NEURIPS 2026

Mesh-Free Convolution: Learned Spectral Attenuation and Transport

Continuous, PDE-inspired filters instead of discrete Fourier weights

Niloufar Zakariaei[1] · Manav Manav[2] · Eldad Haber[1]

[1] University of British Columbia · [2] Computational Geosciences Inc.



THE UNIVERSITY OF BRITISH COLUMBIA



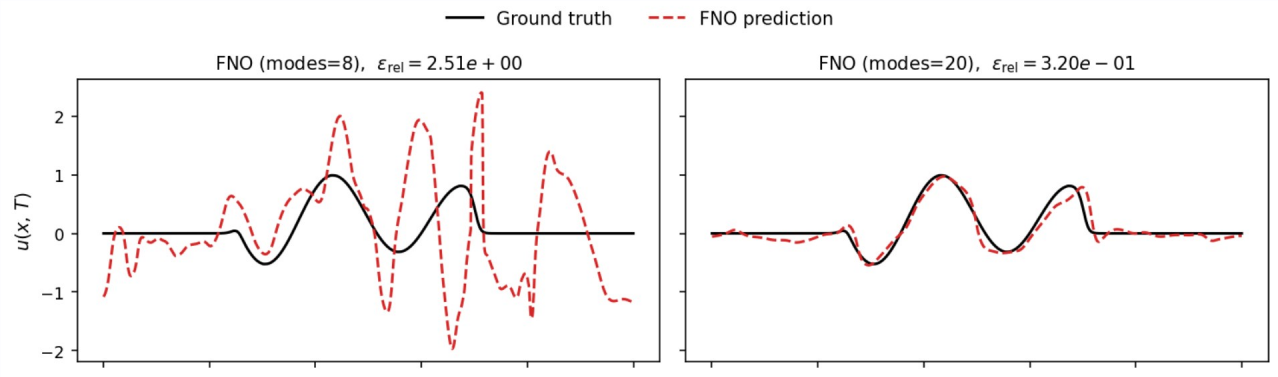
Hard spectral cutoffs break long rollouts

FNO (Fourier Neural Operator) learns one free weight per Fourier mode and keeps only the lowest modes.

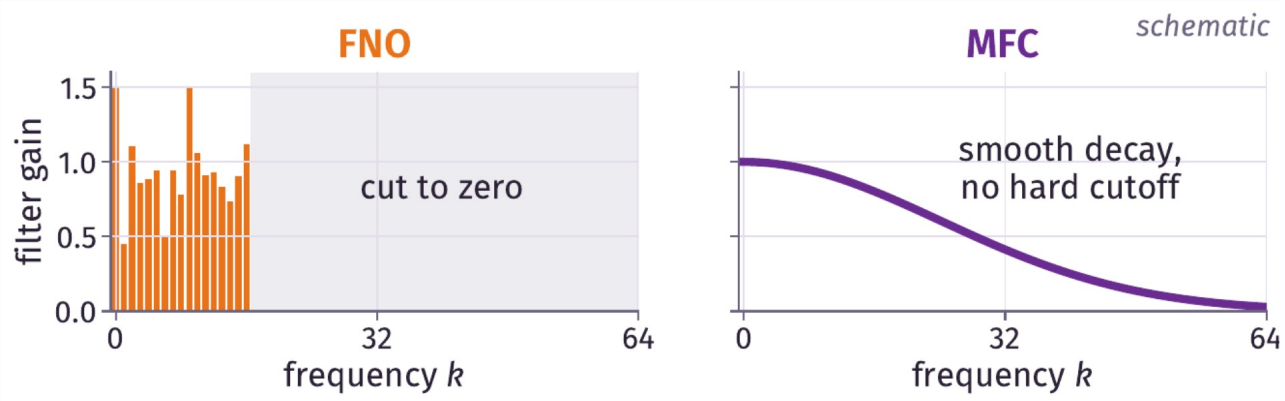
Rolled out step after step, the errors compound: phase drift, ringing, blow-up.

MFC learns a continuous, PDE-inspired filter instead: smooth attenuation, no imposed hard cutoff, sampled on any grid.

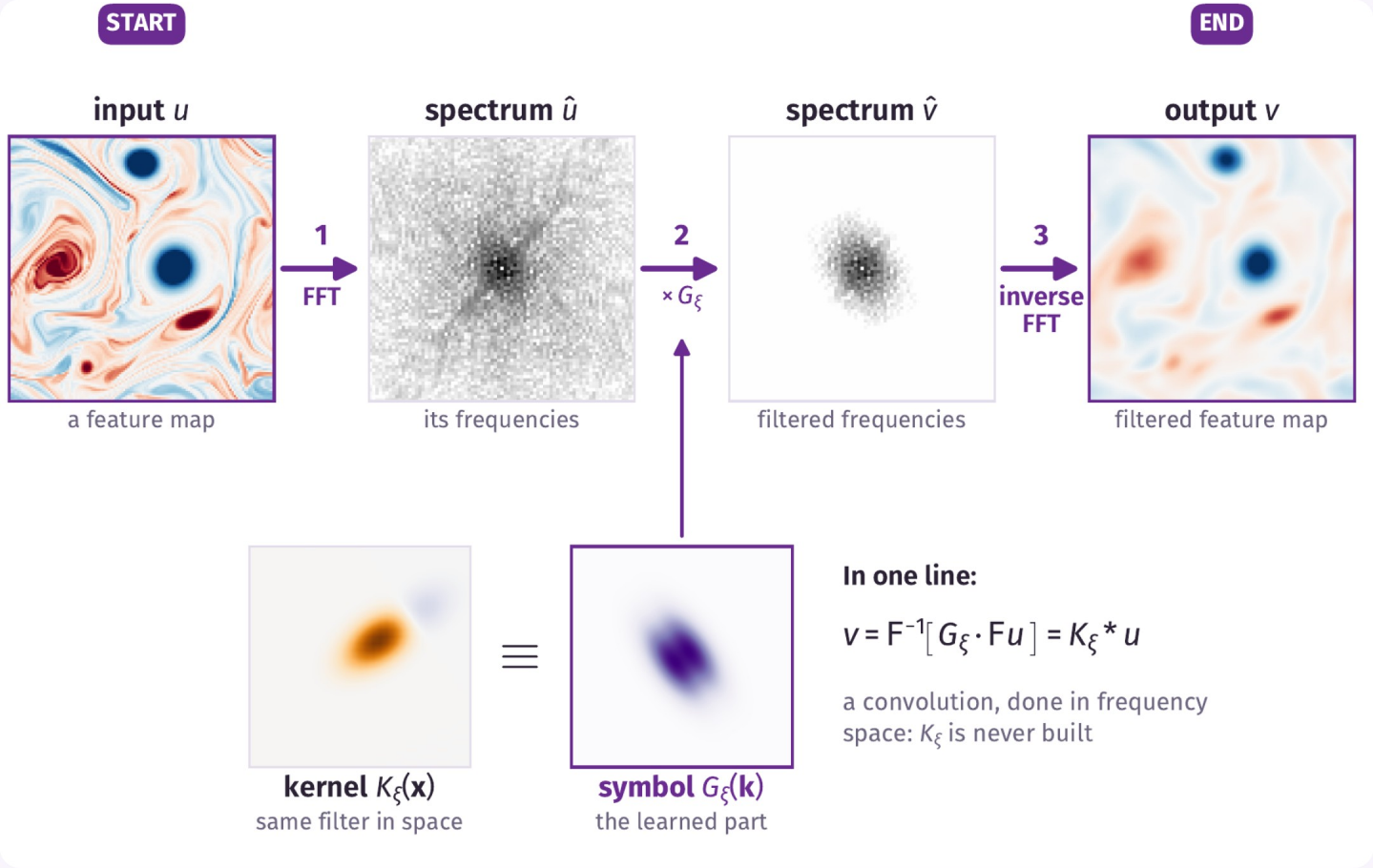
1D advection after many rollout steps: FNO with 8 vs. 20 modes



Filter gain vs. frequency (schematic)



One MFC filter, from input to output



A convolution, done in frequency space:

1. FFT of the input
2. multiply by the learned symbol G
3. inverse FFT

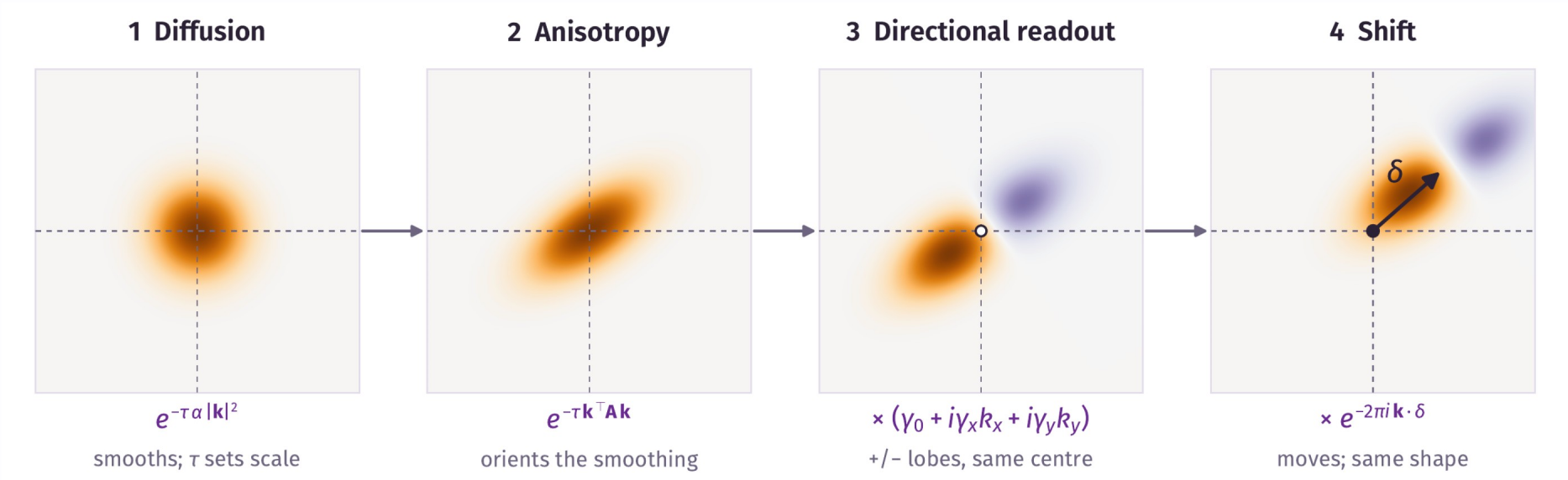
The spatial kernel K is **never built.**

An MFC layer applies one filter per channel, then a 1×1 channel mix.

Illustrative parameters on a real vorticity field.

The learned symbol: a PDE in closed form

$$G_\xi(\mathbf{k}) = \underbrace{e^{-\tau\,\mathbf{k}^\top \mathbf{A}\mathbf{k}}}_{\text{anisotropic diffusion}} \cdot \underbrace{(\gamma_0 + i\boldsymbol{\gamma}\cdot\mathbf{k})}_{\text{directional readout}} \cdot \underbrace{e^{-2\pi i\,\mathbf{k}\cdot\boldsymbol{\delta}}}_{\text{shift (transport)}}$$



No imposed hard spectral cutoff · **mesh-free:** defined for continuous k · **9 parameters** per channel

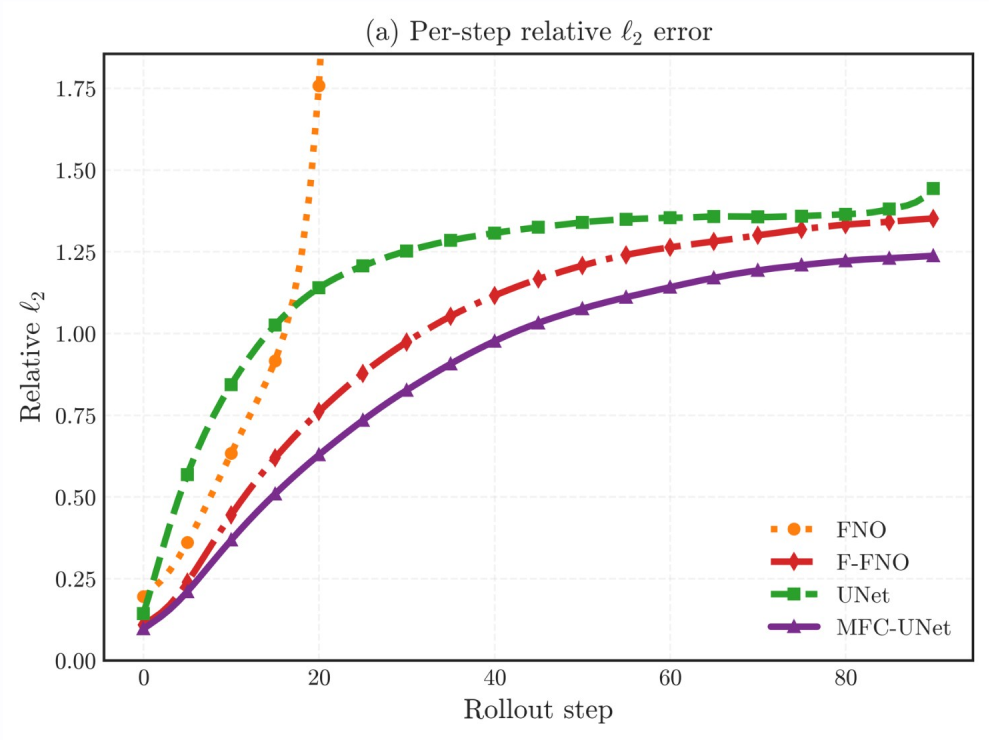
BASELINES

Vanilla FNO diverges in long rollouts. Against its stronger variants, F-FNO and AFNO, MFC still has the lowest long-horizon and zero-shot error.

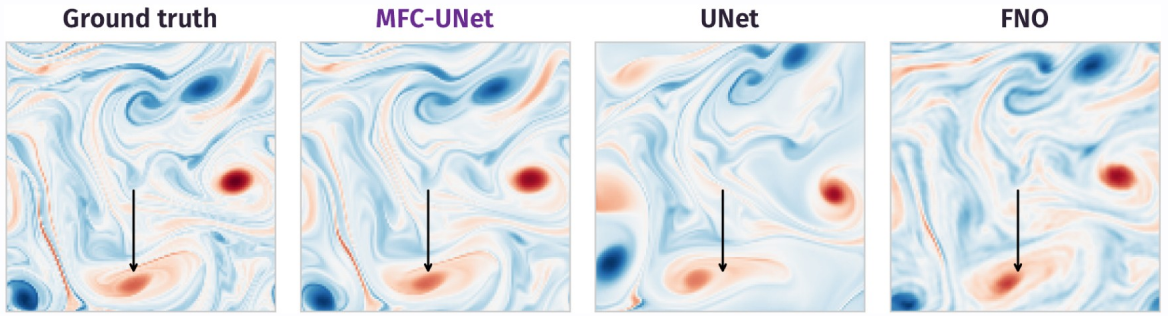
F-FNO: factorized FNO · AFNO: adaptive FNO, the FourCastNet backbone

Stable rollouts: 2D Navier–Stokes

Kolmogorov flow, Re = 5000, 128^2. UNet with MFC layers vs. standard layers, FNO and F-FNO; all trained one step ahead, then rolled out.



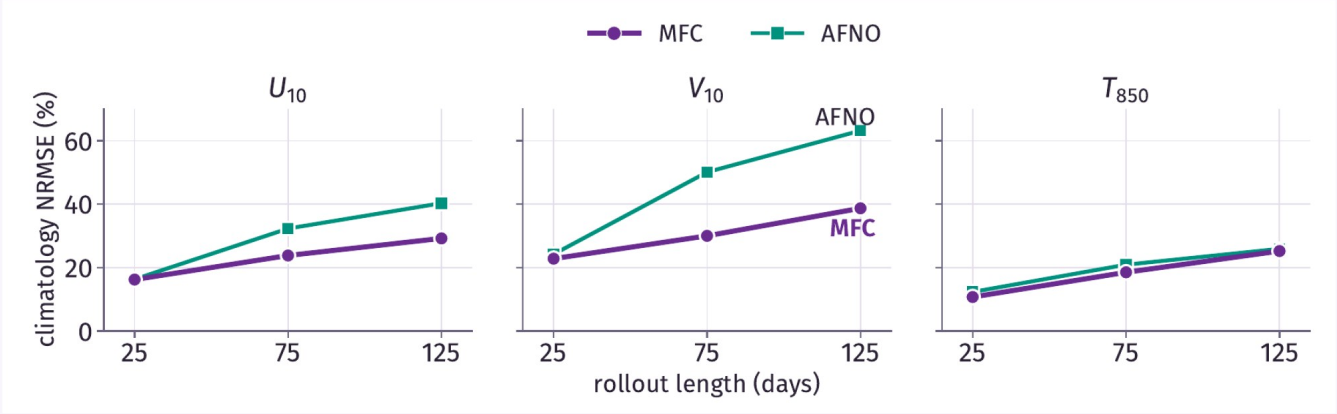
Vorticity after 12 rollout steps

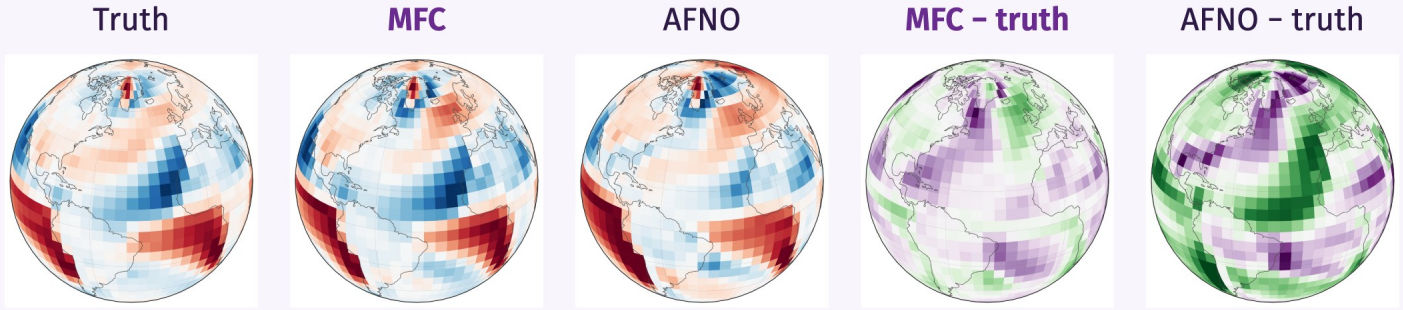


- **MFC: slowest error growth**, below F-FNO
- FNO diverges after about 15–20 steps
- UNet stays bounded by over-smoothing

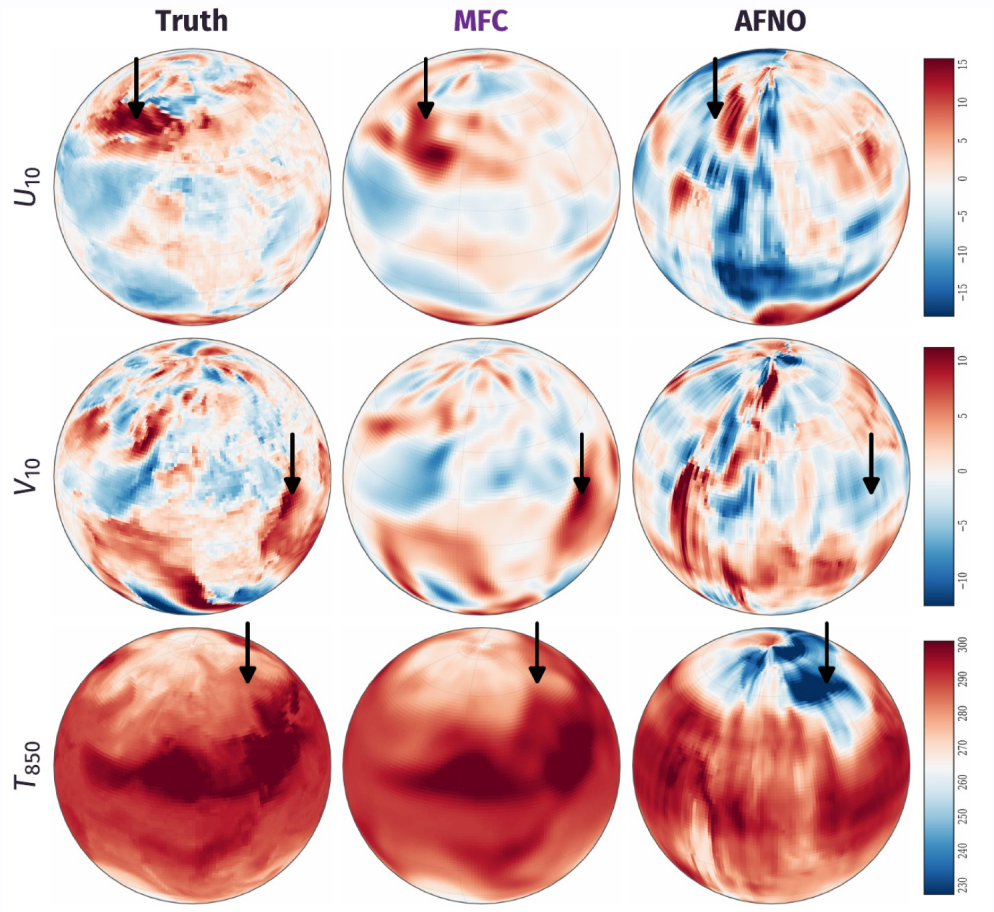
125-day weather rollouts: WeatherBench2

ERA5, 500 six-hour steps from 35 initial conditions. **MFC has the lowest error of the time-mean on every variable and horizon**; FNO diverges in 11/35 runs.





Zero-shot: train coarse, test 3.7× finer



Trained at 64×32, run at **240×121 with no retraining** (50-step rollout; median-improvement case).

- **MFC: lowest error on every variable**, only ×1.02–1.07 vs. its native resolution
- AFNO: up to ×2.97 worse, with grid-scale streaks
- Arrows mark the same spot in each row

Why it transfers: the symbol is a formula in k, so a finer grid simply samples it more densely.

Takeaways

No hard cutoff

A continuous PDE symbol attenuates every frequency smoothly, keeping fine scales over long rollouts.

Mesh-free

The same 9 parameters per channel run on any grid, with no retraining.

Transport-aware

Readout and shift terms model directional, advection-like responses.

Paper: openreview.net/forum?id=QHAqPAgG8a

Code: github.com/NiloufarZK/Mesh-Free-Convolutions

Scope: periodic rectangular grids (FFT); other eigenbases are future work. Stability is proven per layer; long-rollout stability is empirical.