Hi all,

Here is your disruptive ML/AI digest for the day.

Here's what you folks might find interesting today.

TOP PICKS

THEORY AND GENERALIZATION

A Function-Space Dichotomy for Compositional Learning: Exponential Sub-Optimality of the Neural Tangent Kernel

This paper gives a sharp theoretical account of a long-standing empirical fact: kernelized lazy-training limits can fail badly on compositional targets even when finite networks learn them well. The key novelty is the explicit separation between **Fourier complexity**, which governs NTK regression, and **architectural complexity**, which governs norm-controlled finite ReLU networks. That lets the authors prove an exponential sample-complexity gap on iterated sawtooth functions, where NTK needs Omega(4^L) samples while the minimax floor for the architecture class stays polynomial in depth. I think this matters because it moves the "feature learning beats kernels" story from folklore to a concrete function-space dichotomy with rates, not just examples. The sparse-parity experiments also strengthen the claim that the failure mode is tied to compositional structure rather than generic kernel weakness. If you care about implicit bias, representation learning, or when infinite-width theory is actually misleading, this is one of the strongest papers in the batch.

Width-Robust Learnability in Mean-Field Bayesian Neural Networks

This is a conceptually strong result on whether infinite-width mean-field Bayesian networks preserve the same learnability class as polynomial-width networks. The central object, **reduced entropy**, acts like an intensive prior cost for representing a target, and the main theorem says polynomial-sample learnability at infinite width is equivalent to polynomial-sample learnability at polynomial width. That is a nontrivial statement because infinite-width limits often risk introducing artifacts that change the effective inductive bias. The proof idea is also interesting: a subsampling argument that decomposes the mean-field solution into an active data-dependent part and a lazy entropy-dominated part. I would flag this as especially relevant if you are tracking the boundary between feature-learning and lazy regimes, or looking for complexity measures that survive asymptotic limits. It is more foundational than flashy, but the theorem directly addresses a real ambiguity in how people interpret mean-field limits.

Deep Neural Variation Spaces: A Unifying Perspective on Depth and Complexity

This paper proposes a unified function-space formalism for deep norm-constrained networks that is not tied just to homogeneous activations like ReLU. The most provocative claim is the **depth saturation** result in the univariate ReLU setting: once complexity is controlled in the right function-space norm, extra depth only rescales the class by a small constant and does not add real functional diversity. That directly challenges a lot of standard expressivity narratives that rely on parameter counting or unconstrained rescaling across layers. The representer theorem and the directional high-frequency limitation also make the framework more than a rebranding of existing norm-based bounds. I think this is relevant because it tries to separate genuine depth benefits from artifacts of how capacity is measured. If the results hold up, they could be useful for rethinking what kinds of inductive bias depth actually buys under principled complexity control.

QUANTUM ML AND HYBRID PARADIGMS

Provable learning separation for predicting time-evolution of quantum many-body systems

This is the clearest "provable quantum advantage for a learning task" paper in the set, and it does so on a physically motivated supervised problem rather than an artificial oracle setup. The task is to learn quantum many-body dynamics from probe states, times, and observable expectations; the authors then give an efficient quantum learner that identifies the Hamiltonian from short-time data and performs inference using simulation plus classical shadows. The real contribution is the hardness side: they embed BQP-complete computation into long-time dynamics of a low-intersection Feynman-Kitaev clock Hamiltonian and show classical polynomial-time learners would imply a major complexity collapse. That makes the separation much more compelling than typical QML advantage claims, which often lack either a natural task or a robust hardness argument. I would pay attention to this if you care about PAC learning, simulation-based quantum ML, or rigorous boundaries between classical and quantum learnability. It is one of the few papers here that looks genuinely disruptive rather than merely technically solid.

Entanglement as a Structural Complexity Axis: A PAC-Bayesian View of Generalization in Quantum Policies and Value Functions

This paper is notable because it treats **entanglement** not just as expressivity, but as a distinct source of generalization complexity through the Fisher geometry of parameterized quantum circuits. The PAC-Bayesian framing says the relevant complexity is the Fisher effective dimension, and the empirical claim is that increasing entangling connectivity worsens train-test gaps even when parameter count is fixed. That is a useful conceptual shift because parameter counting is especially uninformative in quantum circuit design, and the paper argues for a ranking certificate that can distinguish circuits with equal parameter budgets. I also like that the authors test the mechanism across supervised learning, contextual bandits, value-function approximation, and even on IBM Heron hardware, while being explicit that the multi-step RL evidence is noisier. For researchers interested in theory-grounded quantum RL, this is more valuable than another expressivity paper because it proposes a concrete complexity axis tied to generalization. The main limitation is that the strongest evidence comes from relatively low-variance settings, so the broader RL implications still need more resolution.

BY CATEGORY

NOVEL ARCHITECTURES AND OBJECTIVES

EquiFiLM: Charge-Conditioned Equivariant Force Fields via Feature-wise Linear Modulation

Black Hole Black Boxes: Numerical Black Hole Metrics via AInstein Neural Networks

Orthogonal Dendritic Intrinsic Networks: An Architecture for Significance-Ordered, Orthogonal Latent Spaces

Level-Crossing Density as a Mesh-Free High-Frequency Auxiliary Loss for Implicit Neural Representations

THEORY AND GENERALIZATION

Quantitative Gaussian-Process limits of Tensor Programs

To Retain or to Adapt? Generalizing Continual Learning

No Subspace to Track: Non-Identifiability and Optimizer State in Low-Rank Training

Boosting with List-Decodable Codes

Higher-Order Certified Robustness for Regression

Feature Learning for the High Dimensional Stationary Schodringer Equation with Deep Ritz Method

Separation Capacity of Scattering Networks on Low-Dimensional Datasets

PROBABILISTIC AND STATISTICAL METHODS

EntroPath: Maximum Entropy Path Ensemble Embedding for Manifold Learning

On the convergence of graph Laplacians with a symmetric divergence

A Convex Approximation Framework for Neural Likelihood-Based Bayesian Inverse Problems

Differentially Private Natural Gradient Descent

Learning Sparsest Linear Causal DAGs with Latent Confounders via Higher-Order Cumulants

A unified perspective of Gaussian process approximation for differential equations

QUANTUM ML AND HYBRID PARADIGMS

No remaining papers.

EMERGENT BEHAVIOR AND AGENTIC AI

Beyond Heuristic Tuning: Power-Calibrated LLM Watermarking

Back tomorrow with more!

P.S. In the most recent submission window: cs.LG: 97 papers stat.ML: 20 papers