*The Predictive Power of Computational Astrophysics as a Discovery Tool*
*Proceedings IAU Symposium No. 362, 2023*
*D. Bisikalo, D. Wiebe & C. Boily, eds.*
doi:10.1017/S1743921322001739

# Predicting the Expansion of Supernova Shells Using Deep Learning toward Highly Resolved Galaxy Simulations

**Keiya Hirashima[1]†, Kana Moriwaki[2], Michiko Fujii[1], Yutaka Hirai[3,4,5], Takayuki Saitoh[6] and Junichiro Makino[6,5]**

[1]Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[2]Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[3]Department of Physics, University of Notre Dame, 225 Nieuwland Science Hall, Notre Dame, IN 46556, USA

[4]Astronomical Institute, Tohoku University, 6-3, Aramaki, Aoba-ku, Sendai, Miyagi 980-8578, Japan

[5]RIKEN Center for Computational Science, 7-1-26 Minatojima-Minami-machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

[6]Department of Planetology, Graduate School of Science, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan

**Abstract.** The load imbalance and communication overhead of parallel computing are crucial bottlenecks for galaxy simulations. A successful way to improve the scalability of astronomical simulations is a Hamiltonian splitting method, which needs to identify such regions integrated with smaller timesteps than the global timestep for integrating the entire galaxy. In the case of galaxy simulations, the regions inside supernova (SN) shells require the smallest steps. We developed the deep learning model to forecast the region affected by the SN shell's expansion during one global step. In addition, we identified the particles with small timesteps using image processing. We can identify target particles using our method with a higher identification rate (88 % to 98 % on average) and lower "non-target"-to-"target" fraction (6.4 to 5.5 on average) compared to the analytic approach with the Sedov-Taylor solution. Our method using Hamiltonian splitting and deep learning will improve the performance of extremely high-resolution galaxy simulations.

**Keywords.** Deep Learning, Supernova, Galaxy: formation

## 1. Introduction

Galaxies have formed via several physical processes such as gravitational and hydrodynamic forces, radiative cooling and heating, star formation, supernova explosions, and chemical evolution. Galaxy formation simulations have been performed using $N$-body/hydrodynamics simulations such as smoothed particle hydrodynamics (SPH; Gingold and Monaghan 1977; Lucy 1977), which is a particle-based method using gas particles smoothed with a kernel size depending on the local density.

† E-mail: hirashima.keiya@astron.s.u-tokyo.ac.jp

The resolution of $N$-body/SPH simulations depends on the number of star and gas particles. The number of stars in the Milky-Way galaxy exceeds $10^{10}$. We need more than $10^{10}$ particles to resolve individual stars (star-by-star simulation).

As higher-performance supercomputers have been developed, a higher resolution has been achieved. The highest resolution for large-scale galaxy simulations, IllustrisTNG (Weinberger et al. 2017; Pillepich et al. 2018) with $10^{10}$ particles were now within reach using 25,000 CPU cores.

A crucial bottleneck of galaxy simulations is the communication overhead. In massively parallel computing using more than $\sim 1000$ CPU cores, the communication takes longer than the calculations (see Figure 63 in Springel et al. 2021) simply because the communication costs increase as the parallelization degree increases.

Another problem is the worsening scalability due to the load imbalance. With a better resolution, we can resolve smaller-scale phenomena, and therefore we have to integrate them with a smaller timestep. Unfortunately, this increases the number of integration steps necessary for one simulation since the intended simulation time does not change. Additionally, the particles with small timesteps are a tiny fraction of the entire system. Such a situation worsens the scalability in parallel computing. Thus, we need to solve these problems to achieve star-by-star galaxy simulations.

A successful way to improve the scalability of astronomical simulations is Hamiltonian splitting, which allows us to integrate small-step regions separately. For example, in $N$-body simulations of a cluster in the galaxy Fujii et al. 2007 used each integration scheme for a cluster and the host galaxy derived with the Hamiltonian splitting. This method can also be used for galaxy simulations. Through Hamiltonian splitting, we need to pick up particles in the small-step regions, which must be integrated with small timesteps compared to the global step (the integration timestep for the entire galaxy). In the case of galaxy simulations, the regions inside supernova (SN) shells require the smallest steps. We need to integrate these regions for the decrease of the communication overhead.

The time evolution of SN's shell is given by an analytical solution (Sedov 1959) in isotropic and uniform interstellar medium (ISM). Here, we derive the radius $R$ of a shell with the released energy $E$ in the uniform density $\rho$ at a specific time $t$. Introducing the dimensionless similarity variable $\xi$, the radius $R$ is written as the following;

$$R(t) = \xi \left(\frac{E}{\rho}\right)^{1/5} t^{2/5}. \tag{1.1}$$

However, the real ISM is neither isotropic nor uniform. One way to improve the analytical approach is by dividing the surrounding gas into several regions and applying the analytical solution (Equation 1.1) using each density.

To improve the selection of the small-step regions, we develop a method using a deep learning method. In this study, we propose a computer vision approach to forecast the expansion of SN shells in inhomogeneous ISM using deep learning and to identify "target" particles using image processing. Here, "target" particles are the particles that will have small timesteps in the subsequent global timestep compared to the global time step.

## 2. Data Preparation

Our objective was to develop a deep learning model to predict the shell expansion of SN. We made the training data from hundreds of simulations of a SN explosion in inhomogeneous (turbulent) gas distribution using our SPH code, ASURA-FDPS (Saitoh et al. 2008; Iwasawa et al. 2016). We use SPH particles of the mass of 10 $M_\odot$ and the temperature of 100 K, which is the target resolution of our galaxy simulation. We assume a point source of explosions. The thermal energy of $10^{51}$ erg is injected into the center of gas distributions. The softening parameter is set to be 3 pc.

https://doi.org/10.1017/S1743921322001739 Published online by Cambridge University Press

**Table 1.** The dataset settings for our deep learning model. We made datasets from simulations of SN explosions in the gas cloud. We classify datasets into three types according to the mean density $\bar{\rho}$ and spacial symmetries of gas clouds.

| **Datasets** | $\bar{\rho}$ [cm$^{-3}$] | **Spatial Symmetries** | $t$ [Myr] |
|---|---|---|---|
| Fiducial | $1.864 \times 10^2$ | non-uniform & anisotropic | 0.133 |
| Spherical | $1.864 \times 10^2$ | uniform & isotropic | 0.133 |
| Sparse-Spherical | $3.728 \times 10$ | uniform & isotropic | 0.133 |

**Table 2.** The setup for mathematical morphology operations: "Dilation", "Erosion", "Gradients" and a self-made operator "Majority" (see the text). In the column of iterations, $D_3$, $D_5$, and $E$ are the numbers of iterations varied in the experiments.

| **Operators** | **kernel or threshold** | **iterations** |
|---|---|---|
| Dilation | (3,3) | 1 |
| Gradients | (3,3) | 1 |
| Dilation | (3,3) | $D_3$ (e.g. 1) |
| Dilation | (5,5) | $D_5$ (e.g. 3) |
| Majority | $\geq 2$ | - |
| Erosion | (3,3) | $E$ (e.g. 1) |
| Dilation | (3,3) | $E+1$ (e.g. 2) |
| Majority | $\geq 2$ | - |

**Table 3.** The settings for experiments. In each experiment, we use training data and test data of SN explosions in the gas cloud of the mean density $\bar{\rho}$. Results of simulations are predicted up to the result of forecast horizon $t$ in Table 1. We used the same trained model in Experiment 1, 2, and 3. The names of datasets correspond to those in Table 1.

| **Experiment** | **Training Data** | **test data** |
|---|---|---|
| Experiment 1 | Fiducial | Fiducial |
| Experiment 2 | Fiducial | Spherical |
| Experiment 3 | Fiducial | Sparse-Spherical |

We use a data format of three-dimensional (3D) volume images for deep learning. We perform 300 simulations of supernova explosions described above and obtain 20 snapshots with a timestep $dt = 7.0 \times 10^{-3}$ Myr (Fiducial). The initial conditions are density distributions time-evolved for a specific time, in which gas filamentary structures are formed after a core collapse, from a gas sphere with turbulence. By smoothing particles with SPH kernels of size depending on the local densities, we get 3D volume images composed of $32^3$ voxels† with resolutions of 1.875 pc. A sequence data are converted from each simulation containing 20 volume images.

## 3. Computer Vision Approach

Our deep learning model is based on Memory In Memory network‡ by Wang et al. (2018), which utilizes differential signals effectively to archive high predictability in 2D video prediction. We improved the original MIM in the following two points. First, we increased the internal dimension from two to three to deal with the 3D volume images. Second, the original MIM is the many-to-many model with sequential inputs and outputs. We designed our model as the one-to-many model to have one input and sequential outputs. We call our model 3D-MIM.

We predict the regions affected by a SN explosion by applying image processing. The following processing is performed for individual 2D slices of the 3D volume image. We

† The *voxel* in a 3D data is equivalent to the pixel in a two-dimensional (2D) image.
‡ https://github.com/Yunbo426/MIM

https://doi.org/10.1017/S1743921322001739 Published online by Cambridge University Press

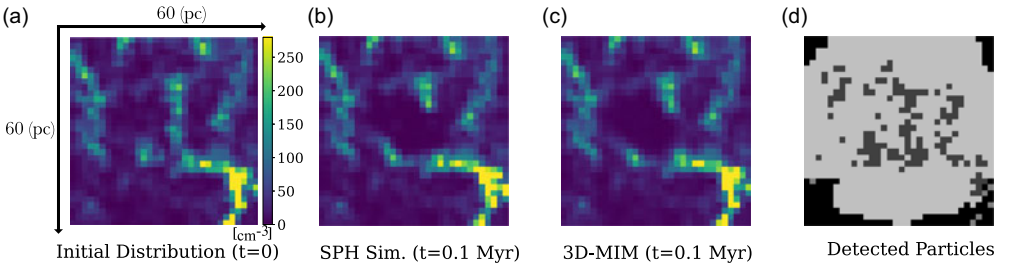


**Figure 1.** An example of forecast results by our deep learning model. One side of each panel corresponds to 60 pc. Color maps show the density distribution. The color bar and scale are the same in all panels. Panel (a) shows the initial distribution. Panel (b) shows the simulation result 0.1 Myrs after the supernova explosion (ground truth). Panel (c) shows the forecast result using our deep learning model. Panel (d) shows the forecast area and distribution of "target" particles. (Dark grey: "target" particles, Pale grey: "forecast area").

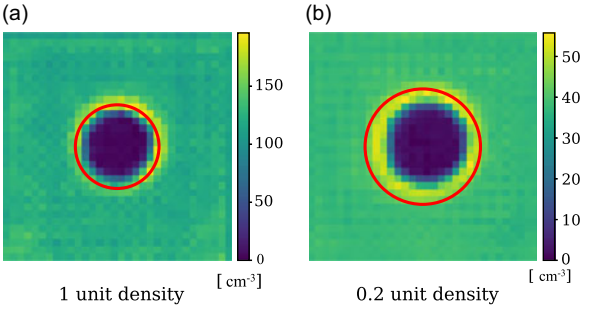


**Figure 2.** The test results of shell expansion 0.1 Myrs after an SN explosion with Experiment 2 (left) and Experiment 3 (right) in Table 3. The predictions are performed by the 3D-MIM. Red circles show the analytical solution (the Sedov-Taylor solution) of the expanding shell's radius.

calculate the pixel-by-pixel quotient of the densities just before the explosion and the densities predicted by the 3D-MIM. Then, we assign 1 to a pixel with a quotient less than 0.9 and 0 otherwise. After the binarization, we find several blobs¶ of pixels assigned 1. Among them, we give 1 to the blob whose barycenter is closest to the source of the explosion and 0 to the other pixels. We then apply three types of mathematical morphology operators: "Dilation", "Erosion", and "Gradients", and a self-made operator: "Majority". To optimize the order of operators applied to the image, we examine several different configurations with the order of operators shown in Table 2.

In practice, we perform the above process on 2D slices along each of the three axes and sum up the resulting three binarized volume images, obtaining a single volume image whose voxels have values from 0 to 3. We then process them using "Majority", which assigns 1 to each voxel with a value more than a threshold. Finally, we identify the particles inside them as those that require small timesteps in the future.

## 4. Results

### 4.1. *Forecasting the Expansion of a Supernova Shell*

Figure 1 shows an example of the forecast result using our deep learning model (Experiment 1; Table 3). With an initial distribution [Figure 1 (a)] is given as the input, our 3D-MIM forecasts the density distribution after 0.1 Myr [Figure 1 (c)].

¶ We call a chunk of four or more connected pixels a blob.

https://doi.org/10.1017/S1743921322001739 Published online by Cambridge University Press

For further examination, we test our model with uniform density distributions for which analytical solutions are obtained using Equation (1.1). Figure 2 shows the test results of Experiments 2 and 3 at 0.1 Myr in Table 3. In these cases, the background gas is uniform and isotropic media. Experiment 2 has the unit density ($1.864 \times 10^2$ cm$^{-3}$, left), i.e., the same mean density as the training data, whereas Experiment 3 has 0.2 times as unit density ($3.728 \times 10$ cm$^{-3}$, right). The red circles show the Sedov solution calculated using Equation (1.1). The forecast result agrees well with the radius predicted by the Sedov solution when the unit density of the input is similar to the trained data. However, the model is not trained with such a uniform distribution. When the input density is smaller than the trained data, the shell expansion becomes smaller than the analytical solution. This indicates that the model has not learned the scaling law. However, it is practically not a problem for our purpose that is to find particles with small timesteps, and therefore we do not need very accurate predictions of the expansion of SN's shells.

### 4.2. *Identification of Target Particles*

To fit our method with particle simulations, we need to extract particles that will evolve with smaller timesteps under the effect of SN from forecast results by deep learning. We call the particles that will require smaller timesteps due to a SN explosion "target" particles and others "non-target" particles. We define the particles of which timesteps are smaller than the global timestep and temperatures are higher than 100 K as "target".

In the analytic approach, we calculate the radius of the shell at a subsequent global timestep in each direction using each mean density of the 20 tetrahedral domains and Equation (1.1) under the icosahedron domain decomposition. On the other hand, in the computer vision approach, we forecast the variation in the spatiotemporal sequence of density distribution due to SN shell using 3D-MIM and decide the region where the density will decrease more than a threshold at the subsequent global timestep. 2D slices made of 3D volume images representing density distributions are processed using some morphology operators to predict the distribution of "target" particles. Figure 1 (d) shows the forecast region and distribution of "target" particles. It can be found that the forecast region covers most "target" particles.

Figure 3 shows the comparison of the detectability of the "target" particles at global timesteps in the (a) analytic and (b) computer vision approaches, respectively. The vertical axis shows the ratio of the number of "non-target" particles to that of "target" particles inside the forecast region. The horizontal axis shows the ratio of the number of the successfully detected "target" particles to all "target" particles.

Compared to the analytic approach, the identification rate of the computer vision approach is better (88 % to 98 % on average and 3 % to 1 % on standard derivation), and the scattering is more negligible (6.4 to 5.5 on average and 0.93 to 0.64 on standard derivation), while the "non-target" ratio is similar for both approaches. This result shows that the computer vision approach can follow the complex change in the gas structure better than the analytic approach.

The computer vision approach itself costs a larger run-time than the analytic approach. However, when combined with the computer vision approach, consistent high-resolution galaxy simulations can be achieved in a realistic time.

## 5. Conclusion

We developed a new algorithm, composed of deep learning and image processing, that improves one step in the challenge of overcoming the difficulties of high parallel computing in high-resolution galaxy simulations. Our new deep learning model successfully forecasted the density distribution after an SN explosion in the non-uniform and

https://doi.org/10.1017/S1743921322001739 Published online by Cambridge University Press

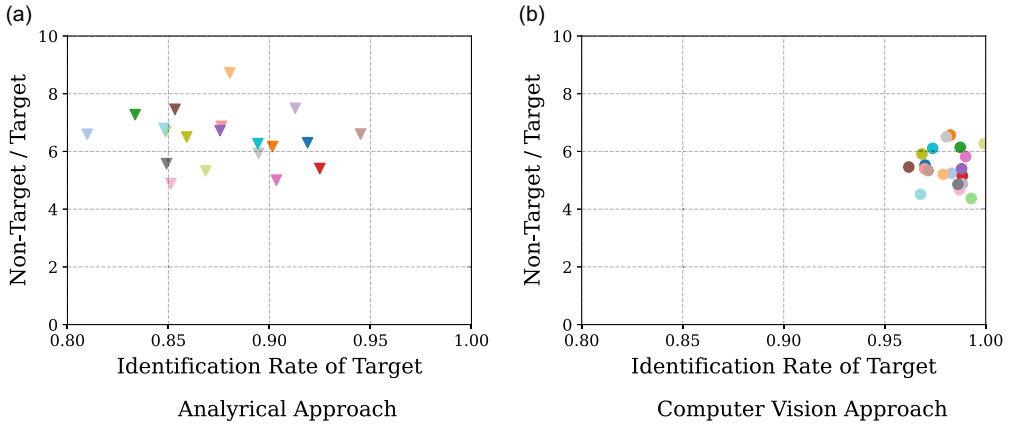


**Figure 3.** The comparison of the (a) analytical and (b) computer vision (CV) approach (Experiment 1) to identify "target" particles for the global timestep $T = 0.1$ Myr. The same color indicates the same simulation data. The vertical axis shows the number ratio of the "non-target" particles to the "target" particles inside the forecast region. The horizontal axis shows the number ratio of the successfully detected "target" particles to all "target" particles.

anisotropy gas cloud until 0.1 Myr. The forecast result for uniform distributions also agrees well with the analytic solution when the unit density of the input is similar to the training data, although the model is not trained using such a uniform distribution. In addition, by image processing, our new algorithm can identify "target" particles while excluding "non-target" particles better than the method based on the analytic solution. The average increases from around 87% to around 98% (as seen in Fig. 3). We are implementing the method under a parallel computing environment and trying to run it within a running time comparative to the analytic approach. We will include this method in our $N$-body/SPH code, ASURA-FDPS (Saitoh et al. (2008); Iwasawa et al. (2016)), and perform a star-by-star galaxy simulation using a massively parallel computer like Fugaku.

## References

Fujii, M., Iwasawa, M., Funato, Y., & Makino, J. 2007, *PASJ*, 59, 1095.
Gingold, R. A. & Monaghan, J. J. 1977, *MNRAS*, 181, 375–389.
Iwasawa, M., Tanikawa, A., Hosono, N., Nitadori, K., Muranushi, T., & Makino, J. 2016, *PASJ*, 68(4), 54.
Lucy, L. B. 1977, *AJ*, 82, 1013–1024.
Pillepich, A., Springel, V., Nelson, D., Genel, S., Naiman, J., Pakmor, R., Hernquist, L., Torrey, P., Vogelsberger, M., Weinberger, R., & Marinacci, F. 2018, *MNRAS*, 473(3), 4077–4106.
Saitoh, T. R., Daisaka, H., Kokubo, E., Makino, J., Okamoto, T., Tomisaka, K., Wada, K., & Yoshida, N. 2008, *PASJ*, 60(4), 667–681.
Sedov, L. I. 1959,.
Springel, V., Pakmor, R., Zier, O., & Reinecke, M. 2021, *MNRAS*, 506(2), 2871–2949.
Wang, Y., Zhang, J., Zhu, H., Long, M., Wang, J., & Yu, P. S. 2018, *arXiv e-prints*, arXiv:1811.07490.
Weinberger, R., Springel, V., Hernquist, L., Pillepich, A., Marinacci, F., Pakmor, R., Nelson, D., Genel, S., Vogelsberger, M., Naiman, J., & Torrey, P. 2017, *MNRAS*, 465(3), 3291–3308.

https://doi.org/10.1017/S1743921322001739 Published online by Cambridge University Press