MoriNet: A Machine Learning-based Mori-Zwanzig Perspective on Weakly Compressible SPH

Rene Winchenbach, Nils Thuerey
Technical University of Munich
Physics-based Simulations Group
Garching, Germany
rene.winchenbach@tum.de, nils.thuerey@tum.de

This work investigates the relationship between weakly compressible and incompressible Smoothed Particle Hydrodynamics (WCSPH and ISPH) using the Mori-Zwanzig formalism to further the development of efficient data-driven models. WCSPH approximates incompressible flows under certain conditions, e.g., low Mach number and density variations. We demonstrate that density fluctuations act as essential variables, primarily influencing the short-time evolution of the fluid, while the long-time behavior converges towards incompressible dynamics. Consequently, we posit that the importance of density as a feature for neural networks trained to emulate WCSPH simulations varies with timescale: dominating at short times and diminishing at longer times. Using the Mori-Zwanzig formalism, we treat particle positions and velocities as the relevant variables, implicitly modeling the density field through a memory term. Our results confirm this timescale dependence, demonstrating its implications for training neural networks and providing a theoretical justification for temporal coarse-graining approaches that simplify the learning task by focusing on longer timescales.

I. Introduction and Related Work

In Computational Fluid Dynamics (CFD), the simulation of incompressible flows, particularly in hydrodynamics, is of significant importance and presents significant challenges, both from a methodological and computational perspective. Weakly Compressible simulation schemes offer a computationally more attractive alternative by introducing small and limited compressibility into the system, which reduces the underlying governing equations from elliptic to hyperbolic. This reduction enables the use of explicit time integration schemes that avoid iterative pressure solvers and thus significantly reduces computational cost and memory requirements. As within the field of Smoothed Particle Hydrodynamics (SPH), weakly compressible formulations have found significant adoption, investigating the relation of these assumptions to the modeled fluid behavior is important.

We investigate this relationship between WCSPH and ISPH from a Mori-Zwanzig perspective, motivated by the need to reduce feature complexity in data-driven machine-learning approaches for surrogate modeling as including density requires not just an additional feature but also requires a recurrent to learn the temporal and spatial evolution of this feature. Within the Mori-Zwanzig formalism, the full dynamics of a system are projected onto a subspace of relevant variables, leading to an effective equation of motion for the relevant variables. A key insight is that even if the original system is Markovian, i.e., the current state fully determines the next state of the system, the reduced system can be non-Markovian, requiring a memory term to capture the influence of past states Woodward et al. 2023.

We now propose that for fluid dynamics, specifically for Weakly Compressible Lagrangian systems, particle positions and velocities are the relevant variables, whereas the density field can be modeled through a history of these variables. This is motivated by the idea that weak compressibility is an approximation of an incompressible system where the density field is constant and equal to the rest density. We investigate this idea by training neural networks to make short-term and long-term predictions of fluid dynamics using a WCSPH-based dataset and compare the behavior with and without memory terms. Before going into further details on our research, we will briefly recap neural networks and SPH.

Machine Learning has gained significant attention over the past few years due to the computational resources available today enabling novel and scalable solutions in many fields, including CFD (Alkin et al. 2024). The learning task in many Partial Differential Equation (PDE) based problems involves learning a surrogate solver to replace a classical solver as a primarily drop-in replacement; however, training such models can be challenging. However, adopting machine learning-based solutions to these problems has been relatively slow as, e.g., verifying that machine learning models will generalize properly to unseen initial conditions is an unsolved problem. Notably, addressing these limitations requires not just empirical solutions by increasing computational resources but also theoretical insights.

Graph Neural Networks (GNNs) (Sanchez-Gonzalez et al. 2020) are a natural fit for Lagrangian simulations, especially SPH, as the SPH kernel operations can be directly interpreted as message-passing steps in a GNN (Winchenbach et al. 2024). Particularly, Continuous Convolutions (CConv) (Ummenhofer et al. 2019) offer a generalized formulation of Convolutional Neural Networks (CNNs) that uses an underlying GNN-like architecture and closely maps to the SPH formalism, which makes it a strong inductive bias. Accordingly, this bias enables

much smaller networks to perform similarly to GNNs with significantly more parameters (Ummenhofer et al. 2019). However, learning Lagrangian fluid dynamics is still a challenging task, especially within the context of engineering (Alkin et al. 2024).

SPH is a Lagrangian simulation method initially developed in an astrophysics context as a purely Lagrangian approach (Monaghan 1994) and has since been adopted in a broad range of applications. SPH, at its core, works by using particles with positions $\mathbf{x}_i$, velocities $\mathbf{u}_i$, and constant masses m_i and tracking them over time. Operations are then performed as convolutions using a spatially compact kernel function $W(\mathbf{r}, h)$ (Dehnen et al. 2012) modeling the interaction of two neighboring particles based on their distance x_{ij}, where $|\mathbf{x}_{ij}| \geq h$ results in no interaction. An SPH operation then is a continuous convolution for a position in space $\mathbf{x}$ based on *neighboring* particles j as

$$\langle q\langle(\mathbf{x}\rangle) = \sum_j \frac{m_j}{\rho_j} q_j W\left(\mathbf{x}_j - \mathbf{x}, h\right) \tag{1}$$

The numerical accuracy of the overall approach heavily relies on the accuracy of the individual convolutions where non-isotropic particle distributions may lead to issues such as gradients of constants being non-zero (Price 2012), i.e., $\nabla 1 \neq 0$. Consequently, ensuring that the particle distribution remains of high quality has a direct and tangible influence on the accuracy of the overall scheme, and high-order accuracy is only possible with high-quality distributions (Vacondio et al. 2021). However, this generally requires moving away from a purely Lagrangian to a quasi-Lagrangian approach (Oger et al. 2016), creating several key differences. While in a purely Lagrangian context, the density field could be modeled based on the current distribution of mass, i.e., we apply the SPH operator to the masses as

$$\rho(\mathbf{x}) = \sum_j m_j W\left(\mathbf{x}_j - \mathbf{x}, h\right), \tag{2}$$

this is not possible in a quasi-Lagrangian system where there is a difference between the physical particle velocity and their actual trajectory via a transport velocity. Consequently, the density field in such a system needs to necessarily be modeled as a time-dependent problem, i.e., we initialize the system at some initial density ρ^0 and integrate the density in time via the momentum equation $\frac{d\rho}{dt} = \nabla \cdot \mathbf{u}$ (Sun et al. 2019). However, this requires adding the density as both an input to the system and the integration of density as a task for the solver, making the task of learning a surrogate solver significantly more challenging.

II. Density as Hidden Information

To train our neural networks we used a GNN-based architecture, see Winchenbach et al. 2024, with 7 message passing layers and 64 vertex features per layer, where we added an input and output decoder similar to Brandstetter et al. 2022 to improve the network, and overall used approximately 400 thousand parameters. As a dataset, we use random initially divergence-free velocity fields in a periodic domain in 2D ($[-1, 1]^2$) with 128^2 particles with a maximum initial velocity of $1\frac{m}{s}$ and, simulated using δ^+-SPH with a symplectic Euler integration scheme using a fixed timestep of $\Delta t = 1$ms. The dataset, solver and neural network framework are available as open source at https://github.com/tum-pbs/SFBC. The learning task then is to, given the current positions $\mathbf{x}_i^t$ and velocities $\mathbf{u}_i^t$, predict the positions and velocities after n timesteps, i.e., $\mathbf{x}_i^{t+n}$ and $\mathbf{u}_i^{t+n}$, where n is the temporal coarse-graining factor. We then trained a networks for different temporal coarse-graining

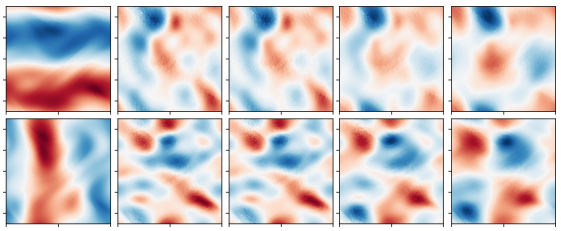
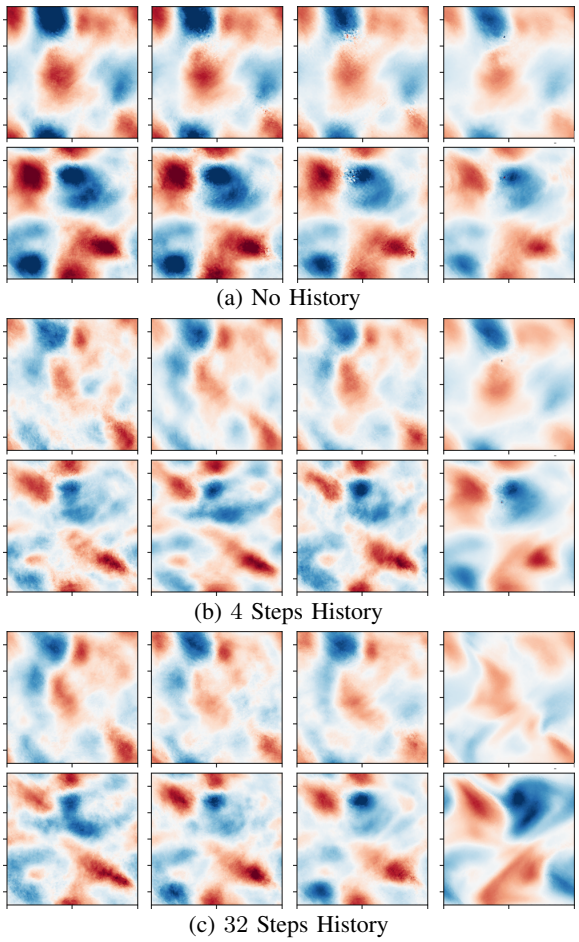


Fig. 1: This figure shows the input velocity field (leftmost column) and the accumulated acceleration f.l.t.r. for a temporal coarse-graining of 1, 2, 4, and 16 : 1 with color indicating negative (blue) to positive (red) values.



(a) No History

(b) 4 Steps History

(c) 32 Steps History

Fig. 2: Network predictions for different amounts of history length and temporal coarse-graining scale. Each subfigure color codes acceleration in x (top) and y (bottom row) with columns f.l.t.r. representing temporal coarse-graining of 1, 2, 4 and 16 : 1.

ratios and history lengths, where a history length of H was realized by adding the previous H velocities of each particle as additional features.To verify our network setups, we also trained a reference network that is given the density field as input explicitly, and trained the same network on data generated on the same initial conditions but with an incompressible SPH scheme, both yielded good results but are beyond our scope.

Our Hypothesis is based on the Mori-Zwanzig perspective where a network g, that is not given the current density field ρ^t, for a given combination of temporal coarse-graining n and history length H can only converge to the ground truth behavior if a sufficiently long history H is provided for the given ratio n. Furthermore, due to the weak compressibility argument, we expect that the behavior of a network that is not given a sufficiently long history would learn the long-term behavior instead of the short-term behavior, as the long-term behavior should be an approximation of the behavior for constant density. We evaluated each trained network for the same input state (see Fig. 1) and compared the predictions, (see Fig. 2).

Based on the results for no history length, see Fig. 1a, we can clearly observe that the predicted acceleration is, structurally, very similar for all temporal coarse-graining ratios, where for a ratio of $n = 16$ the predicted acceleration is very similar to the ground truth data. Using a history length of 4, see Fig. 1b, we observe significantly different predictions for the different coarse-graining ratios where the network trained on $n = 2$ gives a close approximation of the ground-truth, whereas $n = 1$ does not yield a good approximation. Note that the results for $n = 16$ show a decrease in approximation quality as the network is provided with significantly more input features that are not relevant, which from a Mori-Zwanzig perspective is similar to adding noise to the system, making the learning task harder. Increasing the history length to 32 exacerbates this behavior but yields a good approximation for all other coarse-graining ratios. Overall, these results clearly indicate that there is a tradeoff between history length H and temporal coarse-graining ratio n, providing some empirical evidence for our predictions based on the Mori-Zwanzig formalism and highlight how temporal coarse-graining can be used to avoid considering density evolution over time for long-term predictions.

III. Conclusions and Outlook

Our results clearly show that the importance of density as a feature, and thus as a learning task, significantly varies based on the timescale. Accordingly, learning surrogate neural networks that make short-term predictions, i.e., learning to replicate every timestep of a reference WCSPH trajectory, requires either density information as a feature, and thus integration of density as a learning task, or a history of particle states to recover the density information. When we shift the learning task from short-term predictions to long-term predictions, training a neural network to make these predictions can be done with shorter and shorter histories until the density information can be completely removed as an input feature. Consequently, our research shows that temporal coarse-graining and long-term predictions via neural networks are not just motivated by reducing computational requirements (as they require fewer invocations of the neural network), but that they also simplify the learning task. In the future, we aim to train neural networks to show how the memory length and temporal coarse-graining scale directly influence the behavior of a recurrent neural network as a surrogate solver.

References

Alkin, B., A. Fürst, S. Schmid, L. Gruber, M. Holzleitner, and J. Brandstetter (2024). "Universal Physics Transformers". In: *NeurIPS*.

Brandstetter, J., D. E. Worrall, and M. Welling (2022). "Message Passing Neural PDE Solvers". In: *The Tenth International Conference on Learning Representations, ICLR*. OpenReview.net.

Dehnen, W. and H. Aly (2012). "Improving convergence in smoothed particle hydrodynamics simulations without pairing instability". In: *Monthly Notices of the Royal Astronomical Society* 425.2, pp. 1068–1082.

Monaghan, J. J. (1994). "Simulating free surface flows with SPH". In: *Journal of computational physics* 110.2.

Oger, G., S. Marrone, D. Le Touzé, and M. De Leffe (2016). "SPH accuracy improvement through the combination of a quasi-Lagrangian shifting transport velocity and consistent ALE formalisms". In: *Journal of Computational Physics* 313.

Price, D. J. (2012). "Smoothed particle hydrodynamics and magnetohydrodynamics". In: *Journal of Computational Physics* 231.3, pp. 759–794.

Sanchez-Gonzalez, A., J. Godwin, T. Pfaff, R. Ying, J. Leskovec, and P. Battaglia (2020). "Learning to simulate complex physics with graph networks". In: *International conference on machine learning*. PMLR.

Sun, P., A. Colagrossi, S. Marrone, M. Antuono, and A.-M. Zhang (2019). "A consistent approach to particle shifting in the δ-Plus-SPH model". In: *Computer Methods in Applied Mechanics and Engineering* 348, pp. 912–934.

Ummenhofer, B., L. Prantl, N. Thuerey, and V. Koltun (2019). "Lagrangian fluid simulation with continuous convolutions". In: *International Conference on Learning Representations*.

Vacondio, R., C. Altomare, M. De Leffe, X. Hu, D. Le Touzé, S. Lind, J.-C. Marongiu, S. Marrone, B. D. Rogers, and A. Souto-Iglesias (May 2021). "Grand challenges for Smoothed Particle Hydrodynamics numerical schemes". In: *Computational Particle Mechanics* 8.3, pp. 575–588.

Winchenbach, R. and N. Thuerey (2024). "Symmetric Basis Convolutions for Learning Lagrangian Fluid Mechanics". In: *12th International Conference on Learning Representations*.

Woodward, M., Y. Tian, A. T. Mohan, Y. T. Lin, C. Hader, D. Livescu, H. F. Fasel, and M. Chertkov (2023). "Data-driven Mori-Zwanzig: Approaching a reduced order model for hypersonic boundary layer transition". In: *AIAA SCITECH 2023 Forum*, p. 1624.