

MADEMOG

December 28, 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, 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 numpy as np
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt
from tqdm import tqdm
import copy

style_setup(seaborn_pallete=True)
```

```
[4]: num_mogs_made = np.concatenate([[1], np.arange(2, 22, 2)])
```

```
[5]: pipelines = run_maf_pipeline(train=False, gen=False, test=False, use_mog=True,
    use_maf=False, num_mogs=num_mogs_made)
```

```
100%|
```

```
| 11/11 [00:00<00:00, 2543.26it/s]
```

```
[6]: results = []

N = 40 # batch size (i.e. 1024 * N data points)

for pipe in tqdm(pipelines):
    x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipe.pipes
```

```

x1 = ModelLoaderBlock()(x_ConfigBuilderBlock, x_ModelTrainerBlock)
x2 = DatasetBuilderBlock()(x_ConfigBuilderBlock)
x3 = MADEMOGModelTestingBlock(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%|
| 0/11 [00:00<?, ?it/s]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300bb50ca0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300bb505b0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300bb50820>!
9%|
| 1/11 [00:03<00:30, 3.05s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300bb500a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300bb506d0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300bb50400>!
18%|
| 2/11 [00:05<00:25, 2.87s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300b9c2f40>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b9c2a60>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300ba27070>!
27%|
| 3/11 [00:08<00:23, 2.90s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300ba275e0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300ba27730>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b97d670>!
36%|
| 4/11 [00:11<00:19, 2.85s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300ba273a0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b97dfa0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b97d3a0>!
45%|

```

```

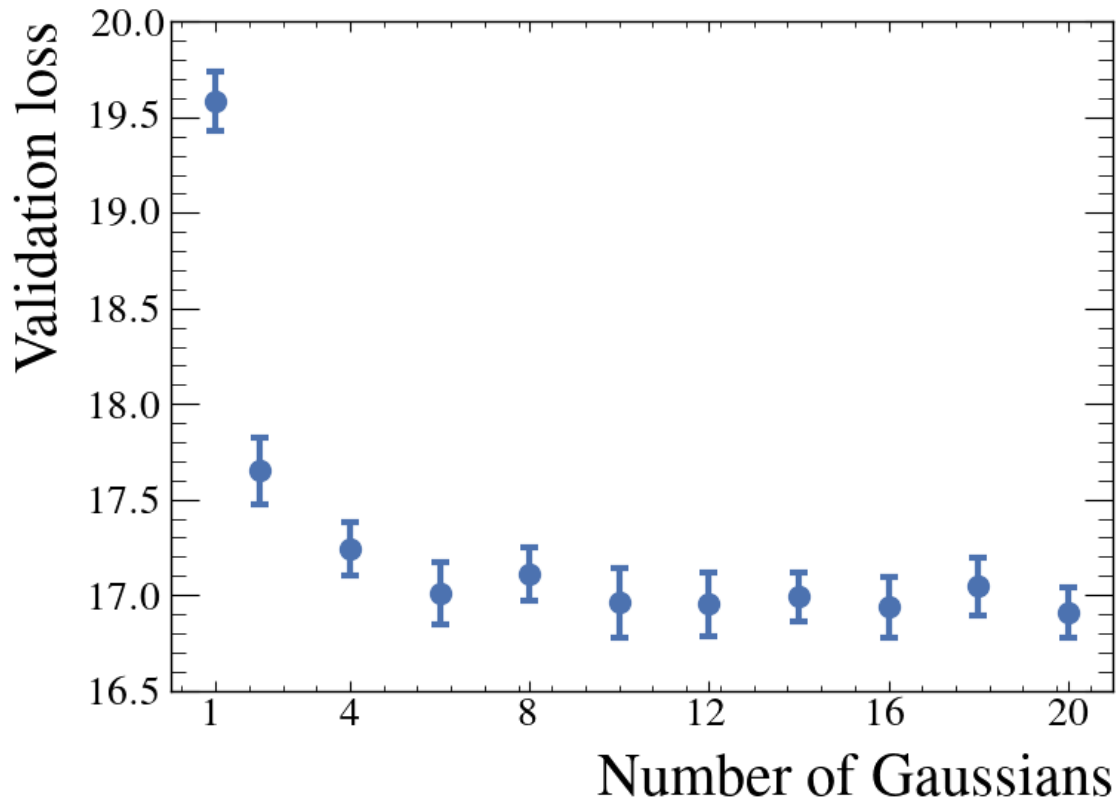
| 5/11 [00:14<00:17, 2.90s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300b986fd0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b986400>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b90bbb0>!
55%|
| 6/11 [00:17<00:14, 2.85s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300b8bc700>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b8bc4f0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b8bc940>!
64%|
| 7/11 [00:20<00:11, 2.87s/it]WARNING:root:fitting #0:
<ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f300b8c6f70>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b8c6ee0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b89d610>!
73%|
| 8/11 [00:23<00:08,
2.93s/it]WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock
object at 0x7f300b8cc130>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b8cc310>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b853b50>!
82%|
| 9/11 [00:26<00:05,
2.91s/it]WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock
object at 0x7f300b853b20>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300b8857f0>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f300b854460>!
91%|
| 10/11 [00:29<00:02,
2.98s/it]WARNING:root:fitting #0: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock
object at 0x7f300bafd4c0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f300bafd340>!
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.MADEMOGModelTestingBlock
object at 0x7f2ff77e2af0>!
100%|
| 11/11 [00:32<00:00, 2.93s/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]: set_size(20)
plt.scatter(num_mogs_made, m_lst, s=100)
plt.errorbar(num_mogs_made, m_lst, yerr=s_lst, capsize=4, ls="none", lw=3,
    ↪ capthick=3)
plt.xlim([0, 21])
plt.xlabel("Number of Gaussians", fontsize=29)
plt.ylabel("Validation loss", fontsize=29)
plt.xticks(num_mogs_made[::2], num_mogs_made[::2])
#plt.minorticks_off()
plt.tight_layout()
plt.savefig(saved + "num_mogs.pdf")
```



```
[9]: # pipeline = run_maf_pipeline(train=False, gen=False, test=False, use_mog=True,
    ↪ use_maf=False,
```

```
# num_mogs=[10], name_str="_rerun_10")[0]

pipeline = run_maf_pipeline(train=False, gen=False, test=False, use_mog=True,
    ↪ use_maf=True,
                                num_mogs=[10], name_str="_rerun_10_2")[0]
```

100%|

| 1/1 [00:00<00:00, 735.97it/s]

```
[10]: # if the classifier training data rescaling and ref/generated reference
    ↪ rescaling do not match we get a better result?
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
    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
0x7f2ff06a8eb0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff06a8f10>!
100%|
          | 10/10 [00:01<00:00, 7.46it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2ff06a8c10>!
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 0x7f2ff06a8f70>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2ff06a8f40>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f300b90ba60>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f300b90bdc0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff06900d0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2ff0690160>!
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
0x7f2ff0690430>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff02c3f70>!
100%|

```

| 10/10 [00:01<00:00, 7.60it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2ff0328be0>!
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 0x7f2ff02e61c0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2ff02e6730>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2ff02e6a00>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff02e6ac0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff02e6b80>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2ff02e6c10>!
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
0x7f2ff031fee0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff01768e0>!
100%|

| 10/10 [00:01<00:00, 7.59it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2ff0176c10>!
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 0x7f2ff018cb50>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2ff018cf70>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2ff018e3d0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff018e490>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff018e550>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at

```

0x7f2ff018e5e0>!
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
0x7f2ff0690430>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff0116100>!
100%|
| 10/10 [00:01<00:00, 7.62it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2ff01160a0>!
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 0x7f2cb1424550>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2cb1424ac0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2cb1424dc0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb1424e80>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb1424f40>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2cb1424fd0>!
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
0x7f2ff031fee0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2cb69ecd90>!
100%|
| 10/10 [00:01<00:00, 7.58it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2cb3e12ac0>!
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 0x7f2cb6a09460>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock

```



```

object at 0x7f2cb6a095b0>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2cb6a09880>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb6a09940>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb6a09a00>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2cb6a09a90>!
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
0x7f2ff0690430>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2cb27f1790>!
100%|
          | 10/10 [00:01<00:00, 7.59it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2cb27f1ac0>!
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 0x7f2cafedb9d0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2cafedbf40>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2cb2800e80>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb2800eb0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2cb2800e50>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2cb2800d60>!
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
0x7f2ff031fee0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2fedf89d00>!
100%|
          | 10/10 [00:01<00:00, 7.59it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2fedf98dc0>!
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 0x7f2fedfa2400>!
 WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2fedfa29a0>!
 WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2fedfa2c70>!
 WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2fedfa2d30>!
 WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2fedfa2df0>!
 WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2fedfa2e80>!
 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 0x7f2ff01d2b20>!
 WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object at 0x7f2fededf040>!
 100%|

| 10/10 [00:01<00:00, 7.60it/s]

WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f2cb281bfd0>!
 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 0x7f2feddfdc70>!
 WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2feddfdfd0>!
 WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2fede004f0>!
 WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2fede005b0>!
 WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2fede00670>!
 WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2fede00700>!
 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

```

0x7f2ff031fee0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2fedd2c730>!
100%|
          | 10/10 [00:01<00:00, 7.60it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2fedd2c850>!
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 0x7f2fedce34c0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2fedce3a60>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2fedce3d30>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2fedce3df0>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2fedce3eb0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2fedce3f40>!
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
0x7f300bb14370>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff07560d0>!
100%|
          | 10/10 [00:01<00:00, 7.65it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2fede6e730>!
WARNING:root:Generated data check...
WARNING:root:nan ERROR -> found 18 invalid values that will be removed
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 ERROR -> found 18 invalid values that
will be removed
WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object
at 0x7f2fede6e040>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2fede6e160>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2fedeabee0>!

```

```

WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2fedeabf10>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2fedeab730>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2fedeab820>!
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
0x7f2cafedbdf0>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2cb3e12dc0>!
100%|
          | 10/10 [00:01<00:00, 7.64it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2cb3e121c0>!
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 0x7f2ff0082160>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2ff0082790>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at
0x7f2ff018e1c0>!
WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff018eb80>!
WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock
object at 0x7f2ff018efa0>!
WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2ff018ef70>!
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
0x7f300bb14370>!
WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2ff06a8c10>!
100%|
          | 10/10 [00:01<00:00, 7.63it/s]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2ff06a8f10>!
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 0x7f2ff766b6a0>!
 WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2ff766be20>!
 WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2ff7662c10>!
 WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2ff7662940>!
 WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2ff76628e0>!
 WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2ff7662310>!
 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 0x7f2cafedbbdf0>!
 WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object at 0x7f2ff768d220>!
 100%|

| 10/10 [00:01<00:00, 7.60it/s]

WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f2ff7578e50>!
 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 0x7f2ff7537970>!
 WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2ff7537880>!
 WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2ff77747f0>!
 WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2ff7774d60>!
 WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2ff7774f10>!
 WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2ff7774e80>!
 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 0x7f300bb14370>!
 WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object at 0x7f2cb2805c10>!

100%|

| 10/10 [00:01<00:00, 7.59it/s]

WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f2cb2805f40>!

WARNING:root:Generated data check...

WARNING:root:nan ERROR -> found 18 invalid values that will be removed

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 ERROR -> found 18 invalid values that will be removed

WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object at 0x7f2ff01cbe20>!

WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2ff01f4280>!

WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2cb27ed490>!

WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2cb27ed4c0>!

WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2cb27ed340>!

WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2cb27ed2b0>!

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 0x7f2cafedbdf0>!

WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object at 0x7f2ff033de50>!

100%|

| 10/10 [00:01<00:00, 7.61it/s]

WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock object at 0x7f2ff033df70>!

WARNING:root:Generated data check...

WARNING:root:nan ERROR -> found 18 invalid values that will be removed

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 ERROR -> found 18 invalid values that will be removed

WARNING:root:fitting #3: <ml_hep_sim.pipeline.blocks.DatasetBuilderBlock object at 0x7f2cafeb4850>!

WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock object at 0x7f2cafeb4dc0>!

WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ModelLoaderBlock object at 0x7f2cafec90d0>!

WARNING:root:fitting #6: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2cafec9190>!

WARNING:root:fitting #7: <ml_hep_sim.pipeline.blocks.ClassifierRunnerBlock object at 0x7f2cafec9250>!
 WARNING:root:fitting #8: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at 0x7f2cafec92e0>!

```
[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]
```

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

```
[31]: 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)

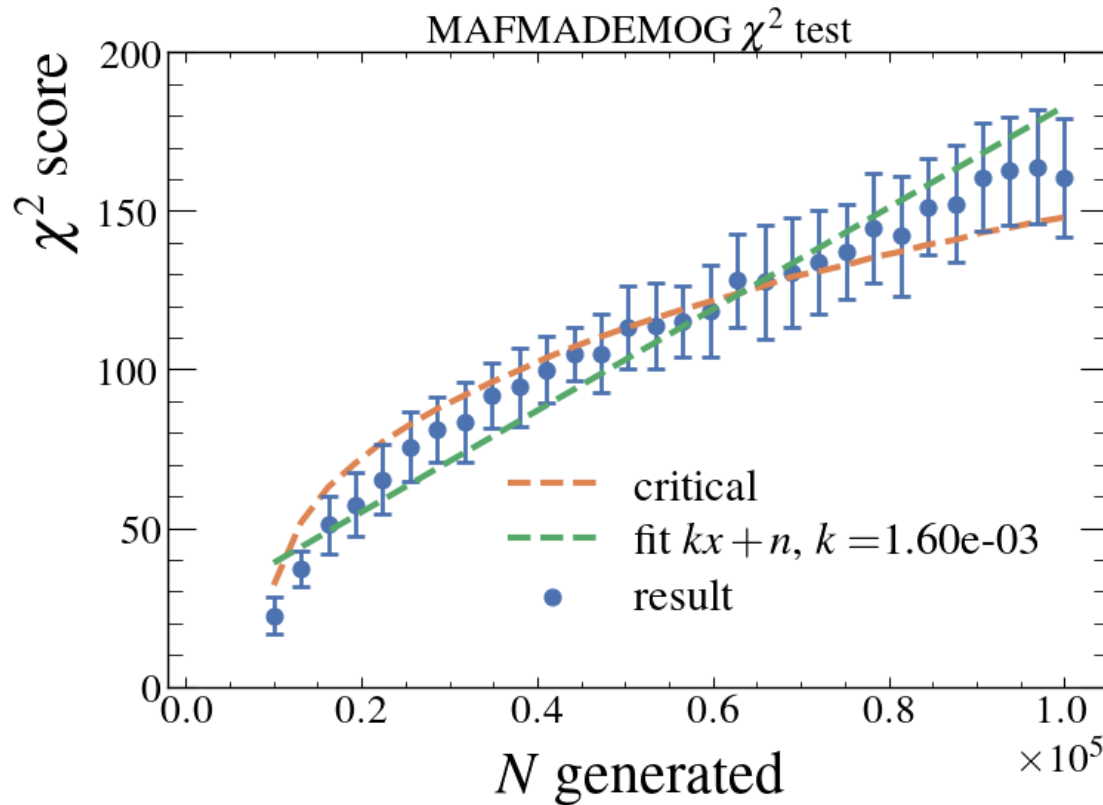
plt.errorbar(N_range, chi2_m.mean(axis=0), yerr=chi2_m.std(axis=0), capsize=4,
             ls="none", lw=2, capthick=2)

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("MAFMADEMOG $\chi^2$ test")

plt.tight_layout()
plt.savefig(saved + "mafmademog_chi2_scaling.pdf")
```



```
[32]: np.diag(pcov)**(1/2)
```

```
[32]: array([7.36595863e-05, 3.41345913e+00])
```

```
[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)

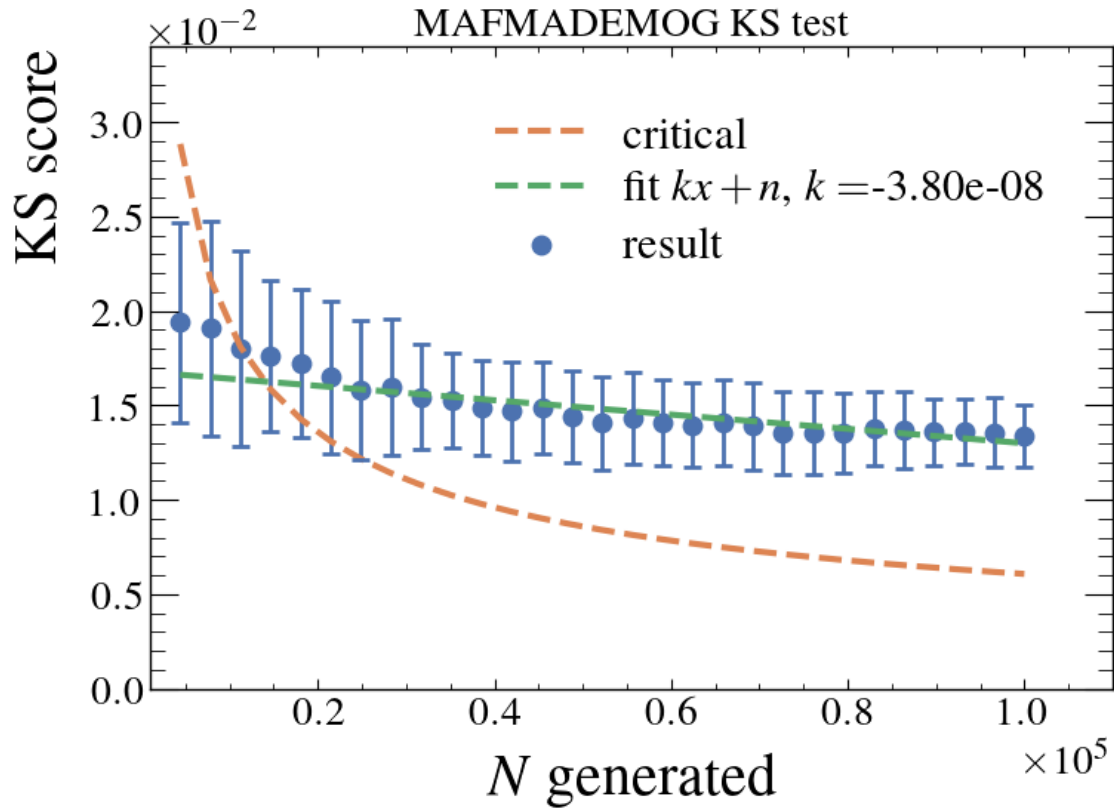
def func(x, k, n):
    return k * x + n

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("MAFMADEMOG KS test")
plt.tight_layout()
plt.savefig(saved + "mafmademog_ks_scaling.pdf")$ 
```

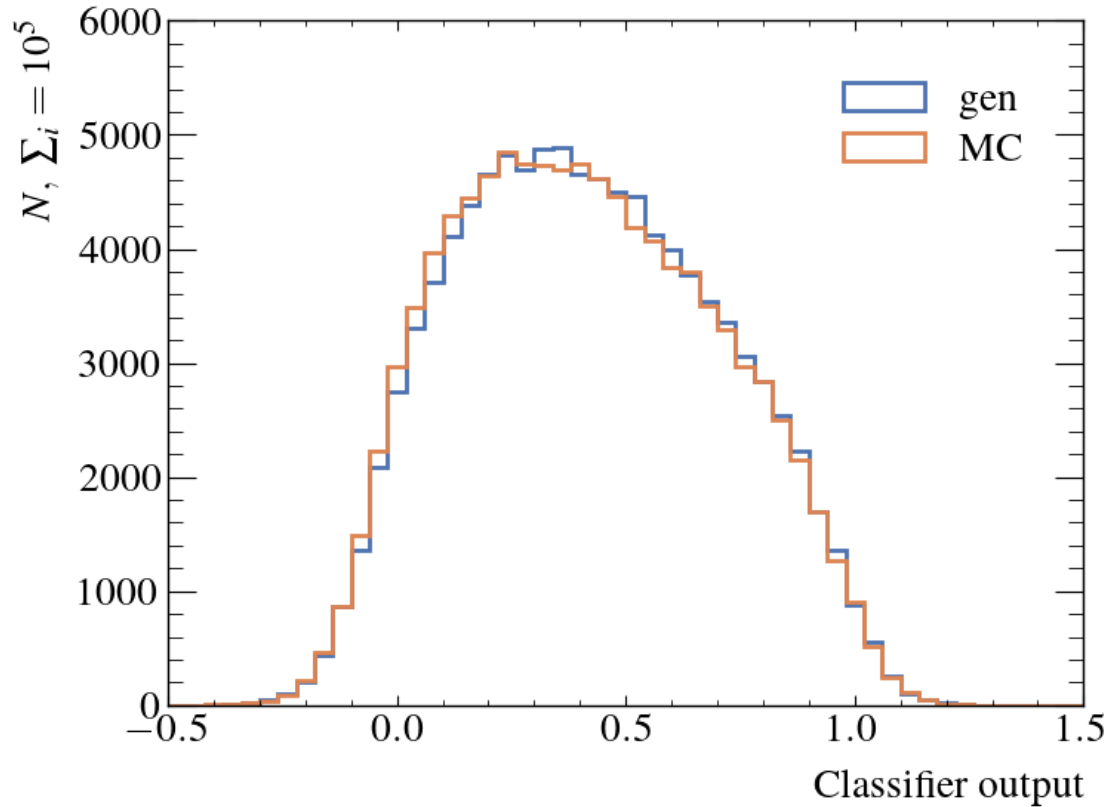


```
[30]: np.diag(pcov)**(1/2)
```

```
[30]: array([4.14325967e-09, 2.95579529e-04])
```

```
[15]: 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("$N, \backslash> \$ \$\sum_i=10^5$")
      plt.legend(["gen", "MC"])
      plt.tight_layout()
      plt.savefig(saved + "mafmademog_class_out.pdf")
```



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

x_ConfigBuilderBlock, _, _, x_ModelTrainerBlock = pipeline.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 0x7f2fed905820>!

```

WARNING:root:fitting #1: <ml_hep_sim.pipeline.blocks.DataGeneratorBlock object
at 0x7f2fed905d00>!
100%|
          | 10/10 [00:11<00:00, 1.13s/it]
WARNING:root:fitting #2: <ml_hep_sim.pipeline.blocks.GeneratedDataVerifierBlock
object at 0x7f2fed905c10>!
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
/home/jan/FMF/masters/ml_hep_sim/data_utils/dataset_utils.py:59: RuntimeWarning:
overflow encountered in exp
  return 1 / (1 + np.exp(-x))
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 0x7f2fed9055b0>!
WARNING:root:fitting #4: <ml_hep_sim.pipeline.blocks.ReferenceDataLoaderBlock
object at 0x7f2fed905880>!
WARNING:root:fitting #5: <ml_hep_sim.pipeline.blocks.ScalingTestBlock object at
0x7f2fed905700>!

```

```
[16]: <ml_hep_sim.pipeline.pipes.Pipeline at 0x7f2fed9051f0>
```

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

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

```
[19]: 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)))
```

```
[20]: 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()
```

```
[21]: 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)

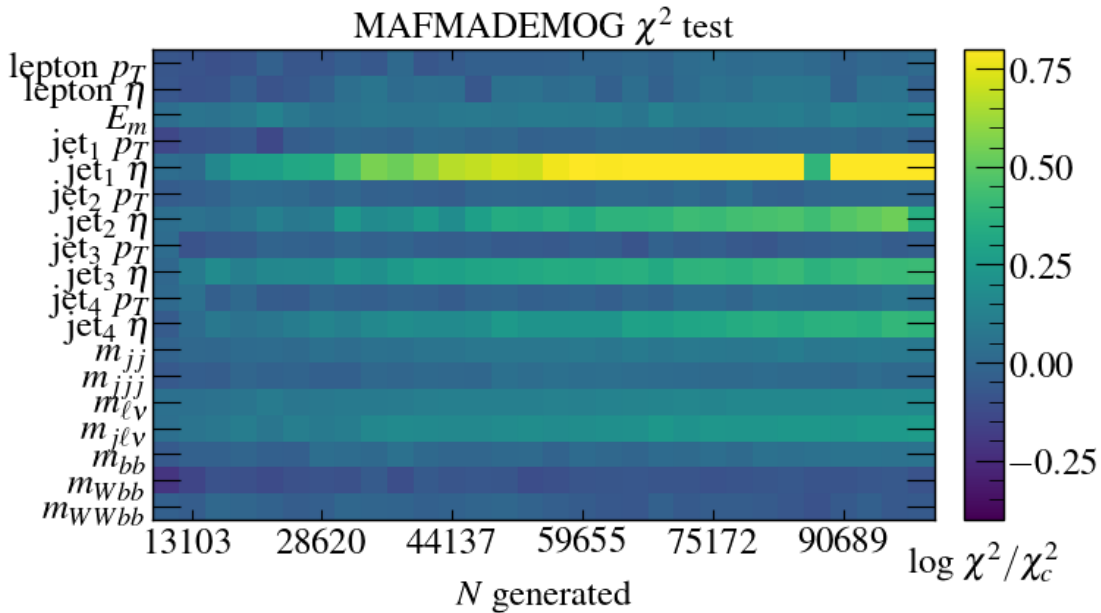
ax.minorticks_off()

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  $\chi^2/\chi_c^2$ ', loc="center")
ax.set_title("MAFMADEMOG  $\chi^2$  test")

plt.tight_layout()
plt.savefig(saved + "imshow_mafmademog_chi2.pdf")
```



```

[22]: 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()

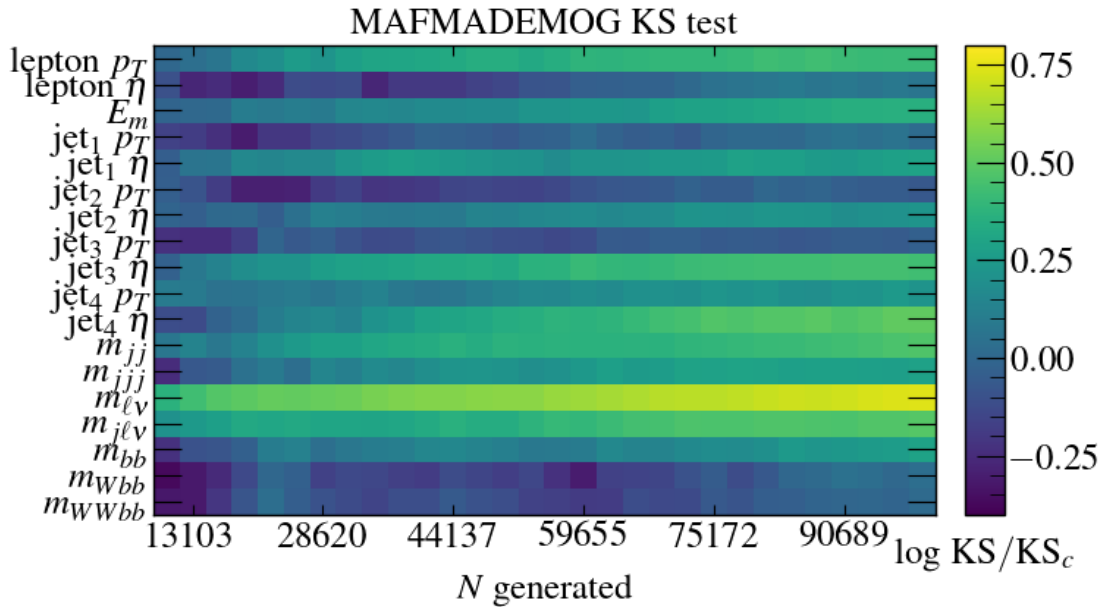
ax.set_xticks(np.arange(0, 30, 1)[1::5])
ax.set_xticklabels(N_range[1::5])

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("MAFMADEMOG KS test")

plt.tight_layout()
plt.savefig(saved + "imshow_mafmademog_ks.pdf")

```



```
[23]: pipeline_mademog = run_maf_pipeline(train=False, gen=False, test=False,
    ↪ use_mog=True, use_maf=False,
    num_mogs=[10], name_str="_rerun_10")[0]
```

100%|

| 1/1 [00:00<00:00, 1458.89it/s]

```
[24]: pipeline_mafmademog = run_maf_pipeline(train=False, gen=False, test=False,
    ↪ use_mog=True, use_maf=True,
    num_mogs=[10], name_str="_rerun_10_2")[0]
```

100%|

| 1/1 [00:00<00:00, 4359.98it/s]

```
[25]: pipeline_maf = run_maf_pipeline(train=False, gen=False, test=False,
    ↪ use_mog=False, use_maf=True,
    num_mogs=[10], name_str="_MAF_only_10")
```

```
[26]: pipelines = [pipeline_maf, pipeline_mafmademog, pipeline_mademog]
```

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

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

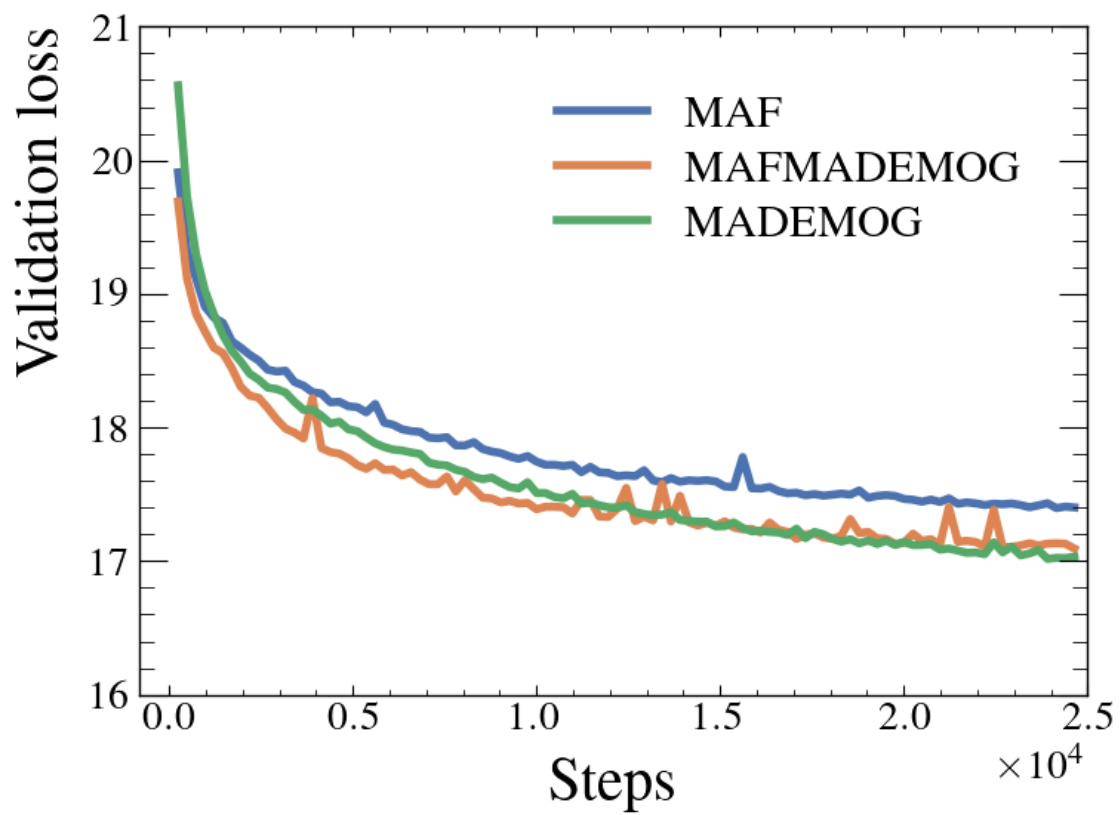
    val_loss = metrics[-1]["val_loss"]
    step = metrics[-2]["step"]
    t = metrics[0]["timestamp"].to_numpy()
    times.append(t[-1] - t[0])
    val_losses.append(val_loss)
    steps.append(step)
```

```
[28]: 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.legend(["MAF", "MAFMADEMOG", "MADEMOG"], 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_mades.pdf")
```



[]: