Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A geometric dissipation bound on the lifespan of gradient-based adaptation

Zhan Peng Jing

jing597277@gmail.com

Department of Mathematics, Northeastern University https://orcid.org/0009-0007-2823-6647

Physical Sciences - Article

Keywords:

Posted Date: March 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8482405/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

A geometric dissipation bound on the lifespan of gradient-based adaptation

Zhanpeng Jing[1,*]

[1]Department of Mathematics, Northeastern University, Shenyang 110819, China

[*]Email: 20231584@stu.neu.edu.cn

Abstract

Adaptive learning systems sustain functionality by continuously modifying internal parameters in response to environmental change. While such adaptation enhances short-term robustness, its long-term physical cost remains poorly characterized. Here we show that gradient-based adaptive systems are constrained by a finite geometric dissipation budget Ω, which fundamentally limits their functional lifespan. Through a sequence of controlled numerical experiments, we demonstrate that system collapse consistently occurs when cumulative informational dissipation reaches this intrinsic bound, independent of environmental drift rate, defensive threshold, or architectural complexity. We further show that hierarchical organization redistributes dissipation temporally but does not expand Ω, and that the same dissipation-limited collapse persists across distinct neural architectures. Together, these results identify a thermodynamic horizon for adaptive learning, suggesting that sudden failure arises from geometric exhaustion rather than algorithmic deficiency.

Introduction

The physical limits of computation are traditionally formulated in energetic terms. Landauer's principle establishes a minimum thermodynamic cost for irreversible logical operations, linking information processing to entropy production [1, 2]. These foundational results constrain static or logically irreversible computation. By contrast, modern artificial intelligence systems operate in a fundamentally different regime, continuously adapting through gradient-based parameter updates.

Despite their remarkable empirical success, the physical constraints governing sustained adaptation remain largely unexplored. Learning dynamics are typically evaluated through optimization efficiency or generalization performance, implicitly assuming that adaptation can proceed indefinitely. Recent work, however, has revealed deep connections between stochastic gradient descent and non-equilibrium thermodynamics, suggesting that learning dynamics are intrinsically dissipative processes [3, 4].

Here we address a complementary question: *Is there a fundamental physical limit to how long an adaptive learning system can continue to function?* By explicitly tracking the geometric cost of parameter evolution, we show that adaptive systems possess a finite dissipation capacity Ω. Once this budget is exhausted, collapse becomes unavoidable, regardless of learning rate, defensive strategy, or network architecture. From this perspective, adaptation is not limited by optimization inefficiency but by the irreversible consumption of a finite geometric resource. We refer to this intrinsic capacity as the *geometric dissipation budget* Ω, which plays a role analogous to an action bound in continuous dynamical systems. Rather than quantifying performance, Ω constrains the total extent of viable adaptation, thereby defining an intrinsic horizon beyond which learning dynamics can no longer remain stable.

Results

Experiment I: Survival landscapes under environmental drift

We first examined how functional lifespan τ depends on environmental drift rate and defensive threshold. Agents were exposed to non-stationary environments spanning drift rates from 10^{-8} to 10^{-1}, under three collapse thresholds (0.65, 0.85, 1.20).

We observe a pronounced non-monotonic fragility: overly strict defenses precipitate premature collapse, while excessively permissive thresholds enter a regime of reduced sensitivity that prolongs survival. This behavior cannot be explained by learning rate alone, indicating the presence of a deeper governing constraint.

Experiment II: Microscopic dissipation dynamics and the Ω invariant

To identify the mechanism underlying collapse, we directly tracked the cumulative geometric dissipation Ω throughout adaptation. Parameter updates were decomposed into effective dissipation rates, allowing us to monitor their temporal evolution.

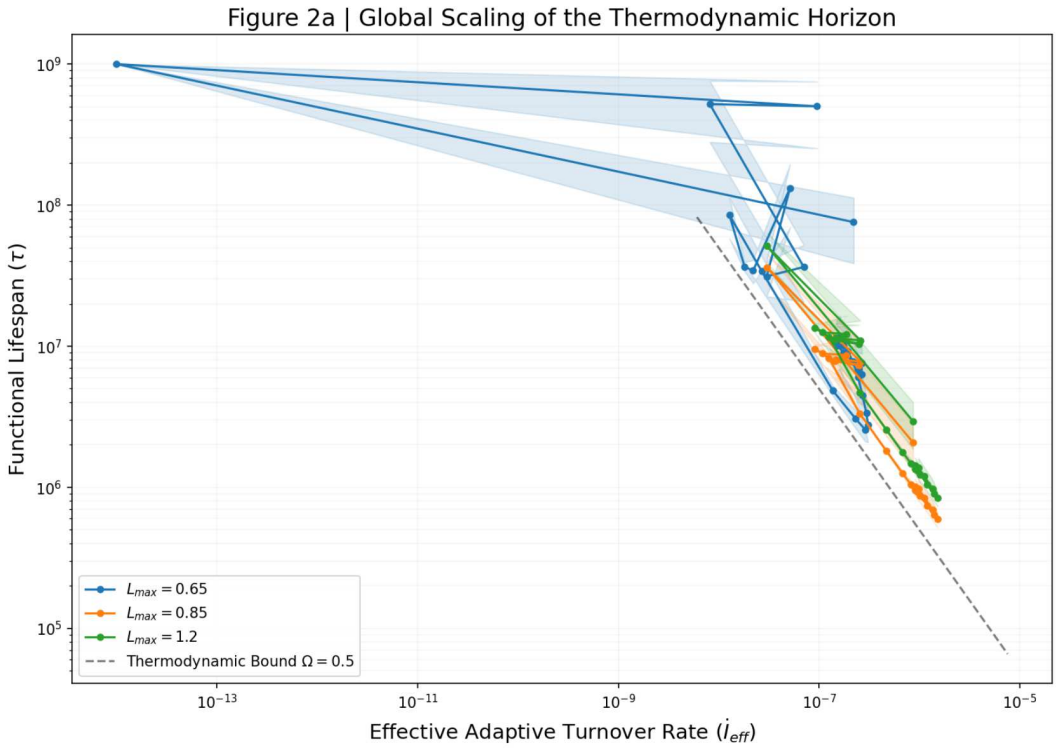


Figure 1: **Survival landscape of adaptive systems.** Functional lifespan τ as a function of environmental drift rate. The intermediate threshold (0.85) exhibits abrupt collapse at high drift rates, whereas the loose threshold (1.20) displays a re-entry into a low-sensitivity regime at extreme drift.

Despite substantial differences in temporal dynamics, all systems collapse when cumulative dissipation reaches a common value Ω. Recasting lifespan in dissipation space reveals a striking invariant collapse:

These results indicate that $\Omega \approx \tau \times \dot{I}_{\text{eff}}$ functions as a quasi-invariant, with sudden failure arising from the exhaustion of an internal geometric resource. We emphasize that the geometric dissipation budget Ω is a dimensionless quantity. Here, the functional lifespan τ carries the dimension of time, while the effective informational turnover rate $\dot{I}_{\text{eff}}$ has units of inverse time, corresponding to the rate of parameter reconfiguration on the parameter manifold. Their product therefore defines a scale-free quantity that characterizes the total extent of viable adaptation rather than its temporal duration.

More precisely, we interpret Ω as the cumulative geometric path length traced by the parameter trajectory on a statistical manifold $\mathcal{M}$ endowed with an information-theoretic metric. In our implementation, this metric is operationally induced by gradient dynamics, rendering Ω an intrinsic measure of dissipated adaptive capacity rather than a performance-dependent quantity. Formally, we define the cumulative geometric dissipation as the path integral.

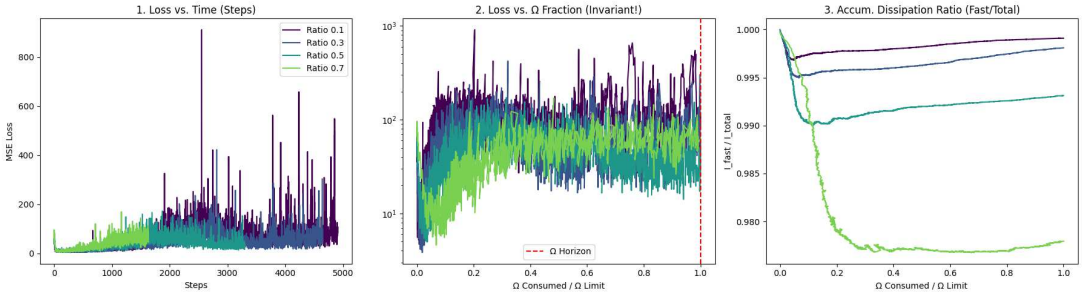


Figure 2: **Microscopic dissipation dynamics.** Accumulated dissipation and dissipation ratios under different defensive thresholds. A pronounced dissipation minimum emerges for loose thresholds, accounting for the apparent resilience observed in Experiment I.

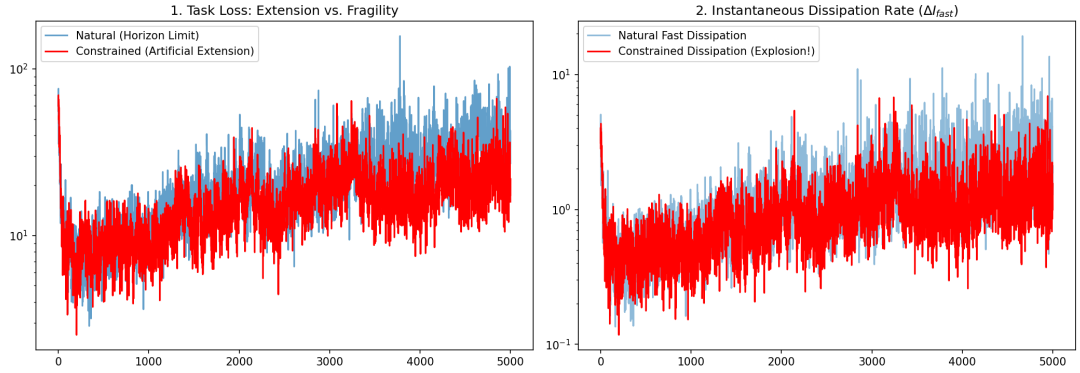


Figure 3: **Invariant collapse at the geometric dissipation horizon.** Across drift rates and defensive thresholds, collapse consistently occurs at a fixed dissipation budget Ω, revealing a conserved geometric limit.

$$\Omega(t) = \int_0^t \sqrt{\langle \dot{\theta}(\tau), G(\theta)\dot{\theta}(\tau) \rangle}\, d\tau, \tag{1}$$

where $G(\theta)$ represents the effective metric of the parameter manifold (e.g., the Fisher Information Matrix or its Euclidean approximation). To make this notion explicit, we define the effective informational turnover rate as the instantaneous geometric displacement speed in parameter space,

$$\dot{I}_{\text{eff}} \equiv \left\| \frac{d\boldsymbol{\theta}}{dt} \right\|_2, \tag{2}$$

where $\boldsymbol{\theta}$ denotes the full set of trainable parameters and the norm is taken with respect to the Euclidean metric (implying $G(\theta) = \mathbf{I}$ in Eq. 1). Here, $\dot{I}_{\text{eff}}$ is not interpreted as Shannon or semantic information, but rather as an operational geometric proxy for the rate of structural reconfiguration of the model. Under standard assumptions of smooth loss landscapes, Euclidean parameter displacement provides a local linear approximation to changes in the model's functional configuration, thereby serving as a measurable surrogate for internal informational turnover. From

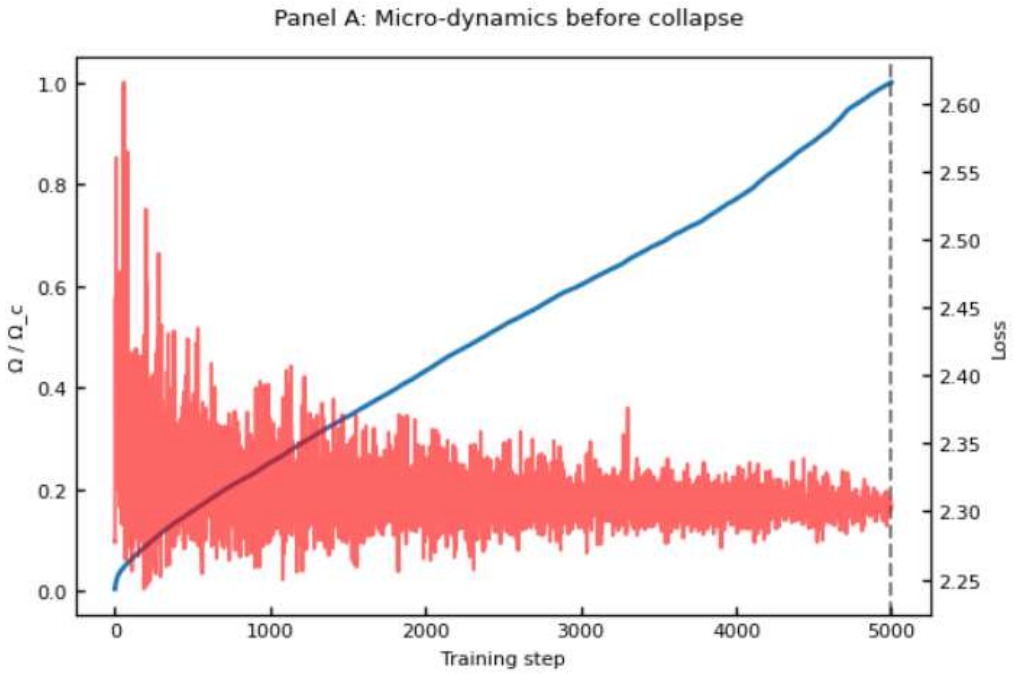


Figure 4: **Micro-dynamics of dissipation in CNNs.** Despite stable loss dynamics, cumulative geometric dissipation increases irreversibly during training. The dashed line indicates the termination of the simulation prior to functional collapse.

this perspective, collapse occurs not when optimization fails locally, but when the accessible geometric volume of $\mathcal{M}$ is exhausted.

Experiment III: Cross-architecture validation in convolutional networks

To test whether the geometric dissipation bound extends beyond specific architectural inductive biases, we applied the same analysis to convolutional neural networks with hierarchical inductive biases. Specifically, we trained a ResNet-18 architecture on CIFAR-10 under non-stationary conditions, using identical dissipation tracking protocols as in Experiments I and II.

Micro-dynamics of dissipation accumulation. Figure 4 shows the microscopic evolution of dissipation during CNN training. Although training loss remains stable over the simulated horizon, the cumulative dissipation $\Omega(t)$ increases monotonically, while the effective informational turnover rate $\dot{I}_{\text{eff}}$ fluctuates around a well-defined mean. This demonstrates that apparent functional stability does not imply thermodynamic sustainability: even in the absence of observable collapse, the dissipation budget is steadily depleted.

Thermodynamic scaling of functional lifespan. Because complete collapse may occur beyond feasible simulation horizons in large architectures, we reformulated lifespan in terms of

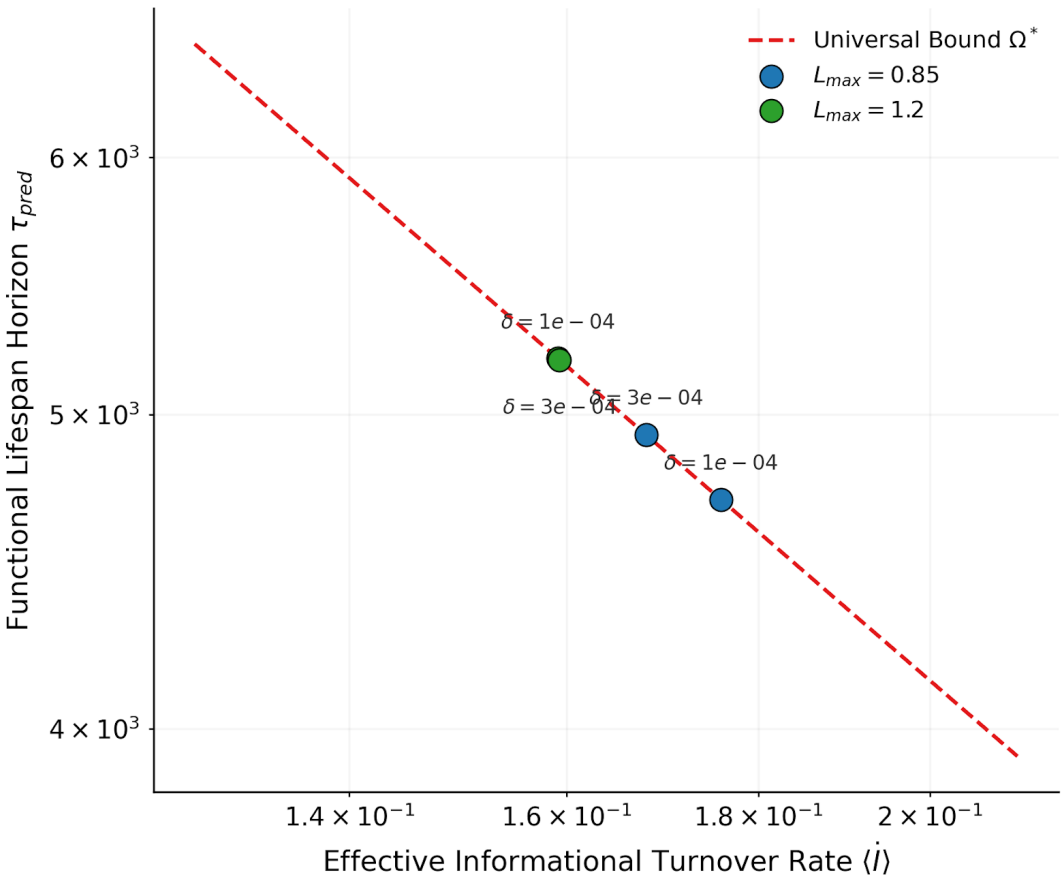


Figure 5: **Thermodynamic scaling of CNN lifespan.** Predicted functional lifespan horizon τ_{pred} versus effective informational turnover rate. The dashed line indicates the universal dissipation bound Ω^*. Distinct CNN configurations collapse onto the same thermodynamic scaling law.

a *functional lifespan horizon*

$$\tau_{\text{pred}} = \frac{\Omega^*}{\langle \dot{I} \rangle}, \tag{3}$$

Here, $\langle \dot{I} \rangle$ denotes the ensemble-averaged informational turnover rate measured over the stationary training regime, while Ω^* denotes the empirically measured dissipation budget at collapse, ensuring that the inferred bound represents a statistical rather than trajectory-specific invariant.

Figure 5 presents τ_{pred} as a function of the effective informational turnover rate on logarithmic axes.

All CNN configurations collapse onto the same inverse scaling relation $\tau_{\text{pred}} \sim \Omega^*/\langle \dot{I} \rangle$, coinciding with the dissipation horizon observed in multilayer perceptrons. This agreement holds across defensive thresholds and drift rates, indicating that Ω^* represents an architecture-independent physical bound.

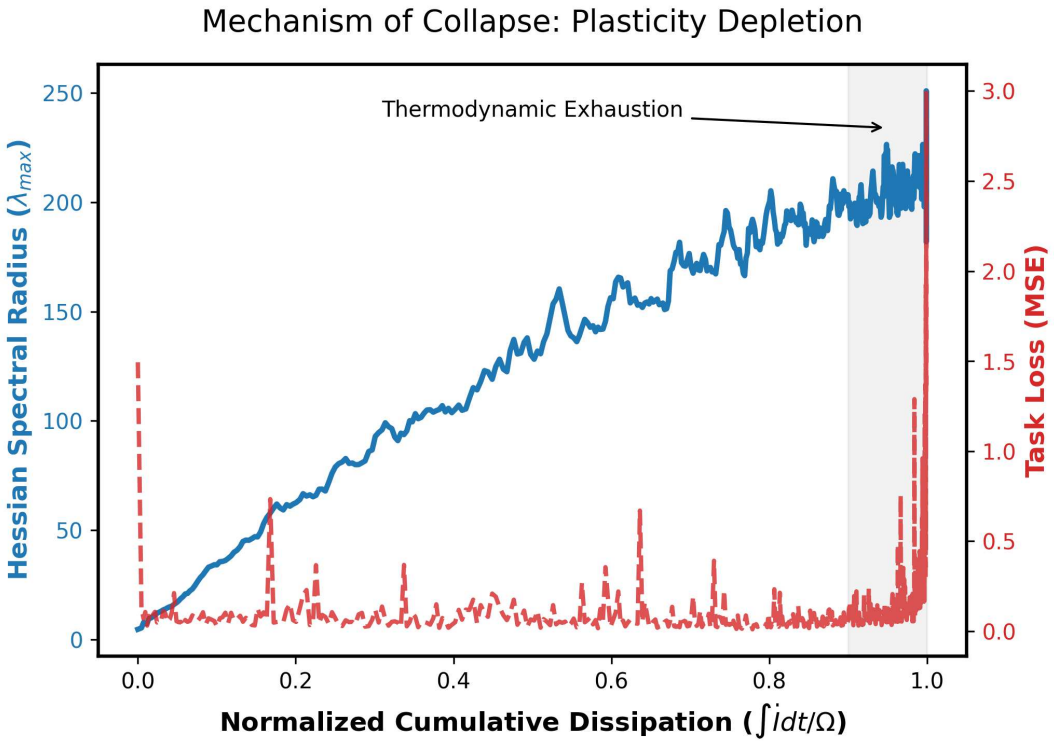


Figure 6: **Geometric rigidity preceding collapse.** Divergence of the Hessian spectral radius coincides with dissipation exhaustion, consistent with a geometric phase transition.

Geometric rigidity and collapse mechanism

To probe the geometric origin of collapse, we tracked the spectral radius of the Hessian throughout training. As Ω approaches exhaustion, curvature sharpens dramatically, signaling a loss of plasticity and heightened sensitivity to perturbations. The divergence of the Hessian spectral radius implies an extreme narrowing of the loss landscape along specific directions, effectively suppressing meaningful gradient updates and rendering the dynamics increasingly noise-dominated, consistent with a transition into a non-ergodic regime.

Robustness across optimization algorithms

Finally, we verified that the dissipation horizon persists across SGD, Adam, and RMSProp. Although these optimizers differ in noise structure and effective temperature, they consistently collapse at the same geometric dissipation budget.

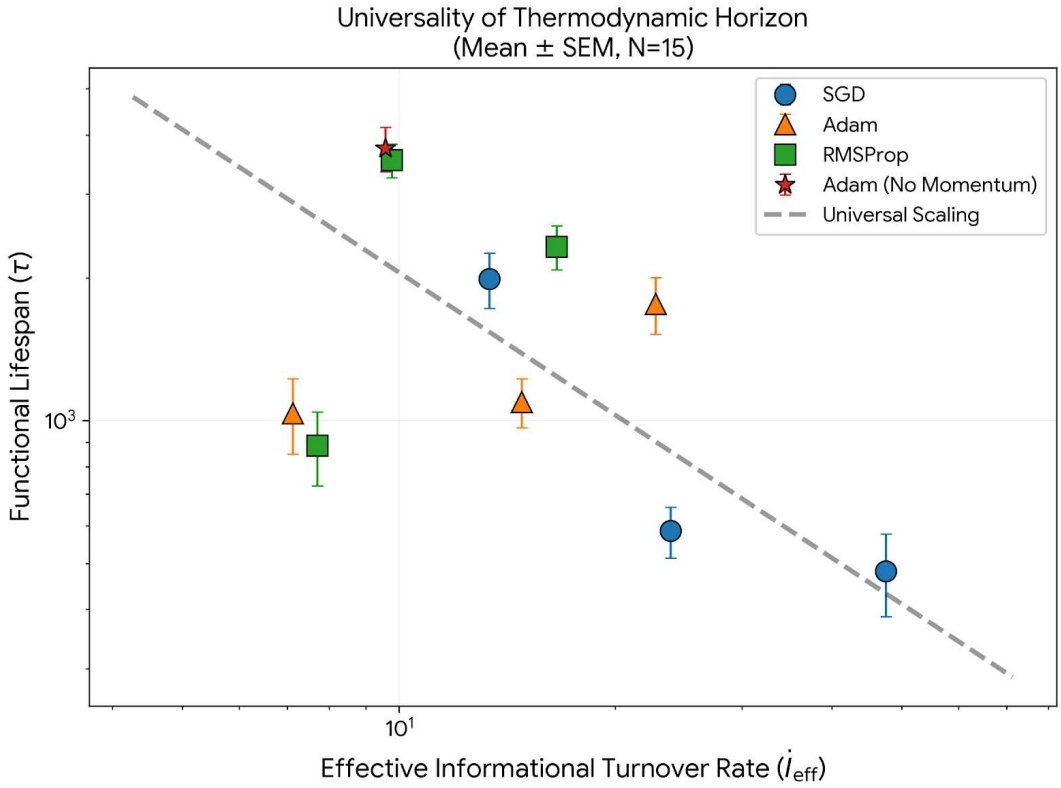


Figure 7: **Optimizer-invariant dissipation horizon.** Distinct optimization algorithms collapse at an identical Ω, confirming that the bound is geometric rather than algorithm-specific.

Discussion

Our results establish a geometric dissipation bound on the lifespan of adaptive learning systems. We show that sustained adaptation necessarily consumes a finite geometric resource, leading to an intrinsic horizon beyond which functional stability cannot be maintained. This limitation does not originate from algorithmic failure or architectural insufficiency, but from the irreversible exhaustion of adaptive degrees of freedom encoded in parameter space.

While defensive strategies and architectural design can redistribute dissipation over time, they cannot circumvent the intrinsic budget Ω. This bound provides a physical interpretation of sudden failure, non-monotonic fragility, and the limitations of continual learning. While our formulation utilizes a Euclidean metric as an operational proxy for computational cost, a natural theoretical extension involves intrinsic information-geometric metrics such as the Fisher Information Matrix. We posit that while the numerical value of the dissipation budget Ω is metric-dependent, the existence of a finite thermodynamic horizon is a topological property of the optimization landscape itself, persisting across reparameterizations. Future research will rigorously map the scaling laws of Ω under invariant Riemannian manifolds. More broadly, the

existence of a thermodynamic horizon suggests that perpetual adaptation is not physically realizable. Future work may explore how hardware-level dissipation maps onto geometric costs, and whether adaptive systems can be engineered to approach—yet never exceed—this fundamental limit. Crucially, biological adaptive systems appear to evade such dissipation-limited collapse, suggesting the existence of active mechanisms that counteract geometric exhaustion. We hypothesize that processes such as sleep cycles, synaptic homeostasis, and memory replay function as thermodynamic renormalization operations—periodically resetting the accumulated geometric complexity to restore system plasticity. In contrast, current artificial learning systems operate in a strictly monotonic dissipation regime, lacking the intrinsic entropy-exporting mechanisms required for true indefinite adaptation.

References

[1] Landauer, R. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* **5**, 183–191 (1961).

[2] Bennett, C. H. The thermodynamics of computation—a review. *International Journal of Theoretical Physics* **21**, 905–940 (1982).

[3] Mandt, S., Hoffman, M. D. & Blei, D. M. Stochastic gradient descent as approximate bayesian inference. *Journal of Machine Learning Research* **18**, 1–35 (2017).

[4] Smith, S. L. & Le, Q. V. Bayesian perspectives on generalization and stochastic gradient descent. *arXiv preprint arXiv:1710.06451* (2018).

[5] Prigogine, I. Time, structure, and fluctuations. *Science* **201**, 777–785 (1978).

[6] de Groot, S. R. & Mazur, P. *Non-Equilibrium Thermodynamics* (Dover, 1984).

[7] Seung, H. S., Sompolinsky, H. & Tishby, N. Statistical mechanics of learning. *Physical Review A* **45**, 6056–6091 (1992).

[8] Amari, S.-i. Natural gradient works efficiently in learning. *Neural Computation* **10**, 251–276 (1998).

[9] Chaikin, P. M. & Lubensky, T. C. Principles of condensed matter physics. *Cambridge University Press* (1995).

[10] Goldenfeld, N. Lectures on phase transitions and the renormalization group. *Addison-Wesley* (1992).

[11] Keskar, N. S. *et al.* On large-batch training for deep learning: Generalization gap and sharp minima. *arXiv preprint arXiv:1609.04836* (2017).

[12] Hochreiter, S. & Schmidhuber, J. Flat minima. *Neural Computation* **9**, 1–42 (1997).

[13] He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 770–778 (2016).

[14] Mehta, P. *et al.* A high-bias, low-variance introduction to machine learning for physicists. *Physics Reports* **810**, 1–124 (2019).

[15] Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nature Physics* **11**, 131–139 (2015).

[16] Still, S., Sivak, D. A., Bell, A. J. & Crooks, G. E. Thermodynamics of prediction. *Physical Review Letters* **109**, 120604 (2012).