

Flattening Entropy from Comparison Geometry

Jonathan R. Landers

Abstract

This short note sketches a continuous geometric extension of *Algorithms as Geodesics: Flattening Entropy and Partial Curvature Removal in Sorting* and the companion note *Learned Monotone Bucketization as Partial Flattening*. In the base framework, flattening entropy is introduced as the coordinate information required to trivialize sorting, and ordered bucketization removes an explicit multinomial block of that entropy. The extension proposed here is to realize this obstruction directly in a continuous uncertainty geometry. Rather than placing the line element on raw array values, we place it on the simplex of order-type distributions. The Shannon entropy of this uncertainty state becomes the disorder potential, and the comparison model is encoded as an admissible line element calibrated so that a unit comparison step can dissipate at most $\log 2$ nats of uncertainty. Under this geometric normalization, the classical comparison lower bound and the ordered-bucketing entropy decomposition emerge as path-length statements. In this sense, flattening entropy appears as the integrated shadow of a nonflattenable comparison geometry.

1. Context

The main paper develops a model-dependent line-element viewpoint in which computation is motion in an admissible geometry, counting sort is a literal flattening on histogram space, and comparison sorting carries flattening entropy $F_{\text{cmp}}(n) = \log(n!) + O(1)$. The learned bucketization note then shows that if a monotone preprocessing map is learned from calibration data, the resulting hybrid runtime decomposes into ideal balanced partial flattening, a KL-type regret term, and learning cost.

The natural next step is to ask whether the comparison obstruction itself can be made to *emerge from the geometry* rather than being supplied externally as a counting argument. The proposal of this note is that the relevant continuous state space is not raw array space but the simplex of uncertainty states over order types. In that space, entropy becomes a potential, binary comparisons become admissible directions, and the comparison lower bound becomes a path-length inequality.

2. Uncertainty geometry on order types

Let

$$\Omega_n := \{\text{all order types of } n \text{ distinct keys}\}, \quad |\Omega_n| = n!$$

Write

$$\mathcal{U}_n := \text{int } \Delta(\Omega_n)$$

for the interior of the probability simplex on Ω_n . A point $q \in \mathcal{U}_n$ is a smooth uncertainty state over order types. The maximally uninformative state is the uniform distribution

$$u_n(\pi) = \frac{1}{n!}, \quad \pi \in \Omega_n,$$

and a fully resolved sorted state is a vertex δ_{π_*} .

Definition 1 (Disorder potential). *The uncertainty potential is the Shannon entropy*

$$\mathcal{H}(q) := - \sum_{\pi \in \Omega_n} q_\pi \log q_\pi.$$

All logarithms are natural.

This potential satisfies

$$\mathcal{H}(u_n) = \log(n!), \quad \mathcal{H}(\delta_{\pi_*}) = 0.$$

Thus the total disorder that must be removed in order to identify a unique sorted order is exactly $\log(n!)$ nats.

3. Comparison-calibrated line element

The main paper begins from a model-dependent line element

$$ds^2 = g_{\mu\nu}^{(A)}(x) dx^\mu dx^\nu,$$

or, more generally, a Finsler cost. The present note instantiates that idea on \mathcal{U}_n for the comparison model.

Definition 2 (Comparison-calibrated geometry). *A comparison-calibrated geometry on \mathcal{U}_n is a line element $g^{(\text{cmp})}$ together with a class of admissible tangent directions such that:*

- (i) *each unit admissible direction corresponds to one binary comparison observable of the form $\mathbf{1}[\pi(i) < \pi(j)]$;*
- (ii) *the entropy gradient is uniformly bounded by*

$$\|\nabla_{g^{(\text{cmp})}} \mathcal{H}(q)\| \leq \log 2 \quad \text{for all } q \in \mathcal{U}_n.$$

Condition (ii) is the geometric calibration of the binary comparison model: a unit admissible step cannot dissipate more than one bit, i.e. $\log 2$ nats, of uncertainty.

Theorem 1 (Geometric emergence of the comparison lower bound). *Let $q : [0, S] \rightarrow \mathcal{U}_n$ be any admissible path in a comparison-calibrated geometry with*

$$q(0) = u_n, \quad q(S) = \delta_{\pi_*}.$$

Then its $g^{(\text{cmp})}$ -length satisfies

$$L_{g^{(\text{cmp})}}(q) \geq \frac{\log(n!)}{\log 2}.$$

Proof. Parameterize the path by arc length, so that $\|\dot{q}(s)\|_{g^{(\text{cmp})}} = 1$ almost everywhere. Then

$$\frac{d}{ds} \mathcal{H}(q(s)) = \langle \nabla_{g^{(\text{cmp})}} \mathcal{H}(q(s)), \dot{q}(s) \rangle_{g^{(\text{cmp})}}.$$

Hence, by Cauchy–Schwarz and the comparison calibration,

$$\left| \frac{d}{ds} \mathcal{H}(q(s)) \right| \leq \|\nabla_{g^{(\text{cmp})}} \mathcal{H}(q(s))\| \|\dot{q}(s)\| \leq \log 2.$$

Integrating from 0 to S gives

$$\mathcal{H}(u_n) - \mathcal{H}(\delta_{\pi_*}) \leq (\log 2)S.$$

Since $\mathcal{H}(u_n) = \log(n!)$ and $\mathcal{H}(\delta_{\pi_*}) = 0$, we obtain

$$\log(n!) \leq (\log 2)S.$$

Because $S = L_{g^{(\text{cmp})}}(q)$ under arc-length parameterization, the result follows. \square

Remark 1. *The theorem does not require that $g^{(\text{cmp})}$ be Euclidean, nor even strictly Riemannian in spirit. A Finsler or control-geometric realization is equally natural. What matters is that the comparison model induces a restricted admissible geometry on uncertainty states, and that binary comparisons calibrate the maximal entropy dissipation rate.*

The main paper's flattening entropy is therefore recovered as a geometric quantity:

$$F_{\text{cmp}}(n) = \log(n!)$$

appears as the minimum entropy drop required to travel from the uniform uncertainty state to a resolved sorted vertex, while the comparison lower bound is the statement that this drop cannot be achieved with path length shorter than $\log(n!)/\log 2$.

4. Ordered bucketization as quotient and fiber

Let $\beta : K \rightarrow \{1, \dots, B\}$ be an ordered bucket map and let the corresponding occupancies on a given input be n_1, \dots, n_B . In the main paper, the exact entropy decomposition is

$$\log(n!) = \log \frac{n!}{\prod_{j=1}^B n_j!} + \sum_{j=1}^B \log(n_j!).$$

In the present geometric picture, this is the decomposition of uncertainty into base and fiber parts.

Theorem 2 (Bucket quotient decomposition). *Fix an ordered bucketization with occupancies n_1, \dots, n_B . Then the uncertainty space factors into:*

(i) *a quotient component, corresponding to cross-bucket order, of entropy*

$$\mathcal{H}_{\text{base}} = \log \frac{n!}{\prod_{j=1}^B n_j!};$$

(ii) *a fiber component, corresponding to unresolved within-bucket order, of entropy*

$$\mathcal{H}_{\text{fiber}} = \sum_{j=1}^B \log(n_j!).$$

Consequently,

$$\mathcal{H}(u_n) = \mathcal{H}_{\text{base}} + \mathcal{H}_{\text{fiber}}.$$

Proof. The ordered bucket labels determine the cross-bucket order exactly. The number of order types compatible with those labels is therefore

$$\prod_{j=1}^B n_j!.$$

Since the total number of order types is $n!$, the number of quotient states resolved by the bucket map is

$$\frac{n!}{\prod_{j=1}^B n_j!}.$$

Taking logarithms yields the stated base entropy. The unresolved within-bucket ambiguity is the product of the factorial fiber sizes, whose logarithm is the stated fiber entropy. Summing the two terms gives the total entropy $\log(n!)$. \square

Thus ordered bucketization is not merely a combinatorial preprocessing step. It is a literal quotienting of the uncertainty geometry. The flattening gain from the main paper is the entropy removed by passing to the base, while the residual comparison problem is the entropy still carried in the fibers.

Corollary 1 (Residual comparison length after bucketization). *For any comparison-based refinement after ordered bucketization,*

$$L_{\text{res}} \geq \frac{1}{\log 2} \sum_{j=1}^B \log(n_j!).$$

In other words, the residual within-bucket entropy is also a residual geometric length lower bound.

5. Connection to learned monotone bucketization

The learned-bucketization note proposes learning a monotone score whose induced occupancies are as balanced as possible under the ambient key distribution. In the present extension, that same idea becomes a statement about residual geometric length.

If $p_j = n_j/n$ and $u_B = (1/B, \dots, 1/B)$, then standard Stirling expansion gives

$$\sum_{j=1}^B \log(n_j!) = n \log \frac{n}{B} + n D_{\text{KL}}(p \| u_B) + \text{lower-order terms.}$$

Hence the residual comparison length obeys

$$L_{\text{res}} \gtrsim \frac{1}{\log 2} \left(n \log \frac{n}{B} + n D_{\text{KL}}(p \| u_B) \right),$$

up to the same lower-order Stirling corrections.

This identifies the KL regret term from the learned-bucketization note with excess residual geometric length beyond the balanced bucket baseline. In this language, learning a good monotone preprocessing chart means learning the cheapest quotient of the uncertainty geometry whose fibers are as small and as balanced as possible.

6. Closing perspective

The original framework introduced flattening entropy as the coordinate information needed to trivialize sorting and proved that ordered bucketization removes a precise multinomial block of it. The present extension suggests that this quantity is not merely a counting surrogate. It can be realized as the entropy drop of a continuous uncertainty geometry whose admissible motions are determined by the comparison model.

In this picture, the obstruction behind comparison sorting is a failure of global flattening internal to the admissible geometry. Counting sort escapes because it passes to a genuinely flat histogram chart. Ordered bucketization performs a partial quotient, removing a base layer of uncertainty while leaving comparison geometry on the fibers. Learned monotone bucketization then becomes the problem of learning a quotient map that minimizes the remaining geometric length.

The broad slogan is:

flattening entropy = integrated entropy drop required by the comparison geometry.

If this viewpoint can be formalized in a fully intrinsic metric or control-geometric language, then the “curvature blockage” of the main paper is no longer metaphorical. It is the continuous manifestation of the impossibility of globally flattening the comparison model while preserving admissible computational motion.