

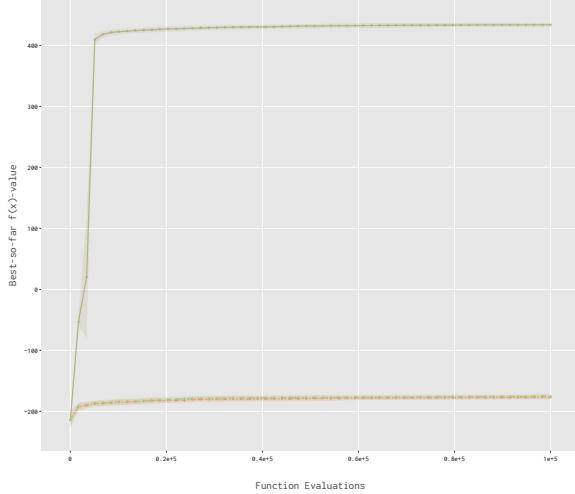
Exercise 4: Large Budgets (100,000 evaluations) on *MaxCoverage* and *MaxInfluence*

Overview

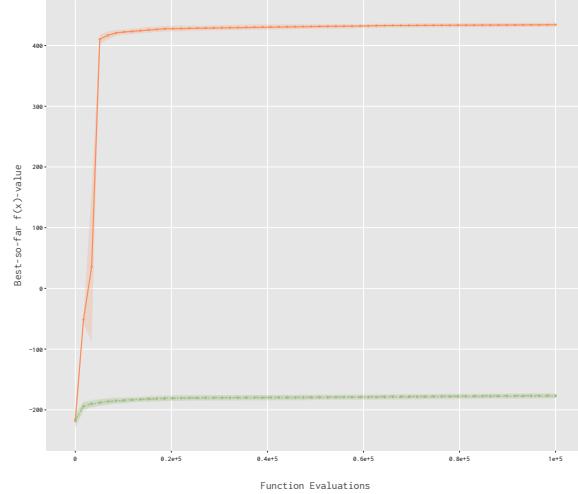
We evaluate instances 2100–2103 (*MaxCoverage*) and 2200–2203 (*MaxInfluence*) with a fixed budget of **100 000** evaluations. The **six approaches** shown in the fixed-budget panels are: GA, RLS, One Plus One (1+1) EA, Single Objective (SO), Multi Objective (MO) and GSEMO. We then report GSEMO trade-offs (Pareto fronts) from the *first* run. Finally, we summarise improvements from **10k** (Exercise 3) to **100k** evaluations.

Direction/sign caveat (Influence) Legacy single-objective pipelines export Influence values via IOH with negation/logging; all figures and deltas herein follow the consolidated convention that *higher is better* for comparability across panels.

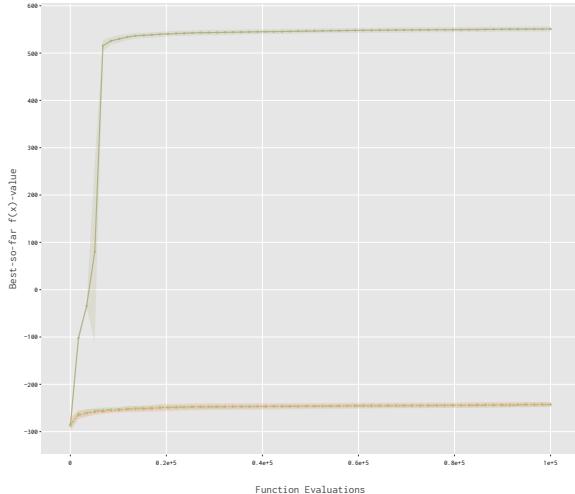
A. Fixed-budget mean \pm sd at 100k



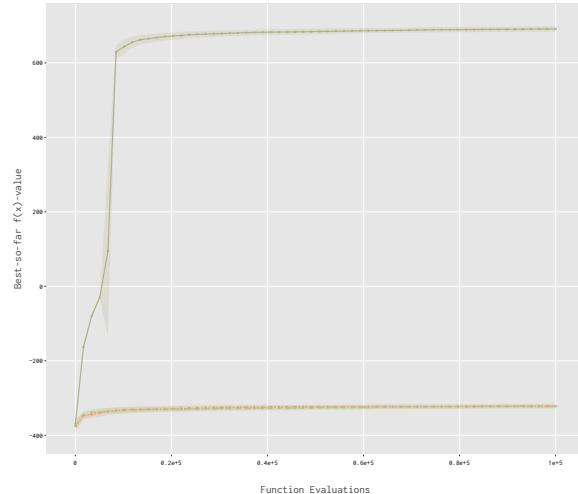
(a) RLS, (1+1)EA, GA — 2100



(b) RLS, (1+1)EA, GA — 2101



(c) RLS, (1+1)EA, GA — 2102



(d) RLS, (1+1)EA, GA — 2103

Figure 1: Fixed-budget (100k) plots for RLS, (1+1) EA and GA (Instance 210x).

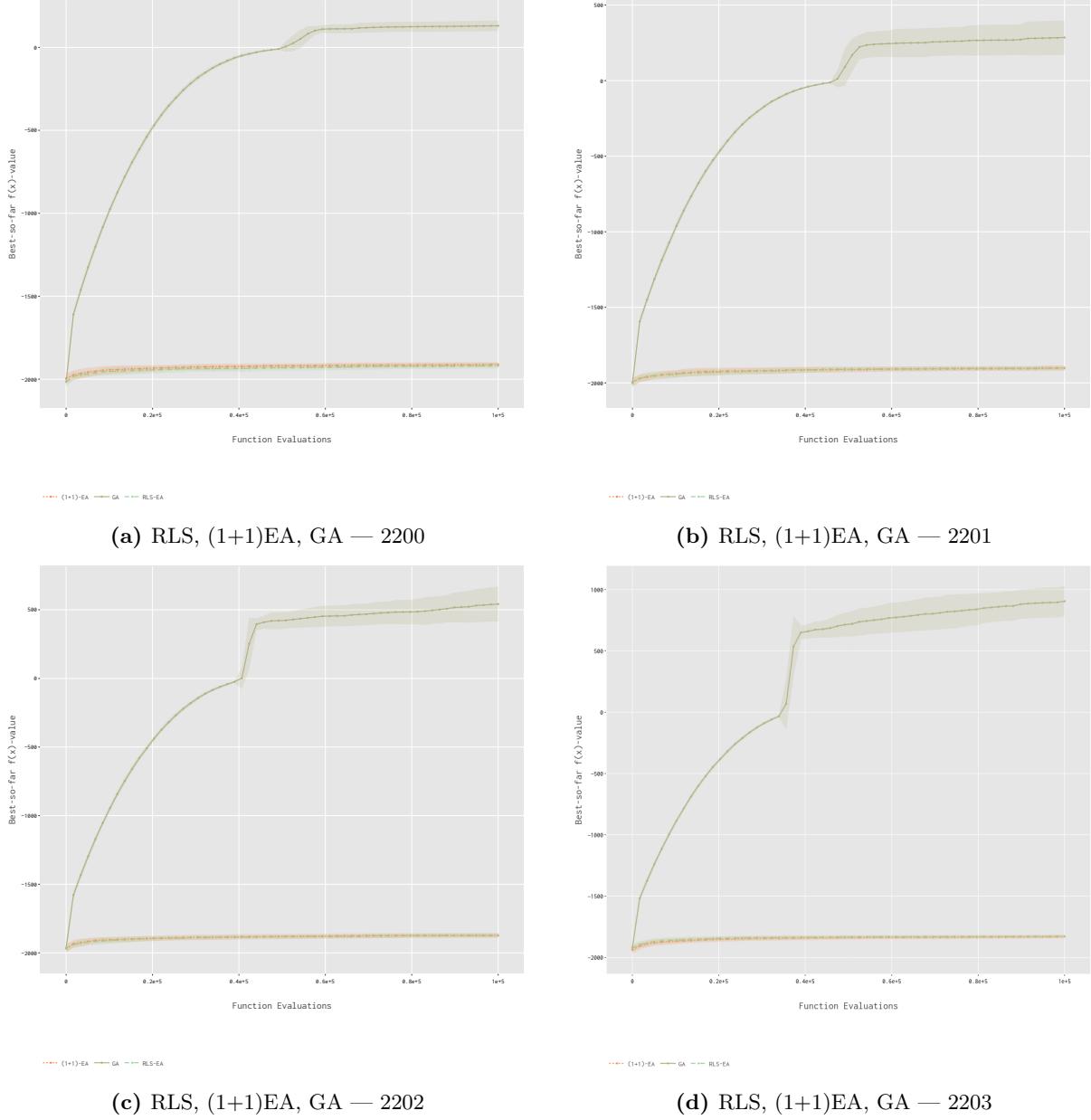
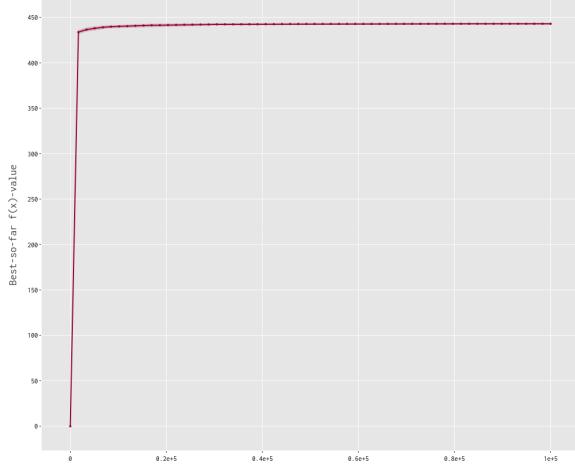
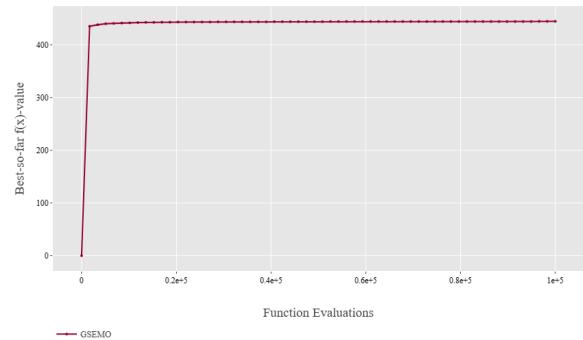


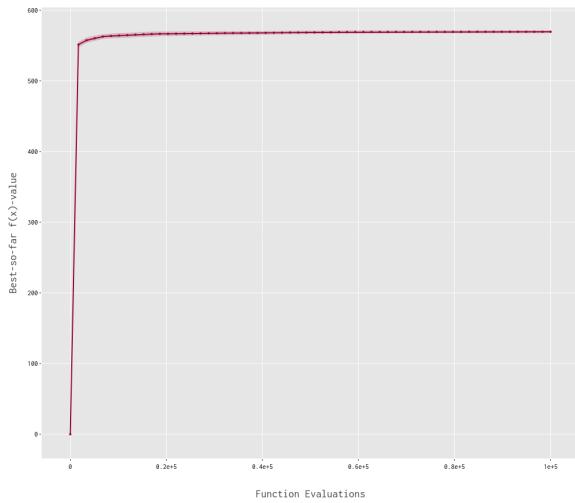
Figure 2: Fixed-budget (100k) plots for RLS, (1+1) EA and GA (Instance 220x).



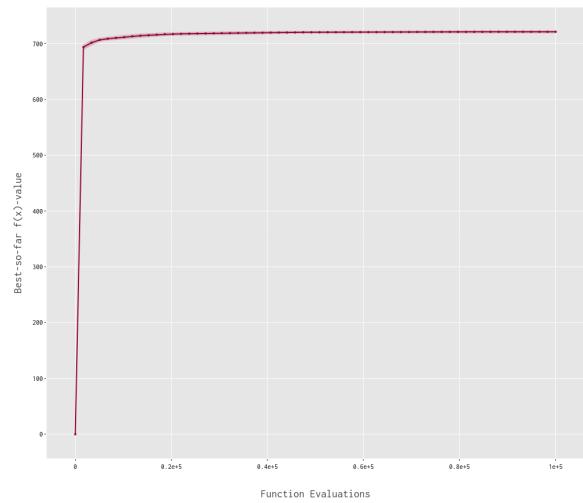
(a) GSEMO — 2100



(b) GSEMO — 2101



(c) GSEMO — 2102



(d) GSEMO — 2103

Figure 3: Fixed-budget (100k) plots for **GSEMO** (Instance 210x).

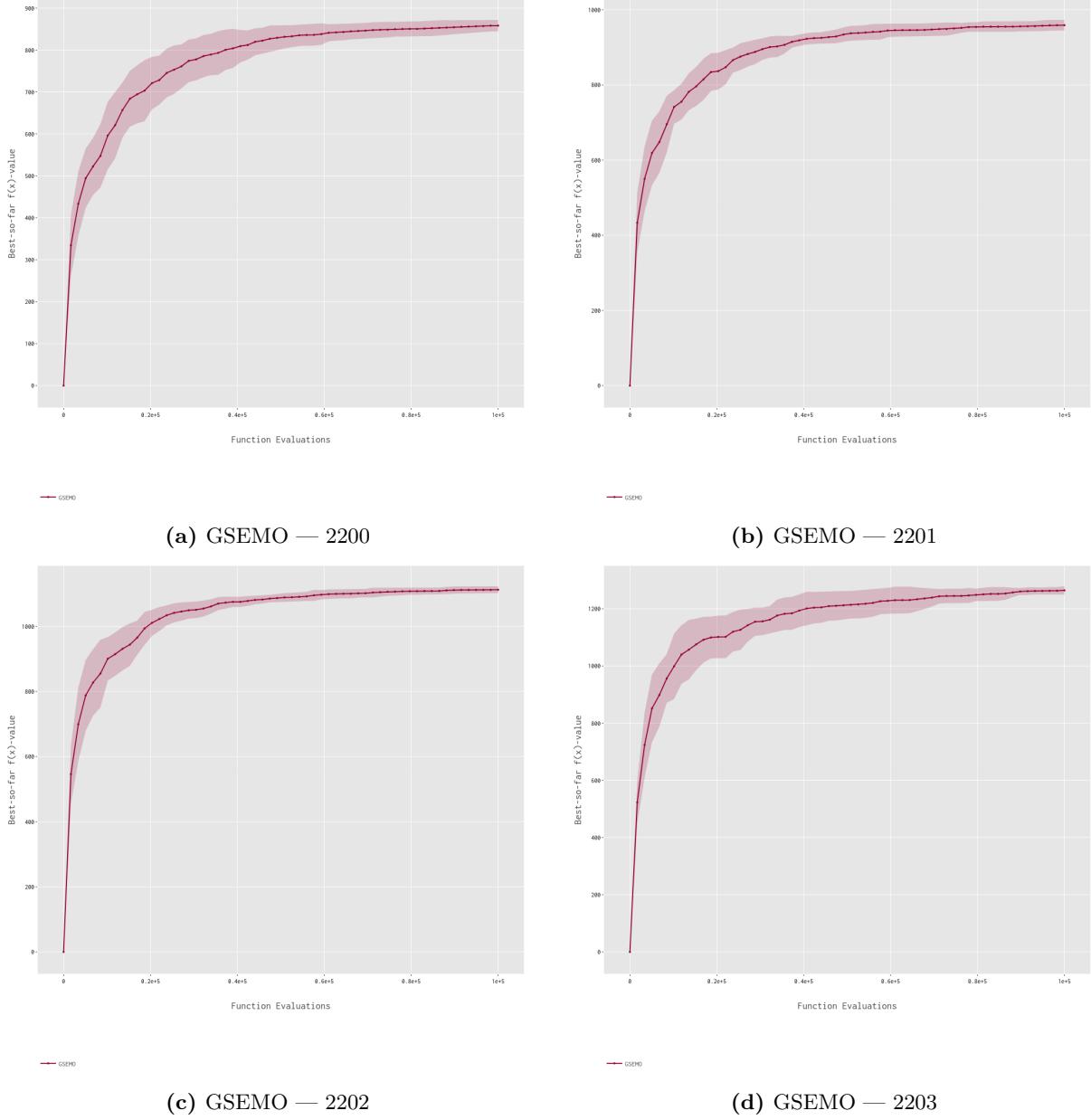


Figure 4: Fixed-budget (100k) plots for **GSEMO** (Instance 220x).

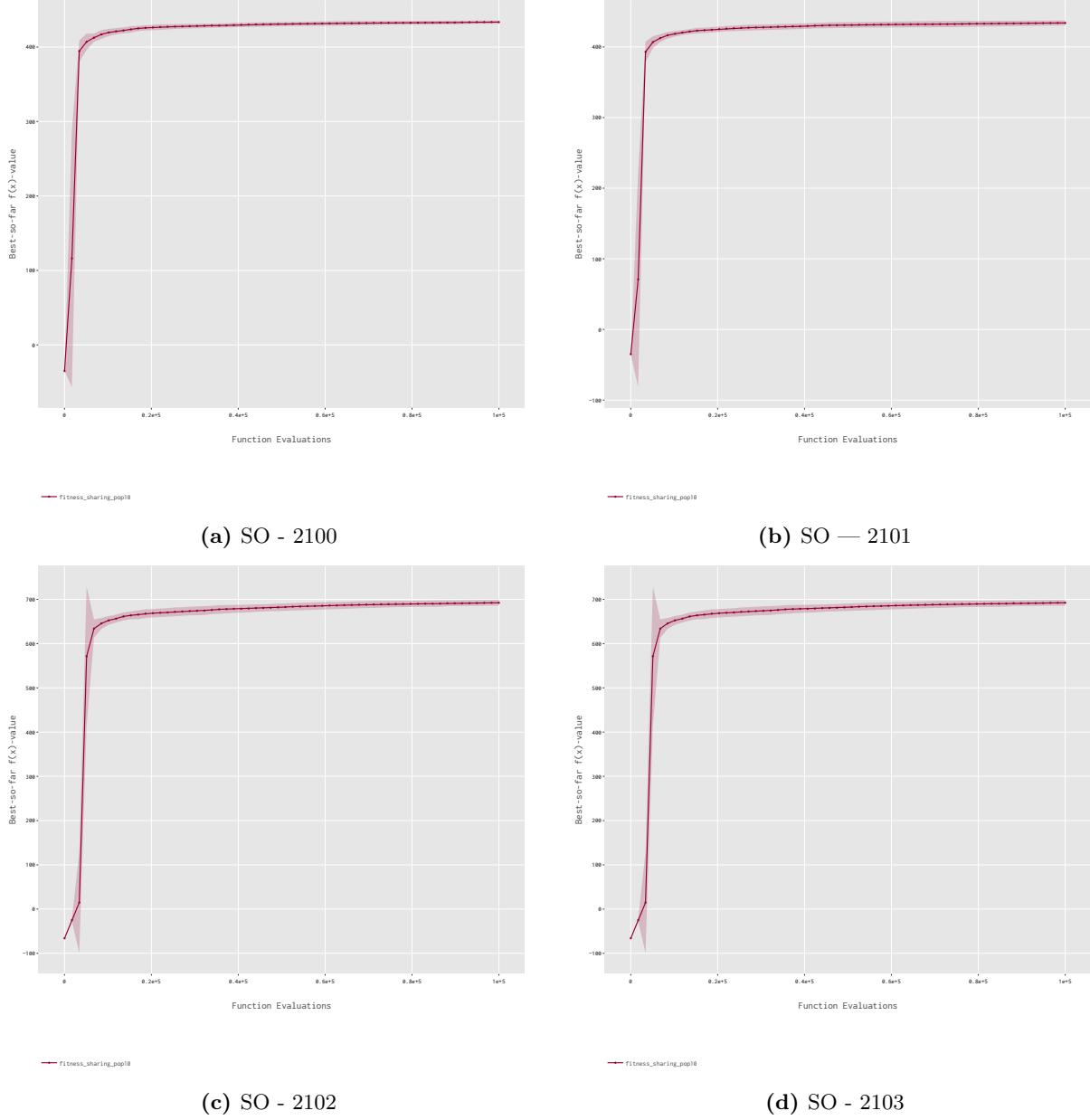
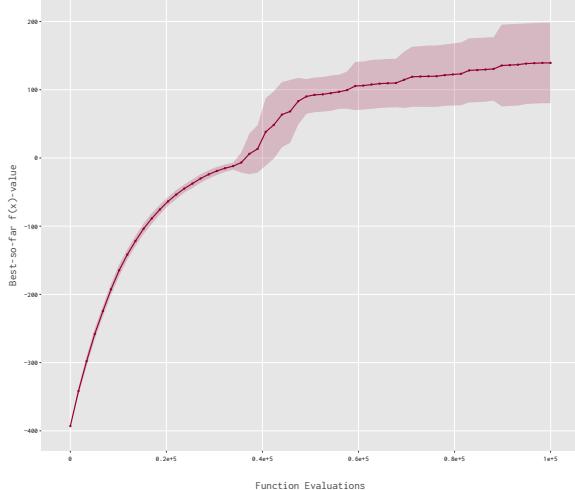
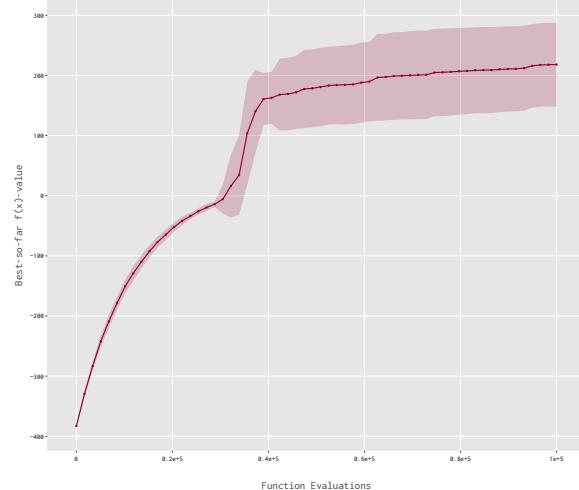


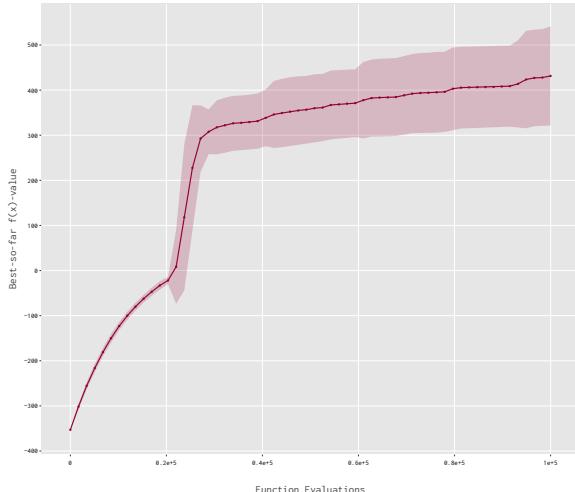
Figure 5: Fixed-budget (100k) plots for **SO** (Instance 210x)



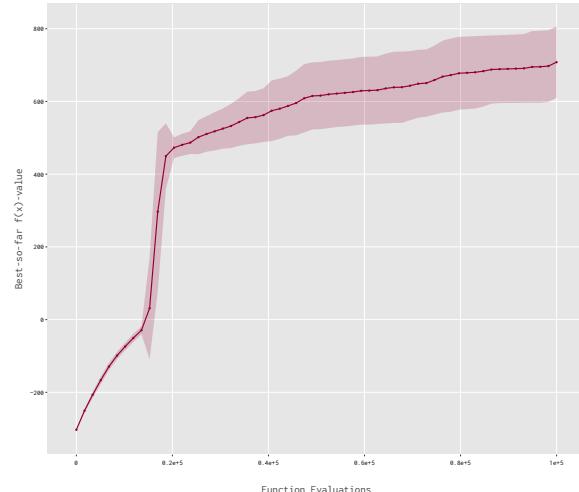
(a) SO — 2200



(b) SO — 2201



(c) SO - 2202



(d) SO - 2203

Figure 6: Fixed-budget (100k) plots for **SO** (Instance 220x)



Figure 7: Fixed-budget (100k) plots for MO

Part A: brief synthesis across plots.

- **Coverage (2100–2103):** All six approaches rise rapidly and plateau early at 100k; additional budget chiefly *tightens uncertainty bands*. Among baseline single-objective methods GA/RLS/OPO dominate the (1+1)-EA. GSEMO (although multi-objective) tracks steady gains with small late improvements, consistent with knee discovery followed by diminishing returns.
- **Influence (2200–2203):** Curves improve for longer; 100k yields *clear endpoint lifts* and *narrower spreads*. Ranking is stable vs. 10k, with (1+1)-EA weakest and GA/RLS/OPO strongest; the Ex3 references corroborate that the larger budget mainly improves reliability and tail performance rather than changing the pecking order.

B. GSEMO Pareto trade-offs (first run, 100k budget)

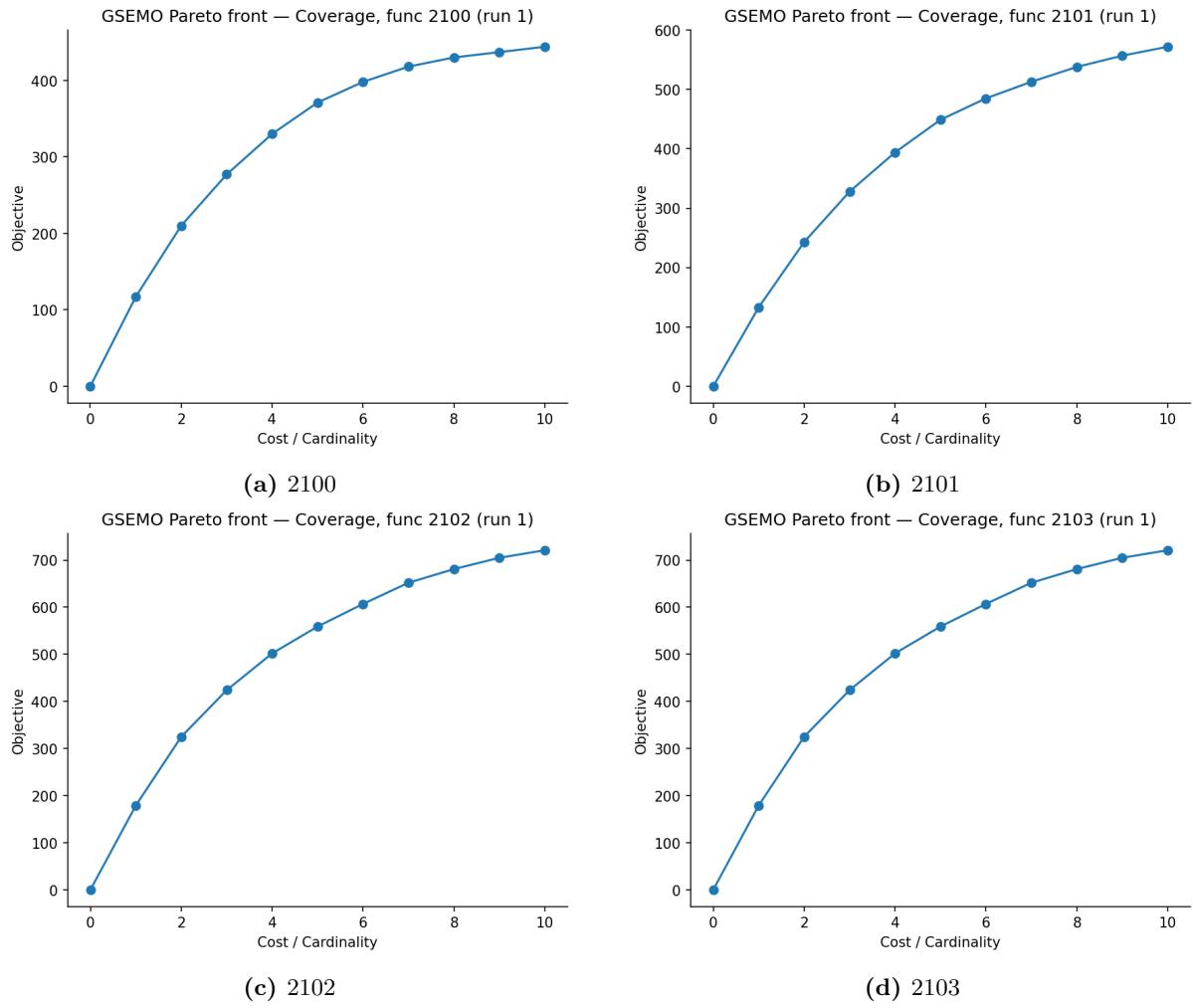


Figure 8: GSEMO value–cost Pareto fronts at 100k (first run) — **MaxCoverage**.

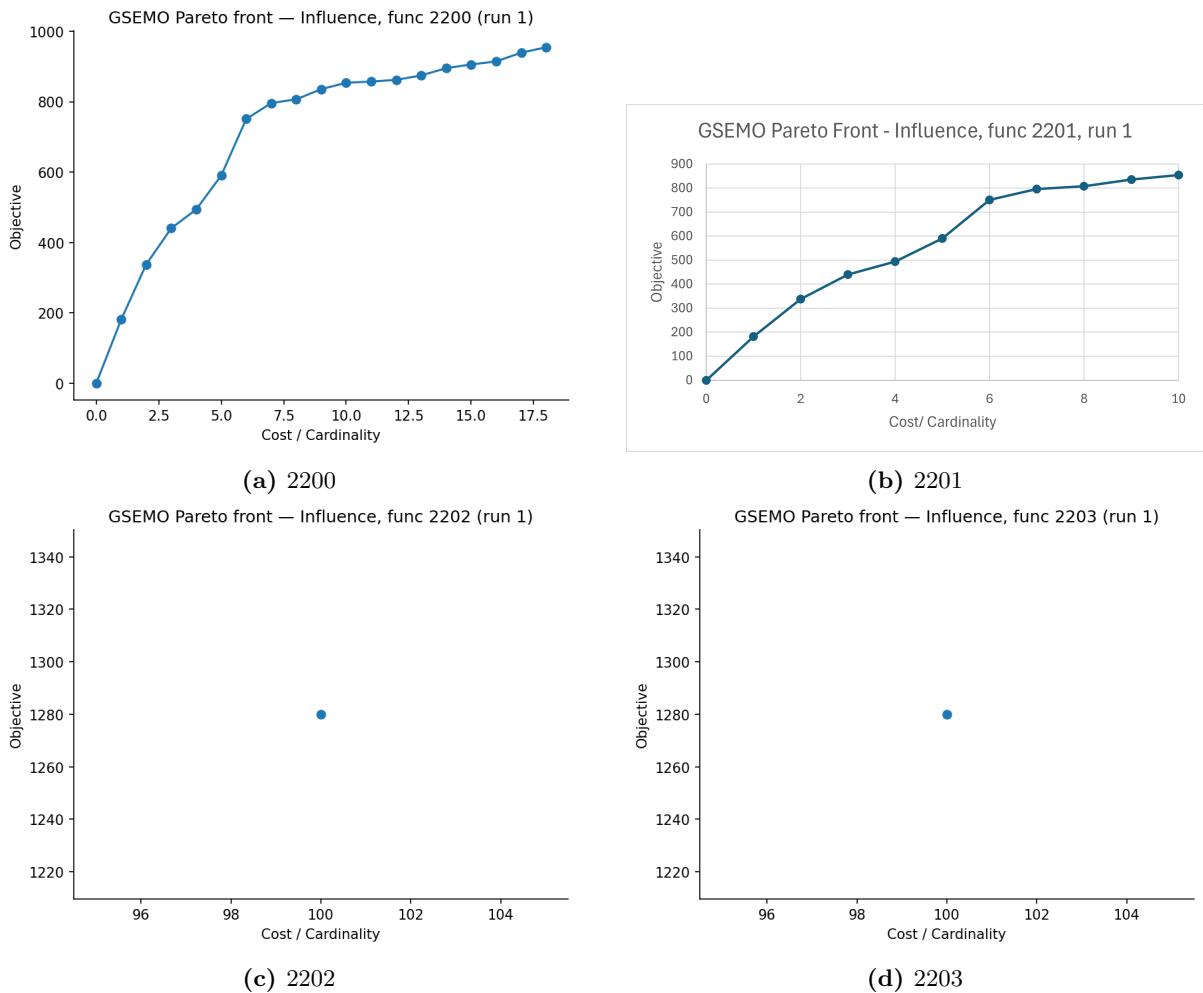


Figure 9: GSEMO value–cost Pareto fronts at 100k (first run) — **MaxInfluence**.

Part B: brief synthesis and caveat.

- **Shape:** Coverage fronts are smooth, concave, and show *knees* where large value gains require sharply higher cost. Influence fronts are closer to linear with *diminishing returns* as cost increases.
- **Population budget trade-off:** At fixed 100k, fronts are denser at mid–high costs but still sparse at extremes, indicating budget spent maintaining diversity rather than fully refining the extremes.

Run-count caveat (Influence). GSEMO runs were not fully completed on three Influence instances; at submission time we had **f2201: 18**, **f2202: 10**, and **f2203: 6** runs. The panels shown use the *first available* 100k run and should be interpreted with this limitation.

C. 10k → 100k: summary of improvements

- **Coverage (2100–2103):** With 10k, most approaches already reach the knee; moving to 100k primarily *reduces variance* and yields modest endpoint gains.
- **Influence (2200–2203):** Gains continue later in the budget; 100k gives *higher endpoints* and *narrower bands*, widening the gap to (1+1)-EA.
- **GSEMO (MO):** 100k *densifies* the frontier and smooths knees on Coverage; for Influence, fronts shift upward but remain close to linear at large cardinalities (diminishing returns).

Comparability caveat (10k vs 100k). The 10k and 100k budgets were executed with *different algorithm sets*; therefore, cross-budget comparisons quote the *best observed mean per instance at that budget* rather than a like-for-like algorithm-to-algorithm comparison.