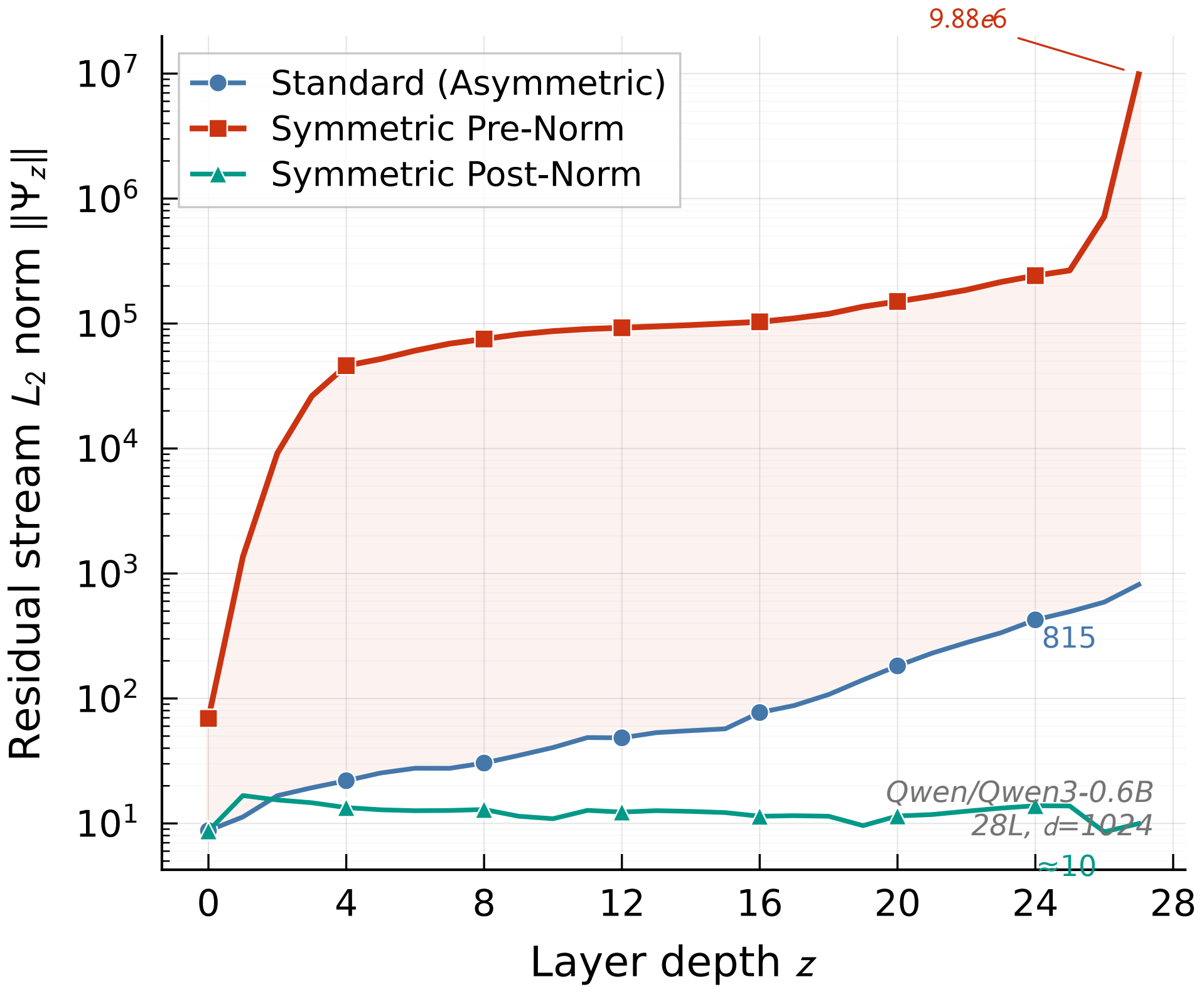
Standard (Asymmetric)
Symmetric Pre-Norm
Symmetric Post-Norm
9.88e6
815
≈10
Qwen/Qwen3-0.6B
28L, d=1024
Residual stream L_2 norm $\|\Psi_z\|$
Layer depth z
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0
4
8
12
16
20
24
28