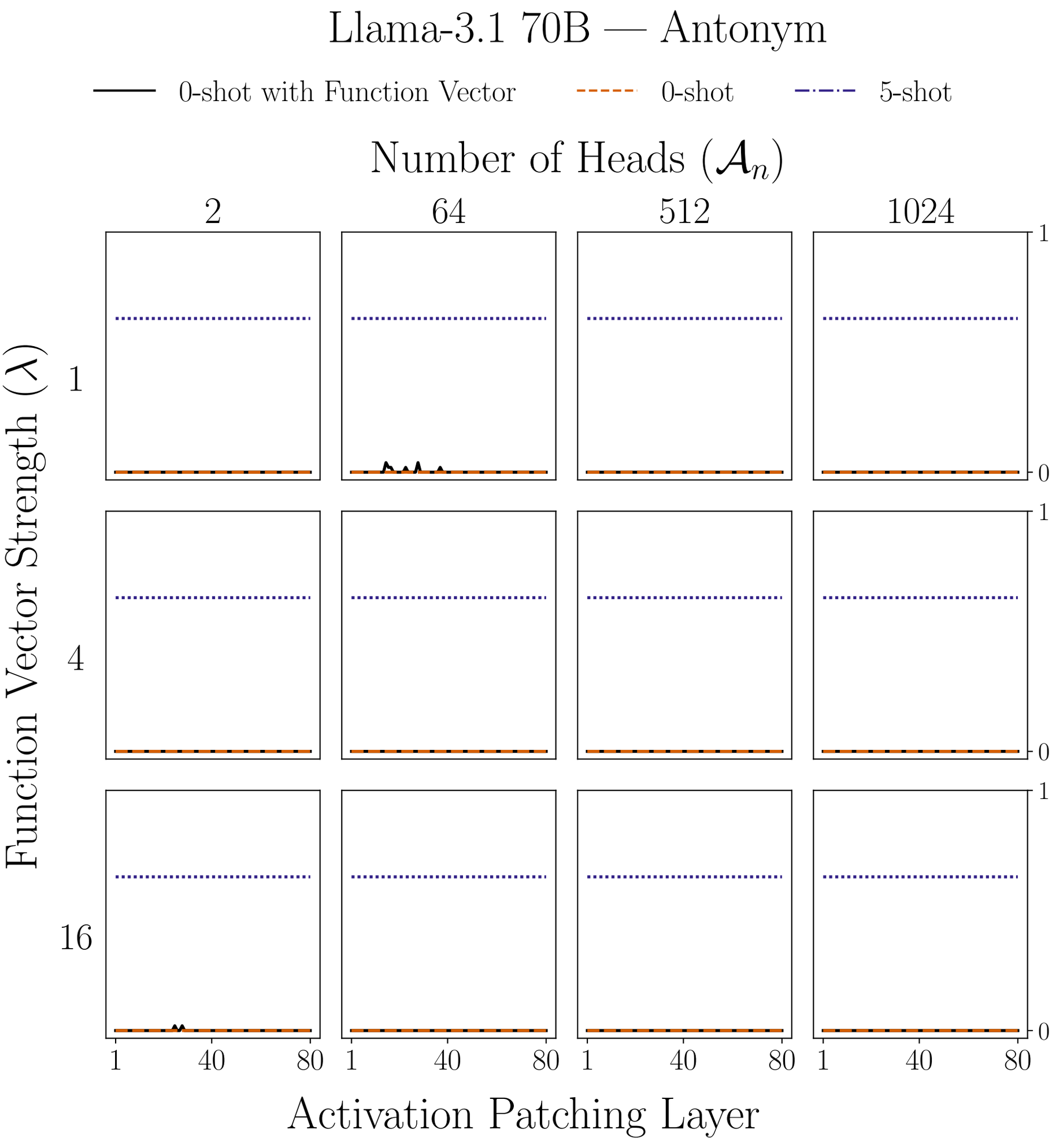
Llama-3.1 70B — Antonym
0-shot with Function Vector
0-shot
5-shot
Number of Heads ($\mathcal{A}_n$)
2
64
512
1024
Function Vector Strength ($\lambda$)
1
4
16
1
0
1
40
80
Activation Patching Layer