

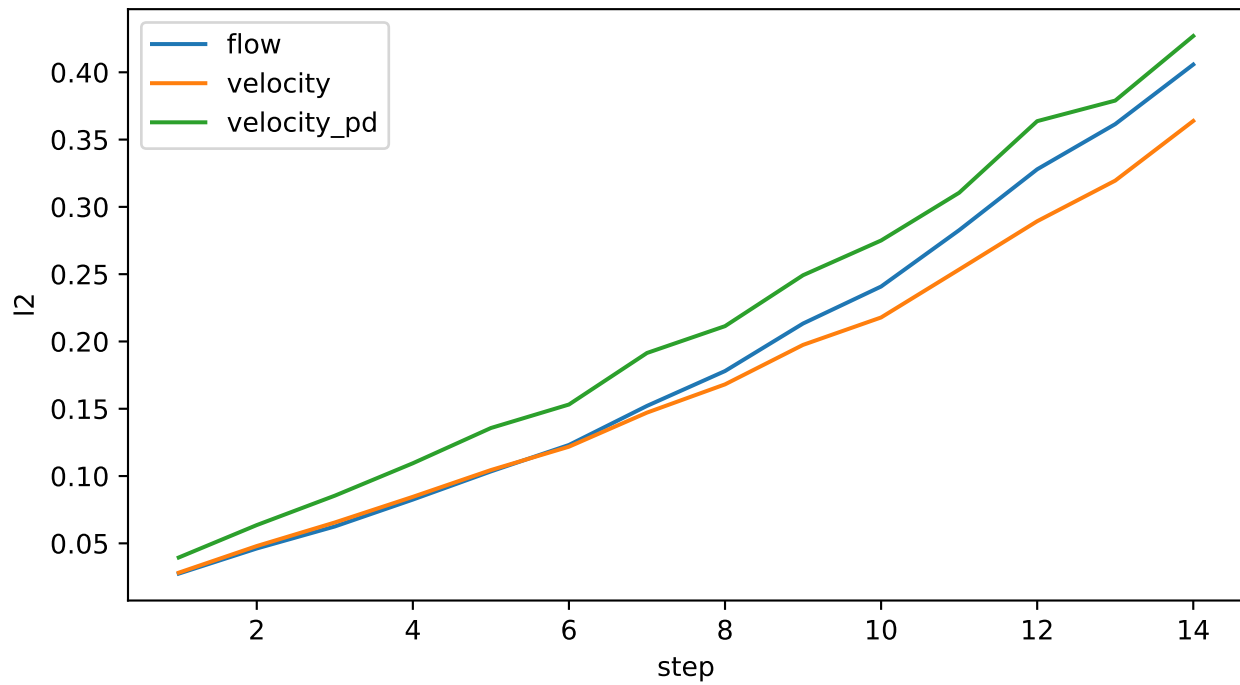
**Evaluation Report**

Models: flow, velocity, velocity\_2d

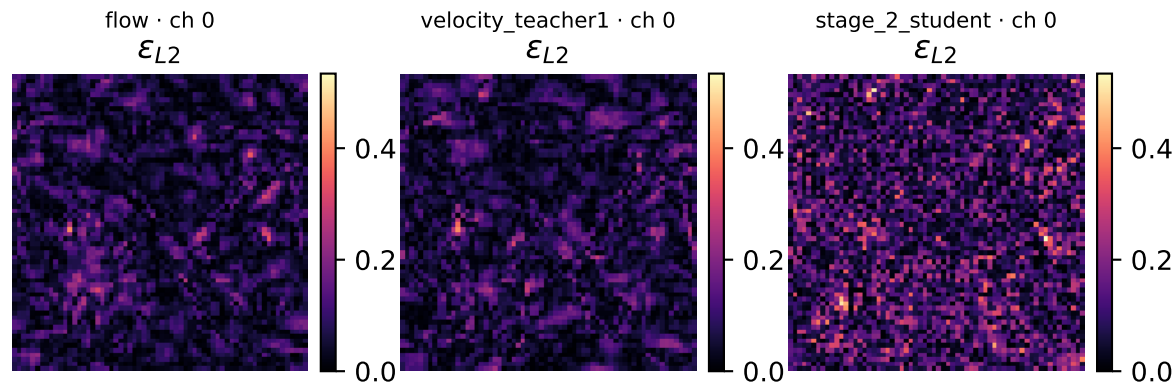
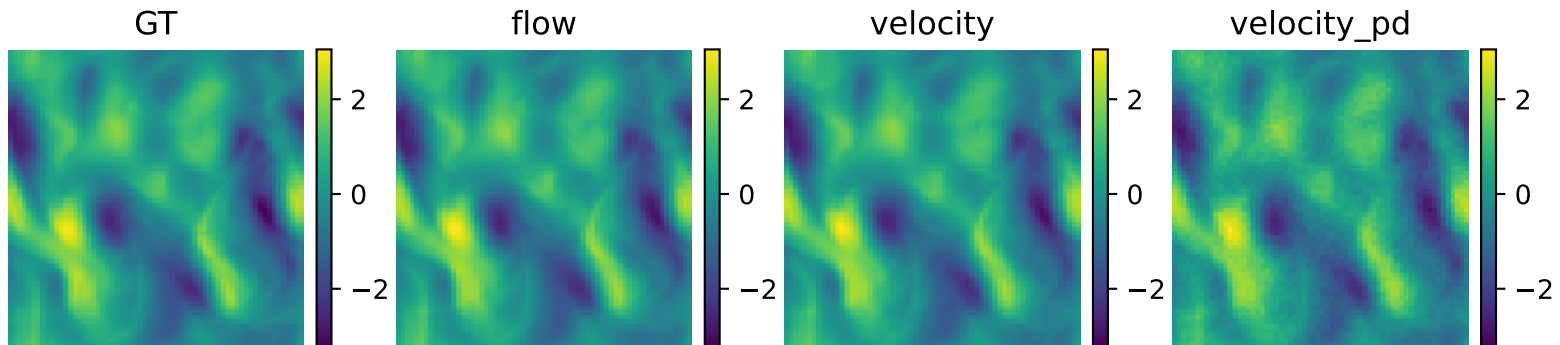
ns2d | 2025\_09\_12T02\_07\_46

mode	model	metric	mean	median	std	count
one_step	flow	l2	0.02852	0.02538	0.01256	128
one_step	velocity	l2	0.02902	0.02494	0.01313	128
one_step	velocity_pd	l2	0.04074	0.03662	0.01293	128
rollout	flow	l2	0.1862	0.1609	0.1277	1386
rollout	velocity	l2	0.1721	0.1517	0.1098	1386
rollout	velocity_pd	l2	0.2139	0.1982	0.1308	1386

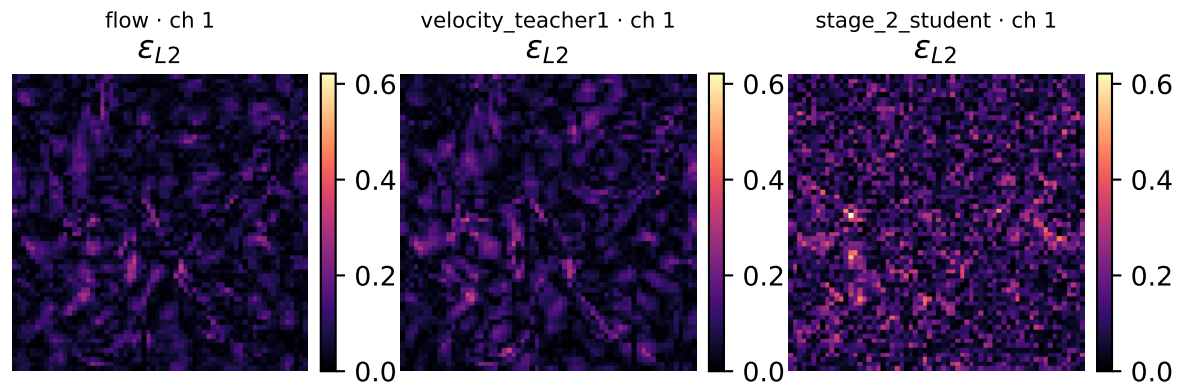
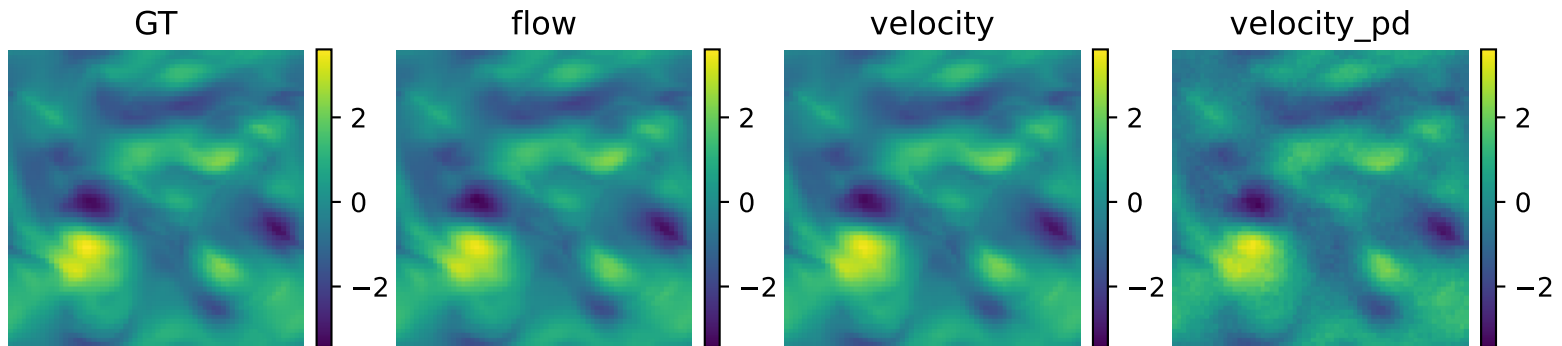
Rollout error curves (l2)



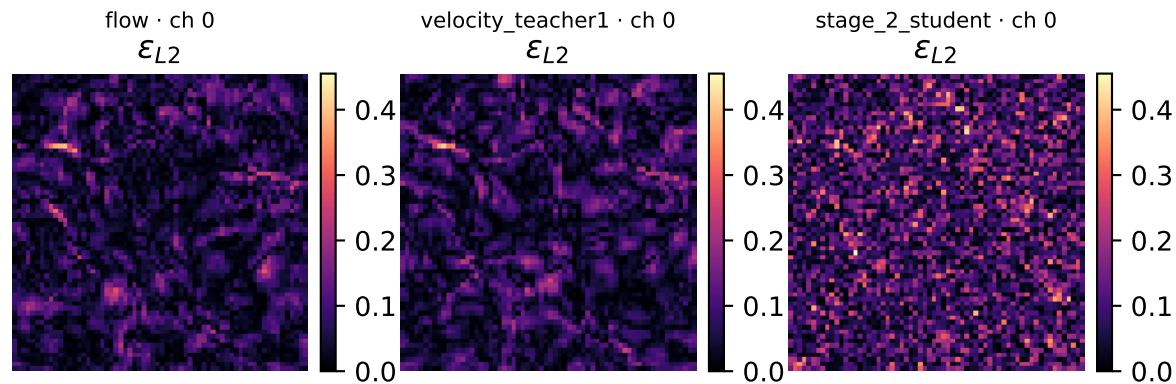
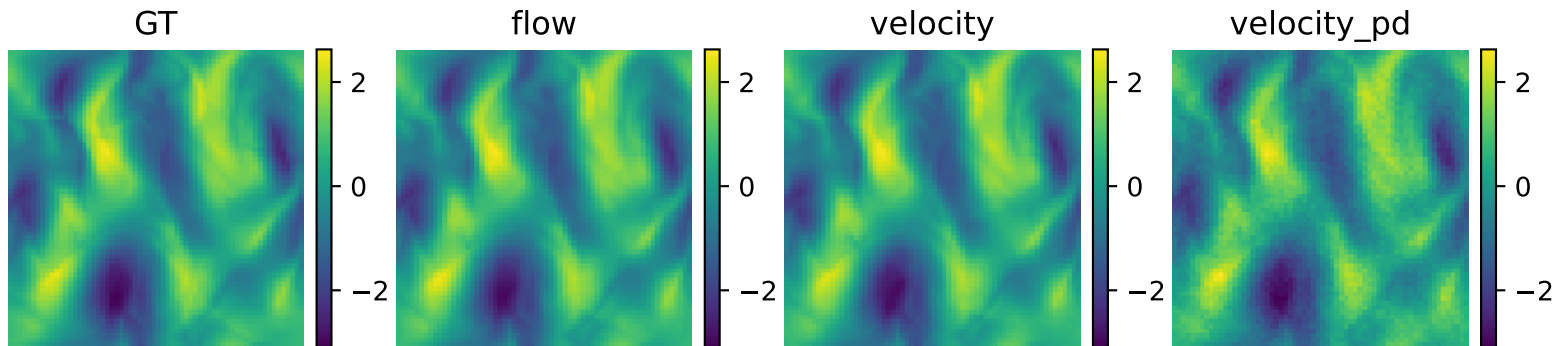
# One-Step Snapshot #0 — ch 0



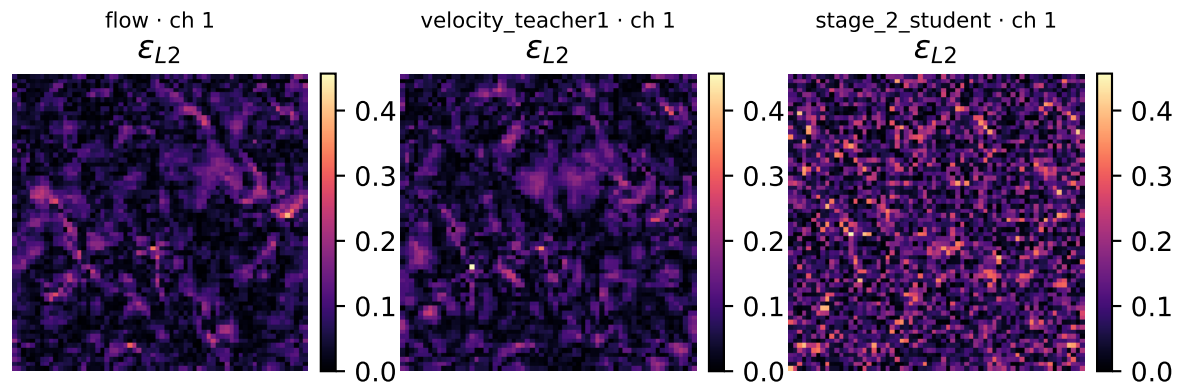
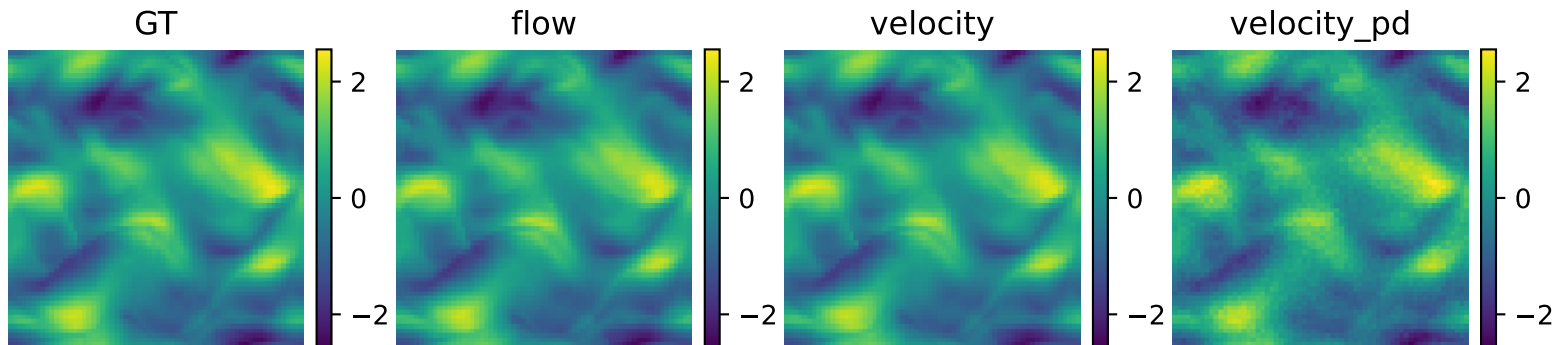
# One-Step Snapshot #0 — ch 1



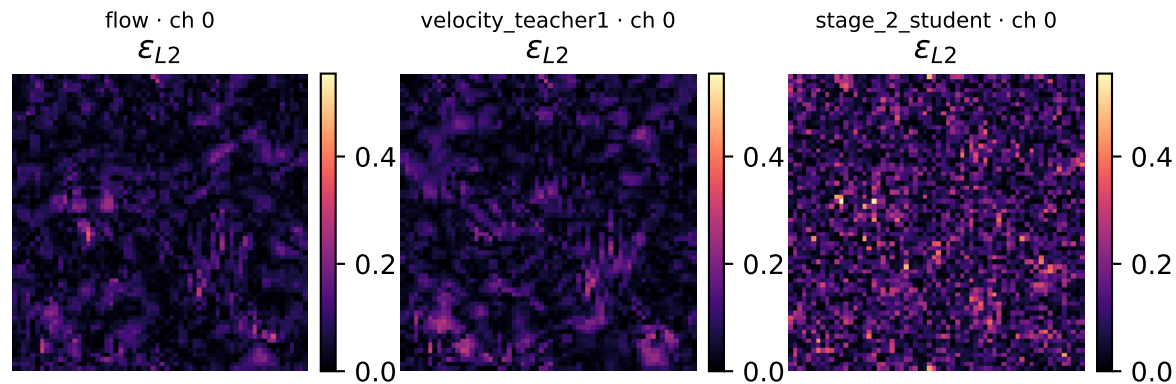
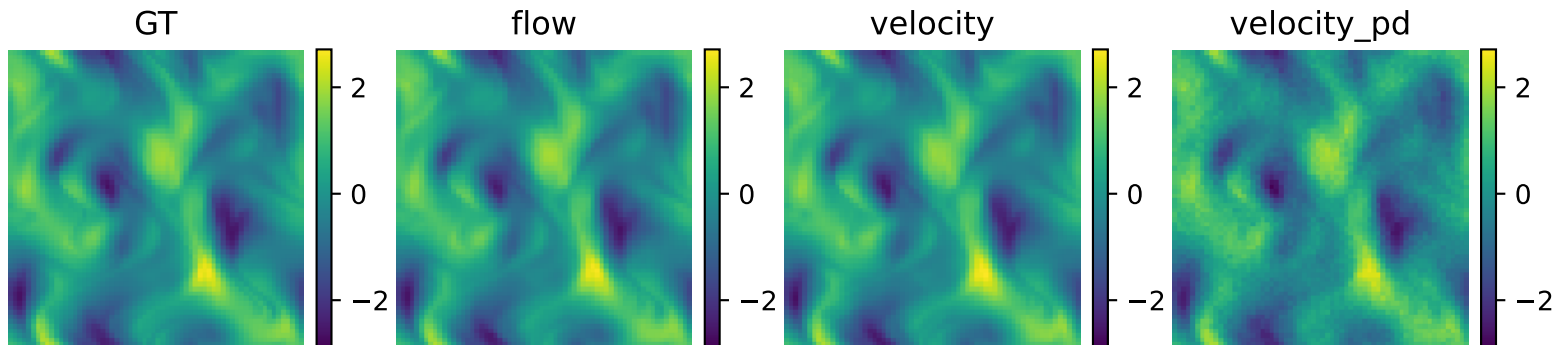
# One-Step Snapshot #1 — ch 0



# One-Step Snapshot #1 — ch 1

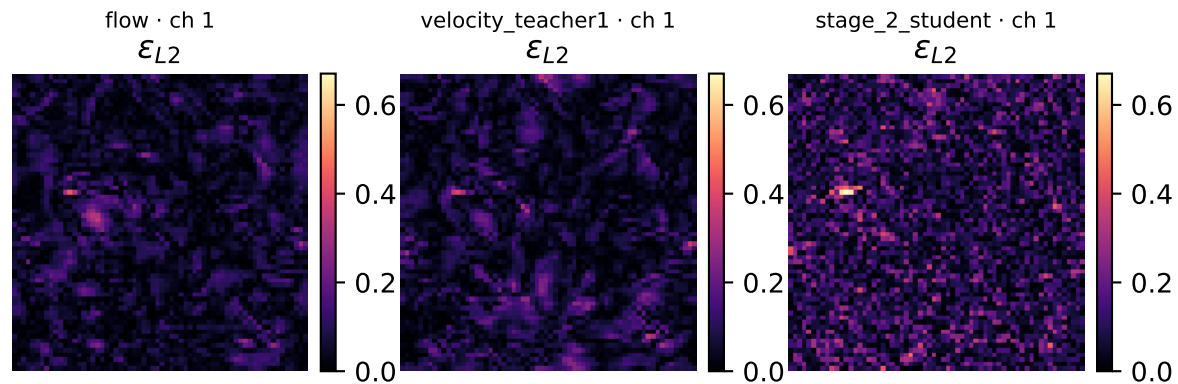
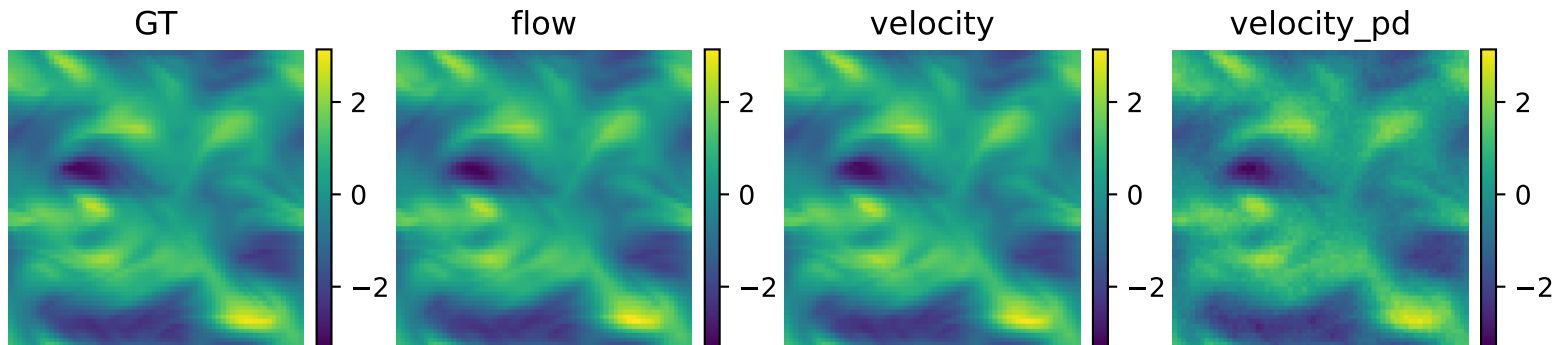


## One-Step Snapshot #2 — ch 0

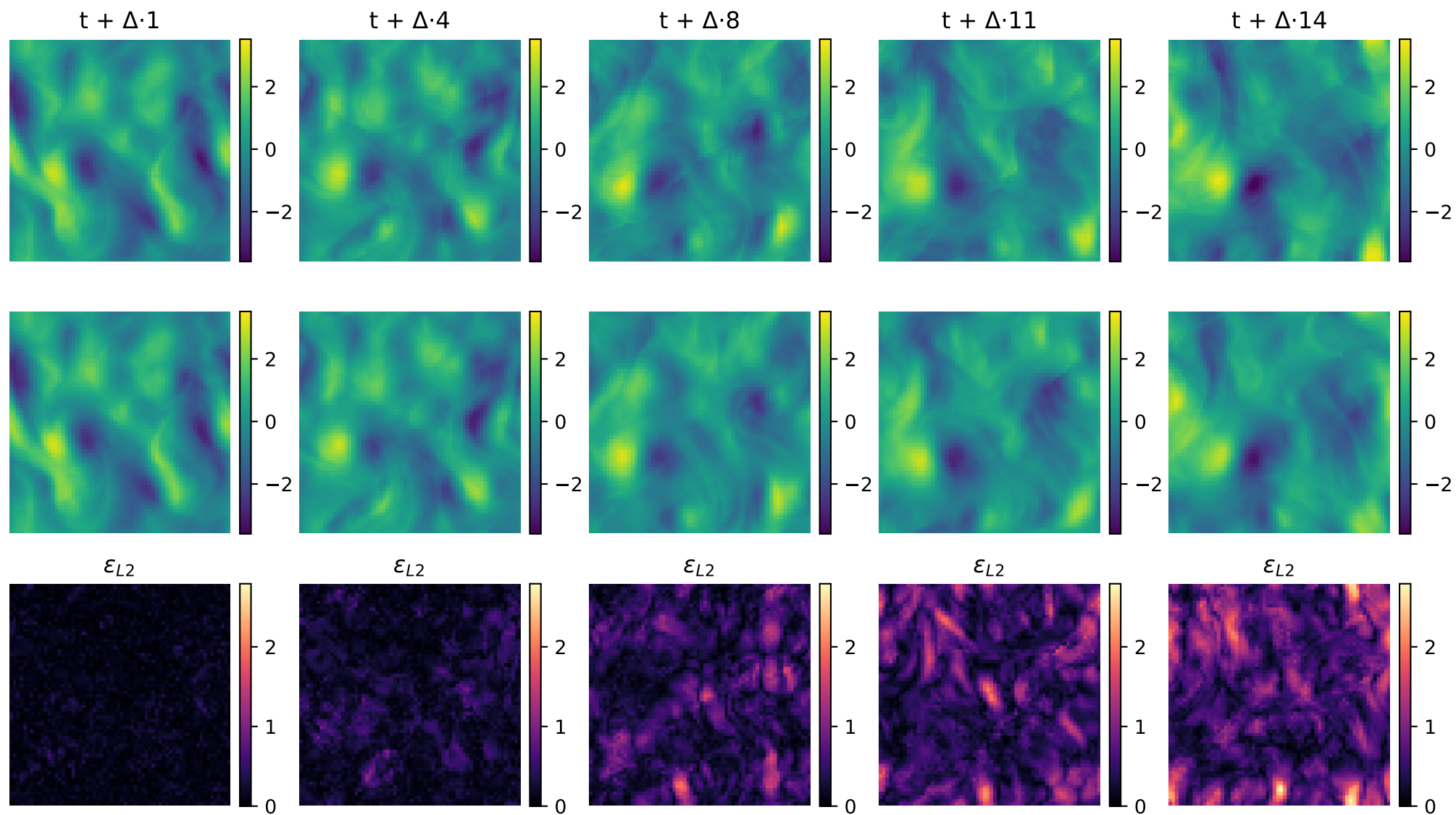




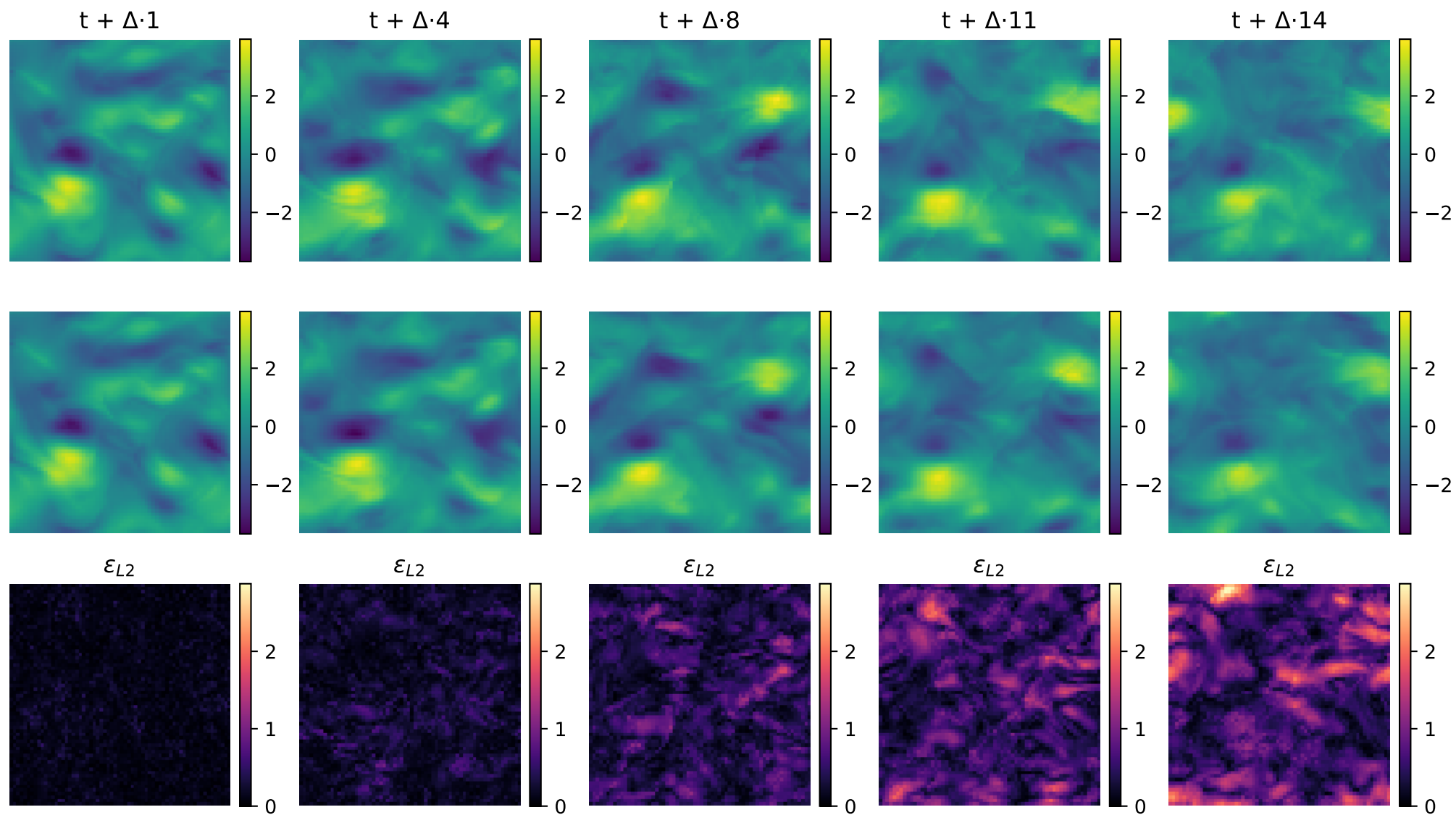
# One-Step Snapshot #2 — ch 1



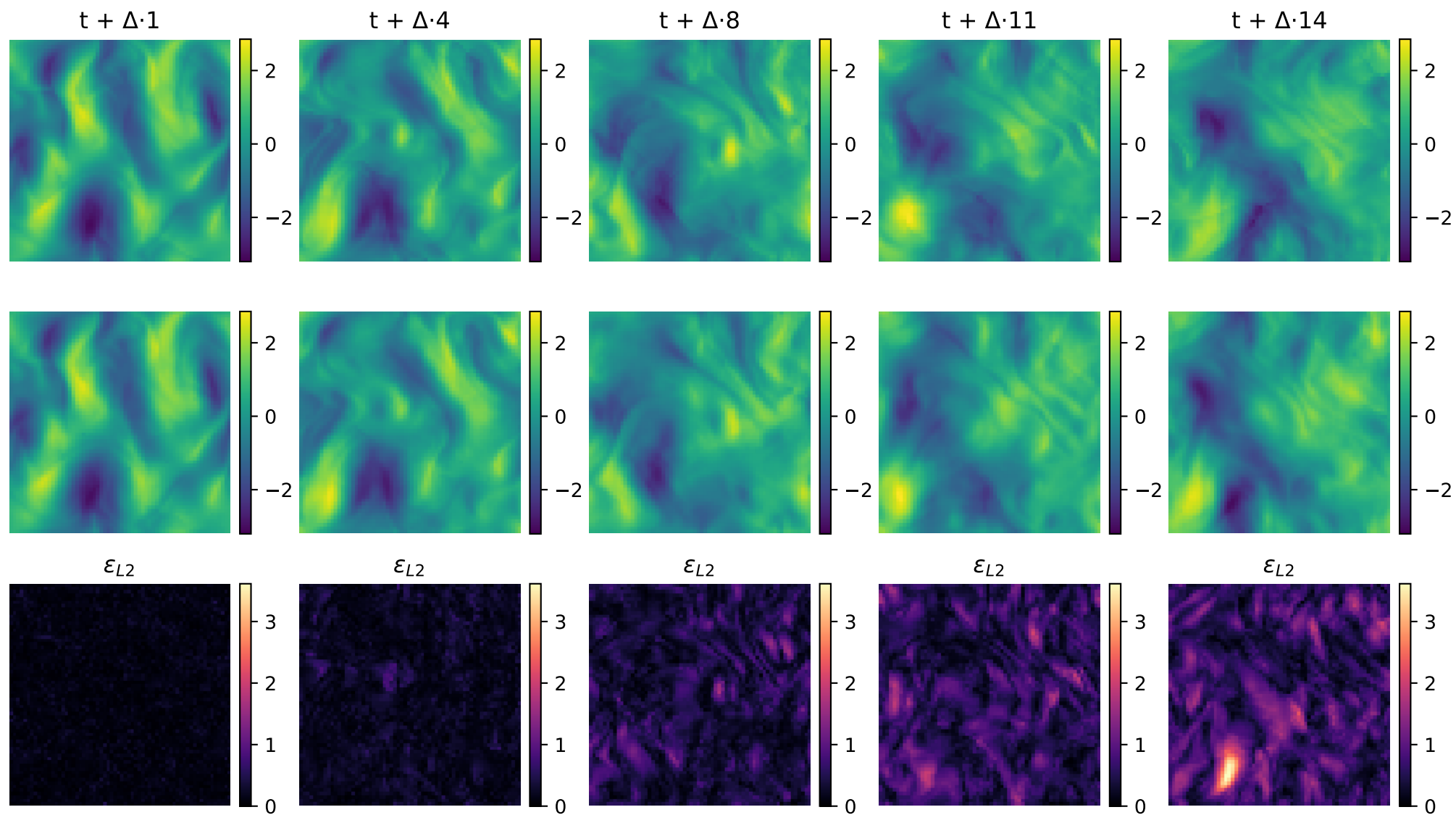
Rollout Snapshot #0 -- ch 0 -- GT (top row) vs flow (middle row)



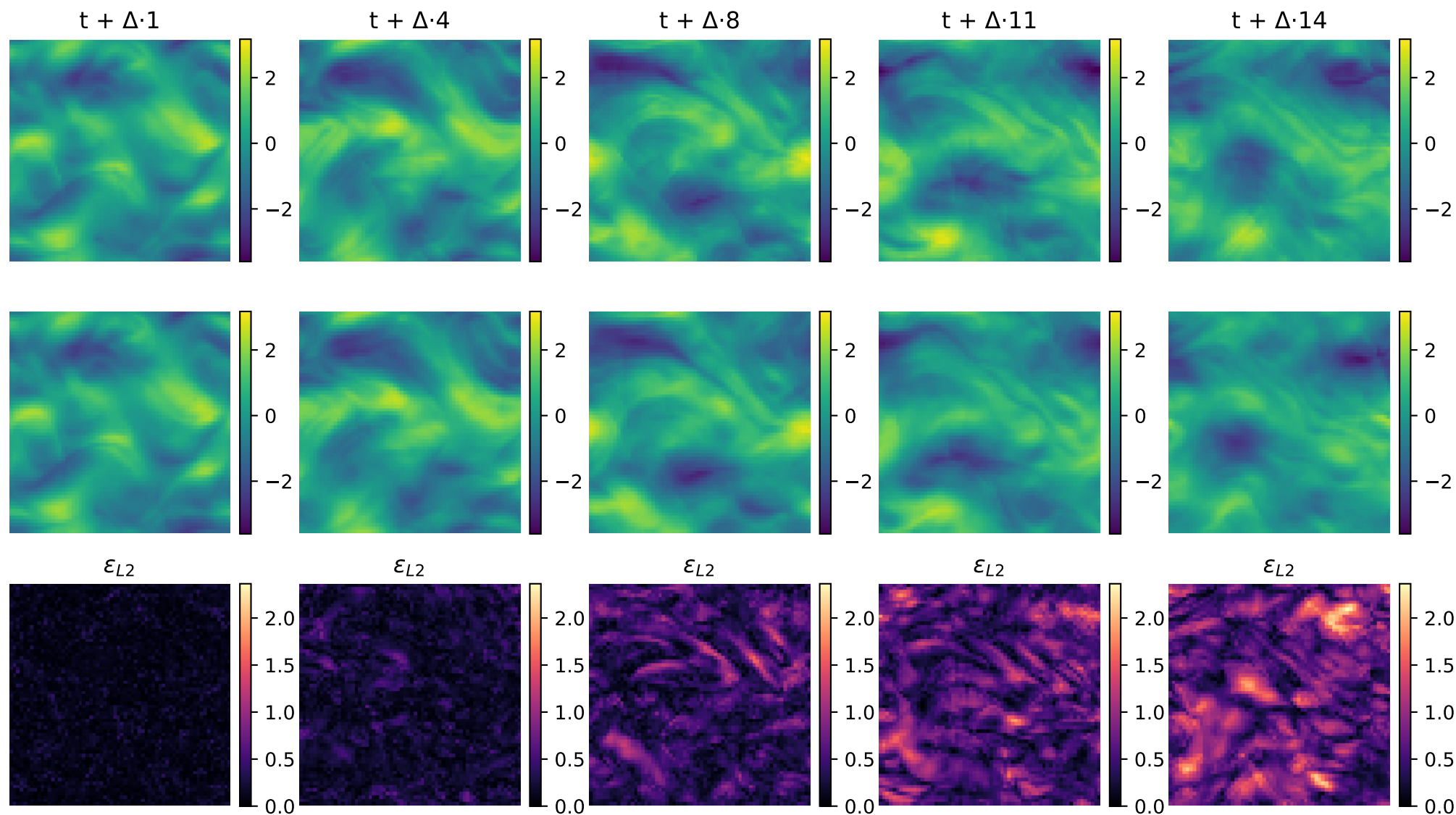
Rollout Snapshot #0 -- ch 1 -- GT (top row) vs flow (middle row)



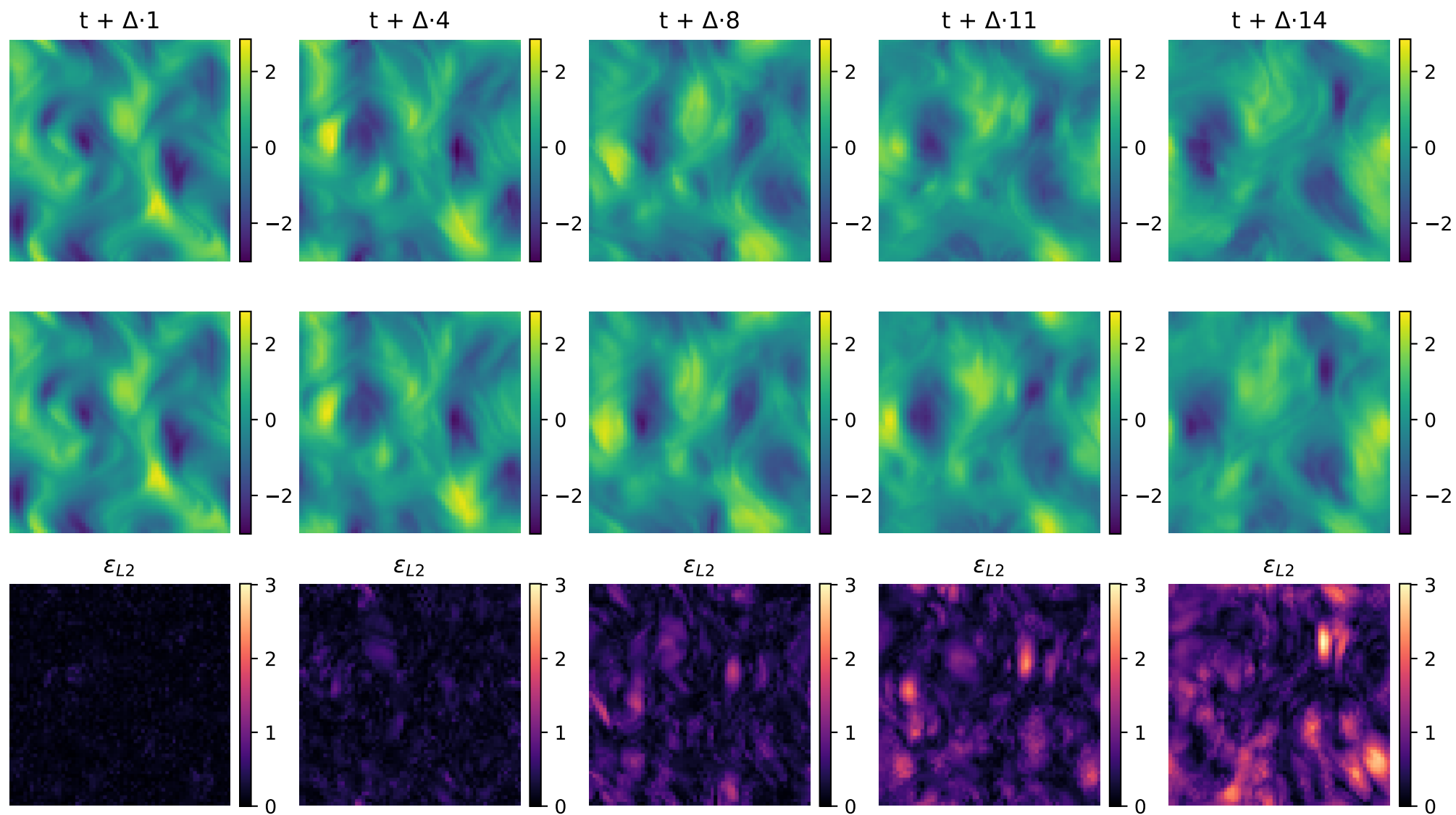
Rollout Snapshot #1 -- ch 0 -- GT (top row) vs flow (middle row)



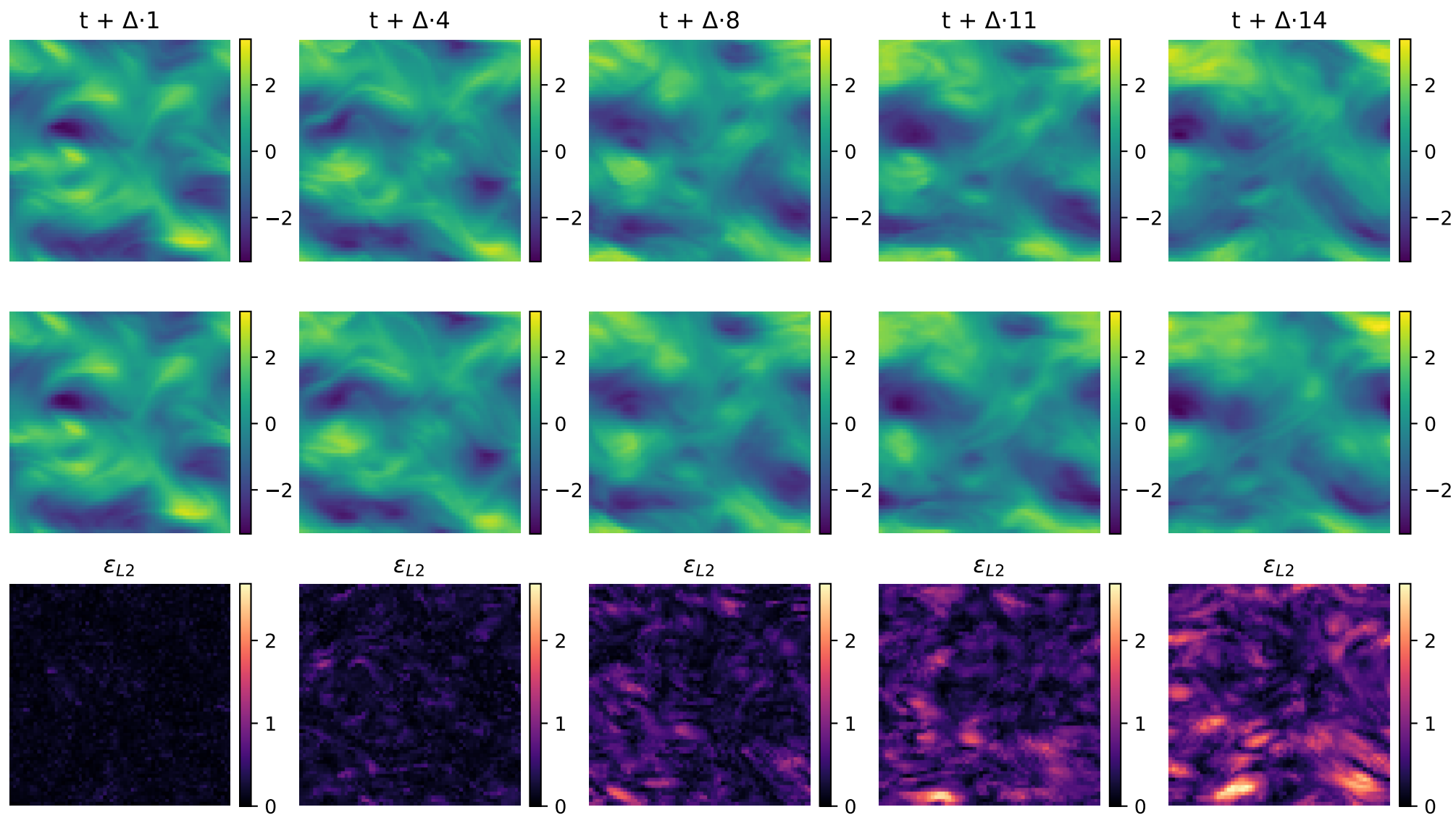
Rollout Snapshot #1 -- ch 1 -- GT (top row) vs flow (middle row)



Rollout Snapshot #2 -- ch 0 -- GT (top row) vs flow (middle row)



Rollout Snapshot #2 -- ch 1 -- GT (top row) vs flow (middle row)



## Evaluation Configuration

```
{
  "common": {
    "device": "cuda",
    "seed": 42,
    "batch_size": 32,
    "samples": 99
  },
  "models": [
    "flow",
    "velocity",
    "velocity_pd"
  ],
  "viz": {
    "enabled": true,
    "n_samples": 3,
    "channels": [
      0,
      1
    ],
    "colorbar": true,
    "rollout_steps": 5
  },
  "one_step": {
    "steps": {
      "velocity": 50,
      "flow": 1,
      "velocity_pd": 6
    },
    "method": {
      "velocity": "midpoint",
      "velocity_pd": "midpoint"
    }
  },
  "rollout": {
    "length": 14,
    "latent_policy": "new_latent",
    "steps": {
      "velocity": 50,
      "flow": 1,
      "velocity_pd": 6
    },
    "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"
}
```