Journal of Physics: Conference Series

PAPER • OPEN ACCESS

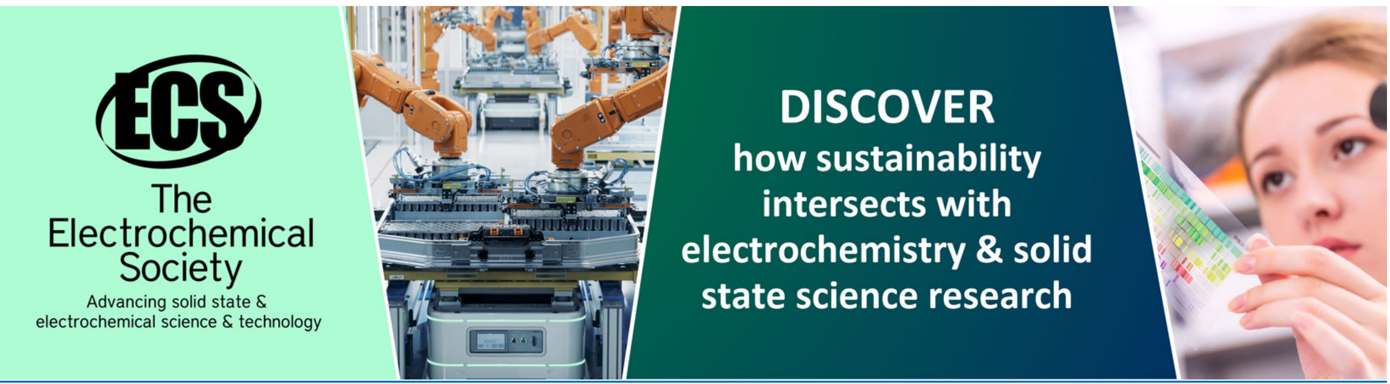
Predicting the Expansion of Supernova Shells for High-Resolution Galaxy Simulations Using Deep Learning

To cite this article: K. Hirashima *et al* 2022 *J. Phys.: Conf. Ser.* **2207** 012050

View the article online for updates and enhancements.

You may also like

- Explicit integration of extremely stiff reaction networks: partial equilibrium methods
M W Guidry, J J Billings and W R Hix
- On the evaluation of Hardy's thermomechanical quantities using ensemble and time averaging
Yao Fu and Albert C To
- Explicit integration of extremely stiff reaction networks: quasi-steady-state methods
M W Guidry and J A Harris



This content was downloaded from IP address 3.138.114.38 on 30/04/2024 at 04:21

Predicting the Expansion of Supernova Shells for High-Resolution Galaxy Simulations Using Deep Learning

K. Hirashima[1], K. Moriwaki[2], M. S. Fujii[1], Y. Hirai[3,4,5], T. Saitoh[6], and J. 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

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

Abstract. Small integration timesteps for a small fraction of the particles become a bottleneck for future galaxy simulations with a higher resolution, especially for massively parallel computing. As we increase the resolution, we must resolve physics on a smaller timescale while the total integration time is fixed as the universe age. The small timesteps for a small fraction of the particles worsen the scalability. More specifically, the regions affected by supernovae (SN) have the smallest timestep in the whole galaxy. Using a Hamiltonian splitting method, we calculate the SN regions with small timesteps using a few thousand CPU cores but integrate the entire galaxy using a shared timestep. For this approach, we need to pick up particles in regions, which will be affected by SN (the target particles) by the next global step (the integration timestep for the entire galaxy) in advance. In this work, we developed the deep learning model to predict the region where the shell due to a supernova explosion expands during one global step. In addition, we identify the target particles using image processing of the density distribution predicted by our deep learning model. Our algorithm could identify the target particles better than the method based on the analytical solution. This particle selection method using deep learning and the Hamiltonian splitting method will improve the performance of galaxy simulations with extremely high resolution.

1. Introduction

Since the timescale of the universe is much longer than that of human beings, numerical simulations are one of the ways to understand the formation history of the universe. Galaxies are a system consisting of billions of stars embedded in a dark matter halo, and the sun is a star in the Milky-Way galaxy, which has been evolved for $\sim 10\,\mathrm{Gyr}$. To simulate galaxy formation,

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

we need to treat multiple complex phenomena such as gravitational and hydrodynamic forces, radiative cooling and heating, star formation, supernova explosions, and chemical evolution.

One commonly used method for the hydrodynamics of galaxy formation is smoothed particle hydrodynamics (SPH) [1, 2], a particle-based method using gas particles smoothed with a kernel size depending on the local density. Stars in galaxies are also modeled using particles (N-body simulations).

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 simulate galaxies resolving individual stars (star-by-star simulation). In addition, we need the mass resolution for gas particles similar to that of stars.

Thanks to the development of supercomputers, a higher mass resolution has been achieved. IllustrisTNG [3, 4] is one of the simulations with the highest resolution. The mass resolution is $8 \times 10^4 M_\odot$ (about 10^{10} particles), and the simulation was performed using 25,000 CPU cores. However, the number of particles is three order-of-magnitude smaller than the number of particles necessary to resolve individual stars, even for Milky-Way-sized galaxies.

The communication overhead is a crucial problem of galaxy simulations using massively parallel computers. In parallel computing with more than thousands of CPU cores, the communication takes longer than the calculations (see Figure 63. in [5]). As we increase the resolution, we have to resolve physics on a smaller timescale, although the total integration time (the universe's age) does not change. The smaller timesteps for a small fraction of the particles worsen the scalability. Thus, we cannot reach star-by-star galaxy simulations without solving this problem.

The regions affected by supernovae (SN) have the smallest timestep in the whole galaxy. Using a Hamiltonian splitting method (e.g., Fujii *et al.* [6] and Saitoh and Makino [7]), we calculate the SN regions with small timesteps using a few thousand CPU cores but integrate the entire galaxy using a shared timestep.

For this approach, we need to pick up particles in regions, which will be affected by the subsequent global step (the integration timestep for the entire galaxy), in advance. We use a deep learning method to predict the region in which the shell due to a supernova explosion expands during one global step.

In this study, we propose a method of deep learning to identify the particles whose future timesteps will be small by using the density distribution just before a supernova explosion as the input data.

2. Methods

2.1. Extended Memory In Memory Network

We employ the Memory In Memory network (MIM) [8], which utilizes differential signals effectively to learn the non-stationary spatiotemporal changes of the video. MIM is composed of two main sequential sectors: the Convolutional Neural Network (CNN) and MIM block. CNNs have been widely used to learn correlations in images [9]. The CNN sector extracts spatial features of the image. MIM block is an extension of Recurrent Neural Network (RNN), in which hidden states are propagated recurrently to lean changes in sequential data. The MIM block makes use of a module that efficiently captures non-stationary changes as well as stationary changes.

We make two improvements to the vanilla MIM to apply for our simulation data. First, we expand the internal dimension of MIM by one dimension so that our model can deal with the three-dimensional physical quantity distribution represented by voxel. We extend the convolutional layer from two dimensions to three. Second, we improve the length of the prediction sequence. While the original MIM [8] performs a many-to-many prediction, in particular, the last ten frames from the initial ten frames, we build a one-to-many prediction model that is commonly used for, for example, the generation of music and video. Our model predicts all the subsequent gas distribution only from the initial one. The schematic diagram of our model is shown in Figure We call our model **Extended-MIM**.1.

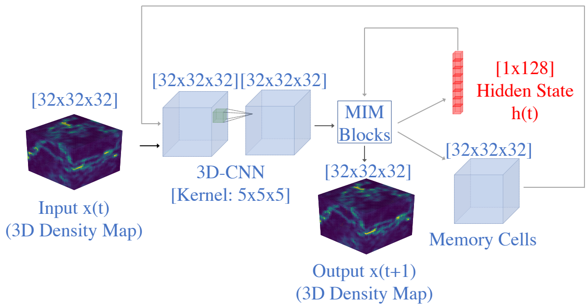


Figure 1. The architecture of our deep learning model.

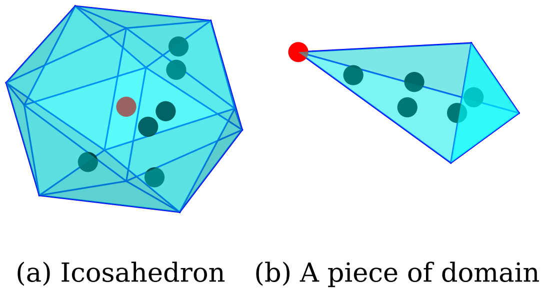


Figure 2. Domain decomposition using icosahedron. Red particle represents the center of the explosion. Darkgreen particles represent SPH particles.

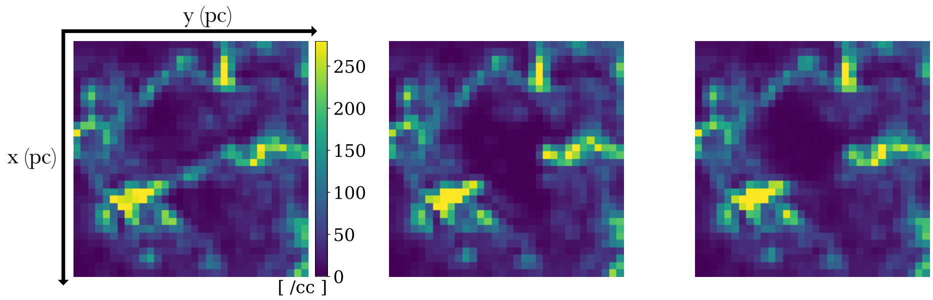
2.2. Data Preparation
As training data, we employ the outputs of simulations of SN explosions in non-uniform gas distributions. We set a density typical for the galactic disc, where the SN explosions occur. Each supernova explosion is simulated using SPH code, ASURA-FDPS (Saitoh et al. in press). By converting the particle data using the SPH kernel (like the Gaussian kernel), we obtain a time series of voxel data composed of 32x32x32 grids. A single training data is composed of time series of 3D density distribution. We generate 14,400 training data with 20 frames by taking 48 different viewing angles for 300 simulations.

2.3. Image Processing
We predict the area affected by a supernova explosion by processing the predicted future density distribution. First, we compute the ratio of the predicted density of each voxel divided by the initial density of the same voxel and collect voxels where the ratio exceeds a certain threshold. Among the regions where four or more of those voxels are connected, we pick up the one whose center of gravity is closest to the voxel where the SN explodes. We assigned a “1” to the voxels in this blob and a “0” to the others.

Another three types of image processing are performed: “Dilate”, “Erode”, and “Gradient”. Dilate assigns 1 to a pixel if at least one of its neighbors has a value of 1. Conversely, Erode assigns 0 to the pixel if it is surrounded by at least one pixel with a value of 0. “Gradient” compares the results of “Dilate” and “Erode” on an image and assigns 1 if the pixel has different values.

2.4. Defining Particles with Domain Decomposition Using Icosahedron
To evaluate our new method to predict the regions affected by SN explosion, we also test a method to determine the region using an analytic solution for the evolution of SN shells. The expansion of the supernova shell is described as a self-similar solution. It approximates the SN as a spherical point explosion in a uniform medium. Here, we derive the radius R of a SN's shell with the released energy E in the uniform density ρ at some time t. Introducing the

Journal of Physics: Conference Series **2207** (2022) 012050 doi:10.1088/1742-6596/2207/1/012050



(a) Initial Condition (t=0) (b) Simulation (t=0.1 Myr) (c) Prediction (t=0.1 Myr)

Figure 3. This figure shows an example of predicted density distributions after a SN using 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 condition just before the supernova explosion in the center. Panel (b) shows the result of the SPH simulation up to a hundred thousand years after the supernova explosion. Panel (c) shows the results predicted using our deep learning model until the same time as panel (b).

dimensionless similarity variable ξ into which the parameters of E and ρ, and the valuables of R and t are combined, the radius R is written as the following;

$$R(t) = \xi \left(\frac{E}{\rho}\right)^{4/5} t^{2/5}. \tag{1}$$

This is known as the Sedov-Taylor solution [10].

Since SN explosions rarely occur in situations under the assumption of uniform density and isotropy, we considered the anisotropy of the density distribution by dividing the region using an icosahedron. First, we align the center of the icosahedron with the center of the explosion. Next, we calculate the average density of the 20 tetrahedral regions formed by connecting each vertex of the icosahedron to the center and calculate the radius of the shell at a specific time in each direction using the 20 average densities and Equation 1. The particle inside each shell radius is identified as a particle with a small timestep in the future and is acquired.

3. Result

3.1. Predicting the shell of the Supernova Explosion

Figure 3 shows the density distributions of the input (panel (a)), simulated using the SPH code until 0.1 Myr (panel (b)), and the predicted using deep learning (panel (c)). We evaluate the performance of the trained model using Mean Structural SIMilarity (MSSIM) [11] and Mean Absolute Percentage Error (MAPE). MSSIM is the index for quality assessment based on the degradation of structural information. MAPE is the index for forecasting accuracy in statistics. When we increased training epochs, these indices converged on a good value. Our model can predict the density distribution sufficiently well based on these indices.

3.2. Identification of Target Particles and Comparison with Analytic Method

We use image processing to determine the voxels in which we expect particles with small timesteps to be included. The table 1 shows the list of parameters we used.

In Fig. 4, we present the relation between the identification rate and the ratio of the number of 'target particles' to the number of 'non-target particles'. The identification rate is defined as detected particles divided by the particles with small timesteps, which is given from the results

Table 1. The list of parameters used for image processing. After binarization, images are processed using three types of morphological operators: "Dilate", "Erode", and "Gradient". Dilate assigns 1 to the pixel if it is surrounded by at least one pixel with a value of 1. Conversely, Erode assigns 0 to the pixel if it is surrounded by at least one pixel with a value of 0. "Gradient" compares the results of "Dilate" and "Erode" on an image, and assigns 1 if the pixel has different values. In the column of iterations, D_x and E mean the number of iterations.

Morphological Operators	kernel or threshold	iterations
Dilate	(3,3)	1
Gradient	(3,3)	-
Dilate	(3,3)	D_3 (e.g. 1)
Dilate	(5,5)	D_5 (e.g. 3)
Majority	≥ 2	-
Erode	(3,3)	E (e.g.1)
Dilate	(3,3)	$E+1$ (e.g.2)
Majority	≥ 2	-

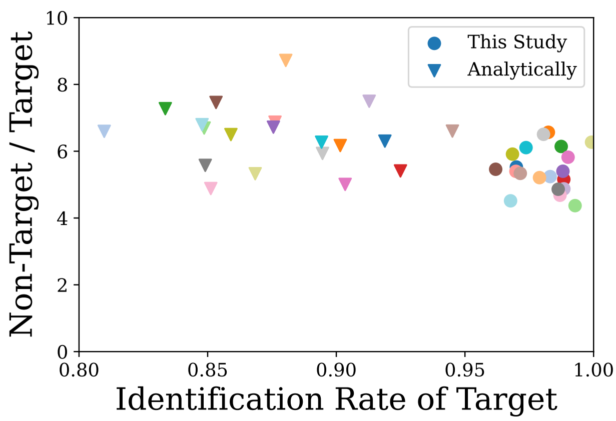


Figure 4. Identification rate of target particles versus non-target-to-target particle rate in the selected region. Circles and triangles indicate the methods using deep learning and analytical solution, respectively. The colors indicate the individual simulation data. The vertical axis shows the ratio of the number of non-target particles to the number of target particles in the enclosed region. The horizontal axis shows the ratio of non-target particles to the target particles in the enclosed region. Here, target particles have a smaller timestep than global timestep and are hotter than 100 K.

of the SPH simulations. Hereafter, 'target' particles mean the particles that will have the small timestep due to a SN by the subsequent global step, whereas 'non-target' particles mean the particles incorrectly identified as target particles. This figure shows that we detected 95% of the target particles.

For comparison, we also show the results using the icosahedron method. If we set the non-target-to-target similar to that of the deep learning method, the identification rate is a maximum of 95%. In addition, the identification rate of the icosahedron method has a large scatter. The identification rate distributes between 80% and 95%. This is probably because the analytic solution fails when the gas density distribution is highly inhomogeneous. Thus, we conclude that our deep learning method is better and more stable compared to an analytic solution.

4. Conclusion

To improve the resolution of galaxy formation simulations using SPH code, we are developing a Hamiltonian splitting scheme, in which only SN regions are integrated with timesteps smaller than that for the entire region (global step). For this method, we developed a new algorithm with a deep learning model to select particles that will have timesteps smaller than the global timestep. Our new deep learning model successfully predicted the region which is affected by a SN explosion in 0.1 Myr. By performing image processing of the predicted density distributions, our new algorithm can identify particles that will have small timesteps. We confirmed that our method could select particles with small timesteps better than that based on a self-similar analytic solution.

We in the future will include this method in our N-body/SPH code, ASURA-FDPS [12, 13], and perform a star-by-star galaxy formation simulation using a massively parallel computer such as Fugaku.

Acknowledgements

Numerical computations were carried out on Cray XC50 CPU-cluster at the Center for Computational Astrophysics (CfCA) of the National Astronomical Observatory of Japan. This work was supported by JSPS KAKENHI Grant Number 19H01933, 19J21379, 20K14532, 21J00153, 21K03614, 21K03633, 21H04499 and Initiative on Promotion of Supercomputing for Young or Women Researchers, Information Technology Center, The University of Tokyo, and MEXT as "Program for Promoting Researches on the Supercomputer Fugaku" (Toward a unified view of the universe: from large scale structures to planets, Revealing the formation history of the universe with large-scale simulations and astronomical big data). MF was supported by The University of Tokyo Excellent Young Researcher Program.

References

[1] Gingold R A and Monaghan J J 1977 *MNRAS* **181** 375–389
[2] Lucy L B 1977 *AJ* **82** 1013–1024
[3] Weinberger R, Springel V, Hernquist L, Pillepich A, Marinacci F, Pakmor R, Nelson D, Genel S, Vogelsberger M, Naiman J and Torrey P 2017 *MNRAS* **465** 3291–3308 (*Preprint* 1607.03486)
[4] Pillepich A, Springel V, Nelson D, Genel S, Naiman J, Pakmor R, Hernquist L, Torrey P, Vogelsberger M, Weinberger R and Marinacci F 2018 *MNRAS* **473** 4077–4106 (*Preprint* 1703.02970)
[5] Springel V, Pakmor R, Zier O and Reinecke M 2021 *MNRAS* **506** 2871–2949 (*Preprint* 2010.03567)
[6] Fujii M, Iwasawa M, Funato Y and Makino J 2007 *PASJ* **59** 1095 (*Preprint* 0706.2059)
[7] Saitoh T R and Makino J 2010 *PASJ* **62** 301 (*Preprint* 0908.1460)
[8] Wang Y, Zhang J, Zhu H, Long M, Wang J and Yu P S 2018 *arXiv e-prints* arXiv:1811.07490 (*Preprint* 1811.07490)
[9] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in Neural Information Processing Systems 25* ed Pereira F, Burges C J C, Bottou L and Weinberger K Q (Curran Associates, Inc.) pp 1097–1105 URL http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf
[10] Sedov L I 1959 *Similarity and Dimensional Methods in Mechanics*
[11] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 *IEEE Transactions on Image Processing* **13** 600–612
[12] Saitoh T R, Daisaka H, Kokubo E, Makino J, Okamoto T, Tomisaka K, Wada K and Yoshida N 2008 *PASJ* **60** 667–681 (*Preprint* 0802.0961)
[13] Iwasawa M, Tanikawa A, Hosono N, Nitadori K, Muranushi T and Makino J 2016 *PASJ* **68** 54 (*Preprint* 1601.03138)