splines

December 29, 2022

```
[1]: import numpy as np
```

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

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

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

```
[4]: 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, set_size
     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

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

     set_size()
     style_setup(seaborn_pallete=True)
```

```
[5]: num_splines = [4, 8, 12, 32]
```

```
[6]: pipelines = run_spline_pipeline(train=False, gen=False, test=False,␣
      ↪num_splines=num_splines)
```

```
100%|
```

```
                                          | 4/4
[00:00<00:00, 3775.25it/s]
```

```python
[7]: val_losses, steps, times = [], [], []

     for pipeline in pipelines:
         pipes = pipeline.pipes
         x1 = ModelLoaderBlock()(*pipes)._run()
         metrics = x1.metrics

         val_loss =metrics[-1]["val_loss"]
         step = metrics[-2]["step"]
         t = metrics[0]["timestamp"].to_numpy()
         times.append(t[-1] - t[0])
         val_losses.append(val_loss)
         steps.append(step)
```
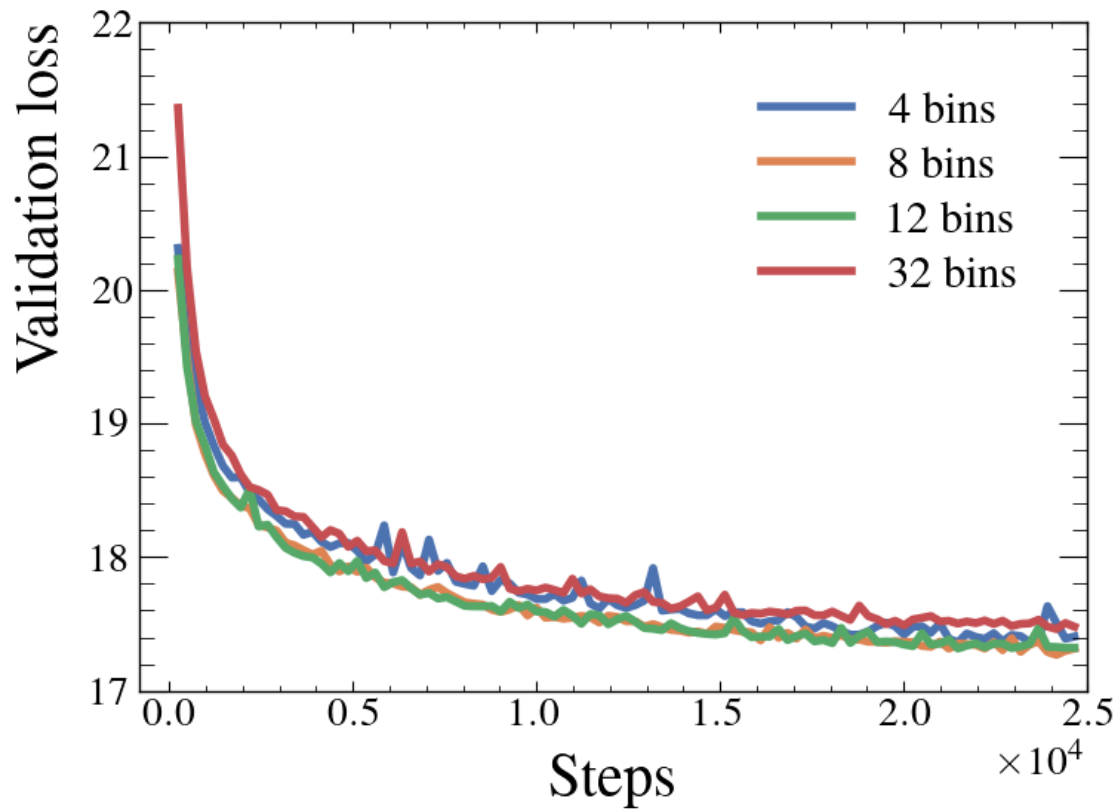
```python
[8]: set_size(s=20)

     plt.plot(steps[0], val_losses[0], lw=4)
     plt.plot(steps[1], val_losses[1], lw=4)
     plt.plot(steps[2], val_losses[2], lw=4)
     plt.plot(steps[3], val_losses[3], lw=4)

     plt.legend(["4 bins", "8 bins", "12 bins", "32 bins"], fontsize=22)
     plt.xlim([-800, 2.5*10**4])
     plt.xlabel("Steps", loc="center", fontsize=29)
     plt.ylabel("Validation loss", fontsize=29)

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

```
[9]: pipeline = run_spline_pipeline(train=False, gen=False, test=False,␣
     ↪num_splines=[32])[0]
```

100%|

    | 1/1
[00:00<00:00, 3912.60it/s]

```
[10]: 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 = pipeline.pipes
```

```python
    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
0x7f5e1479a3d0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e159faa00>!
100%|


                              | 10/10
[00:10<00:00,  1.04s/it]
```

WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f5e1479a430>!
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 0x7f5e1479a460>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f5e1479a4c0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f5e1479a550>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f5e1479a640>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f5e1479a580>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f5e1479a730>!
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 0x7f5e1478fb20>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object at 0x7f5e1569eeb0>!
100%|

                            | 10/10
[00:10<00:00, 1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f5e1466deb0>!
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 0x7f5e156a42e0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f5e15700a30>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f5e157002e0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f5e15700160>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock

```
object at 0x7f5e15700ee0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e15700070>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e14046280>!
100%|
```

| | 10/10
```
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e14046640>!
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 0x7f5e15c4d850>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e15c4d280>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e15cb2ee0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15cb2670>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15cb2a00>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e15cb2c40>!
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e1578d460>!
100%|
```

| | 10/10
```
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e1578d4c0>!
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 0x7f5e15b90850>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e15b90c70>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e15b90f40>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15bae040>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15bae100>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e15bae190>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e15a12b50>!
100%|


                              | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e15a12e80>!
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 0x7f5e15a2cdf0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e15a2a250>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e15a2a670>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15a2a730>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15a2a7f0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e15a2a880>!
```

```
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e15a700a0>!
100%|
```

```
                              | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e15a70160>!
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 0x7f5e15ab0580>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e15ab0af0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e15ab0dc0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15ab0e80>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e15ab0f40>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e15ab0fd0>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e14337a60>!
100%|
```

```
                              | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e14337d90>!
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 0x7f5e142cdca0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e142d0100>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e142d0520>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e142d05e0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e142d06a0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e142d0730>!
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e1422beb0>!
100%|


                              | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e1422bfd0>!
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 0x7f5e141de3d0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e141de970>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5e141dec40>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e141ded00>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5e141dedc0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5e141dee50>!
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
```

```
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de5703850>!
100%|


                                 | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de5703b80>!
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 0x7f5de5718a90>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de5718eb0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de571a310>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de571a3d0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de571a490>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de571a520>!
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de560fee0>!
100%|


                                 | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de55f6c10>!
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 0x7f5de5628160>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de56286d0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de56289a0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de5628a60>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de5628b20>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de5628bb0>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de5522760>!
100%|

                              | 10/10
[00:10<00:00, 1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de55225b0>!
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 0x7f5de54b77f0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de54b7d60>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de54b9070>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de54b9130>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de54b91f0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de54b9280>!
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de5432c70>!

100%|

[00:10<00:00, 1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de5432fa0>!
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 0x7f5de53ca310>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de53ca460>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de53ca730>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de53ca7f0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de53ca8b0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de53ca940>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de53430a0>!
100%|

[00:10<00:00, 1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de5343130>!
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 0x7f5de52db5b0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de52dbb20>!

```
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de52dbdf0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de52dbeb0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de52dbf70>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de52dbfd0>!
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
0x7f5e158f1730>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de51d4a00>!
100%|


                              | 10/10
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de51d4d30>!
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 0x7f5de51e9c40>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de51e9fa0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de51ed4c0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de51ed580>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de51ed640>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de51ed6d0>!
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
0x7f5e14661700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5de50cadc0>!
100%|
```

```
[00:10<00:00,  1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5de50caee0>!
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 0x7f5de507d2e0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5de507d880>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f5de507db50>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de507dc10>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f5de507dcd0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f5de507dd60>!
```

```python
[11]: 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]
```

```python
[12]: N_range = x9.N_range
```

```python
[13]: set_size(20)

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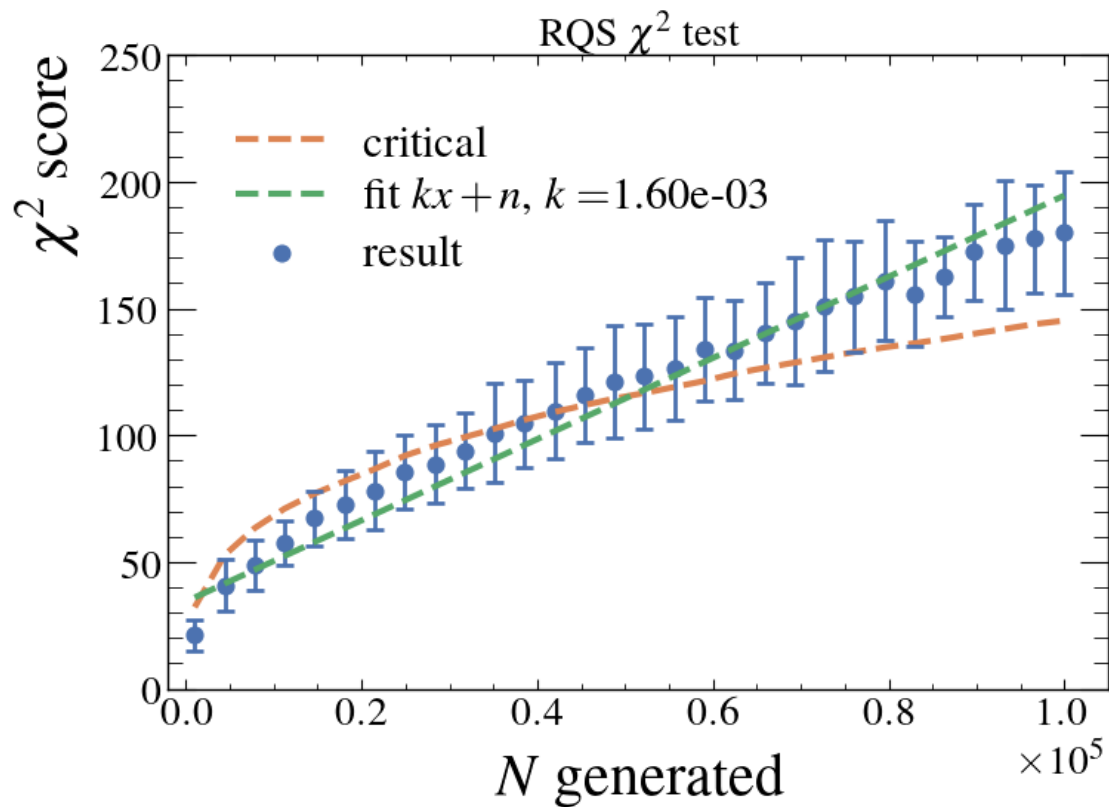
```python
def func(x, k, n):
    return k * x + n

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([-2000, 1.05*10**5])
plt.xlabel("$N$ generated", loc="center", fontsize=29)
plt.ylabel("$\chi^2$ score", fontsize=29)
plt.legend(["critical", f"fit $kx+n$, $k=${popt[0]:.2e}", "result"],␣
 →fontsize=22)
plt.title("RQS $\chi^2$ test")

plt.tight_layout()
plt.savefig(saved + "rqs_chi2_scaling.pdf")
```



```
[14]: set_size(20)

plt.scatter(N_range[1:], ks_m.mean(axis=0)[1:], s=80)
plt.plot(N_range[1:], ks_m_crit.mean(axis=0)[1:], ls='--', c="C1", lw=3)
```

```
plt.errorbar(N_range[1:], ks_m.mean(axis=0)[1:], yerr=ks_m.std(axis=0)[1:],␣
 ↪capsize=4, ls="none", lw=2, capthick=2)

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

popt, pcov = curve_fit(func, N_range[1:], ks_m.mean(axis=0)[1:], sigma=ks_m.
 ↪std(axis=0)[1:])
plt.plot(N_range[1:], func(N_range[1:], *popt), ls='--', c="C2", lw=3)

plt.xlim([1000, 1.1*10**5])
plt.ylim([0, 3.4*10**(-2)])
plt.xlabel("$N$ generated", loc="center", fontsize=29)
plt.ylabel("KS score", fontsize=29)
plt.legend(["critical", f"fit $kx+n$, $k=${popt[0]:.2e}", "result"],␣
 ↪fontsize=22, loc="upper right")
plt.title("RQS KS test")
plt.tight_layout()
plt.savefig(saved + "rqs_ks_scaling.pdf")
```
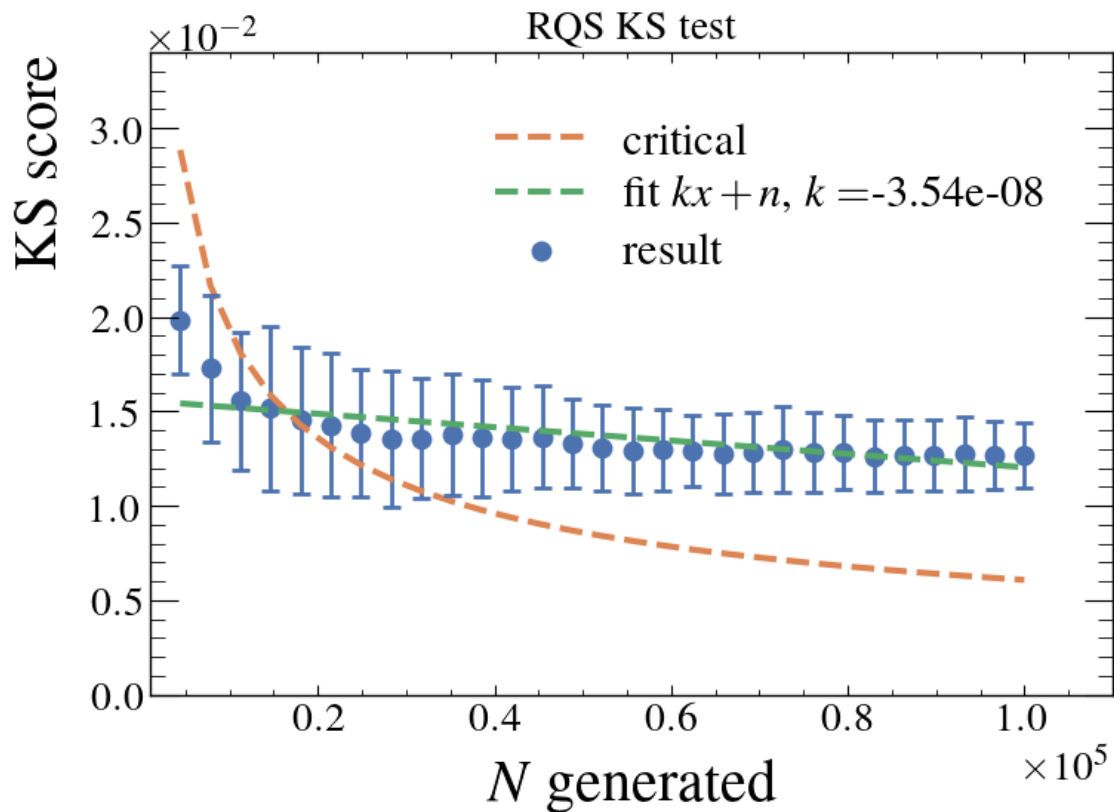
```
[15]:   N = 10 ** 5
        device = "cuda"

        x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipeline.pipes

        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)(x1, x2)

        x4 = DatasetBuilderBlock()(x_ConfigBuilderBlock)
        x5 = ReferenceDataLoaderBlock(device=device)(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
0x7f5e15856580>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f5e158569a0>!
100%|

                                  | 10/10
[00:10<00:00, 1.04s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f5e15856760>!
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 0x7f5e158568e0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f5e15856190>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at

```
                    0x7f5e15856940>!
```

[15]: `<ml_hep_sim.pipeline.pipes.Pipeline at 0x7f5e15856fa0>`

[16]:
```python
N_range = x6.N_range
```

[17]:
```python
res = scaling_pipe_full.pipes[-1].results
```

[18]:
```python
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)))
```

[19]:
```python
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()
```

[20]:
```python
from mpl_toolkits.axes_grid1 import make_axes_locatable

set_size(18)

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=0.2)
cbar = plt.colorbar(im, cax=cax)
cax.set_xlabel('log $\chi^2/\chi^2_c$', loc="center")
ax.set_title("RQS $\chi^2$ test")

plt.tight_layout()
```

```
plt.savefig(saved + "imshow_rqs_chi2.pdf")
```
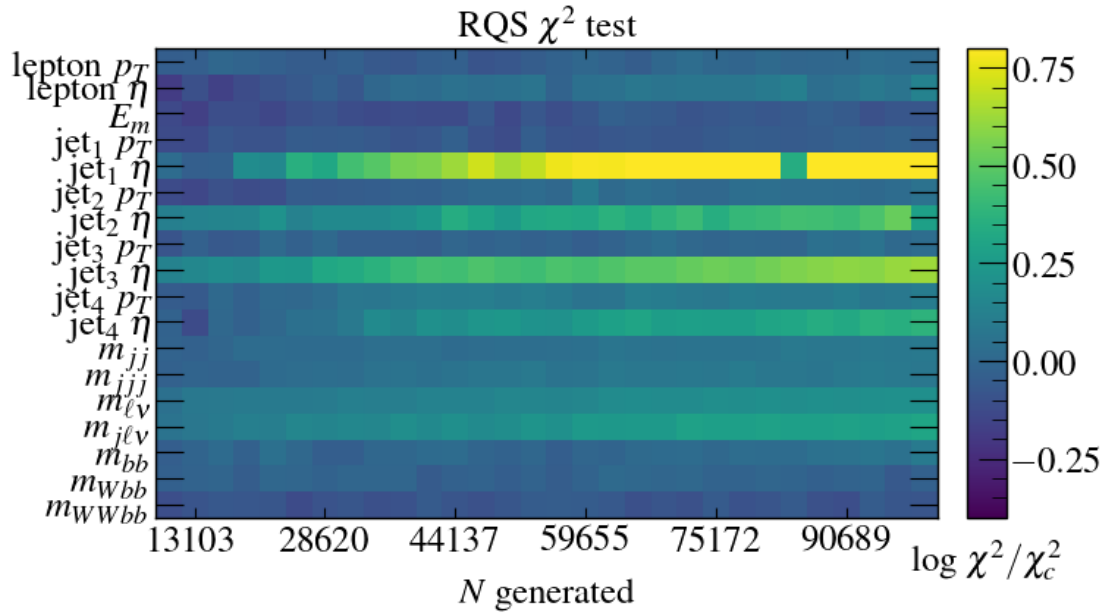

[21]:
```
set_size(18)

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=0.2)
cbar = plt.colorbar(im, cax=cax)
cax.set_xlabel('log KS$/$KS$_c$', loc="center")
ax.set_title("RQS KS test")
```
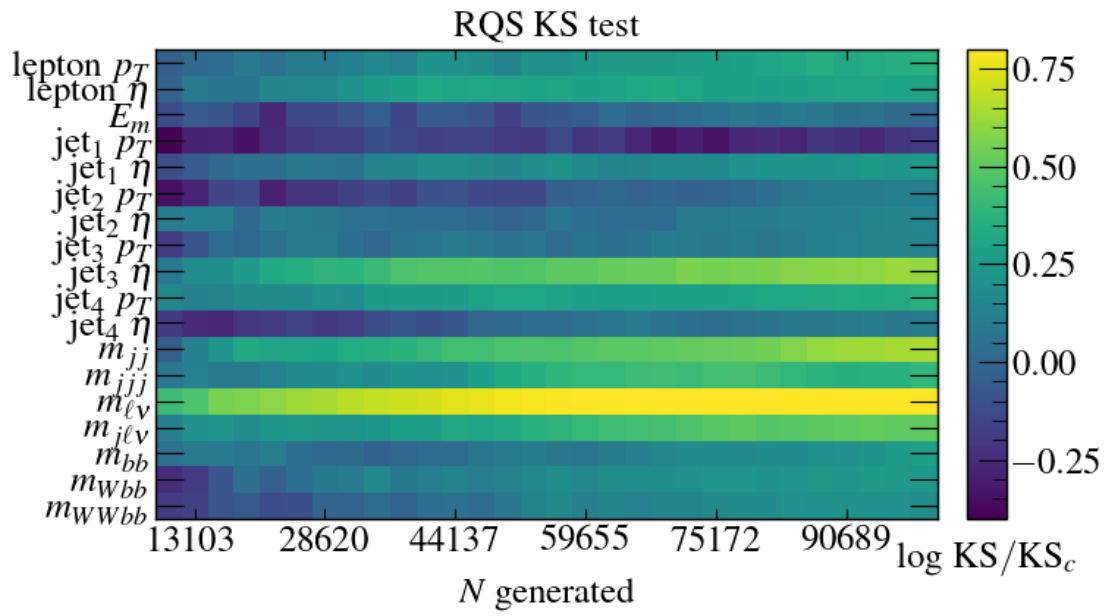
```
plt.tight_layout()
plt.savefig(saved + "imshow_rqs_ks.pdf")
```