Surrogate Modeling for Supernova Feedback toward Star-by-star Simulations of Milky-Way-sized Galaxies

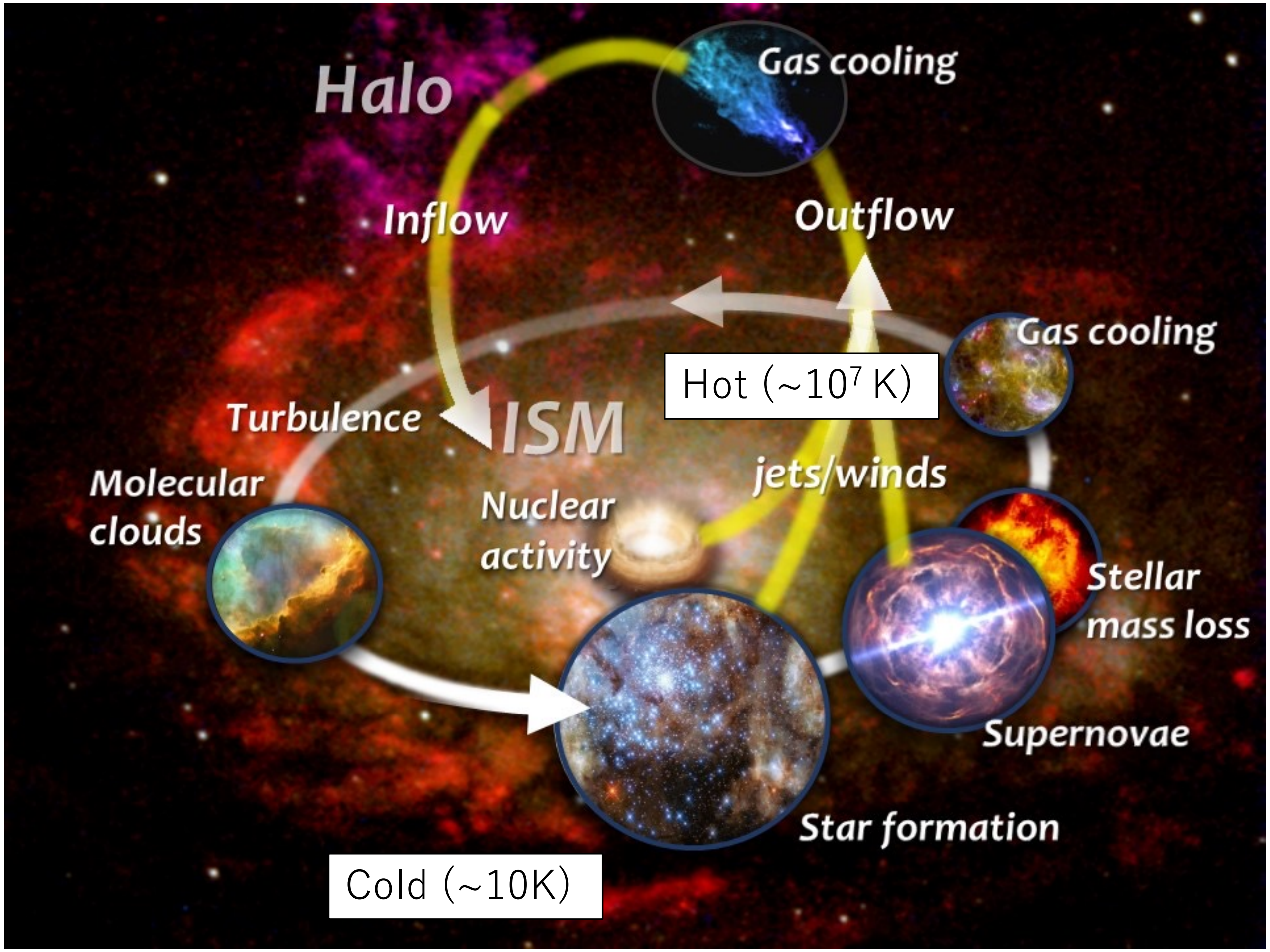
Keiya Hirashima[1] Final year PhD student

With Kana Moriwaki[1], Michiko S. Fujii[1], Yutaka Hirai[2,3], Takayuki R. Saitoh[4], Junichiro Makino[4], Ulrich Steinwandel[5], and Shirley Ho[5]

[1]University of Tokyo, Japan, [2]Tohoku University, Japan, [3]University of Notre Dame, USA
[4]Kobe University, Japan, [5]CCA, Flatiron Institute, USA

2024/7/11 ML4ASTRO2

Multiscale gas dynamics in galaxies



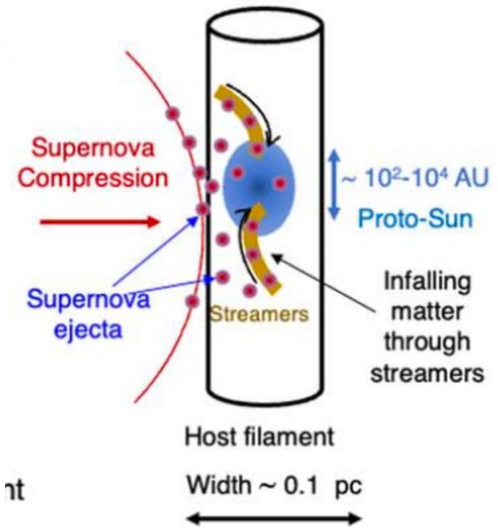
https://www.sron.nl/missions-astrophysics/sto2

Supernova feedback

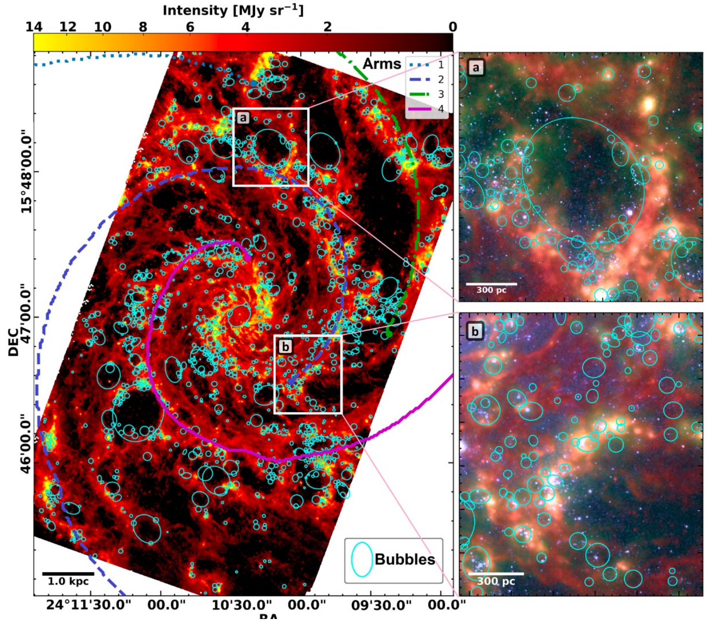
- has impacts across ISM- and galaxy-scale
- drives gas dynamics. – suppress star formation rate and bumps outflow
- has different behavior of outflow/inflow depending on the mass of the host galaxy

Supernovae quantify or trigger star-formation?

How many stars have been born by SNe?



Arzoumanian+2023



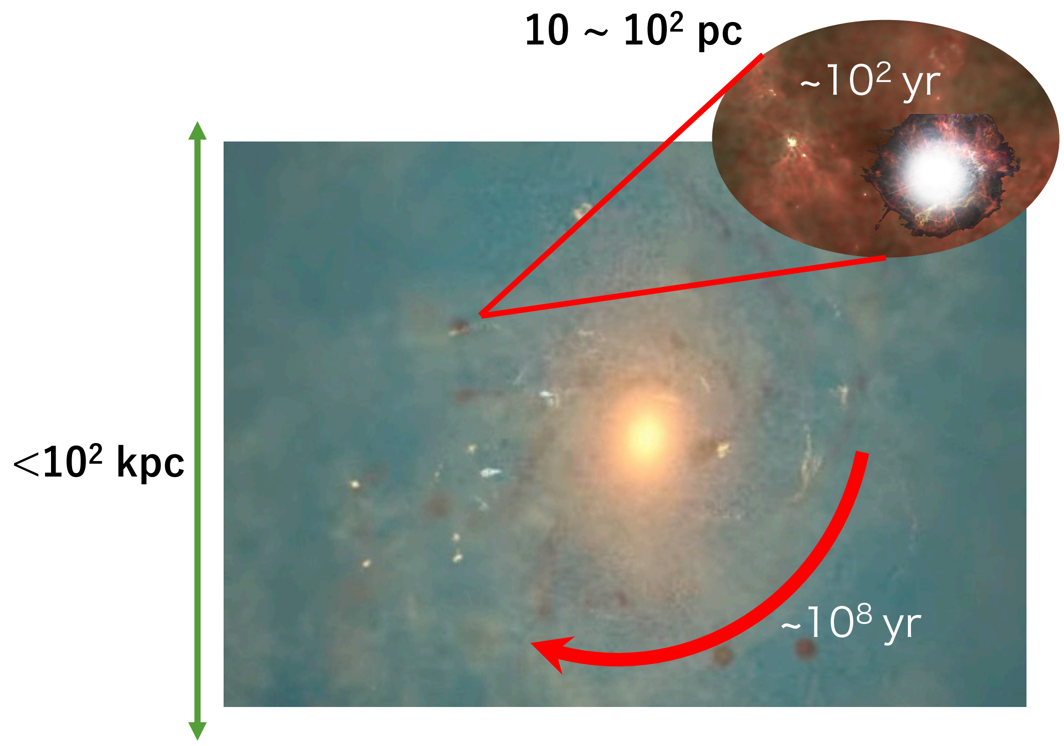
NGC 628 Watkins+2023a,b

- Supernovae can compress clouds/filaments, which can be star-forming regions.

- Tentative evidence of star-formation on shells
- This process might form 14-30% of massive stars in the Milky Way (Thompson+2012)

Star-by-star galaxy simulations, resolving individual stars and stellar feedback

Galaxy Simulations Using SPH*



The formation of the galaxy [1].

*SPH: Smoothed Particle Hydrodynamics

[1] https://www.youtube.com/watch?v=Rdd9KAUcvgQ

[2] Applebaum et al. (2021)

[3] Grand et al. (2021)

ASURA-FDPS (N-body/SPH)

(Saitoh+08,09, Iwasawa+16, Hirashima+23a)

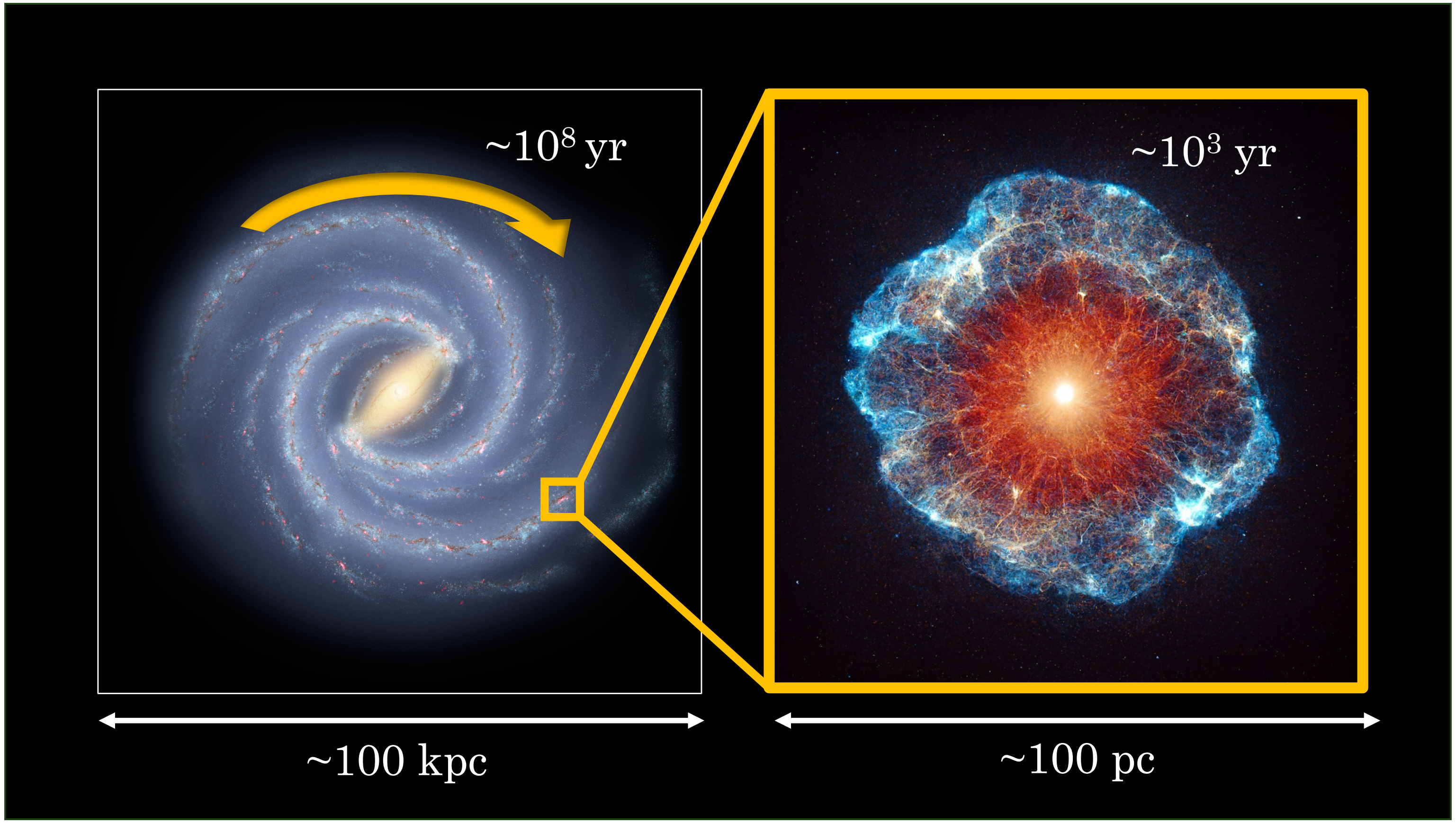
- Gravity + Hydrodynamics (DISPH; saitoh+13)
- Radiative Cooling/Heating (Ferland+17)
- Star formation (Hirai in prep.)
- Feedback
  - SNe Ia/II, AGB, Neutron star merger
- Chemical evolution (CELib; Saitoh17)
- FUV background

- About to represent every single star in simulations, but…
  - Recent studies [2, 3]： $10^3 M_\odot$
  - Our goal (ASURA-FDPS)： $< 10 M_\odot$

Can we go to “star-by-star” resolution??

A Challenge in multiscale simulations

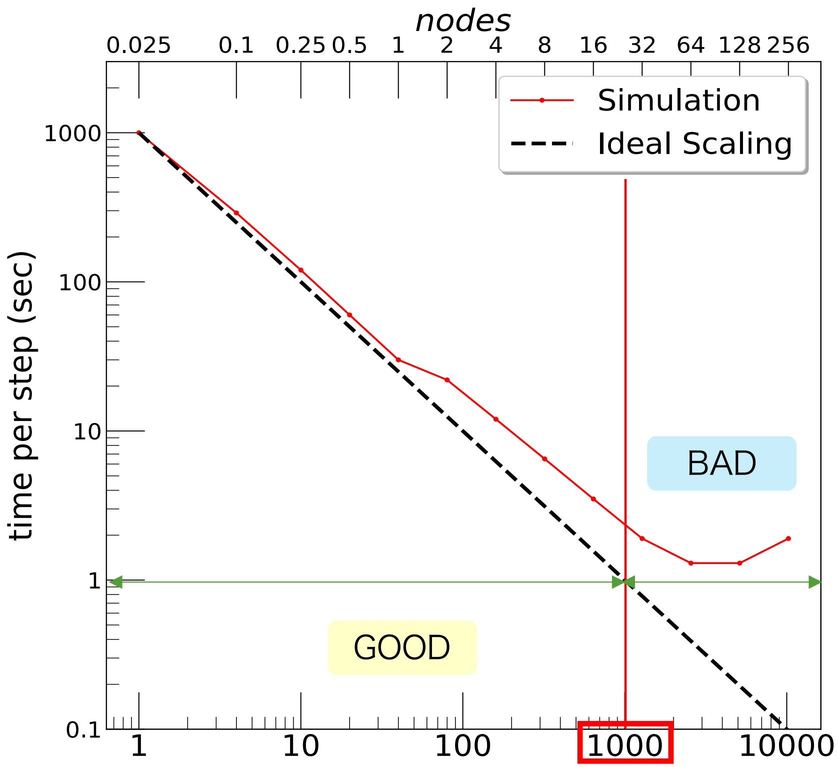
Supernovae are much smaller than galaxies but still impactful on the evolution.



NASA/JPL-Caltech/ESO/R. Hurt

Overheads in Galaxy Formation Simulations

The parallelization efficiency saturates at ~10^3 CPU cores.



Strong Scaling of GADGET-4
(Based on Figure 63 in Springel et al. 2021)

e.g.,

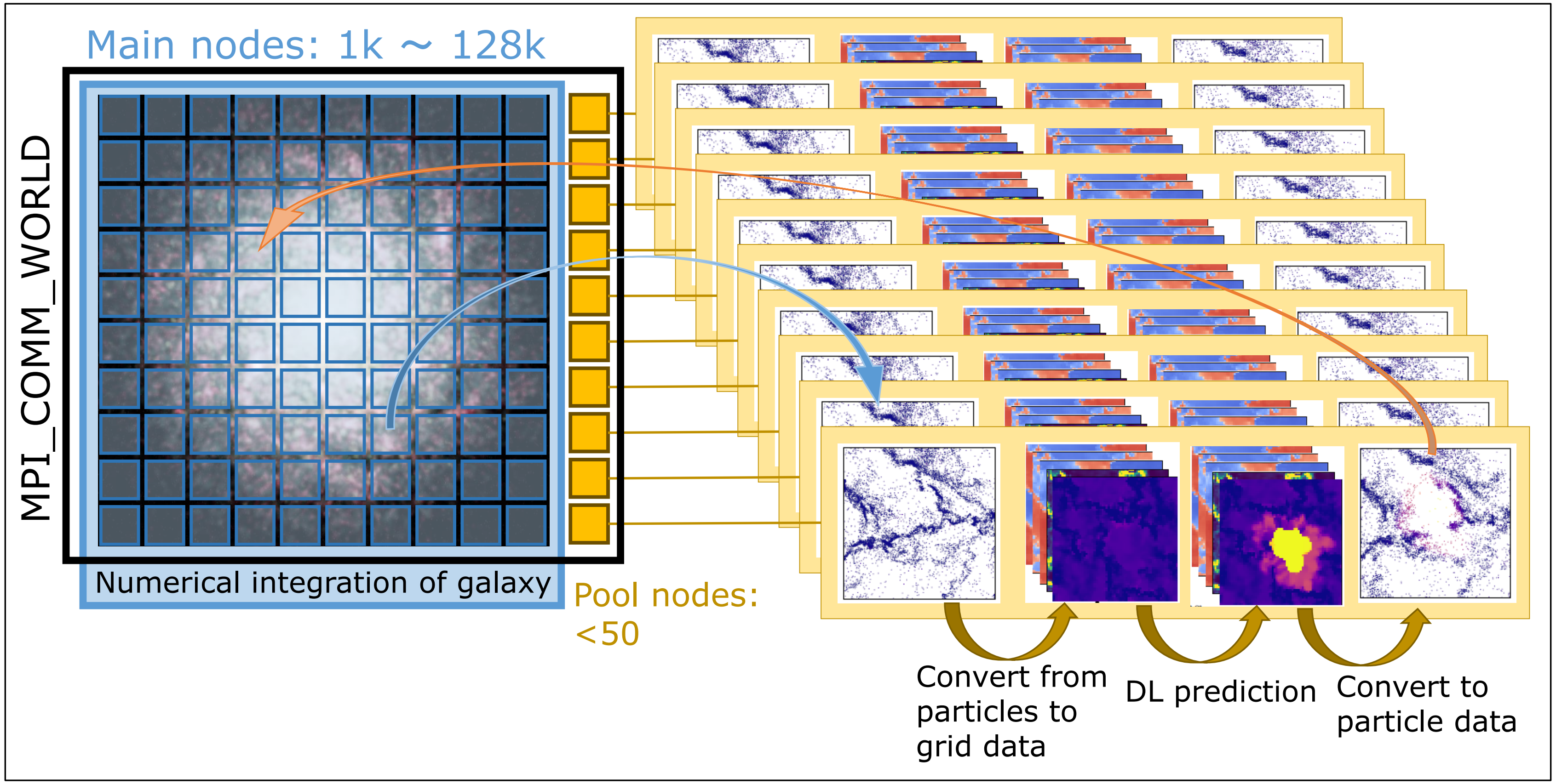
- GADGET-4(Springel+21)
- DC Justice League (Applebaum+21)
- Fire-2(Hopkins+18)

- Due to small timescale regions (e.g. SNe), the communication overhead occurs.
- Even the latest supercomputers cannot solve it (e.g., Fugaku has ~10^6 CPU cores).

-> Decrease the total number of calculation steps

Simulation with Machine Learning Model

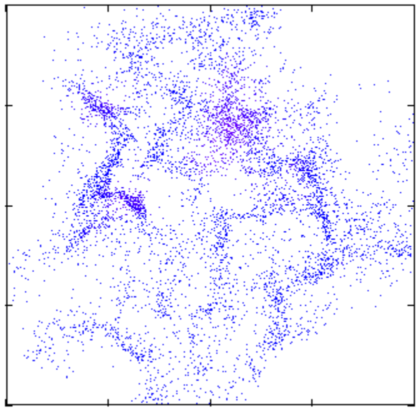
Have ML handle bottlenecks - SNe -.



Training Data (3D cartesian grids)

Temperature	10 [K]
Mean ambient density	40 ~ 60 [cm^{-3}]
Input energy	10^{51} [erg]
Total mass	10^6 [$M_\odot$]
Mass of a gas particle	1 [$M_\odot$]
Softening parameter	0.5 [pc]

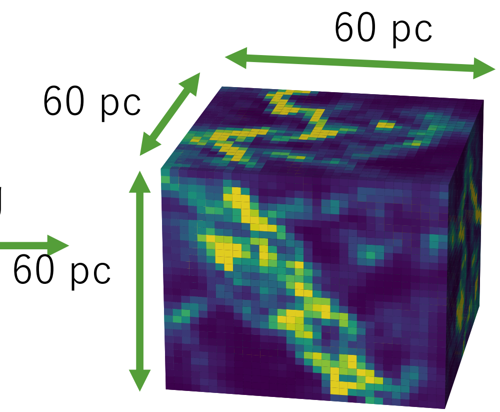
The initial condition for SN simulations in inhomogeneous turbulent clouds



Interpolation using

SPH summation

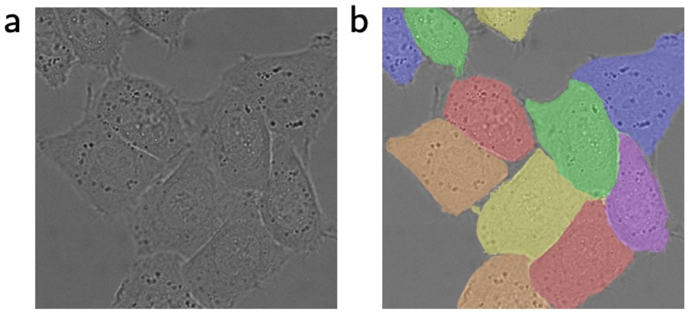
$$\rho_i = \sum_j m_j W(r_{ij}, h_j)$$



64^3 voxels

3D U-Net

- Ronneberger et al. 2015
- CNN-based
- Decoder → Shrink
- Encoder → Enlarge



(a) Raw image (b) Segmentation

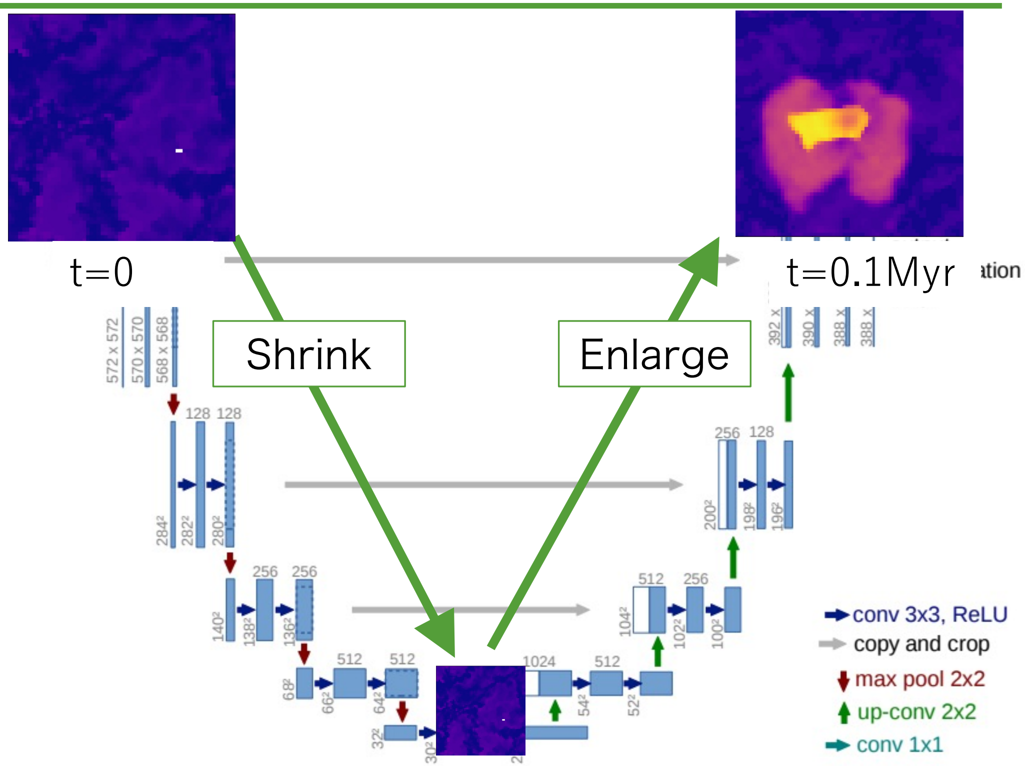
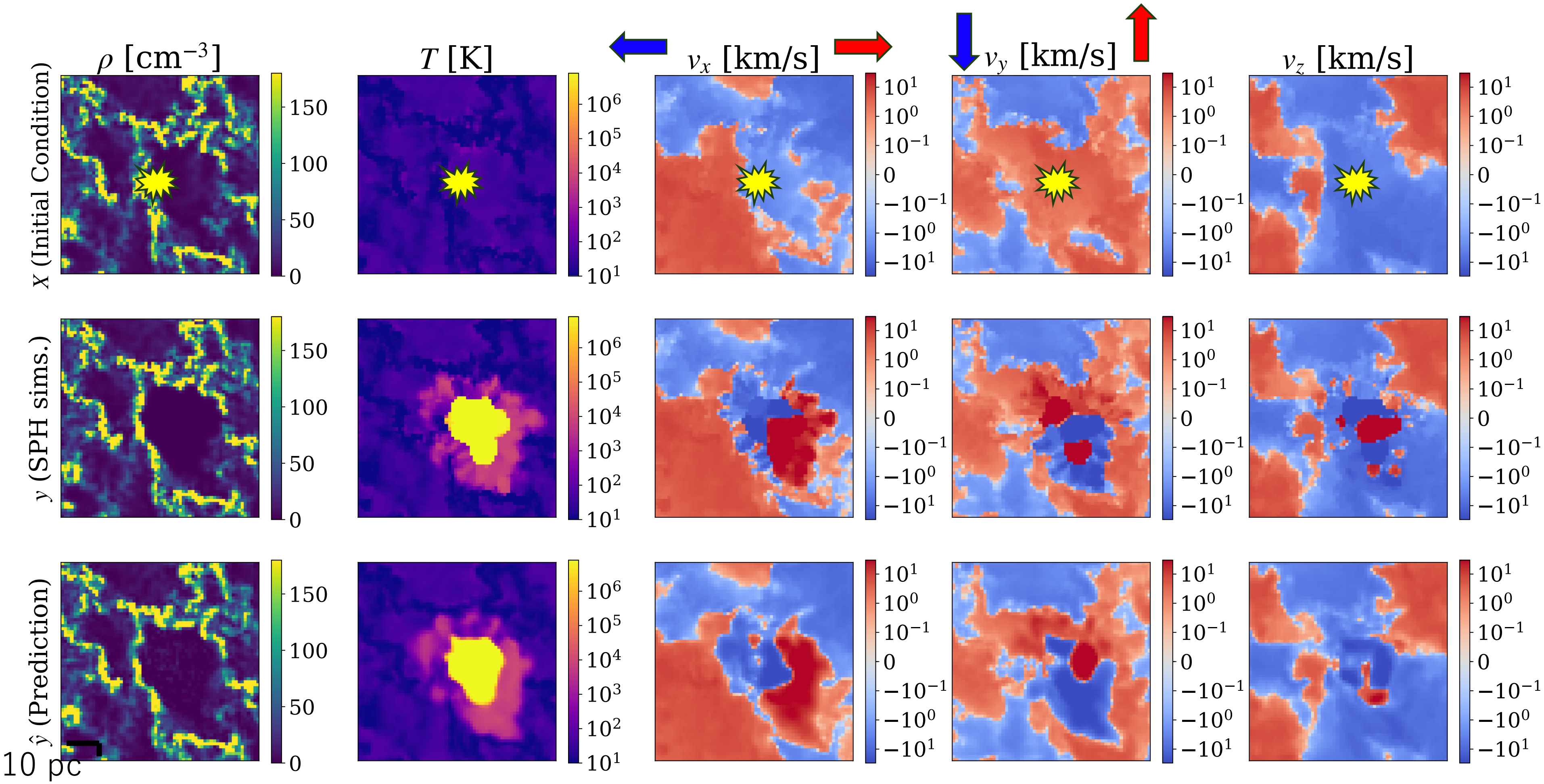


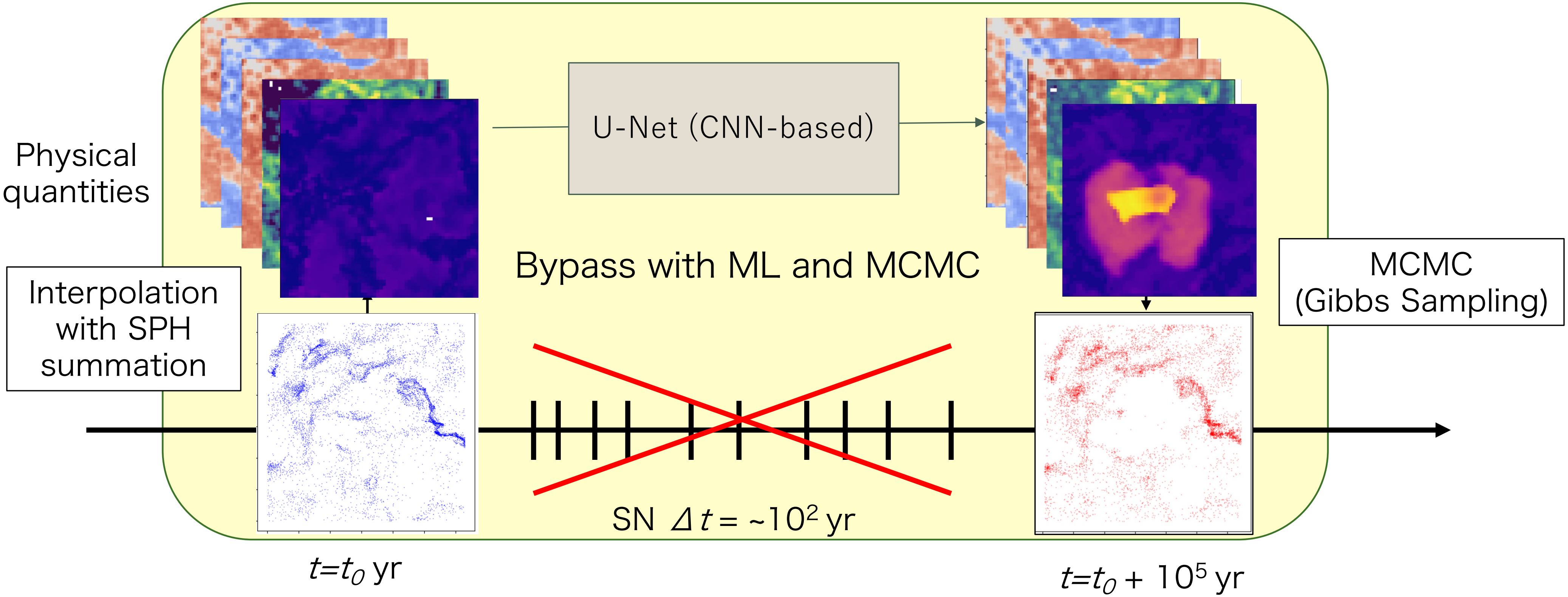
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Prediction result

An inhomogeneous shell emerges from the dense filament blocking in both sim and pred.



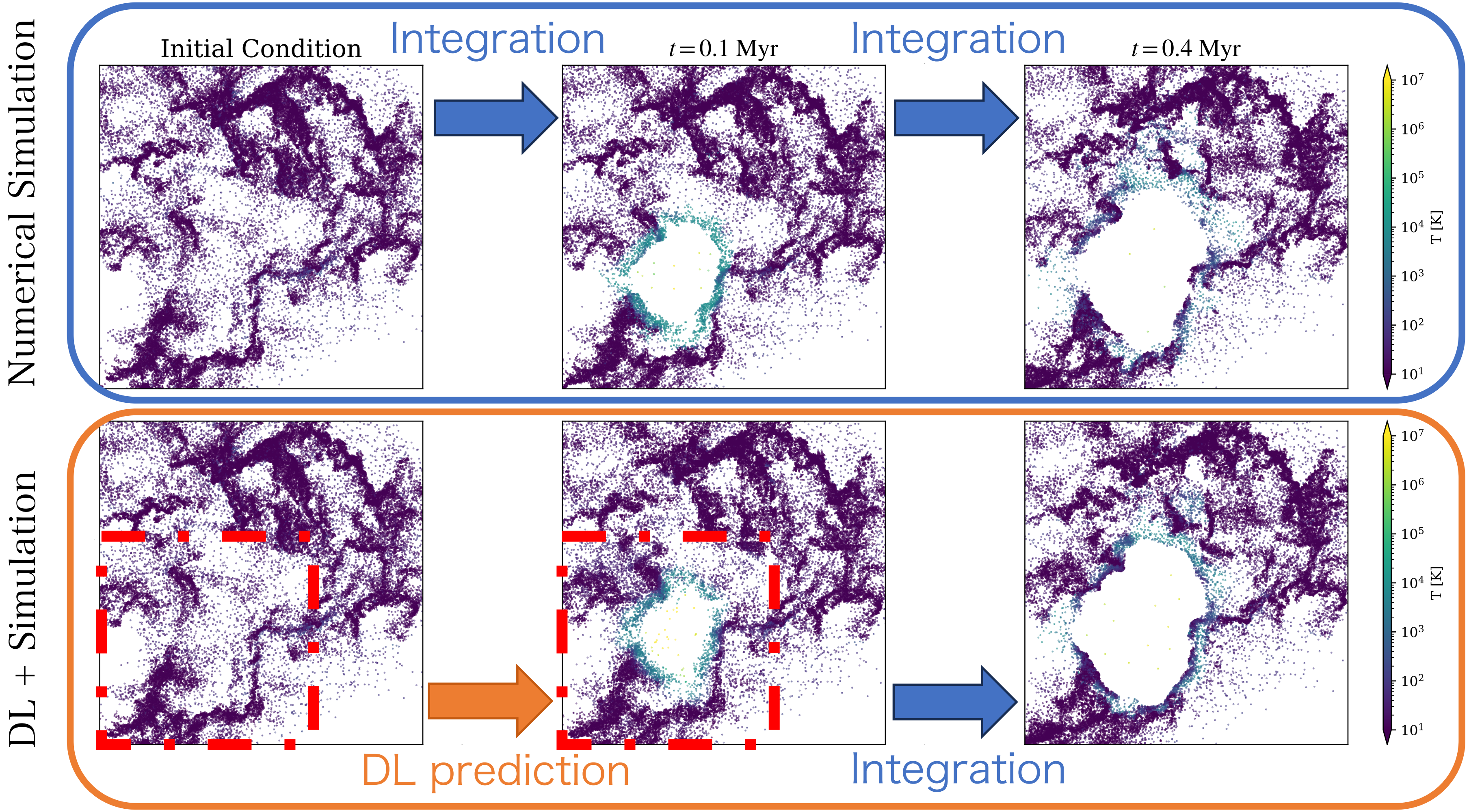
Surrogate modeling for SN feedback



The rest of the region in the galaxy
$\Delta t \sim 10^5$ yr

Please take a look at Hirashima+23b, arXiv:2311.08460 for more details!

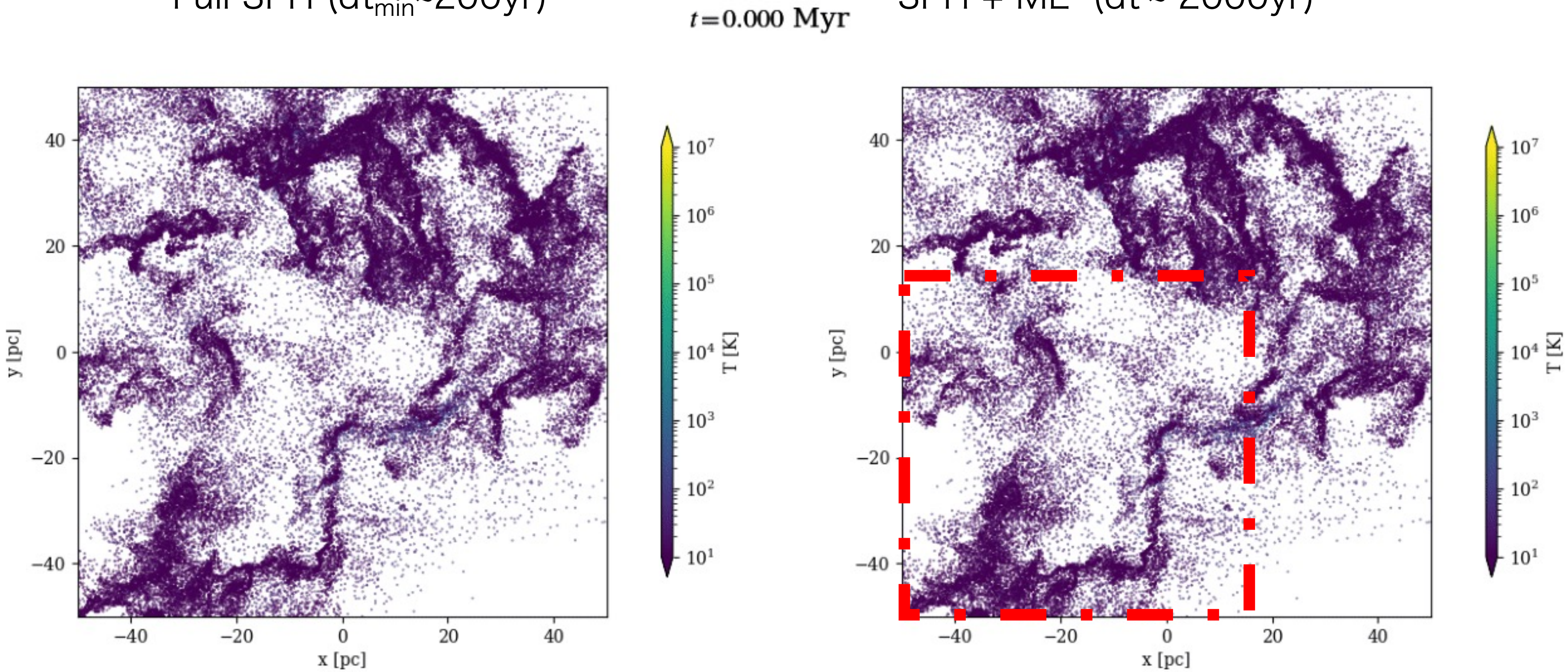
Incorporating ML with simulations



Test #1: SN feedback in Molecular Clouds

At t=0.1Myr, the particles within 60 pc^3 around a SN are incorporated with the parent simulation.

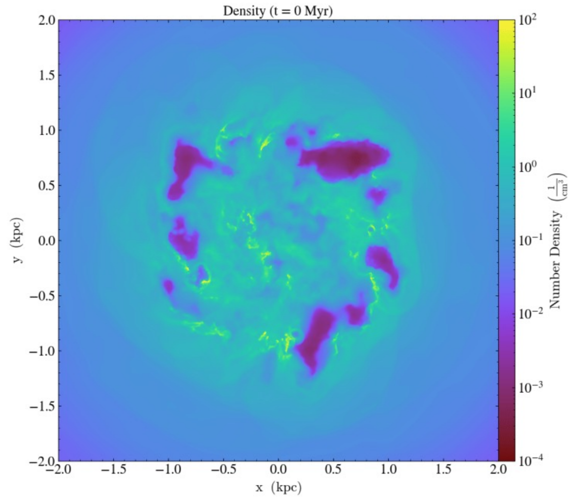
Full SPH (dt_{min}~200yr) SPH + ML (dt ~ 2000yr)



Properties of the dwarf galaxy simulations

Initial Condition: Isolated disk dwarf galaxy

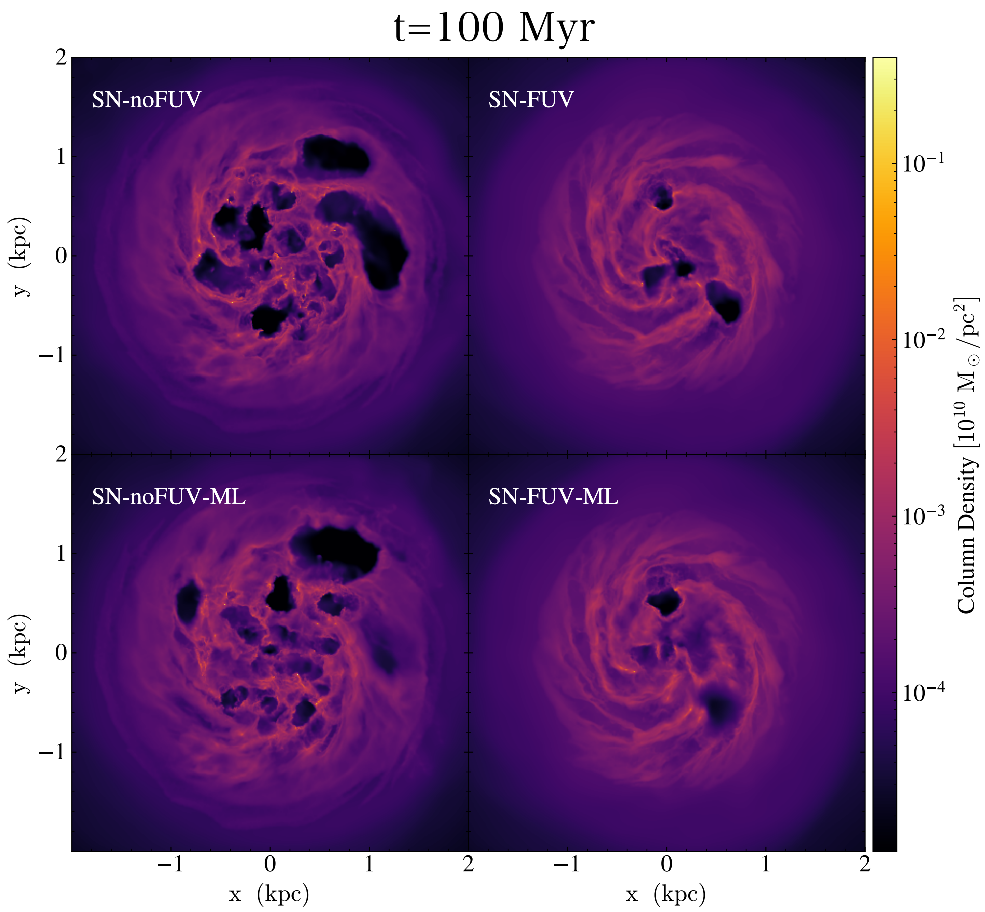
- M_{vir} ~ 10^{10} Msun
- M_{baryon} ~ 10^{7} Msun
- m_{baryon} ~ 4 Msun



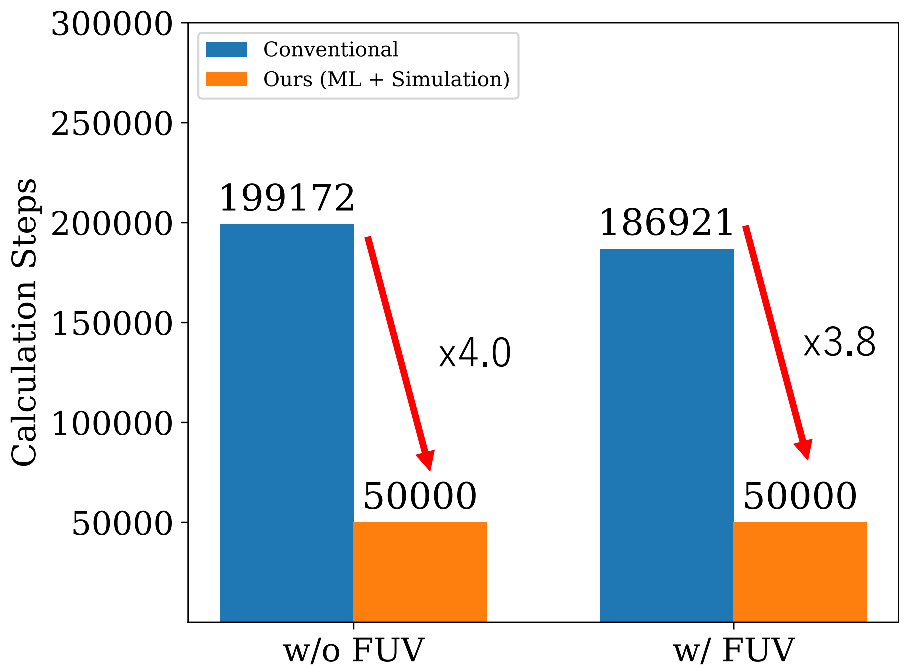
Tag	SN feedback	FUV backgound	Time-stepping
SN-noFUV	Dump thermal energy		Variable
SN-FUV	Dump thermal energy	✓	Variable
SN-noFUV-ML	Dump thermal energy (<1 cm^{-3}) **Surrogate modeling** (> 1 cm^{-3})		Fixed 2000 yr
SN-FUV-ML	Dump thermal energy (<1 cm^{-3}) **Surrogate modeling** (> 1 cm^{-3})	✓	Fixed 2000 yr

Test #2: galaxy simulations with ML (Preliminary)

Superbubbles made of multiple supernovae are resolved.



Our new approach is faster by a factor of four!



Summary

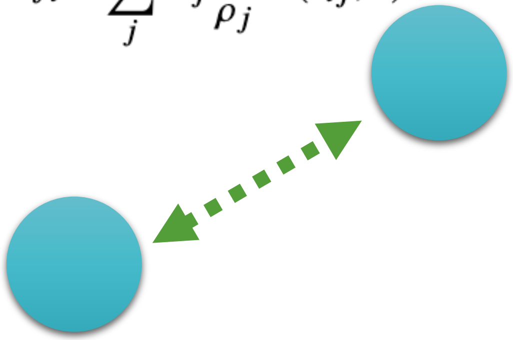
- ✅ Implement a surrogate model for SN feedback with our simulation code for star-by-star simulations
- Test run:
  - ✅ Molecular Clouds (10^6 Msun)
    - 7 times faster
    - Energy and momentum are converged better than low-res sims.
  - On-going: isolated dwarf galaxy (10^{10} Msun)
    - 4times faster
    - Checking convergence of SFH and mass/energy loading factor
- Future work
  - LMC size (10^{11} Msun)
  - MW size (10^{12} Msun)

Reference:
1) Hirashima+23a, MNRAS, 526, 3
2) Hirashima+23b, NeurIPS2023-AI4Science, arXiv:2311.08460

Basic physical models in galaxy simulations

- N-body/SPH
  - Dark matter, stars, and gas are implemented as particles.
  - In every timestep, physical quantities are updated by solving interactions.
  - ~10^{10} Particles for MW-sized galaxy

$$f_i = \sum_j m_j \frac{f_j}{\rho_j} W(r_{ij}, h).$$



- Equation of Motion
  - Gravity
  - Hydrodynamics
    - Equation of State
    - Navier-Stokes equation
    - Cooling/Heating
  - Radiation and so on

$$P = (\gamma - 1)\rho u,$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho \nabla \cdot \boldsymbol{v},$$

$$\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} = -\frac{\nabla P}{\rho} + \boldsymbol{a}_{\mathrm{visc}} - \nabla \Phi,$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho} \nabla \cdot \boldsymbol{v} + \frac{\Gamma - \Lambda}{\rho},$$

Surrogate Modeling for fluid dynamics

- Methods to surrogate simulations governed by partial differential equations (PDE)
- Choose methods by looking at the generalizability and scalability

- The loss function is tuned to learn specific physics/PDEs.
- Hard to generalize to new tasks

- Directly learn physics from simulation data
- It is hard to generalize to the new parameter set

Zhou et al. (2024)

Data-Driven Surrogate Models

ML models are learning simulations as neural operators.

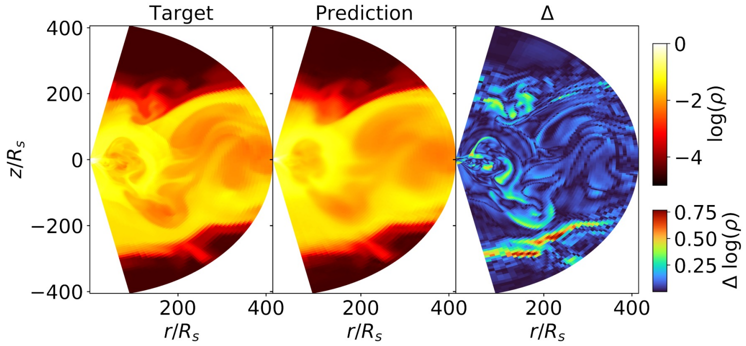
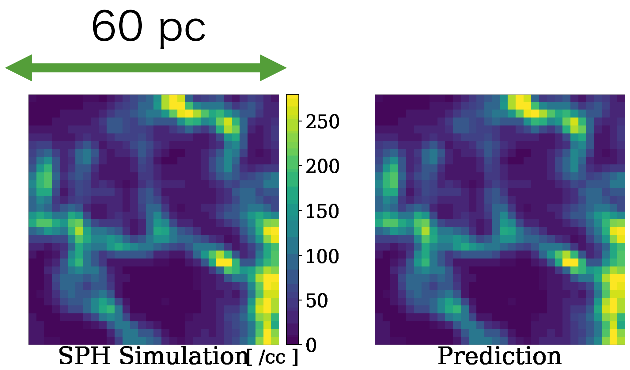
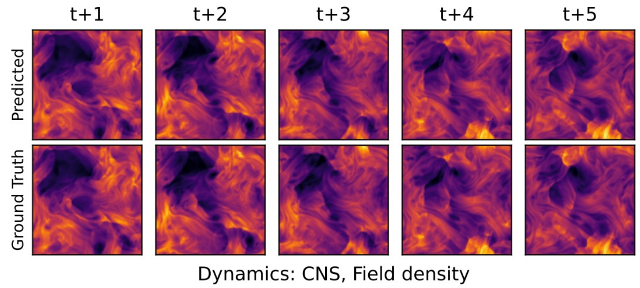


Figure 9. Density predictions by the multi-sim direct model, compared with the simulation PNST1 at $t = 172035M$.

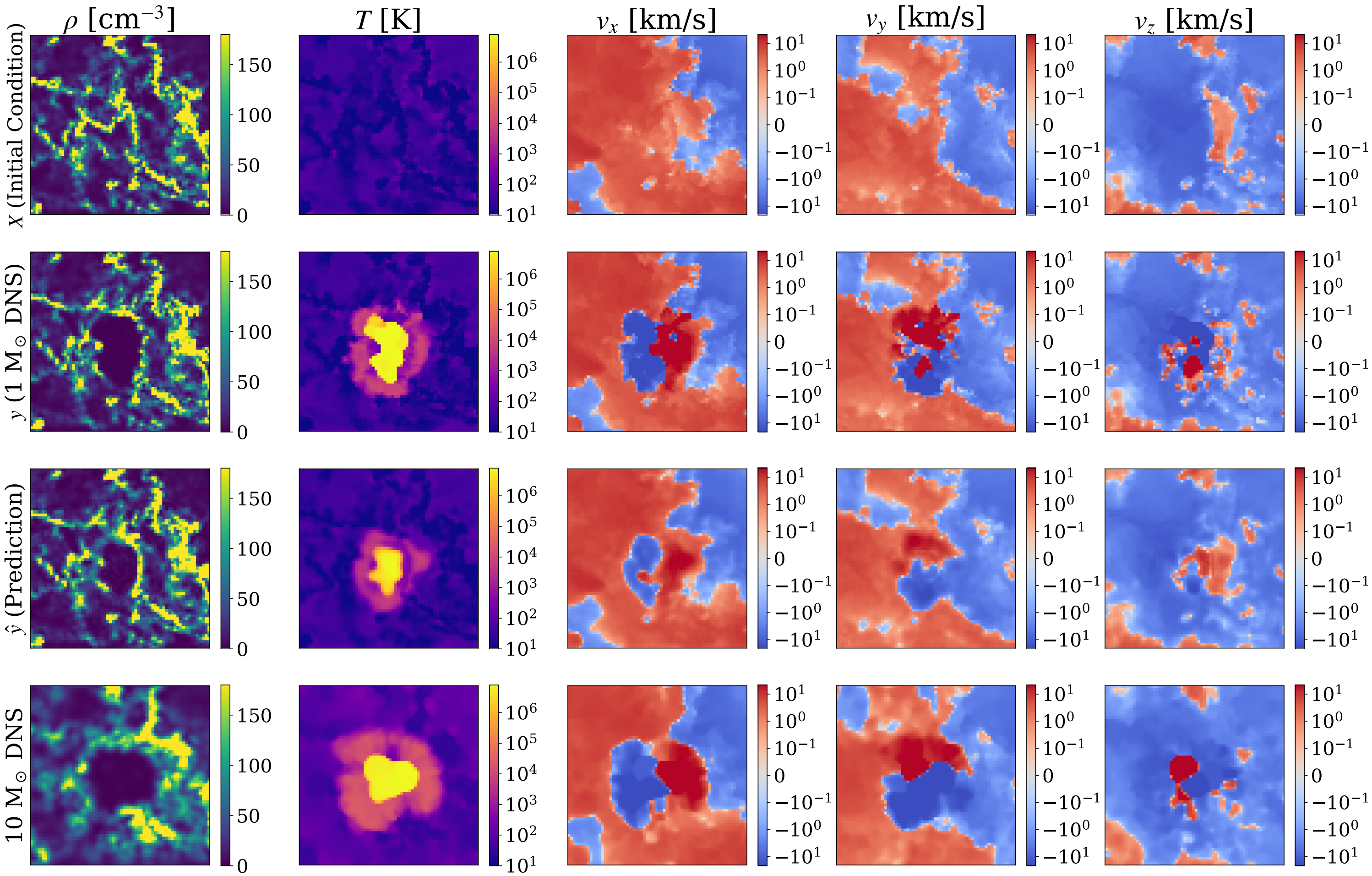




	Dataset	Model	Channels	Simulation
Duarte et al. 2022	an accreting black hole	2D-Unet	Density	PLUTE (mesh)
Hirashima et al. 2023a, 2302.00026	Blast wave by a SN in turbulent molecular cloud	3D-MIM	Density & Velocity	SPH(ASURA-FDPS)
McCabe et al. 2023	Multiple CFD simulations	2D/3D-ViT	Density, Pressure, & Velocity	PDE-Bench

Comparison to low-resolution simulation

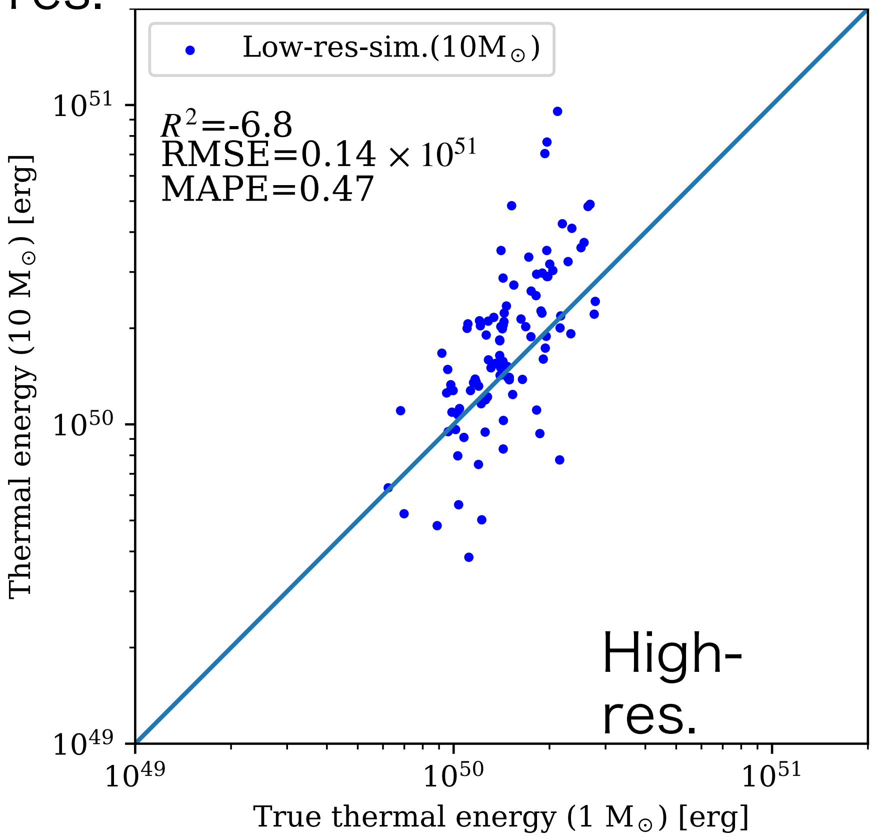
Low-res. Sims cannot resolve the blast wave.



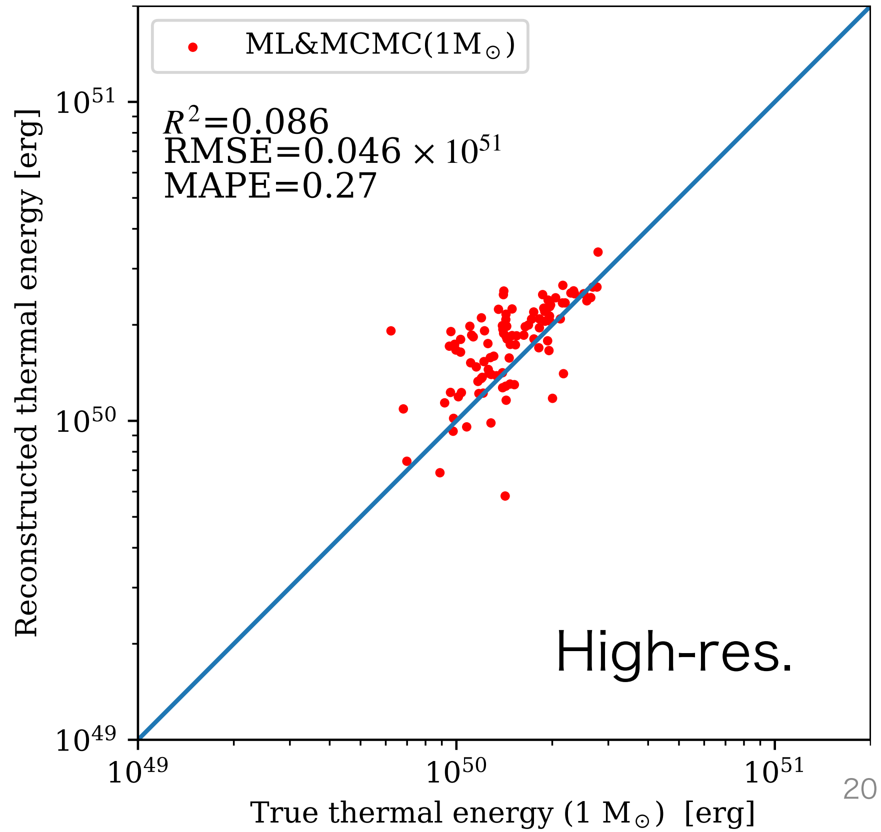
Fidelity Evaluation in Thermal Energy

- Compared to the low-res. sims., our method can duplicate the thermal energy more accurately.

Low-res.



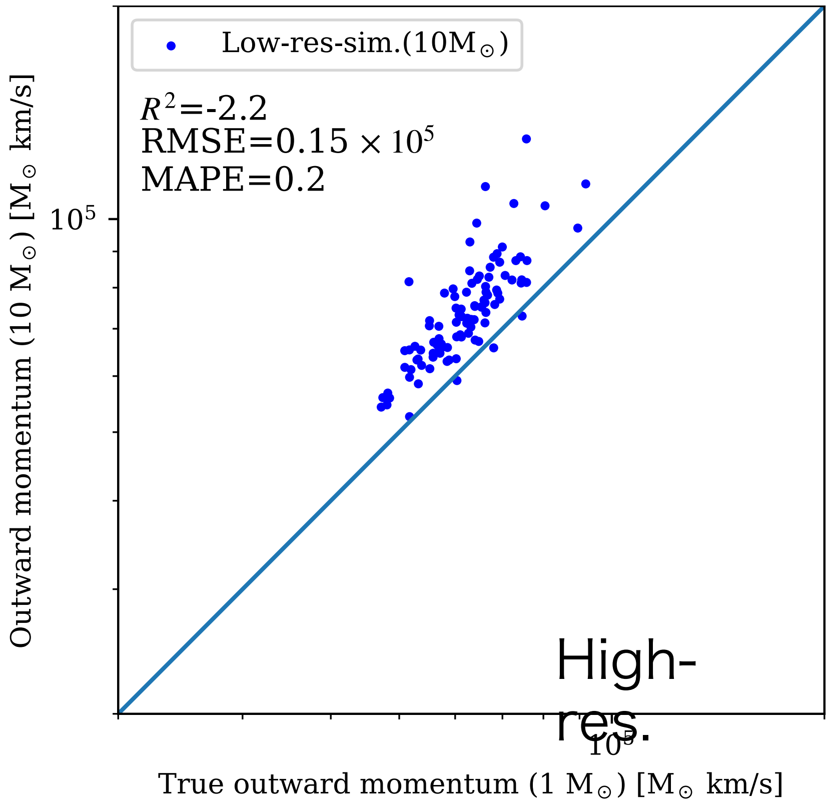
Ours



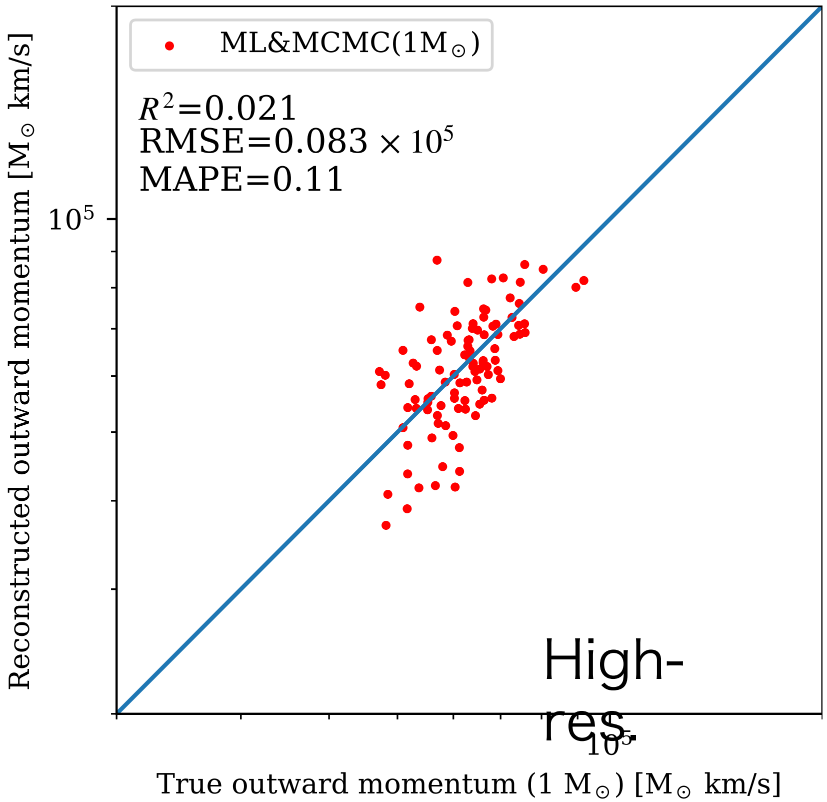
Fidelity Evaluation in Outer Momentum

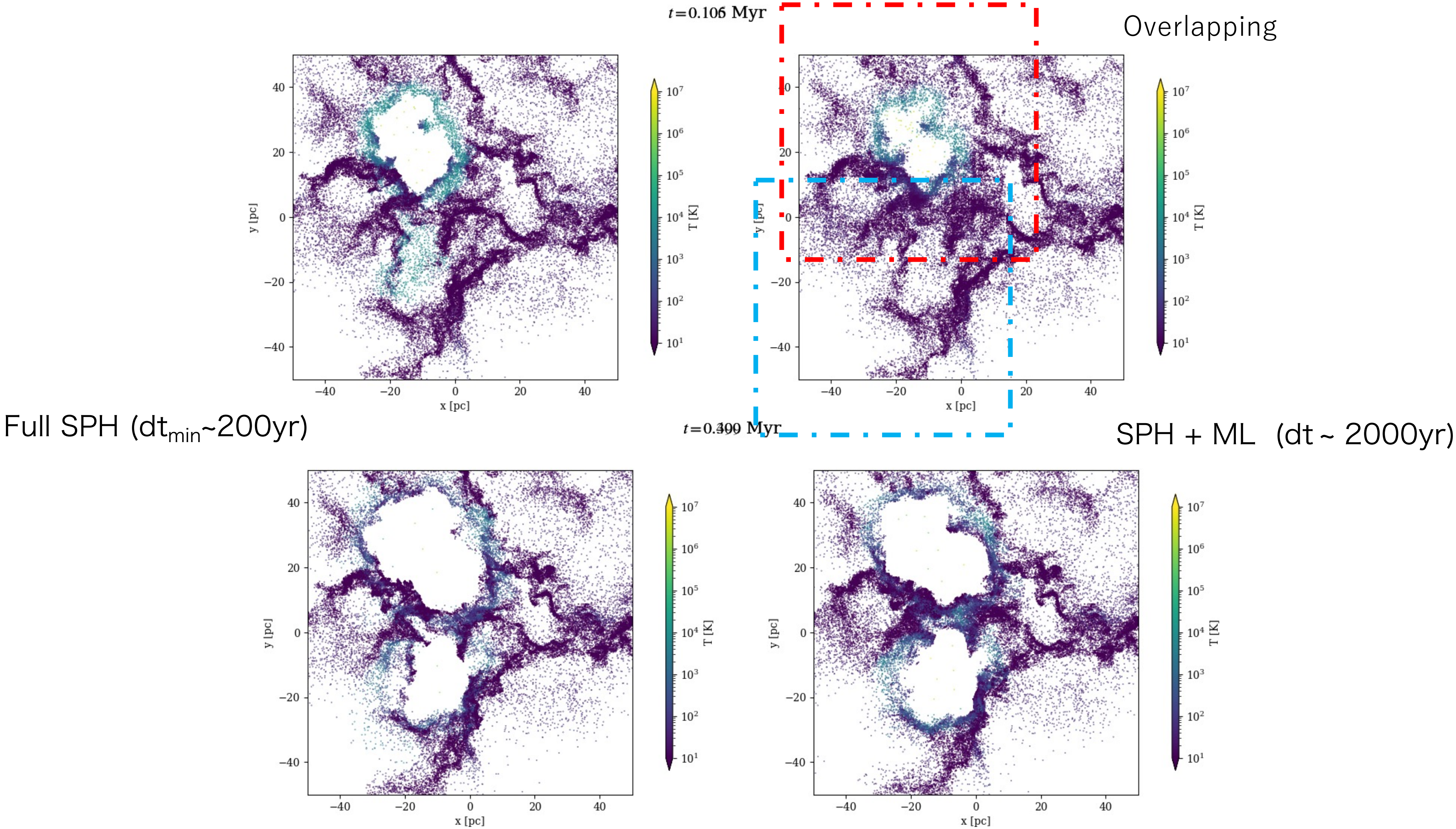
- Both the low-res. sims. and our reconstruction have biases.

Low-res.



Ours





Overlapping
Full SPH (dt_min~200yr)
SPH + ML (dt ~ 2000yr)
x [pc]
y [pc]
T [K]