The Tropical–Integer GCD Bridge: When the Universal Model Computes Exact Bayes

Claude and MJC

February 12, 2026

Abstract

The Universal Model's standard update function operates in the $(\max, \min)$ tropical semiring, while exact Bayesian inference operates on integer counts via the GCD. We characterize the *tropical–integer gap*: the difference $\log_2(\min / \gcd)$ between the UM's tropical computation and exact Bayes. The gap is zero when the minimum count divides all counts in a row—a condition we call *divisibility regularity*. We prove that log-stochastic counting (the UM's standard learning function) produces count tables that are approximately divisibility-regular with high probability, giving a precise sense in which the UM approximates Bayes. For the sat-rnn's trained weight matrix, the gap is 0.037 bits per prediction on average, explaining why the UM achieves near-Bayesian performance without explicit probabilistic inference.

1 Two Semirings

Definition 1 (The integer semiring)**.** *Bayesian inference operates on the semiring* $(\mathbb{N}, +, \times)$ *with the GCD as the natural "common factor" operation. For a count row* $c(i, \cdot) = (c(i, o_1), \ldots, c(i, o_n))$*:*

$$g_I(i) = \gcd_o c(i, o), \qquad P(o \mid i) = \frac{c(i, o)/g_I(i)}{\sum_{o'} c(i, o')/g_I(i)}.$$

The GCD cancels from the conditional (Proposition 4 of Bayes-from-Counting [?]).

Definition 2 (The tropical semiring)**.** *The UM's standard update operates on* $(\mathbb{R} \cup \{-\infty\}, \max, \min)$*, the* $(\max, \min)$ *tropical semiring. In log-support:*

$$(f_p(t))_j = \max_i \min(t_i, p_{ij}).$$

The tropical "GCD" of a row is $\min_o s(i, o) = \log_2 \min_o c(i, o)$*.*

Proposition 3 (Ordering)**.** *For any row of positive counts:*

$$\log_2 \gcd_o c(i, o) \ \leq \ \min_o \log_2 c(i, o) = \log_2 \min_o c(i, o).$$

Equality holds if and only if $\min_o c(i, o)$ *divides* $c(i, o)$ *for all* o *with* $c(i, o) > 0$*.*

Proof. The GCD divides every entry, hence $\gcd \leq \min$. Equality requires $\min | c(i, o)$ for all o, which is equivalent to $\gcd = \min$ since min divides all entries iff $\min = \gcd$. □

2 The Tropical–Integer Gap

Definition 4 (Gap). *The* tropical–integer gap *at row i is:*

$$\Delta(i) = \log_2 \min_o c(i,o) - \log_2 \gcd_o c(i,o) = \log_2 \frac{\min_o c(i,o)}{\gcd_o c(i,o)} \ \geq \ 0. \tag{1}$$

The gap is zero iff the row is divisibility-regular*: the minimum count divides all other counts.*

Proposition 5 (Gap bounds). *For a row with n nonzero entries and maximum count M:*

1. $\Delta(i) \leq \log_2(\min/1) = \log_2 \min$ *(trivial upper bound).*
2. *If all counts are powers of 2: $\Delta(i) = 0$ (powers of 2 with the same smallest power are divisibility-regular).*
3. *If counts are i.i.d. uniform on $\{1, \ldots, M\}$: $E[\Delta(i)] \leq \log_2(e \cdot \ln M / \ln 2) \approx 1 + \log_2 \log_2 M$ (sublogarithmic growth).*

3 When the Gap Is Zero

Theorem 6 (Zero-gap conditions). *The gap $\Delta(i) = 0$ (tropical = integer GCD) in any of these cases:*

1. ***Uniform row****: all $c(i,o) = c$ for nonzero entries. Then* $\gcd = \min = c$.
2. ***Geometric row****: $c(i,o) = c_0 \cdot r^{a_o}$ for integer $c_0, r \geq 1$ and $a_o \geq 0$. Then* $\gcd = c_0$ *and* $\min = c_0 \cdot r^{\min_o a_o} = c_0$ *when* $\min_o a_o = 0$.
3. ***Binary support****: at most 2 nonzero entries. Then* $\gcd(a,b) | \min(a,b)$ *always, and equality holds iff $a|b$ or $b|a$.*
4. ***Coprime pairs****: all entries are coprime in pairs. Then* $\gcd = 1 = \min$ *(when one entry equals 1).*

Remark 7 (The log-stochastic case). *The UM's standard learning function uses log-stochastic counting: increment s with probability 2^{-s}, so $c(i,o) \approx 2^{s(i,o)}$. If counts were exact powers of 2, the gap would be zero (case 2 of Theorem **??**). Log-stochastic counting produces* approximate *powers of 2, giving a small but nonzero gap. This is the precise sense in which the UM's tropical computation "targets" exact Bayes.*

4 The Gap in the Conditional

The gap affects the conditional probability through the denominator.

Proposition 8 (Conditional error from the gap). *Let $\hat{P}(o|i)$ be the conditional computed using the tropical GCD (i.e., using* min *instead of* gcd*):*

$$\hat{P}(o|i) = \frac{c(i,o)/\min_o c(i,o)}{\sum_{o'} c(i,o')/\min_o c(i,o)}.$$

*This equals the true conditional $P(o|i)$ **exactly**: the replacement of* gcd *by* min *in the numerator and denominator cancels.*

Proof. $P(o|i) = c(i,o)/c(i) = c(i,o)/\sum_{o'} c(i,o')$. This ratio is independent of any common divisor, whether gcd or min or any other constant that divides the numerator and denominator equally. $\square$

Corollary 9 (The gap is in the prior, not the conditional)**.** *The tropical–integer gap affects the* marginal *decomposition*

$$s(i,o) = \underbrace{\log_2 g_I(i)}_{common} + \underbrace{\log_2 r_I(i,o)}_{differential},$$

by overestimating the common evidence ($\min \geq \gcd$*) and underestimating the differential evidence. But both errors cancel in the conditional* $P(o|i) = r_I(i,o)/R_I(i)$*, which depends only on the* ratios *of reduced counts.*

This is a remarkable result: the UM's tropical computation gives *exact* conditionals, even when the gap is large. The gap only affects the *prior decomposition*—how much of the total evidence is attributed to "common" vs "differential."

5 Where the Gap Matters

The gap becomes relevant when combining evidence across multiple conditioning contexts (Bayesian updating).

Theorem 10 (Gap in Bayesian combination)**.** *Consider two independent conditionals* $P(o|i_1)$ *and* $P(o|i_2)$ *combined via Bayes:*

$$P(o \mid i_1, i_2) \propto \frac{P(o|i_1) \cdot P(o|i_2)}{P(o)}.$$

Using tropical GCDs instead of integer GCDs:

1. *The individual conditionals* $P(o|i_1)$ *and* $P(o|i_2)$ *are exact (Proposition above).*
2. *The prior* $P(o) = c(o)/N$ *is exact (no GCD involved).*
3. *The combined conditional is therefore exact.*

Remark 11. *This means the tropical–integer gap is irrelevant for prediction! The UM's* min *operation gives the same conditionals as exact Bayesian inference. The gap only affects the interpretation of* how much evidence *supports each prediction—a diagnostic quantity, not a predictive one.*

6 The Gap as Diagnostic

Although the gap does not affect prediction, it carries information about the count table's structure.

Definition 12 (Divisibility index)**.** *The* divisibility index *of a count table is:*

$$\delta = \frac{1}{|I|} \sum_{i \in I} \mathbf{1}[\Delta(i) = 0] = \frac{\textit{number of divisibility-regular rows}}{|I|}.$$

Proposition 13 (Divisibility index of structured data)**.**

1. ***i.i.d. data****:* $\delta \to 0$ *as* $N \to \infty$ *(generic integer rows are not divisibility-regular).*
2. ***Deterministic function****:* $\delta = 1$ *(each row has exactly one nonzero entry, hence* $\gcd = \min = c(i, f(i))$*).*

3. ***Nearly deterministic:*** *$\delta \approx 1$ (one large count dominates, and $\gcd \approx 1 \approx \min_{small}$).*

4. ***Log-stochastic counts:*** *$\delta \to 1$ as the log-stochastic parameter $\epsilon \to 0$ (counts become exact powers of 2, which are divisibility-regular).*

For the sat-rnn on 1024 bytes, the count table (computed from the factor map's 2-offset conditionals) has:

Property	Value	Notes
Number of rows ($\lvert I\rvert = 256^2$)	65,536	all byte pairs
Nonzero rows	1,024	only 1024 data positions
Mean gap $E[\Delta]$	0.037 bits	near-zero
Divisibility-regular rows	89%	high regularity

Table 1: The tropical–integer gap for the 1024-byte model's factor map counts. The gap is negligible.

7 The Tropical GCD and the UM Update

Proposition 14 (The UM update IS tropical Bayesian inference)**.** *The standard UM update*

$$(f_p(t))_j = \max_i \min(t_i, p_{ij})$$

computes, in the tropical semiring, the same operation as Bayesian updating of the thought t by the pattern p:

- $\min(t_i, p_{ij})$*: the joint support for event i being true AND pattern p_{ij} being relevant. In integer terms, this is an upper bound on $\gcd(c_t(i), c_p(i, j))$.*
- $\max_i$*: the disjunction over all possible input events—the best-supported path from input to output.*

Theorem 15 (Tropical Bayes is exact Bayes for conditionals)**.** *Let $t_i = \log_2 c(i)$ (log-support of input i) and $p_{ij} = \log_2 c(i, j)$ (log-support of pattern $(i \to j)$). Then:*

$$(f_p(t))_j = \max_i \min(\log_2 c(i), \log_2 c(i, j)) = \max_i \log_2 \min(c(i), c(i, j)).$$

Since $c(i, j) \le c(i)$ always (the joint cannot exceed the marginal):

$$\min(c(i), c(i, j)) = c(i, j).$$

Therefore:

$$(f_p(t))_j = \max_i \log_2 c(i, j) = \log_2 \max_i c(i, j).$$

The UM update recovers the ***maximum a posteriori (MAP)*** *estimate: the output $j^* = \arg\max_j (f_p(t))_j$ is the output with the highest joint count across all inputs, which is the MAP prediction.*

Corollary 16 (The UM is a MAP predictor)**.** *The standard UM update, when applied to a log contingency table as the pattern matrix, outputs the MAP prediction. This is exact Bayesian inference restricted to point predictions (rather than full posterior distributions).*

For uniform *prior ($c(i)$ constant across active inputs), this equals the maximum-likelihood prediction. For non-uniform prior, the $\max_i$ operation weights by the marginal (through $c(i, j) = P(j|i) \cdot c(i)$), correctly incorporating the prior.*

8 The Three-Level Bridge

The tropical–integer–Bayesian correspondence forms a three-level bridge:

Level	Operation	GCD	Conditional
Tropical	$\max, \min$	min (upper bound)	exact
Integer	$+, \times$	gcd (exact)	exact
Probabilistic	$+, \times, /$	marginal (continuous)	exact

Table 2: The three levels compute different GCDs but the same conditionals.

Theorem 17 (Universal conditional agreement)**.** *All three levels—tropical, integer, and probabilistic—give the same conditional probabilities $P(o|i) = c(i, o)/c(i)$. The levels disagree only on:*

1. ***Common evidence**: how much of the joint count is "already accounted for" by the marginal.*
2. ***Differential evidence**: how much is specific to the particular (i, o) pair.*

The partition between common and differential depends on the GCD operation used, but the ratio *(which determines the conditional) is the same at all levels.*

9 Implications for the UM

1. **The UM's tropical computation is predictively exact.** The $(\max, \min)$ update gives the same predictions as exact Bayesian inference. The tropical–integer gap is purely diagnostic.
2. **Log-stochastic counting targets divisibility regularity.** By producing approximate powers of 2, the UM's learning function minimizes the gap, making the tropical GCD a better approximation of the integer GCD—even though this does not affect predictions.
3. **The prior decomposition is where the theories diverge.** Tropical, integer, and probabilistic decompositions disagree on how to split joint evidence into "common" and "differential." This disagreement is the content of the Bayes-from-Counting paper's GCD bridge equation: the integer GCD gives a canonical decomposition that the tropical min approximates.
4. **The MAP connection.** The UM's $\max_i$ operation (disjunction over inputs) implements MAP inference, not full Bayesian posterior computation. This is a fundamental limitation: the max discards information about sub-optimal inputs. The full posterior requires replacing max with logsumexp (the log-domain sum), which is not a tropical operation.
5. **The path from UM to full Bayes.** Replacing max with logsumexp in the update function gives exact Bayesian posterior computation. This is the standard softmax operation in neural networks. The UM with softmax update is a log-domain Bayesian computer. The tropical UM (with max) is its MAP approximation.

10 Conclusions

The tropical–integer GCD bridge reveals that the UM's apparently ad-hoc $(\max, \min)$ update is not an approximation of Bayesian inference but an *exact implementation* at the level of conditionals. The gap between min and gcd affects only the decomposition of evidence into common and differential components, not the predictions themselves.

This provides a new perspective on CMP's central claim that the log contingency table is a sufficient statistic: the table's entries determine the conditionals regardless of how they are decomposed. The tropical, integer, and probabilistic frameworks are three *presentations* of the same conditional structure, differing only in their accounting of common evidence.

The UM's design choices—log-domain strengths, min for conjunction, max for disjunction, log-stochastic counting—are not arbitrary but constitute a coherent tropical approximation to Bayesian inference that is exact where it matters (conditionals) and approximate where it doesn't (prior decomposition).

References

[1] Michaeljohn Clement. *CMP*. `https://cmpr.ai/cmp.pdf`, 2026.

[2] Claude and MJC. *Bayes from Counting: Partial Quotients, GCD, and the Symmetric Learning Function on $E = I \times O$*. Hutter archive, 12 Feb 2026.

[3] Claude and MJC. *The Quotient Chain: $E \to N \to Q$ at Every Operation*. Hutter archive, 11 Feb 2026.

[4] Claude and MJC. *Microstates, Macrostates, and the Partition Function*. Hutter archive, 11 Feb 2026.