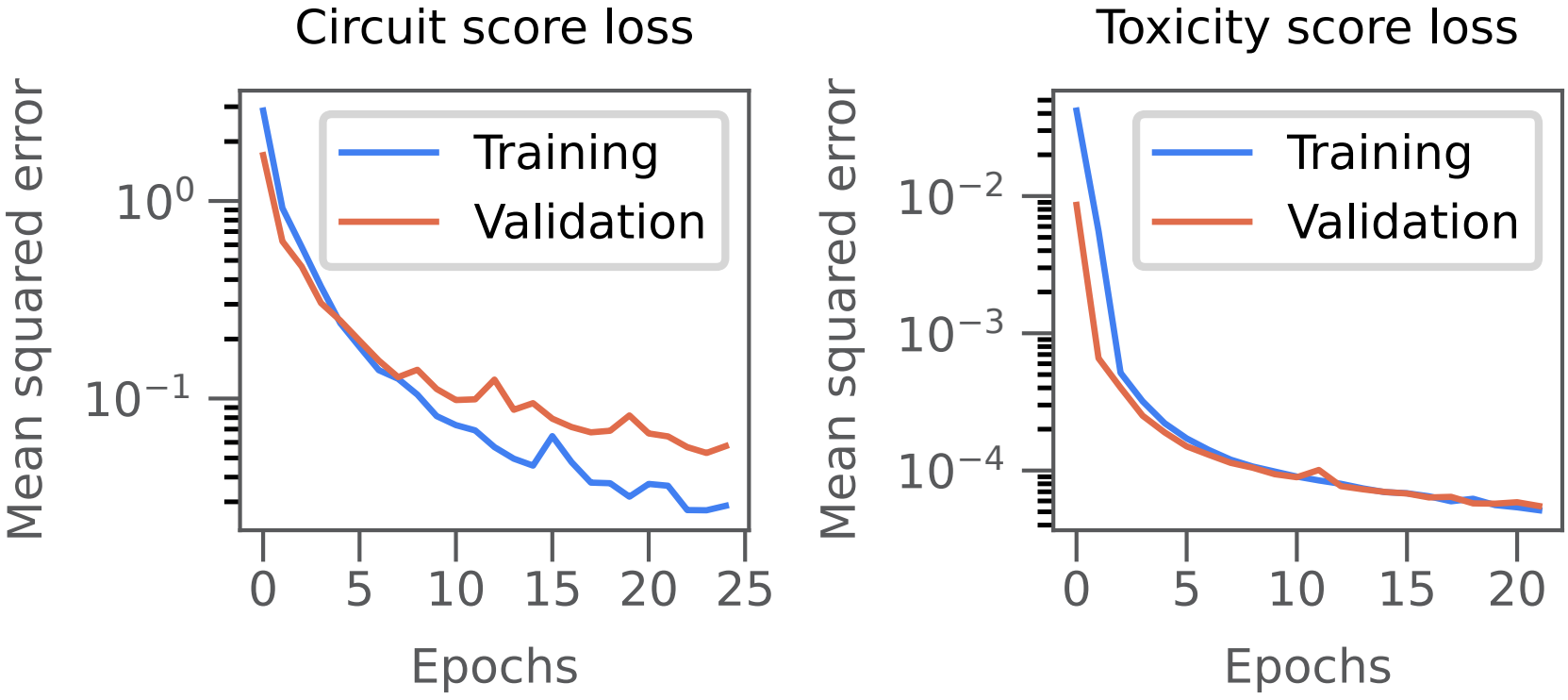
Circuit score loss
Mean squared error
10^0
10^{-1}
Training
Validation
0
5
10
15
20
25
Epochs
Toxicity score loss
Mean squared error
10^{-2}
10^{-3}
10^{-4}
Training
Validation
0
5
10
15
20
Epochs