Super-Resolution without High-Resolution label for Black Hole Simulations

Thomas Helfer
Institute for Advanced Computational Science, Stony Brook University
Stony Brook, NY 11794 USA
thomas.helfer@stonybrook.edu

Thomas D.P. Edwards
William H. Miller III Department of Physics and Astronomy, Johns Hopkins University
(currently working at Meta)
Baltimore, Maryland 21218, USA

Jessica Dafflon
Valence Labs,
6666 Rue Saint-Urbain Montreal, QC H2S 3H1, Canada
London, UK

Kaze W.K. Wong
Department of Applied Mathematics and Statistics, Johns Hopkins University
3400 North Charles Street, Baltimore, MD 21218, USA
Center for Computational Astrophysics, Flatiron Institute,
162 5th Ave, New York,
NY 10010, USA

Matthew Lyle Olson
Intel Labs
Santa Clara CA, USA

Abstract

The generation of high-resolution simulations is essential for advancing our understanding of the universe's most violent events, such as Black Hole mergers. However, generating Black Hole simulations is limited by prohibitive computational costs and scalability issues, reducing the simulation's fidelity and resolution achievable within reasonable time frames and resources. In this work, we introduce a novel method that circumvents these limitations by applying super-resolution techniques *without directly needing high-resolution labels* by leveraging the Hamiltonian and momentum constraints – fundamental equations in general relativity that govern the dynamics of spacetime. Our novel approach addresses the computational inefficiencies of current methods while maintaining the physical accuracy required in numerical relativity simulations. We show that our method creates a two-orders-of-magnitude reduction in numerical error and generalizes to out-of-distribution simulations.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

1 Introduction

The advent of gravitational wave astronomy has heralded a new era in astrophysics, enabling unprecedented insights into some of the universe's most spectacular events, such as Black Hole mergers and neutron star collisions. Numerical relativity (NR) simulations play a crucial role in predicting the waveforms of such phenomena and are essential for the successful analysis of data observed by gravitational wave detectors like LIGO and Virgo Abbott et al. [2016]. As the sensitivity of upcoming detectors (e.g., LISA Amaro-Seoane et al. [2017]) will increase by orders of magnitude, the demand for more accurate and diverse waveforms generated by NR simulations grows exponentially Afshordi et al. [2023]. However, existing numerical methods face challenges in meeting these demands, which could limit the scientific return on the significant financial investments made in these detectors.

Next-generation detectors will require more advanced solutions capable of handling longer, higher-resolution simulations and varied mass ratios for black hole systems. In this publication, we present a super-resolution inspired method that employs a convolutional neural network (NN) and uses constraints from general relativity to make the network physics-aware. The method is designed to be applied to current state-of-are numerical codes, and aims to reduce simulation error and enhance the accuracy of waveform predictions.

Most super-resolution techniques require high-resolutions labels for the training. However, in an NR context to get these high-resolution labels requires us to run expensive simulations, and to avoid this computational cost, we propose a framework that uses a unique loss function derived from general relativity's unique set of physical constraints. These constraints – referred to as Hamiltonian and Momentum constraints – are a crucial part of monitoring the stability of simulations. If they are not fulfilled below a threshold or show fast-growing trends, it is a strong indication of a problem in the simulation.

2 Background: Numerical Relativity

Numerical relativity (NR) provides the computational framework for simulating the complex dynamics of spacetime, such as those observed in Black Hole mergers and gravitational waves. This section outlines the core concepts of numerical relativity, offering a small overview of its theoretical underpinnings. For those interested in a more in-depth exploration, please refer to Baumgarte and Shapiro [2021, 2010], Alcubierre [2008].

2.1 Foundations of General Relativity

The theoretical backbone of NR is Einstein's general theory of relativity Einstein et al. [1916], which is described by the equation

$$G_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{1}$$

This equation describes how matter and energy (encoded in the stress-energy tensor $T_{\mu\nu}$) influence the curvature of spacetime, represented by the Einstein tensor $G_{\mu\nu}$.

The direct application of Einstein's equations in their original form is not feasible in NR simulations due to a lack of distinction between time and space. This challenge is addressed by the Arnowitt-Deser-Misner (ADM) (3+1) decomposition Arnowitt et al. [1959], a mathematical formalism that reformulates Einstein's equations into a set suitable for numerical analysis. To be also numerically stable, we use the standard CCZ4 formulation Alic et al. [2013, 2012].

Einstein's equation in the ADM decomposition gives rise to the constraint equations that we propose for our framework as following

$$\mathcal{H} := R + \frac{2}{3}K^2 - \tilde{A}_{kl}\tilde{A}^{kl} - 16\pi\rho, \tag{2}$$

$$\mathcal{M}_i := \tilde{\gamma}^{kl}\left(\partial_k \tilde{A}_{li} - 2\tilde{\Gamma}^m_{l(i}\tilde{A}_{k)m} - 3\tilde{A}_{ik}\frac{\partial_l \chi}{2\chi}\right) - \frac{2}{3}\partial_i K - 8\pi S_i, \tag{3}$$

where R is the Ricci scalar, χ, K, $\tilde{A}_{ik}$ and $\tilde{\gamma}^{kl}$ are evolution variables, ρ and S_i are energy and momentum density – both describe matter moving on the space manifold (e.g., Neutron Stars, Humans, Photons, cats). These equations $\mathcal{H}$ and $\mathcal{M}_i$ need to be equal to zero to be consistent with general

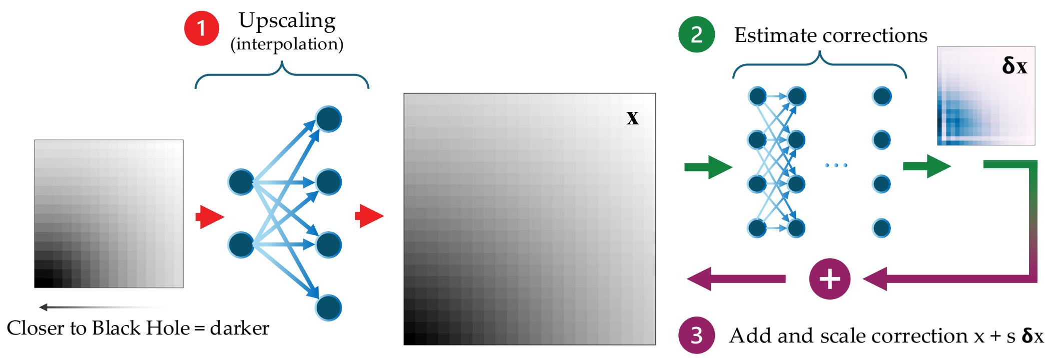

Figure 1: Schematic representation of our framework: (1) We first apply a commonly used interpolation method to up-sample our simulation, then (2) a network takes the up-sampled simulation and produces a correction δx. This correction is then (3) scaled by the hyper-parameter s and added to the up-sampled simulation. The corrected simulation results in reduced constraints violations, leading to an improved simulation.

relativity. However, this is never truly possible in numerical methods as the discretization introduces small errors. Although there are many methods that try to minimize this error by modifying the evolution equations (as was done in the CCZ4 formulation), with our framework, we introduce a physics-aware network to minimize these errors.

3 Methodology

3.1 Loss functions

In contrast to supervised ML methods – where we would use the distance between predicted and ground truth as a loss – here we use the sum of squares of the violation of Eq. 2 and Eq. 3.

$$\mathcal{L}_{\mathrm{GR}} = \sum_j \left(|\mathcal{H}(\mathbf{x}_j)|^2 + \sum_{i=0}^{D} |\mathcal{M}_i(\mathbf{x}_j)|^2 \right) , \tag{4}$$

where D is the number of spatial dimensions (three for our experiments) and $\mathbf{x}_j$ represent positions on the simulation grid. While the $\mathcal{L}_{\mathrm{GR}}$ loss incurs a higher computational cost per iteration, it converges significantly faster. In our experiments, the slower per-iteration time is roughly balanced by fewer required iterations, resulting in similar overall convergence times compared to L_1.

We defined the normalized loss that we used to evaluate the performance of our model and the baseline as

$$\text{Normalized } \mathcal{L}_{\mathrm{GR}} = \frac{\mathcal{L}_{\mathrm{GR}}(\text{Our method})}{\mathcal{L}_{\mathrm{GR}}(\text{Baseline})}. \tag{5}$$

It is important to draw a clear distinction between Physical Informed Neural Networks (PINNs) Raissi et al. [2017], which also use partial differential equations (PDEs) as a loss. While our framework takes as input an approximation of the solution and uses the physical constraints to improve it, PINNs take as input the spatial coordinates and produce the value of the PDE solution at the given coordinate. Furthermore, while PINNs need to be retrained for each new simulation, our method needs to be trained once and can then be inferred on different simulations.

3.2 Dataset generation

Our framework employs GRTeclyn (formerly GRChombo)Radia et al. [2022], Andrade et al. [2021] an established open-source codebase for the NR simulations. We simulate two equal-mass Black Holes, a common reference in numerical relativity [1], similar to how the MNIST dataset is used as a

[1]Data is available at `https://huggingface.co/datasets/anonymous/anonymous`.

benchmark in computer vision. The dataset contains time-evolution, thus including movement of the two black holes.

To be able to fit our simulations in GPU memory we subdivided it into blocks of $(16 \times 16 \times 16)$ points and 25 channels representing different evolution variables (i.e., χ, K, $\tilde{A}_{ik}$, $\tilde{\gamma}_{kl}$, α and β^i). For the model presented in this publication, we use 15G of data, which corresponds to 19208 blocks. We trained on 80% of the data and used 20% to test the in-distribution performance. To test out-of-distribution performance, we also created several independent simulations with increasing Black Hole masses.

3.3 Framework architecture

An overview of the framework architecture can be found in Fig. 1. First we up-sample our simulation from low resolutions using a standard higher order interpolator commonly used throughout Black Holes simulations Schnetter et al. [2004] [2]. This will not only be the input for the neural network that calculates the correction δx, but also will serve as the baseline. As we do not know the scale of our correction priori, we introduce a rescale factor s to reduce floating-point problems. So, our correction is

$$x + s\delta x \ , \tag{6}$$

where x is the vector of all variables $x = (\chi, K, \tilde{A}_{ik}, \dots)$. For our data, we found that $s = 10^{-4}, 10^{-5}$ works well.

Numerical Stability As the Hamiltonian and Momentum constraints (Eq. 2 and 3) are mathematically underspecified (we have 25 variables and only four equations), there are many possible solutions that the system can take. We can address this by introducing masking, changing our correction (Eq. 6) at training to

$$x + m \cdot s\delta x \ , \tag{7}$$

where m is a mask. To clarify why this approach works, consider a scenario where δx is large. If two neighboring points are compared, and one is masked while the other is not, the resulting difference between these points will be significant. This, in turn, leads to large gradients. Since gradients contribute to the $\mathcal{L}_{\mathrm{GR}}$ loss, solutions with large δx are penalized. We aim to avoid such solutions because we want δx to remain small, keeping the system close to its original state x.

Neural Network Details Since translation symmetry is inherently encoded in Einstein's gravity when using a Cartesian grid, convolutional neural networks become a natural choice for our framework. We constructed a convolutional NN composed of 4 hidden layers with 64 channels and ReLU non-linearity at the end of every hidden layer. To keep both input and output sizes the same, we employ padding. Lastly, as the gradients involved in this process can be very small, we use double precision to avoid any underflow issues. For the experiment presented in this paper, this simple architecture is sufficient. However, for larger datasets or more complex tasks, deeper networks with enhancements such as residual connections He et al. [2015] or more advanced architectures like U-Nets Ronneberger et al. [2015] may be necessary to ensure better scaling and performance.

4 Results

Our framework shows an improvement of two orders in magnitude in simulation quality and generalizes to out-of-distribution simulations (Fig. 2). We emphasize that a major advantage of our approach is not requiring any high-resolution simulation as ground truth to train the NN. Our observed increase in performance was obtained by using an NN together with a physics-aware loss.

One limitation of this paper is that our analysis is conducted post-hoc—meaning the simulation generates low-resolution data, and our framework is applied only after the simulation has finished. Ideally, the up-sampling would occur online, with our method integrated into the simulation as it runs, replacing the baseline interpolation within an adaptive mesh codebase. However, implementing an online approach presents additional challenges, such as the engineering effort required to integrate our method with the NR software. These hurdles will be addressed in future research. Additionally,

[2] Code for higher order interpolator in Pytorch is available at `https://github.com/anonymous/anonymous`.

a drawback of our model is that the trained neural network depends on the grid spacing ($\mathrm{d}x$) used during training. Therefore, to operate across the varying spacings present in an adaptive mesh solver, multiple grid spacings would need to be addressed. With our current methods, this would require training separate networks for each grid spacing.

Lastly, while this paper primarily focuses on the newly introduced loss $\mathcal{L}_{\mathrm{GR}}$, a more detailed comparison with supervised methods, including those using L_1 loss for super-resolution tasks, will be presented in an extended version of this work.

Another promising avenue worth exploring is the area of blind super-resolution, which allows for super-resolution without high-resolution labels or explicit degradation models (see Liu et al. [2021] for a detailed overview).

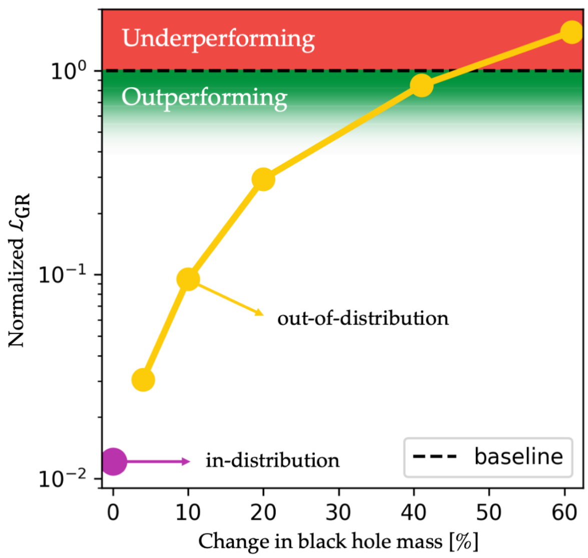

Figure 2: **Our framework (purple and yellow) outperforms the baseline (dotted black) by two orders of maginutude**. In NR simulations, the mass of a Black Hole is a parameter for defining the simulation. We evaluate the loss of the validation set in the in-distribution scenario (purple). However, we aimed to stress-test our framework by varying the Black Hole's mass, enabling us to *evaluate its ability to generalize to out-of-distribution scenarios* (yellow). Remarkably, even with a 41% variation in the Black Hole's mass, our framework still outperforms the baseline.

4.1 Conclusion

In this work, we tested the applicability of deep learning techniques to numerical relativity simulations. We showed that even without using high-resolution simulations as a ground-truth label, we could learn an NN capable of improving the simulation's quality using a physics-aware loss. This advancement will help bridge the current gap in simulation fidelity needed to analyze data from next-generation gravitational wave detectors.

Acknowledgments and Disclosure of Funding

Research reported in this publication was supported by a Postdoctoral Fellowship at the Institute for Advanced Computational Science, Stony Brook University. This work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC1920103. The Flatiron Institute is a division of the Simons Foundation.

References

B. P. Abbott et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 116(24):241103, 2016. doi: 10.1103/PhysRevLett.116.241103.

Niaesh Afshordi et al. Waveform Modelling for the Laser Interferometer Space Antenna. 11 2023.

Miguel Alcubierre. *Introduction to 3+ 1 numerical relativity*, volume 140. OUP Oxford, 2008.

Daniela Alic, Carles Bona-Casas, Carles Bona, Luciano Rezzolla, and Carlos Palenzuela. Conformal and covariant formulation of the z4 system with constraint-violation damping. *Phys. Rev. D*, 85: 064040, Mar 2012. doi: 10.1103/PhysRevD.85.064040. URL `https://link.aps.org/doi/10.1103/PhysRevD.85.064040`.

Daniela Alic, Wolfgang Kastaun, and Luciano Rezzolla. Constraint damping of the conformal and covariant formulation of the z4 system in simulations of binary neutron stars. *Phys. Rev. D*, 88: 064049, Sep 2013. doi: 10.1103/PhysRevD.88.064049. URL `https://link.aps.org/doi/10.1103/PhysRevD.88.064049`.

Pau Amaro-Seoane, Heather Audley, Stanislav Babak, John Baker, Enrico Barausse, Peter Bender, Emanuele Berti, Pierre Binetruy, Michael Born, Daniele Bortoluzzi, Jordan Camp, Chiara Caprini, Vitor Cardoso, Monica Colpi, John Conklin, Neil Cornish, Curt Cutler, Karsten Danzmann, Rita Dolesi, Luigi Ferraioli, Valerio Ferroni, Ewan Fitzsimons, Jonathan Gair, Lluis Gesa Bote, Domenico Giardini, Ferran Gibert, Catia Grimani, Hubert Halloin, Gerhard Heinzel, Thomas Hertog, Martin Hewitson, Kelly Holley-Bockelmann, Daniel Hollington, Mauro Hueller, Henri Inchauspe, Philippe Jetzer, Nikos Karnesis, Christian Killow, Antoine Klein, Bill Klipstein, Natalia Korsakova, Shane L Larson, Jeffrey Livas, Ivan Lloro, Nary Man, Davor Mance, Joseph Martino, Ignacio Mateos, Kirk McKenzie, Sean T McWilliams, Cole Miller, Guido Mueller, Germano Nardini, Gijs Nelemans, Miquel Nofrarias, Antoine Petiteau, Paolo Pivato, Eric Plagnol, Ed Porter, Jens Reiche, David Robertson, Norna Robertson, Elena Rossi, Giuliana Russano, Bernard Schutz, Alberto Sesana, David Shoemaker, Jacob Slutsky, Carlos F. Sopuerta, Tim Sumner, Nicola Tamanini, Ira Thorpe, Michael Troebs, Michele Vallisneri, Alberto Vecchio, Daniele Vetrugno, Stefano Vitale, Marta Volonteri, Gudrun Wanner, Harry Ward, Peter Wass, William Weber, John Ziemer, and Peter Zweifel. Laser interferometer space antenna. 2017. URL `https://arxiv.org/abs/1702.00786`.

Tomas Andrade et al. GRChombo: An adaptable numerical relativity code for fundamental physics. *J. Open Source Softw.*, 6(68):3703, 2021. doi: 10.21105/joss.03703.

R. Arnowitt, S. Deser, and C. W. Misner. Dynamical structure and definition of energy in general relativity. *Phys. Rev.*, 116:1322–1330, Dec 1959. doi: 10.1103/PhysRev.116.1322. URL `https://link.aps.org/doi/10.1103/PhysRev.116.1322`.

Thomas W Baumgarte and Stuart L Shapiro. *Numerical relativity: solving Einstein's equations on the computer*. Cambridge University Press, 2010.

Thomas W Baumgarte and Stuart L Shapiro. *Numerical relativity: starting from scratch*. Cambridge University Press, 2021.

Albert Einstein et al. The foundation of the general theory of relativity. *Annalen Phys*, 49(7):769–822, 1916.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. URL `https://arxiv.org/abs/1512.03385`.

Anran Liu, Yihao Liu, Jinjin Gu, Yu Qiao, and Chao Dong. Blind image super-resolution: A survey and beyond, 2021. URL `https://arxiv.org/abs/2107.03055`.

Miren Radia, Ulrich Sperhake, Amelia Drew, Katy Clough, Pau Figueras, Eugene A. Lim, Justin L. Ripley, Josu C. Aurrekoetxea, Tiago França, and Thomas Helfer. Lessons for adaptive mesh refinement in numerical relativity. *Class. Quant. Grav.*, 39(13):135006, 2022. doi: 10.1088/1361-6382/ac6fa9.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. *arXiv preprint arXiv:1711.10561*, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. URL `https://arxiv.org/abs/1505.04597`.

Erik Schnetter, Scott H. Hawley, and Ian Hawke. Evolutions in 3-D numerical relativity using fixed mesh refinement. *Class. Quant. Grav.*, 21:1465–1488, 2004. doi: 10.1088/0264-9381/21/6/014.