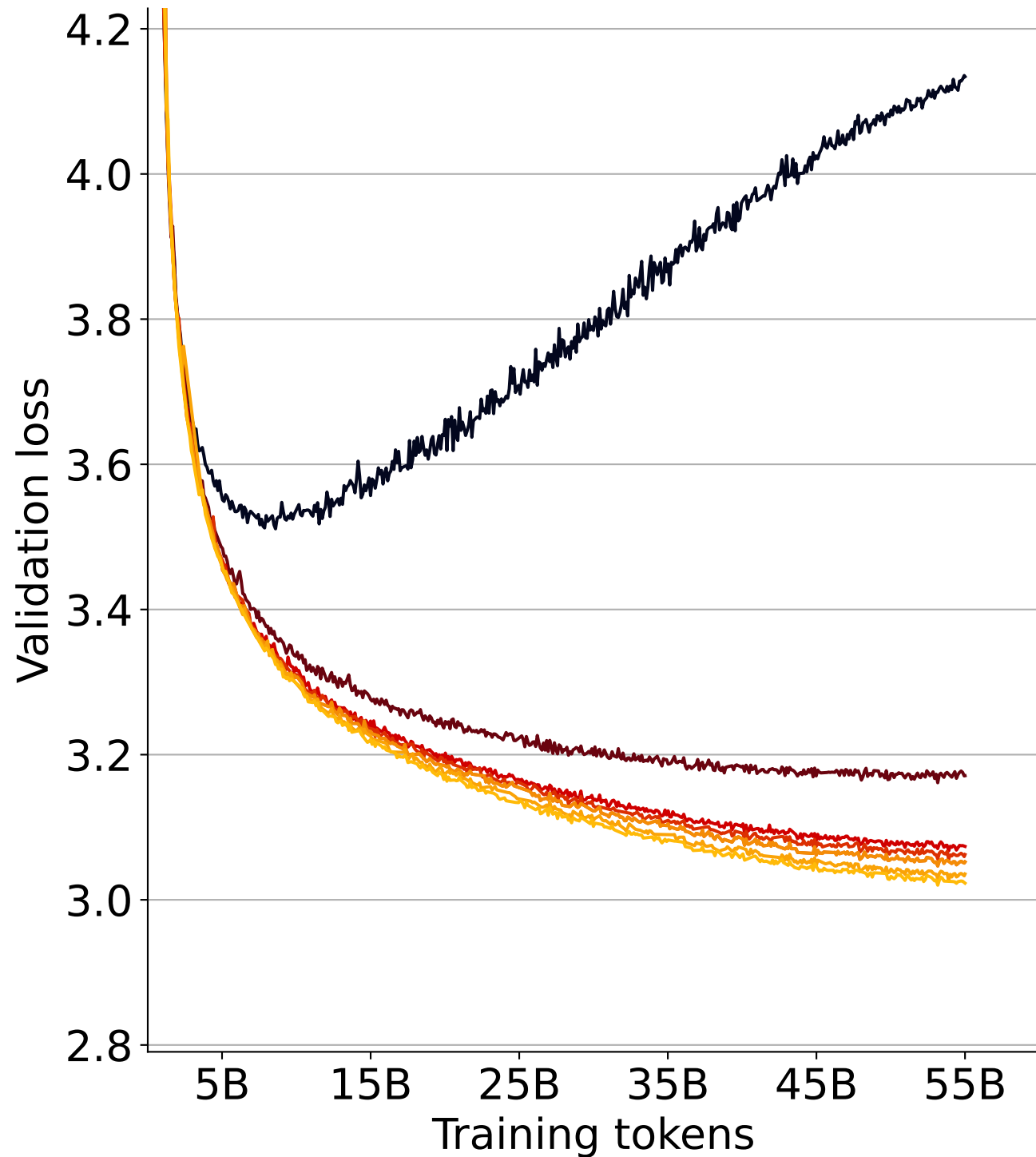
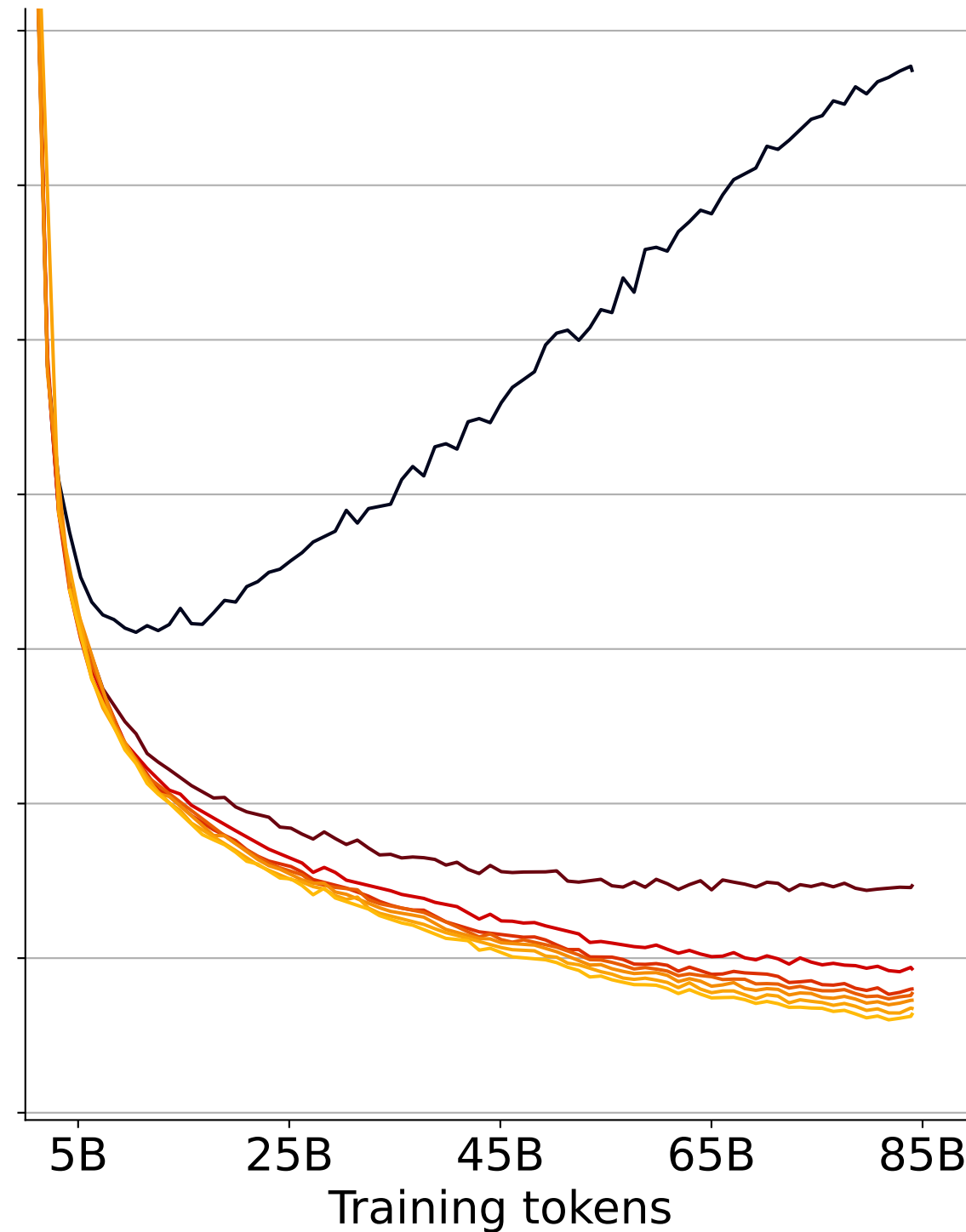


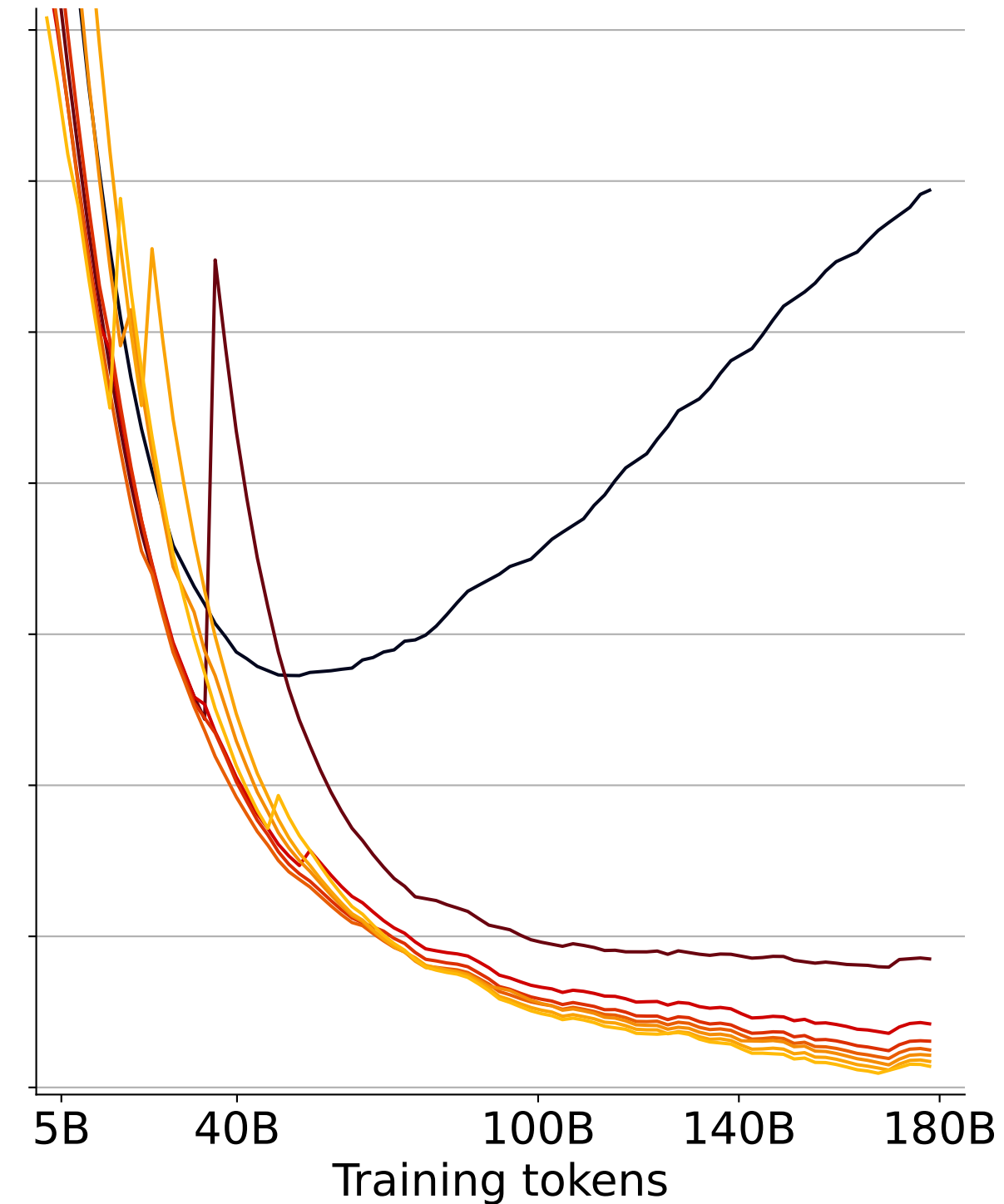
**2.8B parameters trained
for 55B tokens**



**4.2B parameters trained
for 84B tokens**



**8.7B parameters trained
for 178B tokens**



Epochs

1

2

3

4

5

7

14

44