

# Statistical Mechanics Informed Neural Networks for Solar Wind modeling

**Jorge D. Enciso**

*Asuncion, Paraguay*

JORGED.ENCYSO@GMAIL.COM

**Editor:** My editor

## Abstract

The solar wind, a collisionless, quasi-neutral plasma, presents unique challenges for modeling due to its deviation from classical assumptions such as the Maxwell-Boltzmann distribution. Traditional numerical and analytical approaches are computationally expensive and often rely on oversimplified empirical constraints. Recent advancements in machine learning, particularly Physics-Informed Neural Networks (PINNs), offer an efficient alternative for solving partial differential equations while incorporating physical laws. This research introduces a novel framework: Statistical Mechanics Informed Neural Networks, tailored for modeling the kinetic behavior of the solar wind. By leveraging statistical mechanics and avoiding restrictive magnetohydrodynamic assumptions, the proposed methodology directly approximates the solar wind's probability density function. Using data from near-Earth L1-Lagrange point observations, this work demonstrates the feasibility of PINNs in accurately capturing the solar wind's anisotropic and non-thermal dynamics. The findings highlight the potential of integrating statistical physics and machine learning for advanced space plasma modeling, paving the way for more precise solar-terrestrial interaction predictions.

**Keywords:** Physics informed Neural Networks, Solar Wind, Statistical Mechanics, Vlasov's Equation, Maxwell Equations, Ideal MHD

## 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. ? 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 globe 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 ?, numerical modelling ???, and machine learning approaches ??????. All of them with different purposes.

The solar wind presents unique challenges due to its deviation from classical assumptions. It is a quasy-neutral, colisionless, and empirically quasy-isotropic ionized plasma. ? Therefore, the adaptation of plasma dynamics frameworks for solar wind modeling must be carefully analyzed. For instance, the ideal magnetohydrodynamics (Ideal MHD) model is used under the assumption of an ionized plasma that converges to the Boltzmann distribu-

tion (Maxwell-Boltzmann distribution for temperatures), but it's an empirically unfeasible distribution as it misrepresents historical observations on solar wind's behavior.

On the other hand, using numerical approximators to model plasma dynamics ends up being computationally expensive. Take the Ideal MHD as an example:

$$\begin{aligned}
 \nabla \cdot \mathbf{E} &= \frac{\sigma}{\epsilon_0} \\
 \nabla \cdot \mathbf{B} &= 0 \\
 \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0 \\
 \frac{d}{dt} \left( \frac{p}{\rho^\gamma} \right) &= 0 \\
 \mathbf{E} + \mathbf{v} \times \mathbf{B} &= 0 \\
 \frac{\partial \mathbf{B}}{\partial t} &= \nabla \times (\mathbf{v} \times \mathbf{B}) \\
 \mathbf{J} \times \mathbf{B} &= \frac{(\mathbf{B} \cdot \nabla) \mathbf{B}}{\mu_0} - \nabla \left( \frac{B^2}{2\mu_0} \right)
 \end{aligned}$$

where  $\mathbf{E}$  is the electric field,  $\sigma$  the charge density,  $\mathbf{B}$  the magnetic field,  $p$  the pressure,  $\rho$  the mass density,  $v$  the general velocity, and  $\gamma$  the adiabatic index.

Nevertheless, recent research on machine learning computational methods enable novel approaches towards PDE solving and operator learning. These architectures, named Physics Informed Neural Networks, allow learnable solutions for less computational price thanks to modern hardware optimizations and shallower representations with universal approximators (in this case, neural networks). Hence, the scope of this research is to adhere these concepts on a deeper basis, proposing a new methodology within numerical statistical mechanics for solar wind modeling.

To this date, the literature doesn't embark on statistical mechanics informed neural networks for probability density function learning on a solar wind scheme. Therefore, the objective of this work is to leverage statistical mechanics theory and physics informed neural networks to propose a new numerical approximation for the kinetic behavior of the solar wind.

## 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 ? 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. ? It employs the model onto fusion modeling within the collisionless 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 geo-effective 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.

## 2.4 Statistical-Physics-Informed Neural Networks (Stat-PINNs): A Machine Learning Strategy for Coarse-graining Dissipative Dynamics

## 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 collisionality 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 ???, 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:

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

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

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

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

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_\alpha = 0 \quad (1)$$

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 ?. 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 ?.

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

$$\frac{df_\alpha(\mathbf{r}, \mathbf{v}, t)}{dt} = 0$$

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

$$\frac{df_\alpha(\mathbf{r}, \mathbf{v}, t)}{dt} = \frac{\partial f_\alpha}{\partial t} + \nabla_{\mathbf{r}} f_\alpha \frac{\partial \mathbf{r}}{\partial t} + \nabla_{\mathbf{v}} f_\alpha \frac{\partial \mathbf{v}}{\partial t} = 0$$

which indeed is:

$$\frac{df_\alpha(\mathbf{r}, \mathbf{v}, t)}{dt} = \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \mathbf{a} \cdot \nabla_{\mathbf{v}} f_\alpha = 0$$

### 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:

$$0 \leq \frac{d}{dt} \int f_\alpha \log f_\alpha d^3 \mathbf{r} d^3 \mathbf{v}$$

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 collisionless 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:

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

Thus, the entropy of this system is defined as follows:

$$\frac{d}{dt} \int f_\alpha \log f_\alpha d^3 \mathbf{r} d^3 \mathbf{v} = 0 \quad (2)$$

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

This models were 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. ?? 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:

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

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.

## 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 ?. 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 ?. 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 rate scheduler imizer is used to facilitate the convergence to a learning rate plateau. L2 regularizer and dropout is implemented to avoid overfitting.

The loss function consists on Liouville's theorem (Equation 1), the entropy constraint (Equation 2), a normalization term for probability density functions, and a regularizer term.

$$\mathcal{L}_{Liouville}(\mathbf{E}, \mathbf{B}, \mathbf{v}, \mathbf{r}; \theta) = \left( \frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \frac{q(\mathbf{E} + \mathbf{v} \times \mathbf{B})}{m} \cdot \nabla_{\mathbf{v}} f_\alpha \right)^2 \quad (3)$$

$$\mathcal{L}_P(x; \theta) = \left( \int_{x \sim X} f_\alpha(\mathbf{v}, \mathbf{r}, t) d^3 \mathbf{v} d^3 \mathbf{r} - 1 \right)^2 \quad (4)$$

$$\mathcal{L}_S(\mathbf{r}, \mathbf{v}, t) = \left( \frac{d}{dt} \int f_\alpha \log f_\alpha \right)^2 \quad (5)$$

$$\mathcal{L}_{regularizer}(\theta) = \mathbb{E}_{w \in \theta} [w^2] \quad (6)$$

Getting our final objective function:

$$\mathcal{L}[\mathcal{L}_{Liouville}, \mathcal{L}_P, \mathcal{L}_S, \mathcal{L}_{regularizer}] = \alpha_1 \mathcal{L}_{Liouville} + \alpha_2 \mathcal{L}_P + \alpha_3 \mathcal{L}_S + \lambda \mathcal{L}_{regularizer} \quad (7)$$

## 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 statistical mechanics formulations that lead to Vlasov's kinetic model. 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 architecture design with Physics informed Neural Operators (PINOs) ? by embedding Maxwell's equations for data continuization through differentiable splines, the satellite data dependance inhibits the introduction os such powerful frameworks in this case.

## Acknowledgments and Disclosure of Funding

## Appendix A.

## References