LoLa: An Empirical Study of Latent Diffusion Models for Physics Emulation

François Rozet, Ruben Ohana, Michael McCabe, Gilles Louppe, François Lanusse, Shirley Ho

Polymathic FLATIRON INSTITUTE

arXiv:2507.02608
polymathicai/lola

TL;DR We show that latent diffusion models are surprisingly robust to compression (up to 1000×) for physics emulation, and consistently outperform their non-generative and pixel-space counterparts.

Research question

Numerical simulations of **dynamical systems** are ubiquitous in science and engineering but require significant computational resources.

$$x^1 \xrightarrow{x^{i+1} = f_\theta(x^i)} x^2 \longrightarrow \cdots \longrightarrow x^L$$

A widespread strategy is to train a neural network, called neural solver (NS), to approximate the dynamics f_θ orders of magnitude faster than numerical solvers. However, **neural solvers suffer from long-term instability issues** and cannot model uncertainty.

Recently, diffusion models (DMs) were found to mitigate the instability of non-generative emulators. However, they are slow at inference.

For images and videos, it is known that **generating in the latent space of an autoencoder** leads to both quality and efficiency improvements.

But **does this carry to dynamical systems?**

Method	Time
simulator	O(10 s)
NS	56 ms
latent NS	13 ms
DM	O(1 s)
latent DM	84 ms

Methodology

For 3 datasets from **The Well**, we train DCAE-like autoencoders (E_ψ, D_ψ) with **multiple compression rates** (e.g. 64, 256, 1024). We then train DiT-like diffusion models and neural solvers to predict the next $n = 4$ latent states $z^{i+1:i+n}$ given the current state $z^i = E_\psi(x^i)$.

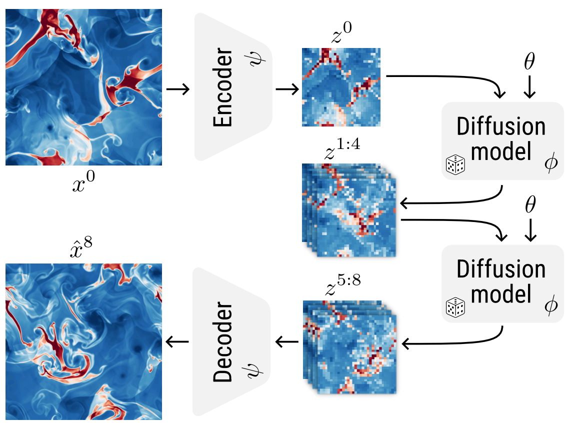

Figure 2. Illustration of latent autoregressive rollout.

Qualitative results

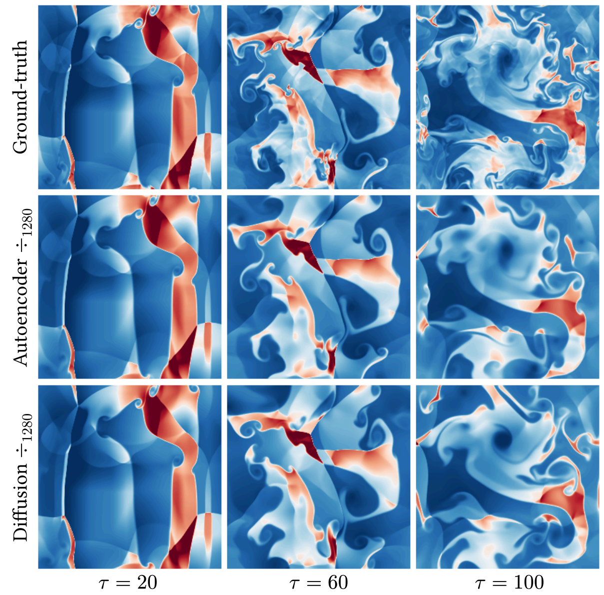

Figure 3. Example of latent emulation for the Euler Quadrants dataset. Even at large compression rates (÷), LDMs **reproduce the dynamics faithfully**, despite significant reconstruction artifacts.

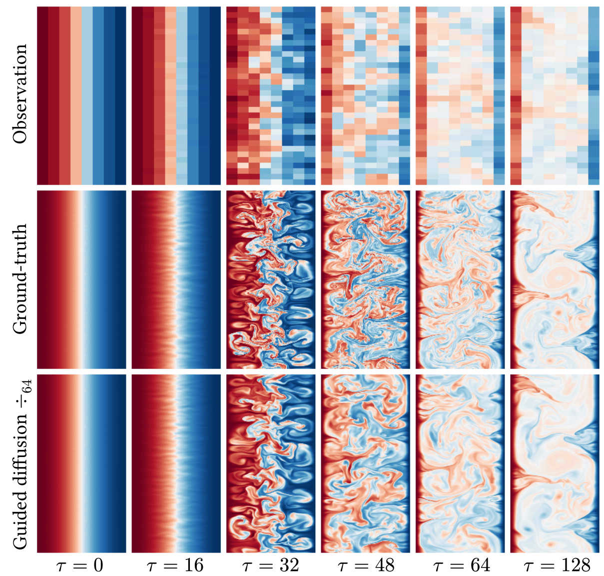

Figure 4. Example of guided emulation for the Rayleigh-Bénard dataset. (L)DMs **enable to incorporate additional information** at inference.

Quantitative results

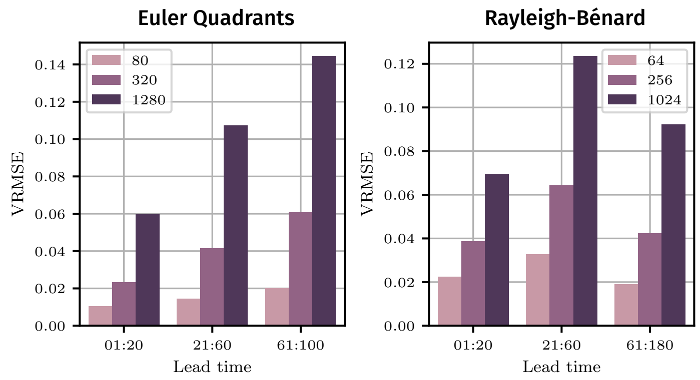

Figure 5. Average VRMSE of the autoencoders' reconstruction. The compression rate has a clear impact on reconstruction quality.

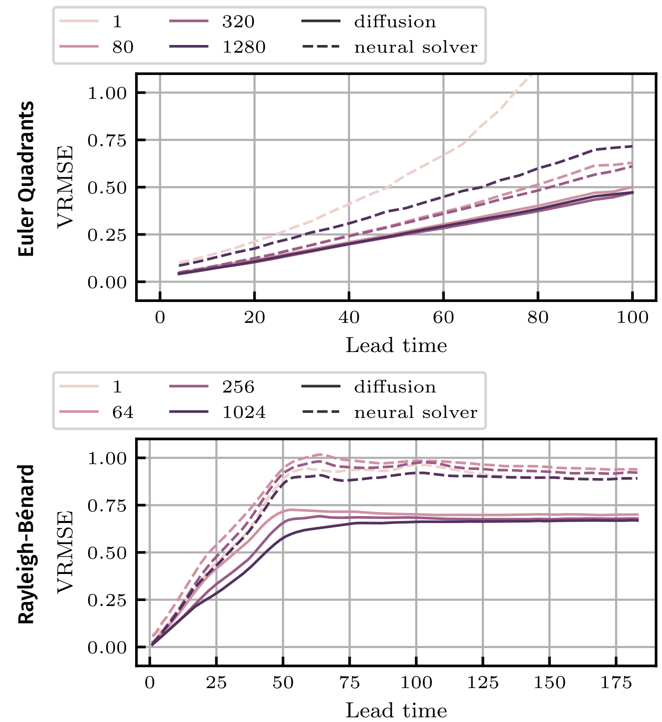

Figure 6. Emulation error grows with the lead time but **increasing the compression rate does not degrade the accuracy** of LDMs.