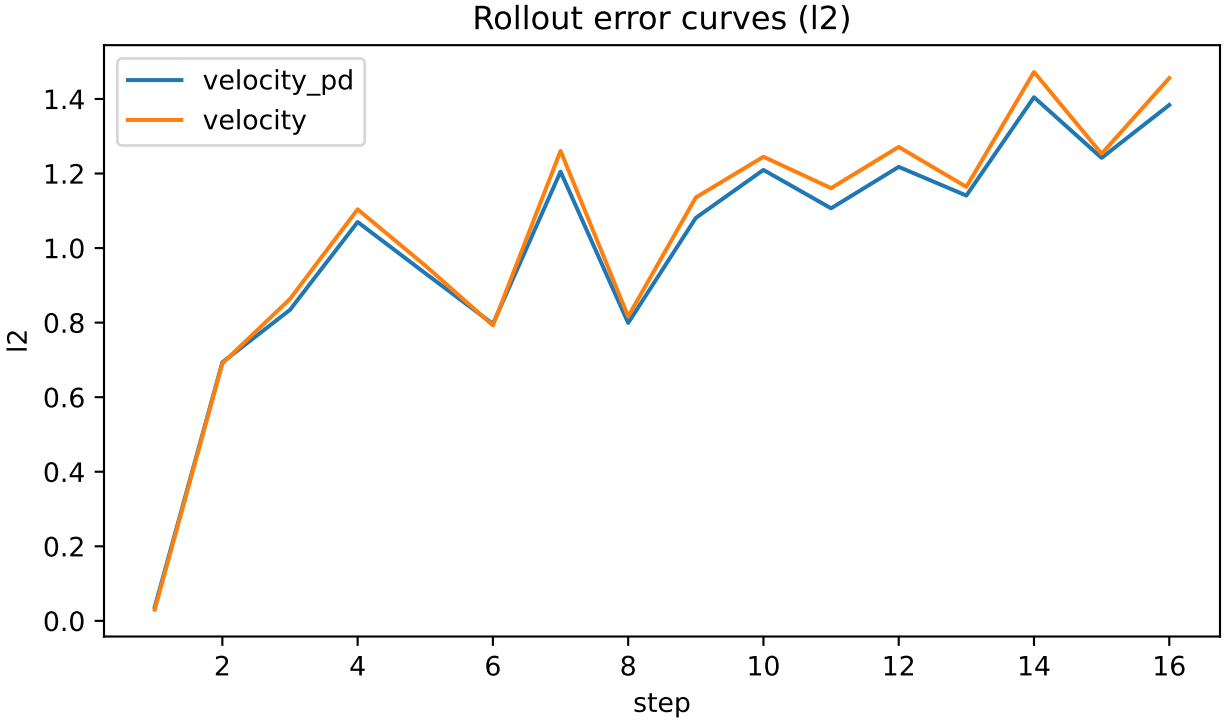
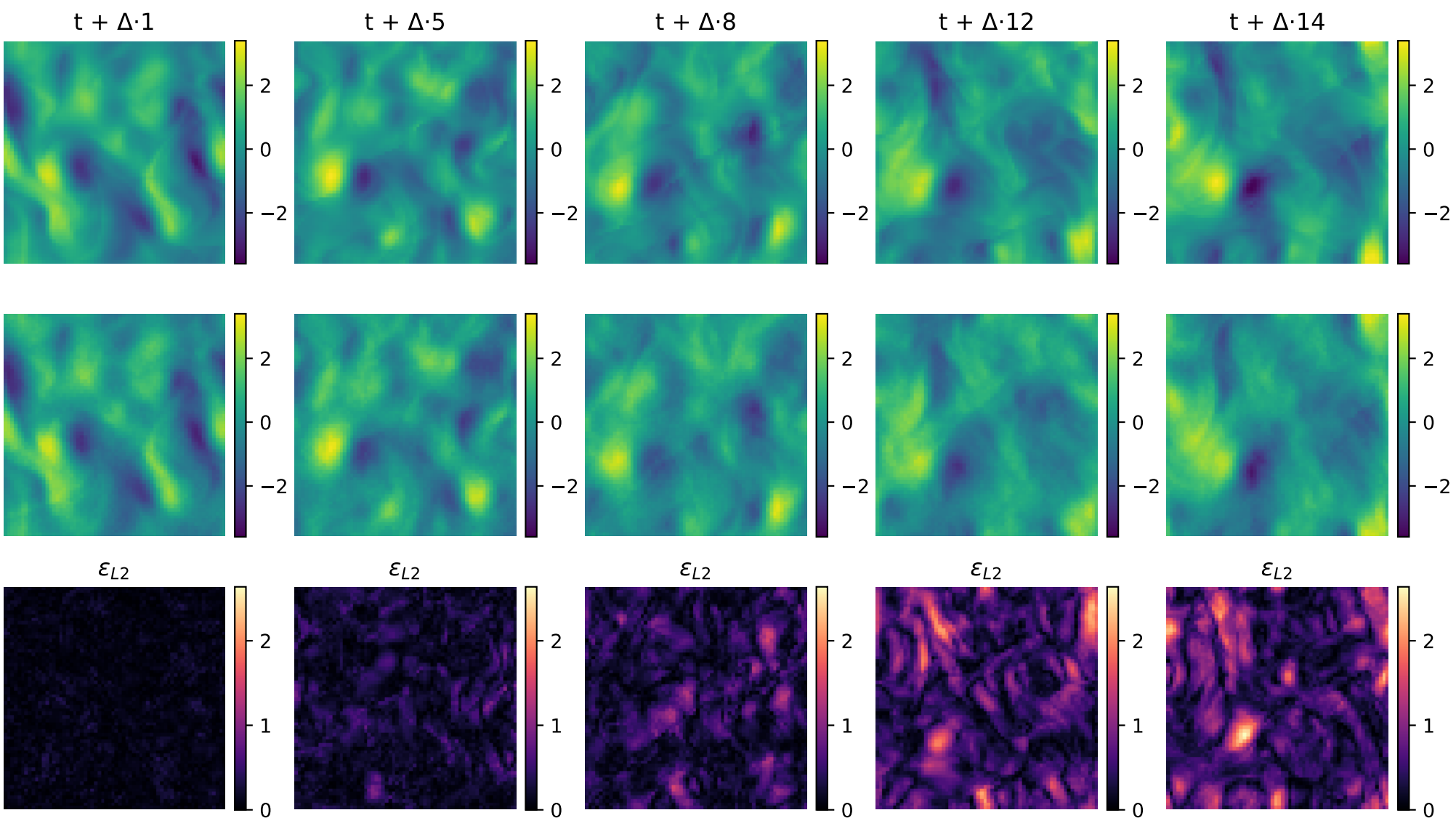
Evaluation Report

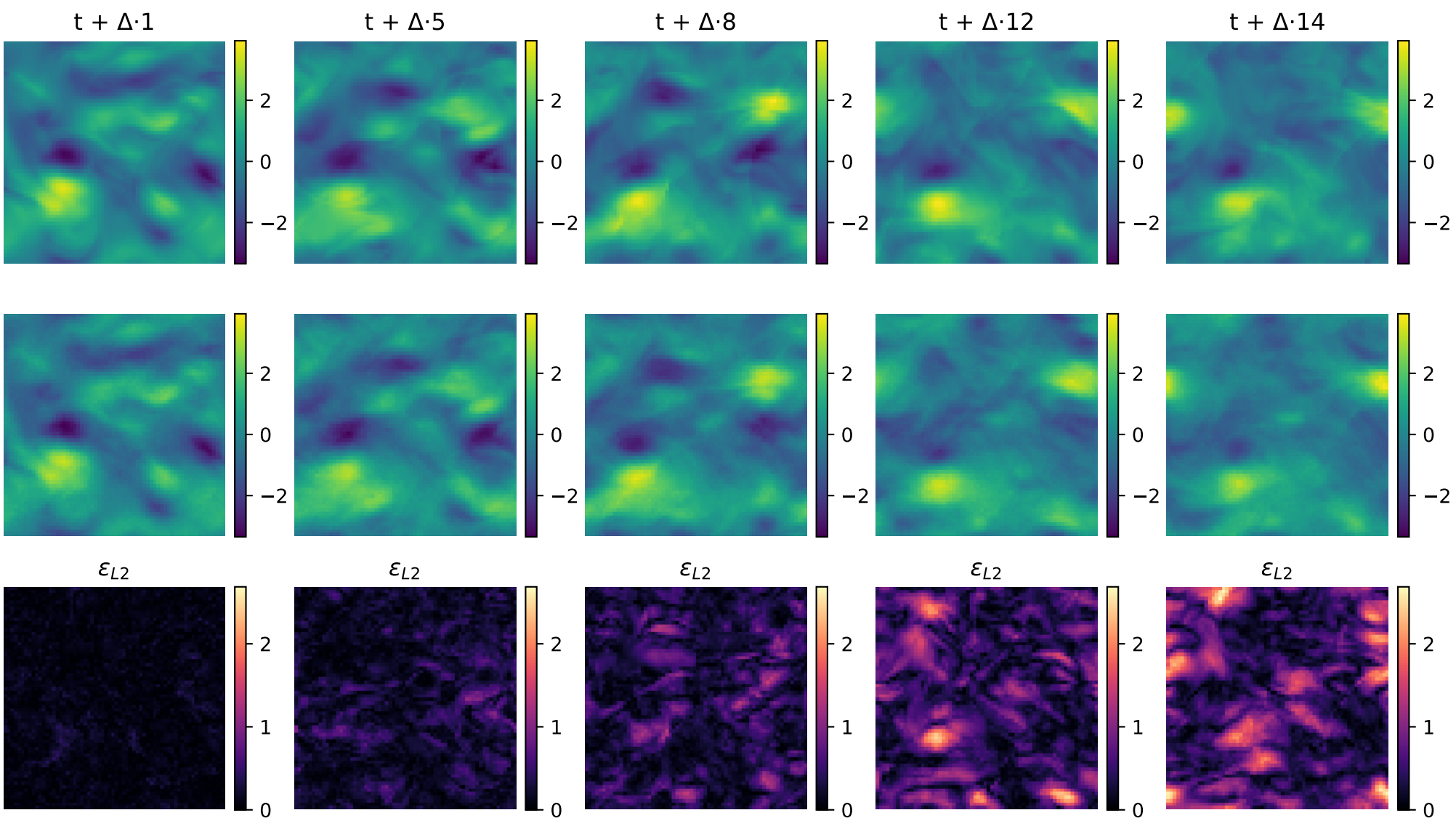
Models: velocity_pd, velocity

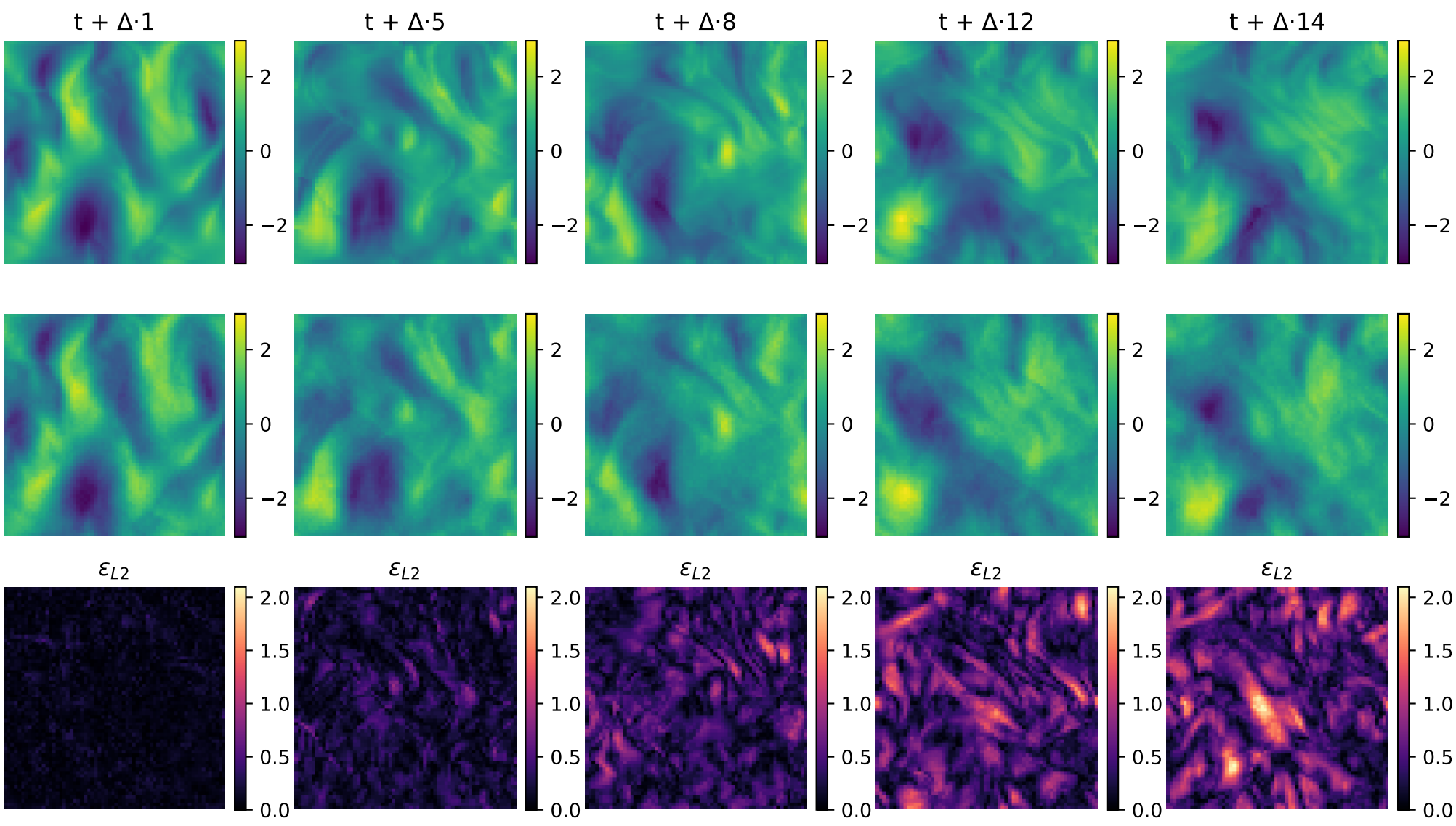
ns2d | 2025_09_12T00_57_48

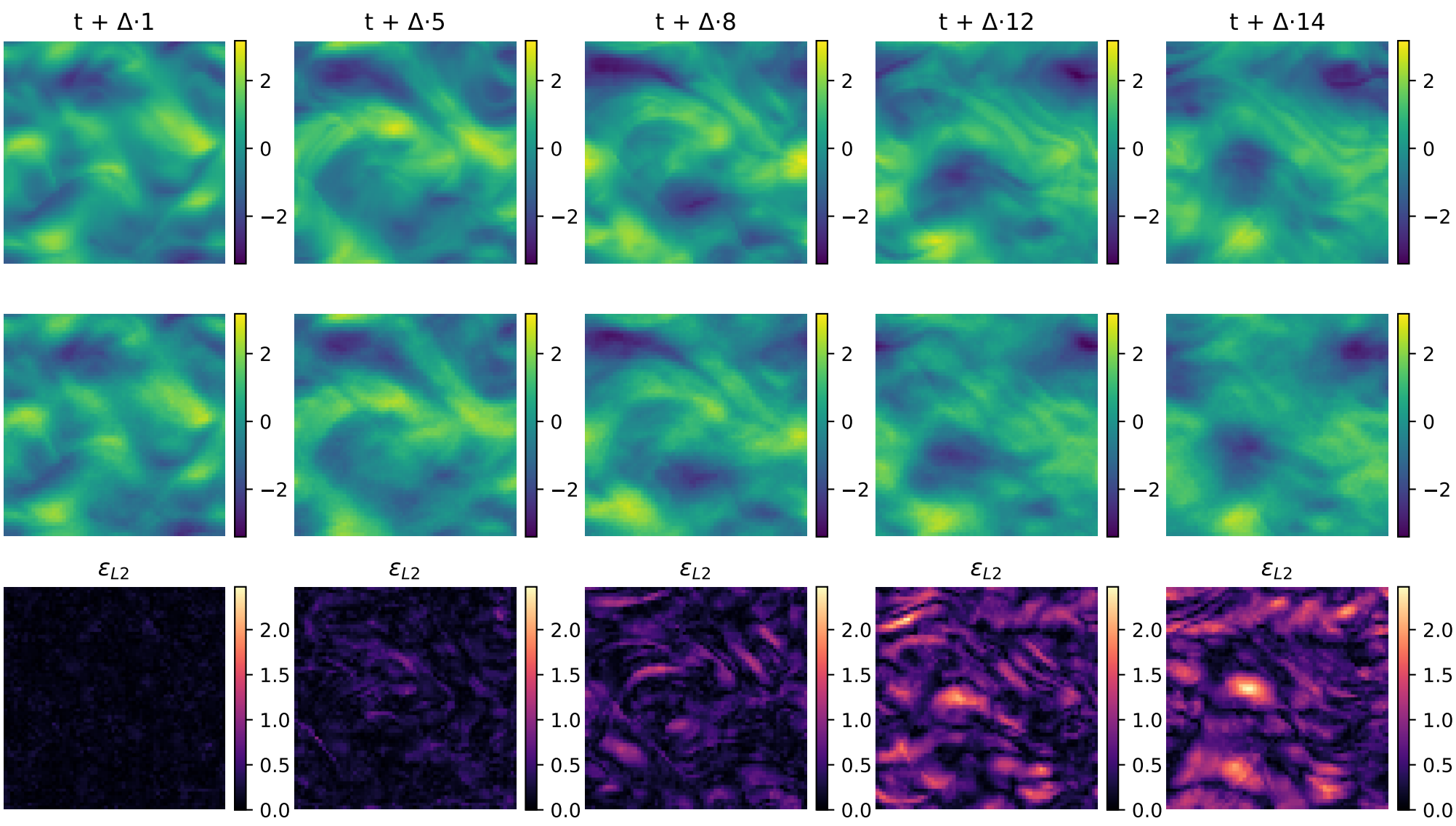
mode	model	metric	mean	median	std	count
rollout	velocity	l2	1.042	1.146	0.5556	256
rollout	velocity_pd	l2	1.01	1.105	0.5226	256

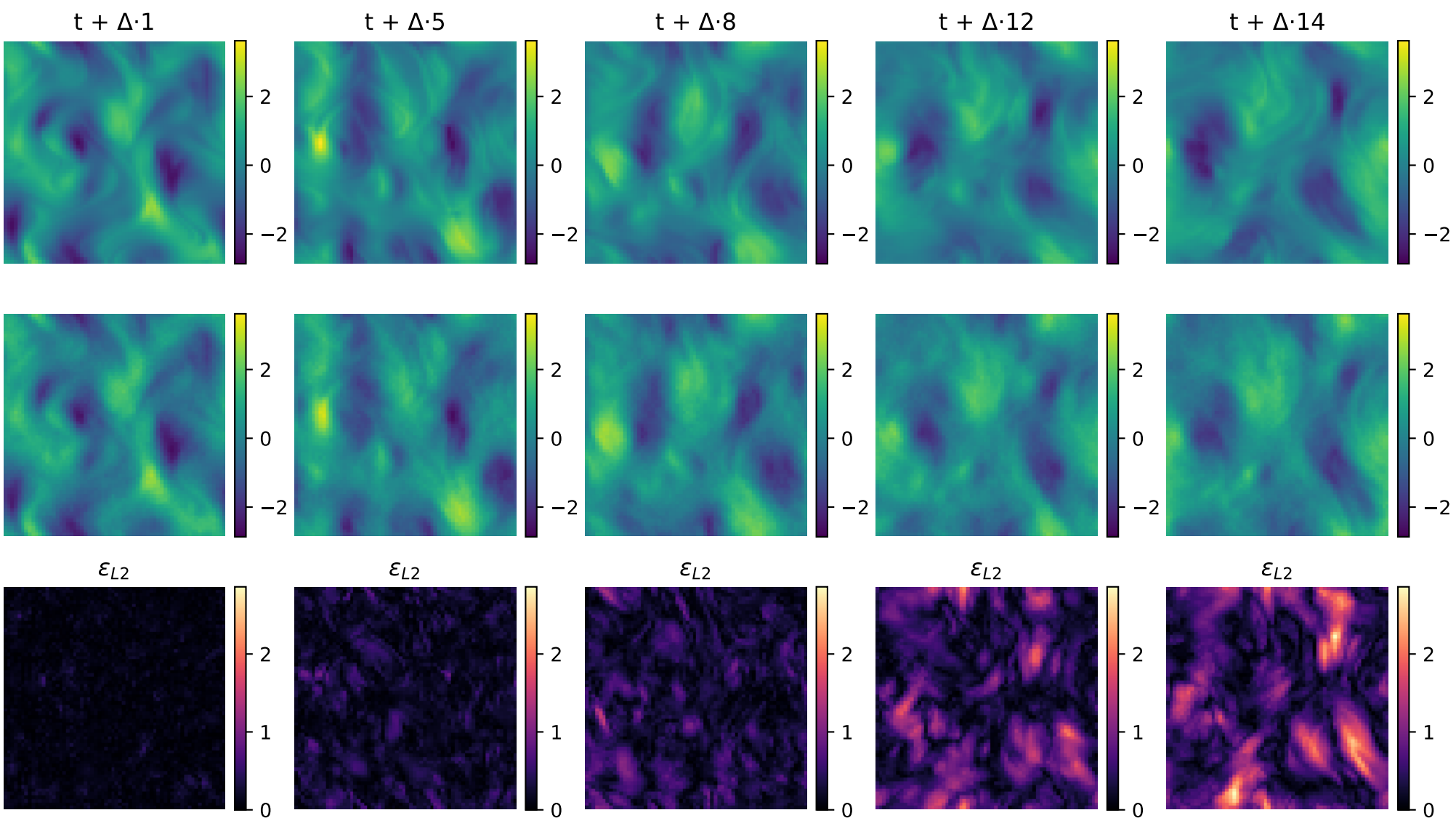

Rollout error curves (l2)
velocity_pd
velocity
l2
step
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
2
4
6
8
10
12
14
16

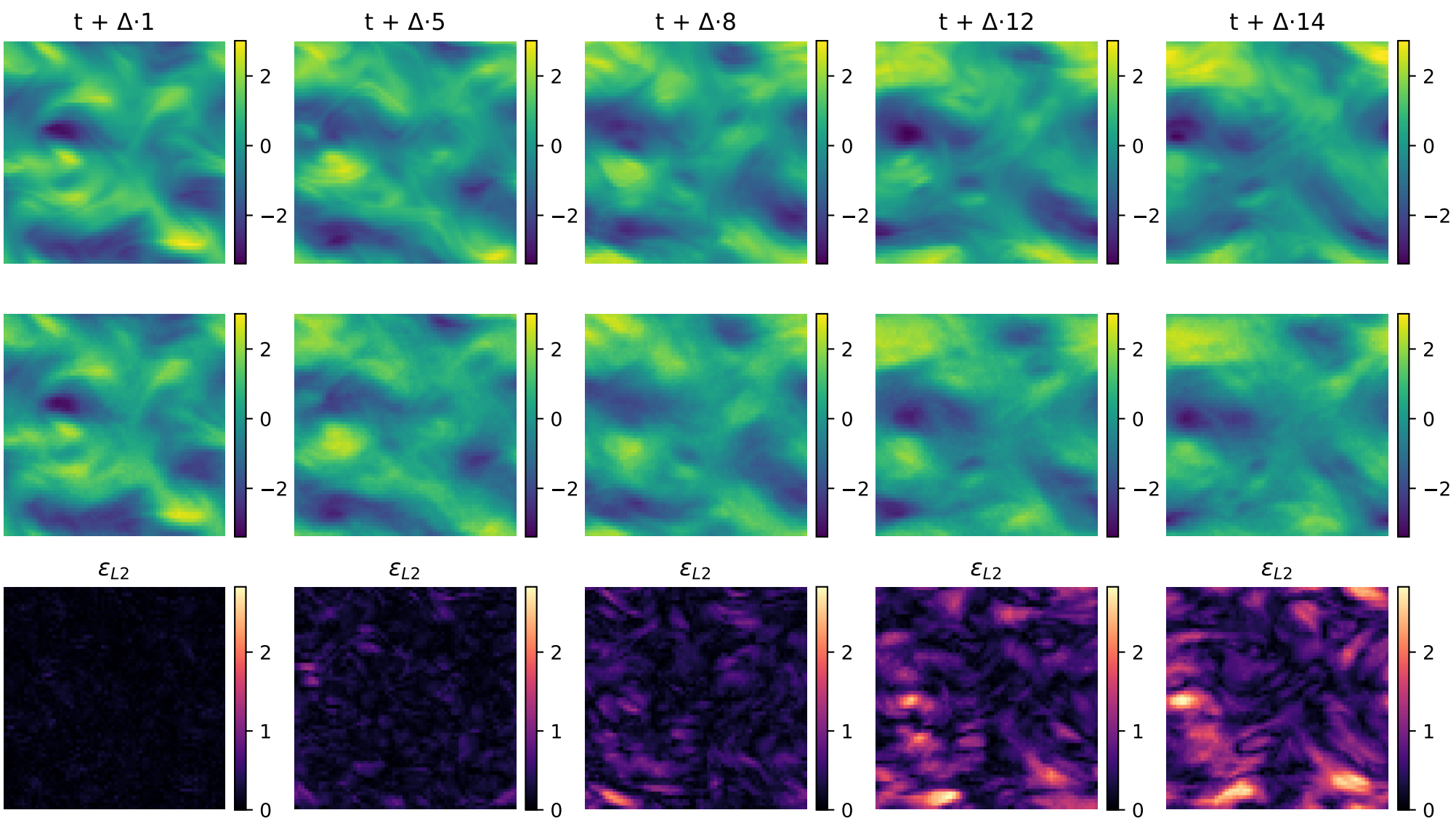

Rollout Snapshot #0 -- ch 0 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
ε_{L2}
2
0
−2
1


Rollout Snapshot #0 -- ch 1 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
2
0
−2
ε_{L2}
1


Rollout Snapshot #1 -- ch 0 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
2
0
−2
ε_{L2}
2.0
1.5
1.0
0.5
0.0


Rollout Snapshot #1 -- ch 1 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
2
0
−2
ε_{L2}
2.0
1.5
1.0
0.5
0.0


Rollout Snapshot #2 -- ch 0 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
2
0
−2
ε_{L2}
1


Rollout Snapshot #2 -- ch 1 -- GT (top row) vs stage_1_student (middle row)
t + Δ·1
t + Δ·5
t + Δ·8
t + Δ·12
t + Δ·14
2
0
−2
ε_{L2}
1

Evaluation Configuration


```
{
  "common": {
    "device": "cuda",
    "seed": 42,
    "samples": 16,
    "batch_size": 32
  },
  "models": [
    "velocity_pd",
    "velocity"
  ],
  "viz": {
    "enabled": true,
    "n_samples": 3,
    "channels": [
      0,
      1
    ],
    "colorbar": true,
    "rollout_steps": 5
  },
  "one_step": {
    "steps": {
      "velocity": 50,
      "flow": 1,
      "velocity_pd": 25
    },
    "method": {
      "velocity": "midpoint",
      "velocity_pd": "midpoint"
    }
  },
  "rollout": {
    "length": 16,
    "latent_policy": "new_latent",
    "steps": {
      "velocity": 50,
      "flow": 1,
      "velocity_pd": 25
    },
    "method": {
      "velocity": "midpoint",
      "velocity_pd": "midpoint"
    },
    "use_sequence_gt_for_metrics": true
  },
  "metrics": {
    "enabled": [
      "l2"
    ],
    "params": {
      "sob": {
        "alpha": 1.0,
        "beta": 0.0
      }
    },
    "use_denormalized": false
  },
  "artifacts": {
    "pdf": {
      "filename": "eval_report.pdf",
      "dpi": 150,
      "max_figs_per_model": 6,
      "include_rollout_curves": true
    },
    "save_csv": true,
    "save_json": true
  },
  "experiment": "ns2d"
}
```