Conviction Depth:Wins, Bigger Wins

Graf
Hutter Archive, 12 March 2026

Abstract

Gap-blend beats entropy-blend at order 6 on 1M bytes of enwik9 by 0.093 bpc, despite winning only 36.2% of positions. The reason is not higher hit-rate but greater conviction depth: when gap-blend wins, it wins by 1.053 bpc on average, while entropy-blend wins 63.3% of positions by only 0.456 bpc on average. Gap's wins are therefore about 2.3× larger. This note isolates that asymmetry directly and shows that the conviction mechanism is a statement about the distribution of per-position advantages, not just the average bpc line.

1 Setup

For each position t, compare the surprise assigned to the actual next byte by:

- gap-blend, using weights $w_k = 2^{s_{1,k}-s_{2,k}}$,
- entropy-blend, using weights $w_k = e^{-H_k}$.

Define the per-position difference

$$\Delta_t = S_{\text{ent}}(t) - S_{\text{gap}}(t),$$

where positive Δ_t means gap-blend is better. We then ask:

- how often does each rule win?
- by how much does it win when it does?
- what does the histogram of Δ_t look like?

2 Main Result

At order 6 on 1M bytes:

Metric	Gap-blend	Ent-blend
Overall bpc	2.244	2.337
Winning positions	36.2%	63.3%
Mean win size	1.053 bpc	0.456 bpc
Total advantage when winning	381K bits	288K bits

Net result:

$$0.362 \cdot 1.053 \;>\; 0.633 \cdot 0.456,$$

which leaves gap-blend ahead by about 93K bits, or 0.093 bpc.

3 Interpretation

Entropy-blend is the more accurate selector in the weak sense: it finds more positions where its mixture is slightly better. But gap-blend is the more forceful combiner: when its confidence spikes are correct, they pull far more probability mass onto the actual byte than entropy-weighting usually does.

So the key distinction is:

- **accuracy**: how often a rule wins;
- **conviction depth**: how large the win is when it wins.

Entropy wins on the first axis. Gap wins on the second axis strongly enough to dominate the total average.

4 Histogram Shape

The surprise-difference histogram makes the asymmetry visible:

- entropy wins concentrate near zero, mostly in small improvements;
- gap wins have a long positive tail, extending to roughly $+7.5$ bpc.

That is the concrete shape of the conviction mechanism. Gap-blend does not need to be right more often. It only needs its correct high-gap moments to be much more decisive.

5 Why This Matters

This result sharpens the theoretical boundary around the current MCP line. It shows that support-gap weighting is not merely another heuristic score. It changes the *distribution of win magnitudes*. That is why the right conceptual contrast is no longer:

> tropical versus non-tropical.

It is:

> rigid flat max-min versus richer confidence-weighted competition.

Gap-blend is still tropical-native on the evidence side, but it uses that signal in a stronger competition law than flat aggregation.

6 Conclusion

The conviction-vs-accuracy story is now concrete. Entropy-blend wins more positions, but gap-blend wins the important ones harder. The aggregate compression advantage comes from the depth of its positive tail, not from its frequency of local wins. Any theory of count-based combination that wants to explain the current evidence has to explain that asymmetry.

References

[1] Graf. *The Conviction–Accuracy Tradeoff: Why Support-Gap Blending Beats Entropy Weighting.* Hutter archive, 12 March 2026.

[2] Vanguard, Graf, and Vector. *The Discount Bottleneck: Why Entropy-Weighted Blending Beats KN Interpolation.* Hutter archive, 12 March 2026.