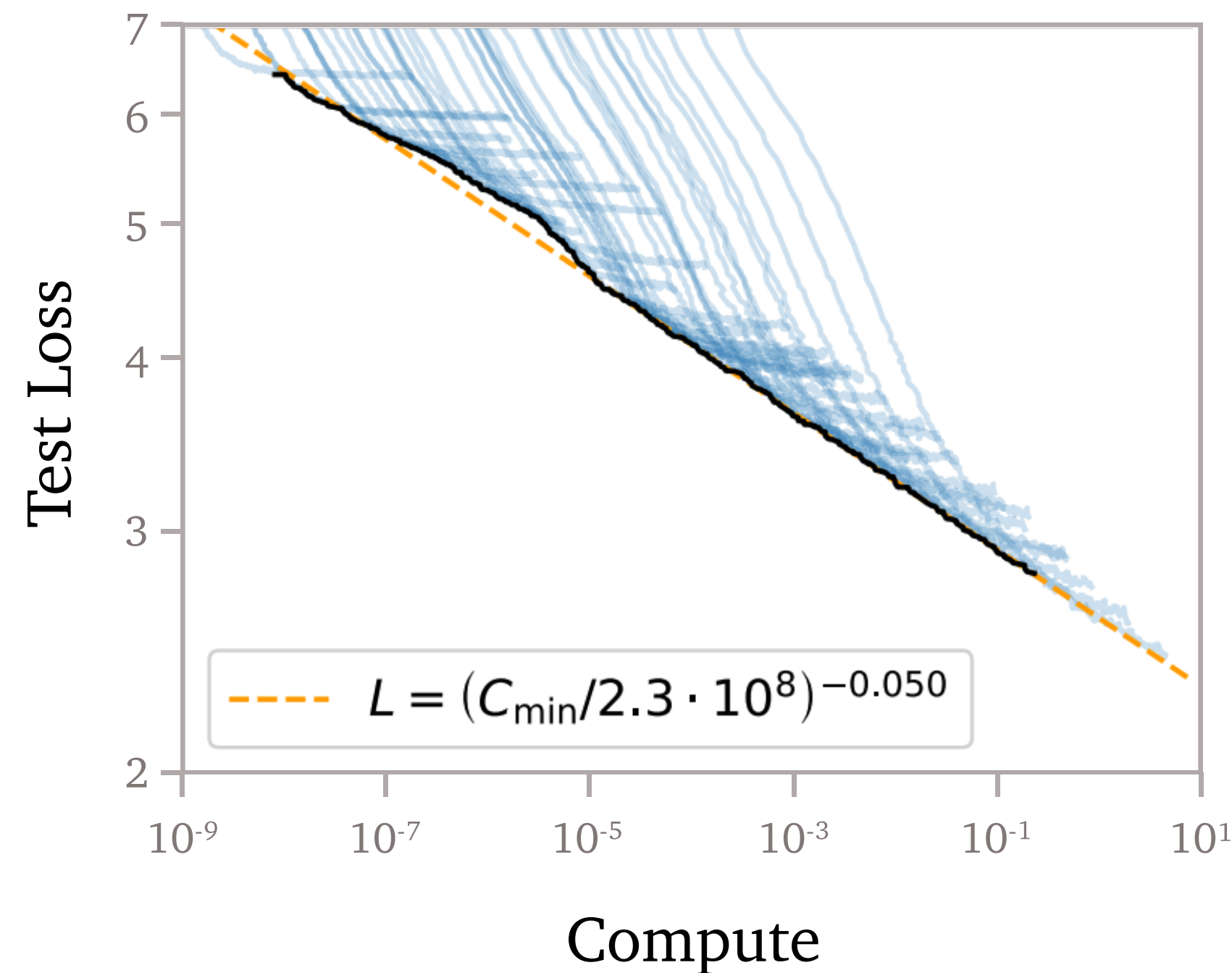


Scaling laws: As compute increases, total loss decreases



Petaflop/s-days*, non-embedding

* A Petaflop/s-days is equivalent to approximately 10^{20} floating point operations

