Hey there,

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

Top Picks

Novel Architectures and Objectives

LpWM: A Case for Sparse Representations in World Models -- *Yann LeCun* is among the authors. This paper challenges the default assumption that dense latent representations are optimal for JEPA-style world models. The key theoretical result shows that nonlinear Lipschitz dynamics can be approximated arbitrarily well by action-conditioned linear dynamics in a sufficiently high-dimensional one-hot latent space, with rollout error vanishing as dimension grows. This motivates a practical relaxation using distributed sparse codes regularized via Rectified Distribution Matching to match a Rectified Generalized Gaussian, yielding non-negative sparse features. The empirical payoff is substantial: sparse LpWM outperforms dense LeWM by up to 57% in planning success at intermediate predictor capacities on PushT, and the learned representations exhibit mode-factored structure where support encodes discrete dynamical regimes and magnitudes capture continuous within-regime state. This is a genuine conceptual departure from the isotropic-Gaussian assumption that pervades JEPA and contrastive representation learning broadly.

Theory and Generalization

One Inverse Step is a Convex Program: Bayes-Limit Calibration of Diffusion Inversion -- This work reframes a single DDIM inversion step as the stationarity condition of an explicitly constructed potential that is strongly convex at the Bayes limit, with modulus exactly determined by the log-SNR gap -- and this holds for every data law, schedule, and point without manifold, reach, or unimodality hypotheses. The paper derives a hypothesis-free certificate of model error: a second solution requires the trained score to violate the posterior-covariance bound by a computable factor. It then separates three consequences -- uniqueness, solver instability via Picard iteration, and the geometry of convergence domains -- and shows that trained DDPM scores on CIFAR-10 and CelebA-HQ-256 violate the unconditional covariance ceiling by 1.26-4.66x, a derived limitation rather than a null result. The analysis connects score-matching error to concrete geometric violations in a way that could serve as a diagnostic tool for diffusion model quality.

A Commutator Framework for Selective Spectral Alignment in Deep Neural Networks -- This develops a finite-width geometric framework for understanding how learned feature geometries are organized, transported, and selectively aligned across layers. The central construction uses three families of commutators -- between gates and covariance, between sensitivities and covariance, and between average gradient outer products and neural feature matrices -- to quantify incompatibility in weight-generated structures. An exact layerwise identity decomposes the sensitivity-covariance commutator into four distinct sources (downstream transport, adjacent-layer imbalance, pointwise sensitivity fluctuations, and nonlinear gate-covariance interactions), and the AGOP-NFM commutator is shown to be a singular-value-weighted transport of the internal commutator. The framework explains why observed feature-side alignment does not determine the internal geometry from which it emerges, and shows that spectral alignment is a layer- and scale-dependent compatibility phenomenon governed by transport, interaction, cancellation, and damping -- not a universal consequence of training.

Probabilistic and Statistical Methods

Provably adaptive sampling with uniform and remasking discrete diffusion models -- This resolves whether the linear dependence of sampling complexity on ambient dimension in discrete diffusion is intrinsic to the uniform forward process. The answer is no: using a leave-one-out

denoiser for uniform and remasking processes, the authors prove that only $O(\mathrm{DTC}(X_0) / \varepsilon)$ discretization steps suffice, where DTC is the dual total correlation of the target distribution. This means sampling complexity is governed by the intrinsic dependence structure rather than ambient dimensionality. The analysis separates discretization error from score-estimation error via a Bayes-optimal auxiliary sampler, and derives an exact information-theoretic representation of discretization error in terms of mutual information between forward-process coordinates at different times. This is a clean theoretical result that clarifies when and why parallel discrete diffusion sampling can avoid the curse of dimensionality.

Until the next announcement!

P.S. In the most recent submission window: cs.LG: 250 papers stat.ML: 31 papers