Hey all,

Here is your disruptive ML/AI digest for the day. Here's what you folks might find interesting today.

Probabilistic and Statistical Methods

Denoising growth complexity: Data geometry and certified schedules for diffusion sampling (Martin J. Wainwright)

This paper introduces the *denoising growth complexity* (DGC), a geometric measure of how the denoising MSE evolves along the Gaussian heat flow, and uses it to derive explicit KL error bounds for Euler discretization of diffusion samplers. The key conceptual contribution is that DGC increments provide *local* per-step control, enabling optimized stepsize schedules that adapt to data geometry -- with provable logarithmic-to-constant separations over single-block schedules for Gaussian mixtures. Wainwright further exploits a martingale structure to build fully data-certified algorithms and connects DGC to information-theoretic quantities (covariance, rate distortion, metric entropy, Poincare constant), recovering and sharpening existing diffusion guarantees. This is a rare paper that bridges the practical gap (certified sampling schedules) with deep theory (information geometry of the heat flow). Wainwright is at Berkeley, and this is a strong conceptual contribution to **probabilistic methods** and **theory**.

Novel Architectures and Objectives

PIKS: Universal Physics-Informed Kernel Methods (Joachim Bona-Pellissier, Giacomo Meanti, Matteo Santacesaria, Lorenzo Rosasco)

PIKS replaces the neural network backbone of physics-informed learning with kernel methods, and crucially proves *universal consistency* for linear differential constraints without requiring the target to lie in the native RKHS -- a relaxation that prior kernel-based PDE solvers could not achieve. The analysis extends classical operator-theoretic tools to the physics-informed setting, yielding finite-sample bounds under source conditions. This matters because PINNs lack a comparable learning theory due to nonconvex optimization, and PIKS offers a tractable alternative that is empirically competitive with both PINNs and finite element methods. The work comes from Rosasco's group (MIT/Genoa), and represents a genuine **novel architecture** grounded in **theory and generalization** rather than empirical tuning.

Theory and Generalization

On the robustness of noisy solutions in non-convex neural networks (Enrico M. Malatesta, Alessandra Passalacqua, Riccardo Zecchina)

This work extends the overlap gap property (OGP) framework from zero-temperature perceptron models to the finite-temperature regime, showing that the frozen 1RSB solution survives at any positive training error. The authors derive a general smoothness criterion for when finite-energy relaxation removes freezing, and show that algorithmically accessible dense regions persist beyond the zero-temperature OGP threshold, up to a new threshold that grows with allowed error. In the teacher-student setting, they demonstrate that these wide finite-energy regions retain good generalization even where zero-temperature recovery is computationally hard. This is a substantive contribution to **theory and generalization**, connecting statistical physics of optimization to practical algorithmic accessibility -- the finite-temperature extension of OGP is conceptually novel and not an incremental result.

A Compositional Theory of Causally Masked Transformers (Franz Nowak, Ryan Cotterell, Reda Boumasmoud)

This paper develops an algebraic formalization of transformer expressivity under *finite precision*, deriving expressivity bounds from the model's implemented dynamics rather than idealized arithmetic. The central object is the memory state computed by attention, and the authors show that different attention types correspond to distinct semigroup classes (definite, R-trivial, locally R-trivial, aperiodic), with tight bounds under a free-wiring assumption. The finite-precision framing is important because rounding and evaluation order change what attention retains, and existing expressivity results that assume exact arithmetic miss these effects. Ryan Cotterell is at Cambridge, and this work provides a rigorous **mechanistic interpretability** and **theory** contribution that grounds transformer expressivity in algebraic structure.

Back tomorrow with more!

P.S. In the most recent submission window: cs.LG: 141 papers stat.ML: 19 papers