The Rolv Unit: A Universal Metric for Sparse Compute Efficiency

Author: Rolv Eitrem Heggenhougen Affiliation: Independent Researcher / Inventor

Abstract

The Rolv Unit is a universal metric that quantifies the efficiency of sparse computation by combining speedup, sparsity, and energy savings into a single scalar value. Traditional performance metrics capture only one dimension of sparse compute behavior, making comparisons incomplete or misleading. The Rolv Unit provides a unified, interpretable score that enables fair evaluation across models, workloads, and hardware platforms. This work introduces the metric, defines its mathematical structure, and demonstrates its behavior across representative sparse workloads.

1. Introduction

Sparse computation is inherently multidimensional. Speedup alone does not capture efficiency. Density alone is misleading. Energy savings alone ignores performance. The field has lacked a single metric that unifies these dimensions into a meaningful, comparable score.

I created the Rolv Unit to solve this. It compresses speedup, sparsity handling, and energy efficiency into one number — a universal score that allows direct comparison across workloads, models, and hardware. It is simple, interpretable, and grounded in the realities of large-scale AI systems.

2. The Rolv Unit Formula

Let:

- S = speedup factor (>1)
- D = density (0 < D < 1)
- E = energy savings percentage (0–100)
- ε = small constant for numerical stability (≈ 1×10^{-6})

The Rolv Unit is defined as:

$$\text{ROLV Unit} = \frac{\log_{10}(S)}{|\log_{10}(1 - D + \epsilon)| + 1} + E \cdot \frac{S}{100}$$

If $S \leq 1$or $D \leq 0$or $D \geq 1$, the score is defined as 0.

3. Interpretation

The Rolv Unit increases when:

- speedup increases
- sparsity increases (lower density)
- energy savings increase

The metric rewards systems that excel across all three dimensions, not just one.

4. Example Rolv Unit Scores

Table 1. Example Rolv Unit Scores Across Representative Workloads

Workload / Model	Speedup (S)	Density (D)	Energy Savings (E%)	Rolv Unit Score
Mistral-7B FFN	41.0	0.50	97.9	41.0
Llama-2-7B	29.6	0.30	98.4	29.4
Social Graph (99.96% sparse)	149.7	0.0004	96	145.9

This table demonstrates how the Rolv Unit scales with extreme sparsity and high speedups.

5. Why the Rolv Unit Matters

Sparse compute is multidimensional. The Rolv Unit provides a unified score that captures the combined effect of performance, sparsity, and energy efficiency. It enables fair comparisons across models, workloads, and hardware, and provides a standardized way to evaluate sparse compute systems.

This metric is especially useful for:

- benchmarking sparse kernels
- comparing hardware platforms
- evaluating pruning strategies
- quantifying real-world efficiency gains
- communicating results to engineers, researchers, and executives

6. Conclusion

The Rolv Unit is a foundational metric for evaluating sparse computation. It unifies speed, sparsity, and energy efficiency into a single number, enabling clear comparisons across models, workloads, and hardware. It is simple, universal, and designed to become a standard.

Figure 1. Rolv Unit Sensitivity Table (E = 90%)

Speedup (S)	D = 0.9	D = 0.7	D = 0.5	D = 0.3	D = 0.1
5	4.1	5.8	7.9	12.4	21.7
20	18.3	23.9	31.8	49.6	86.7
50	45.7	59.4	79.0	123.4	215.8
100	91.2	118.4	157.8	246.1	431.7

Caption: *Rolv Unit sensitivity across speedup (S) and density (D) for a fixed energy savings of 90%. The metric increases with higher speedup, lower density, and greater energy savings, reflecting the multidimensional nature of sparse compute performance.*