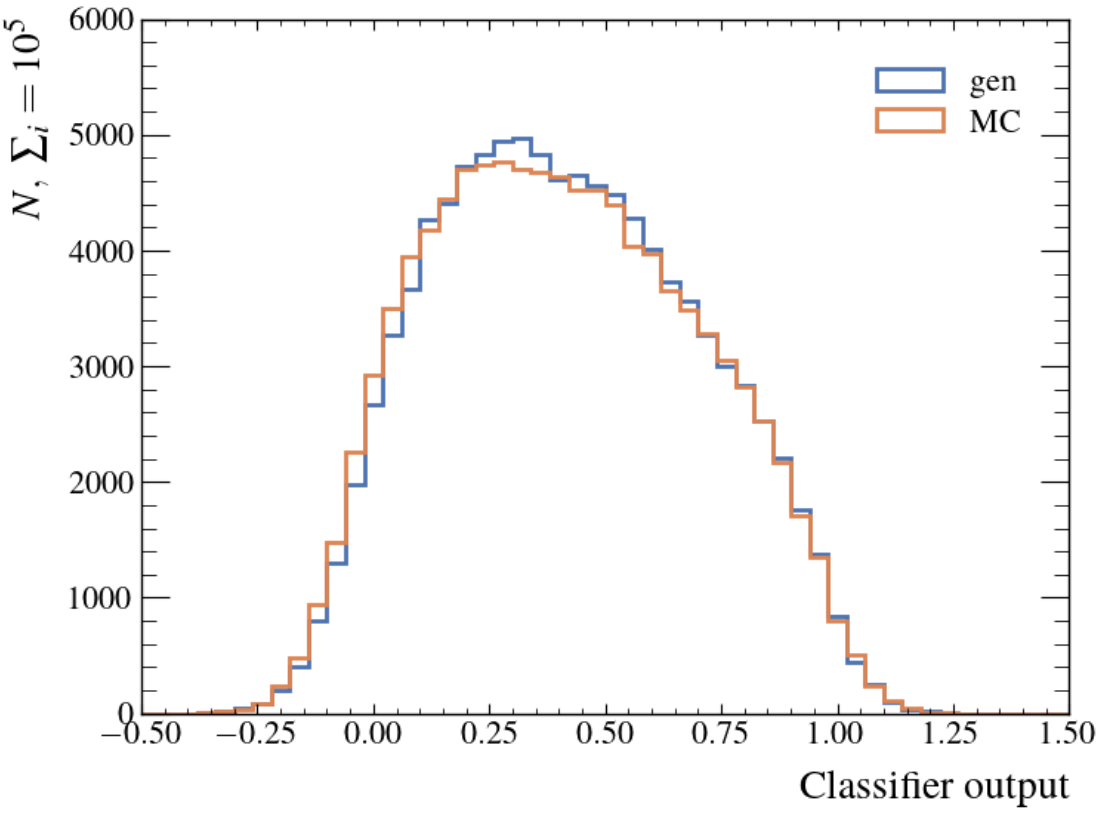
[16]:
```
plt.scatter(N_range, chi2_m.mean(axis=0), s=60)
plt.plot(N_range, chi2_m_crit.mean(axis=0), ls='--', c='C1', lw=3)

plt.errorbar(N_range, chi2_m.mean(axis=0), yerr=chi2_m.std(axis=0), capsize=4,␣
 ↪ls="none", lw=2, capthick=2)

def func(x, k, n):
    return k * x + n

popt, pcov = curve_fit(func, N_range, chi2_m.mean(axis=0), sigma=chi2_m.
 ↪std(axis=0))
plt.plot(N_range, func(N_range, *popt), ls='--', c="C2", lw=3)

plt.xlim([-5000, 1.05*10**5])
plt.xlabel("$N$ generated", loc="center")
plt.ylabel("$\chi^2$ score")
plt.legend(["critical", f"fit $kx+n$, $k=${popt[0]:.2e}", "result"])

plt.tight_layout()
plt.savefig(saved + "realnvp_chi2_scaling.pdf")
```



[17]:
```
plt.scatter(N_range, ks_m.mean(axis=0), s=80)
plt.plot(N_range, ks_m_crit.mean(axis=0), ls='--', c="C1", lw=3)
plt.errorbar(N_range, ks_m.mean(axis=0), yerr=ks_m.std(axis=0), capsize=4,␣
 ↪ls="none", lw=2, capthick=2)

def func(x, k, n):
    return k * x + n

popt, pcov = curve_fit(func, N_range, ks_m.mean(axis=0), sigma=ks_m.std(axis=0))
plt.plot(N_range, func(N_range, *popt), ls='--', c="C2", lw=3)

plt.xlim([-1000, 1.1*10**5])
plt.xlabel("$N$ generated", loc="center")
plt.ylabel("KS score")
plt.legend(["critical", f"fit $kx+n$, $k=${popt[0]:.2e}", "result"])
plt.tight_layout()
plt.savefig(saved + "realnvp_ks_scaling.pdf")
```



[18]:
```
a = x7.results
b = x8.results

plt.hist(b, histtype="step", range=(-0.5, 1.5), bins=50, lw=2)
plt.hist(a, histtype="step", range=(-0.5, 1.5), bins=50, lw=2)
plt.xlabel("Classifier output")
plt.ylabel("$N,\>$ $\sum_i=10^5$")
plt.legend(["gen", "MC"])
plt.tight_layout()
plt.savefig(saved + "realnvp_class_out.pdf")
```



[24]:
```
N = 10 ** 5
device = "cpu"

x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipelines[3][0].pipes

x1 = ModelLoaderBlock()(x_ConfigBuilderBlock, x_ModelTrainerBlock)

x2 = DataGeneratorBlock(N, model_type="flow", chunks=10, device=device)(x1)
x3 = GeneratedDataVerifierBlock(save_data=False, device=device)(x1, x2)

x4 = DatasetBuilderBlock()(x_ConfigBuilderBlock)
x5 = ReferenceDataLoaderBlock()(x4)

x6 = ScalingTestBlock(10000, N, 30)(x5, x3)

scaling_pipe_full = Pipeline()
scaling_pipe_full.compose(x1, x2, x3, x4, x5, x6)
scaling_pipe_full.fit()
```

```
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cabb5e0>!
```

```
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2cabb7f0>!
100%|


           | 10/10 [00:00<00:00, 11.32it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2cabb640>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:Scaled data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cabb790>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cabb820>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f7e2cabb880>!
```

[24]:
```
<ml_hep_sim.pipeline.pipes.Pipeline at 0x7f7e2cabba00>
```

[25]:
```
N_range = x6.N_range
```

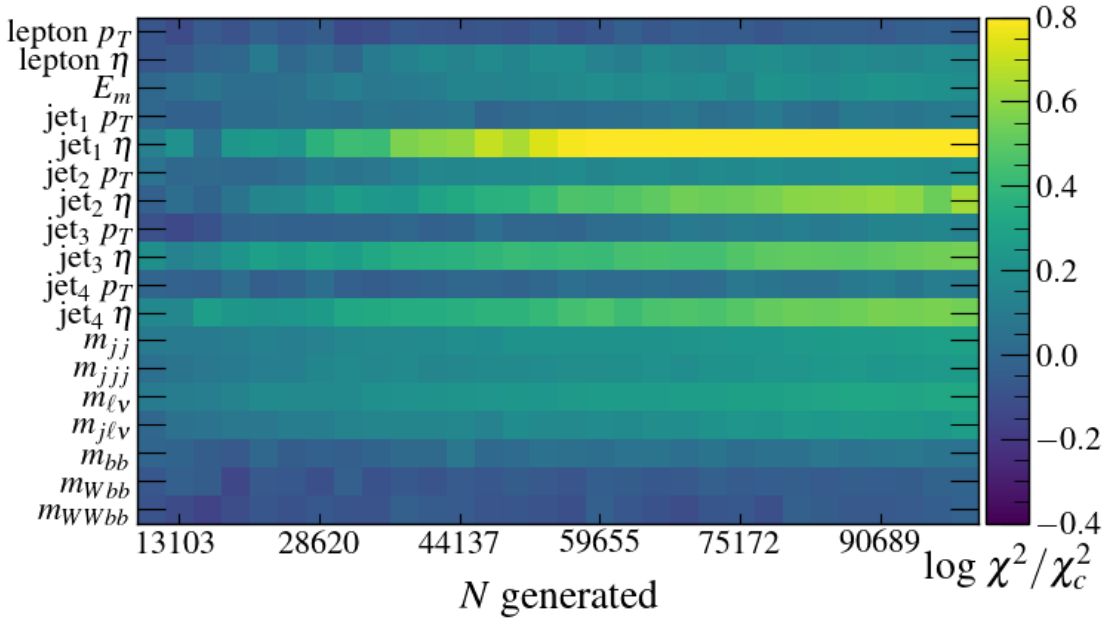
[26]:
```
res = scaling_pipe_full.pipes[-1].results
```

[27]:
```
s_chi2 = np.zeros((18, len(res)))
s_chi2_crit = np.zeros((18, len(res)))
s_ks = np.zeros((18, len(res)))
s_ks_crit = np.zeros((18, len(res)))
```

[28]:
```
for i, r in enumerate(res):
    chi2, ks = r
    s_chi2[:, i] = chi2["chi2"].to_numpy()
    s_ks[:, i] = ks["ks"].to_numpy()
    s_chi2_crit[:, i] = chi2["crit"].to_numpy()
    s_ks_crit[:, i] = ks["crit"].to_numpy()
```

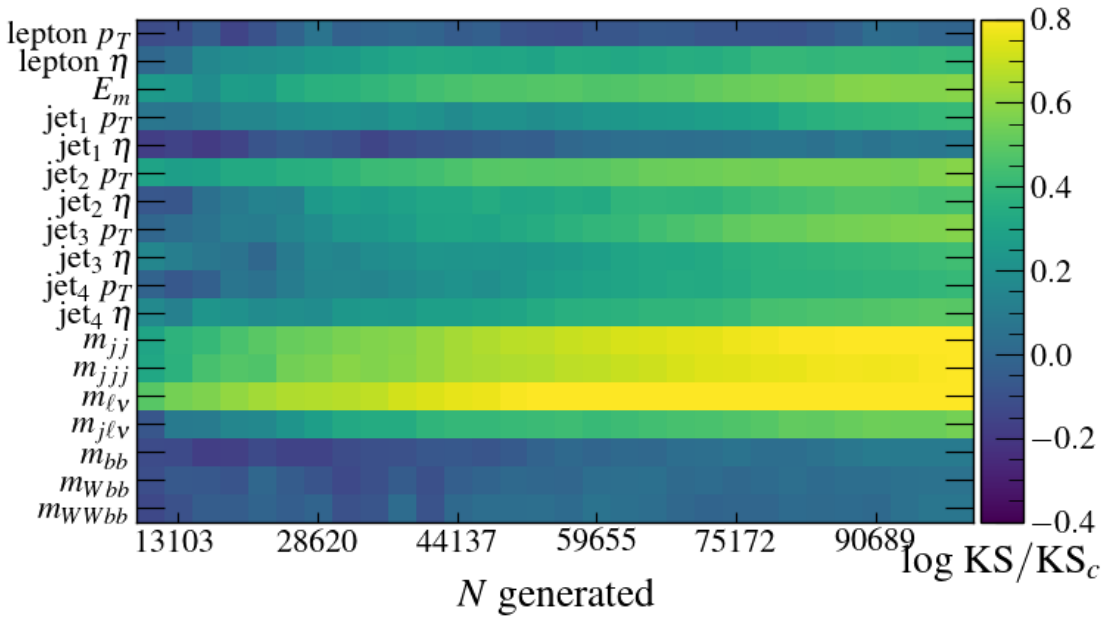
[29]:
```
from mpl_toolkits.axes_grid1 import make_axes_locatable
```

```
plt.figure()
ax = plt.gca()

im = ax.imshow(np.log10(s_chi2 / s_chi2_crit))
im.set_clim(-0.4, 0.8)

ax.set_yticks(np.arange(0, 18, 1))
ax.set_yticklabels(LATEX_COLNAMES)

ax.minorticks_off()

ax.set_xticks(np.arange(0, 30, 1)[1::5])
ax.set_xticklabels(N_range[1::5])

ax.set_xlabel("$N$ generated", loc="center")

divider = make_axes_locatable(ax)
cax = divider.append_axes("right", size="5%", pad=0.05)
cbar = plt.colorbar(im, cax=cax)
cax.set_xlabel('log $\chi^2/\chi^2_c$', loc="center")

plt.tight_layout()
plt.savefig(saved + "imshow_realnvp_chi2.pdf")
```



[30]:
```
plt.figure()
ax = plt.gca()
```

```
im = ax.imshow(np.log10(s_ks / s_ks_crit))
im.set_clim(-0.4, 0.8)

ax.set_yticks(np.arange(0, 18, 1))
ax.set_yticklabels(LATEX_COLNAMES)

ax.minorticks_off()

ax.set_xticks(np.arange(0, 30, 1)[1::5])
ax.set_xticklabels(N_range[1::5])

ax.set_xlabel("$N$ generated", loc="center")

divider = make_axes_locatable(ax)
cax = divider.append_axes("right", size="5%", pad=0.05)
cbar = plt.colorbar(im, cax=cax)
cax.set_xlabel('log KS$/$KS$_c$', loc="center")

plt.tight_layout()
plt.savefig(saved + "imshow_realnvp_ks.pdf")
```



[31]:
```
run_name = "Higgs_Glow"
n = 10

pipeline_path = f"ml_pipeline/{run_name}/"
```

```
pipeline_sig_name = run_name + f"_flow_blocks_{n}_sig_train_pipe"
pipeline_bkg_name = run_name + f"_flow_blocks_{n}_train_pipe"

pipe_sig = Pipeline(pipeline_name=pipeline_sig_name,␣
 ↪pipeline_path=pipeline_path).load().pipes
pipe_bkg = Pipeline(pipeline_name=pipeline_bkg_name,␣
 ↪pipeline_path=pipeline_path).load().pipes
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
```

[32]:
```
pipe_bkg
```

[32]:
```
[<ml_hep_sim.pipeline.blocks.ConfigBuilderBlock at 0x7f7e2c9965e0>,
 <ml_hep_sim.pipeline.blocks.ModelBuilderBlock at 0x7f7e2c996610>,
 <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock at 0x7f7ceab80af0>,
 <ml_hep_sim.pipeline.blocks.ModelTrainerBlock at 0x7f7ceab80df0>]
```

[33]:
```
N = 50

x1 = ModelLoaderBlock()(pipe_sig[0], pipe_sig[-1], pipe_sig[1])
x2 = ModelLoaderBlock()(pipe_bkg[0], pipe_bkg[-1], pipe_bkg[1])

x3 = DatasetBuilderBlock()(pipe_sig[0])
x4 = DatasetBuilderBlock()(pipe_bkg[0])

x5 = CouplingModelTestingBlock(N, mean=False)(x4, x1, pipe_sig[0])
x6 = CouplingModelTestingBlock(N, mean=False)(x4, x2, pipe_bkg[0])
```

[34]:
```
pipe = Pipeline()
pipe.compose(x1, x2, x3, x4, x5, x6)
pipe.fit()
```

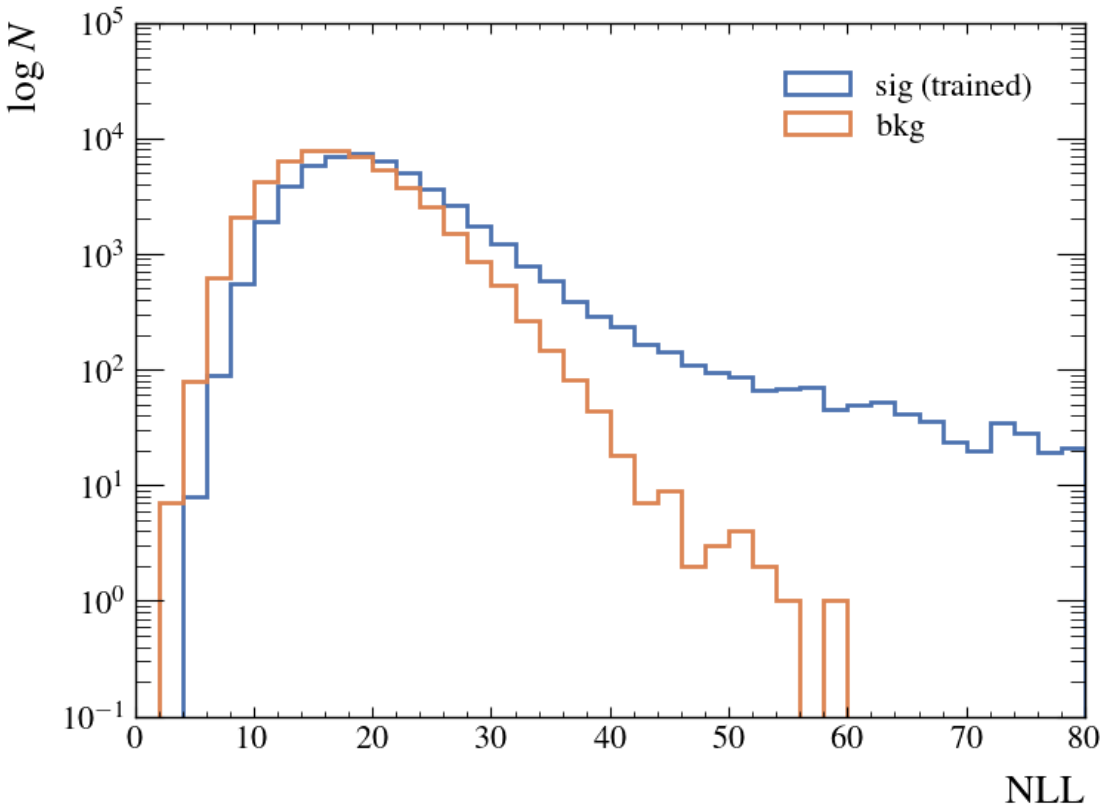
```
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c8c84c0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2c8c85b0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2c8c84f0>!
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2c8c8340>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e2c8c8a90>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.CouplingModelTestingBlock
object at 0x7f7e2c8c8970>!
```

```
/home/jan/ML/ML_venv/lib/python3.8/site-
packages/pytorch_lightning/core/datamodule.py:423: LightningDeprecationWarning:
DataModule.prepare_data has already been called, so it will not be called again.
In v1.6 this behavior will change to always call DataModule.prepare_data.
  rank_zero_deprecation(
/home/jan/ML/ML_venv/lib/python3.8/site-
packages/pytorch_lightning/core/datamodule.py:423: LightningDeprecationWarning:
DataModule.setup has already been called, so it will not be called again. In
v1.6 this behavior will change to always call DataModule.setup.
  rank_zero_deprecation(
```

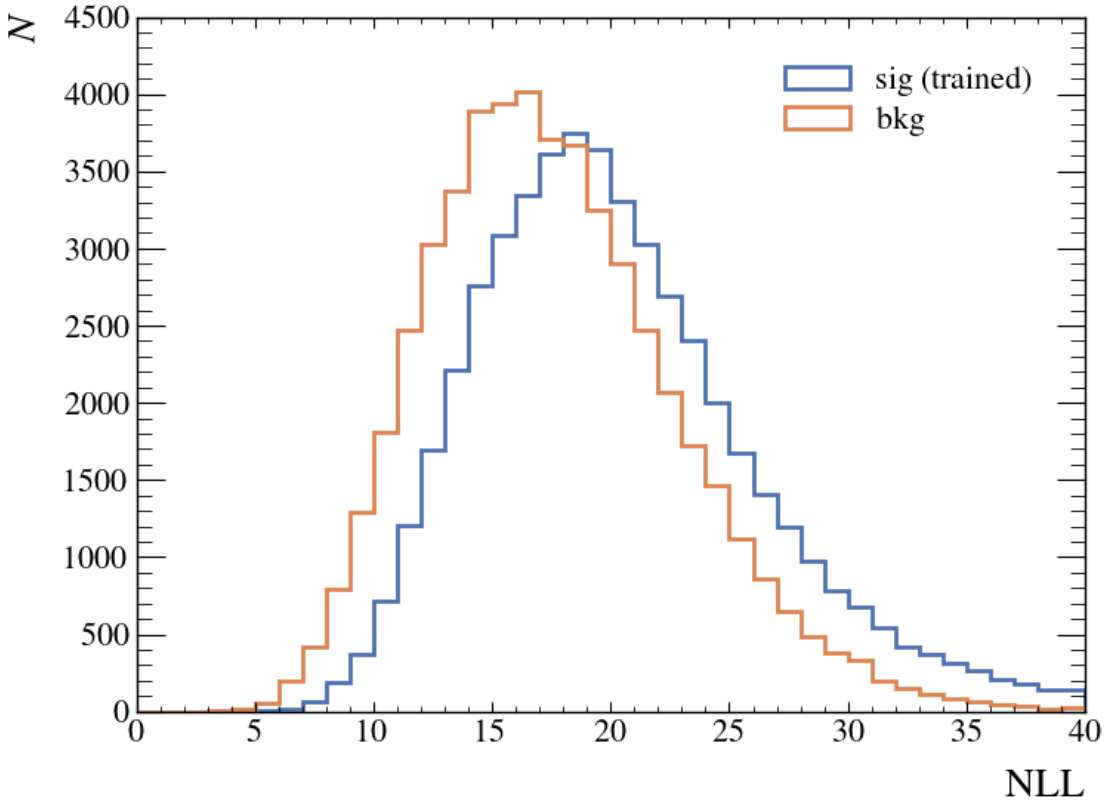
[34]:
```
<ml_hep_sim.pipeline.pipes.Pipeline at 0x7f7cea93c070>
```

[35]:
```
sig_dist = pipe.pipes[-2].results[-1].flatten().numpy()
bkg_dist = pipe.pipes[-1].results[-1].flatten().numpy()
```

[36]:
```
plt.hist(sig_dist, histtype="step", bins=40, range=(0, 80), lw=2)
plt.hist(bkg_dist, histtype="step", bins=40, range=(0, 80), lw=2)
plt.yscale("log")
plt.xlabel("NLL")
plt.ylabel("log $N$")
plt.legend(["sig (trained)", "bkg"])
plt.savefig(saved + "flow_class_realnvp_log.pdf")
plt.tight_layout()
```



[37]:
```
plt.hist(sig_dist, histtype="step", bins=40, range=(0, 40), lw=2)
plt.hist(bkg_dist, histtype="step", bins=40, range=(0, 40), lw=2)
plt.xlabel("NLL")
plt.ylabel("$N$")
plt.legend(["sig (trained)", "bkg"])
plt.savefig(saved + "flow_class_realnvp.pdf")
plt.tight_layout()
```



[38]:
```
run_name = "Higgs_Glow"
n = 10

pipeline_path = f"ml_pipeline/{run_name}/"
pipeline_sig_name = run_name + f"_flow_blocks_{n}_sig_train_pipe"
pipeline_bkg_name = run_name + f"_flow_blocks_{n}_train_pipe"

pipe_sig = Pipeline(pipeline_name=pipeline_sig_name,␣
 ↪pipeline_path=pipeline_path).load().pipes
pipe_bkg = Pipeline(pipeline_name=pipeline_bkg_name,␣
 ↪pipeline_path=pipeline_path).load().pipes
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
WARNING:root:Number of composed and loaded pipes did not match! Loading
anyway…
```

[39]:
```
N = 10**5
device = "cpu"

x1 = ModelLoaderBlock()(pipe_sig[0], pipe_sig[-1], pipe_sig[1])
x2 = ModelLoaderBlock()(pipe_bkg[0], pipe_bkg[-1], pipe_bkg[1])

x3 = DataGeneratorBlock(N, model_type="flow", chunks=10, device=device)(x1)
x4 = GeneratedDataVerifierBlock(save_data=False, device=device,␣
↪rescale_data=False)(x1, x3)

x5 = DataGeneratorBlock(N, model_type="flow", chunks=10, device=device)(x2)
x6 = GeneratedDataVerifierBlock(save_data=False, device=device,␣
↪rescale_data=False)(x2, x5)

config = copy.deepcopy(pipe_sig[0].config)
config["datasets"]["data_name"] = "higgs_bkg"
config["datasets"]["data_params"]["subset_n"] = [10 ** 5, 10 ** 5, 10 ** 6]

x71 = DatasetBuilderBlock(config=config)()
x81 = ReferenceDataLoaderBlock(rescale_reference="logit_normal")(x71)

config = copy.deepcopy(pipe_sig[0].config)
config["datasets"]["data_name"] = "higgs_sig"
config["datasets"]["data_params"]["subset_n"] = [10 ** 5, 10 ** 5, 10 ** 6]

x72 = DatasetBuilderBlock(config=config)()
x82 = ReferenceDataLoaderBlock(rescale_reference="logit_normal")(x72)

class_run_name = "Higgs_resnet_classifier_train_pipeline"
class_train_pipeline = Pipeline(pipeline_name=class_run_name,␣
↪pipeline_path="ml_pipeline/")
class_train_pipeline.load()

x9 = ModelLoaderBlock(device=device)(class_train_pipeline.pipes[0],␣
↪class_train_pipeline.pipes[-1])
x10 = ClassifierRunnerBlock(save_data=False)(x4, x9) # sig gen
x11 = ClassifierRunnerBlock(save_data=False)(x6, x9) # bkg gen

x12 = ClassifierRunnerBlock(save_data=False)(x81, x9) # MC bkg
x13 = ClassifierRunnerBlock(save_data=False)(x82, x9) # MC sig
```

```
WARNING:root:Number of composed and loaded pipes did not match! Loading
```

anyway…

[40]:
```
pipe = Pipeline()
pipe.compose(x1, x2, x3, x4, x5, x6, x71, x81, x72, x82, x9, x10, x11, x12, x13)
pipe.fit()
```

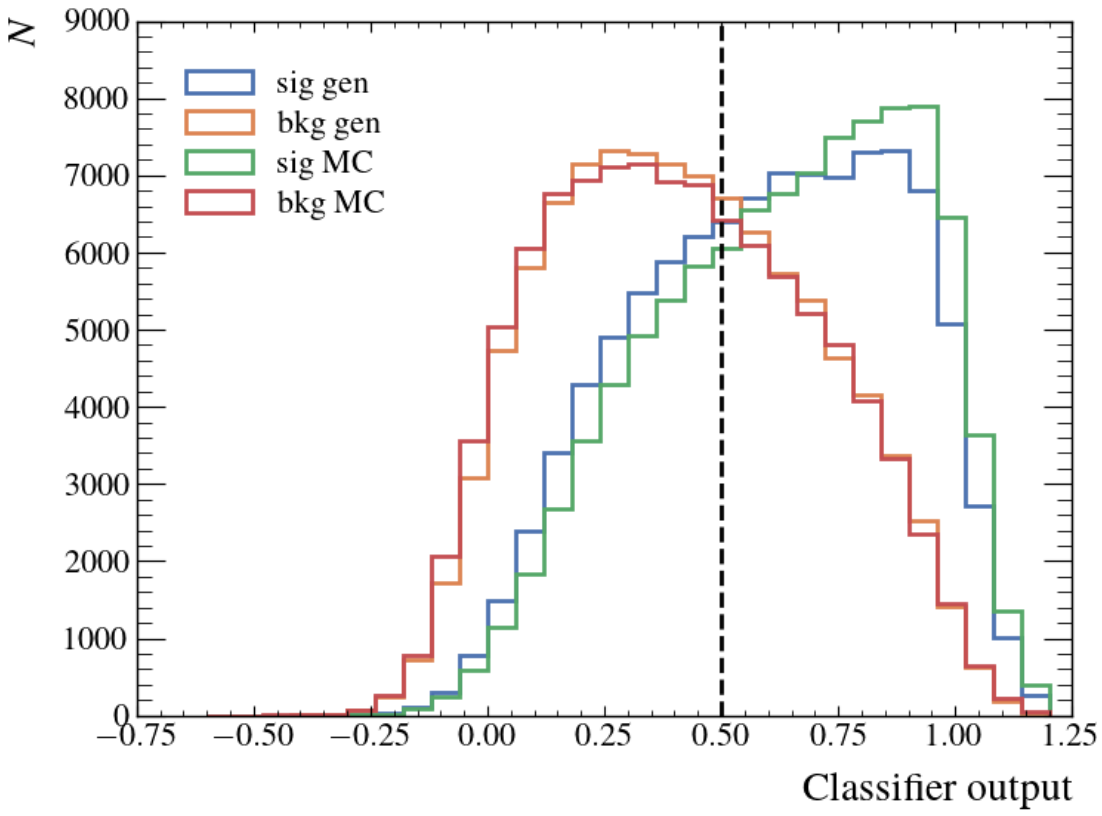
```
WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cadf280>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7e2cadf820>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2cadf670>!
100%|


           | 10/10 [00:00<00:00, 11.35it/s]
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2cadf100>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f7e2cadfc40>!
100%|


           | 10/10 [00:00<00:00, 11.32it/s]
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f7e2cadf9d0>!
WARNING:root:Generated data check…
WARNING:root:nan OK
WARNING:root:pos-inf OK
WARNING:root:neg-inf OK
WARNING:root:pos-inf or neg-inf OK
WARNING:root:pos-inf or neg-inf or nan OK
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cadf790>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cadf1f0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f7e2cadfb50>!
WARNING:root:fitting #9: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f7e2cadf9a0>!
WARNING:root:fitting #10: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f7cc83798e0>!
```

```
WARNING:root:fitting #11: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cadfa00>!
WARNING:root:fitting #12: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7cc8379af0>!
WARNING:root:fitting #13: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cadfe20>!
WARNING:root:fitting #14: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f7e2cadf580>!
```

[40]:
```
<ml_hep_sim.pipeline.pipes.Pipeline at 0x7f7e80294310>
```

[41]:
```
sig_gen = pipe.pipes[-4].results
bkg_gen = pipe.pipes[-3].results
sig_mc = pipe.pipes[-1].results[:len(sig_gen)]
bkg_mc = pipe.pipes[-2].results[:len(sig_gen)]
```

[42]:
```
plt.hist(sig_gen, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)
plt.hist(bkg_gen, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)
plt.hist(sig_mc, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)
plt.hist(bkg_mc, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)

plt.legend(["sig gen", "bkg gen", "sig MC", "bkg MC"], loc="upper left")
plt.axvline(0.5, c="k", ls='--')
plt.ylabel("$N$")
plt.xlabel("Classifier output")
plt.tight_layout()
plt.savefig(saved + "class_gen_mc.pdf")
```

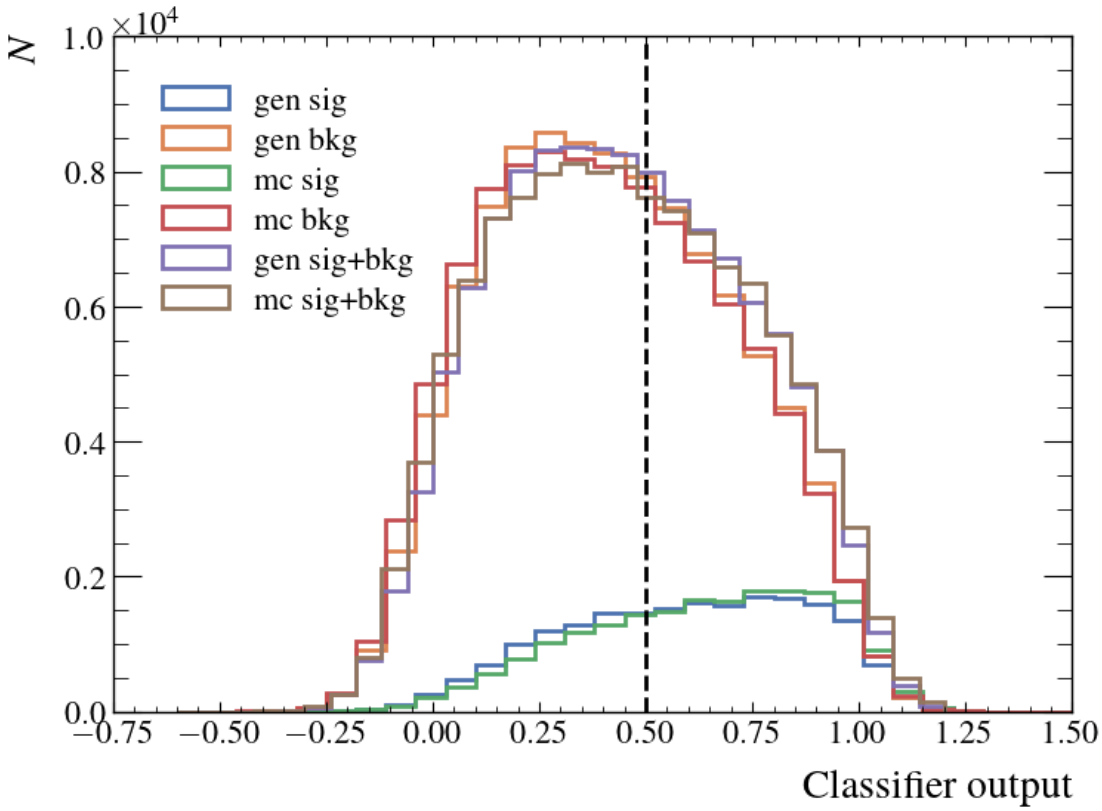


[43]:
```
N = len(sig_gen)
sig_per = 0.2
N_sig = int(N * sig_per)

sig_gen = sig_gen[:N_sig]
sig_mc = sig_mc[:N_sig]
```

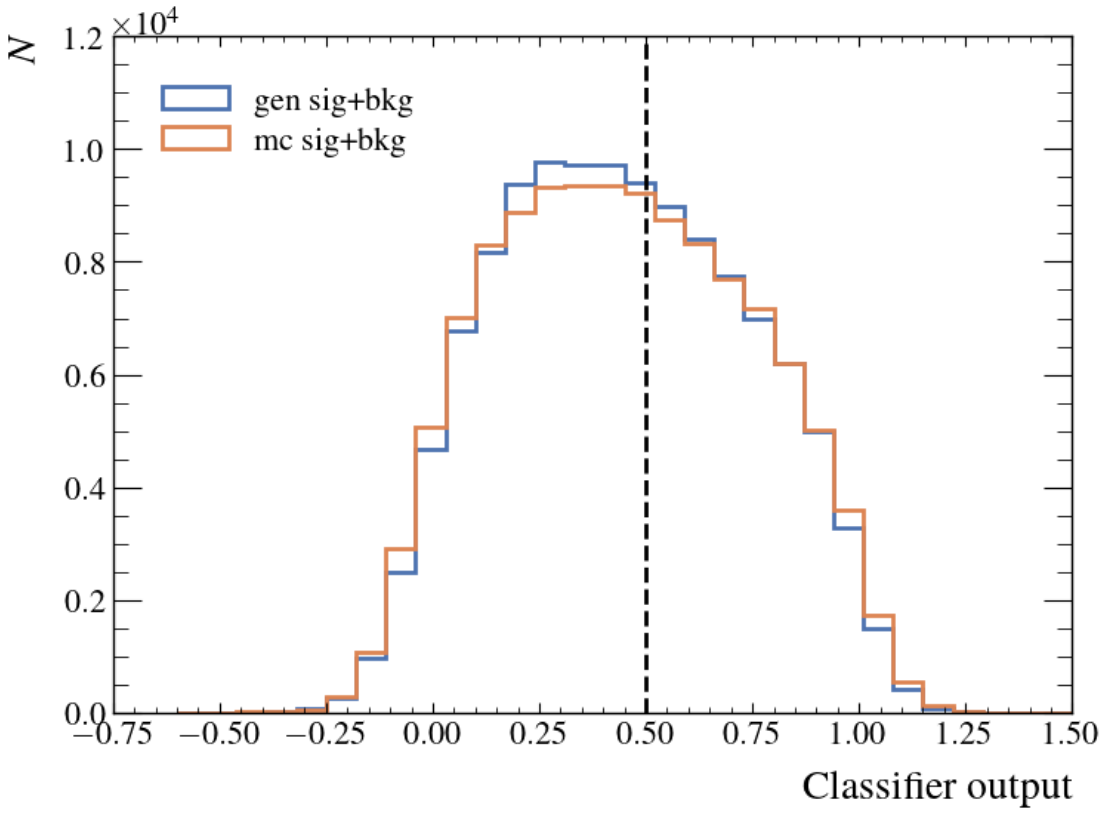
[44]:
```
sig_bkg_gen = np.concatenate([sig_gen, bkg_gen])
sig_bkg_mc = np.concatenate([sig_mc, bkg_mc])
```

[45]:
```
plt.hist(sig_gen, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)
plt.hist(bkg_gen, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)
plt.hist(sig_mc, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)
plt.hist(bkg_mc, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)

plt.hist(sig_bkg_gen, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)
plt.hist(sig_bkg_mc, histtype="step", bins=30, range=(-0.6, 1.2), lw=2)

plt.legend(["gen sig", "gen bkg", "mc sig", "mc bkg", "gen sig+bkg", "mc␣
 ↪sig+bkg"], loc="upper left")
plt.axvline(0.5, c="k", ls='--')
```

```
plt.ylabel("$N$")
plt.xlabel("Classifier output")
plt.tight_layout()
plt.savefig(saved + "class_gen_mc_comp.pdf")
```



[46]:
```
plt.hist(sig_bkg_gen, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)
plt.hist(sig_bkg_mc, histtype="step", bins=30, range=(-0.6, 1.5), lw=2)

plt.legend(["gen sig+bkg", "mc sig+bkg"], loc="upper left")
plt.axvline(0.5, c="k", ls='--')
plt.ylabel("$N$")
plt.xlabel("Classifier output")
plt.tight_layout()
```


$\times 10^4$
N
1.2
1.0
0.8
0.6
0.4
0.2
0.0
gen sig+bkg
mc sig+bkg
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
1.25
1.50
Classifier output

[ ]: