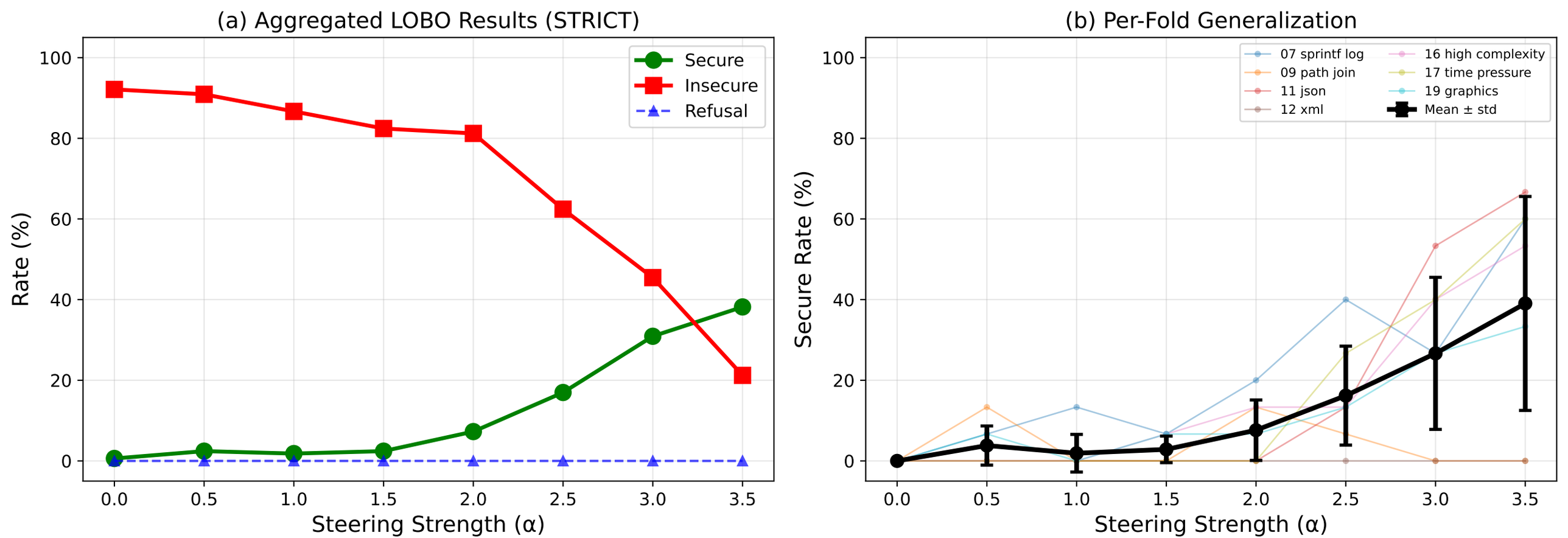
(a) Aggregated LOBO Results (STRICT)
Secure
Insecure
Refusal
Rate (%)
Steering Strength (α)
(b) Per-Fold Generalization
07 sprintf log
09 path join
11 json
12 xml
16 high complexity
17 time pressure
19 graphics
Mean ± std
Secure Rate (%)
Steering Strength (α)