mathematics

MDPI

Article

Depth Fragility and Skeletal Universality: Decoupling Topology and Function in Deep Neural Networks

Quang Nguyen [1,2,*], Hai Ha Pham [3], Davide Cassi [4,5] and Michele Bellingeri [4,5,*]

[1] Department of Physics, International University, VNU-HCM, Ho Chi Minh City, Vietnam
[2] Viet Nam National University, Ho Chi Minh City, Vietnam
[3] Department of Mathematics, International University, Ho Chi Minh City, Vietnam; phha@hcmiu.edu.vn
[4] Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, Italy; davide.cassi@unipr.it
[5] INFN, Gruppo Collegato di Parma, Italy
* Correspondence: nquang@hcmiu.edu.vn (Q.N.); michele.bellingeri@unipr.it (M.B.)

Abstract

Deep neural networks (DNNs) are traditionally analyzed as black-box function approximators, yet their internal structure exhibits phase transitions characteristic of complex physical systems. In this study, we investigate *topological–functional decoupling*—the phenomenon whereby a network retains full graph connectivity while losing computational function—in trained neural networks through the lens of percolation theory. By subjecting three distinct architectures (Shallow, Deep, and Wide MLPs) to a unified edge-pruning analysis on Fashion-MNIST, we uncover a fundamental divergence between structural integrity and computational capacity in this experimental setting. We report three key phenomena observed in these experiments: (1) the *zombie network* state under stochastic pruning, where the system retains global connectivity ($P_\infty \approx 1.0$) yet suffers a catastrophic functional collapse (accuracy falls below 50% of baseline at prunning ratio $p_f \approx 0.35$–0.68 depending on depth), proves that graph reachability does not imply computational capability; (2) *depth fragility*, where increased network depth triggers multiplicative signal decay (the avalanche effect), rendering deep architectures exponentially more vulnerable to random edge removal than shallow ones ($p_f^{\text{deep}} \approx 0.35$ vs. $p_f^{\text{shallow}} \approx 0.68$); and (3) *scale-free universality*, observed under magnitude-based pruning, where a robust functional skeleton maintains accuracy near the baseline (∼89%) up to extreme sparsity ($p_f \approx 0.85$–0.95) across all three architectures. Robustness stems not from holographic redundancy in the overall connection count but from the emergent *heavy-tailed rich-club* organization of weight magnitudes—a sparse set of high-magnitude synapses that form the functional backbone of the network, decoupled from the redundant topological mass. These findings offer new physical constraints for the design of resilient neuromorphic hardware.

Keywords: deep learning; percolation theory; network robustness; functional skeleton; depth fragility

MSC: 82B43

Academic Editor: Ioannis Tsoulos

Received: 18 March 2026
Revised: 18 April 2026
Accepted: 21 April 2026
Published: 24 April 2026

1. Introduction

The architecture of a deep neural network (DNN) [1] is, at its core, a complex directed graph where information flows through a sea of synaptic weights. From the perspective of statistical physics, these networks function as interacting systems where learning induces

a transition from high-entropy randomness to a low-entropy, structured state [2]. While the structural properties of complex networks have been extensively studied in various domains—ranging from financial markets [3–6] to social interactions [7–11]—the precise relationship between a DNN's physical topology and its functional capacity remains an open and profound question in the field of artificial intelligence safety and complexity.

Classical percolation theory, a cornerstone of statistical mechanics, predicts that the integrity of a network depends on the existence of a giant connected component (GCC) [12–16]. Extensive research on network robustness has demonstrated that system stability is highly sensitive to specific nodes or links removal strategies, whether in scale-free networks [13,17–19], weighted social graphs [10,20–22], or epidemic spreading models [23–25]. In these classical systems, if the fraction of removed nodes or edges exceeds a critical threshold p_c, the system undergoes a second-order phase transition, shattering into isolated clusters and terminating global functionality. In the context of AI, one might intuitively assume that the siphoning of classification accuracy would strictly follow the topological disintegration of the underlying DNN connectivity.

The Missing Topological Link: Previous empirical studies have long established that neural networks exhibit a remarkable degree of redundancy. Seminal works like *Optimal Brain Damage* [26], *Deep Learning network attack* [27,28], and *Deep Compression* [29] demonstrated that up to 50–80% of connections can be removed with minimal accuracy loss. However, these findings were largely interpreted through an engineering lens, focusing on compression rates and computational efficiency. The underlying *topological mechanism* that enables this resilience remains an open question in the physics of AI. The network does not collapse because its functional capacity is not distributed uniformly across all connections; rather, it is concentrated in a sparse subset of high-magnitude weights that form a robust functional backbone. Removing low-weight connections destroys the topological redundancy but leaves the functional core intact.

In this study, we bridge this gap by connecting these engineering observations to the statistical mechanics of scale-free organization in weighted networks. Before we formalize the concept, we introduce it intuitively: the **functional skeleton** is the minimal sparse subgraph, composed of the highest-magnitude synaptic weights, that is sufficient to reproduce the network's task performance. This skeleton is distinct from the broader topological structure of the network (the full set of connections that maintain graph connectivity). We show that the robustness observed in previous studies [29–31] is not a happy accident but a predicted property of the heavy-tailed weight distribution that emerges during training. Just as the World Wide Web is robust to random node failure because of its power-law degree distribution [13], we demonstrate that neural networks are robust to random edge pruning because their "functional connectivity"—defined by edge weight magnitude—follows a similar scale-free organization.

Specifically, we challenge classical intuition by demonstrating remarkable *topological–functional decoupling*: unlike financial or social networks where topological fragmentation often serves as a direct proxy for system failure [32], trained DNNs possess a functional skeleton—a hidden, sparse architecture composed of high-magnitude edge weights that governs the network's functioning. This backbone is fundamentally different from the *redundant connections* that maintain the GCC's size. Crucially, robustness arises not from topology itself (the mere existence of connections) but from the *weight organization* of those connections: the heavy-tailed rich-club structure of weight magnitudes. Topology and function are decoupled in the sense that topological connectivity can be preserved while function collapses (the zombie state), but functional robustness is determined by the subset of topological structures that carries the highest weights.

By analyzing the system under two distinct regimes of perturbation—*stochastic pruning* (random removal of synaptic edges, simulating random breakdown [13]) and *magnitude pruning* (targeted removal of the smallest weights to isolate the skeletal structure)—we reveal, for the first time in this experimental setting, three fundamental phenomena:

1. **The Zombie Network State:** Under random edge pruning, a network can remain almost entirely connected (P_∞, is ≈1.0) yet exhibit a complete functional collapse. This indicates that topological reachability does not imply computational capability, because the surviving random connections carry insufficient signal energy to propagate useful feature representations through the network.
2. **Depth-Induced Fragility:** We uncover a trade-off between expressivity and structural stability. While increasing network depth improves performance, it introduces a severe fragility to stochastic noise. Deep networks suffer a catastrophic "avalanche" collapse at $p_f \approx 0.35$, compared to the gradual decay observed in shallow architectures ($p_f \approx 0.68$).
3. **Universality of the Skeleton in this Setting:** Under magnitude pruning, functional robustness becomes nearly invariant to architecture. Networks of different configurations converge to a similar critical threshold ($p_f \approx 0.85$–0.95), suggesting that the functional skeleton reflects the task's complexity rather than the model's capacity, for the architectures and dataset studied here.

By examining these dynamics across three distinct architectural regimes on Fashion-MNIST, we bridge the gap between graph theory and machine learning. This study provides a rigorous physical interpretation of the lottery ticket hypothesis [33,34]. Ultimately, we suggest that the robustness of trained deep neural networks is not rooted in the sheer quantity of connections but in the emergent heavy-tailed organization of their strongest synapses.

Paper organization: Section 2 describes the dataset, network architectures, pruning protocols, and the order parameters used to quantify the phase transitions. Section 3 presents the empirical results for the three phenomena. Section 4 provides a physical interpretation and links to existing theory. Section 5 summarizes conclusions and directions for future work.

2. Methods

2.1. Notation

Table 1 summarizes the principal symbols used throughout this paper.

Table 1. Summary of principal notation.

Symbol	Definition
$G(V, E)$	Graph representation of a neural network with nodes V (neurons) and edges E (synaptic weights)
W	Weight matrix of a fully connected layer
M	Binary pruning mask; $W' = W \odot M$
p	Edge pruning ratio (fraction of weights removed)
P_∞	Normalized size of the giant connected component (GCC)
p_c	Topological threshold: p at which P_∞ drops below 0.9
p_f	Functional threshold: p at which test accuracy drops below 50% of baseline
$\alpha(p)$	Test accuracy as a function of pruning ratio p
L	Number of hidden layers (network depth)
x_l	Activation vector at layer l
$\mathbb{E}[\|x_L\|^2]$	Expected squared signal norm at the output layer

2.2. Dataset

We run our experiments on Fashion-MNIST [35] (available at https://github.com/zalandoresearch/fashion-mnist accessed on 18 January 2026), a dataset consisting of 60,000 training and 10,000 test grayscale images (28×28 pixels) spanning 10 clothing category classes (T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, and Ankle boot). We selected Fashion-MNIST over the simpler MNIST or the color-dependent CIFAR-10 because it offers an optimal balance: it requires the network to learn non-trivial structural features such as shapes and textures, while allowing diverse architectures to converge to a comparable high-accuracy baseline ($\approx$89%). This creates a controlled experimental environment where performance divergences can be attributed to topological variations rather than data inconsistencies.

2.3. Network Modeling and Architectural Variants

We modeled the neural networks as directed, weighted, layered feed-forward graphs $G(V, E)$, where V represents neurons (nodes) and E corresponds to synaptic weights (edges) among nodes in consecutive layers. An important modeling assumption is noted here: we map each synaptic weight to a binary edge (present or absent) for the purpose of computing topological measures such as the GCC. This mapping is a simplification, as in classical edge percolation, the edges are unweighted; however, our functional experiments use the actual weight magnitudes. The key distinction between the two pruning regimes is that stochastic pruning treats all edges as equivalent (as in classical random edge percolation), whereas magnitude pruning exploits the weight hierarchy. The role of nonlinear activations is discussed in Section 4.

We distinguish three structural regimes:

Shallow-MLP ($784 \to 512 \to 10$): minimal-depth architecture with one hidden layer of 512 nodes.
Deep-MLP ($784 \to 512 \to 256 \to 10$): increased depth leading to multiplicative signal attenuation and enhanced sensitivity to stochastic perturbations.
Wide-MLP ($784 \to 1024 \to 512 \to 10$): increased width at the same depth as the Deep-MLP, introducing topological redundancy that can temporarily buffer random damage.

2.4. Edge Pruning Protocol

Pruning was applied exclusively to the weight matrices W of the fully connected layers. Biases and batch normalization parameters were deliberately preserved to isolate the effect of synaptic connectivity loss on the network graph structure. This is a deliberate simplification: biases and normalization parameters could in principle partially compensate for edge removal, and their inclusion would constitute an interesting extension of this study (see Section 4). Mathematically, pruning is implemented by applying a binary mask M to the weights such that $W' = W \odot M$.

We investigated two distinct pruning regimes:

Stochastic pruning (random edge removal) simulates non-selective synaptic decay or hardware failure by removing connections with uniform probability. This corresponds to classical random edge percolation.
Magnitude pruning (weak edge removal) simulates targeted skeletal extraction by prioritizing the removal of connections with the smallest absolute weight values $|w_{ij}|$.

The pruning procedure, at a given ratio p, proceeds as follows: (i) compute $|w_{ij}|$ for all weights in the target layer; (ii) for stochastic pruning, sample a fraction p of edges uniformly at random without replacement; for magnitude pruning, select the fraction p of edges with the smallest $|w_{ij}|$; (iii) set the selected weights to zero (apply mask M); and (iv) evaluate the resulting network on the test set without any retraining. This post hoc evaluation ensures that the measured robustness reflects the structure learned during training, not adaptation to pruning.

2.5. Training Protocol

Prior to pruning, all networks were trained using the Adam optimizer with a learning rate of 10^{-3} and cross-entropy loss [36,37]. We replaced fixed-epoch training with an accuracy stopping mechanism: training was sustained until each architecture reached a stable accuracy plateau exceeding 89%. This high-performance baseline ensures that networks have fully converged to a low-entropy state characterized by a heavy-tailed weight distribution, which is the prerequisite for the decoupling phenomenon.

2.6. Order Parameters and Critical Thresholds

To quantify the phase transitions in the neural graph, we defined two critical percolation thresholds.

The **topological threshold** (p_c) is the edge pruning ratio p at which the normalized GCC size drops below 0.9, marking the onset of physical fragmentation in the graph.

The **functional threshold** (p_f) is the edge pruning ratio at which the test accuracy $\alpha(p)$ drops below 50% of the baseline accuracy. We selected a relative drop of 50%—rather than a total collapse to the random guessing baseline (≈10%)—to identify the onset of *functional collapse*: the point at which the network loses its generalized prediction capability and is no longer performing meaningfully above chance. This threshold marks the beginning of the collapse transition, analogous to identifying the onset of percolation rather than complete fragmentation. We acknowledge that alternative choices of threshold (e.g., 25% or 75% drop) would shift the absolute p_f values but would not qualitatively alter the ordering of architectures or the existence of the zombie phase; this sensitivity is a known limitation and should be considered when comparing across studies.

3. Results

3.1. Phase 1: Emergence of the Heavy-Tailed Skeleton

Post-training analysis reveals a fundamental structural convergence across diverse architectures. As shown in Figure 1, the synaptic weights of the "Shallow, Deep, and Wide" models converge toward a "heavy-tailed distribution", a characteristic hallmark of self-organizing complex systems [38].

Log–log plot analysis confirms a broad scaling behavior: while the vast majority of connections are suppressed toward zero-forming the redundant, non-functional connections, a sparse set of high-magnitude weights forms the heavy-tailed upper region. This emerging hierarchy is the prerequisite for the decoupling phenomenon, creating a high-contrast separation between the network's logical backbone and its redundant connections, independent of the network's depth or width. We note that the log–log plots shown here are indicative of heavy-tailed behavior; a rigorous statistical test of the power-law hypothesis (e.g., the Clauset–Shalizi–Newman method [39]) would be required to formally establish universality, and we flag this as a direction for future work.

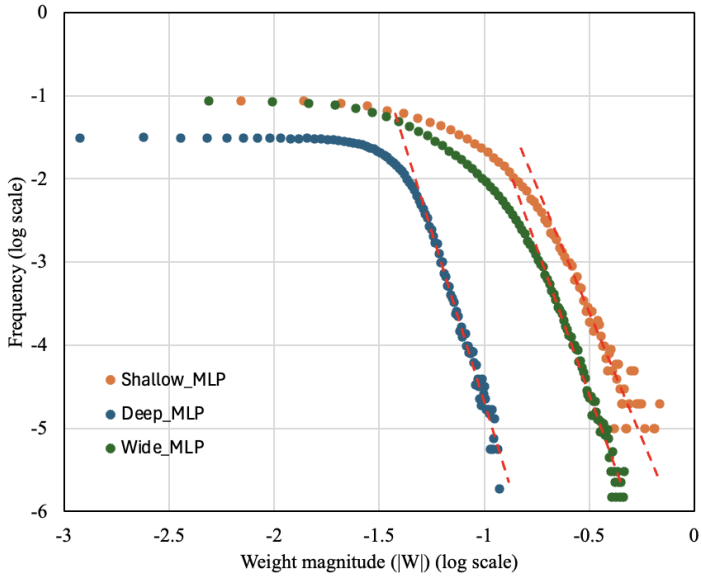

Figure 1. Heavy-tailed weight distributions. Log–log plots of the post-training weight magnitude frequencies for Shallow-MLP (orange), Deep-MLP (blue), and Wide-MLP (green) architectures on Fashion-MNIST. All three models exhibit a heavy-tailed distribution with a sparse set of high-magnitude weights (red dashed lines indicate indicative fits). This behavior is consistent with scale-free organization and motivates the functional skeleton hypothesis.

3.2. Phase 2: The Connectivity Plateau

Contrary to classical results regarding targeted attacks on scale-free networks, our analysis of the topological phase transition (Figure 2) reveals extreme structural robustness for both pruning strategies. The normalized GCC size remains stable at $P_\infty \approx 1.0$ for both random and magnitude pruning up to extreme sparsity levels ($p \approx 0.95$). This *connectivity plateau* is a direct consequence of the high density inherent to fully connected layers: in a graph where every node in layer l connects to every node in layer $l+1$, the redundancy is so high that breaking global paths requires the simultaneous removal of nearly all edges.

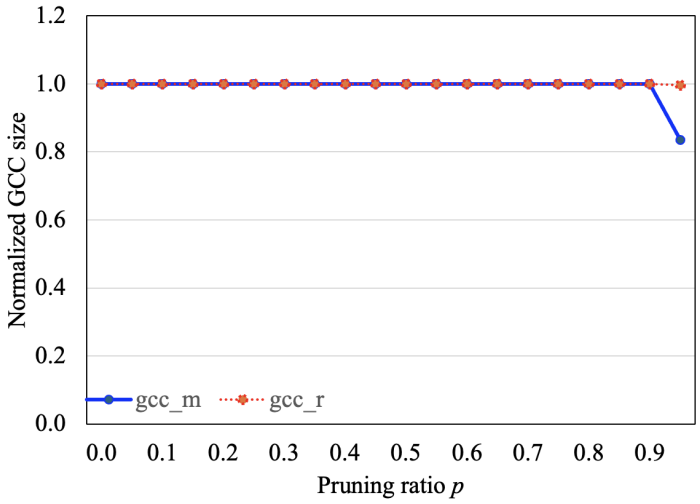

Figure 2. Evolution of the giant connected component (GCC) size under increasing edge pruning ratios. Stochastic (random) pruning (gcc_r) maintains a near-complete global connectivity (GCC ≈ 1.0) even at extreme sparsity levels, whereas magnitude-based pruning (gcc_m) triggers a topological phase transition near $p \approx 0.95$.

It is important to distinguish *graph connectivity* (the existence of at least one path from input to output nodes, measured by P_∞) from *signal propagation* (the ability of the network to carry meaningful feature information through those paths). The zombie phenomenon arises precisely because these two properties decouple: paths exist but carry degraded signals.

3.3. Phase 3: The Zombie Network and Functional Decoupling

The core discovery emerges from comparing topological integrity (Figure 2) with functional performance (Figure 3). Under stochastic pruning, we observe the **zombie network** state: in the range $0.2 < p < 0.8$, the network possesses a near-perfect physical structure ($P_\infty \approx 1.0$), yet its functional capacity decreases progressively and collapses when p approaches 0.8. The signal fails to propagate not because paths are absent but because random pruning indiscriminately destroys the high-magnitude weights that carry the bulk of the feature information while leaving low-weight connections intact. The surviving connections form a topologically connected but informationally incoherent substrate; in formal terms, the expected signal norm at the output layer decays as $\mathbb{E}[\|x_L\|^2] \approx (1-p)^L \cdot \mathbb{E}[\|x_0\|^2]$ (see Section 4), driving the signal-to-noise ratio to zero before the GCC begins to fragment.

In contrast, magnitude pruning demonstrates that the functional skeleton is a strict subset of the full edge set. While the GCC remains intact, magnitude pruning selectively preserves the high-information pathways, maintaining accuracy near the baseline (∼89%) throughout the entire connectivity plateau. Functional collapse only occurs when the pruning ratio exceeds $p \approx 0.8$.

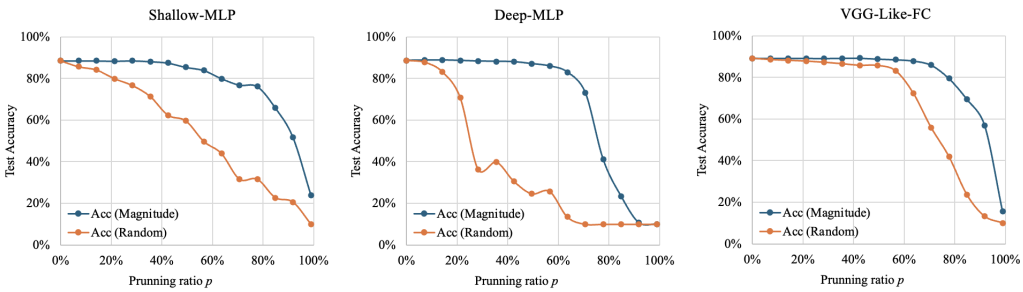

Figure 3. Depth fragility vs. skeletal universality on Fashion-MNIST. Functional robustness (test accuracy) under magnitude pruning (blue) and random pruning (orange) across three architectures. **Left—Shallow-MLP:** gradual, approximately linear decay under random noise. **Center—Deep-MLP:** catastrophic avalanche effect under random pruning; accuracy collapses at $p_f \approx 0.35$ due to multiplicative signal decay. **Right—Wide-MLP:** increased width delays random collapse to $p_f \approx 0.60$ before the same depth-induced fragility sets in. Under magnitude pruning (blue), all three networks maintain near-baseline accuracy up to $p > 0.9$, revealing the architectural invariance of the skeleton.

3.4. Impact of Structural Complexity: The Depth Fragility Paradox

Table 2 compares the critical thresholds across architectures. Under stochastic pruning, the functional threshold p_f shifts dramatically with network depth. The Shallow-MLP exhibits a gradual decay ($p_f \approx 0.68$) because damage in a single-layer transformation is additive. The Deep-MLP suffers a catastrophic exponential collapse ($p_f \approx 0.35$): noise propagates multiplicatively through depth, destroying the inference signal (the avalanche effect). The Wide-MLP delays this collapse ($p_f \approx 0.60$) by utilizing width as a redundant buffer, but eventually succumbs to the same nonlinear breakdown.

Under magnitude pruning, this fragility disappears. All three architectures converge to a similar threshold ($p_f \approx 0.85$–0.95), indicating that the functional skeleton is invariant to the architectural container within this experimental setting.

Table 2. Critical thresholds on Fashion-MNIST. Depth fragility is quantified by the ratio $p_f^{\text{shallow}}/p_f^{\text{deep}} \approx 1.94$ under random pruning, showing that the Deep-MLP collapses nearly twice as fast. Magnitude pruning restores robustness across all architectures.

Architecture	Config	Initial Acc.	Strategy	p_f (Func.)	p_c (Topo.)
Shallow-MLP	Narrow/Shallow	∼89.0%	Magnitude	0.95	0.94
			Random	0.68	0.98
Deep-MLP	Narrow/Deep	∼89.0%	Magnitude	0.85	0.94
			Random	0.35	0.98
Wide-MLP	Wide/Deep	∼89.0%	Magnitude	0.92	0.95
			Random	0.60	0.97

4. Discussion

4.1. Scale-Free Universality and the Functional Skeleton

The empirical consistency observed across the Shallow-MLP, Deep-MLP, and Wide-VGG architectures reveals that the "functional skeleton" is not an artifact of specific configurations but a fundamental topological property of trained neural representations. The statistical foundation for this robustness is the heavy-tailed weight distribution ubiquitously formed during training across all studied architectures (Figure 1). This mechanism explains the observed results ($p_f \approx 0.85$–0.95 for all models). Regardless of whether the initial container is a compact shallow network or a massive Wide-VGG model, magnitude pruning collapses them all onto a similarly sized functional backbone. This suggests that for a fixed task like Fashion-MNIST, there exists an intrinsic minimal description length-required to encode the decision boundaries. Furthermore, this resilience aligns with the geometry of the loss landscape; as noted by Li et al. [40], magnitude pruning effectively navigates the network along the principal axes of wide local minima, ensuring the functional output remains invariant to perturbations in non-critical directions.

We emphasize that the term "universality" is used here in the informal sense: we observe that the same qualitative behavior (heavy-tailed weights, skeletal convergence) emerges across three different architectures on one dataset. A formal claim of universality in the statistical physics sense—requiring critical exponents, finite-size scaling, and formal derivation of thresholds [16]—would require experiments across many more architectures, datasets, and tasks, and this is left for future work.

4.2. The Zombie Phase and Depth-Induced Fragility

The transition to a zombie network is governed by the physics of signal propagation. We model the expected squared activation norm at the output layer under random pruning of fraction p as:

$$\mathbb{E}[\|x_L\|^2] \approx (1-p)^L \cdot \mathbb{E}[\|x_0\|^2], \tag{1}$$

where L is the number of layers. This simplified model assumes independent uniform pruning across layers and ignores nonlinear activations, which redistribute the signal in a depth-dependent manner in practice. It should therefore be interpreted as a qualitative physical picture rather than a quantitative prediction. Nevertheless, it captures the essential mechanism: for deep networks ($L \geq 3$), the $(1-p)^L$ factor decays exponentially with depth, creating an avalanche effect where feature information attenuates into background noise long before reaching the classifier, even when the GCC remains intact.

Regarding nonlinear activations: ReLU activations introduce an additional multiplicative factor related to the fraction of active neurons, which would modify the decay rate. A more accurate signal propagation analysis incorporating activations is an important direction for future work. Similarly, biases and normalization parameters were excluded

from pruning in this study to isolate the effect of synaptic connectivity. Whether these components can partially compensate for edge removal—and whether including them in the pruning protocol changes the observed thresholds—is an open question that future experiments should address.

4.3. Implications for Hardware Design and Limitations

The finding that functional robustness is determined by a small set of high-magnitude synapses—rather than the global connection count—suggests a design principle for neuromorphic hardware: protection efforts should be concentrated on preserving high-magnitude skeletal pathways rather than maintaining indiscriminate global redundancy. We emphasize, however, that this is currently a hypothesis derived from our experimental observations; it has not been directly validated in hardware experiments, and such validation would be an important step before engineering recommendations can be made.

Several limitations of this study should be noted explicitly:

- Experiments were conducted on a single dataset (Fashion-MNIST) and three MLP architectures. Modern architectures (e.g., Transformers, ResNets, convolutional networks with attention) may exhibit different behavior.
- The signal decay model ignored nonlinear activations.
- Biases and normalization parameters were excluded from pruning.
- Formal statistical tests of the power-law hypothesis for weight distributions are not provided.
- Critical exponents and finite-size scaling analysis, standard in statistical physics, were not derived.

Addressing these limitations in future work would substantially strengthen the generality of the conclusions.

5. Conclusions

In this study, we established a quantitative framework linking the statistical mechanics of graph pruning to the learning dynamics of deep neural networks. By subjecting Shallow, Deep, and Wide MLP configurations to a unified percolation analysis on Fashion-MNIST, our investigation yields three primary conclusions.

First, we demonstrated the existence of a zombie network phase.For dense, overparameterized networks, topological percolation is a trivial condition; the giant connected component remains robust ($P_\infty \approx 1.0$) even at extreme sparsity ($p > 0.95$). However, we showed for the first time in this context that *reachability does not imply computability*: under random pruning, networks retain global connectivity yet fail to process information. This proves that the network's knowledge is concentrated within a low-entropy functional skeleton defined by the weight magnitude hierarchy, not distributed holographically across all connections.

Second, we uncovered and quantified depth fragility. While increasing network depth enhances computational capacity, it introduces severe vulnerability to stochastic noise. Deep architectures suffer a catastrophic avalanche effect ($p_f \approx 0.35$) compared to the gradual decay of shallow networks ($p_f \approx 0.68$)—a factor of $\sim 2\times$ difference. This confirms that error propagation is multiplicative in deep systems and that topological redundancy (width) offers only a temporary buffer.

Third, we characterized the invariance of the functional skeleton within this experimental setting. Magnitude pruning reveals that regardless of architecture, the functional skeleton required to solve Fashion-MNIST converges to the same critical threshold ($p_f \approx 0.85$–0.95), suggesting that magnitude pruning acts as a topological renormalization operator collapsing complex architectures to the intrinsic minimal description length of the task.

These findings suggest that future resilient AI hardware should prioritize the protection of high-magnitude skeletal pathways. By viewing neural networks through the lens of phase transitions, we conclude that intelligence survives not through the quantity of connections but through the precise hierarchical organization of the strong few.

Author Contributions: Conceptualization, Q.N.; methodology, D.C.; software, H.H.P.; validation, Q.N.; writing—original draft, M.B.; writing—review & editing, Q.N., H.H.P., D.C. and M.B.; visualization, M.B.; supervision, Q.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this article are openly available [35].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
2. Zdeborová, L. Understanding deep learning is also a job for physicists. *Nat. Phys.* **2020**, *16*, 602–604. [CrossRef]
3. Tumminello, M.; Aste, T.; Di Matteo, T.; Mantegna, R.N. A tool for filtering information in complex systems. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10421–10426. [CrossRef] [PubMed]
4. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **1999**, *11*, 193–197. [CrossRef]
5. Bonanno, G.; Caldarelli, G.; Lillo, F.; Mantegna, R.N. Topology of correlation-based minimal spanning trees in real and model markets. *Phys. Rev. E* **2003**, *68*, 046130. [CrossRef]
6. Garas, A.; Argyrakis, P.; Havlin, S. The structural role of weak and strong links in a financial market network. *Eur. Phys. J. B* **2008**, *63*, 265–271. [CrossRef]
7. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [CrossRef]
8. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [CrossRef]
9. Newman, M.E. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [CrossRef]
10. Onnela, J.P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Argollo de Menezes, M.; Kaski, K.; Kertész, J. Analysis of a large-scale weighted network of one-to-one human communication. *New J. Phys.* **2007**, *9*, 179. [CrossRef]
11. Bellingeri, M.; Bevacqua, D.; Scotognella, F.; Alfieri, R.; Nguyen, Q.; Montepietra, D.; Cassi, D. Link and node removal in real social networks: A review. *Front. Phys.* **2020**, *8*, 228. [CrossRef]
12. Newman, M. *Networks*; Oxford University Press: Oxford, UK, 2018.
13. Albert, R.; Jeong, H.; Barabási, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [CrossRef]
14. Callaway, D.S.; Newman, M.E.; Strogatz, S.H.; Watts, D.J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **2000**, *85*, 5468. [CrossRef] [PubMed]
15. Kitsak, M.; Ganin, A.A.; Alderson, D.L.; Fowler, M.S.; Linkov, I. Stability of a giant connected component. *Phys. Rev. E* **2018**, *97*, 012309. [CrossRef] [PubMed]
16. Stauffer, D.; Aharony, A. *Introduction to Percolation Theory*; CRC Press: Boca Raton, FL, USA, 2018.
17. Cohen, R.; Erez, K.; Ben-Avraham, D.; Havlin, S. Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.* **2000**, *85*, 4626. [CrossRef] [PubMed]
18. Pajevic, S.; Plenz, D. The organization of strong links in complex networks. *Nat. Phys.* **2012**, *8*, 429–436. [CrossRef]
19. Nguyen, N.K.K.; Nguyen, Q.; Pham, H.; Tr, N.T.; Alfieri, R.; Cassi, D.; Bellingeri, M. Analytics solution for the robustness of incomplete information scale-free networks. *Phys. Lett. A* **2025**, *565*, 131123. [CrossRef]
20. Barrat, A.; Barthelemy, M.; Pastor-Satorras, R.; Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3747–3752. [CrossRef]
21. Onnela, J.P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; Barabási, A.L. Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7332–7336. [CrossRef]
22. Bellingeri, M.; Bevacqua, D.; Sartori, F.; Turchetto, M.; Scotognella, F.; Alfieri, R.; Nguyen, N.; Le, T.; Nguyen, Q.; Cassi, D. Considering weights in real social networks: A review. *Front. Phys.* **2023**, *11*, 1152243. [CrossRef]
23. Pastor-Satorras, R.; Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **2001**, *86*, 3200. [CrossRef]
24. Newman, M.E. Spread of epidemic disease on networks. *Phys. Rev. E* **2002**, *66*, 016128. [CrossRef] [PubMed]

25. Chen, Y.; Paul, G.; Havlin, S.; Liljeros, F.; Stanley, H. Finding a Better Immunization Strategy. *Phys. Rev. Lett.* **2008**, *101*, 058701. [CrossRef] [PubMed]
26. LeCun, Y.; Denker, J.; Solla, S. Optimal brain damage. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1989; Volume 2.
27. Akhtar, N.; Mian, A. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. *IEEE Access* **2018**, *6*, 14410–14430. [CrossRef]
28. Yuan, X.; He, P.; Zhu, Q.; Li, X. Adversarial Examples: Attacks and Defenses for Deep Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 2805–2824. [CrossRef]
29. Han, S.; Pool, J.; Tran, J.; Dally, W. Learning both weights and connections for efficient neural network. In *Advances in Neural Information Processing Systems (NeurIPS)*; MIT Press: Cambridge, MA, USA, 2015; Volume 28.
30. Lazarevich, I.; Kozlov, A.; Malinin, N. Post-training deep neural network pruning via layer-wise calibration. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021.
31. Frantar, E.; Alistarh, D. SparseGPT: Massive Language Models Can Be Accurately Pruned in One-Shot. In Proceedings of the 40th International Conference on Machine Learning (ICML), PMLR 2023, Honolulu, HI, USA, 23–29 July 2023; Volume 202, pp. 10323–10337.
32. Nguyen, N.K.K.; Nguyen, Q.; Pham, H.H.; Le, T.T.; Nguyen, T.M.; Cassi, D.; Scotognella, F.; Alfieri, R.; Bellingeri, M. Predicting the Robustness of Large Real-World Social Networks Using a Machine Learning Model. *Complexity* **2022**, *2022*, 3616163. [CrossRef]
33. Frankle, J.; Carbin, M. The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. In Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.
34. Tanaka, H.; Kunin, D.; Yamins, D.L.; Ganguli, S. Pruning neural networks without any data by iteratively conserving synaptic flow. In *Advances in Neural Information Processing Systems (NeurIPS)*; MIT Press: Cambridge, MA, USA, 2020.
35. Xiao, H.; Rasul, K.; Vollgraf, R. Fashion-MNIST: A Novel Image Dataset for Benchmarking Machine Learning Algorithms. *arXiv* **2017**, arXiv:1708.07747. [CrossRef]
36. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
37. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
38. Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [CrossRef]
39. Clauset, A.; Shalizi, C.R.; Newman, M.E. Power-law distributions in empirical data. *SIAM Rev.* **2009**, *51*, 661–703. [CrossRef]
40. Li, H.; Xu, Z.; Taylor, G.; Studer, C.; Goldstein, T. Visualizing the loss landscape of neural nets. In *Advances in Neural Information Processing Systems (NeurIPS)*; MIT Press: Cambridge, MA, USA, 2018.