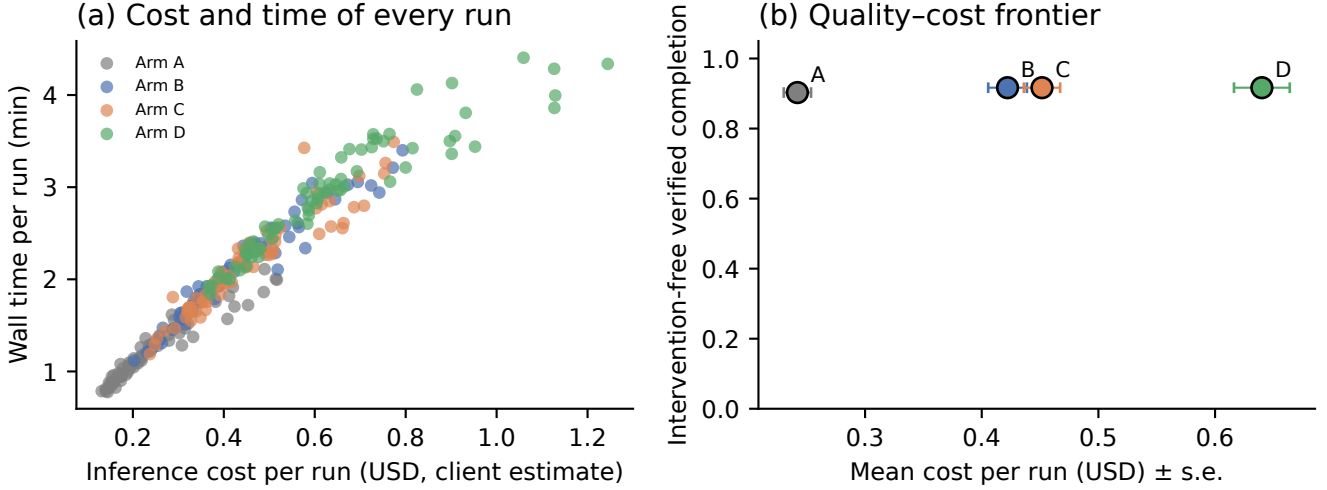

(a) Cost and time of every run
Arm A
Arm B
Arm C
Arm D
Wall time per run (min)
1
2
3
4
0.2
0.4
0.6
0.8
1.0
1.2
Inference cost per run (USD, client estimate)
(b) Quality–cost frontier
Intervention-free verified completion
0.0
0.2
0.4
0.6
0.8
1.0
A
B
C
D
0.3
0.4
0.5
0.6
Mean cost per run (USD) ± s.e.