Greedy Boundary Expansion for $\lambda > 1$ on a Path: A Standalone Proof Note

Standalone note for internal understanding

August 27, 2025

Abstract

We present a clean, self-contained proof that the greedy "expand the boundary that decreases the loss the most; stop when neither expansion helps" strategy is correct for a natural class of losses when selecting an interval $[i, j]$ on a fixed path. The note is independent from the main paper and can be compiled on its own.

1 Setting

Let $P = (v_1, \ldots, v_k)$ be the (fixed) path connecting two support vertices. Any candidate solution corresponds to an interval $[i, j]$ with $1 \leq i \leq j \leq k$. For a parameter $\lambda > 1$, consider a loss of the form

$$L_\lambda(i, j) \;=\; \sum_{u \in V} \varphi_u\big(\,\mathrm{dist}(u, [i, j])\big) \;+\; \lambda\, \psi(j - i), \tag{1}$$

where $\mathrm{dist}(u, [i, j])$ is the graph distance from u to the interval $[i, j]$ on P (zero if $u \in [i, j]$), each $\varphi_u : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is convex and nondecreasing, and $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is convex. These mild assumptions cover hinge-type or Huber-type terms on distances plus a convex length penalty.

Define the one-step expansions and their deltas:

$$\begin{aligned} \Delta_\ell(i \mid j) &:= L_\lambda(i-1, j) - L_\lambda(i, j) && \text{(expand left boundary)} \\ \Delta_r(j \mid i) &:= L_\lambda(i, j+1) - L_\lambda(i, j) && \text{(expand right boundary).} \end{aligned}$$

The greedy rule chooses the boundary with the smaller (more negative) delta; it stops when both $\Delta_\ell(i \mid j) \geq 0$ and $\Delta_r(j \mid i) \geq 0$.

2 Key structural lemmas

Lemma 1 (Discrete convexity in each coordinate)**.** *Fix j. The function $f(i) := L_\lambda(i, j)$ is discretely convex in i, i.e., $f(i-1) - 2f(i) + f(i+1) \geq 0$ wherever defined. Symmetrically, for fixed i, $g(j) := L_\lambda(i, j)$ is discretely convex in j.*

Sketch. For a fixed j, $h_u(i) := \mathrm{dist}(u, [i, j])$ is a piecewise-linear (on $\mathbb{Z}$) "V-shaped" function of i: it is 0 when $i \leq \mathrm{pos}(u) \leq j$ and increases linearly with slope that does not decrease as i moves away. Hence h_u is discretely convex. Since φ_u is convex and nondecreasing, the composition $\varphi_u \circ h_u$ is discretely convex; summing over u preserves discrete convexity. Finally, $i \mapsto \lambda\, \psi(j - i)$ is convex (composition of convex with affine), hence discretely convex. The proof for j is identical by symmetry. □

Corollary 1 (Monotonicity of one-step deltas)**.** *As i decreases (expanding further to the left), the first difference $\Delta_\ell(i \mid j) = f(i-1) - f(i)$ is nondecreasing; in particular, $\Delta_\ell(i-1 \mid j) \geq \Delta_\ell(i \mid j)$. As j increases (expanding further to the right), $\Delta_r(j \mid i) = g(j+1) - g(j)$ is nondecreasing; in particular, $\Delta_r(j+1 \mid i) \geq \Delta_r(j \mid i)$.*

Proof. For a discretely convex function, forward (or backward) first differences are nondecreasing. Apply Lemma 1 to f and g. □

3 Optimality of the stopping criterion and greedy correctness

Theorem 1 (Stopping criterion implies optimality over monotone expansions)**.** *If at* (i, j) *we have* $\Delta_\ell(i \mid j) \geq 0$ *and* $\Delta_r(j \mid i) \geq 0$*, then for any* (i', j') *reachable by a finite sequence of left/right expansions (i.e.,* $i' \leq i$ *and* $j' \geq j$*),*

$$L_\lambda(i', j') - L_\lambda(i, j) \;\geq\; 0.$$

Proof. Any (i', j') in the expansion cone is obtained via a finite chain of unit expansions. The total change in loss is the sum of the corresponding one-step deltas along the chain. By Corollary 1, each subsequent left-step delta is at least $\Delta_\ell(i \mid j) \geq 0$, and each subsequent right-step delta is at least $\Delta_r(j \mid i) \geq 0$. Summing nonnegative terms yields a nonnegative total change. □

Theorem 2 (Greedy expansion is correct and terminates)**.** *Starting from any* (i_0, j_0)*, repeatedly expand the boundary with the smaller (more negative) delta while* $\min\{\Delta_\ell(i \mid j), \Delta_r(j \mid i)\} < 0$*, and stop when both deltas are nonnegative. The algorithm terminates in finitely many steps and the final* (i^*, j^*) *satisfies the condition of Theorem 1, hence is optimal among all monotone expansions from* (i_0, j_0)*.*

Proof. Each step strictly decreases the loss by at least $-\min\{\Delta_\ell, \Delta_r\} > 0$. Since the length penalty $\lambda\,\psi(j-i)$ is convex in the interval length, the deltas increase along expansions (Corollary 1) and eventually become nonnegative, so the process stops. At termination, Theorem 1 applies. □

Remark 1 (Towards global optimality)**.** *If the array* $L_\lambda(i, j)$ *satisfies the Monge (quadrangle) inequality*

$$L_\lambda(i, j) + L_\lambda(i+1, j+1) \;\leq\; L_\lambda(i+1, j) + L_\lambda(i, j+1) \quad (i < j),$$

then L_λ *is* $L^\natural$*-convex on the index lattice, and local optimality with respect to unit moves implies global optimality. Verifying Monge for a specific instantiation of* (1) *can be done directly from preprocessed count/distance formulas.*

4 What to verify in your concrete model

To specialize the above to your exact loss, (i) write explicit one-step deltas Δ_ℓ, Δ_r using the preprocessed counts and distance sums on the path boundary layers; (ii) check the discrete second differences in i and in j are nonnegative (Lemma 1); (iii) optionally, check the Monge inequality if a global optimality guarantee is desired beyond monotone expansions.

Appendix: Explicit Δ formulas in the paper's notation

This appendix mirrors the notation used in your paper and gives closed forms for one-step deltas.

Notation recap (from your paper). Spine vertices: $x'_1, \ldots, x'_n$ with labels $y_k \in \{-1, +1\}$. Tree distance is $d(\cdot, \cdot)$. Preprocessed quantities:

$$N_\pm^{>i} = 2\,|\{j > i : y_j = \pm 1\}|, \qquad N_\pm^{\leq i} = 2\,|\{j \leq i : y_j = \pm 1\}|,$$

$$d_\pm^{>i} = \sum_{\substack{j>i \\ y_j=\pm 1}} \big(d(x'_i, x'_j) + d(x'_i, x_j)\big), \qquad d_\pm^{\leq i} = \sum_{\substack{j\leq i \\ y_j=\pm 1}} \big(d(x'_j, x'_{i+1}) + d(x_j, x'_{i+1})\big).$$

Loss is $L_\lambda(u, v) = f(u, v) - \lambda\, d(u, v)$.

General pair (x'_i, x'_j)**,** $i < j$**,** $y_i \neq y_j$**.** Consistent with your preprocessing, we use

$$extif(y_i, y_j) = (-1, +1): \qquad f(x'_i, x'_j) = d_-^{\geq i} + d_+^{\leq (j-1)},$$

$$extif(y_i, y_j) = (+1, -1): \qquad f(x'_i, x'_j) = d_+^{\geq i} + d_-^{\leq (j-1)}.$$

One-step expansions and deltas. Define

$$\Delta_{\text{left}}(i \mid j) := L_\lambda(x'_{i-1}, x'_j) - L_\lambda(x'_i, x'_j), \qquad \Delta_{\text{right}}(j \mid i) := L_\lambda(x'_i, x'_{j+1}) - L_\lambda(x'_i, x'_j).$$

Each delta splits as "noise change" minus λ times the change of margin. On the spine,

$$d(x'_{i-1}, x'_j) - d(x'_i, x'_j) = d(x'_{i-1}, x'_i), \qquad d(x'_i, x'_{j+1}) - d(x'_i, x'_j) = d(x'_j, x'_{j+1}).$$

Case A: $(y_i, y_j) = (-1, +1)$. Then $f(x'_i, x'_j) = d_-^{\geq i} + d_+^{\leq (j-1)}$ and

$$\begin{aligned}
\Delta_{\text{left}}(i \mid j) &= \left[d_-^{\geq (i-1)} - d_-^{\geq i}\right] - \lambda\, d(x'_{i-1}, x'_i) \\
&= \left[N_-^{\geq i}\, d(x'_{i-1}, x'_i) + 2\, d(x'_{i-1}, x'_i) + d(x_i, x'_i)\right] - \lambda\, d(x'_{i-1}, x'_i) \\
&= \left(N_-^{\geq i} + 2 - \lambda\right) d(x'_{i-1}, x'_i) + d(x_i, x'_i),
\end{aligned}$$

$$\begin{aligned}
\Delta_{\text{right}}(j \mid i) &= \left[d_+^{\leq j} - d_+^{\leq (j-1)}\right] - \lambda\, d(x'_j, x'_{j+1}) \\
&= \left[N_+^{\leq (j-1)}\, d(x'_j, x'_{j+1}) + 2\, d(x'_j, x'_{j+1}) + d(x_j, x'_j)\right] - \lambda\, d(x'_j, x'_{j+1}) \\
&= \left(N_+^{\leq (j-1)} + 2 - \lambda\right) d(x'_j, x'_{j+1}) + d(x_j, x'_j).
\end{aligned}$$

The bracketed increments follow directly from your $O(1)$ update rule for $d_\pm^{\leq \cdot}$ (mirrored to the right side) and the fact that crossing index i (resp. j) adds the spoke $d(x_i, x'_i)$ (resp. $d(x_j, x'_j)$) and one more spine segment.

Case B: $(y_i, y_j) = (+1, -1)$. Then $f(x'_i, x'_j) = d_+^{\geq i} + d_-^{\leq (j-1)}$ and

$$\begin{aligned}
\Delta_{\text{left}}(i \mid j) &= \left[d_+^{\geq (i-1)} - d_+^{\geq i}\right] - \lambda\, d(x'_{i-1}, x'_i) \\
&= \left(N_+^{\geq i} + 2 - \lambda\right) d(x'_{i-1}, x'_i) + d(x_i, x'_i),
\end{aligned}$$

$$\begin{aligned}
\Delta_{\text{right}}(j \mid i) &= \left[d_-^{\leq j} - d_-^{\leq (j-1)}\right] - \lambda\, d(x'_j, x'_{j+1}) \\
&= \left(N_-^{\leq (j-1)} + 2 - \lambda\right) d(x'_j, x'_{j+1}) + d(x_j, x'_j).
\end{aligned}$$

Monotonicity and stopping. As i decreases, $N_\pm^{\geq i}$ is nondecreasing and the added terms are nonnegative, hence Δ_{left} is nondecreasing along successive left expansions. Symmetrically, as j increases, $N_\pm^{\leq (j-1)}$ is nondecreasing, so Δ_{right} is nondecreasing. Therefore, if at (i, j) both deltas are ≥ 0, every further expansion has nonnegative delta, so no monotone expansion can improve the loss. Greedy expansion (pick the more negative delta) strictly decreases the loss until both deltas become nonnegative.