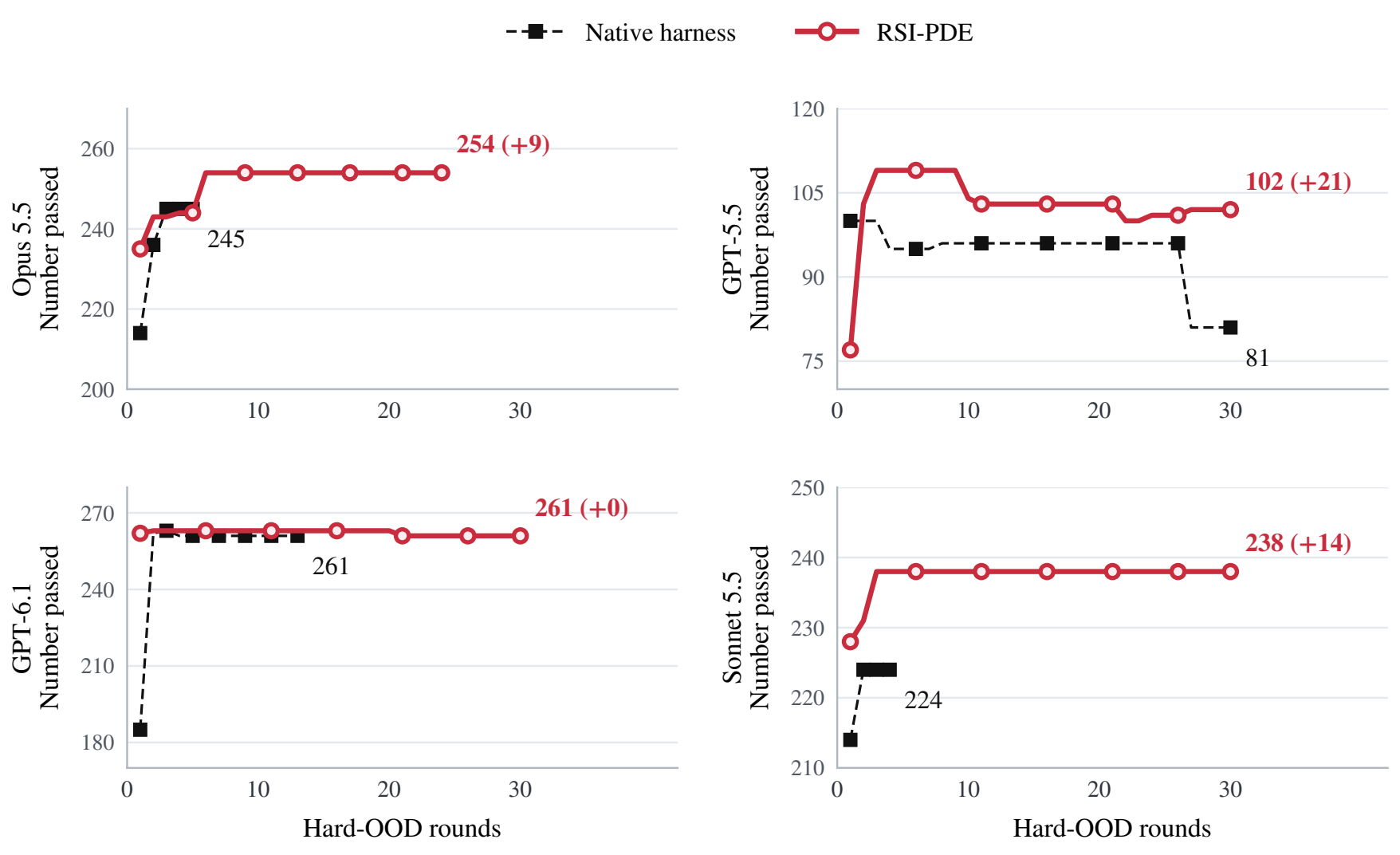

Native harness
RSI-PDE
Opus 5.5
Number passed
254 (+9)
245
GPT-5.5
Number passed
102 (+21)
81
GPT-6.1
Number passed
261 (+0)
261
Sonnet 5.5
Number passed
238 (+14)
224
Hard-OOD rounds
Hard-OOD rounds