Approximating Latent Manifolds in Neural Networks via Vanishing Ideals

Nico Pelleriti, Max Zimmer, Elias Wirth, Sebastian Pokutta

Zuse Institute Berlin (ZIB) Technische Universität Berlin (TUB)

Berlin Mathematics Research Center MATH+

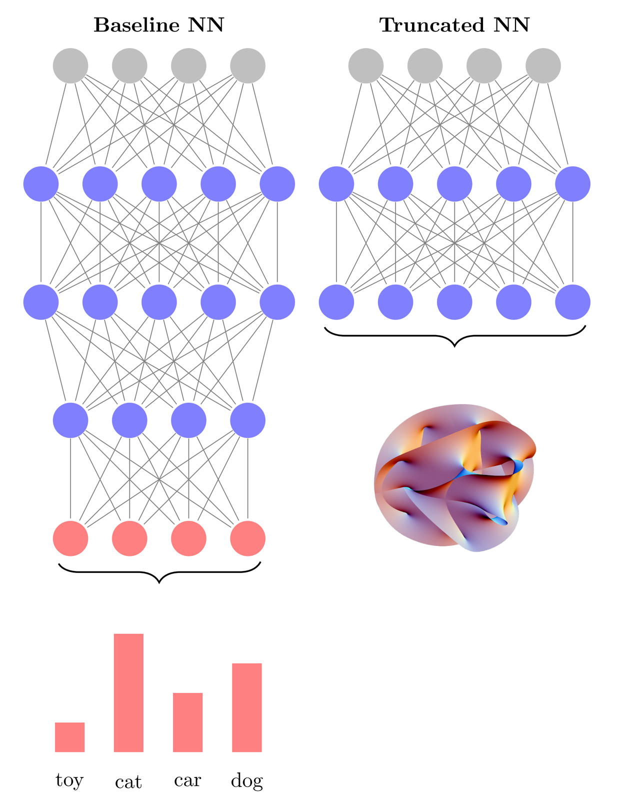
ZIB ZUSE INSTITUTE BERLIN

Motivation

- Deep neural networks (NNs) learn powerful latent representations.
- Manifold hypothesis: data lies on low-dimensional manifolds.
- Algebraic geometry: manifolds can be described by vanishing polynomials.
- Can we use vanishing ideals to characterize NN latent spaces?
- **Goal:** Efficiently approximate class manifolds in NN latent spaces using vanishing ideal generators.
- **Approach:** Replace final NN layers with a polynomial layer built from vanishing ideal generators.
- **Benefits:** Fewer parameters, competitive accuracy, improved throughput, theoretical guarantees.

Latent Manifolds

- Deep networks' final layers output class probabilities.
- Intermediate layers reveal **latent manifolds** in feature space.
- These manifolds capture the underlying structure of each class.
- Understanding this motivates our algebraic (vanishing ideal) approach.



Visualization of a learned latent manifold in neural network feature space.

Method Overview

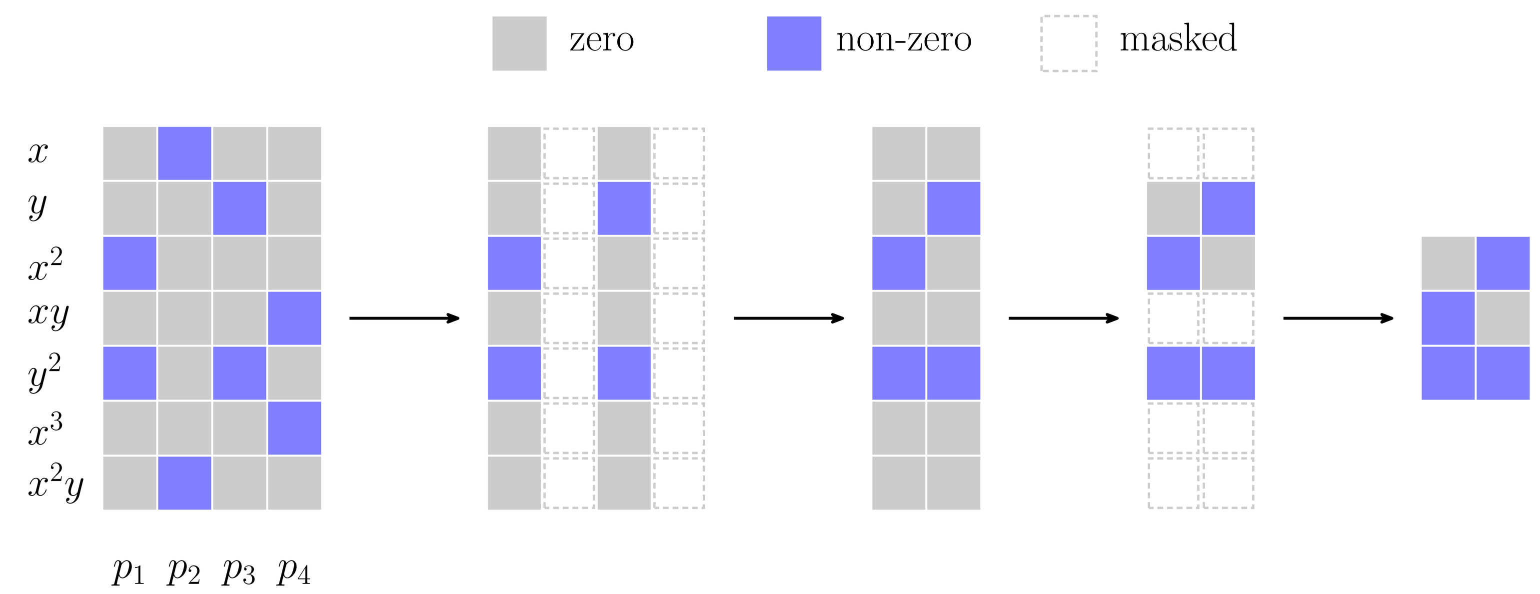
Pipeline:

1. Train or use a pretrained NN; truncate at intermediate layer.
2. For each class, extract latent activations $\mathbf{Z}^k$.
3. Compute generators of the vanishing ideal for $\mathbf{Z}^k$.
4. Prune/select most discriminative generators.
5. Build a polynomial layer: $\mathbf{z} \mapsto (|p_1^k(\mathbf{z})|, \ldots, |p_{n_k}^k(\mathbf{z})|)$.
6. Add linear classifier on top.

Key Techniques:

- Dimensionality reduction (PCA)
- Stochastic vanishing ideal algorithms (OAVI, ABM)
- Pruning for discriminative power and efficiency

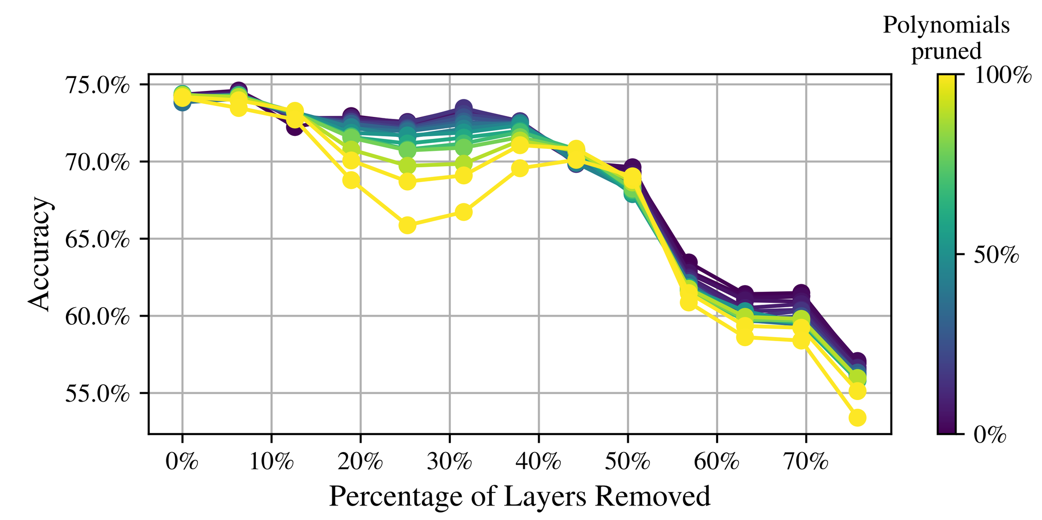
Pruning Approach



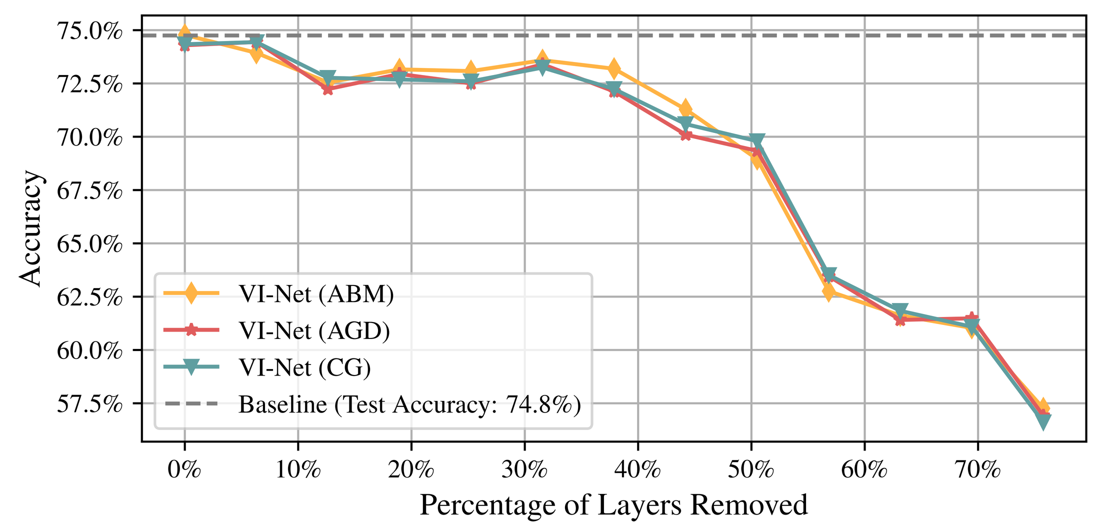
Schematic: Prune sparse generators to reduce model size and improve efficiency.

Experiments: CIFAR-10/100

- **Competitive accuracy:** Matches or exceeds baseline with fewer parameters.
- **Parameter efficiency:** Up to 75% reduction in parameters.
- **Throughput:** Up to 1.4× higher images/sec due to smaller model.



3D accuracy vs. pruning and truncation (CIFAR-100).



Accuracy vs. truncation for different methods (CIFAR-100).

Model	Total	Acc (%)	Thruput
VI-Net34-Tiny	3.52M	71.9	42.1k
VI-Net34-Small	5.88M	74.0	37.6k
VI-Net34-Med	14.6M	74.7	29.3k
VI-Net34-Large	19.3M	74.8	27.9k
ResNet34	20.4M	74.8	28.3k

Accuracy, total parameters, and throughput (img/s) for main models.

% Layers	Monomials	Polys	Time (s)
76	7,625	188,743	3,369
50	5,196	53,675	82
44	1,990	30,004	18
25	196	12,978	7
7	237	12,997	9

Monomials, polynomials, and training time vs. layers removed.

Algebraic Foundations

- **Vanishing ideal:** $\mathscr{I}(\mathbf{Z}) = \{p \in \mathscr{P} : p(\mathbf{z}_i) = 0 \; \forall \mathbf{z}_i \in \mathbf{Z}\}$
- **Algebraic variety:** Set of common zeros of a finite set of polynomials.
- **Tognoli Theorem:** Every compact smooth manifold is diffeomorphic to a real algebraic variety.
- **Approximate vanishing:** Allow small error ψ for noisy data.

Theory and Guarantees

- **Spectral complexity:** Measures the capacity of a feedforward neural network in terms of the product and structure of its weight matrices.
- **Definition:** $R_\phi = \left(\prod_{i=1}^L \rho_i \|\mathbf{W}_i\|_2\right) \left(\sum_{i=1}^L \frac{\|\mathbf{W}_i^T\|_{2,1}^{2/3}}{\|\mathbf{W}_i\|_2^{2/3}}\right)^{3/2}$
- **Main idea:** Encode the polynomial layer using a fixed custom activation function that allows computation of a fixed number of arbitrary monomials up to a certain degree.
- **Main theorem (simplified):** By choosing the degree, number, and sparsity of vanishing ideal generators, we can ensure the spectral complexity of our model is strictly lower than the original network.

Conclusion and Outlook

- Vanishing ideals provide a principled way to characterize NN latent manifolds.
- Polynomial layers can replace deep NN tails with little loss in accuracy.
- **Future:** Apply to larger models, more complex data, and explore interpretability.
- **Future:** Explore sampling from the algebraic varieties characterized by vanishing ideals.