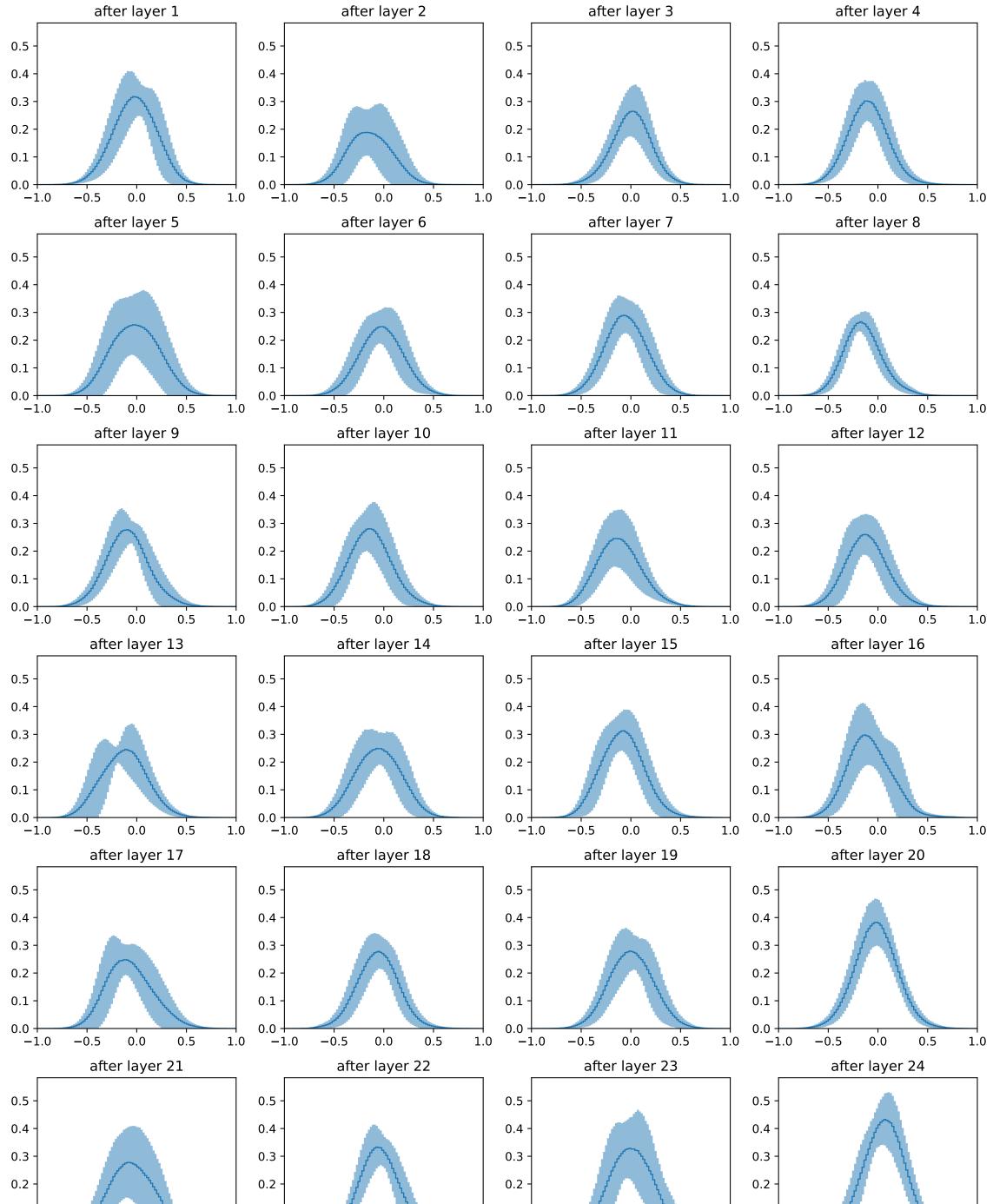
attention_logits histograms (randomtext_bert-large-uncased_params-trained_dense-off_layers-24) after layer 1 after layer 3 after layer 2 0.5 0.5



0.1

0.0

1.0

-1.0

-0.5

0.0

0.5

0.1

0.0

-1.0

-0.5

0.0

0.5

0.1

0.0

-1.0

-0.5

0.0

0.5

1.0

0.1

0.0

-1.0

-0.5

0.0

1.0

0.5

1.0