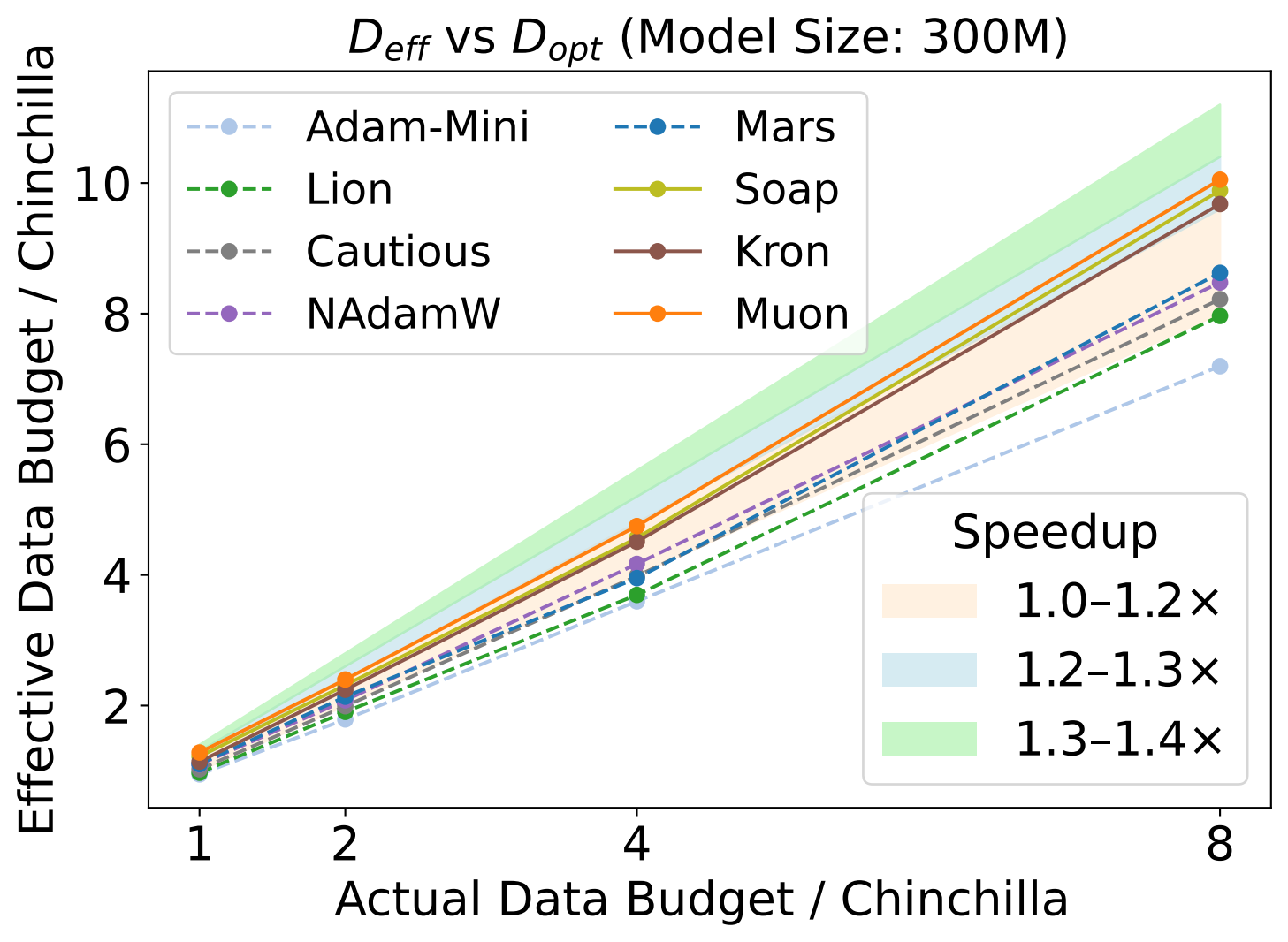
D_{eff} vs D_{opt} (Model Size: 300M)
Adam-Mini
Mars
Lion
Soap
Cautious
Kron
NAdamW
Muon
Speedup
1.0–1.2×
1.2–1.3×
1.3–1.4×
Effective Data Budget / Chinchilla
Actual Data Budget / Chinchilla
2
4
6
8
10
1
2
4
8