

# Late Breaking Results: Input Loss Curvature as a Predictor of Sample Vulnerability to Hardware-Noise

Deepika Sharma, Deepak Ravikumar, Chih-Hsing Ho, Sangamesh D. Kodge and Kaushik Roy

*School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA*

{sharm444, dravikum, ho216, skodge, kaushik}@purdue.edu

**Abstract**—Analog in-memory computing (AIMC) accelerators can deliver significant energy efficiency over conventional architectures, but their accuracy is limited by device and circuit-level non-idealities. While prior work has characterized the effects of these non-idealities at model or layer granularity, their impact on individual samples remains largely unexplored. In this work, we show that an input’s vulnerability to such non-idealities can be strongly predicted by its loss curvature, a metric capturing how sharply the loss changes under small input perturbations. Across multiple models, datasets, and non-idealities, our experiments reveal a strong positive correlation between input loss curvature and analog non-ideality-induced failures, with failure rates increasing substantially for high-curvature samples. These findings uncover a previously overlooked, sample-level dimension of hardware robustness and suggest new opportunities for input-aware strategies for addressing non-idealities.

**Index Terms**—analog in-memory computing, non-idealities, input loss curvature

## I. INTRODUCTION

Analog in-memory computing (AIMC) accelerators promise substantial energy and throughput benefits for deep neural network inference, particularly in edge-constrained environments [1]–[3]. However, these gains come at the cost of dealing with the uncertainty induced by device and circuit-level non-idealities such as write variability, conductance drift, read noise, and Digital-to-Analog/Analog-to-Digital (DAC/ADC) quantization [4]. Prior works focus on modeling analog non-idealities, quantifying their effect on model accuracy, and mitigating these effects through compensation [5] or hardware-aware training [6]–[9].

In this work, we take a data-centric perspective and uncover a key pattern: not all inputs are equally sensitive to hardware noise, and an input’s susceptibility aligns closely with the *curvature* of the loss landscape surrounding that input. *Curvature*, as defined in [10], [11], measures how sharply the model loss changes under small perturbations to the input [11]–[13]. Here, we investigate whether this input-sensitivity metric also predicts how a sample responds to model perturbations introduced by analog non-idealities. We find a strong positive correlation, i.e., samples with higher input curvature are considerably more likely to fail when subjected to realistic analog noise.

To the best of our knowledge, this is the first empirical study establishing a direct relationship between sample loss curvature and failure under analog non-idealities. Using controlled simulations on IBM’s AIHWKit [14], we select specific non-idealities and analyze their impact on various input samples. We observe a clear upward trend in failure rate as curvature

increases, indicating that sample-level geometry plays a critical role in robustness to hardware noise. These findings expose a previously overlooked, sample-level dimension of hardware robustness and motivate data-aware robustness strategies such as curvature-informed calibration and noise mitigation, and hybrid analog/digital routing, where high-curvature samples are selectively redirected to reliable digital paths. By shifting focus from model- to sample-level behavior, this work can potentially enable a more workload-aware hardware–algorithm co-design.

## II. EXPERIMENTAL SETUP

**Models and datasets.** We conduct experiments on ResNet-18 [15] and ResNet-50 models trained on CIFAR-10/100 [16], Tiny-ImageNet [17], [18], and ImageNet-1K [19] datasets. For ImageNet, we use the official pretrained models from `torchvision`. For the remaining datasets, we train the models from scratch following standard recipes, and select the best checkpoint on the validation set for downstream analysis.

**Curvature computation.** Following [11], we estimate per-sample input loss curvature using Hutchinson’s trace estimator. Given an input-label pair  $(x, y)$ , we draw  $n$  random vectors  $v_i$  from the Rademacher distribution (each entry independently chosen from  $\{-1, +1\}$ ), and perturb the input by a small step size  $h$ . For each  $v_i$ , we measure how the loss gradient changes under this perturbation, and define

$$\text{curv}(x) = \frac{1}{n} \sum_{i=1}^n \|\nabla_x (L(x + hv_i, y) - L(x, y))\|_2, \quad (1)$$

We use  $h = 10^{-3}$  and  $n = 10$  across all experiments and compute curvature once using the final full-precision model and do not aggregate across epochs, consistent with [12].

**Analog inference with non-idealities.** We simulate analog in-memory inference using IBM’s AIHWKit [7], [14]. All convolutional and linear layers are mapped to analog crossbar tiles using a standard inference-oriented configuration with digital biases, symmetric weight remapping, and per-layer weight clipping. First, we construct an “ideal” analog model in which no circuit-level noise, device variation, or quantization is present as a baseline. We then select three representative types of non-idealities:

- **ADC quantization (ADC):** finite-precision conversion at the tile outputs, modeled using a 9-bit uniform ADC.
- **Static weight noise (WNoise):** small readout deviations in the effective programmed weights, capturing short-term conductance variation.

TABLE I  
ANALOG-ONLY FAILURE RATE (FR) IN THE LOWEST AND HIGHEST CURVATURE DECILES (D1 AND D10) FOR RESNET-18/50 ON CIFAR10/100, TINY-IMAGENET AND IMAGENET-1K UNDER THE FULL “ADC + WNOISE + PCM” PROFILE

Model	Dataset	D1 FR (%)	D10 FR (%)
ResNet-18	CIFAR-10	0.14	3.4
	CIFAR-100	1.34	13.75
	Tiny-Imagenet	26.45	56.62
	ImageNet-1K	73.69	94.93
ResNet-50	CIFAR-10	2.67	29.74
	CIFAR-100	8.54	44.4
	Tiny-Imagenet	41.48	69.45
	ImageNet-1K	97.76	99.3

- **Device noise and drift (PCM):** a calibrated model of phase-change-memory (PCM) programming noise, read noise, and temporal conductance drift.

We use AIHWKit’s PCM model, but the analysis is generally applicable to other resistive memories.

These effects are toggled independently, yielding  $2^3$  combinations: ideal, each independent non-ideality, pairwise combinations, and a full “ADC + WNoise + PCM” profile. For every configuration, analog weights are programmed once, and inference is performed without retraining or compensation.

**Failure rate analysis.** For each model, dataset, and non-ideality profile, we run inference and record the indices of misclassified samples in the train and test splits. We define *analog-only* failures as samples that are correctly classified under full precision (digital) inference but misclassified under a given analog profile. To study how robustness varies with curvature, we sort all samples by curvature, partition them into ten equal-sized bins (deciles), and compute the analog-only failure rate for each bin.

### III. RESULTS

Table I reports the analog-only failure rates in the lowest and highest curvature deciles (D1 and D10) for all evaluated dataset–model pairs under the full non-ideality profile. Across every setting, high-curvature samples (D10) fail more often than low-curvature samples (D1), establishing input loss curvature as a strong and consistent predictor of vulnerability.

Absolute failure rates increase substantially as we move from CIFAR to Tiny-ImageNet and ImageNet. This behavior is expected, since larger and more complex datasets, combined with analog inference performed without calibration or compensation, result in a large number of failures across the whole dataset. Even in this saturated regime, input loss curvature preserves its discriminative power. For example, ResNet-50 on ImageNet has an analog-only failure rate of 97.76% for D1 vs. 99.3% for D10. Deeper models such as ResNet-50 also show higher overall degradation because non-idealities accumulate across more layers, yet the curvature-driven separation between D1 and D10 remains pronounced.

To illustrate the full trend, Fig. 1 plots failure rate versus curvature decile for ResNet-18 on CIFAR-100 across individual non-ideality profiles. The monotonic relationship between curvature and analog failure rate persists across ADC quantization, static weight noise, and PCM device noise, although the

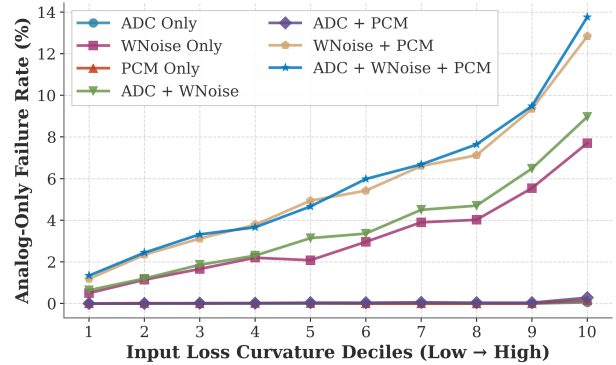


Fig. 1. Analog-only failure rate across curvature deciles under various non-ideality profiles for Resnet-18 trained on CIFAR-100.

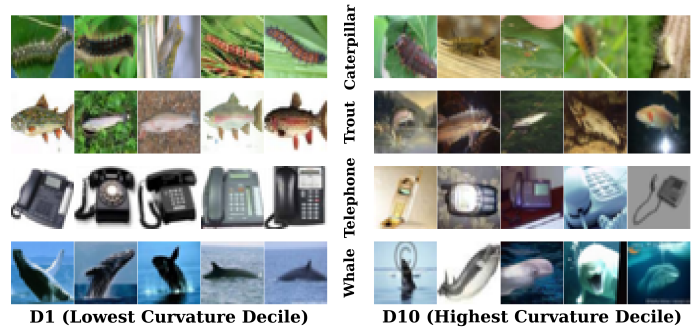


Fig. 2. Example low and high curvature samples from CIFAR 100

absolute scales differ. These results demonstrate that input loss curvature captures a fundamental sample-level fragility that is largely agnostic to the specific non-ideality.

Finally, Fig. 2 highlights qualitative differences between low and high curvature samples. Low-curvature images tend to be clean, centered, and unambiguous, whereas high-curvature samples often exhibit blur, occlusion, unusual viewpoints, or atypical scale. These visual factors might be a cause of their higher sensitivity to analog noise.

### IV. DISCUSSION AND CONCLUSION

Our results show that input loss curvature, a metric computed solely from the dataset and the trained model, strongly predicts which inputs are vulnerable to analog non-idealities. This work demonstrates that hardware robustness is not uniform across samples, but instead follows the structure already present in the model’s loss landscape. This perspective opens several promising research directions. If sample vulnerability is predictable, then analog accelerators could prioritize compensation, precision allocation, or retraining effort on the minority of high-curvature samples that dominate failure rates. Conversely, low-curvature inputs may tolerate more aggressive quantization or hardware noise, creating opportunities for dynamic or sample-aware hardware–algorithm co-design as a promising path for improving the accuracy–efficiency trade-off in future analog AI systems.

#### ACKNOWLEDGMENT

This work was supported in part by DARPA/SRC JUMP CoCoSys and the US DoE.

## REFERENCES

- [1] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, and E. Eleftheriou, "Memory devices and applications for in-memory computing," *Nature nanotechnology*, vol. 15, no. 7, pp. 529–544, 2020.
- [2] S. Yu, H. Jiang, S. Huang, X. Peng, and A. Lu, "Compute-in-memory chips for deep learning: Recent trends and prospects," *IEEE circuits and systems magazine*, vol. 21, no. 3, pp. 31–56, 2021.
- [3] R. Khaddam-Aljameh, M. Stanisavljevic, J. F. Mas, G. Karunaratne, M. Braendli, F. Liu, A. Singh, S. M. Müller, U. Egger, A. Petropoulos *et al.*, "Hermes core—a 14nm cmos and pcm-based in-memory compute core using an array of 300ps/lsb linearized cco-based adcs and local digital processing," in *2021 Symposium on VLSI Circuits*. IEEE, 2021, pp. 1–2.
- [4] V. Joshi, M. Le Gallo, S. Haefeli, I. Boybat, S. R. Nandakumar, C. Piveteau, M. Dazzi, B. Rajendran, A. Sebastian, and E. Eleftheriou, "Accurate deep neural network inference using computational phase-change memory," *Nature communications*, vol. 11, no. 1, p. 2473, 2020.
- [5] C. Mackin, M. J. Rasch, A. Chen, J. Timcheck, R. L. Bruce, N. Li, P. Narayanan, S. Ambrogio, M. Le Gallo, S. Nandakumar *et al.*, "Optimised weight programming for analogue memory-based deep neural networks," *Nature communications*, vol. 13, no. 1, p. 3765, 2022.
- [6] G. Charan, A. Mohanty, X. Du, G. Krishnan, R. V. Joshi, and Y. Cao, "Accurate inference with inaccurate rram devices: A joint algorithm-design solution," *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, vol. 6, no. 1, pp. 27–35, 2020.
- [7] M. Le Gallo, C. Lammie, J. Büchel, F. Carta, O. Fagbohunge, C. Mackin, H. Tsai, V. Narayanan, A. Sebastian, K. El Maghraoui *et al.*, "Using the ibm analog in-memory hardware acceleration kit for neural network training and inference," *APL Machine Learning*, vol. 1, no. 4, 2023.
- [8] M. J. Rasch, C. Mackin, M. Le Gallo, A. Chen, A. Fasoli, F. Odermatt, N. Li, S. Nandakumar, P. Narayanan, H. Tsai *et al.*, "Hardware-aware training for large-scale and diverse deep learning inference workloads using in-memory computing-based accelerators," *Nature communications*, vol. 14, no. 1, p. 5282, 2023.
- [9] M. J. Rasch, D. Moreda, T. Gokmen, M. Le Gallo, F. Carta, C. Goldberg, K. El Maghraoui, A. Sebastian, and V. Narayanan, "A flexible and fast pytorch toolkit for simulating training and inference on analog crossbar arrays," in *2021 IEEE 3rd international conference on artificial intelligence circuits and systems (AICAS)*. IEEE, 2021, pp. 1–4.
- [10] S.-M. Moosavi-Dezfooli, A. Fawzi, J. Uesato, and P. Frossard, "Robustness via curvature regularization, and vice versa," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9078–9086.
- [11] I. Garg, D. Ravikumar, and K. Roy, "Memorization through the lens of curvature of loss function around samples," in *International Conference on Machine Learning (ICML)*. PMLR, 2024, pp. 15 083–15 101.
- [12] D. Ravikumar, E. Soufleri, and K. Roy, "Curvature clues: Decoding deep learning privacy with input loss curvature," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 37, pp. 20 003–20 030, 2024.
- [13] D. Ravikumar, E. Soufleri, A. Hashemi, and K. Roy, "Unveiling privacy, memorization, and input curvature links," in *Proceedings of the 41st International Conference on Machine Learning (ICML)*, 2024, pp. 42 192–42 212.
- [14] M. Rasch, T. Gokmen, D. Moreda, M. Le Gallo, and K. El Maghraoui, "Ibm analog hardware acceleration kit," *GitHub*, v. 0.8. 0. <https://github.com/IBM/aihwkit>, 2023.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [16] A. Krizhevsky, G. Hinton *et al.*, "Learning multiple layers of features from tiny images," 2009.
- [17] Y. Le and X. Yang, "Tiny imagenet visual recognition challenge," *CS 231N*, vol. 7, no. 7, p. 3, 2015.
- [18] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.
- [19] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.