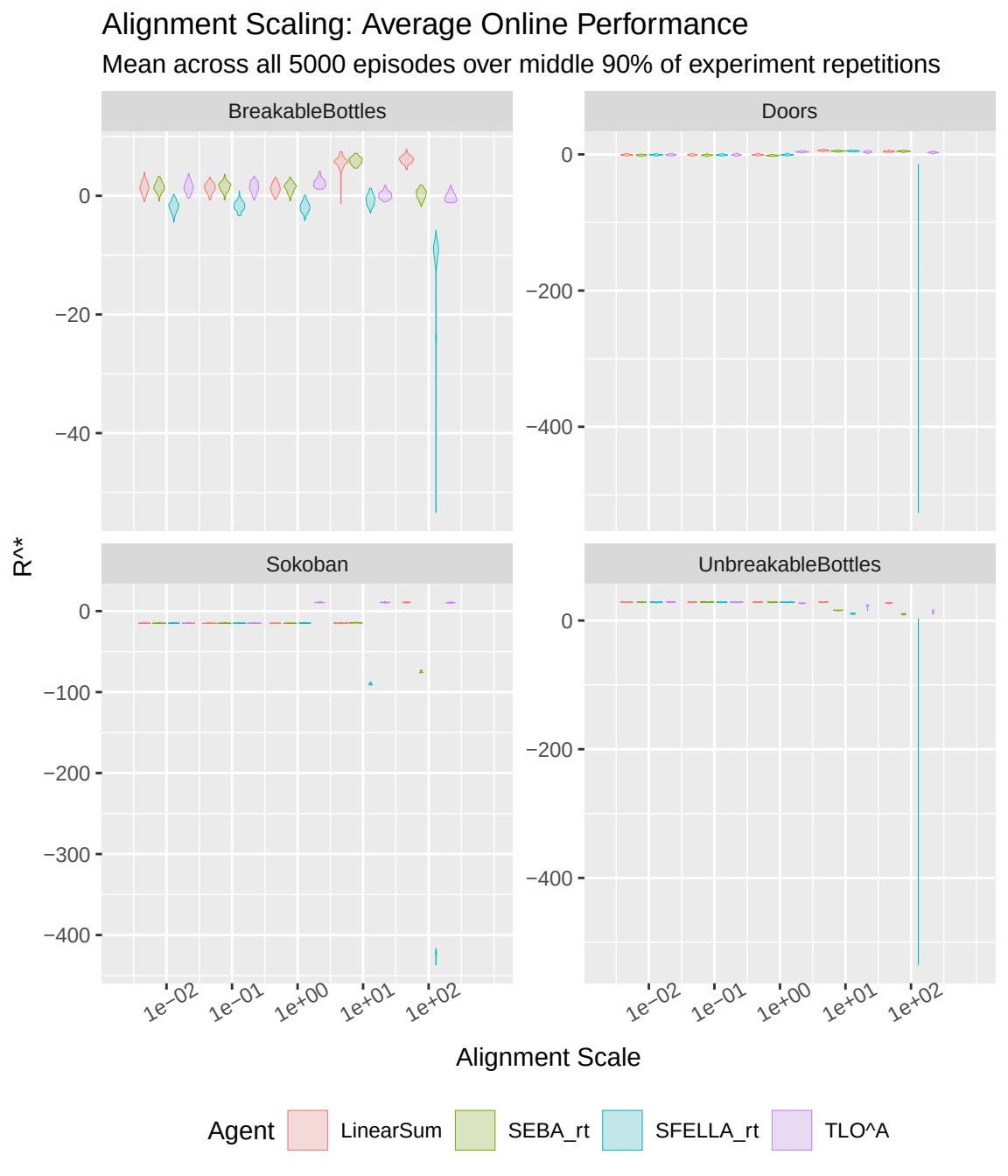
Alignment Scaling: Average Online Performance
Mean across all 5000 episodes over middle 90% of experiment repetitions
BreakableBottles
Doors
Sokoban
UnbreakableBottles
R^*
0
−20
−40
−200
−400
−100
−300
1e−02
1e−01
1e+00
1e+01
1e+02
Alignment Scale
Agent
LinearSum
SEBA_rt
SFELLA_rt
TLO^A