ENMA: Tokenwise Autoregression for Generative Neural PDE Operators

Armand Kassaï Koupaï, Lise Le Boudec, Louis Serrano, Patrick Gallinari

NEURAL INFORMATION PROCESSING SYSTEMS

TLDR; We propose to solve parametric PDEs using a continuous autoregressive generative model operating in a compressed latent space.

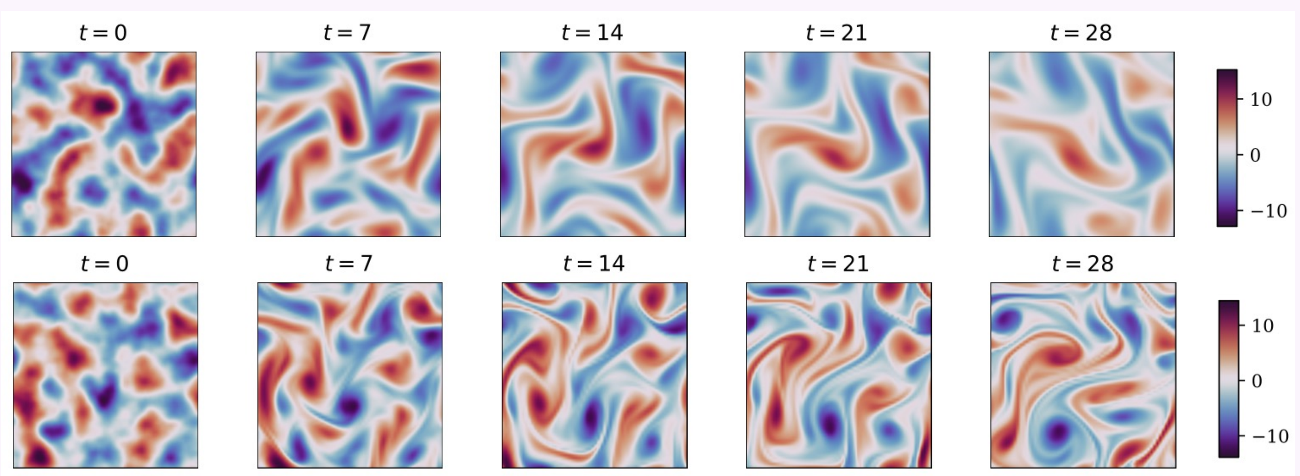
code paper page

1.Problem formulation

- **Solve parametric PDEs**

$$
\begin{aligned}
\mathcal{N}[u; c, f](x,t) &= 0, && for\ (x,t) \in \Omega \times (0,T] \\
\mathcal{B}[u; b](x,t) &= 0, && for\ (x,t) \in \partial\Omega \times (0,T] \\
u(x,0) &= u^0(x), && for\ x \in \Omega
\end{aligned}
$$

Where $\mathcal{N}$ and $\mathcal{B}$ denote differential operators, Ω is the spatial domain and T is the time horizon. c refers to the PDE parameters, f is the forcing term, b represents the boundary conditions and u^0 is the initial condition.



Viscosity 5e-3 (top) vs Viscosity 7e-4 (bottom)

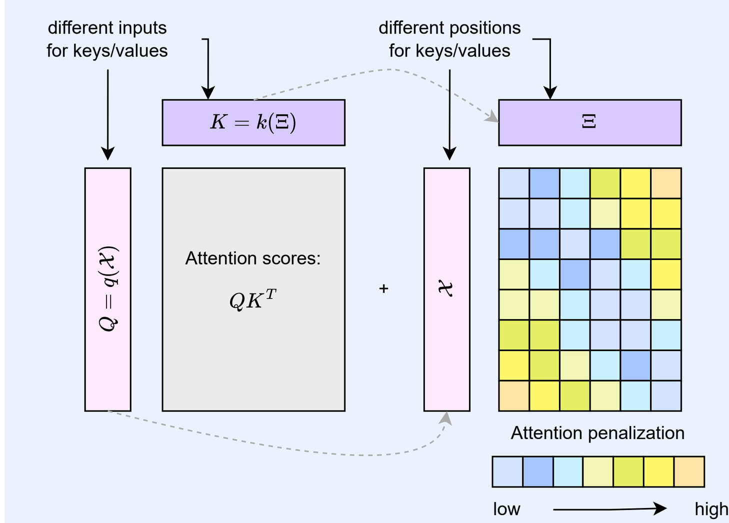
- **3 properties**

Robust to changes in the initial conditions

Robust to changes in the discretization grid

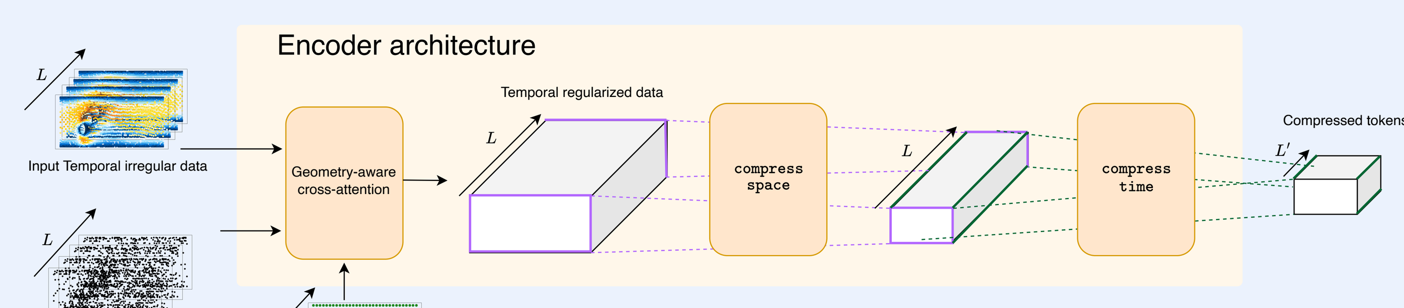
Robust to changes in the PDE parameters

2.Encoding strategy

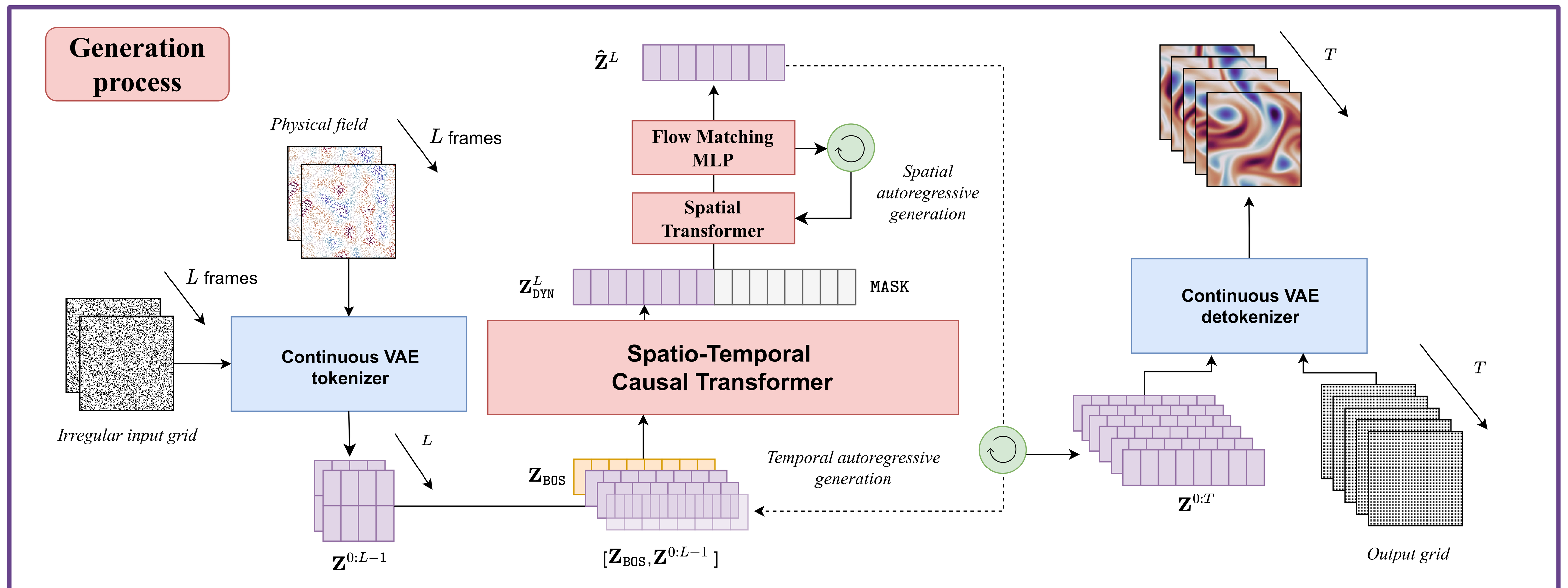


1. Interpolation

2. Compression



ENMA can handle irregular grids and proposes an efficient encoding strategy.



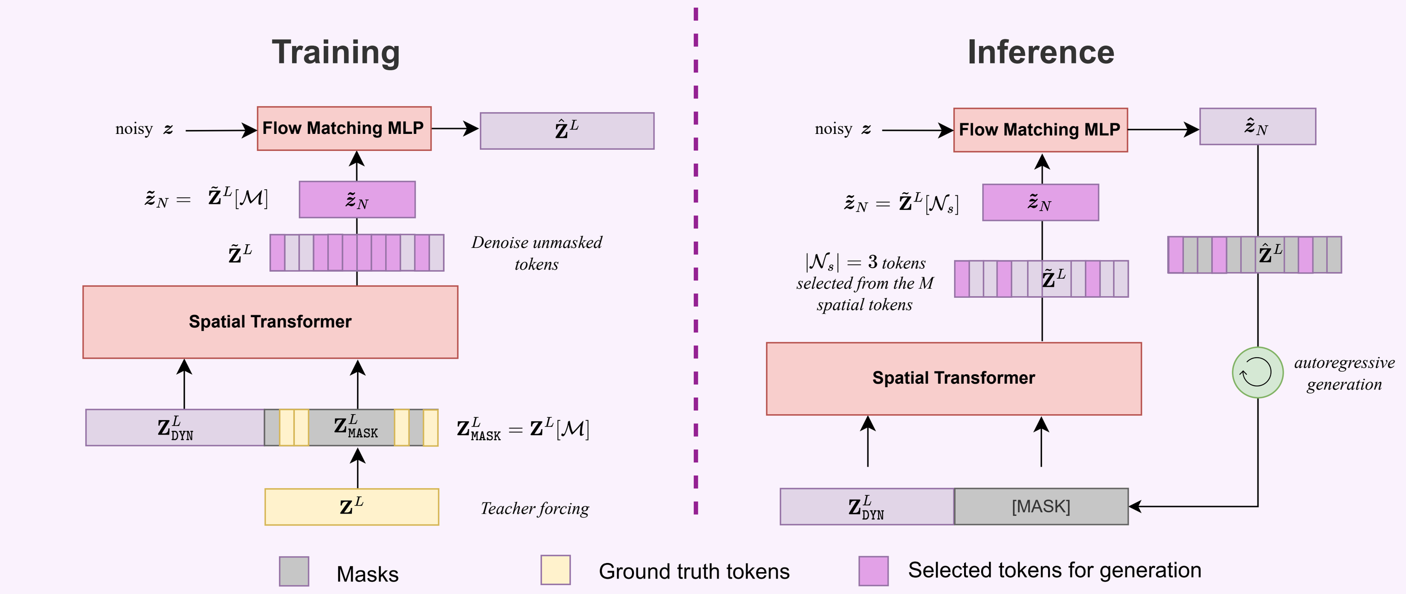
3.Generative process

1. Extract spatio-temporal representation

We extract the spatio-temporal representation of a trajectory by using a **causal attention mechanism** Z_{DYN}^{L} = CausalTransformer(Z_{BOS},$Z^{0:L-1}$), where Z_{BOS} is a **learned begin-of-sequence token**. Z_{DYN}^{L} can be seen as a **latent context capturing the dynamics**, from the observed time-steps [0,L−1], to predict the next step L.

2. Auto-regressive spatial generation

To perform **tokenwise continuous autoregression**, ENMA employs a masked decoding scheme conditioned on the context Z_{DYN}^{L}. This is implemented using a **spatial transformer**, combined with a lightweight MLP that **models the per-token ouput distribution.**



ENMA is a generative autoregressive model that operates continuously in the latent space.

4.Experiments

1. Encoder-Decoder quality

$\downarrow \mathcal{X}_{te}$	Dataset → / Model ↓	*Vorticity* Reconstruction	Time-stepping	Compression rate
$\pi = 100\%$	OFormer	9.99e-1	1.00	×0.125
	GINO	5.63e-1	9.83e-1	×8
	AROMA	1.45e-1	1.13	×8
	CORAL	4.50e-1	9.85e-1	×2
	ENMA	**9.20e-2**	**2.62e-1**	**×15**
$\pi = 50\%$	OFormer	9.99e-1	1.00	-
	GINO	5.69e-1	9.91e-1	-
	AROMA	1.64e-1	1.14	-
	CORAL	4.93e-1	9.85e-1	-
	ENMA	**9.90e-2**	**2.68e-1**	-
$\pi = 20\%$	OFormer	9.99e-1	1.00	-
	GINO	5.90e-1	1.04	-
	AROMA	2.29e-1	1.14	-
	CORAL	7.59e-1	9.87e-1	-
	ENMA	**1.37e-1**	**3.11e-1**	-

Table 1: Comparison of ENMA and encoder-decoder baselines.

ENMA can encode and decode at any resolution.

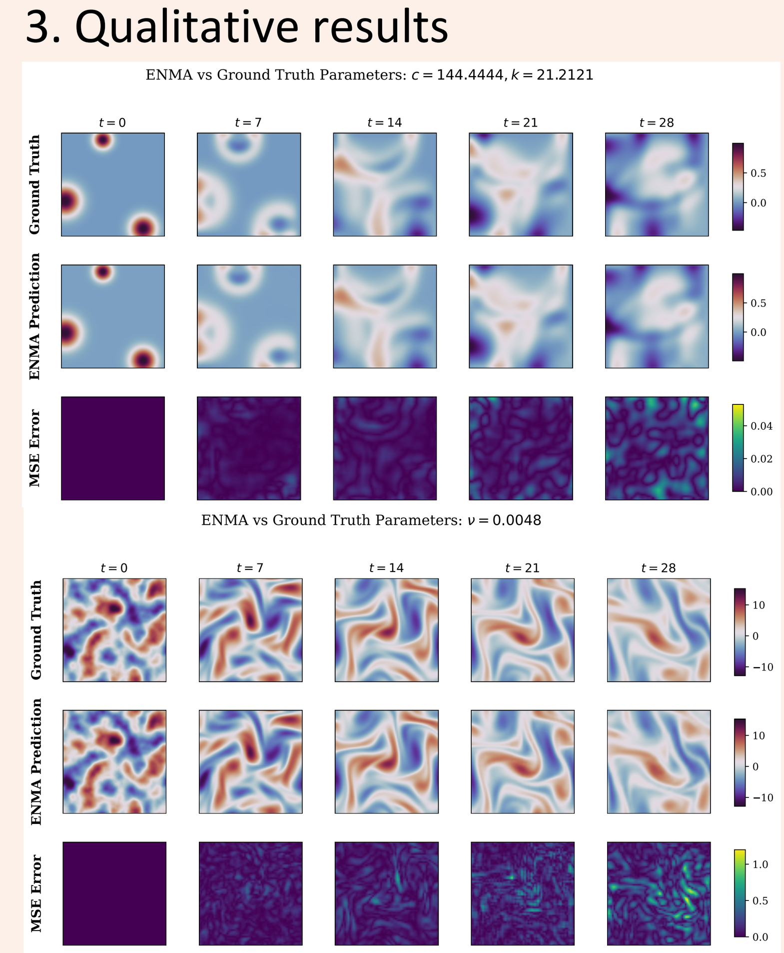
2. Generative process to solve PDEs

Setting ↓	Dataset → / Model ↓	Combined In-D	Combined Out-D	Gray-Scott In-D	Gray-Scott Out-D	Wave In-D	Wave Out-D
Temporal Conditioning	FNO	0.133	26.634	0.0504	0.192	0.691	2.643
	BCAT	0.268	0.928	0.0374	0.1571	0.219	0.538
	AVIT	0.0567	0.305	0.0426	0.168	0.157	0.588
	AR-DiT	0.295	1.797	0.369	0.499	1.117	7.522
	Zebra	0.0182	2.197	0.0421	0.182	**0.140**	**0.315**
	ENMA	**0.00786**	**0.102**	**0.034**	**0.144**	0.145	0.489
Initial Value Problem	In-Context ViT	0.579	1.364	0.069	0.194	0.172	0.624
	[CLS] ViT	0.096	1.160	0.048	0.219	0.556	1.021
	Zebra	0.0478	0.963	**0.044**	**0.1218**	0.169	**0.352**
	ENMA	**0.0156**	**0.330**	0.048	0.134	**0.154**	0.502

Table 2: Comparison of ENMA and baselines on 2 tasks.

ENMA demonstrates strong performance on various difficult benchmarks.

3. Qualitative results



4. Generative capabilities of ENMA

Model	FPD ↓	Precision ↑	Recall ↑
Zebra	1.03×10^{-1}	0.77	**0.86**
ENMA (ours)	9.50×10^{-3}	**0.79**	0.78

Table 3: Generative metrics on the Combined dataset. Lower FPD and higher Precision/Recall indicate better quality and diversity.

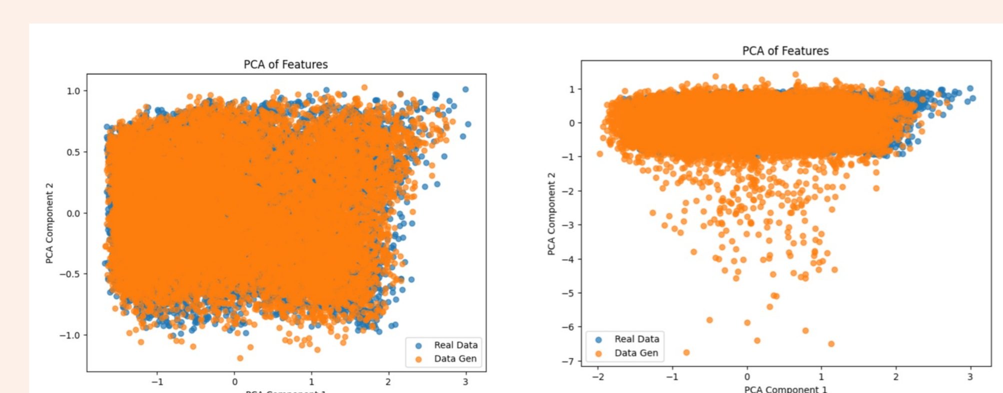


Figure 1: PCA projections of CNN features from generated (orange) and real (blue) trajectories at the final timestep with ENMA (left) and Zebra (right).

5 Uncertainty quantification with ENMA

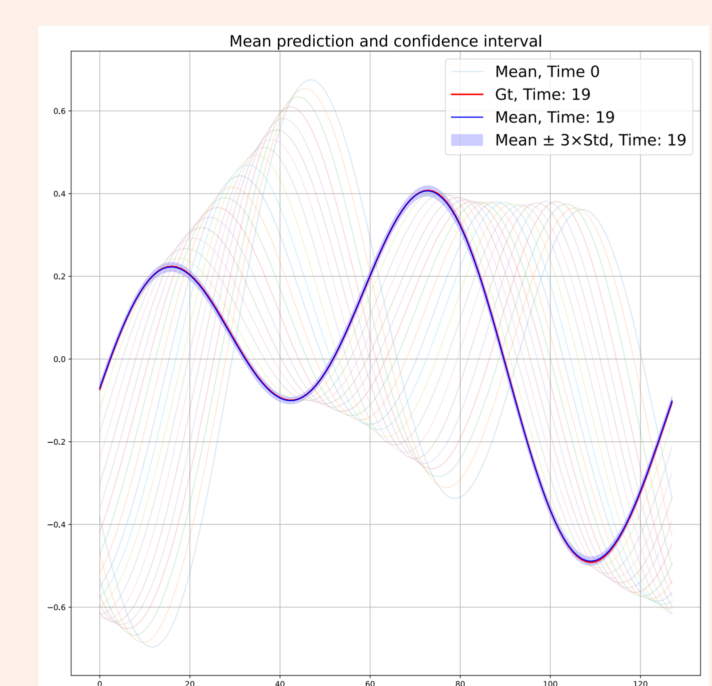


Figure 2: Uncertainty quantification using ENMA. Multiple trajectories are sampled, and the final time step is used to compute the pointwise mean (blue), standard deviation (shaded), and ground truth (red).