

Ordinal Geometry Complements Reconstruction

Diagnosing Planning with Compressed Value Functions

Shiheng Zhang University of Washington



Paper



Code

NeurIPS 2026 Atlanta

In hard topology, local action order predicts planning success beyond reconstruction error.

Theory shows how dense small action gaps weaken reconstruction's control of local order.

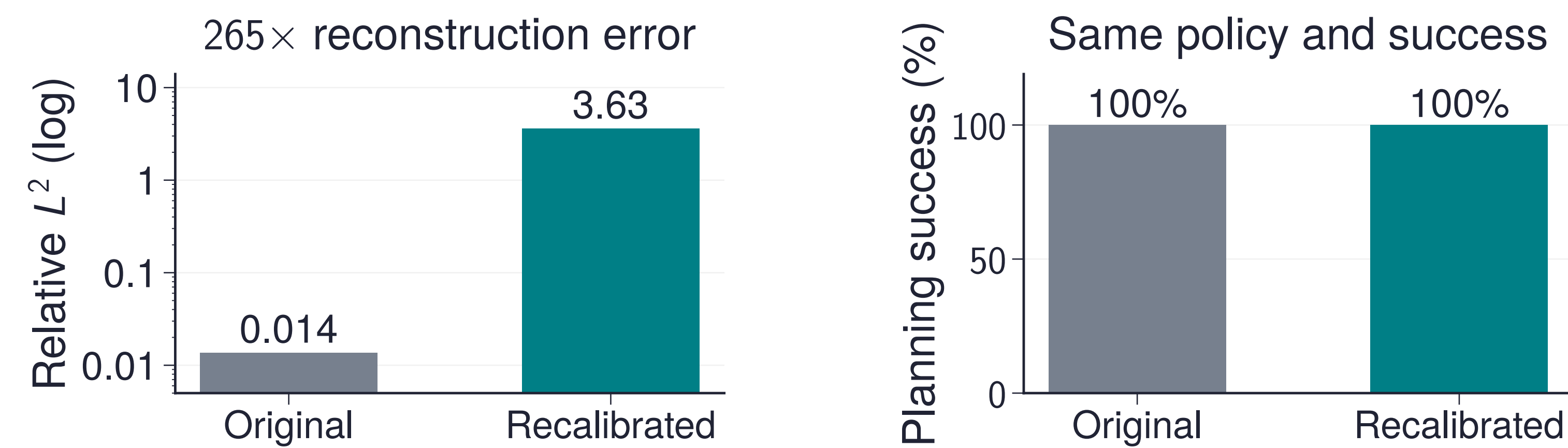
1 Error magnitude is not enough

A greedy planner picks the action with the lowest score:

$$\hat{Q}(x, u; g) = c(x, u) + \hat{V}(f(x, u), g).$$

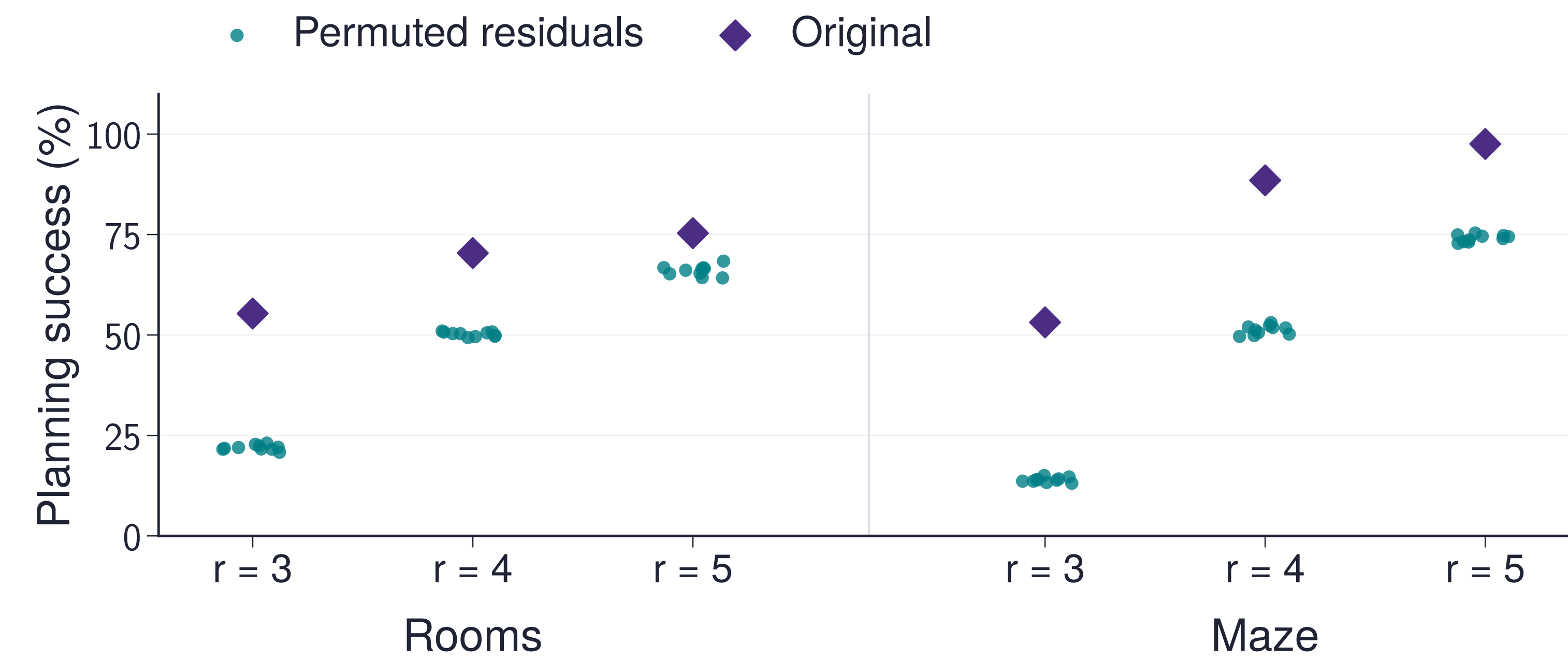
$\hat{V}$ approximates the optimal cost V^* . Here r denotes SVD rank or the learned latent dimension.

Monotone recalibration (20×20 wall grid)



On a uniform-cost wall grid, monotone recalibration raises L^2 by **265×** while preserving every selected action and planning success.

Fixed-norm shell permutation (15×15 grids)

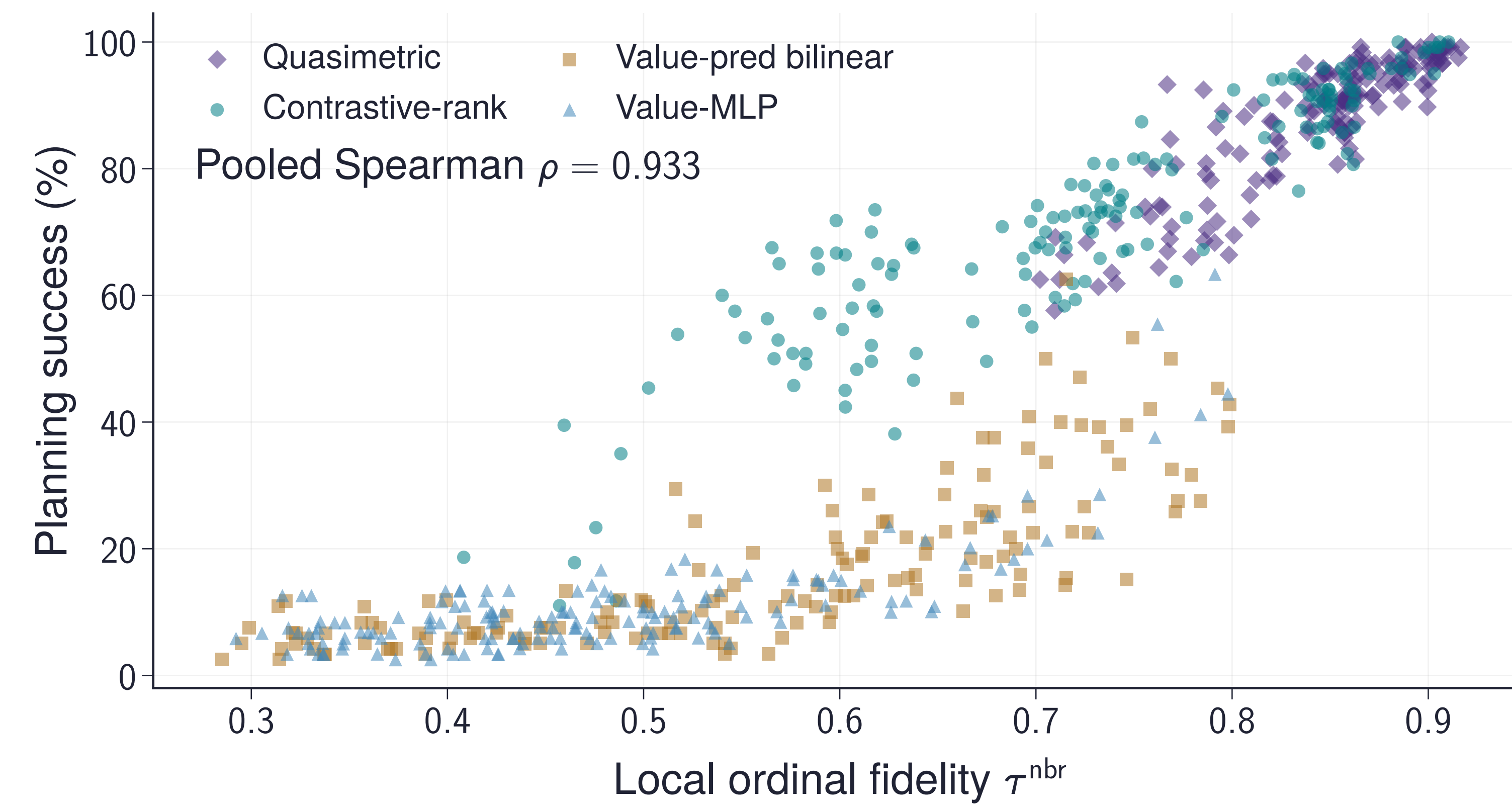


Permuting signed residuals within exact goal-distance shells preserves L^2 and L^∞ . Yet **all 60 trials** reduce τ^{nbr} , action agreement, and success.

Two grids, three ranks, ten permutations per cell. Residual placement, including local order and smoothness, matters.

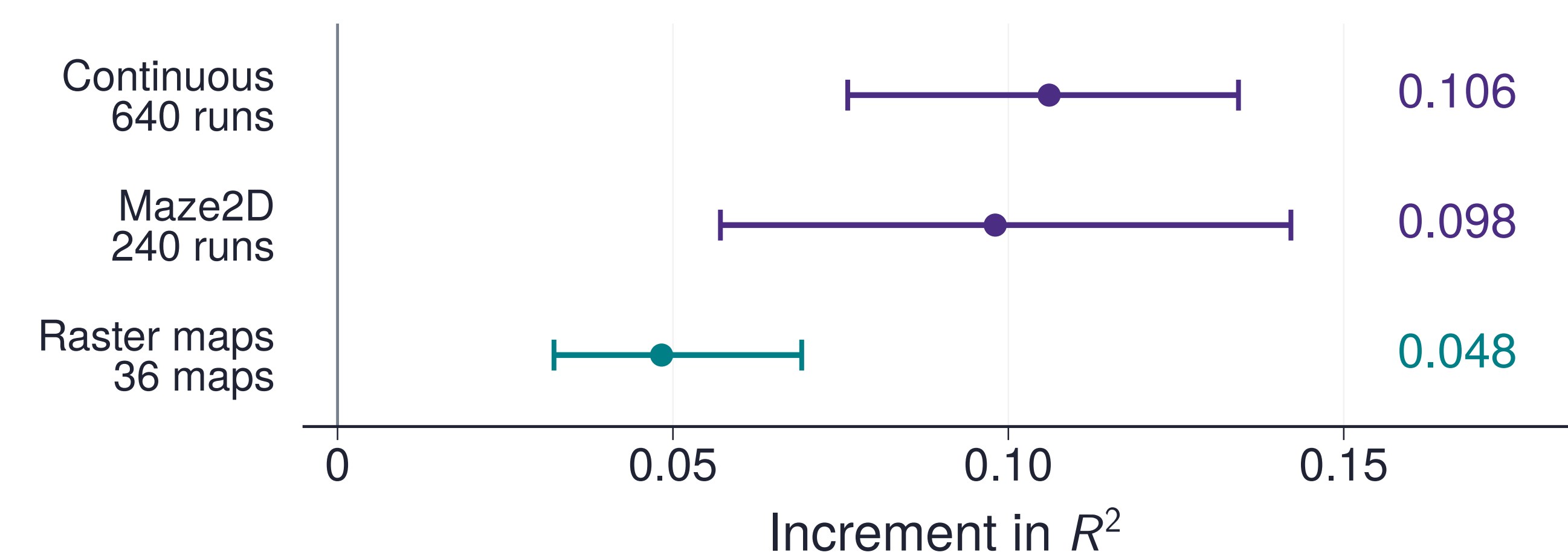
2 Local order complements L^2

τ^{nbr} : Kendall τ_b between true and estimated one-step scores over each state's actions, averaged over state-goal pairs. Higher is better.



640 runs, 8 navigation environments. Four neural objectives, four latent dimensions, five seeds. Each point is one trained representation.

Added explanatory power for success



ΔR^2 from adding τ^{nbr} beyond L^2 and fixed effects; bars show 95% bootstrap intervals. Continuous/Maze2D control for environment. Raster maps also control for objective, training-set size, checkpoint, and seed.

Action agreement checks only the greedy action and predicts greedy success more strongly. τ^{nbr} scores the full local ranking. Coordinate-input increments are mainly between objectives.

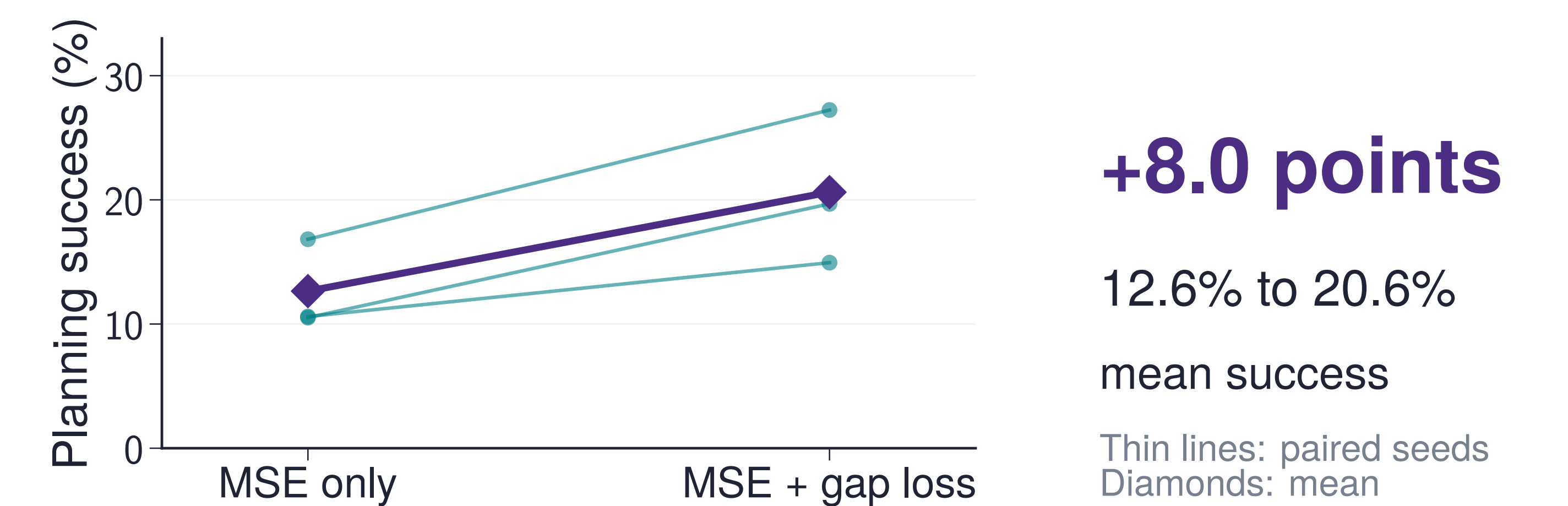
3 Gap-weighted inversions bound excess cost

Small errors can reverse near-tied actions. A wrong choice incurs excess cost according to its **true action gap** $\Delta_t(u)$.

$$R_t^{\text{gap}} = \sum_{u \neq u_t^*} \Delta_t(u) \mathbf{1}\{\hat{Q}_t(u) \leq \hat{Q}_t(u_t^*)\}$$
$$J_{\hat{\pi}} - V^* = \sum_{t < T} \beta_t \leq \sum_{t < T} R_t^{\text{gap}}.$$

For finite successful rollouts in deterministic shortest-path problems. u_t^* is optimal; β_t is the chosen action's Bellman excess.

With exact gaps, a gap-weighted loss improves greedy planning



A smooth gap-weighted surrogate augments MSE. **3/3 paired seeds improve.** $\lambda = 3$ selected on wall_drift, then evaluated on circle_drift. Value-MLP, $r = 16$; exact gaps and training goals.

Replication with high-dimensional context

36 fresh maps; 441-dimensional obstacle rasters. Shared CNN-MLP models, 160 checkpoints.

$\Delta R^2 = \mathbf{0.048}$, 95% CI [0.032, 0.069].

36/36 leave-one-map-out estimates are positive.

Wrong-map input lowers success by **10.6 points**.

Known dynamics. The wrong-map control verifies raster use in the 384-map MSE model across four seeds.

Coordinate-only success is comparable: 0.538 vs. 0.526 with raster input.

For offline checkpoint selection: report L^2 , local order τ^{nbr} , and top-1 action agreement together.