Planner-Admissible Graph-PDE Value Extensions for Sparse Goal-Conditioned Planning

Anonymous authors
Paper under double-blind review

Abstract

Sparse goal-conditioned planning with few cost-to-go labels can be cast as graph-PDE Dirichlet extension: extend labels on a small boundary Γ_g to unlabelled vertices so that greedy rollouts reach the goal. We argue the key criterion is *planner admissibility*, not pointwise error. Within the graph p-Laplacian family, harmonic averaging ($p = 2$) often fails to preserve local greedy orderings, whereas high-p and Absolutely Minimal Lipschitz Extension (AMLE) / Lipschitz-extremal extensions enter a high-success regime. AMLE is the simple, analyzable $p = \infty$ endpoint of this family, with a local midrange update and a clean comparison / Lipschitz theory.

Our main theoretical statement is a **planner-admissibility certificate**: under the operational argmin-Q planner, if at every state visited by the surrogate-greedy rollout the local value error stays below half the true action gap, then the rollout reaches the goal (Theorem 3.3). AMLE instantiates the certificate through a comparison-principle fill-distance bound (Corollary 3.5); harmonic can violate it because its neighbour rankings are boundary-label-weighted hitting-probability rankings, not shortest-path rankings (Lemma 3.8, Lemma 3.9). All p-Laplacian endpoints share a structural maximum principle (Lemma 3.6): no strict interior sinks across $p \in [2, \infty]$. An operator-level identity (Proposition 3.10) shows that shortest-path distance is locally AMLE-compatible on the geodesically extendable interior, while harmonic-compatibility requires a non-generic degree-balance condition $n_+ = n_-$ (Corollary 3.11). A refinement-stable seven-node construction is a positive-margin witness for a local harmonic-ordering failure mode that persists under graph subdivision.

On 120 AntMaze layout-derived graph configurations (Fu et al., 2020), aggregate rollout success is 0.584 for $p = 2$, 0.903 for $p = 4$, 0.973 for $p = 8$, 0.982 for a fixed-budget $p = 16$ solver, and 0.970 for AMLE. Large finite-p rows are not used for exact endpoint ranking because solver certification is incomplete. On the rollout-weighted decision scope, AMLE reduces the low neighbour-Kendall-τ_{nbr} tail from 0.064 to 0.015 and the true cost-to-go gap of the surrogate-chosen action from 0.049 to 0.006.

1 Introduction

Goal-conditioned reinforcement learning (GCRL) supports agents that reach designated states by consuming a goal-conditioned value function $V^\star(s, g)$ on a state-transition graph (Schaul et al., 2015; Andrychowicz et al., 2017; Eysenbach et al., 2019). In practice this value function is observed at only a small subset of state-goal pairs — a sparse boundary $\Gamma_g \subset V$ of cost-to-go labels — and must be extended to unlabelled interior vertices before the agent can plan. Many extension principles are available: harmonic averaging (Zhu et al., 2003), smoothness-regularised graph p-Laplacians (Kyng et al., 2015; Calder, 2018), and Lipschitz-extremal completions (Aronsson, 1967; Peres et al., 2009). Existing analyses of value approximation justify greedy policies through action-value ordering or weighted L^p loss bounds (Singh & Yee, 1994; Munos, 2003), but graph semi-supervised learning objectives optimise extension smoothness or pointwise fit. Our criterion is the induced graph planner: which extension principle preserves the local action ordering on which one-step decisions depend?

We frame this as a graph partial differential equation (graph-PDE) Dirichlet extension problem. The state-transition graph $G = (V, E, w)$ has positive edge weights $w(s, y)$ encoding transition costs; the boundary $\Gamma_g \subset V$ carries the observed labels with $g \in \Gamma_g$ as the goal; the surrogate $\hat{V}$ extends these labels to the interior $V \setminus \Gamma_g$. The planner used throughout this paper, matching the graph-search rule of Search-on-Replay-Buffer-style methods (Eysenbach et al., 2019), is the operational argmin-Q rule

$$T_{\hat{V},g}(s) \;=\; \arg\min_{y \sim s} \big[w(s, y) + \hat{V}(y, g)\big],$$

with consistent tie-breaking and absorbing g. Within the graph p-Laplacian family, the natural endpoints are harmonic averaging at $p = 2$ and the Absolutely Minimal Lipschitz Extension (AMLE) at $p = \infty$, with intermediate p interpolating between them; the question is which members of the family preserve the action ordering well enough for the argmin-Q rollout to reach the goal from interior starts.

Our main theoretical result is a **planner-admissibility certificate** (Theorem 3.3): if local surrogate error stays below half the true action gap along the realised rollout, the greedy rollout reaches the goal. AMLE instantiates the certificate via a fill-distance comparison bound (Corollary 3.5); harmonic can violate it because its rankings are harmonic-measure rankings rather than shortest-path rankings (Lemma 3.9). All p-Laplacian endpoints share a structural maximum principle (Lemma 3.6): no strict interior local extrema across $p \in [2, \infty]$. The AMLE-vs-harmonic separator therefore lives in the local action-ordering layer.

Empirically, on 120 paired AntMaze layout-derived configurations (Fu et al., 2020) (61,440 rollouts per method), aggregate rollout success is 0.584 ± 0.230 for $p = 2$ and $\mathbf{0.970} \pm \mathbf{0.061}$ for $p = \infty$, a $+\mathbf{38.6}$ pp paired lift (95% run-bootstrap CI $[+34.9, +42.3]$). Intermediate $p \in \{4, 8, 16\}$ cluster with AMLE in the 0.90–0.98 regime, with the few-pp $p = 16$ vs $p = \infty$ lift solver-tolerance-bound rather than a converged endpoint ranking. The mechanism audit in Section 5 localises the gap to operator geometry: 54.8% of rollout decisions sit in geometry where the true distance d_g is locally AMLE-compatible but harmonic-incompatible (Proposition 3.10 and Corollary 3.11); 99.3% of harmonic inversions concentrate there, AMLE corrects 93.1% of inversions, and the per-state half-gap test of Lemma 3.2 (the local hypothesis underlying Theorem 3.3) certifies 67.0% of inversion events. The ordering and mechanism audits are decision-scope diagnostics, not direct upper bounds on $\phi(\hat{V})$.

The local mechanism is illustrated by a seven-node graph G_7 (Example 3.12) with sparse-label boundary $\Gamma_g = \{0, 7\}$ and labels $Y_g(0) = 0$, $Y_g(7) = 3$. At decision state $s = 4$ the true greedy neighbour is 3, but harmonic assigns $\hat{u}_2(1) = 36/29 < \hat{u}_2(3) = 39/29$ (*wrong*), whereas AMLE assigns $\hat{V}_\infty(3) = 1 < \hat{V}_\infty(1) = 4/3$ (*correct*); the inversion persists under uniform graph subdivision (Corollary B.5). An adversarial 4×4 subgraph search complements this construction with mirror cases where harmonic is greedy-perfect and AMLE degenerates on a plateau, so the seven-node example illustrates a mechanism rather than a global comparison.

Contributions. We formulate sparse GCRL value completion as a planner-admissible graph-PDE extension problem and read the surrogate choice as a p-family geometry question. Specifically:

- We prove a local action-gap certificate (Theorem 3.3) and its AMLE fill-distance instantiation (Corollary 3.5) translating sparse-label density into greedy-rollout success.
- We establish harmonic anti-admissibility through harmonic-measure rankings (Lemma 3.9) and an operator-level mismatch with shortest-path distance (Proposition 3.10, Corollary 3.11).
- We validate the resulting p-family transition on 120 AntMaze graph layouts via the main rollout experiment, the ordering audit, the mechanism audit, and the failure-mode decomposition.
- Adaptive label selection, closed-loop continuous control, and certified intermediate-p solvers are open directions deferred to §6.

2 Related work and positioning

Harmonic / Laplacian SSL and low-label degeneracy. The classical harmonic SSL approach fits a function agreeing with labels on a boundary set and minimising quadratic Dirichlet energy on the interior (Zhu

et al., 2003; Zhou et al., 2003), generalised by manifold and graph-Laplacian regularisation (Belkin & Niyogi, 2003; Belkin et al., 2006; Bengio et al., 2006; Smola & Kondor, 2003). At very sparse labellings the harmonic minimiser is known to degenerate to a near-constant spike pattern (Nadler et al., 2009); properly-weighted graph Laplacians restore well-posedness (Calder & Slepčev, 2020; Calder et al., 2023). Harmonic is the $p = 2$ endpoint of our p-family; its low-label degeneracy is one structural reason planner admissibility should not be checked through pointwise harmonic accuracy.

AMLE, p-Laplacian, and Lipschitz learning. The absolutely minimising Lipschitz extension (AMLE) of Aronsson (1967) coincides in the continuum with the viscosity solution of the infinity-Laplace equation, made precise via tug-of-war games (Peres et al., 2009; Sheffield & Smart, 2012; Manfredi et al., 2012) and characterised by comparison principles (Jensen, 1993; Juutinen & Shanmugalingam, 2006; Le Gruyer, 2007); convergent numerics: Oberman (2005). The graph p-Laplacian interpolates harmonic ($p = 2$) and AMLE ($p = \infty$); efficient algorithms, game-theoretic unification, and consistency / convergence-rate analyses are surveyed in (Kyng et al., 2015; Calder, 2018; Bungert et al., 2023; Elmoataz et al., 2017; El Alaoui et al., 2016; Slepčev & Thorpe, 2019; Calder, 2019; Roith & Bungert, 2023; García Trillos & Murray, 2020). Provably convergent fast ℓ_p-regression solvers for intermediate p (Adil et al., 2019; 2024) are relevant to §6.

Eikonal viscosity solutions on networks. Discrete eikonal / shortest-path solvers include fast marching (Sethian, 1996), Dijkstra-style discretisations (Tsitsiklis, 1995), and heat-flow alternatives (Crane et al., 2013); viscosity theory on metric graphs is developed in (Schieborn & Camilli, 2013; Camilli & Marchi, 2013; Achdou et al., 2013). Proposition 3.10 is a distinct statement: shortest-path distance also satisfies the graph-AMLE midrange identity on geodesically extendable interiors. We do *not* claim that the sparse-label AMLE extension is the network eikonal solution.

GCRL and positioning. GCRL value learning (Schaul et al., 2015; Andrychowicz et al., 2017; Pong et al., 2018) on D4RL AntMaze (Fu et al., 2020) combines with planning-aware architectures (Tamar et al., 2016; Eysenbach et al., 2019; Nasiriany et al., 2019) and spectral / Laplacian viewpoints (Mahadevan & Maggioni, 2007; Dayan, 1993; Stachenfeld et al., 2017). Existing graph-PDE / Lipschitz work studies interpolation, regularisation, and continuum limits but not which extension avoids greedy-planner failure modes; existing GCRL work does not study sparse graph-label completion. The theme that value error matters for greedy control only through action ordering goes back (Singh & Yee, 1994; Williams & Baird, 1993; Munos, 2003; 2007; Farahmand et al., 2010); we sit at this intersection. Our planner-admissibility theory selects within the graph p-Laplacian family via a local action-margin certificate (Theorem 3.3), with the local action-ordering and operator-compatibility properties (Lemma 3.9, Proposition 3.10) as the separating axes; the family shares a structural maximum principle (Lemma 3.6). We test on D4RL AntMaze *layout graphs*; closed-loop continuous control is deferred (§6).

3 Planner admissibility under the GCRL argmin-Q planner

We develop the *planner-admissibility certificate* (§3.2): a local action-margin condition along the realised rollout sufficient for the $\hat{V}$-greedy planner to reach the goal. It is the local graph-planning analogue of classical value-error-to-greedy-policy bounds (Williams & Baird, 1993; Singh & Yee, 1994; Munos, 2003), with the global Bellman-residual norm replaced by per-state local error ϵ_s and the discount-factor condition by a one-step action gap $\Delta_s^\star$. Section 3.3 instantiates the certificate for AMLE via a sparse-label fill-distance bound; the remaining subsections explain how harmonic *can* violate the certificate in planner-relevant local decisions through harmonic-measure rather than shortest-path rankings, while both endpoints share an interior maximum principle.

3.1 Operational setup

Let $G = (V, E, w)$ be a finite, connected, undirected graph with positive symmetric edge weights, $\partial \subset V$ a non-empty boundary, and $g \in \partial$ a goal vertex. The true cost-to-go $V^\star(\cdot, g) : V \to \mathbb{R}_{\geq 0}$ is the graph

shortest-path cost-to-go: $V^\star(g,g) = 0$ and, for every $s \in V \setminus \{g\}$,

$$V^\star(s,g) \;=\; \min_{y:y\sim s}\big[w(s,y) + V^\star(y,g)\big] \tag{1}$$

(the Bellman / shortest-path identity used in the proofs below). We assume $V^\star$ is known on ∂, and a surrogate $\hat{V} : V \to \mathbb{R}$ extends the boundary condition $\hat{V}|_\partial = V^\star|_\partial$ to the interior.

We adopt the operational argmin-Q planner of the experiments (§5). Define the surrogate action value $Q_{\hat{V},g}(s,y) := w(s,y) + \hat{V}(y,g)$ on each edge $\{s,y\} \in E$, and the deterministic successor

$$T_{\hat{V},g}(s) \;:=\; \arg\min_{y:y\sim s} Q_{\hat{V},g}(s,y),$$

with a fixed consistent tie-breaking rule (lexicographic on vertex indices throughout this paper). The goal g is absorbing: $T_{\hat{V},g}(g) := g$. Other boundary vertices are *not* absorbing — the planner continues from any $z \in \partial \setminus \{g\}$ along the same rule. Starting from $s_0 \in V$, the rollout is the orbit $s_{t+1} = T_{\hat{V},g}(s_t)$ for $t \geq 0$. Since V is finite and $T_{\hat{V},g}$ is deterministic, the orbit eventually enters a finite set $\rho(s_0) \subseteq V$ that is either $\{g\}$ (operational success) or a non-goal directed cycle (operational failure).

The *operational failure rate* is $\phi(\hat{V}) := |\{s_0 \in V : \rho(s_0) \neq \{g\}\}|/|V \setminus \{g\}|$.

Lemma 3.1 (Operational basin partition)**.** *Under the deterministic argmin-Q planner above on a finite connected graph, every $s_0 \in V$ has $\rho(s_0)$ equal to $\{g\}$ or to a directed cycle in $V \setminus \{g\}$. Equivalently, V partitions into the goal basin $\{s_0 : \rho(s_0) = \{g\}\}$ and the operational failure set $F(\hat{V}) := \{s_0 : \rho(s_0) \neq \{g\}\}$.*

Proof. The forward orbit on a finite state space with deterministic $T_{\hat{V},g}$ eventually repeats; with g absorbing, the limit set is either $\{g\}$ or a cycle in $V \setminus \{g\}$. □

In Section 5, the main rollout experiment records per-start outcomes as `reached` ($\rho(s_0) = \{g\}$) or `loop` ($s_0 \in F(\hat{V})$); the `loop` share equals $\phi(\hat{V})$.

3.2 The planner-admissibility certificate

We first state the certificate for an arbitrary boundary-pinned surrogate $\hat{V}$; §3.3 then lets the goal-specific sparse label set Γ_g play the role of ∂ and specialises the surrogate to the AMLE extension. The certificate is built from three local ingredients at each non-goal state s: the true Bellman-optimal neighbour set $A^\star(s,g)$, the true local action gap $\Delta_s^\star$, and the surrogate local error ϵ_s.

Write $N(s) := \{y \in V : y \sim s\}$ for the neighbour set of s and $d(s) := |N(s)|$ for the vertex degree, and write the true one-step action value as

$$Q^\star(s,y;g) \;:=\; w(s,y) + V^\star(y,g).$$

Define the true Bellman-optimal neighbour set at each non-goal state s:

$$A^\star(s,g) \;:=\; \arg\min_{y\sim s} Q^\star(s,y;g),$$

and the true *action gap*

$$\Delta_s^\star \;:=\; \begin{cases} \min\limits_{b\in N(s)\setminus A^\star(s,g)} \big[Q^\star(s,b;g) - \min\limits_{a\in A^\star(s,g)} Q^\star(s,a;g)\big], & N(s)\setminus A^\star(s,g) \neq \emptyset, \\ +\infty, & N(s) = A^\star(s,g). \end{cases}$$

(The $+\infty$ branch is the trivial case where every neighbour is optimal; the certificate then holds vacuously.) For a surrogate $\hat{V}$, the local error on $N(s)$ is

$$\epsilon_s \;:=\; \max_{y\sim s}\big|\hat{V}(y,g) - V^\star(y,g)\big|.$$

Edge costs cancel in the argmin-Q comparison, so $|Q_{\hat{V}}(s,y) - Q^\star(s,y;g)| = |\hat{V}(y,g) - V^\star(y,g)| \le \epsilon_s$: the value-level local error ϵ_s directly controls Q-rankings at s. This cancellation is why the certificate below is a single state-level condition on $\hat{V}$ rather than on $Q_{\hat{V}}$.

The local-error / action-gap pair $(\epsilon_s, \Delta_s^\star)$ enters the theory through a single-state action-ordering lemma; its top-1 conclusion is the theorem-critical part, and a pairwise-inversion bound is a diagnostic by-product used by the ordering-audit table in Section 5.

Lemma 3.2 (Local action-ordering preservation)**.** *At every non-goal $s \in V \setminus \{g\}$, if $\epsilon_s < \Delta_s^\star/2$ then*

$$\arg\min_{y\sim s} Q_{\hat{V},g}(s,y) \;\subseteq\; A^\star(s,g).$$

For the Kendall-τ diagnostic statement that follows, assume in addition that $d(s) \ge 2$ and fix a deterministic tie-breaking convention for true $Q^\star$-ties (tied pairs are counted in $M_{s,g}(0)$). The top-1 set-admissibility conclusion above does not *require tie-free true actions. The neighbour-Kendall-τ at (s,g) defined by $\tau_{\mathrm{nbr}}(s,g) := 1 - \frac{4\,I_{s,g}}{d(s)(d(s)-1)}$ with $I_{s,g}$ the count of pairwise inversions of the surrogate ordering relative to the true ordering also satisfies the gap-dependent bound*

$$\tau_{\mathrm{nbr}}(s,g) \;\ge\; 1 - \frac{4\,M_{s,g}(2\epsilon_s)}{d(s)(d(s)-1)},$$

where $M_{s,g}(\eta) := \#\{(i,j) : i<j, |Q_i^\star - Q_j^\star| \le \eta\}$ counts small-gap pairs of neighbours.

Proof sketch. For $a \in A^\star(s,g)$, $b \in N(s) \setminus A^\star(s,g)$, the true gap is at least $\Delta_s^\star$ while the surrogate perturbs each Q-value by at most ϵ_s, so $Q_{\hat{V},g}(s,b) - Q_{\hat{V},g}(s,a) \ge \Delta_s^\star - 2\epsilon_s > 0$. Hence $\arg\min_y Q_{\hat{V},g}(s,y) \subseteq A^\star(s,g)$. The same gap argument applied pair-by-pair yields the Kendall-τ bound. Full proof in Appendix B. □

Single-state action-ordering lifts to a rollout-level admissibility certificate.

Theorem 3.3 (Planner-admissibility certificate)**.** *Consider the deterministic $\hat{V}$-greedy rollout $s_0, s_1, \ldots$ on G with positive edge costs and absorbing g. If for every non-goal state s_t visited before the rollout either reaches g or repeats a non-goal state,*

$$\epsilon_{s_t} \;<\; \Delta_{s_t}^\star/2,$$

then $T_{\hat{V},g}(s_t) \in A^\star(s_t,g)$ at every such step, $V^\star(s_t,g)$ strictly decreases along the rollout, and the rollout reaches g in finitely many steps.

Proof sketch. Lemma 3.2 at each s_t gives $s_{t+1} \in A^\star(s_t,g)$, so $V^\star(s_{t+1},g) = V^\star(s_t,g) - w(s_t,s_{t+1})$ strictly decreases; with positive edge costs, the strict-descent sequence reaches $V^\star = 0$ (i.e., g) in at most $\lceil V^\star(s_0,g)/w_{\min}\rceil$ steps, where $w_{\min} := \min_{\{s,y\}\in E} w(s,y) > 0$. Finiteness excludes non-goal cycles. □

Theorem 3.3 is the central operational certificate: *rollout success follows from a local action-margin condition along the actual trajectory.* Its single-state version Lemma 3.2 is the verification tool used to certify the AMLE endpoint and to flag harmonic failures in subsequent subsections.

The ordering audit (Kendall-τ) and mechanism audit (local action margin) are decision-scope diagnostics for the certificate's local ingredients, not direct upper bounds on $\phi(\hat{V})$ — which the main rollout experiment measures directly. For a start distribution μ on V, writing $\phi_\mu(\hat{V}) := \Pr_{s_0\sim\mu}[\rho(s_0) \ne \{g\}]$ for the corresponding failure probability, the path-existential inequality $\phi_\mu(\hat{V}) \le \Pr_{s_0\sim\mu}[\exists s_t : \epsilon_{s_t} \ge \Delta_{s_t}^\star/2]$ and a bad-tail mass diagnostic (Corollary B.4) are recorded in Appendix B.

The certificate's content is now fully laid out at the operational level. We turn next to its first concrete instantiation: the AMLE endpoint of the graph p-Laplacian family, where a comparison-principle fill-distance bound controls ϵ_s directly from sparse-label density.

3.3 AMLE instantiation of the certificate

We specialise to unit-cost graphs ($w \equiv 1$) with boundary the goal-specific labelled set $\Gamma_g \subset V$, $g \in \Gamma_g$, and observed labels $Y_g : \Gamma_g \to \mathbb{R}$ satisfying $\|Y_g - V^\star(\cdot, g)\|_{\infty,\Gamma_g} \le \epsilon_{\text{lab}}$. The AMLE extension $\hat{V}_\infty := \mathcal{A}_{\Gamma_g}(Y_g)$ is the unique boundary-pinned fixed point of the midrange operator $\mathcal{A}[u](x) := \frac{1}{2}(\min_{y\sim x} u(y) + \max_{y\sim x} u(y))$ on $V \setminus \Gamma_g$ (Sheffield & Smart, 2012; Kyng et al., 2015; Aronsson, 1967; Jensen, 1993; Juutinen & Shanmugalingam, 2006); it controls extremal slopes rather than averaging values. Under $w \equiv 1$ the planner reduces to $T_{\hat{V}_\infty,g}(s) = \arg\min_{y\sim s} \hat{V}_\infty(y)$, so the admissibility conditions of §3.2 are value-level conditions on $\hat{V}_\infty$ alone.

Lemma 3.4 (AMLE local extension error)**.** *Let $A \subseteq V \setminus \Gamma_g$. Define the* AMLE *residual of the true cost-to-go on A:*

$$\eta_\infty(V^\star; A, \Gamma_g) \;:=\; \left\|\mathcal{A}_{\Gamma_g}(V^\star(\cdot, g)|_{\Gamma_g}) - V^\star(\cdot, g)\right\|_{\infty,A}.$$

Then

$$\left\|\hat{V}_\infty(\cdot, g) - V^\star(\cdot, g)\right\|_{\infty,A} \;\le\; \epsilon_{\text{lab}} + \eta_\infty(V^\star; A, \Gamma_g).$$

If, in addition, $V^\star(\cdot, g)$ is L_g-Lipschitz on $A \cup \Gamma_g$ with fill distance $h_{\Gamma_g}(A) := \max_{x\in A} \min_{z\in\Gamma_g} d_G(x, z)$, then $\eta_\infty(V^\star; A, \Gamma_g) \le 2L_g\, h_{\Gamma_g}(A)$, and in particular

$$\left\|\hat{V}_\infty - V^\star\right\|_{\infty,A} \;\le\; \epsilon_{\text{lab}} + 2L_g\, h_{\Gamma_g}(A).$$

Proof. See Appendix B (proof of Lemma 3.4). □

Combining Lemma 3.4 (with $A = N(s_t)$) and Theorem 3.3 yields the AMLE admissibility certificate.

Corollary 3.5 (AMLE planner-admissibility certificate)**.** *If along the* AMLE*-greedy rollout from s_0,*

$$\epsilon_{\text{lab}} + 2L_g\, h_{\Gamma_g}(N(s_t)) \;<\; \Delta^\star_{s_t}/2$$

for every non-goal state s_t visited before termination, then the AMLE*-greedy rollout from s_0 reaches g.*

Proof. Lemma 3.4 applied with $A = N(s_t)$ gives $\epsilon^{\text{AMLE}}_{s_t} \le \epsilon_{\text{lab}} + 2L_g h_{\Gamma_g}(N(s_t))$, which under the hypothesis is strictly below $\Delta^\star_{s_t}/2$. Theorem 3.3 concludes. □

Corollary 3.5 is the *sparse-label-density-to-greedy-correctness* chain for AMLE: small enough $\epsilon_{\text{lab}} + 2L_g h_{\Gamma_g}$ relative to the local action gap along the rollout implies success. The bound is per-state along the realised rollout; a global-uniform variant (Corollary B.1) is in Appendix B. The harmonic certificate is not in general satisfied. We first record the p-Laplacian maximum principle shared by all endpoints (§3.4), then develop the harmonic-side mechanism that violates the certificate (§3.5) and the operator-level explanation (§3.6).

3.4 Maximum principle for the p-Laplacian family

Exact p-Laplacian Dirichlet solutions admit no strict interior local extrema, a structural property shared across $p \in [2, \infty]$ and inherited in particular by both harmonic and AMLE (Lemma 3.6).

Define $M(u) := \{m \in V \setminus \Gamma_g : u(m) < u(s') \text{ for every } s' \sim m\}$ as the strict-interior-min set of any $u : V \to \mathbb{R}$.

Lemma 3.6 (p-Laplacian maximum principle: no strict interior extrema)**.** *Let $u : V \to \mathbb{R}$ be an exact interior solution, with pinned Dirichlet boundary $u|_{\Gamma_g} = Y_g$, of any one of the three graph p-Laplacian completion problems:*

(i) harmonic, $p = 2$: $\sum_{y\sim x}(u(y) - u(x)) = 0$;

(ii) finite graph p-Laplacian, $2 < p < \infty$: $\sum_{y\sim x} |u(y) - u(x)|^{p-2}(u(y) - u(x)) = 0$;

(iii) AMLE *$/\, p = \infty$: $u(x) = \frac{1}{2}\big(\min_{y\sim x} u(y) + \max_{y\sim x} u(y)\big)$.*

Then u has no strict interior local minimum and no strict interior local maximum. In particular $M(u) = \emptyset$.

Proof. See Appendix B (proof of Lemma 3.6). □

Remark 3.7 (Operational consequence: interior cycles are plateau cycles)**.** Let u be an exact solution of any of the three Dirichlet problems in Lemma 3.6 on a finite connected unweighted graph G. At every interior $s \in V \setminus \Gamma_g$ there exists $y \sim s$ with $u(y) \le u(s)$ (else $s \in M(u)$), so the argmin-Q planner satisfies $u(T_{u,g}(s)) \le u(s)$: u is monotone non-increasing along interior rollout segments. A directed cycle contained in $V \setminus \Gamma_g$ therefore forces u to be constant on the cycle (plateau cycle), and no strict-decrease interior cycle exists for any $p \in \{2\} \cup (2, \infty) \cup \{\infty\}$. Cycles touching $\Gamma_g \setminus \{g\}$ are not controlled by this interior argument — pinned boundary values are unaffected by the max principle, so $T_{u,g}(z)$ for $z \in \Gamma_g \setminus \{g\}$ may increase u — and are separately tracked by the *boundary-touching cycle basin*

$$C^{\Gamma}(\hat{V}) \;:=\; \{s_0 \in V : \rho(s_0) \cap (\Gamma_g \setminus \{g\}) \neq \emptyset\} \;\subseteq\; F(\hat{V}),$$

the sub-class of operational failures whose limit cycle passes through some labelled non-goal vertex. The empirical failure-mode decomposition of §5.4 shows that this boundary-touching subclass is aggregate-dominant under both surrogates. The family separator therefore lies on the local-ordering and operator-compatibility axes of subsequent subsections.

We next turn to the local-ordering mechanism by which harmonic can violate the certificate in planner-relevant decisions: boundary-label-weighted hitting probabilities (§3.5), with the operator-level reason (§3.6) that shortest-path distance is locally compatible with the AMLE midrange but not with harmonic averaging unless degree-balanced.

Finite-sweep refinement. For finite-iteration AMLE solvers, a residual–margin bound (Lemma B.3 in Appendix B) controls the strict-min depth at any approximate interior local minimum by the midrange residual; this is the AMLE-side numerical refinement used by the AMLE iteration audit.

3.5 Harmonic anti-admissibility certificate

The harmonic endpoint admits the classical random-walk representation from harmonic graph SSL (Zhu et al., 2003), which locates where its neighbour ranking can disagree with the shortest-path ranking.

Lemma 3.8 (Harmonic-measure representation of neighbour rankings)**.** *Let h be the harmonic extension of $y : \partial \to \mathbb{R}$ on a connected unweighted graph. With (X_t) the simple random walk and $\omega_x(z) := \mathbb{P}_x(X_{\sigma_\partial} = z)$ the harmonic measure (where $\sigma_\partial := \inf\{t \ge 0 : X_t \in \partial\}$ is the boundary-hitting time; optional stopping gives $h(x) = \sum_z \omega_x(z) y_z$), for any decision state s with neighbours $a, b \sim s$,*

$$h(a) - h(b) \;=\; \sum_{z \in \partial} \big(\omega_a(z) - \omega_b(z)\big) y_z.$$

For the goal-specific labelled boundary Γ_g with labels Y_g, write $\hat{u}_2$ for the harmonic ($p = 2$) extension of Y_g to V and $\omega_v(z) := \mathbb{P}_v(X_{\sigma_{\Gamma_g}} = z)$ for the harmonic measure from v on Γ_g (the instantiation of Lemma 3.8 with $\partial = \Gamma_g$).

Lemma 3.9 (Harmonic local anti-admissibility certificate)**.** *Fix a decision state s with $A^\star(s, g) \subseteq N(s)$ the true Bellman-optimal neighbour set, and let $b \in N(s) \setminus A^\star(s, g)$ be any strictly suboptimal competitor. The harmonic-greedy planner at s strictly prefers b over every $a \in A^\star(s, g)$ if and only if*

$$w(s, b) + \hat{u}_2(b) \;<\; w(s, a) + \hat{u}_2(a) \qquad \text{for every } a \in A^\star(s, g),$$

equivalently, via Lemma 3.8,

$$w(s, b) - w(s, a) \;+\; \sum_{z \in \Gamma_g} \big(\omega_b(z) - \omega_a(z)\big)\, Y_g(z) \;<\; 0 \qquad \text{for every } a \in A^\star(s, g).$$

On unit-cost graphs ($w \equiv 1$), this reduces to $\sum_z (\omega_b(z) - \omega_a(z)) Y_g(z) < 0$ for every $a \in A^\star(s, g)$.

Proof. Direct: $Q_{\hat{u}_2,g}(s,b) < \min_{a\in A^\star} Q_{\hat{u}_2,g}(s,a)$ is the displayed inequality; substitute $\hat{u}_2(v) = \sum_z \omega_v(z) Y_g(z)$. □

Harmonic ranking at s is determined by differences in boundary *hitting* probabilities against boundary labels — a random-walk averaging ordering, not a shortest-path ordering. A truly shorter branch may have larger hitting probability on a high-cost label and so be ranked higher; harmonic averaging therefore does not by itself guarantee planner-admissible local ordering. AMLE controls extremal slopes rather than averaging hitting probabilities, which is what enables the local sup-error chain through Lemma 3.4 and Corollary 3.5.

3.6 Operator-level compatibility with shortest-path distance

Write $d_g(x) := d_G(x,g)$ for the *unweighted* shortest-path distance (this subsection works on the unit-cost graph metric, consistent with the AMLE-instantiation scope of §3.3), and call x *geodesically extendable away from* g if some $y_+ \sim x$ has $d_g(y_+) = d_g(x) + 1$; let V_g^{ext} denote the set of such vertices.

Proposition 3.10 (Distance is graph-AMLE on the extendable interior)**.** *For any $x \in V$ with $x \neq g$:*

(a) $\min_{y\sim x} d_g(y) = d_g(x) - 1$*;*

(b) $\max_{y\sim x} d_g(y) \in \{d_g(x)-1, d_g(x), d_g(x)+1\}$*;*

(c) d_g *satisfies the* AMLE *midrange identity* $d_g(x) = \mathcal{A}[d_g](x)$ *if and only if* $\max_{y\sim x} d_g(y) = d_g(x)+1$*, equivalently iff* $x \in V_g^{\text{ext}}$*.*

*In particular, at every $x \in V_g^{\text{ext}}$, d_g satisfies both the discrete Bellman / eikonal identity $d_g(x) = 1 + \min_{y\sim x} d_g(y)$ and the graph-*AMLE *midrange identity $d_g(x) = \mathcal{A}[d_g](x)$.*

Proof. See Appendix B (proof of Proposition 3.10). □

Proposition 3.10(c) is a fixed-point statement about d_g itself; the sparse-label AMLE extension is a separate object with approximation error controlled by Lemma 3.4, and the global Bellman / eikonal identity $d_g(x) = 1 + \min_y d_g(y)$ holds at every $x \neq g$ on all of V. The AMLE midrange compatibility of d_g on V_g^{ext} should not be identified with solving the full eikonal / shortest-path problem from the sparse-label data.

Corollary 3.11 (Harmonic residual of the shortest-path distance)**.** *Partition $N(x) = N_+(x) \sqcup N_0(x) \sqcup N_-(x)$ by $d_g(y) - d_g(x) \in \{+1, 0, -1\}$ with sizes $n_\sigma(x)$. Then*

$$\tfrac{1}{\deg(x)} \sum_{y\sim x} d_g(y) - d_g(x) \;=\; \tfrac{n_+(x) - n_-(x)}{\deg(x)}.$$

So d_g satisfies the harmonic-averaging identity at x iff $n_+(x) = n_-(x)$, whereas by Proposition 3.10 it satisfies the AMLE *midrange identity iff $n_+(x) \geq 1$.*

Proof. By Proposition 3.10(a)-(b), $d_g(y) - d_g(x) \in \{-1, 0, +1\}$ for every $y \sim x$, so the trichotomy partition $N(x) = N_+ \sqcup N_0 \sqcup N_-$ is exhaustive. Summing $d_g(y) - d_g(x)$ over $y \sim x$ gives $(+1)n_+(x) + (0)n_0(x) + (-1)n_-(x) = n_+(x) - n_-(x)$; dividing by $\deg(x)$ yields the identity. □

The compatibility conditions differ in scope: AMLE asks only that some outward geodesic continuation exists through x, whereas harmonic requires a non-generic degree-balance $n_+(x) = n_-(x)$ that fails at junctions, wall-adjacent vertices, and asymmetric branching points. Where it fails, harmonic pulls d_g away from its true value by a residual of order $|n_+ - n_-|/\deg(x)$ toward the over-represented neighbour class; together with Lemma 3.8, this is the structural source of the AMLE-vs-harmonic difference on sparse-label graph-PDE planning.

3.7 Combined local separation and mechanism witness

The two certificates combine into a local separation statement (Proposition B.2 in Appendix B): on decision states where the AMLE local-admissibility condition $\epsilon_s^{\mathrm{AMLE}} < \Delta_s^\star/2$ and the harmonic anti-admissibility inequality both hold, AMLE-greedy is correct and harmonic-greedy is wrong. The mechanism audit reports the rollout-weighted firing rates: harmonic inversion 5.6%, AMLE correction of inversions 93.1%, and certified correction (clause (i) fires) 67.0%. The 93.1%–67.0% gap is AMLE-correctness via mechanisms outside the sufficient local-error condition — e.g. the Proposition 3.10 exact-distance match on V_g^{ext}.

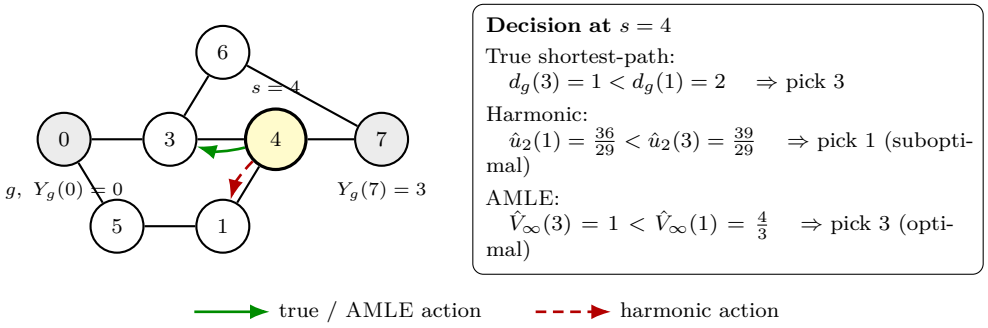

Figure 1: Seven-node mechanism witness. At $s = 4$, the true shortest-path action and AMLE choose neighbour 3, while harmonic chooses neighbour 1. The right panel gives the local inequalities that determine the choices. This is a positive-margin mechanism witness, not a universal-dominance claim.

Example 3.12 (Harmonic-measure mismatch on a 7-node graph)**.** On the seven-node graph G_7 defined in Appendix B (goal $g = 0$, sparse-label boundary $\Gamma_g = \{0, 7\}$, labels $Y_g(0) = 0$, $Y_g(7) = 3$; Figure 1), at decision state $s = 4$ with $A^\star(4, g) = \{3\}$ and gap $\Delta_4^\star = 1$:

$$\hat{u}_2(1) = \tfrac{36}{29} < \hat{u}_2(3) = \tfrac{39}{29} \quad \text{(harmonic, } \textit{wrong}\text{)}, \qquad \hat{V}_\infty(3) = 1 < \hat{V}_\infty(1) = \tfrac{4}{3} \quad \text{(AMLE, } \textit{correct}\text{)}.$$

Harmonic realises clause (ii) of Proposition B.2 with harmonic-measure values $\omega_1(7) - \omega_3(7) = -1/29$. At $s = 4$, AMLE is correct through the operator-level exact match $\hat{V}_\infty(3) = V^\star(3, g) = 1$ of Proposition 3.10 ($3 \in V_g^{\mathrm{ext}}$) rather than through clause (i)'s sufficient local-error condition, which gives $\epsilon_4^{\mathrm{AMLE}} = 2/3 > 1/2 = \Delta_4^\star/2$. This matches the mechanism audit in Section 5: the local-error test certifies 67.0% of harmonic-inversion events, while AMLE corrects 93.1% overall — a ~ 26 pp gap consisting of AMLE corrections via operator-level mechanisms outside the sufficient local-error condition. Full algebra and the formal mechanism-scope remark are in Appendix B.

The inversion of Example 3.12 persists under uniform k-subdivision with $\lambda_k = k$ (Corollary B.5 in Appendix B; the analogous AMLE statement is Lemma B.6 there), so refining the graph does not rescue harmonic. An adversarial 4×4 subgraph search (Appendix A.10) also exhibits mirror cases where harmonic is greedy-perfect and AMLE degenerates on interior plateaus adjacent to a high-cost boundary vertex; the distributional question of which surrogate wins on a graph family is settled empirically by the §5 aggregate.

The certificate framework and its two-sided mechanism reading are now in place. Section 4 catalogues the canonical graph-PDE surrogates used in the experiments under this admissibility lens, situating harmonic and AMLE as the $p = 2$ and $p = \infty$ endpoints of a larger family.

4 Graph-PDE surrogate family

Section 3 cast greedy-rollout failure as a non-goal directed cycle in the argmin-Q successor map (Lemma 3.1) and identified planner admissibility as the design criterion for sparse-label value extension. We now situate each member of the graph-PDE surrogate family under this lens. Every exact p-Laplacian extension satisfies the maximum principle of §3.4 (Lemma 3.6): $M(\hat{u}) = \emptyset$, and interior operational cycles are uniformly plateau cycles (Remark 3.7; boundary-touching cycles tracked separately by C^Γ). The differentiating axes

— local action-ordering preservation (Lemma 3.2), operator-level compatibility with shortest-path distance (Proposition 3.10), and the harmonic-measure mismatch criterion (Lemma 3.8) — separate the family members and drive the planner-admissibility verdicts in Table 1. The eikonal row is included as an *exact-distance / oracle-adjacent reference*, not as a sparse-label completion method: it requires full graph and goal-boundary knowledge, and Proposition 3.10 clarifies the local AMLE-vs-eikonal relationship without identifying the sparse-label AMLE extension with the full eikonal solve.

Table 1: Planner-admissibility taxonomy for the graph-PDE surrogates studied in this paper. Operational rollout (Lemma 3.1): every start either reaches g or enters a non-goal cycle in $F(\hat{V})$. Maximum-principle column: $M(\hat{V}) = \emptyset$ implies every interior cycle is a plateau cycle (boundary-touching cycles tracked by C^{Γ}, Remark 3.7). Algorithm column: standard solver class.

p	Equation	Admissibility status	Algorithm	Empirical
$p = 2$ (harmonic)	$\Delta_2 u = 0$	$M = \emptyset$ via Lem. 3.6; plateau cycles in F unconstrained; anti-admissibility via Lem. 3.9	$\tilde{O}(\|E\|)$ Laplacian solver	§5 baseline, 0.584
$2 < p < \infty$	$\Delta_p u = 0$	$M = \emptyset$ via Lem. 3.6; no per-state admissibility certificate	IRLS / Picard / L-BFGS-B; no fast intermediate-p solver	finite-p family sweep
$p = \infty$ (AMLE)	midrange	$M = \emptyset$ via Lem. 3.6; finite-sweep via Lem. B.3; admissibility via Cor. 3.5	midrange fixed-point, $O(\|E\|)$ / sweep	§5 main, 0.970
eikonal (oracle)	$\|\nabla u\| = 1$	exact distance (full-graph oracle)	Dijkstra / FMM	*reference only*

Among the sparse-label methods, only AMLE admits a local planner-admissibility certificate (Corollary 3.5) under a fill-distance hypothesis; harmonic has a complementary local anti-admissibility certificate (Lemma 3.9); finite-p inherits the shared maximum principle without an analogous per-state ordinal guarantee. The empirical $p = 2 \to \infty$ progression in §5.3 reads as a geometry transition (averaging $\to$ Lipschitz-extremal), with the few-pp $p = 16$ vs $p = \infty$ lift solver-tolerance-bound rather than a converged endpoint ranking.

5 Empirical phase diagram on D4RL AntMaze graph geometry

5.1 Setup

We evaluate the planner-admissible graph-PDE family on the **D4RL AntMaze graph geometry** (Fu et al., 2020), a standard GCRL benchmark. Each AntMaze layout (medium and large) is converted to an unweighted graph G by sampling reachable cell coordinates at refinement $r \in \{4, 8, 12\}$; this gives approximate medium grids 18×24, 36×48, 54×72, and large grids 36×48, 72×96, 108×144. The labelled set $\Gamma_g \subset V$ is a sparsely sampled subset with goal- reachability values; we sweep label fraction $\mathrm{lf} \in \{0.02, 0.05, 0.08, 0.12\}$ over seeds $\{54, 55, 56, 57, 58\}$.

For each $(\mathrm{maze}, r, \mathrm{lf}, \mathrm{seed})$ configuration we fit a harmonic ($p = 2$) and an AMLE ($p = \infty$) surrogate to the observed labels and run the operational argmin-Q planner from 512 uniformly sampled start–goal evaluation pairs, for a total of 120 paired configurations and 61,440 rollouts per method. The main rollout experiment records per-pair outcomes as success at goal (`reached`) versus rollout loop failure (`loop`). A `loop` outcome is exactly the operational failure event $\rho(s_0) \neq \{g\}$ of Lemma 3.1; the main rollout experiment does not directly record per-loop cycle structure, but a reconstruction from main-rollout raw outputs (the failure-mode decomposition, Section 5.4) decomposes each loop into the interior-cycle subclass and the boundary-touching subclass C^{Γ} of Remark 3.7.

Table 2: Main rollout phase diagram, aggregated across all 120 paired $(\text{maze}, r, \text{lf}, \text{seed})$ configurations.

method	success mean ± sd	run-bootstrap 95% CI	loop share
harmonic $p = 2$	0.584 ± 0.230	$[0.543, 0.623]$	41.6%
AMLE $p = \infty$	$\mathbf{0.970 \pm 0.061}$	$[\mathbf{0.959}, \mathbf{0.980}]$	**3.0**%

5.2 Main rollout phase diagram

The paired lift is $+\mathbf{38.6} \pm \mathbf{20.9}$ pp with run-bootstrap 95% CI $[+\mathbf{34.9}, +\mathbf{42.3}]$ pp (Wilson 95% on AMLE eval-pair success $[0.969, 0.971]$). AMLE reduces the rollout-loop-failure share from 41.6% to 3.0%. This is a direct measurement of the operational failure rate $\phi(\hat{V})$ of Lemma 3.1: $\hat{V}_\infty$ leaves $\sim 3\%$ of starts in non-goal cycles, against harmonic's $\sim 42\%$. The run-bootstrap CI shows the aggregate lift is stable across resampled configurations. The C^Γ subclass of boundary-touching cycles is decomposed separately via the failure-mode decomposition (Section 5.4).

Resolution and label-fraction slices. Finer AntMaze graph extraction amplifies the harmonic gap: at $r \in \{4, 8, 12\}$ the lift is $\{+14.3, +43.7, +57.8\}$ pp (harmonic $\{0.850, 0.534, 0.368\}$ versus AMLE $\{0.993, 0.971, 0.946\}$). The harmonic gap also persists across the label-fraction sweep $\text{lf} \in \{0.02, 0.05, 0.08, 0.12\}$ with lift $\{+39.0, +44.3, +40.9, +30.1\}$ pp; AMLE saturates to 1.000 at $\text{lf} = 0.12$, consistent with Lemma 3.4's fill-distance bound. Full per-cell tables in Appendix A.1. The resolution sweep is aggregate evidence complementary to (but distinct from) the fixed-graph Corollary B.5.

Ordering audit. On the rollout-weighted decision scope, AMLE reduces the low-τ_{nbr} tail rate from 0.064 to 0.015, the mean true gap of the surrogate-chosen action from 0.049 to 0.006, and the disagreement rate from 0.025 to 0.003; top-1 agreement rises from 97.5% to 99.7%. These are decision-scope diagnostics for the Lemma 3.2 local-ordering condition on each method's rollout distribution, not upper bounds on $\phi(\hat{V})$ (which the main rollout experiment measures directly). Scope-resolved tables (`eval_rollouts` vs `all`): Appendix A.2.

5.3 PDE-family sweep and mechanism audit

Finite-p family sweep and solver audit. The finite-p rows support the family-level geometry transition: on the full 120-cell grid, rollout success is $p = 2$: 0.584; $p = 4$: 0.903; $p = 8$: 0.973; $p = 16$: 0.982; $p = \infty$: 0.970. We do not treat the few-pp $p = 16$ lift over AMLE as a converged endpoint ranking: under the finite-p sweep budget, L-BFGS-B reports convergence on 95/10/0 of 120 cells at $p = 4/8/16$, and the solver audit shows that the usable high-p solvers (L-BFGS-B and fixed-point Picard) reproduce high $p = 16$ success while disagreeing on residual / certification; a capped Newton-CG stress test fails at $p = 16$. Finite high-p methods are therefore strong empirical members of the same high-success regime, while AMLE remains the simple analyzable endpoint with no finite-p tuning parameter. Full p-family sweep, solver-audit, and p-family summary tables (including the p-family low-τ_{nbr} tail trend) are in Appendices A.4, A.7, and A.9.

Mechanism audit. Table 3 reports the rollout-weighted local decision audit on the 120-cell grid. Like the ordering audit, these are decision-scope diagnostics, not upper bounds on $\phi(\hat{V})$. More than half of visited decision scopes lie in AMLE-compatible / harmonic-incompatible local geometry ($n_+(x) \neq n_-(x)$, Cor. 3.11), and 99.3% of harmonic inversions on the AntMaze rollout distribution concentrate there. Conditional on harmonic inversions, AMLE corrects 93.1% of cases; the AMLE local certificate $\epsilon_s^{\text{AMLE}} < \Delta_s^\star/2$ fires on 67.0% of inversions, so the 93.1%–67.0% gap is the mass of AMLE corrections via mechanisms outside the sufficient local-error condition (consistent with Proposition 3.10's operator-compatibility picture). The concentration is empirical, not a structural theorem.

Table 3: Mechanism-audit rollout-weighted summary (full 120-cell AntMaze grid). Geometry classes from d_g on the greedy decision scope; inversion / correction rows conditioned on non-tied true-best decisions. *Certified correction* is $\epsilon_s^{\mathrm{AMLE}} < \Delta_s^\star/2$ firing (Lem. 3.2), as fraction of harmonic-inverted states (20,135/30,037). Per-resolution table: App. A.8.

diagnostic	value
AMLE-compatible / harmonic-incompatible local geometry	54.8%
harmonic inversion rate among non-tied decisions	5.6%
harmonic inversions in that geometry class	99.3%
AMLE correction given harmonic inversion	93.1%
certified correction given harmonic inversion	67.0%

5.4 Failure-mode decomposition

The failure-mode decomposition reruns the main rollout experiment on the same 61,440 start–goal pairs and classifies each operational failure $\rho(s_0) \neq \{g\}$ by limit-cycle structure: the interior subclass $F \setminus C^\Gamma$ (limit cycle in $V \setminus \Gamma_g$) versus the boundary-touching limit-cycle subclass C^Γ of Remark 3.7. The reconstruction's total failure count matches the main rollout experiment's `loop` count exactly on every one of the 120 configurations. The pooled split is harmonic 41.6% = 1.6% interior + 40.0% C^Γ (96.1% of failures in C^Γ) and AMLE 3.0% = 0.34% interior + 2.66% C^Γ (88.6%). C^Γ is aggregate-dominant under both surrogates, confirming the Remark 3.7 scope caveat that the p-Laplacian maximum principle (Lemma 3.6) controls interior cycles only. The interior-cycle density under AMLE is an order of magnitude below harmonic; the auxiliary strict-minimum diagnostic points in the same direction (cases with any strict interior minimum of $\hat{u}$ drop from 116/120 under harmonic to 41/120 under AMLE), but the decomposition does not by itself separate tolerance-driven strict sinks from plateau-cycle structure, so the appropriate conclusion is numerical-scope caution rather than a structural failure theorem.

Reference baselines and AMLE iteration audit. Nearest-label Voronoi on the same sparse Γ_g reaches only 0.032 ± 0.011, 93.8 pp below AMLE, so PDE-style smoothing is important in this sparse-label graph setup. An oracle Dijkstra solve with the full-graph $V^\star$ reaches 1.000 ($\sim$ 3 pp above AMLE), the full-information upper bound. The AMLE iteration audit (24 cells, $\{50, \ldots, 5000\}$ midrange sweeps) shows AMLE success $0.707 \to 1.000$ monotonically with the AMLE-vs-harmonic gap *widening* ($+30.8 \to +60.1$ pp), ruling out an under-iterated-AMLE reading. Full tables in Appendices A.5–A.6.

These results leave the open problems of §6: adaptive label selection, certified intermediate-p solvers, and extensions beyond the present unweighted graph extraction.

6 Discussion, limitations, and open problems

Scope. The evaluation (Section 5) is on the *graph extraction* of D4RL AntMaze layouts, with label fractions $\mathrm{lf} \in [0.02, 0.12]$ on unweighted unit-cost graphs. The full continuous-control setting — MuJoCo Ant dynamics, low-level locomotion, partial observability, and closed-loop replanning — is not evaluated. Within the admissibility framework itself, the planner-admissibility certificate (Theorem 3.3) is stated for general positive edge costs $w(s, y)$, but the AMLE admissibility instantiation (Corollary 3.5) is specialised to the unweighted graph setting in §3.3, and the harmonic comparison uses the unweighted graph-harmonic representation of §3.5 (Lemma 3.9); weighted-edge and directed-graph extensions of the differentiating mechanisms, and the very-sparse limit $\mathrm{lf} \to 0$, are not tested.

Open problems. The $p = 2$ Laplacian solver and $p = \infty$ midrange iteration are mature endpoints, but $p \in (2, \infty)$ sits in a solver gap; recent fast ℓ_p-regression (Adil et al., 2019; 2024) could turn the family sweep into a solver-certified endpoint comparison. Theoretical refinements of Lemma 3.2 (sharper fill-distance dependence; expected-ordinal-fidelity statements under structured action-gap distributions), adaptive sparse-label selection, and per-decision calibrated certificates (Vovk et al., 2005; Angelopoulos & Bates, 2023) are follow-on directions deferred to future work.

References

Yves Achdou, Fabio Camilli, Alessandra Cutrì, and Nicoletta Tchou. Hamilton–Jacobi equations constrained on networks. *NoDEA Nonlinear Differential Equations and Applications*, 20(3):413–445, 2013. doi: 10.1007/s00030-012-0158-1.

Deeksha Adil, Richard Peng, and Sushant Sachdeva. Fast, provably convergent IRLS algorithm for p-norm linear regression. In *Advances in Neural Information Processing Systems 32*, pp. 14166–14177, 2019.

Deeksha Adil, Rasmus Kyng, Richard Peng, and Sushant Sachdeva. Fast algorithms for ℓ_p-regression. *Journal of the ACM*, 71(5):1–45, 2024. doi: 10.1145/3686794.

Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay. In *Advances in neural information processing systems*, 2017.

Anastasios N. Angelopoulos and Stephen Bates. Conformal prediction: a gentle introduction. *Foundations and Trends in Machine Learning*, 16(4):494–591, 2023. doi: 10.1561/2200000101.

Gunnar Aronsson. Extension of functions satisfying Lipschitz conditions. *Arkiv för Matematik*, 6(6):551–561, 1967.

Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, 2003. doi: 10.1162/089976603321780317.

Mikhail Belkin, Partha Niyogi, and Vikas Sindhwani. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research*, 7:2399–2434, 2006.

Yoshua Bengio, Olivier Delalleau, and Nicolas Le Roux. Label propagation and quadratic criterion. In *Semi-Supervised Learning*, pp. 193–216. MIT Press, 2006.

Leon Bungert, Jeff Calder, and Tim Roith. Uniform convergence rates for Lipschitz learning on graphs. *IMA Journal of Numerical Analysis*, 43(4):2445–2495, 2023. doi: 10.1093/imanum/drac048.

Jeff Calder. The game-theoretic p-Laplacian and semi-supervised learning with few labels. *Nonlinearity*, 32 (1):301–330, 2018.

Jeff Calder. Consistency of Lipschitz learning with infinite unlabeled data and finite labeled data. *SIAM Journal on Mathematics of Data Science*, 1(4):780–812, 2019. doi: 10.1137/18M1199241.

Jeff Calder and Dejan Slepčev. Properly-weighted graph Laplacian for semi-supervised learning. *Applied Mathematics & Optimization*, 82(3):1111–1159, 2020. doi: 10.1007/s00245-019-09637-3.

Jeff Calder, Dejan Slepčev, and Matthew Thorpe. Rates of convergence for Laplacian semi-supervised learning with low labeling rates. *Research in the Mathematical Sciences*, 10(1), 2023. doi: 10.1007/s40687-022-00371-x.

Fabio Camilli and Claudio Marchi. A comparison among various notions of viscosity solution for Hamilton–Jacobi equations on networks. *Journal of Mathematical Analysis and Applications*, 2013. doi: 10.1016/j.jmaa.2013.05.015.

Keenan Crane, Clarisse Weischedel, and Max Wardetzky. Geodesics in heat: a new approach to computing distance based on heat flow. *ACM Transactions on Graphics*, 32(5):1–11, 2013. doi: 10.1145/2516971.2516977.

Peter Dayan. Improving generalization for temporal difference learning: the successor representation. *Neural Computation*, 5(4):613–624, 1993. doi: 10.1162/neco.1993.5.4.613.

Ahmed El Alaoui, Xiuyuan Cheng, Aaditya Ramdas, Martin J. Wainwright, and Michael I. Jordan. Asymptotic behavior of ℓ_p-based Laplacian regularization in semi-supervised learning. In *Proceedings of the 29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 879–906, 2016.

Abderrahim Elmoataz, Xavier Desquesnes, and Matthieu Toutain. On the game p-Laplacian on weighted graphs with applications in image processing and data clustering. *European Journal of Applied Mathematics*, 28(6), 2017. doi: 10.1017/S0956792517000122.

Benjamin Eysenbach, Ruslan Salakhutdinov, and Sergey Levine. Search on the replay buffer: bridging planning and reinforcement learning. In *Advances in Neural Information Processing Systems 32*, 2019.

Amir-massoud Farahmand, Rémi Munos, and Csaba Szepesvári. Error propagation for approximate policy and value iteration. In *Advances in Neural Information Processing Systems 23*, pp. 568–576, 2010.

Justin Fu, Aviral Kumar, Ofir Nachum, George Tucker, and Sergey Levine. D4RL: Datasets for deep data-driven reinforcement learning. In *arXiv preprint arXiv:2004.07219*, 2020.

Nicolás García Trillos and Ryan W. Murray. A maximum principle argument for the uniform convergence of graph Laplacian regressors. *SIAM Journal on Mathematics of Data Science*, 2(3):705–739, 2020. doi: 10.1137/19M1245372.

Robert Jensen. Uniqueness of Lipschitz extensions: minimizing the sup norm of the gradient. *Archive for Rational Mechanics and Analysis*, 123(1):51–74, 1993. doi: 10.1007/BF00386368.

Petri Juutinen and Nageswari Shanmugalingam. Equivalence of AMLE, strong AMLE, and comparison with cones in metric measure spaces. *Mathematische Nachrichten*, 279(9–10):1083–1098, 2006. doi: 10.1002/mana.200510411.

Rasmus Kyng, Anup Rao, Sushant Sachdeva, and Daniel A Spielman. Algorithms for Lipschitz learning on graphs. In *Proceedings of the 28th Annual Conference on Learning Theory*, volume 40 of *PMLR*, pp. 1190–1223, 2015.

Erwan Le Gruyer. On absolutely minimizing Lipschitz extensions and PDE $\Delta_\infty u = 0$. *NoDEA Nonlinear Differential Equations and Applications*, 14(1):29–55, 2007.

Sridhar Mahadevan and Mauro Maggioni. Proto-value functions: a Laplacian framework for learning representation and control in Markov decision processes. *Journal of Machine Learning Research*, 8: 2169–2231, 2007.

Juan J. Manfredi, Mikko Parviainen, and Julio D. Rossi. Dynamic programming principle for tug-of-war games with noise. *ESAIM: Control, Optimisation and Calculus of Variations*, 2012. doi: 10.1051/cocv/2010046.

Rémi Munos. Error bounds for approximate policy iteration. In *Proceedings of the 20th International Conference on Machine Learning*, pp. 560–567, 2003.

Rémi Munos. Performance bounds in L_p-norm for approximate value iteration. *SIAM Journal on Control and Optimization*, 46(2):541–561, 2007. doi: 10.1137/040614384.

Boaz Nadler, Nathan Srebro, and Xueyuan Zhou. Semi-supervised learning with the graph Laplacian: the limit of infinite unlabelled data. In *Advances in Neural Information Processing Systems 22*, pp. 1330–1338, 2009.

Soroush Nasiriany, Vitchyr H. Pong, Steven Lin, and Sergey Levine. Planning with goal-conditioned policies. In *Advances in Neural Information Processing Systems 32*, 2019.

Adam M. Oberman. A convergent difference scheme for the infinity Laplacian: construction of absolutely minimizing Lipschitz extensions. *Mathematics of Computation*, 74(251):1217–1230, 2005.

Yuval Peres, Oded Schramm, Scott Sheffield, and David B. Wilson. Tug-of-war and the infinity Laplacian. *Journal of the American Mathematical Society*, 22(1):167–210, 2009. doi: 10.1090/S0894-0347-08-00606-1.

Vitchyr H. Pong, Shixiang Gu, Murtaza Dalal, and Sergey Levine. Temporal difference models: model-free deep RL for model-based control. In *International Conference on Learning Representations*, 2018.

Tim Roith and Leon Bungert. Continuum limit of Lipschitz learning on graphs. *Foundations of Computational Mathematics*, 23(2):393–431, 2023. doi: 10.1007/s10208-022-09557-9.

Tom Schaul, Daniel Horgan, Karol Gregor, and David Silver. Universal value function approximators. In *International conference on machine learning*, 2015.

Dirk Schieborn and Fabio Camilli. Viscosity solutions for Eikonal equations on topological networks. *Calculus of Variations and Partial Differential Equations*, 46(3-4):671–686, 2013. doi: 10.1007/s00526-012-0498-z.

James A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996. doi: 10.1073/pnas.93.4.1591.

Scott Sheffield and Charles K Smart. Vector-valued optimal Lipschitz extensions. *Communications on Pure and Applied Mathematics*, 65(1):128–154, 2012. doi: 10.1002/cpa.20391.

Satinder P. Singh and Richard C. Yee. An upper bound on the loss from approximate optimal-value functions. *Machine Learning*, 16(3):227–233, 1994. doi: 10.1023/A:1022693225949.

Dejan Slepčev and Matthew Thorpe. Analysis of p-Laplacian regularization in semi-supervised learning. *SIAM Journal on Mathematical Analysis*, 51(3):2085–2120, 2019. doi: 10.1137/17M115222X.

Alexander J. Smola and Risi Kondor. Kernels and regularization on graphs. In *Learning Theory and Kernel Machines*, volume 2777 of *Lecture Notes in Computer Science*, pp. 144–158. 2003. doi: 10.1007/978-3-540-45167-9_12.

Kimberly L. Stachenfeld, Matthew M. Botvinick, and Samuel J. Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20:1643–1653, 2017. doi: 10.1038/nn.4650.

Aviv Tamar, Yi Wu, Garrett Thomas, Sergey Levine, and Pieter Abbeel. Value iteration networks. In *Advances in Neural Information Processing Systems 29*, 2016.

John N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. *IEEE Transactions on Automatic Control*, 40(9):1528–1538, 1995. doi: 10.1109/9.412624.

Vladimir Vovk, Alex Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer, 2005. doi: 10.1007/b106715.

Ronald J. Williams and Leemon C. Baird. Tight performance bounds on greedy policies based on imperfect value functions. Technical Report NU-CCS-93-14, Northeastern University, College of Computer Science, 1993.

Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf. Learning with local and global consistency. In *Advances in Neural Information Processing Systems 16*, 2003.

Xiaojin Zhu, Zoubin Ghahramani, and John D. Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions. In *Proceedings of the 20th International Conference on Machine Learning*, 2003.

A Audit tables for the empirical claims

This appendix collects the compact tables behind the empirical claims used in Section 5. The entries below are migrated from the locked main rollout, ordering-audit, subdivision-verification, finite-p, baseline, solver, adversarial-search, mechanism-audit, and failure-decomposition artifacts. File-level provenance is recorded in the supplement manifest.

A.1 Main rollout per-configuration AntMaze grid

Tables 4 and 5 expand the main rollout phase diagram from Section 5.2. Each row aggregates five seeds, each seed contains 512 start/goal rollout pairs, and the methods are paired on the same graph and evaluation pairs. The lift column is AMLE success minus harmonic success in percentage points. Loop columns report the method-specific fraction of rollout outcomes recorded as `loop`.

Table 4: Main rollout per-configuration grid for AntMaze medium graph geometry (five seeds per row; 512 rollouts per seed).

r	lf	harmonic success	AMLE success	lift (pp)	harmonic loop	AMLE loop
4	0.02	0.849	0.977	+12.8	0.151	0.023
4	0.05	0.860	0.996	+13.6	0.140	0.004
4	0.08	0.802	0.996	+19.4	0.198	0.004
4	0.12	0.876	1.000	+12.4	0.124	0.000
8	0.02	0.533	0.875	+34.1	0.467	0.125
8	0.05	0.547	1.000	+45.3	0.453	0.000
8	0.08	0.580	0.982	+40.2	0.420	0.018
8	0.12	0.681	1.000	+31.9	0.319	0.000
12	0.02	0.231	0.794	+56.3	0.769	0.206
12	0.05	0.306	0.973	+66.7	0.694	0.027
12	0.08	0.448	0.998	+55.0	0.552	0.002
12	0.12	0.587	1.000	+41.3	0.413	0.000

Table 5: Main rollout per-configuration grid for AntMaze large graph geometry (five seeds per row; 512 rollouts per seed).

r	lf	harmonic success	AMLE success	lift (pp)	harmonic loop	AMLE loop
4	0.02	0.831	0.979	+14.7	0.169	0.021
4	0.05	0.850	0.993	+14.4	0.150	0.007
4	0.08	0.834	1.000	+16.6	0.166	0.000
4	0.12	0.895	1.000	+10.5	0.105	0.000
8	0.02	0.458	0.932	+47.3	0.542	0.068
8	0.05	0.412	0.982	+57.0	0.588	0.018
8	0.08	0.468	1.000	+53.2	0.532	0.000
8	0.12	0.591	1.000	+40.9	0.409	0.000
12	0.02	0.167	0.854	+68.7	0.833	0.146
12	0.05	0.261	0.952	+69.1	0.739	0.048
12	0.08	0.386	0.996	+61.0	0.614	0.004
12	0.12	0.561	1.000	+43.9	0.439	0.000

A.2 Ordering audit by decision scope

Table 6 reports the neighbour Kendall-τ_{nbr} ordering audit under the two decision scopes summarised in Section 5. The `all` scope evaluates local ordering over all sampled neighbour decisions in the reconstructed graph cases. The `eval_rollouts` scope weights decisions by the greedy rollouts used in the main rollout experiment.

Table 6: Ordering audit by decision scope. Deltas are AMLE minus harmonic, so negative values are favorable for `tau_lt_05_rate`, `mean_beta_true_gap`, and `positive_gap_rate`; positive values are favorable for `best_agree_rate` and `tau_mean`.

scope	metric	harmonic	AMLE	delta
`all`	`tau_lt_05_rate`	0.0986	0.0747	−0.0239
`all`	`best_agree_rate`	0.9703	0.9831	+0.0128
`all`	`mean_beta_true_gap`	0.0594	0.0338	−0.0255
`all`	`positive_gap_rate`	0.0297	0.0169	−0.0128
`all`	`tau_mean`	0.7388	0.8029	+0.0641
`eval_rollouts`	`tau_lt_05_rate`	0.0638	0.0149	−0.0489
`eval_rollouts`	`best_agree_rate`	0.9754	0.9969	+0.0215
`eval_rollouts`	`mean_beta_true_gap`	0.0491	0.0062	−0.0429
`eval_rollouts`	`positive_gap_rate`	0.0246	0.0031	−0.0215
`eval_rollouts`	`tau_mean`	0.7672	0.8271	+0.0599

A.3 Subdivision verification table

Table 7 records the script-verified subdivision sweep used by Example 3.12 and Corollary B.5. The true greedy branch from coarse state 4 is branch 3 for all k. Harmonic chooses branch 1 for all tested subdivisions, while AMLE chooses branch 3.

Table 7: Subdivision-stability verification on the seven-node counterexample. Margins are the method-specific first-step value gap favoring the listed branch; all verification flags are true in the archived subdivision-verification summary.

subdivision k	true branch	harmonic branch / margin	AMLE branch / margin	verified
1	3	1 / 0.1034	3 / 0.3333	yes
2	3	1 / 0.1034	3 / 0.3333	yes
4	3	1 / 0.1034	3 / 0.3333	yes
8	3	1 / 0.1034	3 / 0.3333	yes
16	3	1 / 0.1034	3 / 0.3333	yes

The subdivision-verification table supports only the fixed-graph refinement-stability statement in Example 3.12 / Corollary B.5. It is not used to claim that the main rollout experiment's changing-resolution AntMaze sweep is a theorem consequence of the seven-node example.

A.4 Finite-p Laplacian sweep

Table 8 aggregates the finite-p family sweep by resolution r, supporting the §5.3 finite-p baseline discussion. Each (r, p) cell aggregates two mazes, four label fractions, and five seeds (40 configurations); the $p = 2$ and $p = \infty$ columns are main rollout references (harmonic and AMLE from Table 4–5 aggregated identically).

L-BFGS-B optimizer convergence flags (scipy `success` on the p-Dirichlet energy minimisation) per p across all 120 cells are: $p = 4$: 95/120 converged (mean $\|\nabla E_p\|_\infty = 0.0016$); $p = 8$: 10/120 (0.0082); $p = 16$: 0/120 (0.023). Mean per-cell max midrange-style p-residual at termination (maximum across cells in parentheses): $p = 4$: 0.04 (0.07); $p = 8$: 0.29 (0.40); $p = 16$: 0.61 (0.77). These flags indicate the energy-minimisation

Table 8: Finite-p graph Laplacian success aggregated by resolution r. Means $\pm$ standard deviations over the 40-configuration aggregate at each r. $p = 2$ and $p = \infty$ columns are main rollout references.

r	$p = 2$ (harm. ref.)	$p = 4$	$p = 8$	$p = 16$	$p = \infty$ (AMLE ref.)
4	0.850	0.977 ± 0.023	0.973 ± 0.024	0.972 ± 0.025	0.993
8	0.534	0.890 ± 0.100	0.980 ± 0.031	0.987 ± 0.016	0.971
12	0.368	0.841 ± 0.126	0.967 ± 0.046	0.987 ± 0.012	0.946

routine did not reach scipy's internal tolerance at high p within the iteration budget, although the resulting surrogates still produce high greedy success. See §5.3 for interpretation.

A.5 Reference baselines: nearest-label and full-graph Dijkstra

Table 9 reports the reference baselines aggregated by resolution r, comparing sparse-label AMLE (main rollout reference) against a nearest-label Voronoi assignment (using the same sparse Γ_g as AMLE) and a full-graph oracle Dijkstra. Each (r, method) cell aggregates the same 40-configuration grouping used in Table 8.

Table 9: Reference baselines aggregated by resolution r. Nearest-label uses the same sparse boundary as AMLE; oracle Dijkstra uses full-graph shortest-path information.

r	AMLE (sparse, main rollout)	nearest-label	oracle Dijkstra
4	0.993	0.038 ± 0.011	1.000 ± 0.000
8	0.971	0.031 ± 0.009	1.000 ± 0.000
12	0.946	0.027 ± 0.009	1.000 ± 0.000

The sparse-label AMLE-vs-oracle gap is 0.7 pp at $r = 4$, 2.9 pp at $r = 8$, and 5.4 pp at $r = 12$; the gap grows with graph size but remains small relative to the +93.9 pp aggregate margin AMLE holds over nearest-label.

A.6 AMLE iteration audit

Table 10 reports the AMLE iteration audit: on 24 representative cells (two mazes, $r \in \{8, 12\}$, $\mathrm{lf} \in \{0.02, 0.08\}$, three seeds), AMLE is rerun at seven iteration budgets with frozen random state across budgets (paired Γ_g and start–goal pairs). Each row aggregates 24 cell-level success / loop / residual values.

Table 10: AMLE iteration audit on a 24-cell representative subset; means $\pm$ standard deviations over the 24 cells. $\|R_\infty\|_\infty$ is the maximum midrange residual at solver termination, averaged over cells.

iterations	success	loop share	mean $\|R_\infty\|_\infty$	max $\|R_\infty\|_\infty$
50	0.707 ± 0.196	0.293	5.5×10^{-2}	1.1×10^{-1}
100	0.808 ± 0.166	0.192	3.2×10^{-2}	6.8×10^{-2}
200	0.894 ± 0.116	0.106	1.7×10^{-2}	3.4×10^{-2}
500	0.970 ± 0.049	0.030	6.2×10^{-3}	1.7×10^{-2}
1000	0.996 ± 0.014	0.004	2.3×10^{-3}	9.5×10^{-3}
2000	1.000 ± 0.002	0.000	7.3×10^{-4}	3.0×10^{-3}
5000	1.000 ± 0.000	0.000	9.2×10^{-5}	3.8×10^{-4}

The 95%-of-converged-success stabilisation threshold sits at iteration 500 (median across cells), matching the approximate iteration budget used by the main rollout runner. The main rollout AMLE aggregate 0.970 is therefore consistent with a partially- converged solver; longer iteration drives the observed rollout failure rate to zero on this 24-cell subset.

A.7 Finite-p solver audit

Table 11 reports the solver audit supporting the §5.3 finite-p caveat. The 24-cell representative subset (mazes $\in$ {medium, large}, $r \in \{8, 12\}$, lf $\in \{0.02, 0.08\}$, seeds $\in \{54, 55, 56\}$) is solved at each $p \in \{4, 8, 16\}$ by three solvers: L-BFGS-B with a 50,000-iteration cap and ∇E_p tolerance 10^{-6}; fixed-point Picard on the p-Laplacian update at 5000 outer sweeps with relaxation $\omega = 0.05$ and update-tolerance 10^{-6}; and Newton-CG with a 20-iteration cap (an intentionally tight budget chosen as a curvature-stress diagnostic, not a competitive solver).

Table 11: Finite-p solver audit on a 24-cell representative subset. "conv." is the count of cells the solver reports as converged ($n = 24$). $\|R_p\|_\infty$ is the L^∞ norm of the p-Laplacian update residual at termination, averaged over cells. Newton-CG with a 20-iteration cap is included as a curvature-stress diagnostic and is not a competitive solver at high p.

p	solver	success	conv.	mean $\|R_p\|_\infty$
4	L-BFGS-B (50,000 cap)	0.829 ± 0.125	24/24	0.019
4	fixed-point (5000 sweeps)	0.827 ± 0.127	10/24	0.003
4	Newton-CG (20 cap)	0.829 ± 0.125	5/24	0.036
8	L-BFGS-B (50,000 cap)	0.958 ± 0.049	23/24	0.178
8	fixed-point (5000 sweeps)	0.965 ± 0.047	11/24	0.002
8	Newton-CG (20 cap)	0.908 ± 0.105	0/24	0.834
16	L-BFGS-B (50,000 cap)	0.982 ± 0.015	16/24	0.436
16	fixed-point (5000 sweeps)	0.997 ± 0.004	15/24	0.00013
16	Newton-CG (20 cap)	0.551 ± 0.219	0/24	1.646

The two well-behaved solvers (L-BFGS-B at the 50,000 cap and fixed-point at 5000 sweeps) report similar success on $p \in \{4, 8\}$ and qualitatively reproduce the finite-p family sweep's high $p = 16$ success. They disagree quantitatively at $p = 16$ in a revealing way: L-BFGS-B reports gradient-converged solutions ($\|\nabla E_p\|_\infty \sim 5 \cdot 10^{-4}$) whose p-Laplacian update residual is still order 0.4, while fixed-point drives the update residual to $\sim 10^{-4}$ but does not certify the gradient. Newton-CG with a 20-iteration cap fails to converge anywhere at $p \in \{8, 16\}$, yielding success values that are solver artefacts. Per the §5.3 interpretation, this cross-solver disagreement at $p = 16$ means the finite-p family sweep lift of $p = 16$ over $p = \infty$ cannot be read as a planner-admissibility advantage of finite p; rather, the $p = \infty$ endpoint is the principled choice because its local update rule does not suffer the high-p ill-conditioning that prevents a clean comparison across solvers.

A.8 Operator-compatibility mechanism audit

The mechanism audit recomputes the ordering-audit rollout-weighted decision scopes on the full 120-cell main rollout grid and evaluates the true shortest-path distance d_g under two local operators (the operator-level objects of Corollary 3.11 and Proposition 3.10). The local geometry class is AMLE-compatible but harmonic-incompatible when the AMLE midrange residual $\mathcal{A}[d_g](x) - d_g(x)$ vanishes while the harmonic averaging residual $\frac{1}{\deg(x)} \sum_{y \sim x} d_g(y) - d_g(x)$ is non-zero over the greedy neighbor scope, i.e. the $n_+(x) \neq n_-(x)$ regime of Corollary 3.11. Harmonic inversions are counted only when the true-best neighbor is non-tied and harmonic strictly ranks a worse neighbor ahead of it; AMLE correction means that AMLE selects a true-best neighbor on that same decision. The certified-correction column is the single-state AMLE local-admissibility test from Lemma 3.2 firing, i.e. $\epsilon_s^{\mathrm{AMLE}} < \Delta_s^\star/2$ at the inverted state (with AMLE error bounded as in Corollary 3.5), expressed as a fraction of harmonic-inverted states. Harmonic anti-admissibility is the Lemma 3.9 certificate and is satisfied by definition on inversion events.

The primary AMLE-compatible / harmonic-incompatible geometry is most common at the coarsest resolution, but harmonic inversions remain present at all resolutions and become more often certified by the local-gap condition as r increases. The audit is deliberately local and distributional: it supports the operator-mismatch reading of the AntMaze rollouts, not universal AMLE dominance.

Table 12: Aggregate operator-compatibility and inversion audit. Geometry rows are over all rollout-weighted decisions (2,481,029 total). Inversion and correction rows are conditioned on 534,916 non-tied true-best decisions, except where the denominator is explicitly the harmonic-inversion count.

diagnostic	count / denominator	rate
both-compatible geometry	1,079,361/2,481,029	43.5%
AMLE-compatible / harmonic-incompatible geometry	1,358,673/2,481,029	54.8%
both-incompatible geometry	42,995/2,481,029	1.7%
harmonic inversions	30,037/534,916	5.6%
inversions in primary geometry	29,821/30,037	99.3%
AMLE corrections of harmonic inversions	27,977/30,037	93.1%
certified corrections of harmonic inversions	20,135/30,037	67.0%

Table 13: Mechanism audit by graph resolution r. All entries are percentages; correction and certificate rates are conditional on harmonic inversions at the given resolution.

r	primary geometry	harmonic inversion	AMLE correction	certified correction
4	84.4	4.7	88.4	39.8
8	56.6	5.5	94.8	71.7
12	38.5	6.5	94.9	81.6

A.9 p-family ordering summary

The p-family ordering summary is summary-only: it merges the ordering-audit local-ordering diagnostics for $p = 2$ and $p = \infty$ with the finite-p family-sweep summaries for $p \in \{4, 8, 16\}$ on the same 120 case keys. It does not compute finite-p local-gap certificates, and the $p = 16$ rows retain the finite-p sweep caveat that L-BFGS-B was run at the 700-iteration budget without scipy optimizer convergence.

Table 14: Summary-only p-family ordering table over the 120-cell main rollout grid. The local-ordering columns are scope-normalized ordering-audit / finite-p sweep diagnostics; low-τ_{nbr} is `tau_lt_05_rate`.

p	success	loop	low-τ_{nbr}	best agree	mean gap	mean τ_{nbr}
2	0.584	0.416	0.081	0.973	0.054	0.753
4	0.903	0.097	0.051	0.987	0.026	0.775
8	0.973	0.027	0.044	0.987	0.025	0.777
16	0.982	0.018	0.043	0.986	0.027	0.776
∞	0.970	0.030	0.045	0.990	0.020	0.815

The family summary supports the same qualitative transition as the p-family sweep and solver audit: the high-p cluster has much lower local-ordering tail mass than $p = 2$. It is not monotone enough to be treated as a finite-p theorem. Across the 120 paired cases, success is monotone nondecreasing in only 33/120 cases, mean τ_{nbr} is monotone nondecreasing in 17/120, and $p = 16$ lies between $p = 8$ and $p = \infty$ in success on 57/120 cases. These failures are reported as diagnostics, not anomalies requiring post-hoc filtering.

A.10 Adversarial AMLE-versus-harmonic subgraph search

The adversarial subgraph search performs a random search for small-graph instances on which AMLE underperforms harmonic, complementing the Example 3.12 witness in the opposite direction. Search candidates are random connected induced subgraphs of 4×4, 5×4, and 5×5 unweighted lattices; for each candidate, the boundary Γ_g is set to the goal vertex plus 1–4 additional non-goal vertices labelled by their true shortest-path distance, and both surrogates are run to update-norm tolerance 10^{-8}. Of 588 tested 4×4 candidates with random seed 20260510, 5 exhibit harmonic-greedy success 1.000 versus AMLE-greedy success ≤ 0.83. The mechanism is common across the five witnesses: AMLE produces a high-value plateau on a pair

of interior vertices whose only nearby labelled vertex is the high-cost boundary, and the plateau inverts the local greedy ordering between the plateau vertices.

Table 15 summarises the five witnesses. Full edge lists, boundary values, and per-vertex surrogate values for each are archived with the adversarial-search summary.

Table 15: Adversarial AMLE-bad / harmonic-good witnesses on 4×4 subgraphs. $|V|$ is the candidate's vertex count, goal is the labelled goal vertex, $|\partial|$ is the number of labelled boundary vertices, and the success columns count greedy rollouts launched from each non-boundary vertex. Failures are listed by interior-vertex label.

#	$\lvert V\rvert$	goal	$\lvert\partial\rvert$	harm. succ.	AMLE succ.	AMLE failures
1	13	10	5	1.000	0.833	$\{12, 13\}$
2	14	6	4	0.769	0.692	$\{8, 9, 12, 13\}$
3	13	15	4	1.000	0.833	$\{1, 5\}$
4	16	3	5	1.000	0.800	$\{8, 12, 13\}$
5	13	2	5	0.750	0.667	$\{7, 11, 14, 15\}$

Witnesses #1, #3, #4 have harmonic at the perfect-success ceiling and AMLE strictly below; witnesses #2 and #5 show both surrogates failing on overlapping interior sets, with harmonic still strictly better. These five small-graph witnesses are existence proofs that the converse direction (AMLE-bad, harmonic-good) is realisable on lattice subgraphs with random sparse boundaries; combined with Example 3.12, they establish that neither AMLE nor harmonic universally dominates on sparse-boundary graph planning. The main rollout AntMaze aggregate (+38.6 pp paired lift in favour of AMLE) is therefore a *distributional* statement about AntMaze graph geometry, not a universal-dominance theorem.

B Proofs of §3 statements

This appendix gathers proof bodies for the theorems, lemmas, propositions, and corollaries stated in Section 3. Each is keyed by its label in the main text.

Statement and proof of Corollary B.1

Corollary B.1 (End-to-end ordinal fidelity, restated)**.** *Combining Lemmas 3.4 and 3.2 with $A = N(s)$,*

$$\epsilon_s \;\le\; \epsilon_{\text{lab}} + \eta_\infty(V^\star; N(s), \Gamma_g) \;\le\; \epsilon_{\text{lab}} + 2L_g\, h_{\Gamma_g}(N(s)),$$

and, for any non-leaf state s ($d(s) \ge 2$),

$$\tau^{\text{AMLE}}_{\text{nbr}}(s, g) \;\ge\; 1 - \frac{4\, M_{s,g}(2\epsilon_{\text{lab}} + 4L_g\, h_{\Gamma_g}(N(s)))}{d(s)(d(s)-1)}.$$

Proof. Immediate from Lemma 3.4 applied with $A = N(s)$ (controlling ϵ_s in terms of fill distance) and Lemma 3.2 (converting ϵ_s to a neighbour-Kendall-τ bound). □

Proof of Lemma 3.6 (p-Laplacian maximum principle)

Proof. We prove the no-strict-minimum claim; the no-strict-maximum claim is identical with inequalities reversed. Suppose for contradiction that $x \in V \setminus \Gamma_g$ is a strict interior local minimum of u, i.e. $u(y) > u(x)$ for every $y \sim x$.

Case $p = 2$ (harmonic). Strict minimality gives $u(y) - u(x) > 0$ for every $y \sim x$, so $\sum_{y\sim x}(u(y) - u(x)) > 0$, contradicting the harmonic equation $\sum_{y\sim x}(u(y) - u(x)) = 0$.

Case $2 < p < \infty$. Strict minimality gives both $|u(y) - u(x)| > 0$ and $u(y) - u(x) > 0$ for every $y \sim x$, so $|u(y) - u(x)|^{p-2}(u(y) - u(x)) > 0$. Hence $\sum_{y\sim x} |u(y) - u(x)|^{p-2}(u(y) - u(x)) > 0$, contradicting $\Delta_p u(x) = 0$.

*Case $p = \infty$ (*AMLE*).* Strict minimality gives $\min_{y\sim x} u(y) > u(x)$ and $\max_{y\sim x} u(y) > u(x)$, so the midrange satisfies $\frac{1}{2}(\min_y u(y) + \max_y u(y)) > u(x)$, contradicting the AMLE identity $u(x) = \frac{1}{2}(\min_y u(y) + \max_y u(y))$. □

Statement and proof of Proposition B.2 (combined local separation)

Proposition B.2 (Combined AMLE-vs-harmonic local separation)**.** *Fix a decision state s with true optimal neighbour set $A^\star(s,g)$. Suppose*

*(i) (*AMLE local admissibility, Lemma 3.2*)* $\epsilon_s^{\mathrm{AMLE}} := \max_{y\sim s} |\hat{V}_\infty(y,g) - V^\star(y,g)| < \Delta_s^\star/2$*, and*

*(ii) (*harmonic local anti-admissibility, Lemma 3.9*) there exists $b \in N(s) \setminus A^\star(s,g)$ with $Q_{\hat{u}_2,g}(s,b) < \min_{a\in A^\star(s,g)} Q_{\hat{u}_2,g}(s,a)$, equivalently $w(s,b) - w(s,a) + \sum_{z\in\Gamma_g}(\omega_b(z) - \omega_a(z))Y_g(z) < 0$ for every $a \in A^\star(s,g)$.*

*Then $T_{\hat{V}_\infty,g}(s) \in A^\star(s,g)$ (*AMLE*-greedy correct) and $T_{\hat{u}_2,g}(s) \notin A^\star(s,g)$ (harmonic-greedy wrong).*

Proof. Clause (i) and Lemma 3.2 give $\arg\min_y Q_{\hat{V}_\infty,g}(s,y) \subseteq A^\star(s,g)$. Clause (ii) gives $\min_y Q_{\hat{u}_2,g}(s,y) \le Q_{\hat{u}_2,g}(s,b) < \min_{a\in A^\star} Q_{\hat{u}_2,g}(s,a)$, so $T_{\hat{u}_2,g}(s) \notin A^\star(s,g)$. □

Statement and proof of Lemma B.3 (residual–margin bound)

For a surrogate $\hat{u} : V \to \mathbb{R}$ with $\hat{u}|_{\Gamma_g} = Y_g$, write $R_\infty(x;\hat{u}) := \mathcal{A}[\hat{u}](x) - \hat{u}(x)$ for the midrange residual, and $\gamma(s^*) := \min_{s'\sim s^*} \hat{u}(s') - \hat{u}(s^*)$ for the strict-min depth at any approximate strict interior local minimum s^* of $\hat{u}$.

Lemma B.3 (Residual–margin bound for finite-sweep AMLE)**.** *If $\|R_\infty(\cdot;\hat{u})\|_\infty \le \delta$ then $\gamma(s^*) \le \delta$ at every approximate strict interior local minimum s^*.*

Proof. Let s^* be an approximate strict interior local minimum of $\hat{u}$ with $\gamma(s^*) > 0$, so $\min_{s'\sim s^*} \hat{u}(s') = \hat{u}(s^*) + \gamma(s^*)$ and every neighbor exceeds $\hat{u}(s^*)$ by at least $\gamma(s^*)$, whence $\max_{s'\sim s^*} \hat{u}(s') \ge \hat{u}(s^*) + \gamma(s^*)$. Therefore

$$\mathcal{A}[\hat{u}](s^*) \;=\; \tfrac{1}{2}\big(\min_{s'\sim s^*} \hat{u}(s') + \max_{s'\sim s^*} \hat{u}(s')\big) \;\ge\; \hat{u}(s^*) + \gamma(s^*),$$

so $R_\infty(s^*;\hat{u}) \ge \gamma(s^*)$. Combining with $\|R_\infty\|_\infty \le \delta$ gives $\gamma(s^*) \le \delta$. □

Proof of Lemma 3.4 (AMLE local extension error)

Proof sketch. By the boundary-pinned AMLE comparison principle on graphs (Kyng et al., 2015; Bungert et al., 2023), $\|\mathcal{A}_{\Gamma_g}(z) - \mathcal{A}_{\Gamma_g}(z')\|_\infty \le \|z - z'\|_{\infty,\Gamma_g}$. Apply with $z = Y_g$ and $z' = V^\star|_{\Gamma_g}$ to obtain $\|\hat{V}_\infty - \mathcal{A}_{\Gamma_g}(V^\star|_{\Gamma_g})\|_\infty \le \epsilon_{\mathrm{lab}}$; triangle inequality with η_∞ gives the first bound. For the fill-distance bound, AMLE preserves the Lipschitz constant of the boundary data; for any $x \in A$ and a closest labeled $z \in \Gamma_g$, both $\mathcal{A}_{\Gamma_g}(V^\star|_{\Gamma_g})$ and $V^\star$ change by at most $L_g\, h_{\Gamma_g}(A)$ along a path from z to x, so their difference at x is at most $2L_g\, h_{\Gamma_g}(A)$. □

Proof of Lemma 3.2 (gap-dependent local ordinal fidelity)

Proof. For any pair of neighbours $y, y' \in N(s)$, $Q_{\hat{V},g}(s,y) - Q_{\hat{V},g}(s,y')$ differs from $Q^\star(s,y;g) - Q^\star(s,y';g)$ by exactly $(\hat{V}(y,g) - V^\star(y,g)) - (\hat{V}(y',g) - V^\star(y',g))$, which has magnitude at most $2\epsilon_s$ (the proof uses only this local sup error bound and not any AMLE-specific structure). Hence if $|Q^\star(s,y;g) - Q^\star(s,y';g)| > 2\epsilon_s$, the surrogate ordering on (y,y') matches the true ordering and the pair is not an inversion. Therefore $I_{s,g} \le M_{s,g}(2\epsilon_s)$, and the Kendall-tau identity gives the displayed bound.

For the action-ordering preservation part, pick any $a \in A^\star(s,g)$ and any $b \in N(s) \setminus A^\star(s,g)$. By definition of $A^\star$ and $\Delta_s^\star$, $Q^\star(s,b;g) - Q^\star(s,a;g) \ge \Delta_s^\star > 2\epsilon_s$; the same gap argument forces $Q_{\hat{V},g}(s,b) > Q_{\hat{V},g}(s,a)$. Hence $\arg\min_y Q_{\hat{V},g}(s,y) \subseteq A^\star(s,g)$. □

Statement and proof of Corollary B.4 (bad-tail mass)

Corollary B.4 (Bad-tail mass under local-gap regularity, restated)**.** *Let μ_{dec} be the planner's decision distribution over (s, g) pairs with $d(s) \geq 2$, and define the small-gap pair fraction $G_{s,g}(\eta) := M_{s,g}(\eta)/\binom{d(s)}{2}$. Suppose $G_{s,g}(\eta) \leq C_0\eta^\alpha$ uniformly in $(s, g) \sim \mu_{\mathrm{dec}}$ for some $\alpha, C_0 > 0$ (a hypothesis on the local action-gap CDF, not a derived property). Then for any threshold $\theta < 1$,*

$$\mu_{\mathrm{dec}}\big\{\tau_{\mathrm{nbr}} \leq \theta\big\} \;\leq\; \frac{2^{\alpha+1}C_0}{1-\theta}\,\mathbb{E}_{\mu_{\mathrm{dec}}}[\epsilon_s^\alpha].$$

Proof. By Lemma 3.2, $1 - \tau_{\mathrm{nbr}}(s, g) \leq \frac{4}{d(d-1)}M_{s,g}(2\epsilon_s) = 2G_{s,g}(2\epsilon_s) \leq 2C_0(2\epsilon_s)^\alpha = 2^{\alpha+1}C_0\epsilon_s^\alpha$. Markov inequality on $1 - \tau_{\mathrm{nbr}}$ at level $1 - \theta$ gives the bound. □

Proof of Proposition 3.10 (distance is graph-AMLE on the extendable interior)

Proof. Part (a). By the unit-graph triangle inequality, $|d_g(y) - d_g(x)| \leq 1$ for every $y \sim x$, so $\min_{y\sim x} d_g(y) \geq d_g(x) - 1$. Let $x = x_0, x_1, \ldots, x_k = g$ be a shortest path with $k = d_g(x)$; then x_1 is a neighbor of x with $d_g(x_1) = d_g(x) - 1$, giving the reverse inequality.

Part (b). By the same triangle inequality, $\max_{y\sim x} d_g(y) \leq d_g(x) + 1$, and $\max \geq \min = d_g(x) - 1$. The case $\max = d_g(x) - 1$ does occur at non-leaf vertices in cut-locus regions (e.g., a diamond graph: g at distance 0, two intermediate neighbors at distance 1, and a vertex x at distance 2 adjacent to both intermediates — every neighbor of x is closer to g).

Part (c). By (a), $\min_{y\sim x} d_g(y) = d_g(x) - 1$ always. Setting $\mathcal{A}[d_g](x) = d_g(x)$ and using (b),

$$d_g(x) \;=\; \tfrac{1}{2}\big((d_g(x) - 1) + \max_{y\sim x} d_g(y)\big) \iff \max_{y\sim x} d_g(y) = d_g(x) + 1 \iff x \in V_g^{\mathrm{ext}}.$$

The Bellman / eikonal identity at $x \neq g$ follows directly from (a). □

Statement and proof of Corollary B.5 (subdivision equivariance of strict harmonic rankings)

Corollary B.5 (Subdivision-equivariance of strict harmonic rankings)**.** *Strict harmonic neighbour rankings of the type described by Lemma 3.8 are preserved under uniform k-subdivision of G with label scaling $\lambda_k = k$ (metric-preserving). In particular, the inversion of Example 3.12 persists at every $k \geq 1$. The analogous statement on the* AMLE *side is Lemma B.6 below, so the discriminator between the two endpoints lies in Lemma 3.9 vs Corollary 3.5, not in subdivision-stability.*

Proof. Define $\tilde{h}^{(k)}$ on $G^{(k)}$ by $\tilde{h}^{(k)}(v) = \lambda_k h(v)$ at every original vertex and by linear interpolation along subdivided edges. At an inserted vertex v_j of an edge $v \to z$ (with $1 \leq j \leq k - 1$), the two $G^{(k)}$-neighbours are v_{j-1} and v_{j+1}, and linear interpolation gives $\tilde{h}^{(k)}(v_j) = (\tilde{h}^{(k)}(v_{j-1}) + \tilde{h}^{(k)}(v_{j+1}))/2$, i.e. the harmonic averaging identity is satisfied on the inserted path. At an original vertex $v \in V \setminus \partial$, the $G^{(k)}$-neighbours are the first inserted vertices v_1 on each edge $v \to z$ ($z \sim v$ in G); by linear interpolation, $\tilde{h}^{(k)}(v_1) = \lambda_k h(v) + (\lambda_k h(z) - \lambda_k h(v))/k$. Averaging over the original-graph neighbours $z \sim v$:

$$\tfrac{1}{\deg(v)}\sum_{z\sim v}\tilde{h}^{(k)}(v_1) \;=\; \lambda_k h(v) + \tfrac{\lambda_k}{k\deg(v)}\sum_{z\sim v}(h(z) - h(v)) \;=\; \lambda_k h(v),$$

since h is harmonic at v in G. Hence $\tilde{h}^{(k)}$ satisfies the harmonic averaging identity at every interior vertex of $G^{(k)}$ with the scaled boundary data $\lambda_k y$, and by uniqueness of the harmonic extension $h^{(k)} = \tilde{h}^{(k)}$.

For the first-step branch difference at any original decision state s with original neighbours a, b: the $G^{(k)}$-neighbours of s are the first inserted vertices $a^{(1)}, b^{(1)}$ along edges $\{s, a\}, \{s, b\}$ respectively, with $\tilde{h}^{(k)}(a^{(1)}) = \lambda_k h(s) + \lambda_k(h(a) - h(s))/k$ and analogously for $b^{(1)}$. The Q-value difference (under unit edge cost in $G^{(k)}$) is $\hat{Q}^{(k)}_{a^{(1)}} - \hat{Q}^{(k)}_{b^{(1)}} = \tilde{h}^{(k)}(a^{(1)}) - \tilde{h}^{(k)}(b^{(1)}) = (\lambda_k/k)(h(a) - h(b))$. Taking $\lambda_k = k$ gives the k-independent margin $h(a) - h(b)$. □

Statement and proof of Lemma B.6 (AMLE midrange subdivision equivariance)

Lemma B.6 (AMLE midrange equivariance under uniform subdivision)**.** *Let u_∞ be the exact graph-*AMLE *on G with boundary ∂ and labels y. Uniformly subdivide each edge into a length-k path and scale boundary labels by $\lambda_k > 0$. The linear interpolation of $\lambda_k u_\infty$ along each subdivided edge satisfies the midrange equation on the subdivided graph $G^{(k)}$, so $u_\infty^{(k)}(v) = \lambda_k u_\infty(v)$ at every original vertex. Consequently, for original neighbour candidates a, b of an original decision state s, the first-step* AMLE *branch difference satisfies $\hat{Q}^{(k),\mathrm{AMLE}}_{a^{(1)}} - \hat{Q}^{(k),\mathrm{AMLE}}_{b^{(1)}} = (\lambda_k/k)(u_\infty(a) - u_\infty(b))$; with $\lambda_k = k$ the margin is independent of k and any strict* AMLE *neighbour ranking on G persists for every $k \geq 1$.*

Proof. Linear interpolation along a degree-2 inserted path satisfies the midrange identity automatically (with neighbours of an inserted vertex at adjacent path-positions, min and max are the two neighbour values, so the midrange equals their average, which equals the linearly-interpolated value). At an original vertex v, the $G^{(k)}$-neighbours are the first inserted vertices v_1 along each edge $v \to z$, with values $\lambda_k u_\infty(v) + \lambda_k(u_\infty(z) - u_\infty(v))/k$. The midrange of these values over $z \sim v$ equals $\lambda_k u_\infty(v) + (\lambda_k/k) \cdot \frac{1}{2}(\min_z u_\infty(z) - u_\infty(v) + \max_z u_\infty(z) - u_\infty(v))$, which equals $\lambda_k u_\infty(v)$ exactly when u_∞ satisfies the midrange identity at v in G. Hence $u_\infty^{(k)} = \lambda_k u_\infty$ on original vertices by uniqueness of the AMLE extension. The branch-difference formula follows by the same first-inserted-vertex computation as in the harmonic case. □

The seven-node graph G_7 (Example 3.12)

Define $G_7 = (V_7, E_7)$ by

$$V_7 = \{0, 1, 3, 4, 5, 6, 7\}, \qquad E_7 = \big\{\{0,3\}, \{0,5\}, \{5,1\}, \{1,4\}, \{3,4\}, \{3,6\}, \{6,7\}, \{4,7\}\big\}.$$

The goal is $g = 0$ and the sparse-label boundary is $\Gamma_g = \{0, 7\}$ with observed labels $Y_g(0) = 0$ and $Y_g(7) = 3$. Shortest-path distances to g are $V^\star(3, g) = V^\star(5, g) = 1$ and $V^\star(1, g) = V^\star(4, g) = V^\star(6, g) = 2$, so the decision state $s = 4$ has neighbour set $N(4) = \{1, 3, 7\}$, true Bellman-optimal neighbour set $A^\star(4, g) = \{3\}$, and action gap $\Delta_4^\star = 1$.

Harmonic Dirichlet solve. Solving $\Delta_2 \hat{u}_2 = 0$ on $V_7 \setminus \Gamma_g$ with $\hat{u}_2|_{\Gamma_g} = Y_g$ yields, at the decision-relevant neighbours of $s = 4$,

$$\hat{u}_2(1) = \tfrac{36}{29}, \qquad \hat{u}_2(3) = \tfrac{39}{29}.$$

Via Lemma 3.8 the harmonic measures of neighbours $1, 3$ on the non-goal boundary $\{7\}$ are $\omega_1(7) = 12/29$ and $\omega_3(7) = 13/29$, so $\omega_1(7) - \omega_3(7) = -1/29 < 0$. The harmonic-greedy planner therefore prefers 1 over 3 at $s = 4$, instantiating clause (ii) of Proposition B.2.

AMLE midrange solve. The midrange fixed point $\hat{V}_\infty = \mathcal{A}_{\Gamma_g}(Y_g)$ takes the values $\hat{V}_\infty(3) = 1$ and $\hat{V}_\infty(1) = 4/3$ at the decision-relevant neighbours. In particular $\hat{V}_\infty(3) = V^\star(3, g) = 1$ exactly: $3 \in V_g^{\mathrm{ext}}$ since the neighbour 4 has $V^\star(4, g) = 2 = V^\star(3, g) + 1$, so Proposition 3.10 applies and the midrange identity $\hat{V}_\infty(3) = \mathcal{A}[\hat{V}_\infty](3)$ holds at the true value. The AMLE-greedy planner picks 3 correctly.

Mechanism-scope remark. At $s = 4$, the local sup errors are $\epsilon_4^{\mathrm{harm}} = 22/29 \approx 0.76$ and $\epsilon_4^{\mathrm{AMLE}} = 2/3 \approx 0.67$; both exceed $\Delta_4^\star/2 = 0.5$, so the Lemma 3.2 sufficient condition fails at this state for both surrogates. Harmonic nevertheless mis-ranks because its error symmetrises across the two non-goal neighbours of $s = 4$ (the harmonic-measure averaging effect of Lemma 3.8); AMLE remains greedy-correct through the operator-level exact match $\hat{V}_\infty(3) = V^\star(3, g) = 1$ of Proposition 3.10 instead. G_7 illustrates the harmonic side of Proposition B.2 with positive margin, not clause (i): clause (i)'s AMLE-correctness hypothesis is sufficient but not necessary.

Specialisation to G_7 (subdivision). Applying the harmonic case of Corollary B.5 to the G_7 data ($h(1) - h(3) = -3/29$ with $\lambda_k = k$) gives wrong-ordering margin $-3/29$ at every k; applying the AMLE case (Lemma B.6) to $u_\infty(1) - u_\infty(3) = 1/3$ gives correct-ordering margin $1/3$ at every k. These are the G_7-specific numbers used in §3.7.