Message Passing is all you need!

Lucas Tesán[1*], Mikel M. Iparraguirre[1*], David González[1], Pedro Martins[1,2] and Elias Cueto[1]

[1]ESI Group-UZ Chair of the National Strategy of AI.
Aragón Institute of Engineering Research, Universidad de Zaragoza, Zaragoza, Spain.

[2] Aragonese Foundation for Research and Development (ARAID), Aragón, Spain
* Equal Contributors

Abstract

This study presents a novel approach to deepening the physical understanding of message passing architectures within simulations of physical systems. By tailoring the design to the underlying nature of hyperbolic, parabolic, and elliptic partial differential equations (PDEs), the method ensures effective information propagation throughout the computational domain. This alignment between Graph Neural Network (GNN) architecture and the governing physical principles enhances both the accuracy and robustness of simulations, enabling more efficient and high-fidelity modeling across diverse physical regimes.

Neural PDEs Solvers

In his renowned essay "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," Eugene Wigner (1902–1995) pondered the mysterious and astonishing power of mathematics to describe and predict real-world phenomena. This capability has solidified the dominance of traditional numerical methods as indispensable tools for solving complex differential systems. Partial differential equations (PDEs), serving as mathematical translations of physical phenomena, often lack trivial or computationally tractable solutions.

However, the advent of artificial intelligence has abruptly shifted this paradigm. The optimization of solutions from data—understood as discrete representations of physical behaviors—combined with the universal approximation capabilities of neural networks, has unlocked unprecedented potential in simulation and efficiency. The use of neural networks to solve PDEs is not new, with physics-informed neural networks [1] being perhaps its most prominent example. Nevertheless, the search for neural simulation techniques that rely less on initial conditions, boundary constraints, and PDE parameters has driven the development of novel methodologies. Among the most promising are those based on geometric deep learning, as they can mimic the cell or mesh structures typical of classical discretization methods (finite elements, finite volumes, finite differences, etc.) [2].

Message-passing neural PDE solvers, as the state-of-the-art in geometric learning, offer remarkable flexibility in obtaining highly accurate and mesh-independent solutions. They replicate the local structure of most physical conservation law equations [3–7], functioning as a spatial operator capable of capturing even the most nonlinear behaviors. However, their hyperparameterization is complex, and their interpretability remains an even greater challenge. Few studies delve into understanding the nature of their efficiency, and even fewer attempt to relate their learning mechanisms to the intrinsic properties of the systems they aim to model.

Message Passing

Message passing in Graph Neural Networks (GNNs) refers to the process by which nodes exchange information through their connecting neighbors, given a predefined connectivity. While this mechanism has proven effective in practice, the theoretical basis for how message-passing algorithms reconstruct solutions remains only partially understood.

Some studies, such as [4], show that message passing can approximate classical finite difference schemes. However, their approach often requires many more iterations per time step than theoretically expected, resulting in nearly dense stiffness matrices.

Other empirical findings suggest that the messages exchanged during propagation encode

meaningful physical information about the problem. As a result, it has been argued that the number of iterations per time step must be sufficient to ensure that this information reaches all parts of the mesh [8].

Results

In this work, we demonstrate that the number of iterations required to achieve the maximum accuracy provided by the network depends primarily on the type of equation governing the physical phenomenon. For hyperbolic equations, an information wave propagates over time through both the physical medium and the network. Drawing inspiration from the Courant-Friedrichs-Lewy (CFL) condition in finite difference schemes, we establish that the necessary condition for attaining asymptotic network accuracy is ensuring the number of message iterations per time step keeps information ahead of the physical wave propagating through the network. If the physical wave outpaces the information propagation, the network's outputs become invalid. Conversely, increasing the number of iterations per time step results in highly stable and precise solutions.

For parabolic and elliptic equations, physical changes propagate instantaneously. For example, in Fourier's heat transfer equation, a heat flux immediately induces a temperature rise across the entire domain, with the magnitude of the increase decaying with distance from the flux application point. Consequently, we show that for elliptic or parabolic systems, the number of message iterations per time step must be sufficient to propagate information to the farthest relevant node in the network. Beyond this threshold, additional iterations do not enhance accuracy, as demonstrated in Fig. 1.

The study also reveals that generalization depends on whether message passing iterations (MPIs) match the problem's causal structure. Physics-guided problems (with local interactions) generalize well to new geometries (Fig 2.), while geometry-guided problems require sufficient MPIs to cover the domain—otherwise, performance degrades on larger or unseen domains.

Finally, we demonstrate that the latent space of these complex architectures contains interpretable features, which we correlate with key behavioral patterns observed in our study, as illustrated in Fig. 3. This opens a new path toward interpretability in models once deemed too chaotic to analyze—previously dismissed as mere 'black boxes.'

Conclusions

This study explores how message-passing iterations in Graph Neural Networks affect their ability to model physical systems governed by Partial Differential Equations. The findings show that MPIs are not just a computational tool but a physics-aligned process that significantly impacts accuracy, generalization, and stability.

Overall, this work provides a rigorous framework for designing physics-aware GNNs, linking the message passing algorithm to the underlying physics and geometry of the system. These insights offer practical guidelines for architecture selection and hyperparameter tuning, advancing GNNs in scientific machine learning and physics simulations.

REFERENCES

[1] KARNIADAKIS, George Em, et al. Physics-informed machine learning. Nature Reviews Physics, 2021, vol. 3, no. 6, p. 422-440.

[2] BRONSTEIN, Michael M., et al. Geometric deep learning: Going beyond Euclidean data. IEEE Signal Processing Magazine, 2017, vol. 34, no. 4, p. 18-42.

[3] ZHAO, Yingxue, et al. A review of graph neural network applications in mechanics-related domains. Artificial Intelligence Review, 2024, vol. 57, no. 11, p. 315.

[4] BRANDSTETTER, Johannes, WORRALL, Daniel y WELLING, Max. Message passing neural PDE solvers. arXiv preprint arXiv:2202.03376, 2022.

[5] PFAFF, Tobias, et al. Learning mesh-based simulation with graph networks. In: International Conference on Learning Representations, 2021.

[6] LAM, Remi, et al. Learning skillful medium-range global weather forecasting. Science, 2023, vol. 382, no. 6677, p. 1416-1421.

[7] BODNAR, Cristian, et al. A foundation model for the earth system. Nature, 2025, p. 1-8.

[8] ALON, Uri y YAHAV, Eran. On the bottleneck of graph neural networks and its practical implications. arXiv preprint arXiv:2006.05205, 2020.

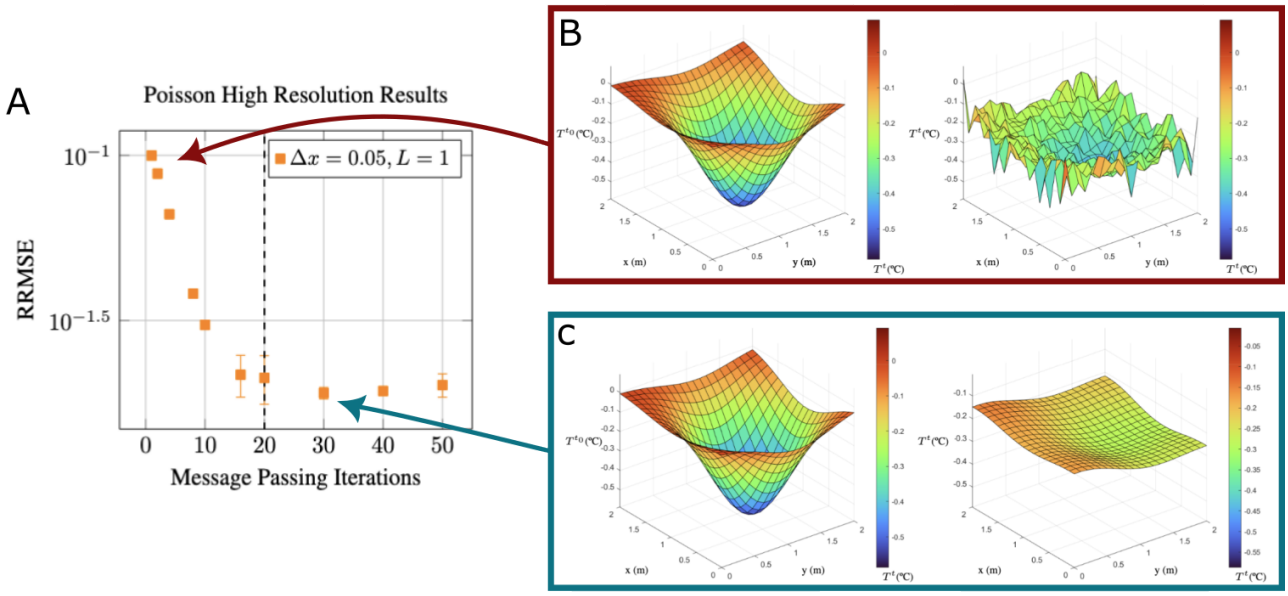

Figure 1: Solution of a diffusion system (parabolic). (A) Relative error as a function of the number of message steps used in each time increment. The vertical dashed line marks the lower bound proposed for this problem. If the number of message steps applied is less than the proposed number (B), the predictions of the neural solver fail immediately and are corrupted within a few time steps. On the contrary, if the number of message passes is larger than the limit proposed in this work (C), the results are stable for a large number of time integration steps. Time steps 3 and 7 are shown both in (C) and (B).

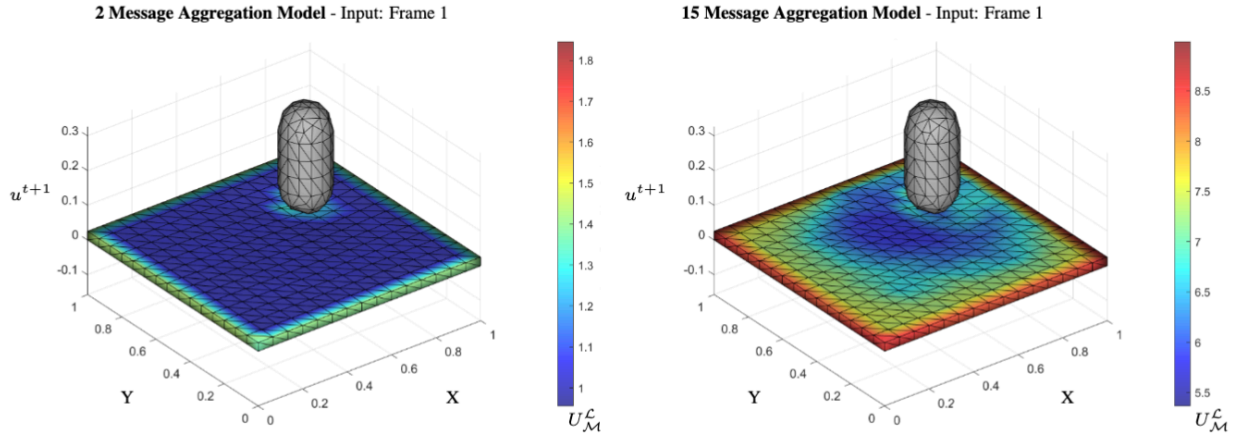

Figure 2: Visualization of latent pseudo-attention maps during the first prediction step of the plate elliptic collision problem. The color scale represents relative normalized values: values closer to one indicate minimal change in the latent representation during the aggregation process, while higher values reflect significant variation. The model with 15 aggregation steps exhibits stronger activation around the hole region and demonstrates a better understanding of the actuator's influence. In contrast, the model with only 2 aggregation steps shows poor representation, with central plate regions lacking information from both boundary conditions and the actuator.

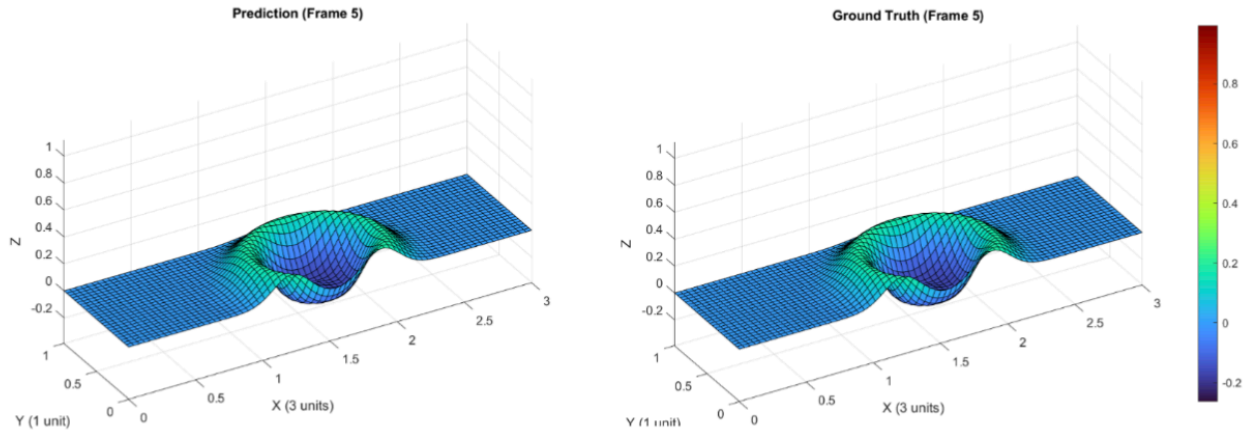

Figure 3: Extrapolation results of the hyperbolic neural solver on an unseen domain. Although the model was trained on a 1 × 1 1×1 plate, it successfully generalizes to a domain extended threefold in the X-axis and 1.5 times in temporal rollout. This capability is enabled by the local nature of the Graph Neural Network (GNN), allowing it to extrapolate effectively in both space and time despite the larger, untrained configuration.