Showerflow – variations on a theme
Part II

2024/08/29

Previously; Variations tested

Decided to train a bunch of models, and see what timings look like in more realistic settings.

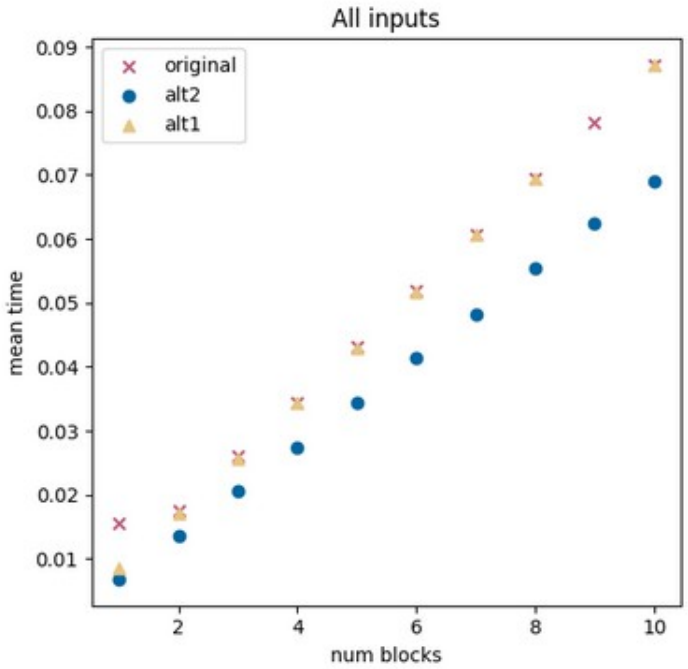
Original;
Num blocks times
- affine then permute
- affine then permute
- affine then permute
- permute then spline
- affine then permute
- affine then permute
- affine then permute

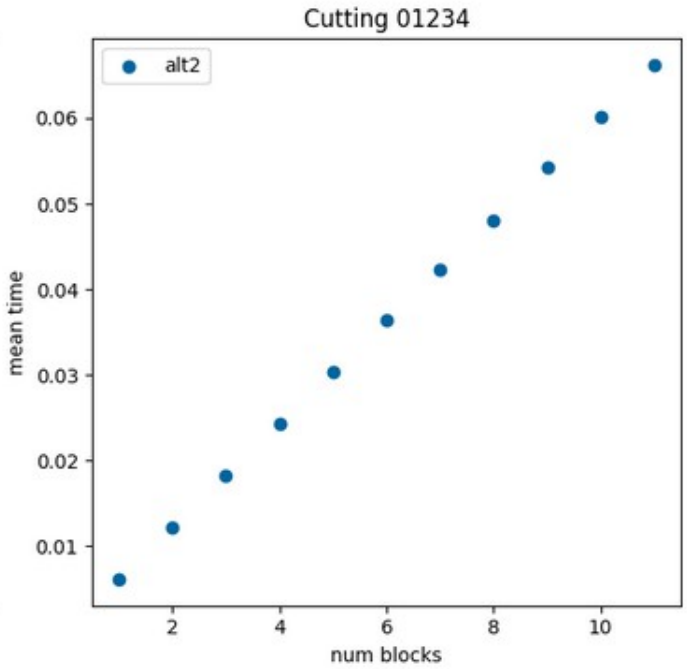
alt1;
Num blocks times
- affine then permute
- affine then permute
- affine then permute
- spline then permute
- affine then permute
- affine then permute
- affine then permute

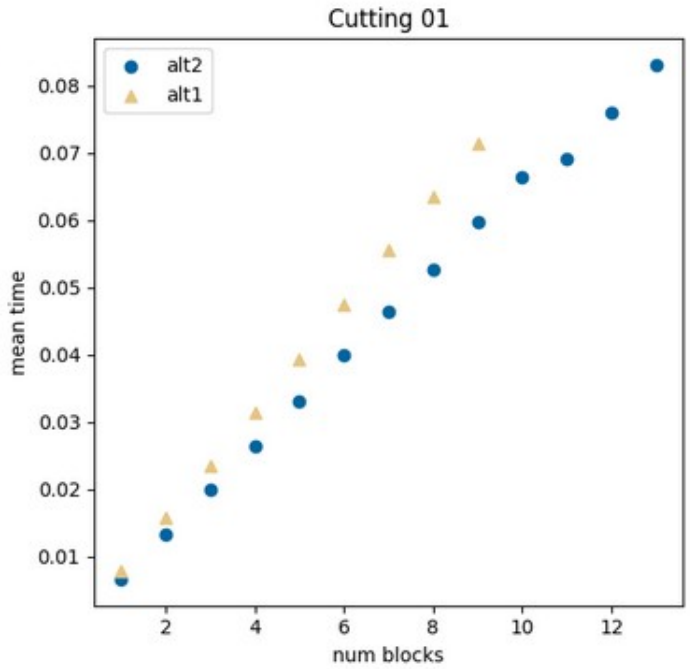
alt2;
Num blocks times
- affine then permute
- affine then permute
- affine then permute
- affine then permute
- affine then permute
- affine then permute
- affine then permute

Previously; Times for variations

Took as many timings as I could in 48 hours. Tested at a range of incident energies.



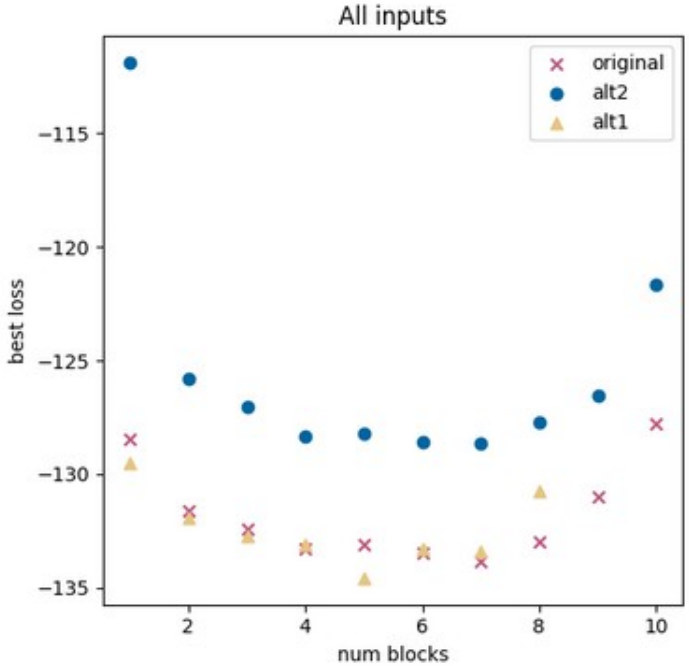


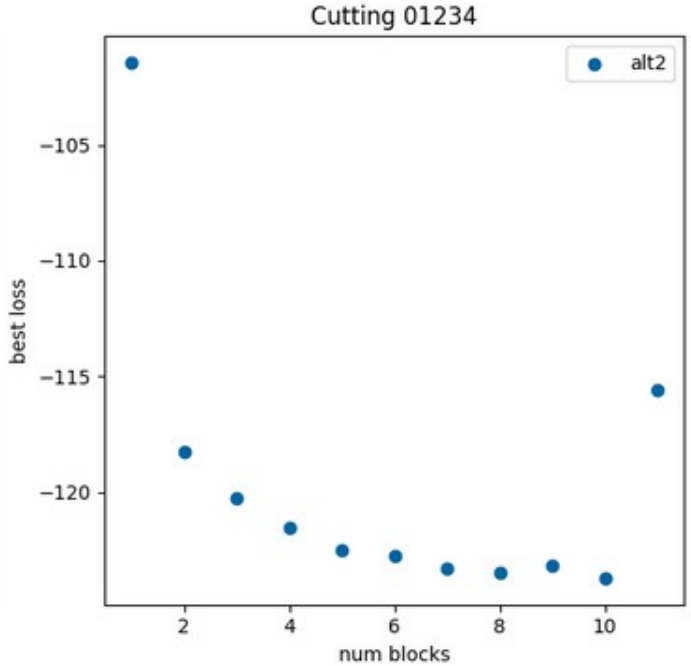


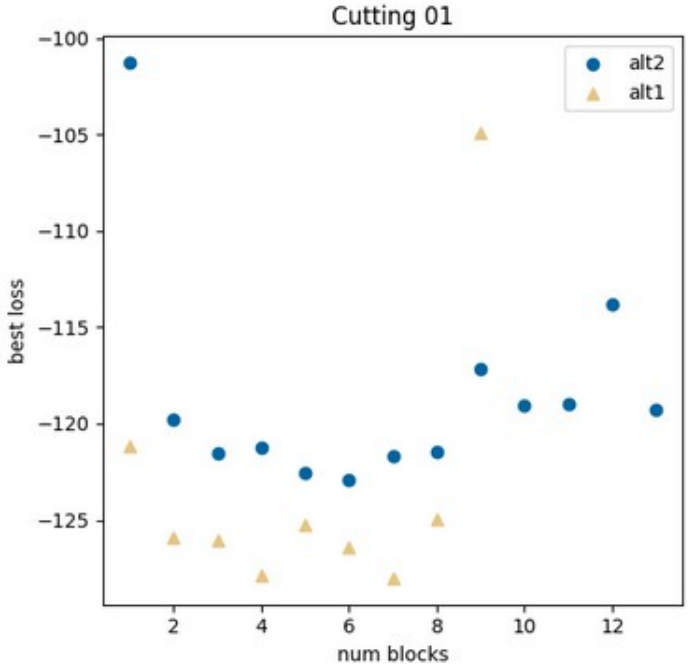
Previously; Negative log prob

Can use this relative to models with the same number of inputs.

I need to make some physics stats to compare between models.

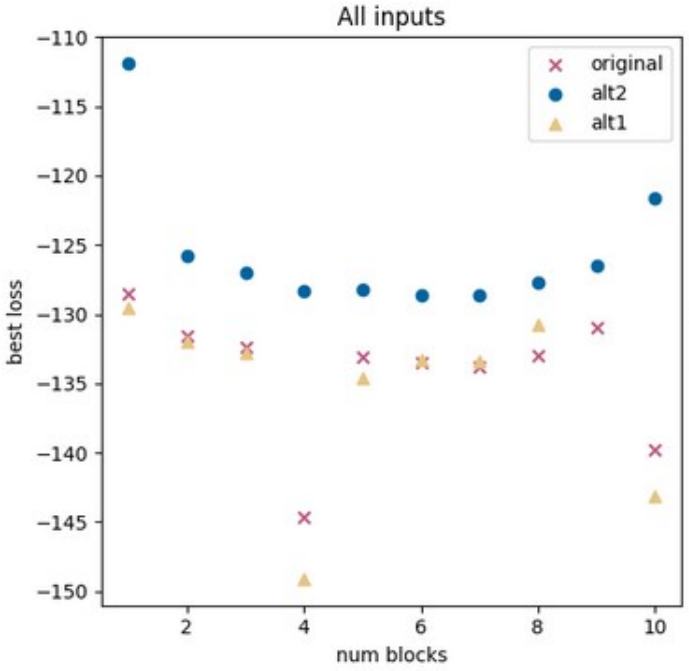


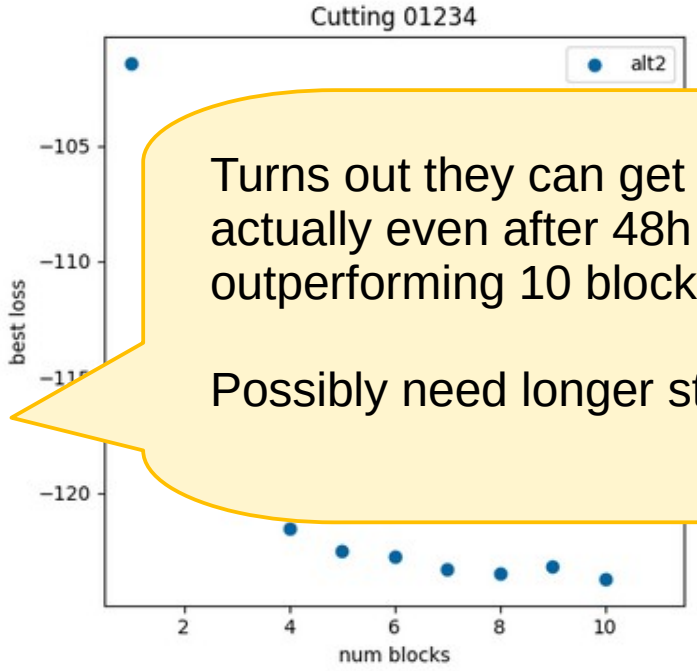


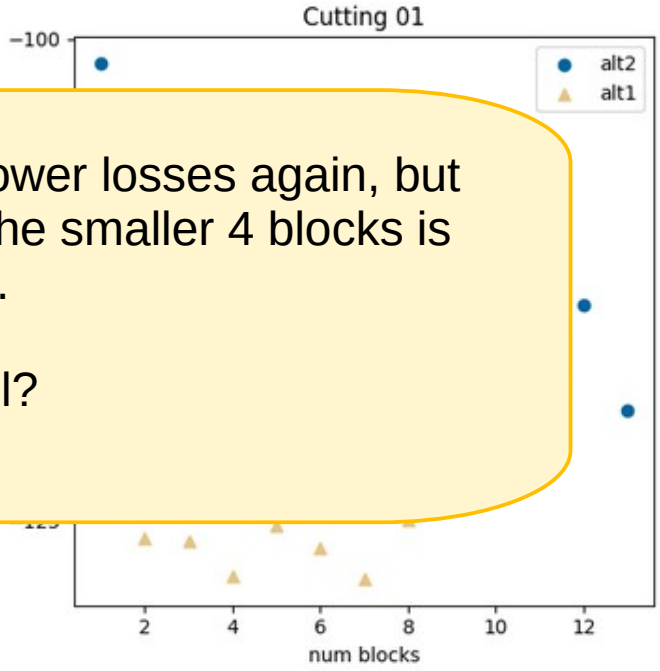


Updated; Negative log prob

Everyone wanted to know where the log prob plateaus, so I took alt1 and original at 4 blocks and 10 blocks and ran them for 48h each.







Turns out they can get lower losses again, but actually even after 48h the smaller 4 blocks is outperforming 10 blocks.

Possibly need longer still?

Negative log prob

```
ShowerFlow_alt2_nb4_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-128.3624267578125 5848
ShowerFlow_alt2_nb10_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-121.65310516357422 3709
ShowerFlow_original_nb10_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-139.7427490234375 54399
ShowerFlow_original_nb4_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-144.6334991455078 48041
ShowerFlow_alt1_nb10_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-143.1125061035156 60558
ShowerFlow_alt1_nb4_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-149.15553588867186 89737
```

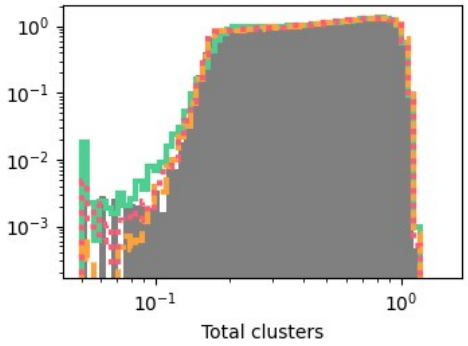
Negative log prob

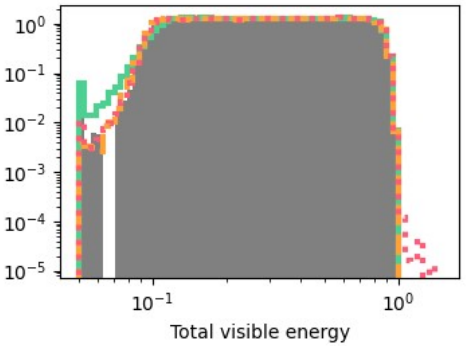
```
ShowerFlow_alt2_nb4_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-128.3624267578125 5848
ShowerFlow_alt2_nb10_inputs36893488147419103231_best_data.txt
Lowest loss         Epoch number
-121.65310516357422 3709
ShowerFlow_original_nb10_i
Lowest loss         Epoch n
-139.7427490234375 54399
ShowerFlow_original_nb4_in
Lowest loss         Epoch n
-144.6334991455078 48041
ShowerFlow_alt1_nb10_input
Lowest loss         Epoch n
-143.1125061035156 60558
ShowerFlow_alt1_nb4_inputs
Lowest loss         Epoch number
-149.15553588867186 89737
```

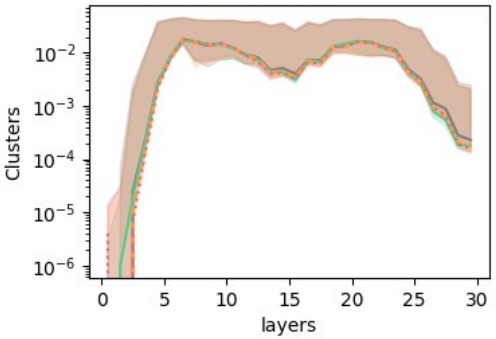
```
Aug 28 11:35 ShowerFlow_original_nb10_inputs36893488147419103231_best_data.txt
Aug 28 11:40 ShowerFlow_original_nb4_inputs36893488147419103231_best_data.txt
Aug 28 11:41 ShowerFlow_original_nb4_inputs36893488147419103231_latest.pth
Aug 28 11:41 ShowerFlow_original_nb10_inputs36893488147419103231_latest.pth
Aug 28 13:00 ShowerFlow_alt1_nb10_inputs36893488147419103231_best_data.txt
Aug 28 15:49 ShowerFlow_alt1_nb4_inputs36893488147419103231_best_data.txt
Aug 28 16:21 ShowerFlow_alt1_nb4_inputs36893488147419103231_latest.pth
Aug 28 16:22 ShowerFlow_alt1_nb10_inputs36893488147419103231_latest.pth
```

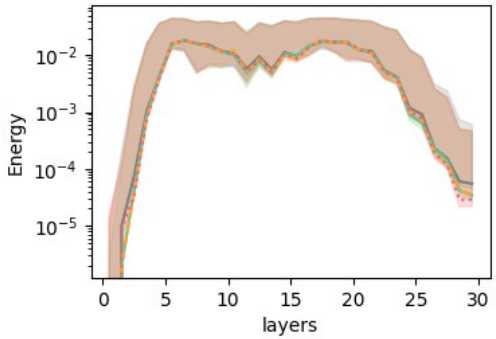
From this alt1 saturated but original may still be learning.

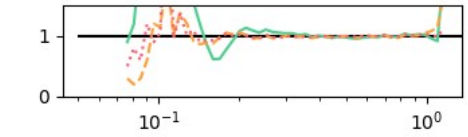
However, I'm not so sure the difference matters….

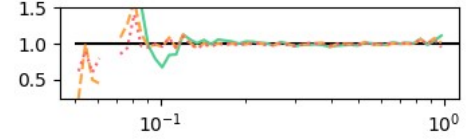


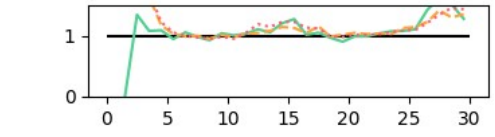


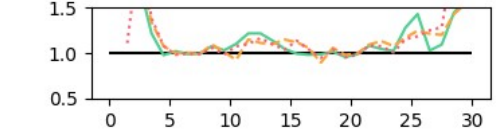


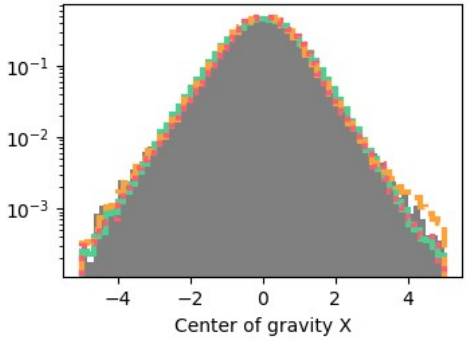


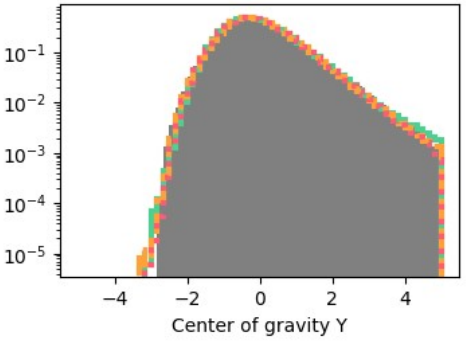


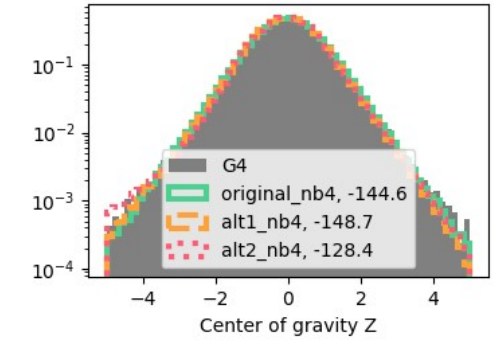


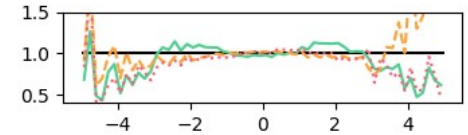


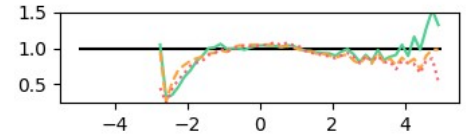


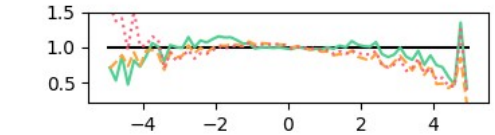






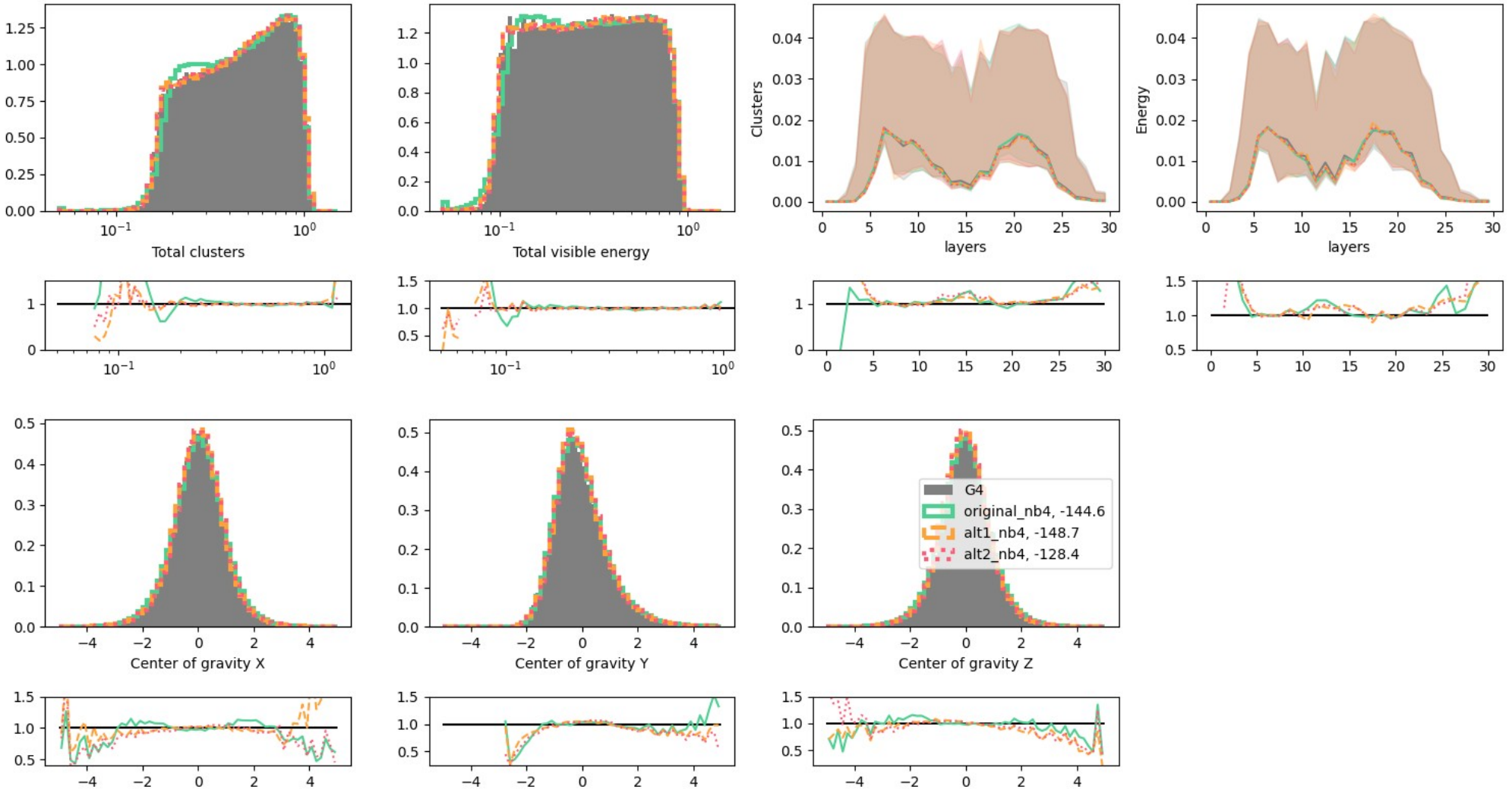




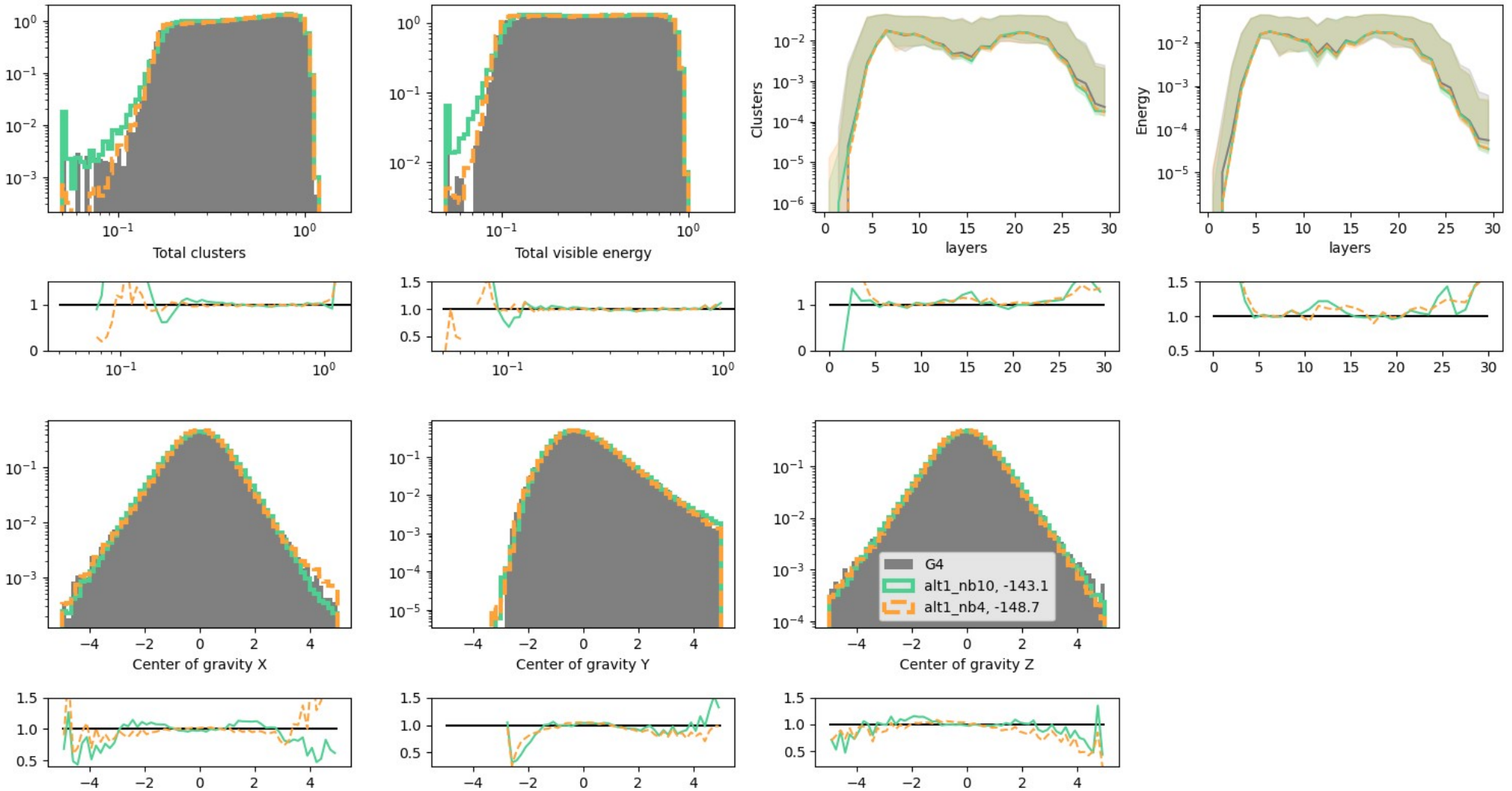


Actual behavior is very good for all choices.

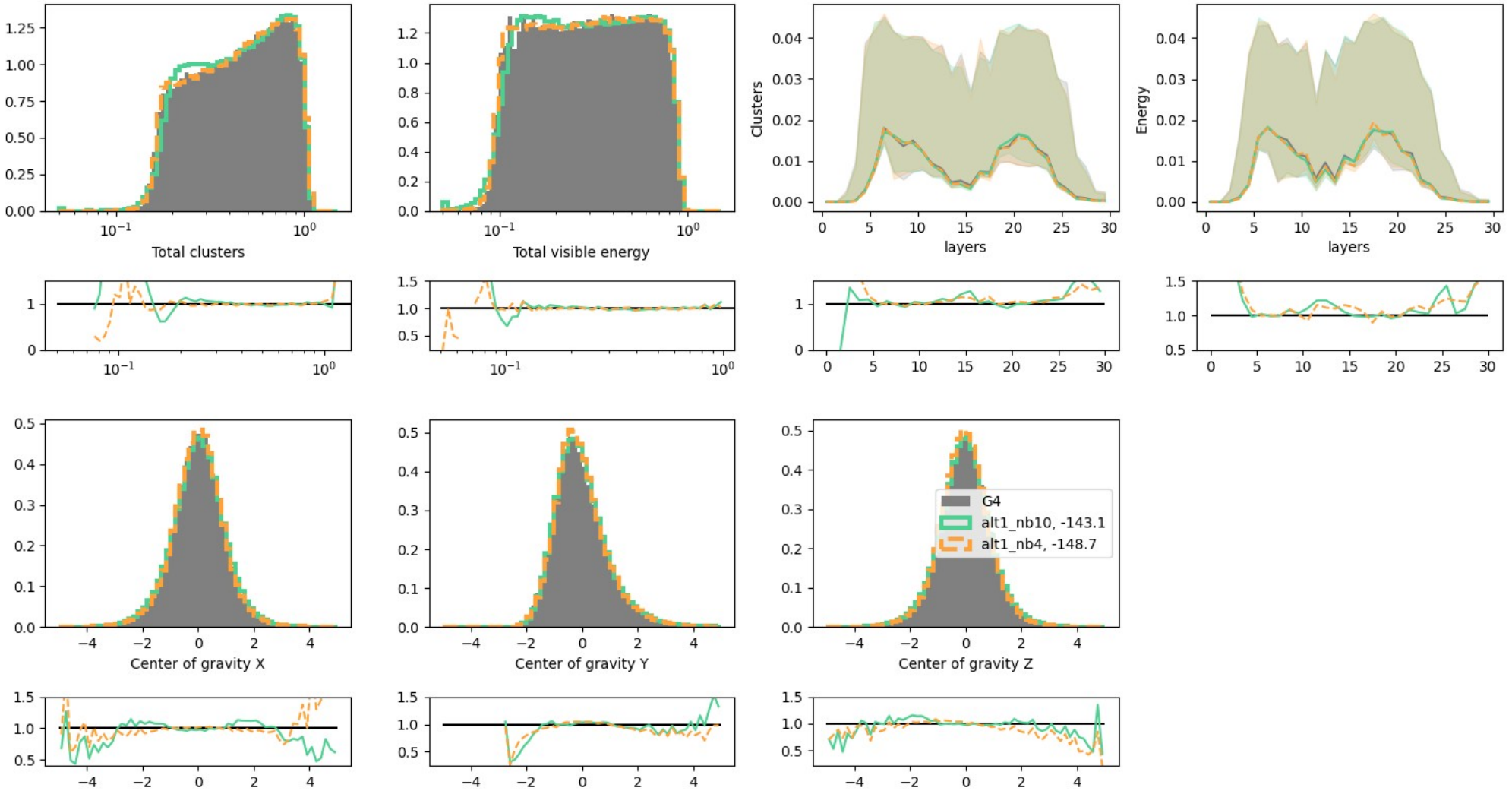
The distributions



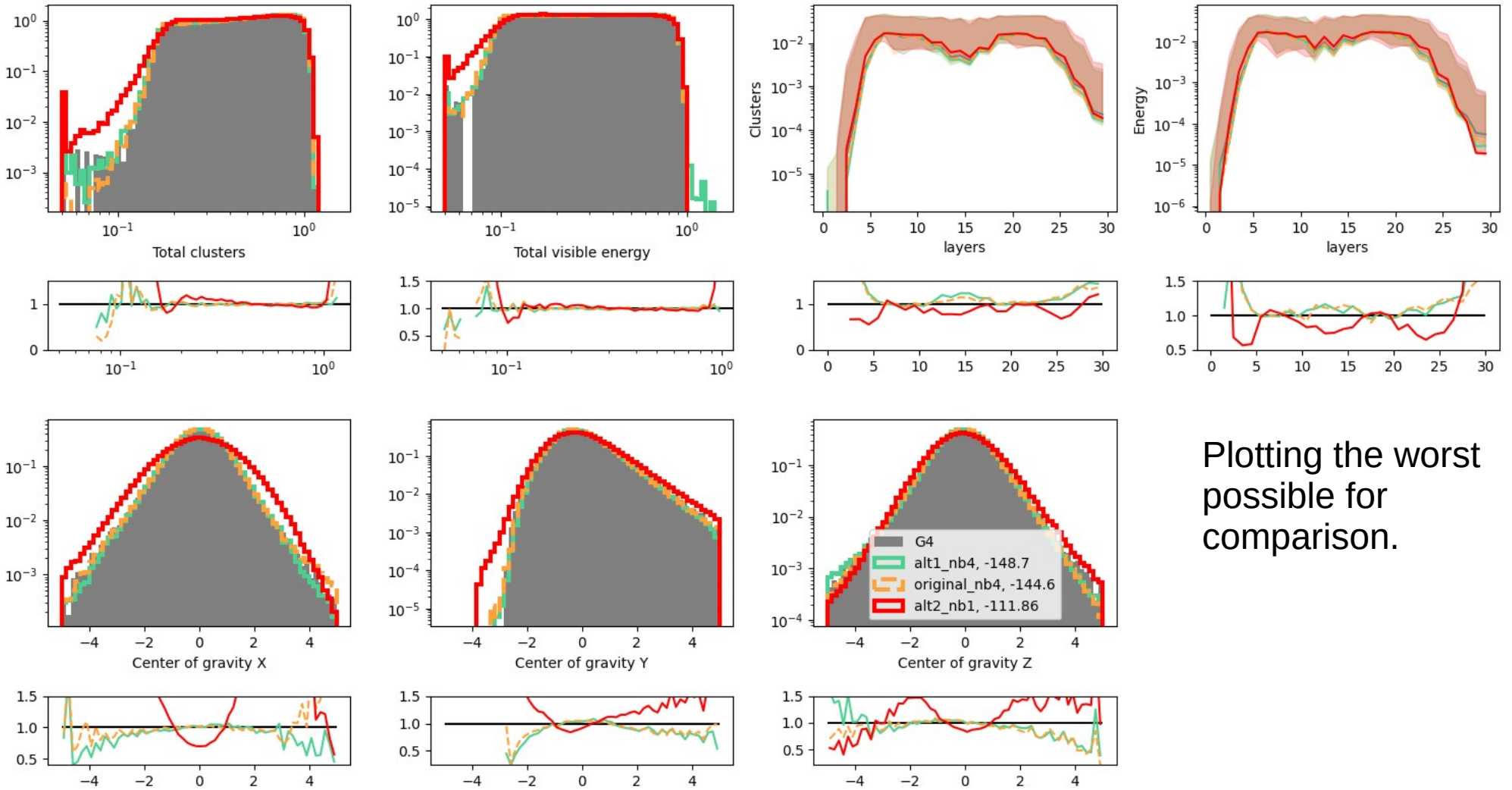
The distributions



The distributions

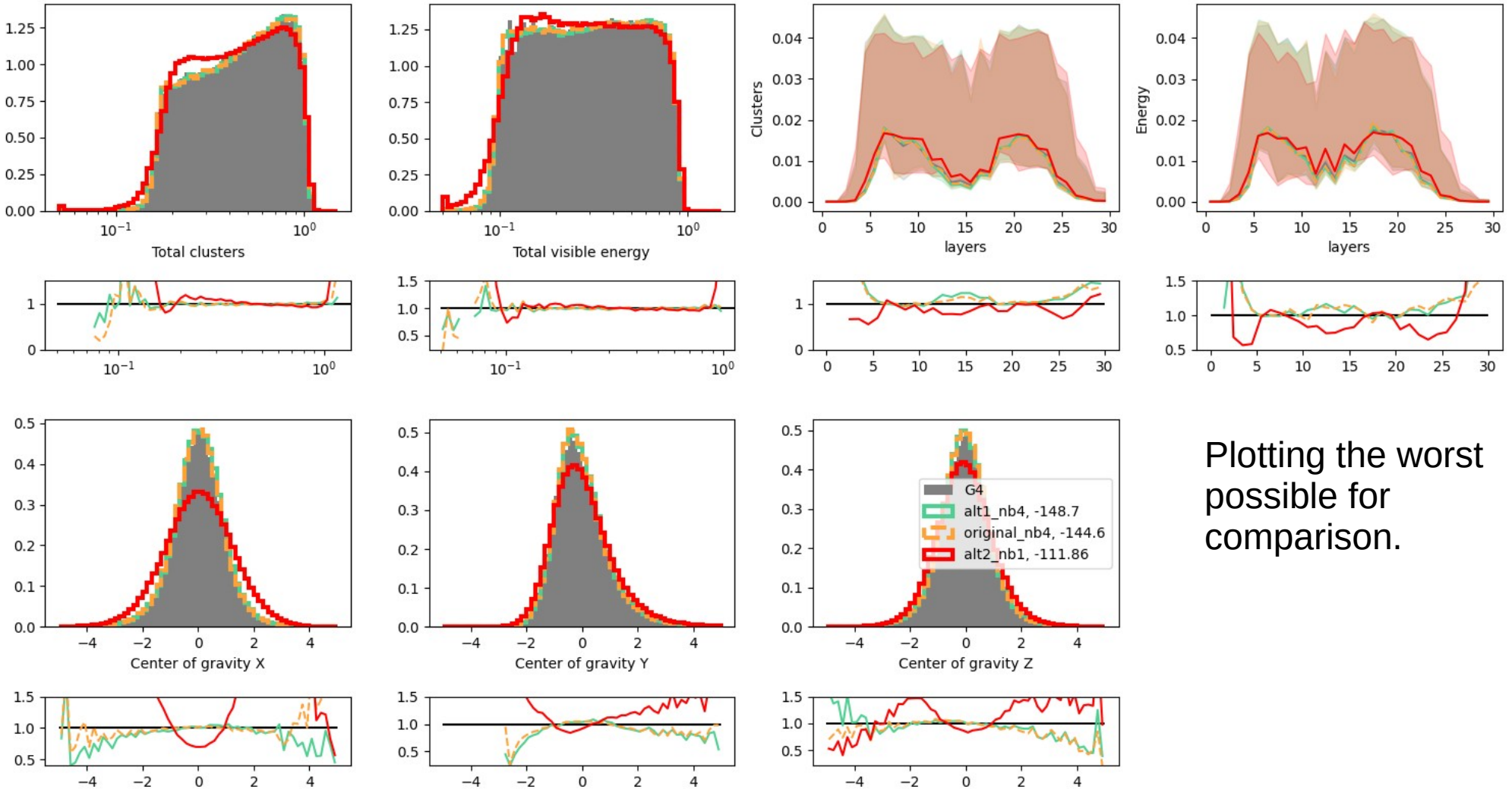


The distributions



Plotting the worst possible for comparison.

The distributions



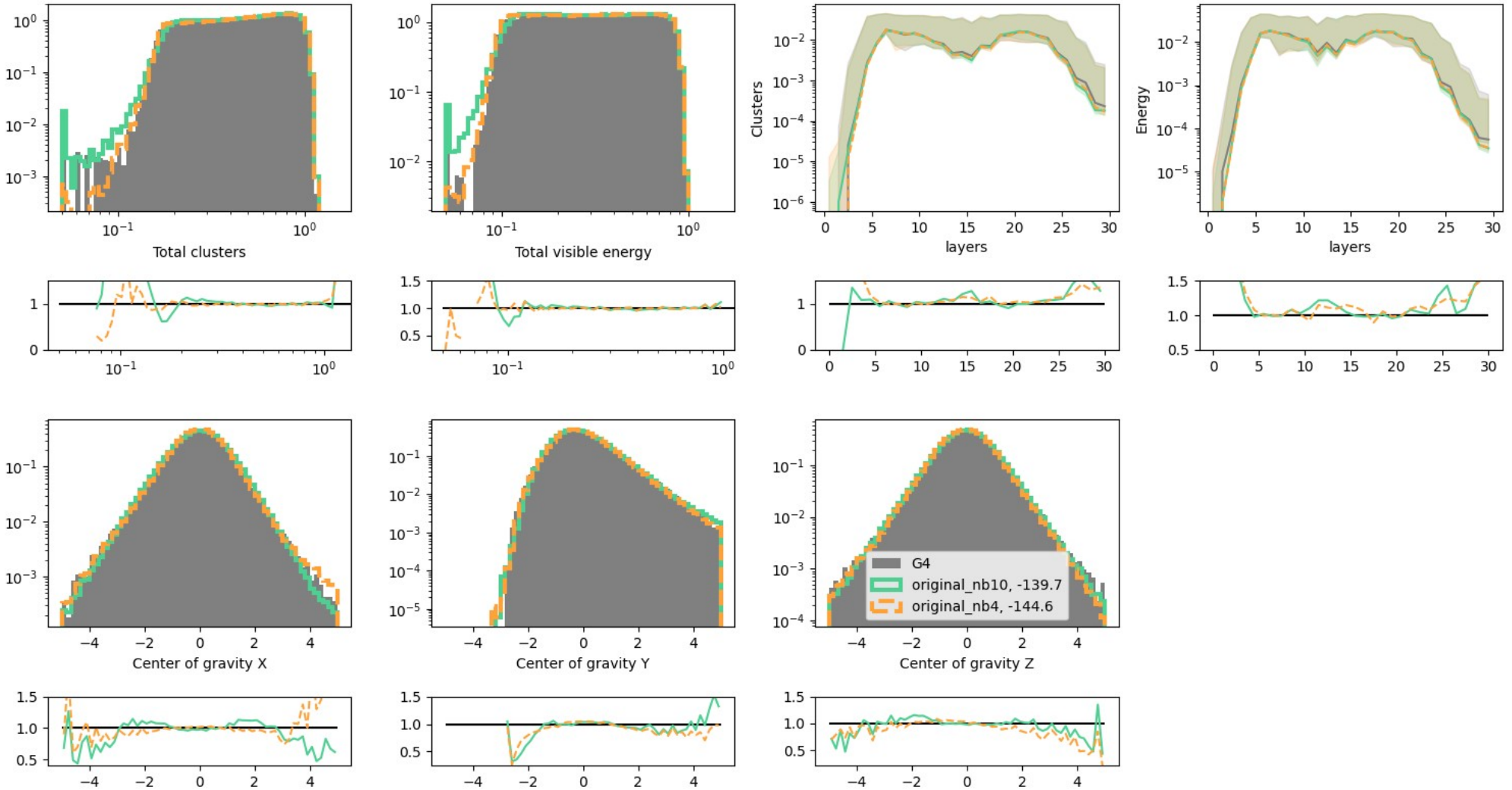
Plotting the worst possible for comparison.

The distributions

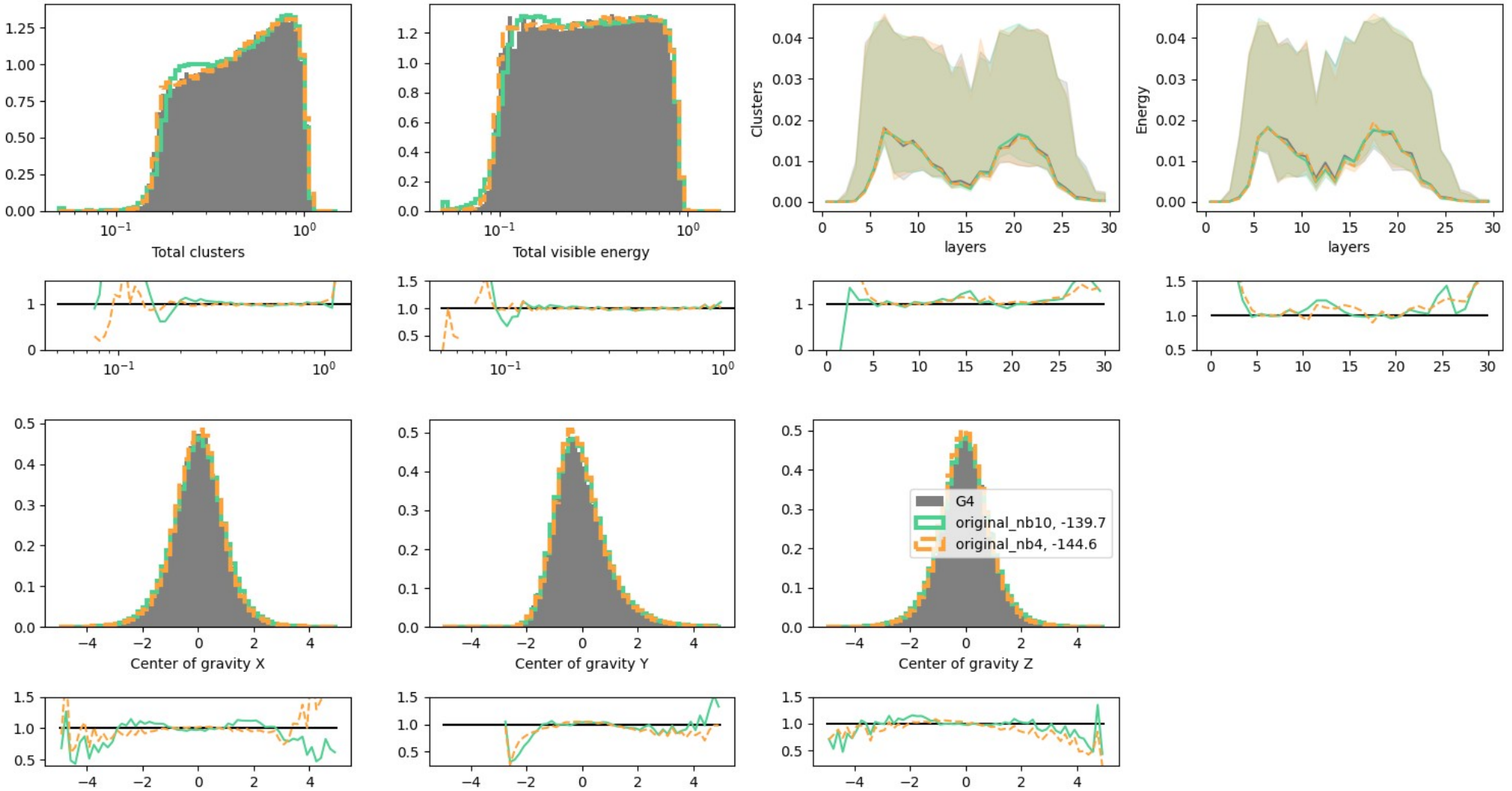
Further comparisons

- Correlations between variables matter too.
- Lots of dimensions, hard to make ratios. Gets messy to plot.
- Maybe at this point an NN like the one that Simon has investigated is the right tool for comparing these models?

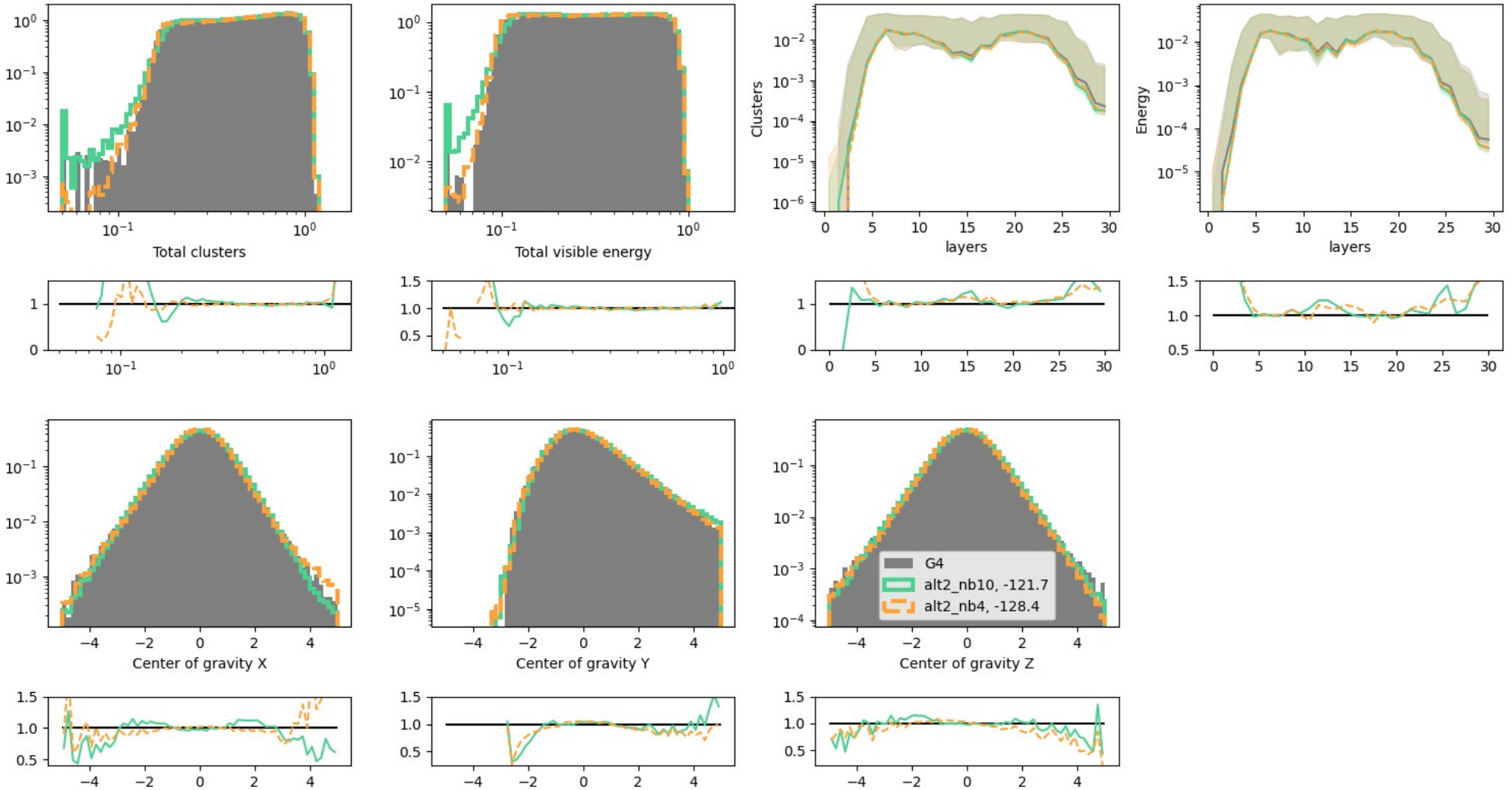
Backup



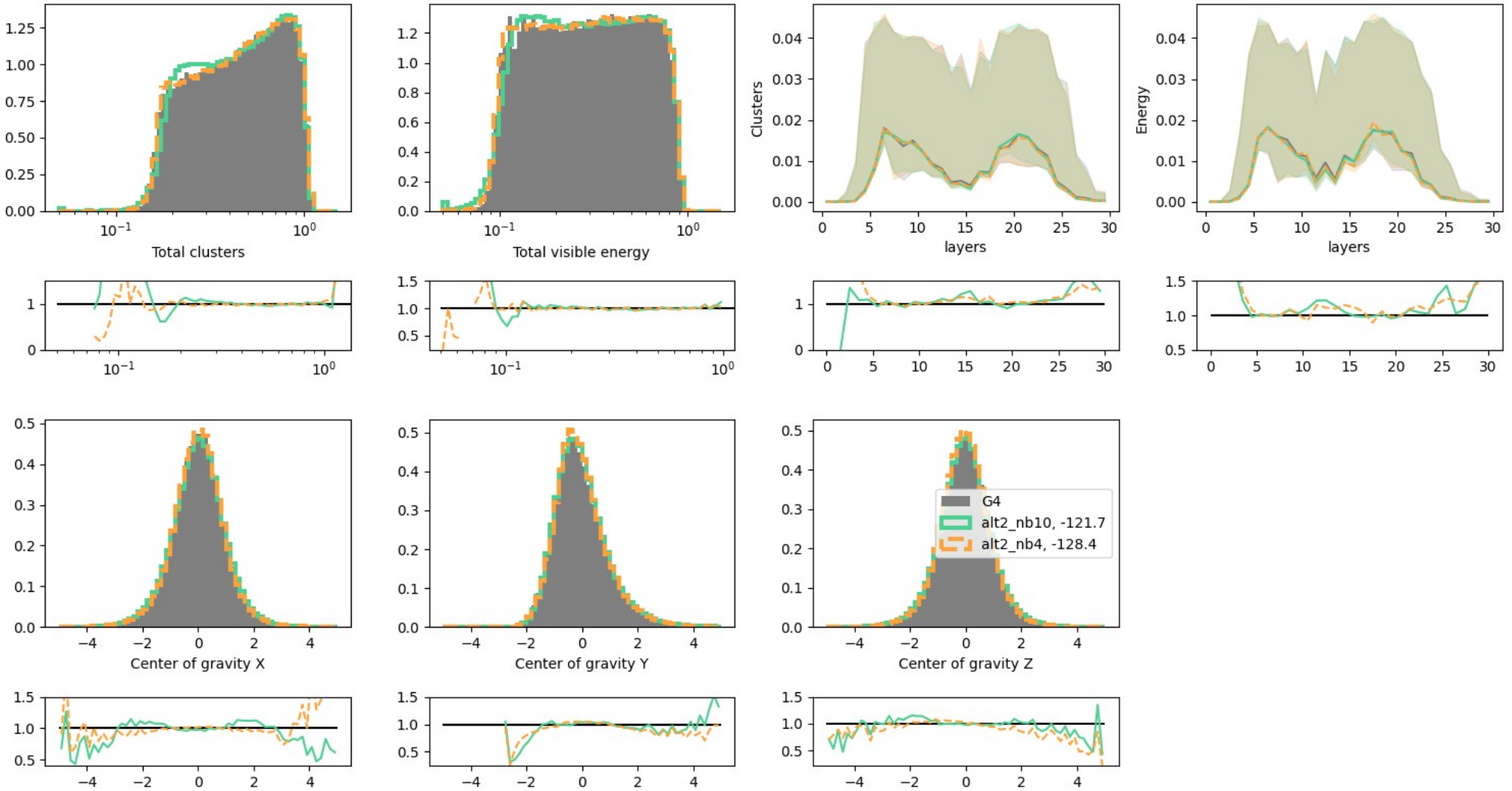
The distributions



The distributions



The distributions



The distributions

Previously; Times for variations

Still a little strangeness at incident energy = 0.1

But small enough that I won't investigate.