Statistical Mechanics Informed Neural Networks for Solar Wind modeling

Jorge D. Enciso

November 13, 2024

Abstract

Contents

1 Introduction

The solar wind phenomena is a vastly known event that arises from ionized solar outbursts into the stellar medium, commonly stimulating planets' magnetosphere. [5] The first record of a related event dates from 1859, popularly named "The Carrington event" after the English astronomer Richard Carrington (1826, 1875), who settled the intuition upon solar flares and geomagnetic fluctuations. Individuals from all around the glove sighted the Northern Lights, subtle variations in atmosphere's hues' tones, caused by ionized particles flowing through the ionosphere.

Several attempts to model Solar Wind's dynamics can be empirically pinpointed : analytical methods [2], numerical modelling [4, 11, 18], and machine learning approaches [1, 8, 10, 14, 15]. All of them with different purposes.

Traditional models and methodologies neglect the anysotropic nature of the Solar Wind. The Magnetohydrodynamics Model assumes a Boltzmann-Maxwellian distribution, doted with standard thermodynamical behavior. On the other hand, the Solar Wind is a collisionless plasma, precluding its inclusion solely within those boundaries. Thus, the present work emphasizes the usage of learnable neural operators engineered to analytically and empirically fit kinetic descriptions.

Moreover, kinetic frames based on Maxwellian electromagnetism tend to be computationally expensive for the microscopic and macroscopic frame:

$$n_\alpha(\mathbf{r}, t) := \int f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3 v \tag{1}$$

$$\tau(r, t) := \sum q_\alpha n_\alpha \tag{2}$$

$$\mathbf{u}_\alpha(\mathbf{r}, t) := \frac{1}{n_\alpha(\mathbf{r}, t)} \int v f_\alpha(\mathbf{r}, v, t) d^3 v \tag{3}$$

$$j(\mathbf{r}, t) := \sum q_\alpha n_\alpha \mathbf{u}_\alpha \tag{4}$$

where n_α is the particle density, u_α the particle's drift velocity, $\tau(r, t)$ is the charge density, $j(r, t)$ is the current density, and $f_\alpha(r, v, t)$ is the Boltzmann's density function.

Yet modern physics informed machine learning research schemes enable faster inference time and less computational load: embedding systems' equa-

3

tions within the loss function and constraining network's backbone based on empirical approaches.

2 Related Work

2.1 Full compressible 3D MHD simulation of solar wind

On the matter of Solar Wind modeling, several numerical approaches are developed to understand near-Sun physics given certain boundary conditions. This research [11] creates a magnetohydrodynamical modeling scheme to understand these interactions under certain restrictions that approach the conditions from the transition from the photosphere to the heliospheric distance of 27 solar radii. In this case, the current work is seeking more general descriptors of the Solar Wind, with real L1-Lagrange data as a cornerstone for the modeling scheme to finally get a kinetic description of the phenomena.

2.2 Physics informed Neural Networks applied to the description of wave-particle resonance in kinetic simulations of fusion plasmas

Regarding Vlasov-Maxwell modeling with physics informed neural networks, this paper constitutes the first machine learning approach intersecting both this theory and the novel PDE solving scheme proposed by PINNs. [9] It employs the model onto fusion modeling within the colisionless constraint. It constitutes an important basis for the present paper, offering a perspective of the efficiency and satisfaction offered by using these methods on alternative fields.

They use the VOICE code to numerically solve the Vlasov-Poisson system of equations, and use this data as ground truth for the Physics Informed Neural Network. Given that it's an Integro-Differential equation, they advocate for the usage of I-PINN (integrable-PINN) [] which uses the fundamental theorem of Calculus to approach the integrable part. This integration approach will be used by the present work as well to compute several constraints tied to statistical mechanics.

2.3 Physics-Informed Neural Networks for Solar Wind Prediction

This research [**?**] is regarded as one of the first physics informed approaches towards geoeffective solar wind prediction. It uses L1-Lagrange spacecraft readings and enforces the Ohm's Law for an ideal plasma, following the Ideal MHD model for space plasma. It uses a wide variety of model architectures (GRU, LSTM, 2D CNN ResNets, etc.) to test their effectiveness under this task.

3 Kinetic models and general constraints

In order to fully assume the magnetohydrodynamical model, interactions must converge to the Maxwell-Boltzmann thermodynamical distribution for certain conditions, implying a high colisionality degree, contrary to empirical observations of the Solar Wind. Therefore, the utilization of such a constrictive framework must be carefully evaluated before setting it up as a modeling constraints.

Moreover, trying to answer quasy-isotropic behavior on near-earth L1-Lagrange point, different research papers assume the bi-Maxwellian or Kappa temperature distributions [13, 16, 19], carrying out the Maxwell-Boltzmann thermodynamical distribution. Nevertheless, this assumption seems overly ideal. As a possible option, the inclusion of numerical distributions could bring insights into the interactions between plasma ionized particles.

That's why, this work departs from that approach, seeking the description of Solar Wind kinetic model without magnetohydrodynamical systems nor temperature distributions, fully describing these interactions with a physics informed data-driven structure as a substitution of empirical assumptions.

3.1 Boltzmann's Equation

In the context of kinetic plasma theory, various statistical descriptors are applied, provided they make appropriate assumptions about particle interactions. Boltzmann's equation is often used to describe the evolution of the particle distribution function, $f_\alpha(\mathbf{r}, \mathbf{v}, t)$, by considering the effects of particle collisions:

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_\mathbf{r} f_\alpha + \mathbf{a} \cdot \nabla_\mathbf{v} f_\alpha = \left(\frac{\delta f}{\delta t}\right)_{\text{col}} \tag{5}$$

The left side of the equation describes the rate of change of f_α due to the particles' motion and external forces, while the term $\left(\frac{\delta f_\alpha}{\delta t}\right)_{col}$ on the right side represents the change in f_α due to collisions.

This formulation is essential for systems where collisions are frequent and lead to a Maxwell-Boltzmann distribution under equilibrium conditions. However, for plasmas like the solar wind, where collisions are rare, a collisionless description is more appropriate.

3.2 Vlasov's Equation

The Vlasov equation is the collisionless analog of the Boltzmann equation, making it a more suitable kinetic model for the solar wind. It describes the evolution of the particle distribution function in a collisionless plasma, where interactions are primarily governed by external forces rather than particle collisions:

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_\mathbf{r} f_\alpha + \frac{\mathbf{F}}{m} \cdot \nabla_\mathbf{v} f_\alpha = 0 \tag{6}$$

In this context, the system is typically dominated by electromagnetic interactions. Thus, substituting with the Lorentz force is the ideal case:

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_\mathbf{r} f_\alpha + \frac{q}{m}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_\mathbf{v} f_\alpha = 0 \tag{7}$$

This form of Vlasov equation is widely used to capture the dynamics of collisionless plasmas, as it allows us to account for anisotropies in the solar wind, such as those arising from velocity space instabilities and non-thermal distributions of particle energies [17]. Unlike ideal MHD models, which assume that the Lorentz force term can be neglected, the Vlasov equation includes this term to account for deviations from isotropy and Maxwellian distributions in the solar wind, requiring numerical approximation techniques for practical solutions [6].

Note that Vlasov's equation is tightly related to statistical mechanics through Liouville's theorem:

$$\frac{\mathrm{d}f_\alpha(\mathbf{r}, \mathbf{v}, \mathrm{t})}{\mathrm{d}t} = 0 \tag{8}$$

If we compute use the chain rule along the phase and time space, we end up with the following expression:

$$\frac{\mathrm{d}f_\alpha(\mathbf{r}, \mathbf{v}, \mathrm{t})}{\mathrm{d}t} = \frac{\partial f_\alpha}{\partial t} + \nabla_\mathbf{r} f_\alpha \frac{\partial \mathbf{r}}{\partial t} + \nabla_\mathbf{v} f_\alpha \frac{\partial \mathbf{v}}{\partial t} = 0 \tag{9}$$

which indeed is:

$$\frac{\mathrm{d}f_\alpha(\mathbf{r}, \mathbf{v}, \mathrm{t})}{\mathrm{d}t} = \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_\mathbf{r} f_\alpha + \mathbf{a} \cdot \nabla_\mathbf{v} f_\alpha = 0 \tag{10}$$

3.3 Entropy

According to the second law of thermodynamics, the entropy of a closed system cannot decrease over time; it may increase or remain constant. In the kinetic framework, the entropy production rate is given by the time derivative of the entropy functional:

$$0 \leq \frac{\mathrm{d}}{\mathrm{d}t} \int f_\alpha \log f_\alpha d^3\mathbf{r} d^3\mathbf{v} \tag{11}$$

where $_\alpha \log f_\alpha$ represents the local entropy density in phase space. In a collisionless plasma, this entropy can remain constant if the system maintains phase-space coherence, but under realistic conditions, such as wave-particle interactions, entropy may increase as energy cascades through different scales. This formalism highlights how irreversible processes in the plasma lead to an increase in degrees of freedom, a key consideration for accurately modeling solar wind behavior.

For this case particularly, the colisionless nature of the plasma and Liouville's theorem assumes no change in entropy, if we use Leibniz method and compute the time derivative of the integrant:

$$\frac{d}{dt} f_\alpha \log f_\alpha = \frac{df_\alpha}{dt} \log f_\alpha + \frac{f_\alpha}{f_\alpha} \frac{df_\alpha}{dt} = 0 + 0 = 0 \tag{12}$$

4 Physics Informed Machine Learning

Usually, solely data-driven methods require huge volumes of high-quality data with black box algorithms that prevent interpretability, a key factor for physical modeling. The most recent research advocate for physical informed

7

virtual loss function terms [], leading off to the vastly known field of Physics Informed Neural Networks.

This models where created as numerical methods for solving Partial Differential Equations that embed physical modeling. It further extended to all fields that required PDE solving as a computationally efficient numerical alternative to other solvers.

4.1 Approximating the density function with PINNs

The most recent research, using PINNs, develop a way to not just learn functions, but to learn functional operators: Lagrangian, Hamiltonian, etc. [3,7] The first research creates a neural network that takes the general coordinates q and $\dot{q}$ as input, to provide the lagrangian as an output, taking advantage of the auto-differentiation capabilities of modern machine learning frameworks to enforce the euler-lagrange equation with gradient-descent methods:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0 \tag{13}$$

The current work extrapolates this concept to the field of density functions, embedding it into the training phase by ensuring the normalized nature of a density function, the second law of thermodynamics, and systems governed by kinetic theory following Liouville's theorem.

4.2 Neural Operators for discretization-invariant inference

Recent research on operator learning lead to the creation of Physics informed Neural Operators (PINOs), a set of weighted integral operators that are sequencially evaluated through the network.

5 Data

5.1 DSCOVR: Deep Space Climate Observatory

DSCOVR, a joint mission between NASA and the National Oceanic and Atmospheric Administration, is a crucial observational platform for monitoring space weather [12]. Launched in 2015, DSCOVR's primary mission

is to monitor and provide advanced warning of potentially hazardous space weather events such as solar flares and coronal mass ejections that could impact Earth.

It is equipped with two key instruments for measuring both energetic particle incidence and magnetic field parameters: the Faraday cup and the magnetometer from the PlasMag instrument [12]. The readings from these two sensors are crucial for virtually analyzing plasma dynamics near the L1 Lagrange point. These readings will be used as part of the core model data due to their real-time availability.

5.2   ACE: Advanced Composition Explorer

ACE, launched in 1997, provides continuous measurements of the solar wind and interstellar particles. It is equipped with several instruments designed to study the composition of solar and galactic particles, which are crucial for understanding the space weather environment. ACE's data helps in predicting geomagnetic storms and contributes to our understanding of the heliosphere.

5.3   WIND

The WIND spacecraft, launched in 1994, is part of the Global Geospace Science initiative. It provides comprehensive measurements of the solar wind, magnetic fields, and energetic particles. WIND's data is essential for understanding the fundamental processes of the solar wind and its interaction with the Earth's magnetosphere.

6   Methods

6.1   Neural Architecture Search (NAS)

As an intermediate step before training, a hyperparameter grid search algorithm based in Bayesian multi-objective optimization is employed to find the neural architecture. The Pareto front is plotted, selecting minimization of neural complexity and loss function.

6.2 Loss function

The neural network is trained using the standard gradient-descent based optimization algorithm: Adam. One Cycle learning optimizer is used to facilitate the learning rate plateau. Weight decay is implemented as a regularizer.

The loss function consists on a virtual term, three integral boundary condition terms, and a l2 regularizer term:

$$\mathcal{L}_{liouville}\left(\mathbf{E}, \mathbf{B}, \mathbf{v}, \mathbf{r}; \theta\right) = \left(\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \frac{q(\mathbf{E} + \mathbf{v} \times \mathbf{B})}{m} \cdot \nabla_{\mathbf{v}} f_\alpha\right)^2 \quad (14)$$

$$\mathcal{L}_{prob}\left(x; \theta\right) = \left(\int_{x \sim X} f_\alpha(\mathbf{v}, \mathbf{r}, t) d^3\mathbf{v} d^3\mathbf{r} - 1\right)^2 \quad (15)$$

$$\mathcal{L}_S(\mathbf{r}, \mathbf{v}, t) = \left(\frac{d}{dt} \int f_\alpha \log f_\alpha\right)^2 \quad (16)$$

$$\mathcal{L}_{regularizer}\left(\theta\right) = \mathbb{E}_{w \in \theta}\left[w^2\right] \quad (17)$$

Getting our final objective function:

$$\mathcal{L}\left[\mathcal{L}_{liouville}, \mathcal{L}_{prob}, \mathcal{L}_S \mathcal{L}_{regularizer}\right] = \alpha_1 \mathcal{L}_{liouville} + \alpha_2 \mathcal{L}_{prob} + \alpha_3 \mathcal{L}_S + \lambda \mathcal{L}_{regularizer} \quad (18)$$

7   Results

8   Discussion

9   Conclusion

This work demonstrated the effectiveness of Physics-Informed Neural Networks to model Solar Wind's behavior tied to diverse empirical and analytical formulations. By embedding the governing physical equations directly into the loss function, we can efficiently train a neural network to approximate the solution of this complex, non-linear system. Future work will focus on improving the kernel design and extending the approach to more sophisticated MHD scenarios, such as the interactions between Solar Wind and Sun's magnetosphere.

References

[1] Yasser Abduallah, Jason TL Wang, Prianka Bose, Genwei Zhang, Firas Gerges, and Haimin Wang. A deep learning approach to dst index prediction. *arXiv preprint arXiv:2205.02447*, 2022.

[2] Kristina Blume, Marc Röger, and Robert Pitz-Paal. Simplified analytical model to describe wind loads and wind-induced tracking deviations of heliostats. *Solar Energy*, 256:96–109, 2023.

[3] Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks, 2020.

[4] Tamas I. Gombosi, Bart van der Holst, Ward B. Manchester, and Igor V. Sokolov. Extended mhd modeling of the steady solar corona and the solar wind. *Living Reviews in Solar Physics*, 15(1), July 2018.

[5] J. T. Gosling. *The Solar Wind*. Elsevier, 2007.

[6] Maxime Grandin, Hyunju K. Connor, Sanni Hoilijoki, Markus Battarbee, Yann Pfau-Kempf, Urs Ganse, Konstantinos Papadakis, and Minna Palmroth. Hybrid-vlasov simulation of soft x-ray emissions at the earth's dayside magnetospheric boundaries, 2023.

[7] Sam Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks, 2019.

[8] Sabrina Guastavino, Katsiaryna Bahamazava, Emma Perracchione, Fabiana Camattari, Gianluca Audone, Daniele Telloni, Roberto Susino, Gianalfredo Nicolini, Silvano Fineschi, Michele Piana, and Anna Maria Massone. Forecasting geoffective events from solar wind data and evaluating the most predictive features through machine learning approaches, 2024.

[9] Jai Kumar, David Zarzoso, Virginie Grandgirard, Jan Ebert, and Stefan Kesselheim. Physics informed neural networks applied to the description of wave-particle resonance in kinetic simulations of fusion plasmas, 2023.

[10] Rong Lin, Zhekai Luo, Jiansen He, Lun Xie, Chuanpeng Hou, and Shuwei Chen. Prediction of solar wind speed through machine learning from extrapolated solar coronal magnetic field. *Space Weather*, 22(6):e2023SW003561, 2024. e2023SW003561 2023SW003561.

[11] Takuma Matsumoto. Full compressible 3D MHD simulation of solar wind. *Monthly Notices of the Royal Astronomical Society*, 500(4):4779–4787, 11 2020.

[12] National Aeronautics and Space Administration. Dscovr mission overview, 2023.

[13] G. Nicolaou, G. Livadiotis, C. J. Owen, D. Verscharen, and R. T. Wicks. Determining the kappa distributions of space plasmas from observations in a limited energy range. *The Astrophysical Journal*, 864(1):3, aug 2018.

[14] AS Parnowski. Regression modeling method of space weather prediction. *Astrophysics and Space Science*, 323(2):169–180, 2009.

[15] Federico Sabbatini and Catia Grimani. Solar wind speed estimate with machine learning ensemble models for lisa, 2023.

[16] David Stansby, Chadi Salem, Lorenzo Matteini, and Timothy Horbury. A new inner heliosphere proton parameter dataset from the helios mission. *Solar Physics*, 293(11), November 2018.

[17] Daniel Verscharen, Benjamin D. G. Chandran, Kristopher G. Klein, and Eliot Quataert. Collisionless isotropization of the solar-wind protons by compressive fluctuations and plasma instabilities. *The Astrophysical Journal*, 831(2):128, November 2016.

[18] Man Zhang, Xueshang Feng, Huichao Li, Ming Xiong, Fang Shen, Liping Yang, Xinhua Zhao, Yufen Zhou, and Xiaojing Liu. Numerical modeling of solar wind and coronal mass ejection in the inner heliosphere: A review. *Frontiers in Astronomy and Space Sciences*, 10, 2023.

[19] I. Zouganelis, M. Maksimovic, N. Meyer-Vernet, H. Lamy, and K. Issautier. A transonic collisionless model of the solar wind. *The Astrophysical Journal*, 606(1):542–554, May 2004.