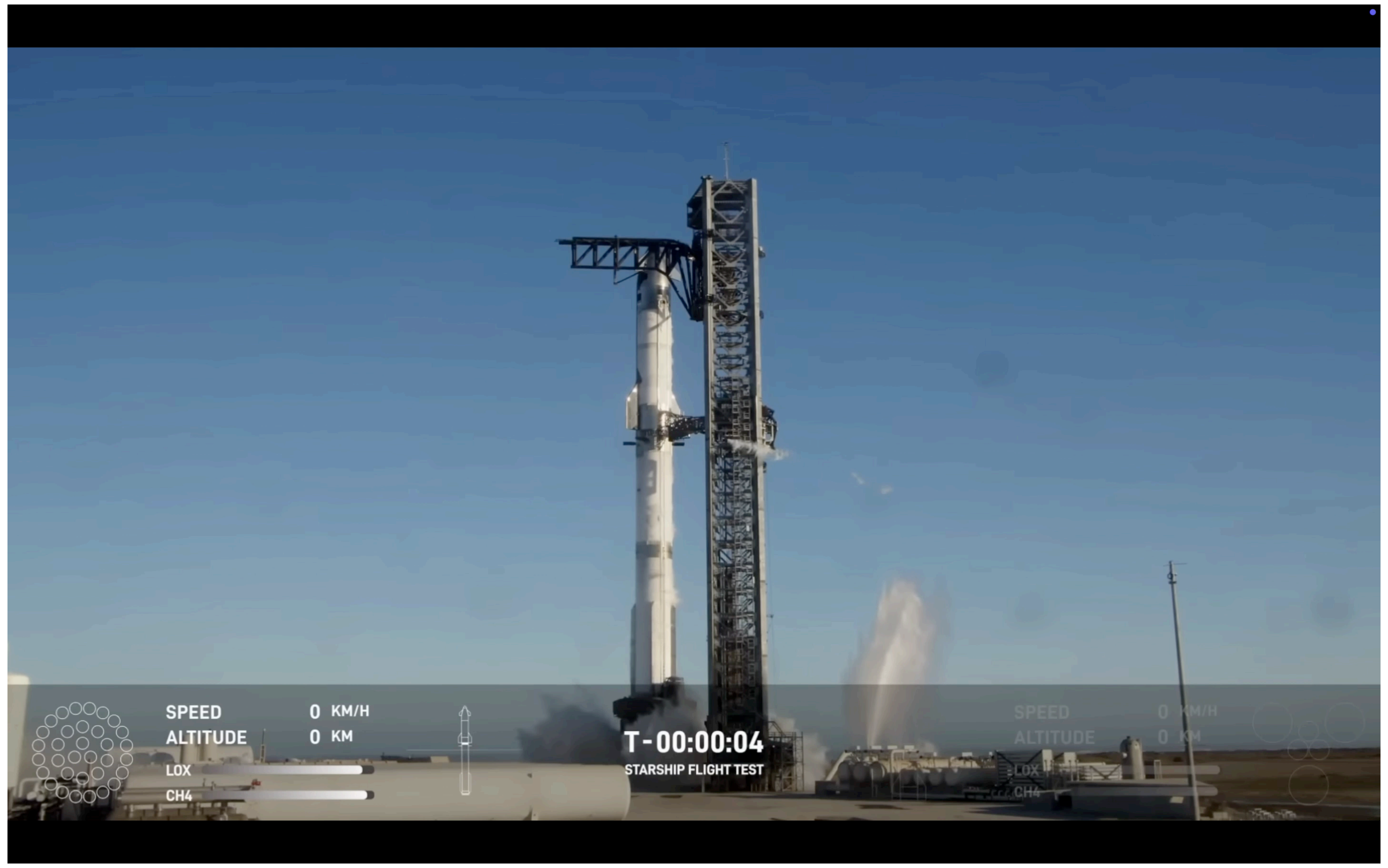
Efficient Physics Informed Dynamic Neural Fluid Fields Reconstruction From Sparse Videos

AIT 2025-02

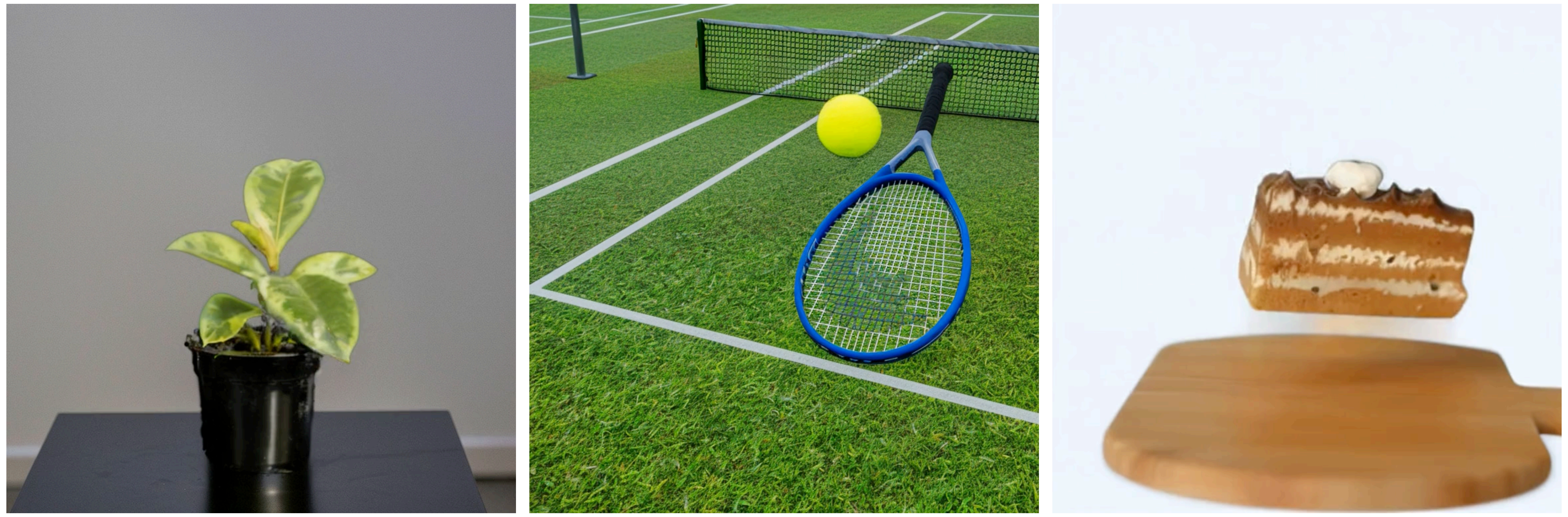
Yangcheng Xiang, Yoshinori Dobashi
Hokkaido University

Background - Physical Dynamics in Videos

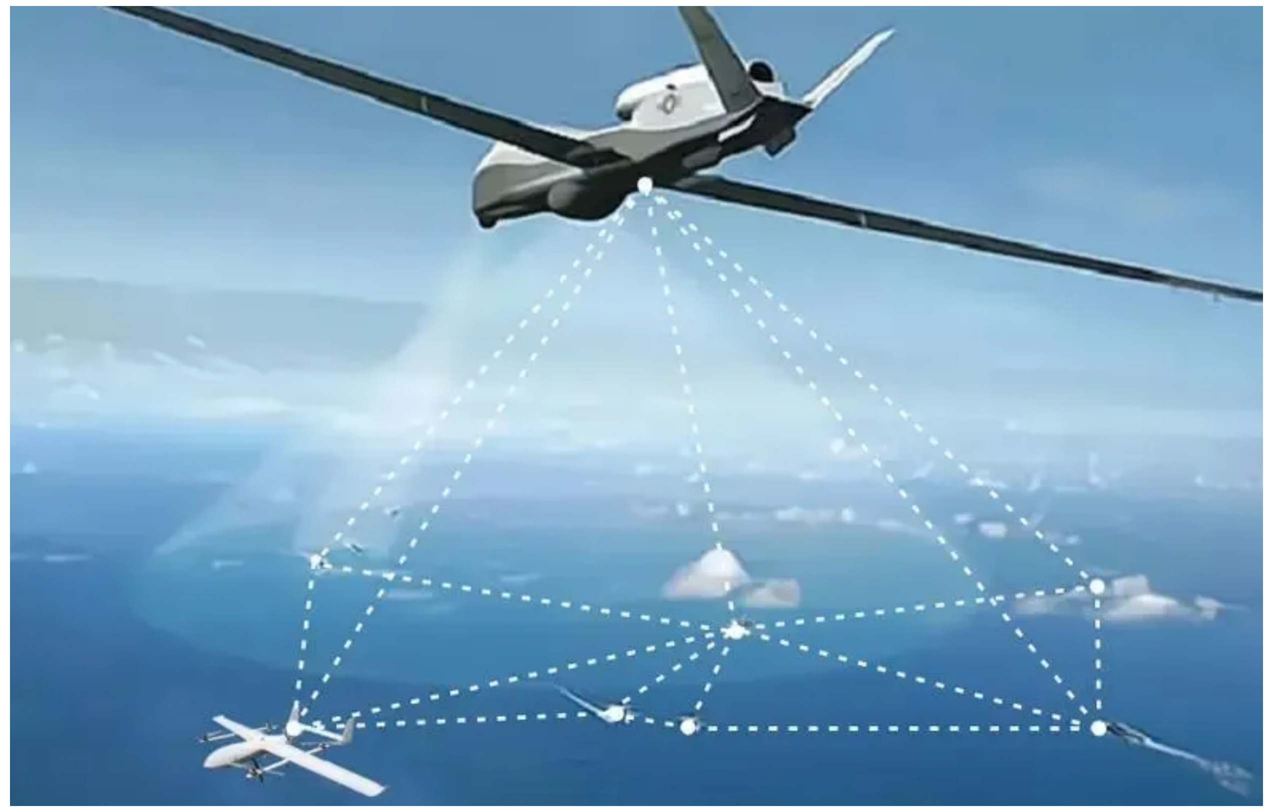
Physical Video Generated by Sora



Starship Launch



PhysMotion: Physics-Grounded Dynamics From a Single Image (Arxiv. 2024.)



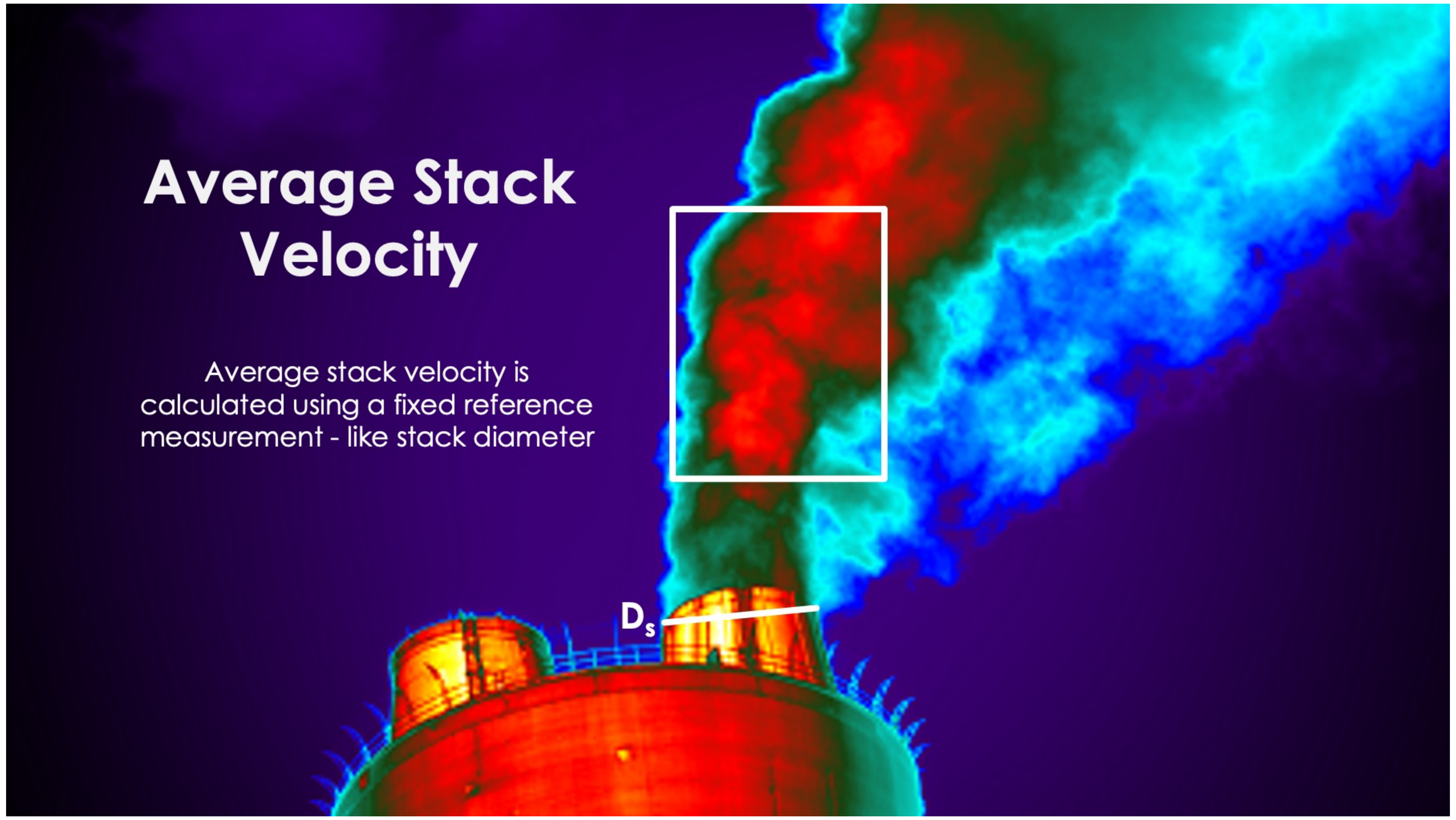
Drone or Robot Swarm

Objective - Estimate Underlying Fluid Motion

Video (View 1)

Video (View 2)

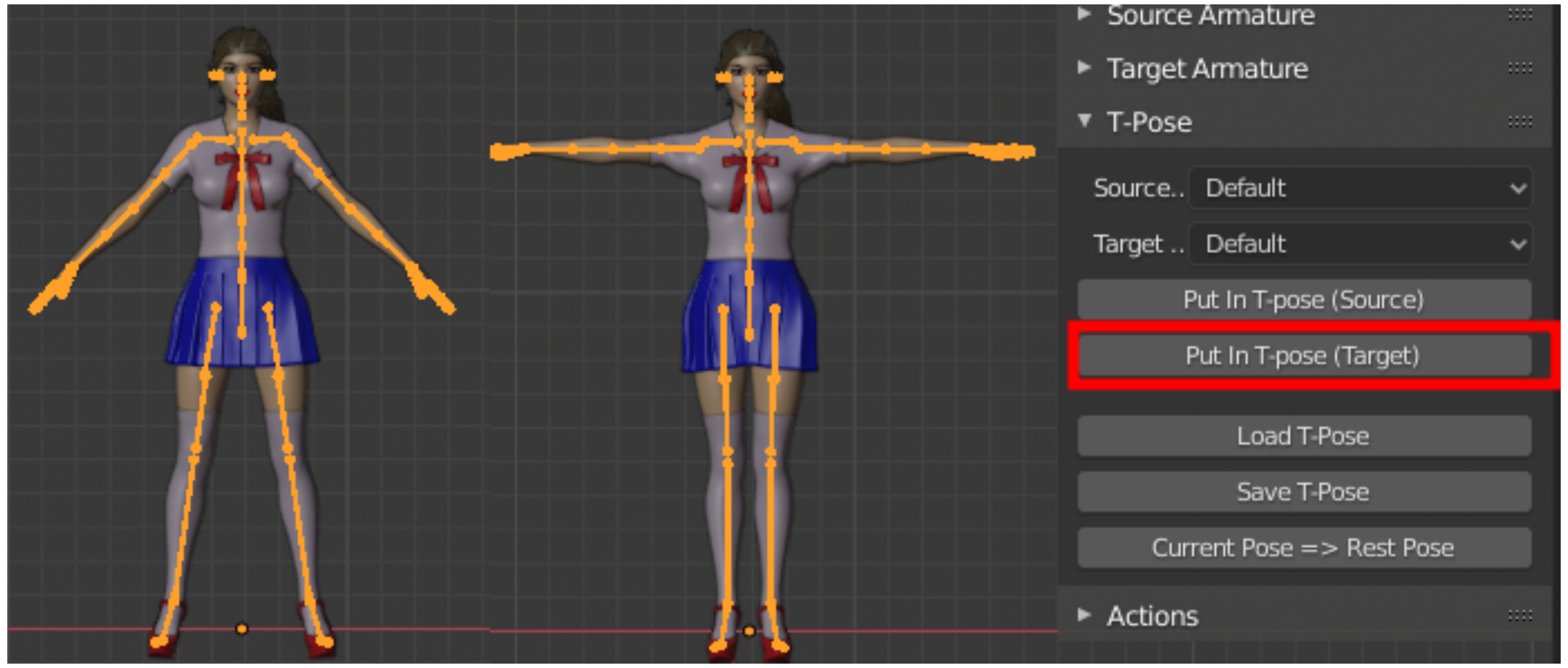
👇 *A Drone's perspective (our objective)* 👇



Correct Physical Dynamics Informations

density, velocity, etc.

Problem Statement



Articulations Dynamics

Rigid Body Dynamics

Dynamics with **LOW** Degrees of freedom

Opaque

Fixed shape

Simple

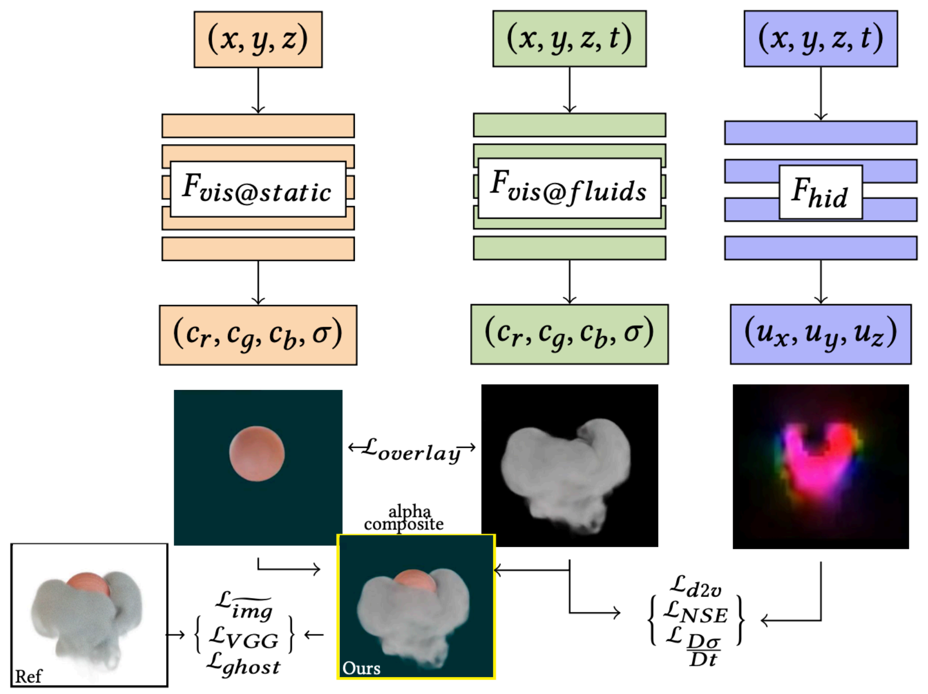
Rocket Plume Dynamics

Dynamics with **HIGH** Degrees of freedom

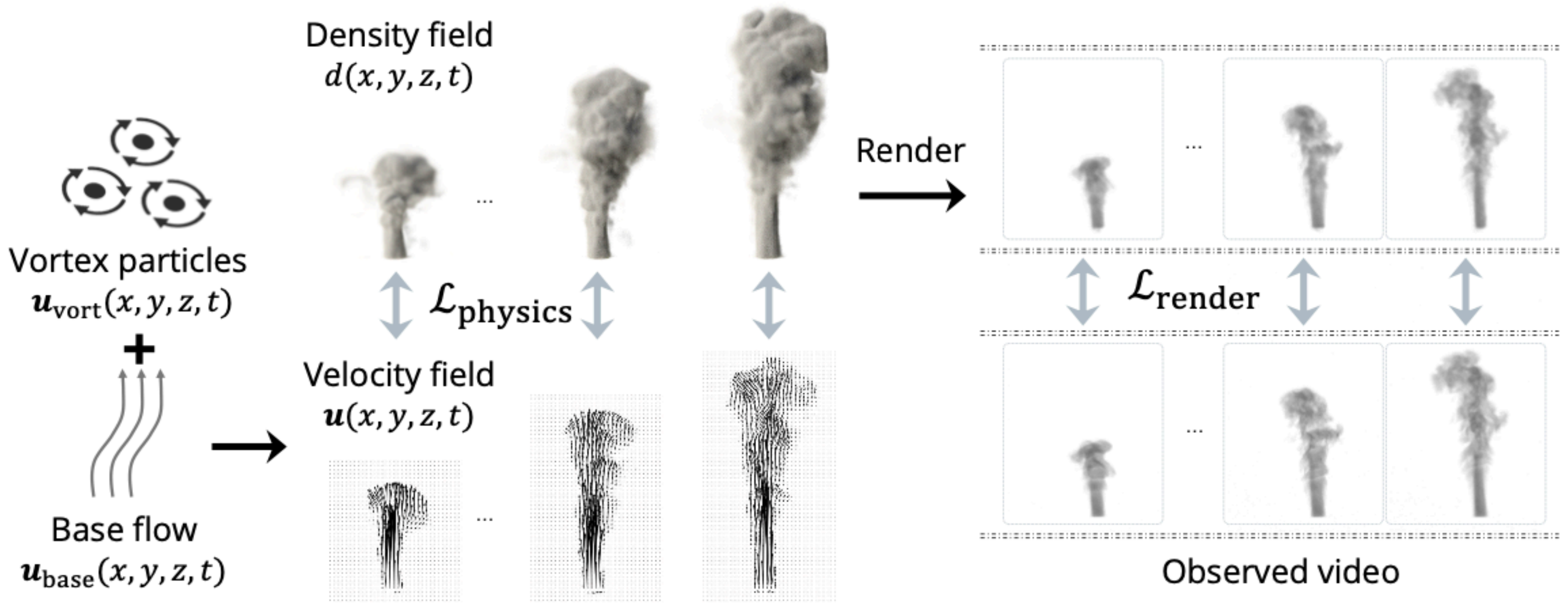
Semi-transparent Shapeless Turbulent

Related Works - Neural Smoke Reconstruction

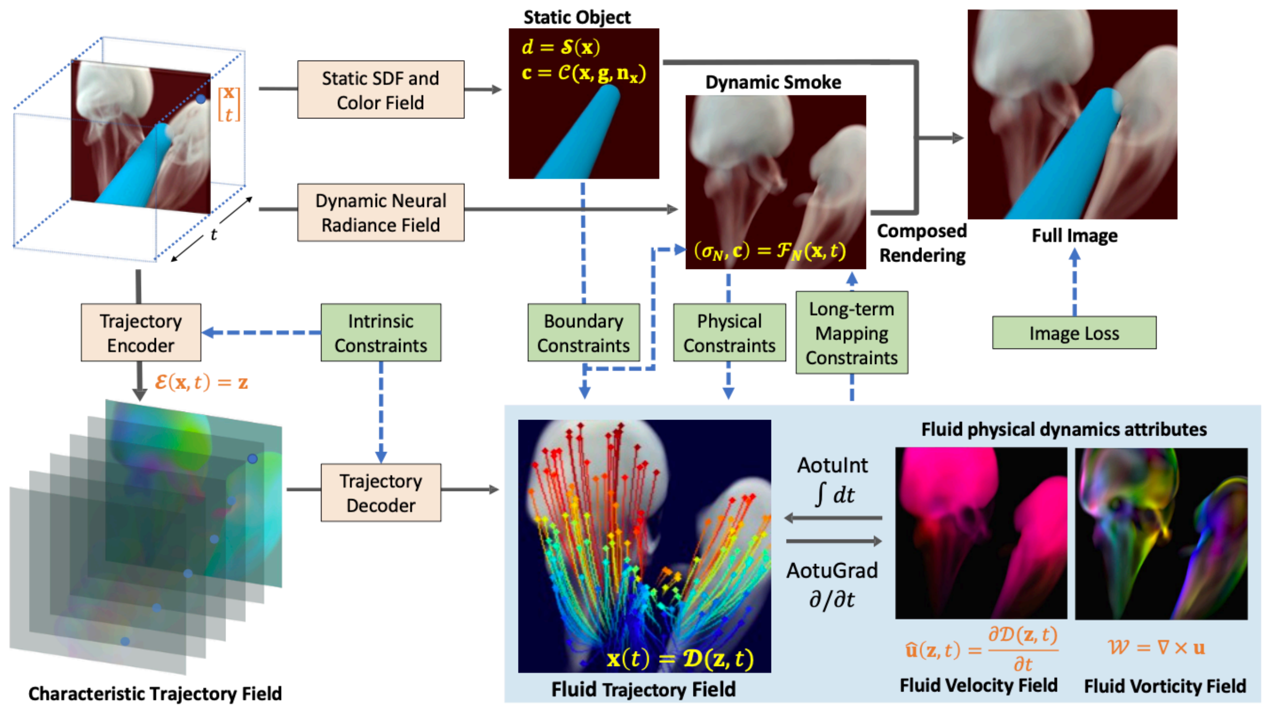
- PINF [M. Chu+ 2022]: Physics Informed Neural Field for smoke reconstruction with sparse data.
- HyFluid [H. Yu+ 2023]: Inferring Hybrid Neural Fluid Fields from Videos
- PICT [Y. Wang+ 2024]: Physics-Informed Learning of Characteristic Trajectories for Smoke Reconstruction



Overview of PINF

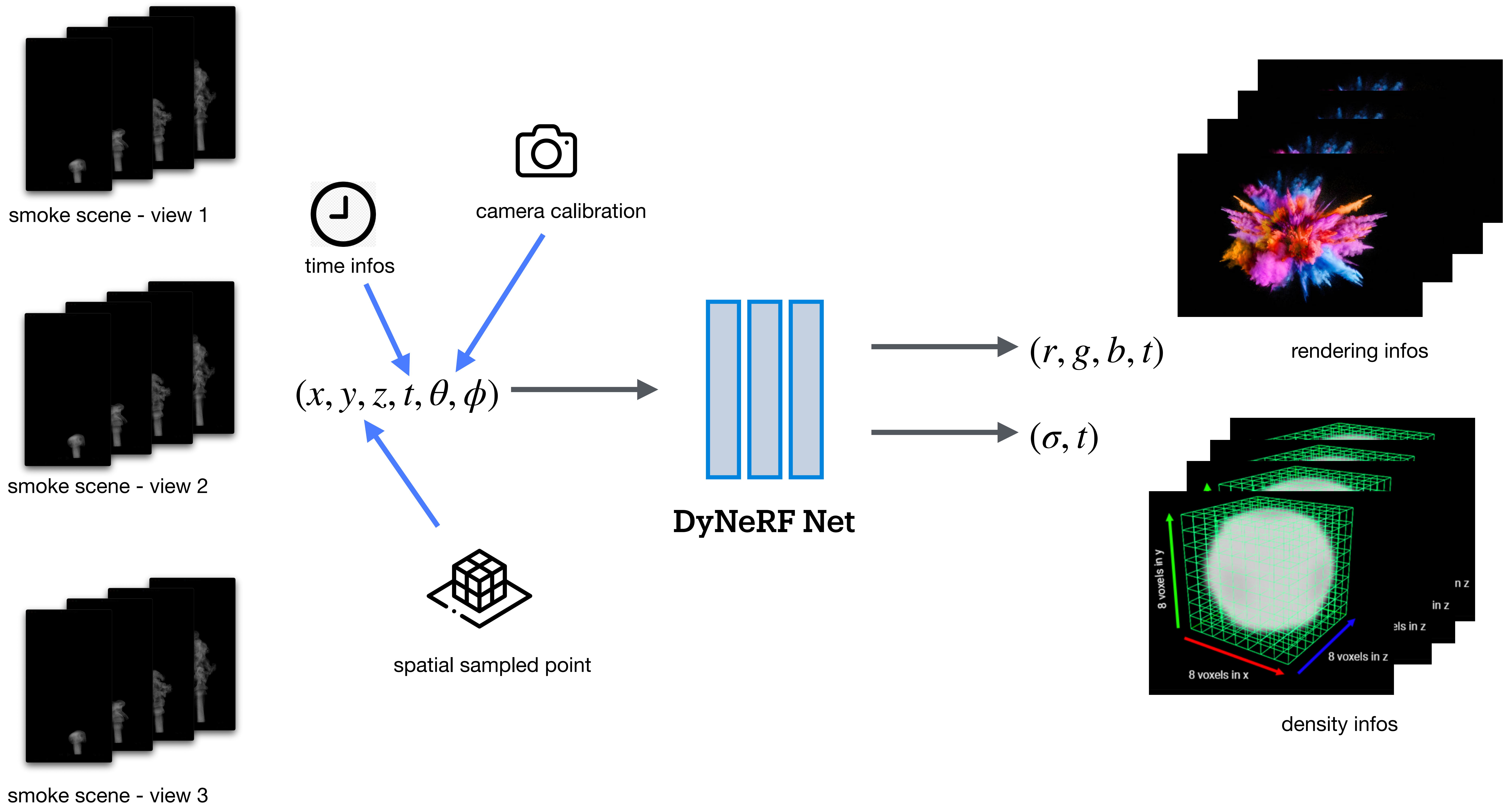


Overview of HyFluid



Overview of PICT

Dynamic Neural Radiance Field（D-NeRF）



Physical Laws for Fluid Dynamics

Navier–Stokes Equation:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{f}$$

(Govern the motion of fluid)

Continuity Equation:

$$\nabla \cdot \mathbf{u} = 0$$

(Maintain mass conservation of fluid)

Density transport equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0.$$

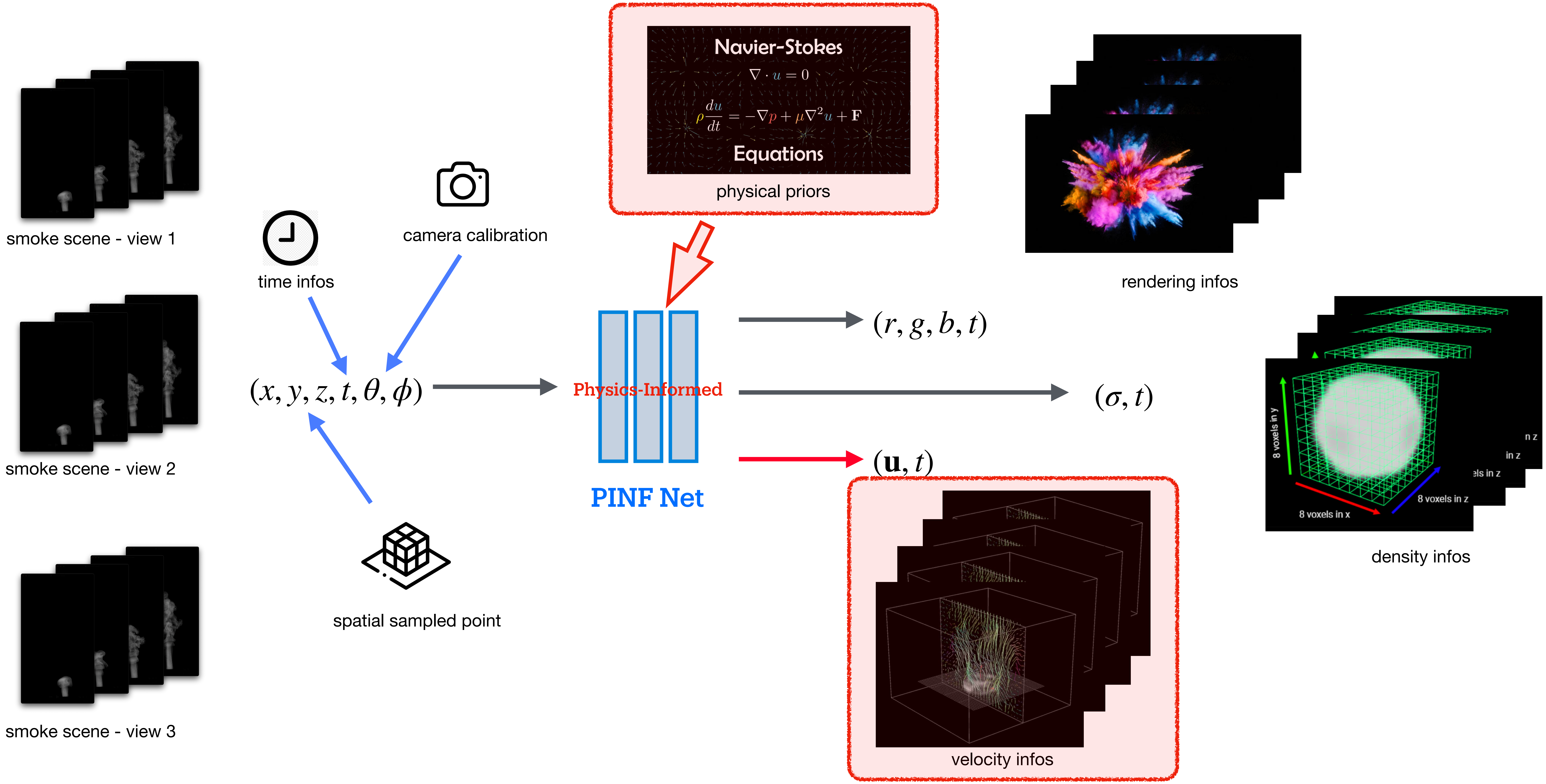
(Govern the density transport of fluid)

Clouds

Fire

Plume

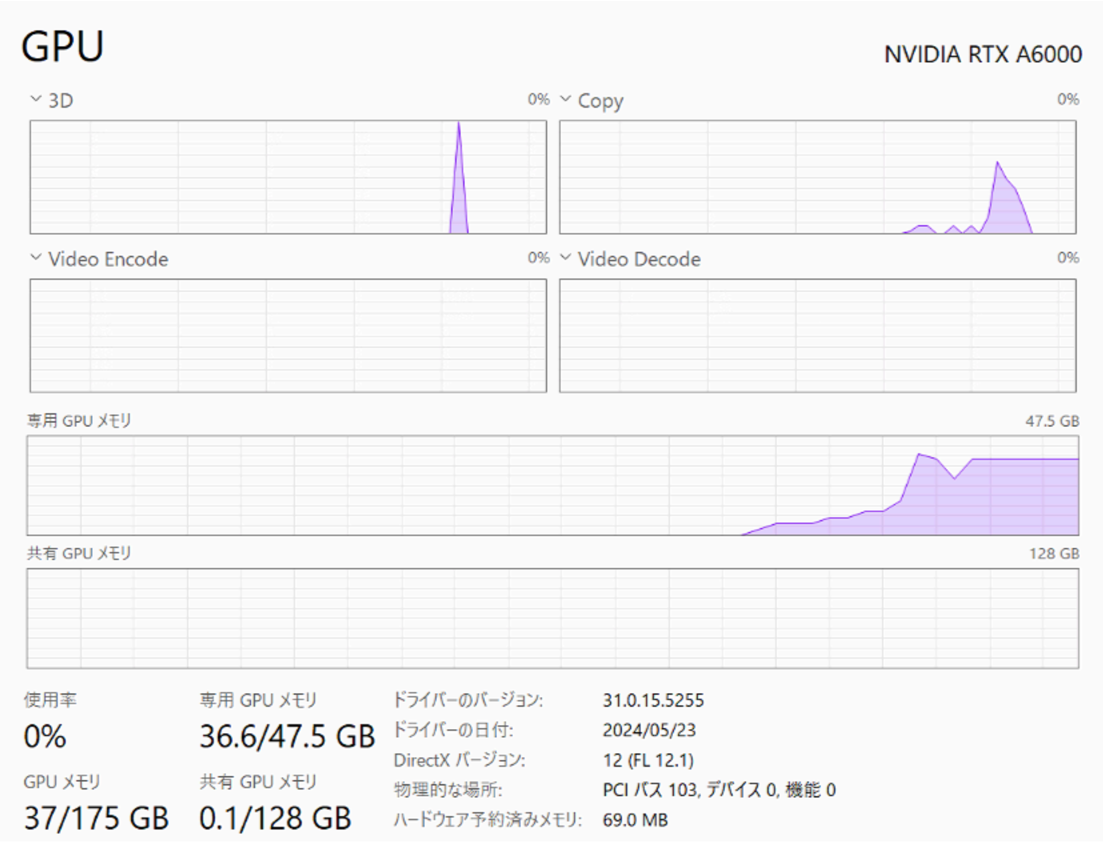
Physics-Informed Neural Fluid Reconstruction (PINF)



Problems on Existing Methods

- PINF [M. Chu+ 2022]:
  - Long training time
  - Blurred reconstructed result, few vorticity details.
  - Inaccurate reconstructed velocity field.
- HyFluid [H. Yu+ 2023]:
  - Extremely memory- and time-consuming during training and inferring.
  - Hard to reconstruct long-duration motion.
- PICT [Y. Wang+ 2024]:
  - Excessive complexity, leading to long training time.

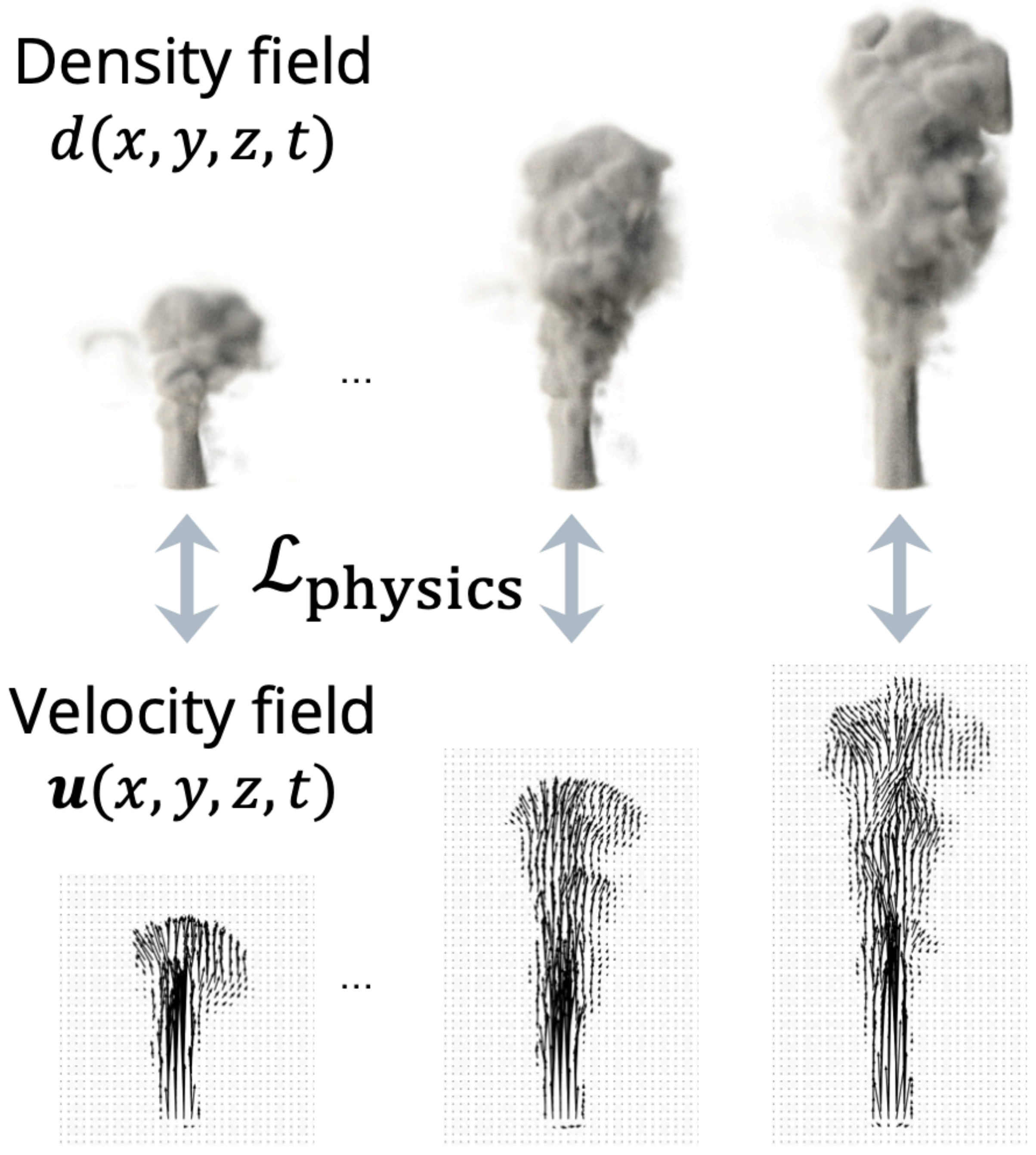
left: ground truth image
right: reconstructed density field by PINF



GPU memory requirement > 36 GB by HyFluid

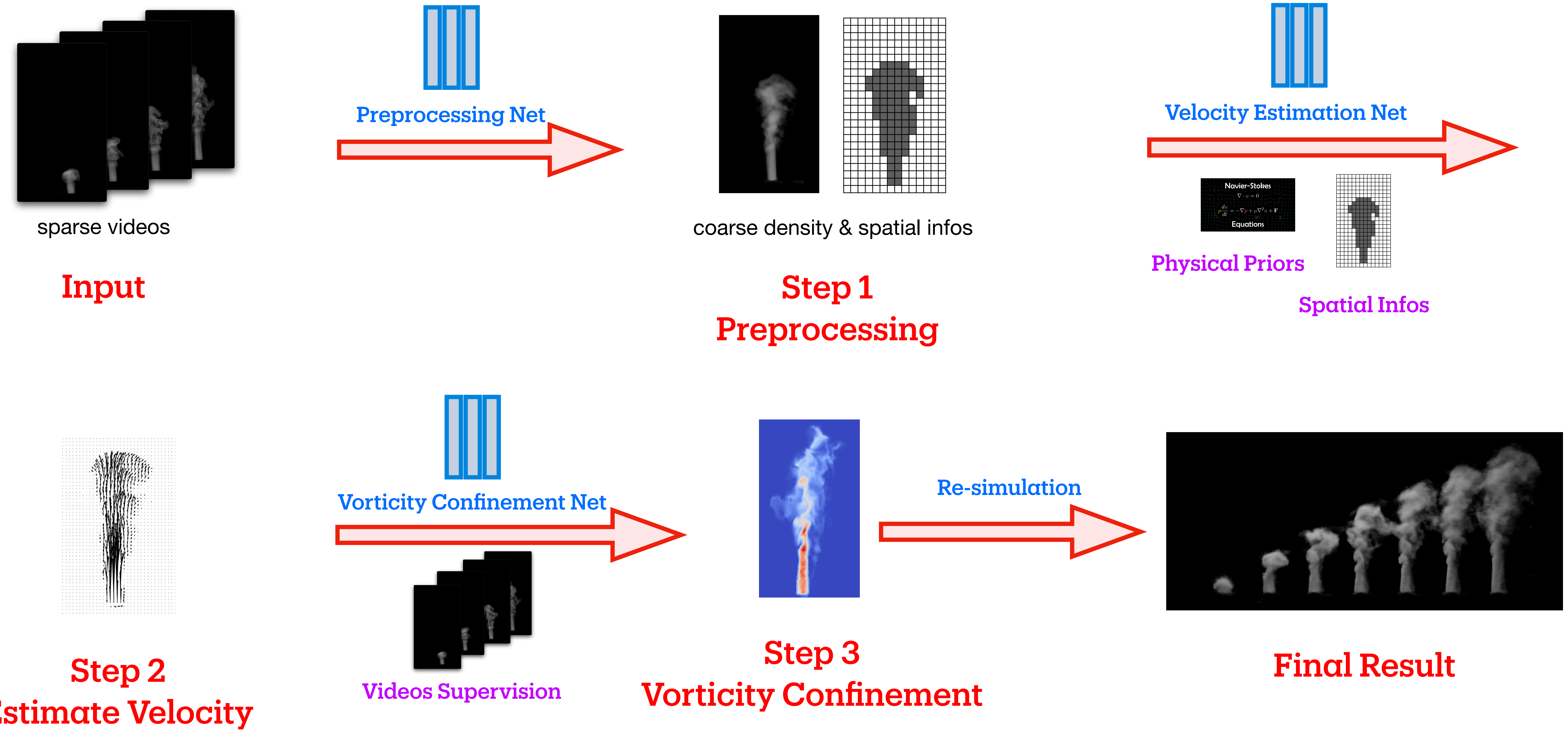
Our Goal

1. **Reconstruction Quality:** more accurate result by sparse videos
   - plausible density field σ
   - reasonable velocity field $\mathbf{u}$
2. **Reconstruction Speed:**
   - faster training & inference time
3. **Long-term Velocity Stability:**
   - support re-simulation for a longer duration.

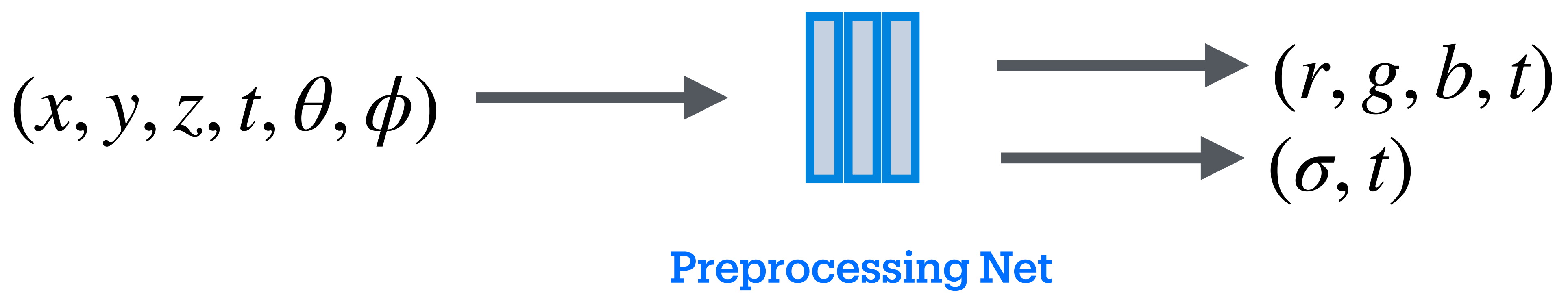


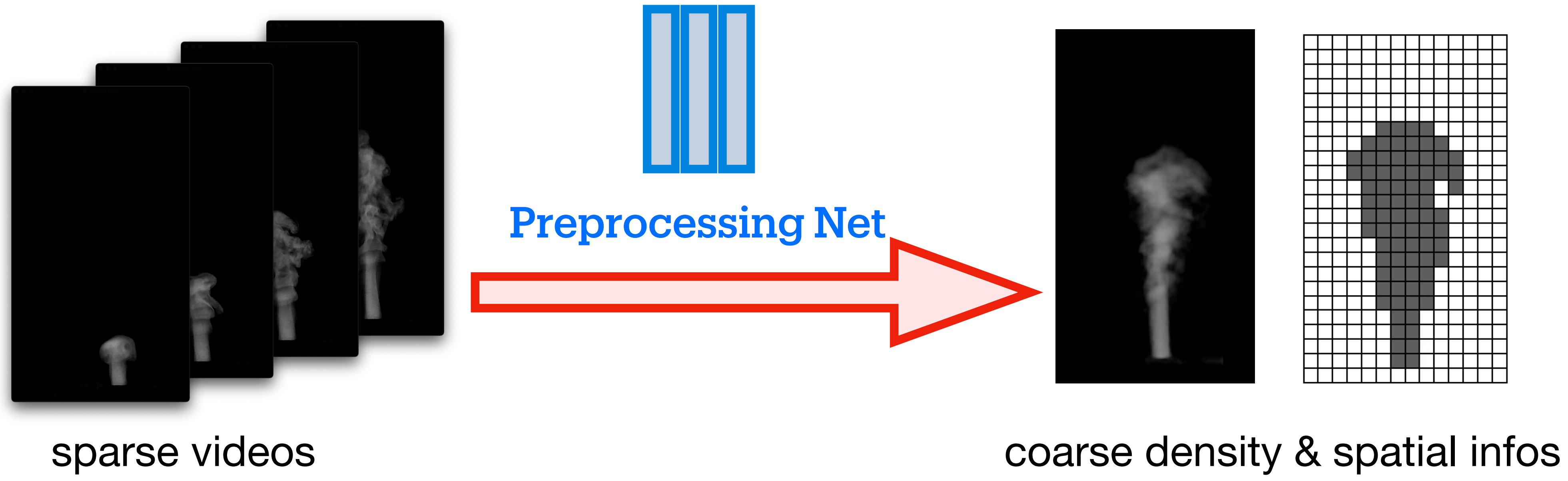
Reconstructed physical fields

Our Method - Overview

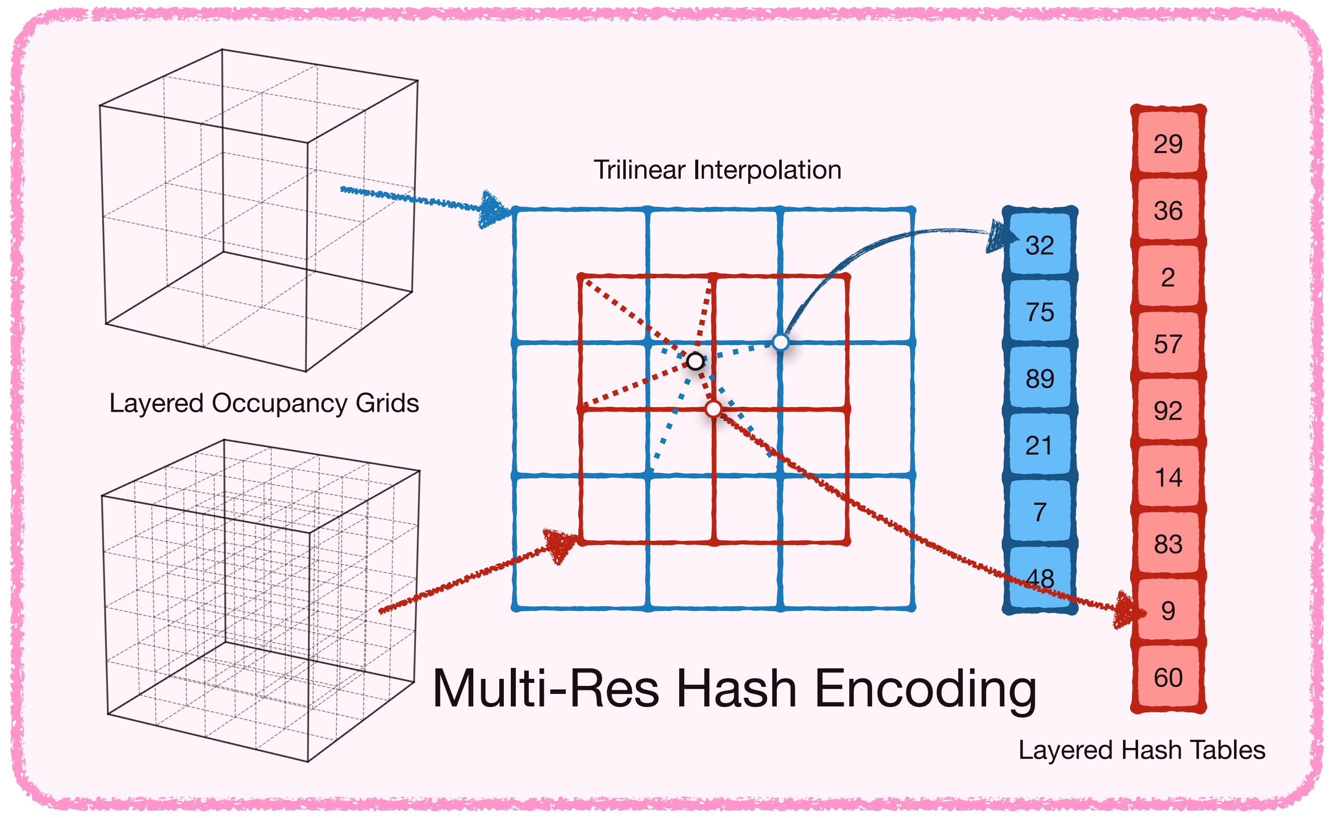


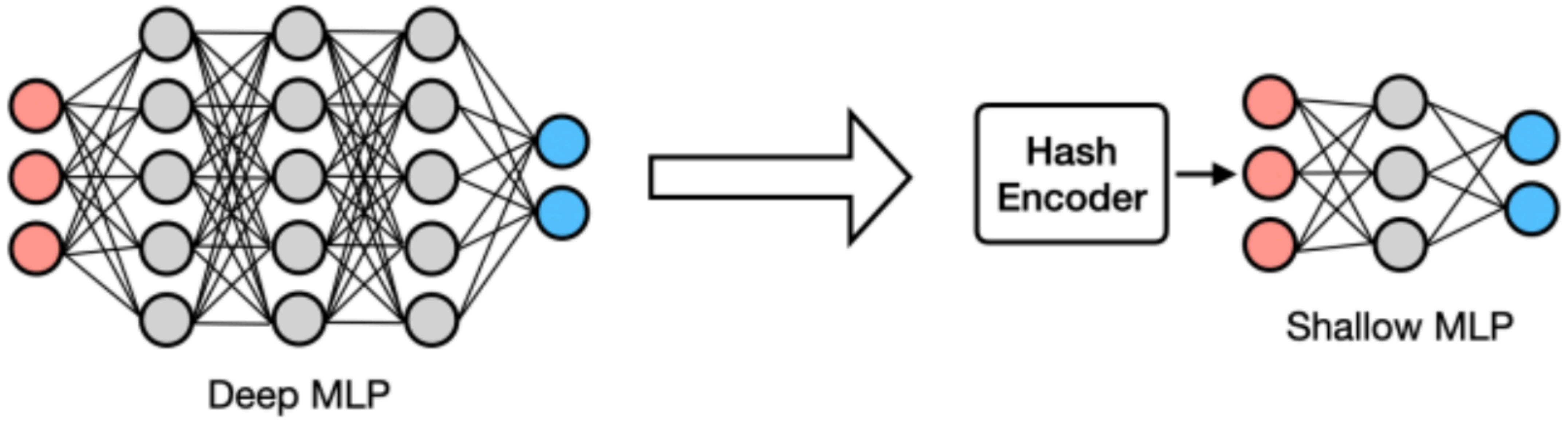
Step 1 - Fast Preprocessing



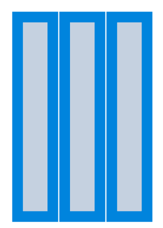


- Volume Rendering Equation
  - $$C(r) = \int_{t_n}^{t_f} T(t)\, \sigma\big(r(t)\big)\, c\big(r(t), d\big)\, dt$$
- RGB Loss (basic loss function for NeRF model)
  - $$\mathscr{L}_{\text{render}} = \mathscr{E}_{o,d}\big[\, C_{\text{render}}(o,d) \; - \; C_{\text{image}}(o,d)\big]^2$$
- Dynamic Multi-Res Hash Encoding to reduce MLP size (from HyFluid)
  - $$h(\mathbf{v}) = \left(\bigoplus_{i=1}^{d} v_i \pi_i \right) \bmod T, \quad \mathbf{v} = (x, y, z, t)\,.$$

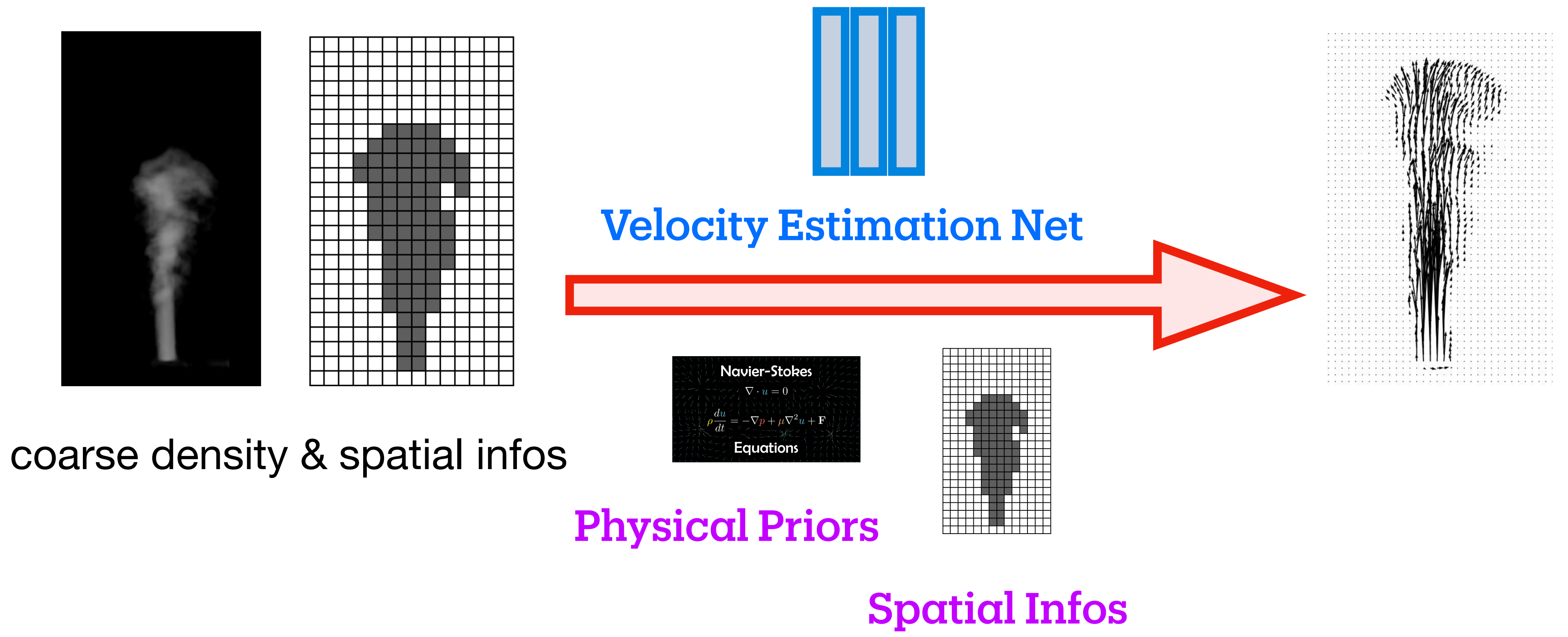




Step 2 - Velocity Estimation

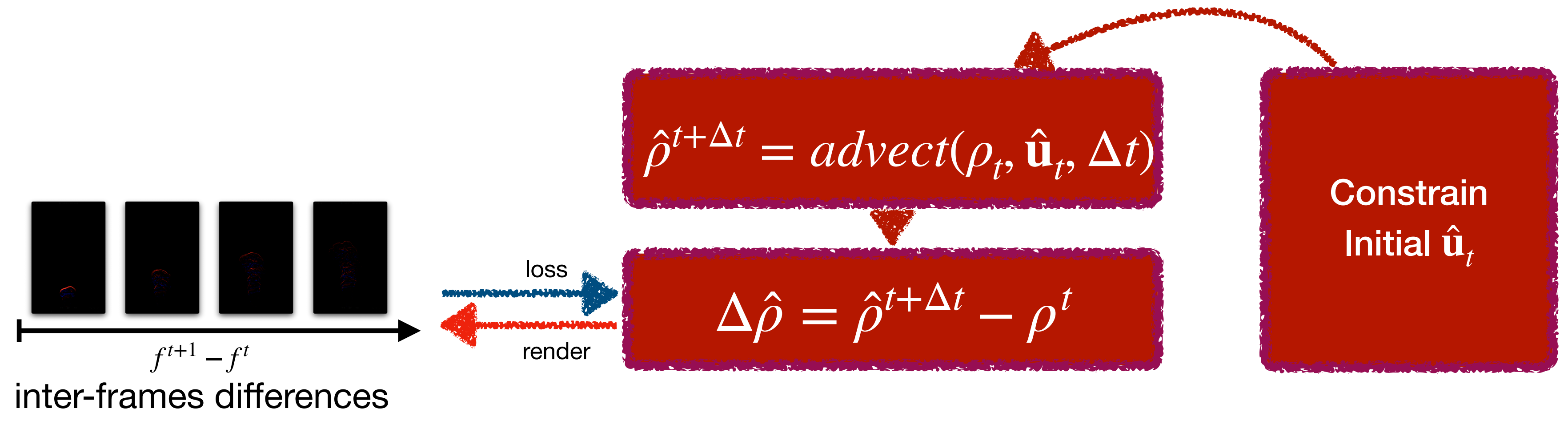
$(x, y, z, t) \longrightarrow$  $\longrightarrow (\mathbf{u}_{base}, t)$

Velocity Estimation Net



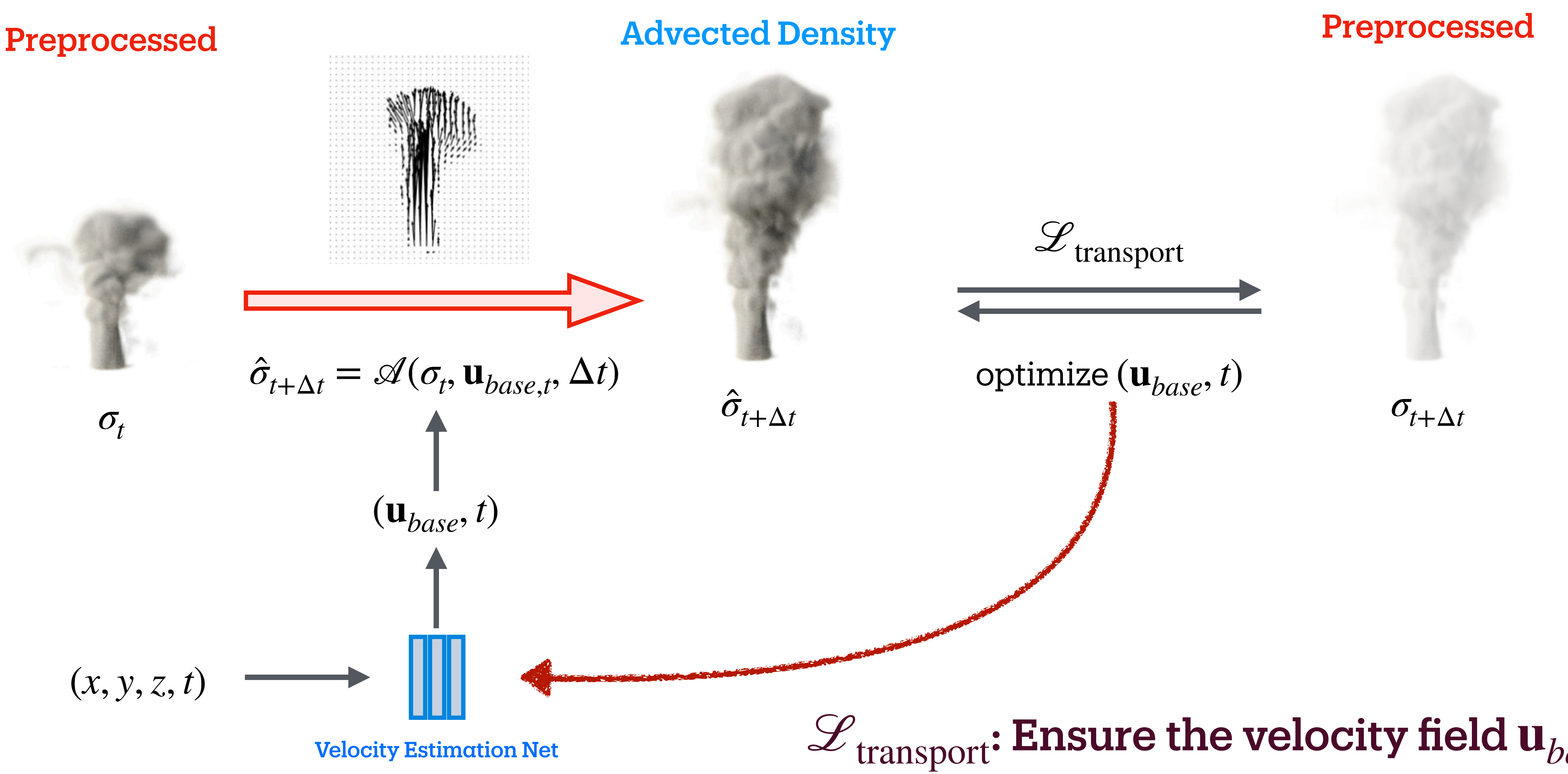
Velocity Estimation Network Structure

- Density Transport Loss (with spatial filter by occupancy grid O)
  - $\mathscr{L}_{\text{transport}}\left(x_{\text{sampled}}\right)$
- Navier-Stokes Equations Loss (from PINF)
  - $\mathscr{L}_{\text{NSE}} = \mathscr{L}_{\text{Navier-Stokes}} + \mathscr{L}_{\text{Div}}$
- Inter-frame Rendering Difference Loss (enhancing long-term physical consistency)
  - $\mathscr{L}_{\text{render}, u} = \mathscr{L}_{\text{advection}} + \lambda_{\text{diff}} \mathscr{L}_{\text{IRD}}$
- Final Velocity Loss
  - $\mathscr{L}_{\textbf{velocity}} = \mathscr{L}_{\text{transport}} + \mathscr{L}_{\text{NSE}} + \mathscr{L}_{\text{render}, u}$



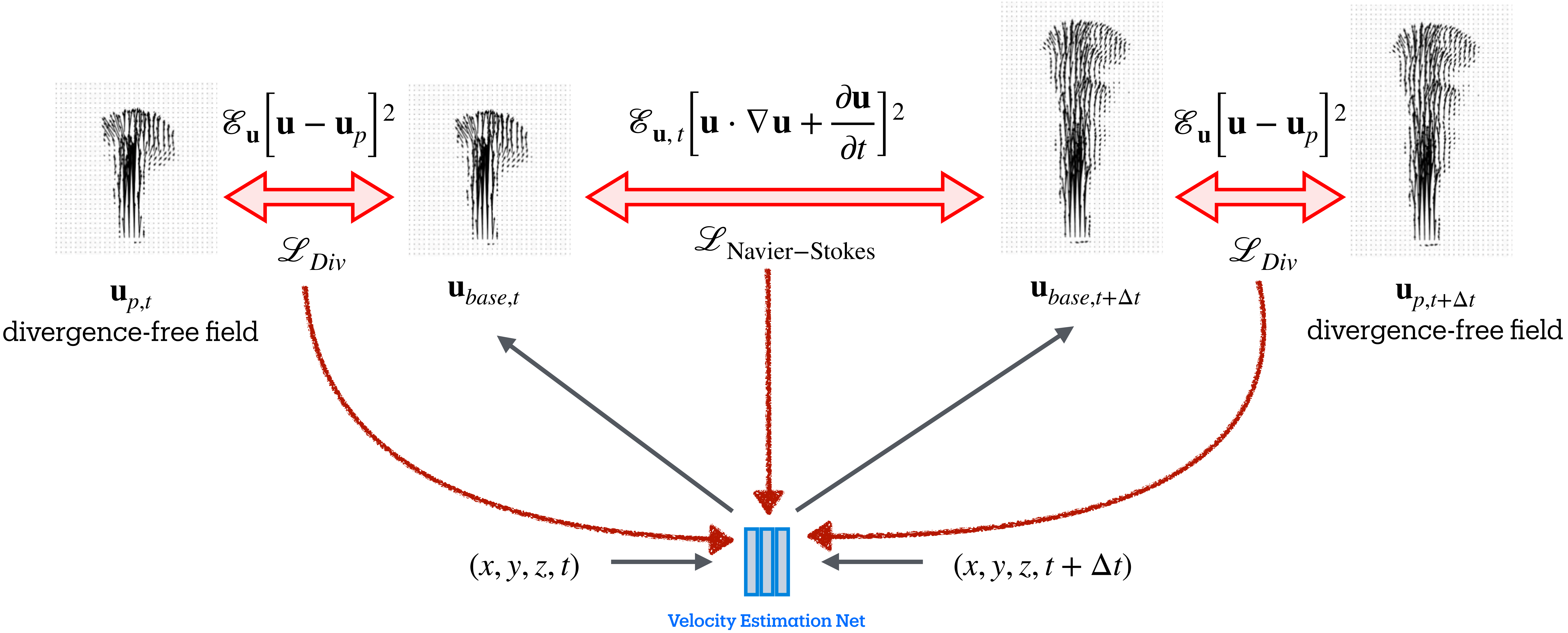
Inter-frame Rendering Difference Loss

Step 2 - Velocity Estimation (Density Transport Loss)



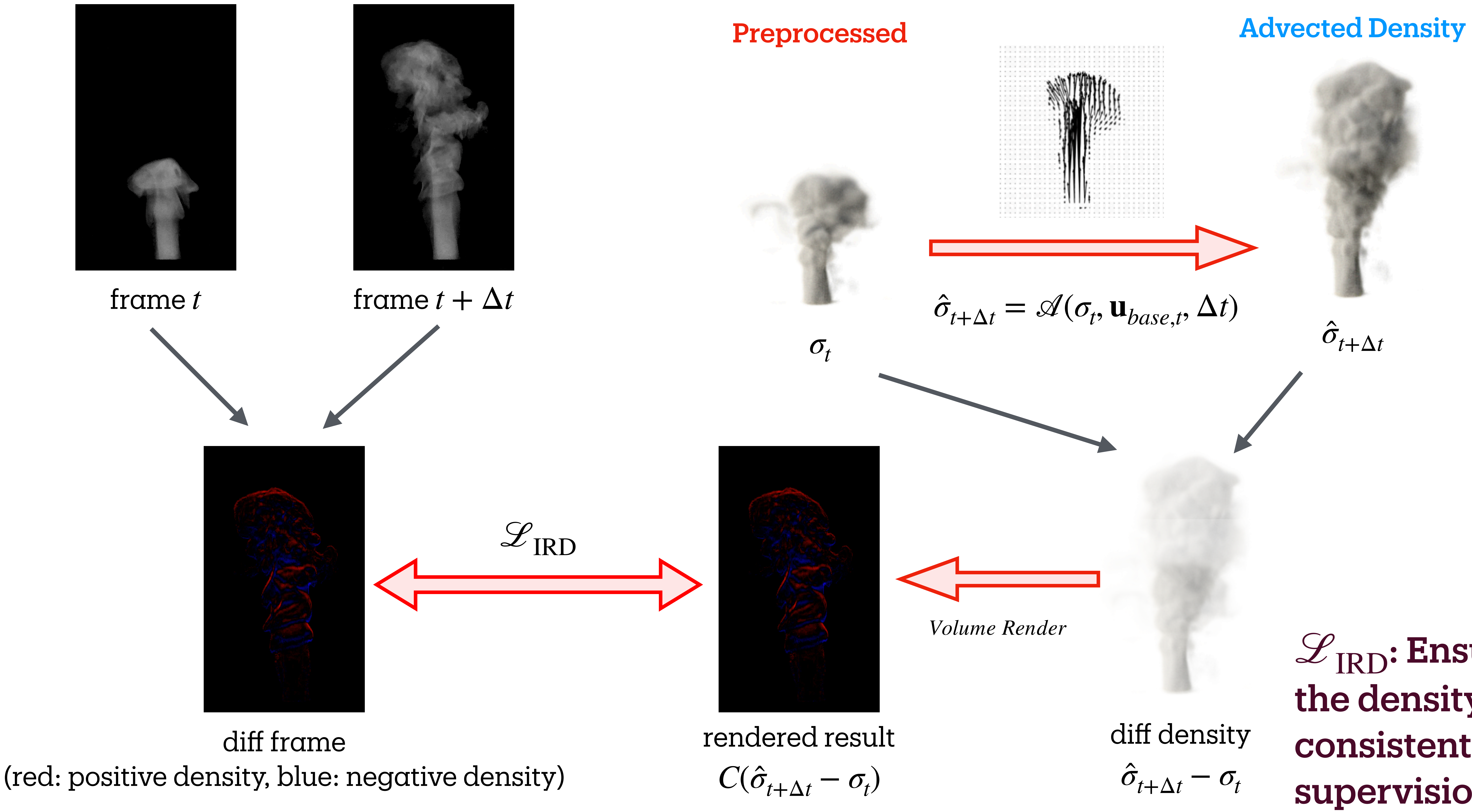
$\mathscr{L}_{\text{transport}}$: **Ensure the velocity field $\mathbf{u}_{base}$ match the motion of the density field σ**

Step 2 - Velocity Estimation (NSE Loss)



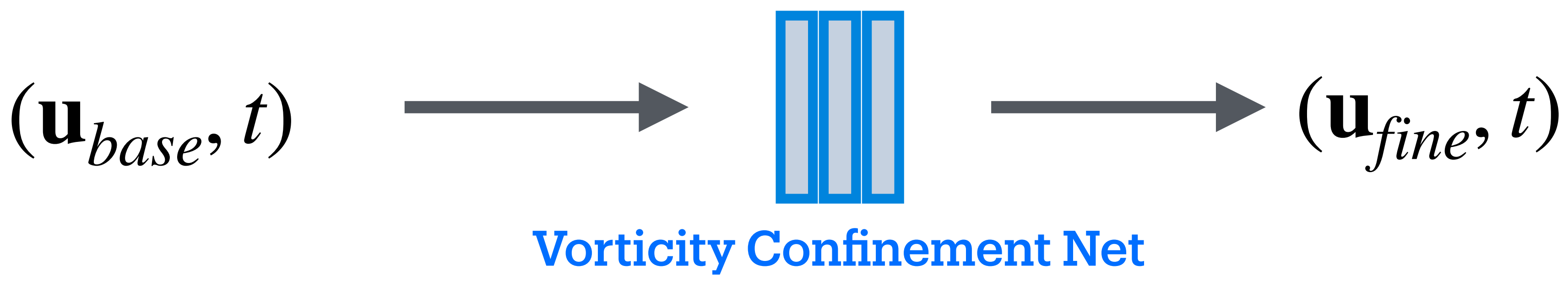
$\mathcal{L}_{\text{NSE}}$: **Ensure the reconstructed velocity field evolves according to the Navier–Stokes equations and remains divergence-free.**

Step 2 - Velocity Estimation (Inter-frame Rendering Difference Loss)

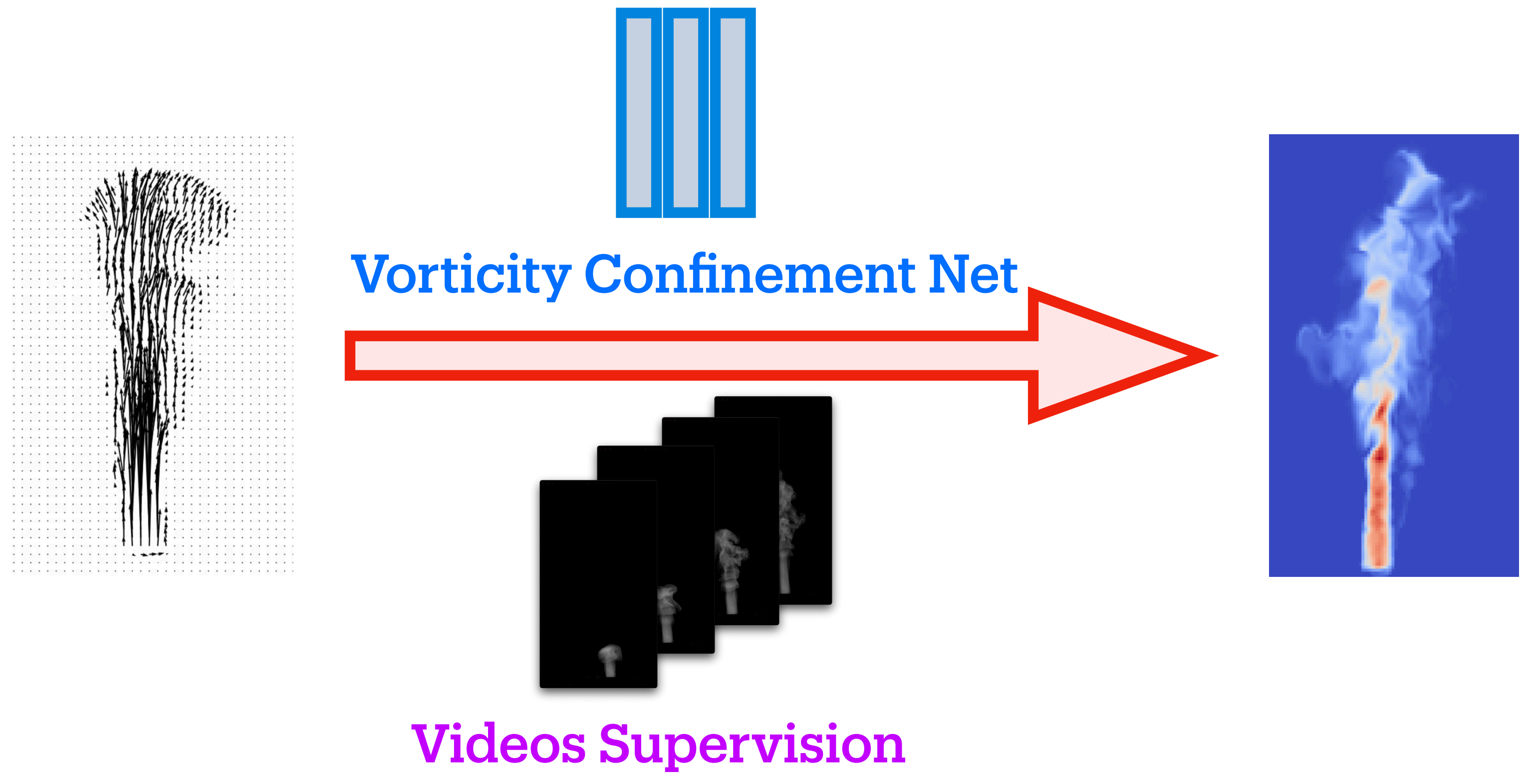


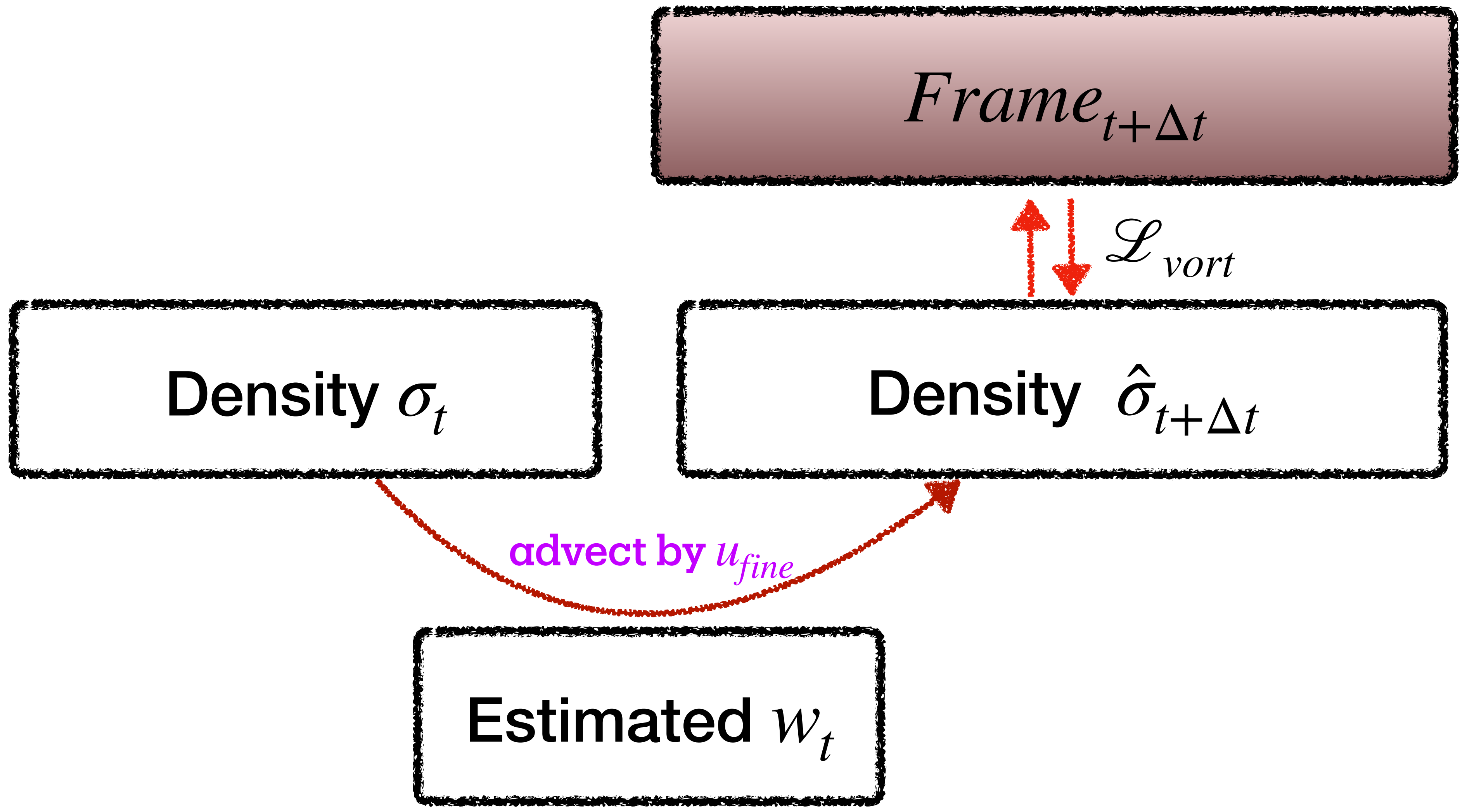
$\mathcal{L}_{\mathrm{IRD}}$: Ensure that the density changes consistent with the supervision video.

Step 3 - Image-based Vorticity Confinement

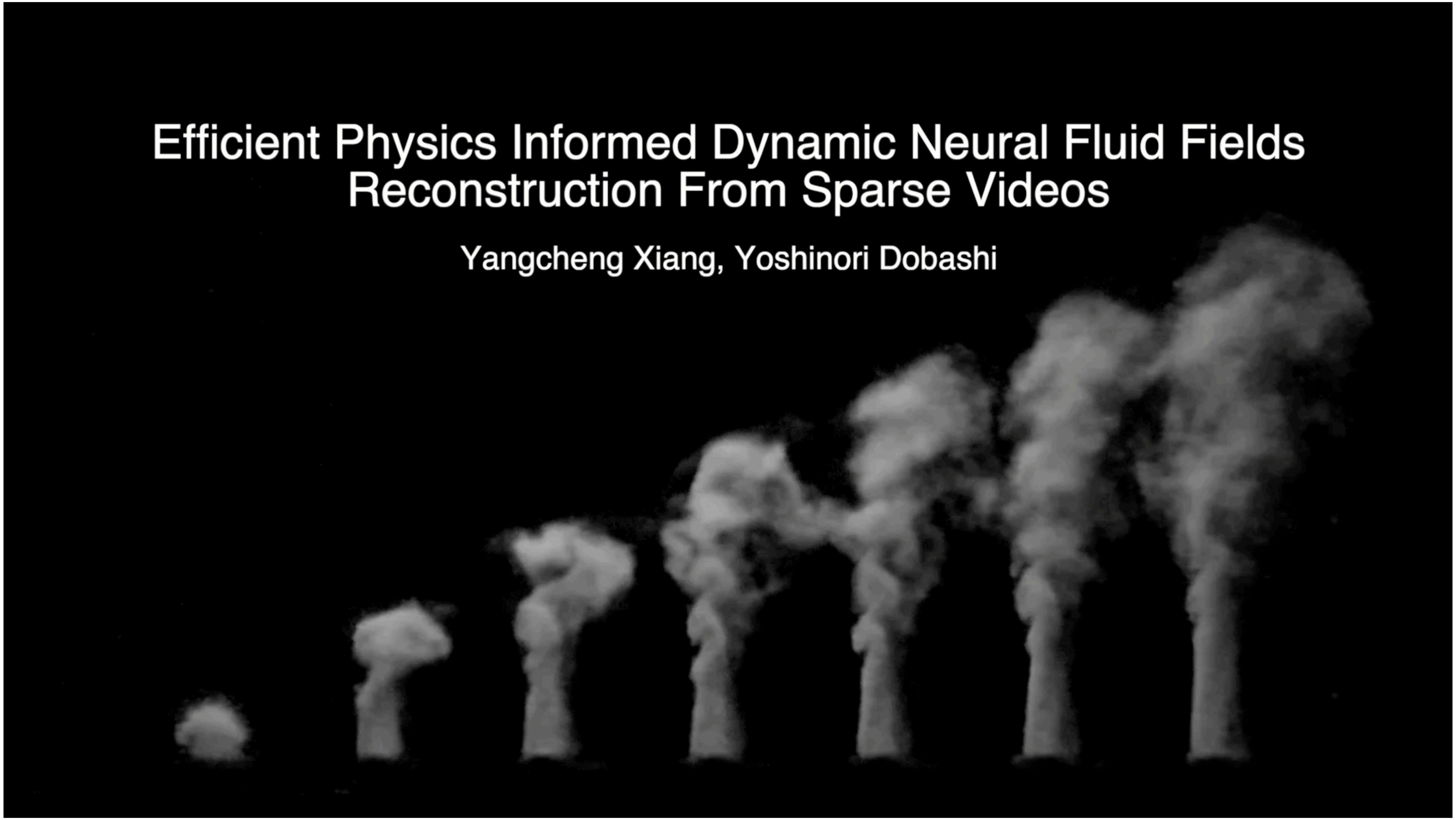


- Helmholtz Hodge decomposition
  - $\mathbf{u}_{\text{base}} = \nabla \phi_{\text{base}} + \nabla \times \mathbf{A}_{\text{base}} + \mathbf{h}_{\text{base}}$
- Enhance the vorticity field
  - $\mathbf{A}_{\text{final},t} = (1 + w_t)\mathbf{A}_{\text{base},t}$
- Assemble the final velocity field
  - $\mathbf{u}_{\text{final}} = \nabla \phi_{\text{base}} + \nabla \times \mathbf{A}_{\text{final}} + \mathbf{h}_{\text{base}}$
- Loss
  - $\mathscr{L}_{\text{vort}} = \mathscr{E}_{o,d}\left[\left|\left|C_{\mathscr{A}(\sigma,\mathbf{u}_{\text{final}})}(o,d) - C_{\text{image}}(o,d)\right|\right|\right]^2$

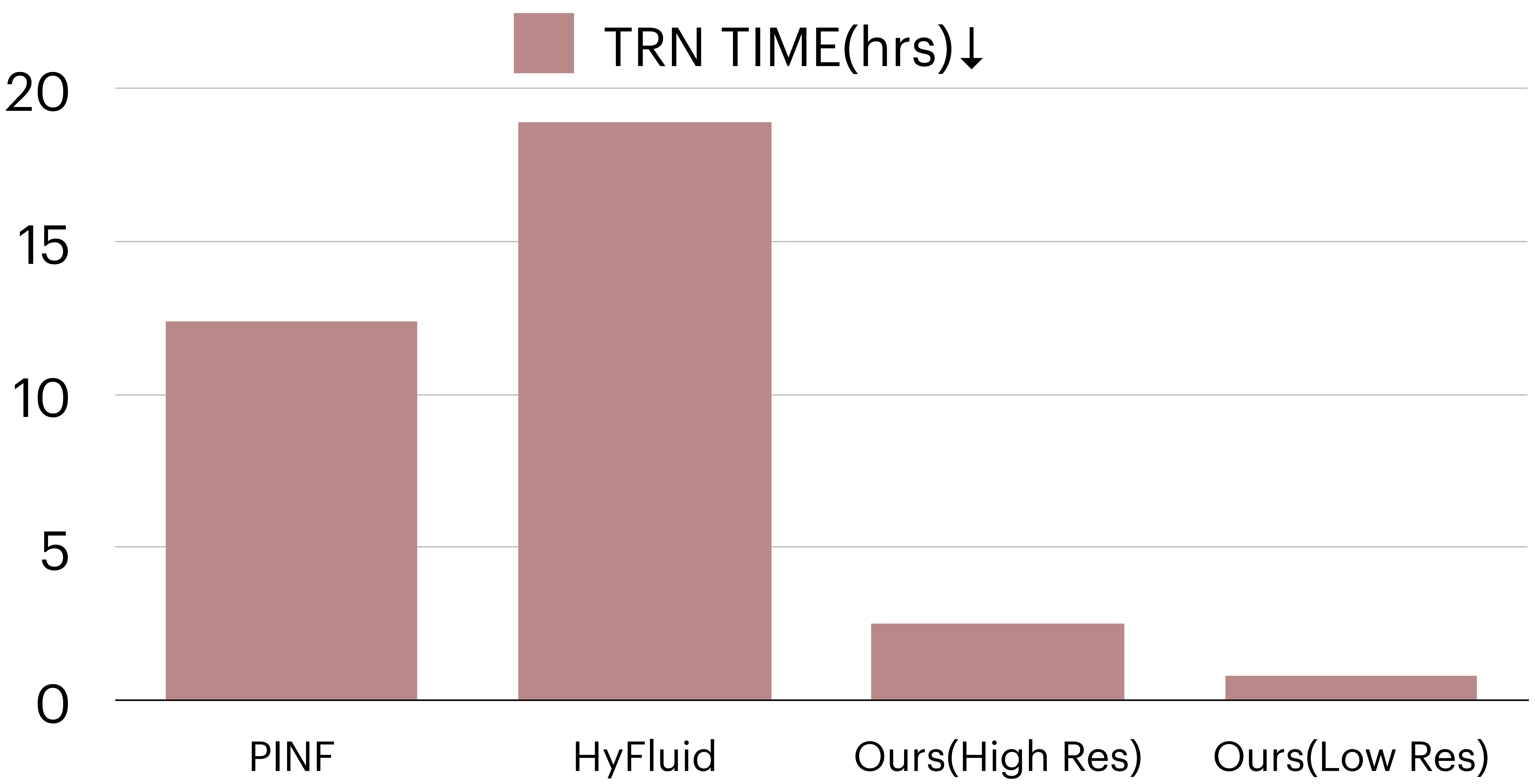


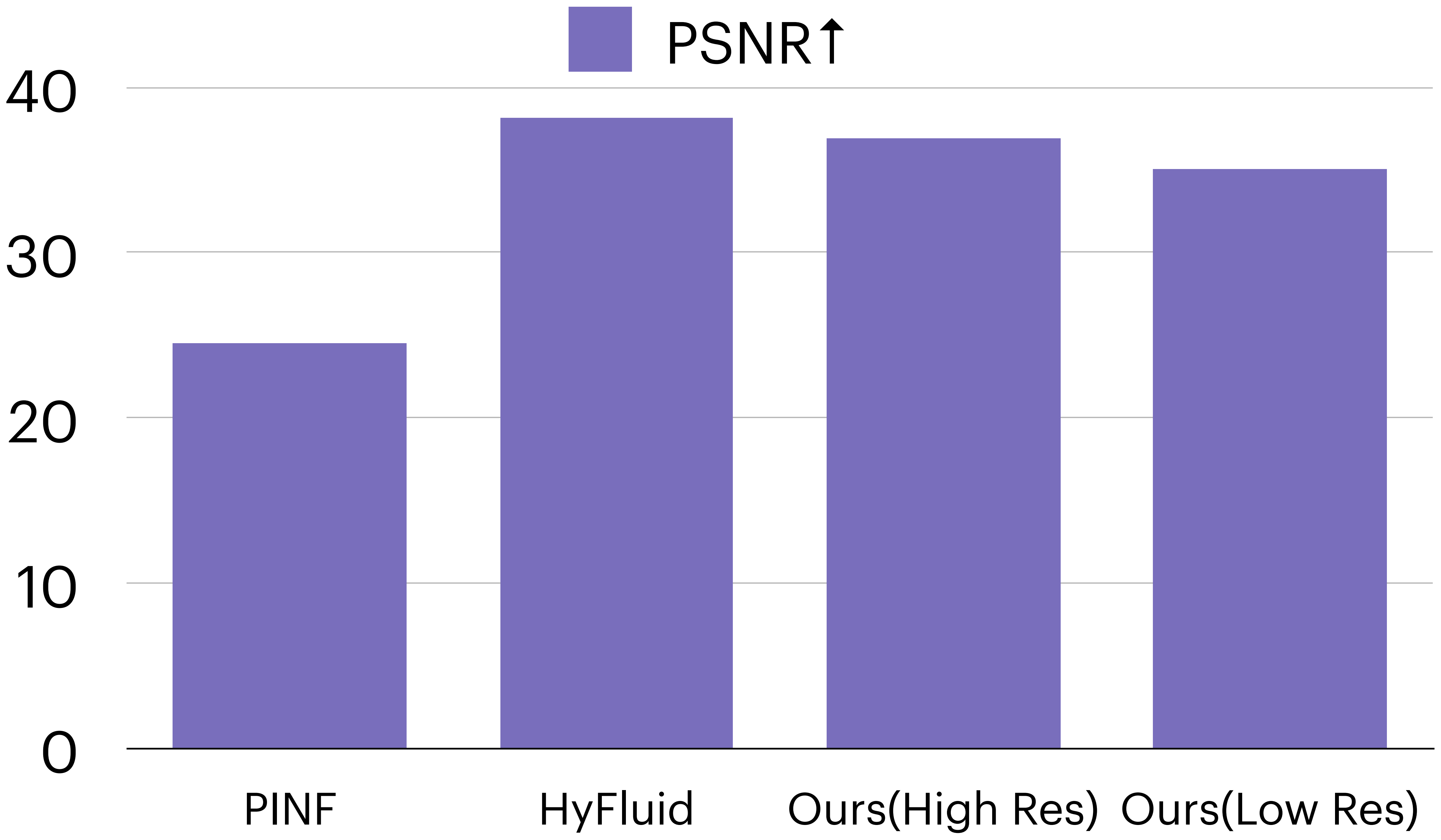


Evaluation - Results



Ablations





Method	Input Video Res	PSNR ↑	$\mathcal{L}_{RGB}$ (AVE.) ↓	TRN TIME	Speedup
PINF	540x960	24.51	0.00191	12.4 hr	1x
HyFluid	540x960	38.12	0.00018	18.9 hr	0.65x
Ours (High Res)	1080x1920	36.88	0.00019	2 hr 30 min	5x
Ours (Low Res)	540x960	35.11	0.00022	49 min	15x

Thanks for your attention.