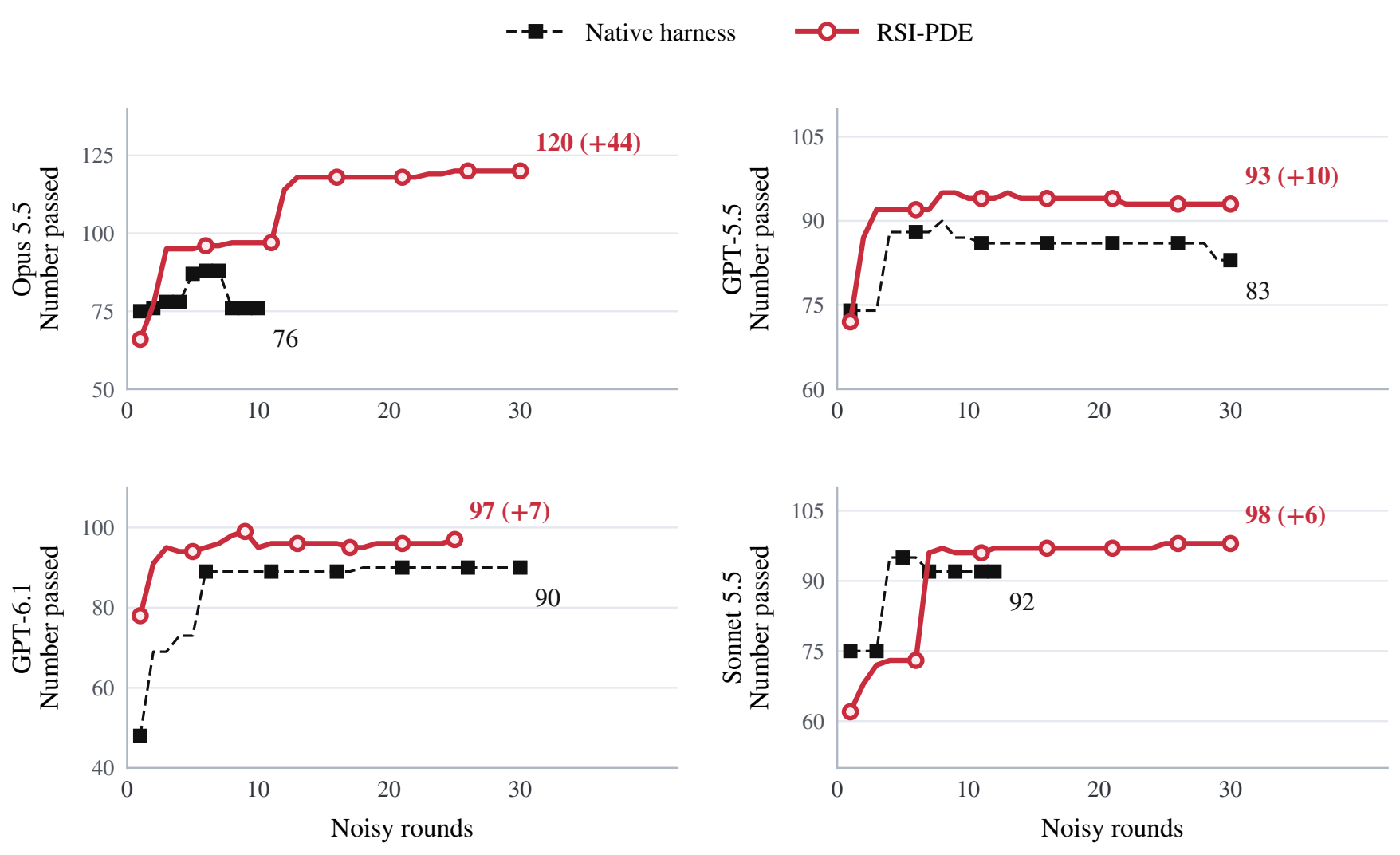

Native harness
RSI-PDE
Opus 5.5
Number passed
120 (+44)
76
GPT-5.5
Number passed
93 (+10)
83
GPT-6.1
Number passed
97 (+7)
90
Sonnet 5.5
Number passed
98 (+6)
92
Noisy rounds
Noisy rounds