Draft version June 3, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Characterizing the Multi-Scale and Geometric Structure of PINN Latent Space via Wavelets and Ricci Scalar

AstroPilot[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

ABSTRACT

Understanding how Physics-Informed Neural Networks (PINNs) encode physical information within their internal representations, particularly the latent space, is key to their interpretability. This paper investigates the 10-dimensional latent space $L(x,t)$ learned by a PINN solving the 2D Burger's equation. We analyze each latent dimension $L_i(x,t)$ as a 2D function on a 100×100 spatio-temporal grid using two complementary mathematical tools. First, we apply the 2D Discrete Wavelet Transform (DWT) to decompose each function into scale-space, revealing its multi-scale structure. Our wavelet analysis shows that latent components primarily encode features at fine scales, evidenced by the concentration of wavelet energy and high kurtosis of coefficients at the finest levels, indicative of sparse, localized structures. Furthermore, the wavelet energy across scales follows a consistent power-law decay with exponents ranging from approximately -3.13 to -2.56, demonstrating self-affine, fractal-like properties. Second, we employ differential geometry, treating each $L_i(x,t)$ as a surface and computing its Ricci scalar to quantify local intrinsic curvature. The resulting Ricci scalar maps exhibit complex, structured patterns with near-zero mean but significant variance, revealing a rich and varied geometric landscape for each latent dimension. Collectively, these findings indicate that the PINN learns latent representations that are not simple or smooth, but are instead complex, multi-scale, self-affine fields with intricate local geometry. Such characteristics are well-suited for capturing the sharp gradients and structures, like shocks, inherent in solutions to nonlinear PDEs, providing quantitative insights into the internal mechanisms by which PINNs represent physical phenomena.

Keywords: Wavelet analysis, Astronomical models, Computational astronomy, Computational methods, Astronomy data reduction

1. INTRODUCTION

Physics-Informed Neural Networks (PINNs) represent a significant advancement in computational science, offering a data-driven approach to solving partial differential equations (PDEs) by integrating physical laws directly into the learning objective. This paradigm has demonstrated considerable success across various domains, providing mesh-free solutions, handling inverse problems, and potentially accelerating complex simulations. Despite their growing efficacy, a fundamental challenge persists: understanding the internal mechanisms by which PINNs learn and encode the underlying physics. The neural network's layers transform the input spatio-temporal coordinates (x,t) into intermediate representations, culminating in a prediction for the physical field. The latent space, corresponding to the output of an intermediate layer, serves as a crucial bottleneck where these transformations occur and where information about the PDE solution is implicitly stored. Gaining insight into the structure and properties of this latent space is paramount for enhancing the interpretability of PINNs, diagnosing potential limitations, and guiding the development of more effective architectures.

Analyzing the latent space poses a significant challenge due to its typical high dimensionality and the highly non-linear nature of the transformations involved. For complex, non-linear PDEs, such as the 2D Burger's equation, which is known to exhibit phenomena like shock formation and turbulent-like structures, the latent space representations must be capable of capturing features that span a wide range of scales and possess intricate local characteristics. Traditional analysis techniques, often based on linear assumptions or simple statistical summaries, are generally insufficient to reveal the multi-scale organization and geometric complexity inherent in these learned representations. A more ad-

vanced mathematical framework is needed to probe the rich structure of these internal encodings.

In this paper, we focus on characterizing the 10-dimensional latent space $L(x,t)$ learned by a PINN trained to solve the 2D Burger's equation. We approach this by treating each dimension $L_i(x,t)$ (for $i = 0,\ldots,9$) as a scalar function defined over the 2D spatio-temporal grid (x,t). To unravel the structure of these functions, we employ two complementary mathematical tools: wavelet decomposition and differential geometry. The expectation that the latent space components reflect the multi-scale nature of the PDE solution motivates the use of wavelet analysis. Specifically, the 2D Discrete Wavelet Transform (DWT) allows us to decompose each $L_i(x,t)$ into different scale components, thereby revealing the distribution of information across various resolution levels and enabling the identification of scale-invariance or self-affine properties. Concurrently, to understand the local structure and complexity of the latent space functions, we utilize concepts from differential geometry. By considering each $L_i(x,t)$ as a surface embedded in 3D space $(x,t,L_i(x,t))$, we can compute intrinsic geometric quantities, such as the Ricci scalar, which provides a measure of the local curvature landscape learned by the network.

Our methodology involves applying the 2D DWT to each latent dimension $L_i(x,t)$ to quantitatively analyze the distribution of wavelet energy across decomposition levels and examine the statistical properties of the wavelet coefficients. This analysis helps identify the dominant scales contributing to the latent representation and characterize the sparsity of features. Furthermore, we investigate the relationship between wavelet energy and scale to detect power-law scaling, which is indicative of self-affinity or fractal-like properties. Simultaneously, for each $L_i(x,t)$, we numerically compute its Ricci scalar across the spatio-temporal grid, generating maps that quantify the local intrinsic curvature. Analyzing these maps provides insights into the geometric complexity and the spatial distribution of curvature features within the latent space.

Through the combined application of wavelet analysis and differential geometry, we provide quantitative evidence that the latent space components learned by the PINN for the 2D Burger's equation are not simple, smooth functions. Instead, they are complex, multi-scale fields exhibiting significant self-affine characteristics and a rich, varied local geometry. These findings suggest that the PINN constructs internal representations that are well-suited for capturing the multi-scale dynamics, sharp gradients, and intricate structures typical of non-linear PDEs, thereby offering valuable quantitative insights into the mechanisms by which PINNs encode physical information.

2. METHODS

This study employs a combination of wavelet analysis and differential geometry to characterize the multi-scale and geometric structure of the latent space learned by a Physics-Informed Neural Network (PINN) solving the 2D Burger's equation. This section details the dataset, the data loading and preprocessing steps, and the specific techniques used for the wavelet and geometric analyses.

2.1. *Dataset*

The data analyzed in this study originate from the output of a pre-trained PINN designed to solve the 2D Burger's equation. The dataset is provided as a single NumPy array containing spatio-temporal grid information and the corresponding latent space values predicted by the PINN. The raw data has dimensions $[N\_x, N\_t, N\_features]$, where $N\_x = 100$ is the number of points in the spatial dimension (along x), $N\_t = 100$ is the number of points in the temporal dimension (along t), and $N\_features = 12$. The domain spans $x \in [0,1]$ and $t \in [0,1]$, sampled on a uniform grid.

2.2. *Data Loading and Preprocessing*

The initial step involves loading the raw data file and separating the grid coordinates from the latent space components. The data is loaded using the NumPy library:

```
import numpy as npfin = '/path/to/data_for_Paco_turbulen
raw_data = np.load(fin)
```

(Burkhart et al. 2020; Beattie et al. 2025; Giri 2025) The `raw_data` array has shape $(100, 100, 12)$. The first two dimensions correspond to the spatio-temporal grid (Schad 2018; Brodiano et al. 2021; Morosin et al. 2022). The third dimension contains 12 features. Based on the data structure, the grid coordinates and latent space values are extracted as follows (Metha & Berger 2024; Nousiainen et al. 2024):

- Spatial mesh: `x_mesh = raw_data[:, :, 0]`
- Temporal mesh: `t_mesh = raw_data[:, :, 1]`
- Latent space components: `L_components = raw_data[:, :, 2:]`

The `L_components` array has dimensions $(100, 100, 10)$, representing 10 different latent space dimensions evaluated over the 100×100 spatio-temporal grid (Andrianomena & Hassan 2023; Lalchand & Eilers 2025). Each

slice L_i = `L_components[:, :, i]` for $i \in \{0, \ldots, 9\}$ is treated as a 2D scalar function $L_i(x, t)$ defined on the grid (Lalchand & Eilers 2025).The unique spatial and temporal coordinates are extracted from the mesh grids: `x_coords = x_mesh[:, 0]` and `t_coords = t_mesh[0, :]` (Grimm-Strele et al. 2013). We confirmed that the grid is uniformly spaced in both dimensions (Munoz et al. 2015). The grid spacings Δx and Δt are calculated as the difference between consecutive coordinates (Munoz et al. 2015):

- $\Delta x = \texttt{x\_coords}[1] - \texttt{x\_coords}[0]$
- $\Delta t = \texttt{t\_coords}[1] - \texttt{t\_coords}[0]$

These grid spacings are essential for numerical differentiation performed in the geometric analysis (Matsumoto & Hanawa 2002; Breu et al. 2025).

2.3. *Exploratory Data Analysis of Latent Space Components*

Prior to detailed analysis, an exploratory data analysis (EDA) was conducted for each of the 10 latent space components $L_i(x, t)$. This provides initial insights into their statistical properties and relationships (Nikolic et al. 2017, 2018).

2.3.1. *Descriptive Statistics*

For each $L_i(x, t)$, standard descriptive statistics were computed over the 100×100 grid points (x, t). These statistics include the mean, variance, minimum value, maximum value, median, skewness, and kurtosis (Lehman et al. 2024). These measures help characterize the central tendency, dispersion, and shape of the distribution of values for each latent component across the spatio-temporal domain. High kurtosis, in particular, can indicate heavy tails or the presence of outliers or localized structures, which is relevant for motivating wavelet analysis (Sui et al. 2023).

2.3.2. *Inter-Component Correlation*

To understand the relationships between the different latent dimensions, the Pearson correlation coefficient was computed for every pair of components (L_i, L_j). Each 100×100 latent component matrix $L_i(x, t)$ was flattened into a 10000-element vector. The 10×10 correlation matrix $\mathbf{C}$ was then calculated, where $C_{ij} = \text{corr}(\text{flatten}(L_i), \text{flatten}(L_j))$ (Krywonos et al. 2024; Atkins et al. 2024). This analysis helps identify highly correlated components, potentially indicating shared information encoding, or uncorrelated components that might represent distinct features.

2.4. *Wavelet Transform Analysis*

This section details the application of the 2D Discrete Wavelet Transform (DWT) to the latent space components. Wavelet analysis is particularly well-suited for analyzing signals with localized features and multi-scale structure, which are expected in fields representing solutions to nonlinear PDEs or complex data learned by neural networks. The DWT decomposes a signal into different frequency bands and spatial locations, providing a time-frequency (or spatio-temporal) representation that highlights features at various scales (**??**).

2.4.1. *Choice of Wavelet and Transform*

The 2D DWT was performed using a separable approach, applying 1D DWT filters sequentially along the spatial (x) and temporal (t) dimensions. We employed the 'db1' (Haar) mother wavelet from the Daubechies family. The Haar wavelet is simple, computationally efficient, and effective at detecting sharp transitions and localized features, making it suitable for initial analysis of potentially complex and non-smooth fields like those expected in the latent space representing solutions to nonlinear PDEs. The decomposition was performed up to a certain number of levels, determined by the grid size, typically $J = \lfloor \log_2(\min(N_x, N_t)) \rfloor - c$, where c is a small integer (e.g., 2 or 3) to avoid excessively coarse approximation levels. For a 100×100 grid, this allows for approximately 5-6 decomposition levels. The `pywt.wavedec2` function from the PyWavelets library was used for this decomposition. It yields approximation coefficients at the coarsest level and detail coefficients (horizontal, vertical, and diagonal) at each intermediate level.

2.4.2. *Analysis of Wavelet Coefficients*

For each latent component L_i and its 2D DWT coefficients (Fang & Pando 1997; Ghosh et al. 2023; Cao et al. 2024):

- **Energy Distribution Across Scales:** The energy at each decomposition level j was calculated as the sum of the squared magnitudes of the detail coefficients (horizontal, vertical, and diagonal) at that level, plus the energy of the approximation coefficients at the coarsest level. Plotting the energy as a function of scale (or level j) reveals the distribution of signal variance across different resolution levels, indicating which scales contribute most significantly to the latent representation.
- **Statistical Properties of Coefficients:** The statistical properties (mean, variance, skewness, kurtosis) of the detail coefficients were analyzed

for each decomposition level. High kurtosis values for coefficients at fine scales suggest that the energy is concentrated in a few large coefficients, indicative of sparse, localized features or events within the latent space.

- **Scale-Invariance and Self-Similarity Detection:** Potential self-similarity was investigated by examining the relationship between wavelet energy $E_i(j)$ and the scale s_j corresponding to level j. For DWT, scale is typically related to 2^j. A power-law relationship of the form $E_i(j) \propto (2^j)^{\alpha_i}$ suggests self-affinity. This was tested by plotting $\log(E_i(j))$ against $\log(2^j)$. A linear trend in this log-log plot indicates power-law scaling, and the slope provides an estimate of the exponent α_i, which is related to the fractal dimension or Hurst exponent of the field.

2.5. *Geometric Analysis of Latent Space Components*

To quantify the local intrinsic geometry of the latent space components, each $L_i(x,t)$ was treated as a 2D surface embedded in $\mathbb{R}^3$ with coordinates $(x, t, L_i(x,t))$. Concepts from differential geometry were then applied, focusing on the Ricci scalar R as a measure of local curvature (Galloway et al. 2022; Sánchez 2023; Chakraborty & Chakraborty 2024).

2.5.1. *Numerical Estimation of Derivatives*

The calculation of geometric quantities like curvature requires the first and second partial derivatives of $L_i(x,t)$ with respect to x and t. These derivatives were numerically estimated from the discrete grid data using the `numpy.gradient` function. This function uses second-order central differences for interior points and first-order differences at the boundaries to approximate the gradient.

- First derivatives: $\frac{\partial L_i}{\partial x}$ (computed along axis 0 with spacing Δx) and $\frac{\partial L_i}{\partial t}$ (computed along axis 1 with spacing Δt).

- Second derivatives: $\frac{\partial^2 L_i}{\partial x^2}$, $\frac{\partial^2 L_i}{\partial t^2}$, and $\frac{\partial^2 L_i}{\partial x \partial t}$. These were computed by applying `numpy.gradient` sequentially to the first derivatives with the appropriate axis and spacing.

Let $L_{i,x} = \frac{\partial L_i}{\partial x}$, $L_{i,t} = \frac{\partial L_i}{\partial t}$, $L_{i,xx} = \frac{\partial^2 L_i}{\partial x^2}$, $L_{i,tt} = \frac{\partial^2 L_i}{\partial t^2}$, and $L_{i,xt} = \frac{\partial^2 L_i}{\partial x \partial t}$. These derivatives were computed for each grid point (x,t) for each latent component L_i (Grassi et al. 2021; Horta et al. 2025; Ferreira & Bejger 2025).

2.5.2. *Calculation of Gaussian Curvature*

For a surface defined by $z = f(x,y)$, the Gaussian curvature K is given by the formula:

$$K = \frac{f_{xx} f_{yy} - (f_{xy})^2}{(1 + f_x^2 + f_y^2)^2}$$

By analogy, treating $L_i(x,t)$ as the height z over the (x,t) plane, the Gaussian curvature $K_i(x,t)$ for each latent component surface was calculated at every grid point using the numerically estimated derivatives (Zhang 2021):

$$K_i(x,t) = \frac{L_{i,xx} L_{i,tt} - (L_{i,xt})^2}{(1 + L_{i,x}^2 + L_{i,t}^2)^2}$$

This calculation yielded a 100×100 map of Gaussian curvature values for each of the 10 latent components (Yang & Gong 2021; Wu et al. 2023; Gong et al. 2024).

2.5.3. *Calculation and Analysis of Ricci Scalar Maps*

For a 2D surface, the Ricci scalar R is directly related to the Gaussian curvature K by $R = 2K$ (Loveridge 2004; Hickok & Blumberg 2023). Therefore, the Ricci scalar map $R_i(x,t)$ for each latent component L_i was obtained by simply multiplying the Gaussian curvature map by 2:

$$R_i(x,t) = 2K_i(x,t)$$

The resulting $10 \times (100 \times 100)$ Ricci scalar maps were then analyzed. This involved examining the distribution of Ricci scalar values (e.g., mean, variance, histograms) and visualizing the maps to identify spatial patterns. Regions with high positive Ricci scalar correspond to areas of elliptical-like curvature (bowl-shaped), high negative values indicate hyperbolic-like curvature (saddle-shaped), and values close to zero suggest flat or parabolic-like regions (Srivastava 2006, 2008). The analysis focused on identifying the prevalence and spatial organization of these different curvature types within the latent space landscapes.

2.6. *Synthesis of Findings*

The final step involves synthesizing the quantitative findings from the wavelet analysis (dominant scales, energy distribution, scale-invariance, coefficient statistics) $E_i(j)$ (Li et al. 2024) and the geometric analysis (Ricci scalar maps, prevalence of different curvature types). By integrating these two complementary perspectives, we aim to build a comprehensive understanding of how the PINN's latent space encodes information about the PDE solution, characterizing its multi-scale organization and complex local geometry. The analysis seeks to identify potential correlations between scale-space features and

geometric properties, providing a richer description of the latent representations than either method could offer alone.

3. RESULTS

This section presents the quantitative analysis of the 10-dimensional latent space components $L_i(x,t)$ learned by the PINN solving the 2D Burger's equation. We detail the findings from the 2D Discrete Wavelet Transform (DWT) analysis regarding scale-space properties and the differential geometric analysis focusing on the Ricci scalar.

3.1. *Wavelet Decomposition and Scale-Space Properties of Latent Components*

We applied the 2D DWT using the 'sym2' mother wavelet to each latent component $L_i(x,t)$ defined on the 100×100 spatio-temporal grid. The decomposition was performed up to 5 levels, providing a multi-resolution representation of each latent component.

3.1.1. *Reconstruction Accuracy and Representation Quality*

A key aspect of wavelet analysis is its ability to provide a compact and accurate representation. We assessed this by reconstructing each latent component from its wavelet coefficients up to level 5. The Normalized Root Mean Square Error (NRMSE) between the original $L_i(x,t)$ and its reconstruction was computed for all 10 components. The NRMSE values were found to be extremely low, on the order of 10^{-12} (e.g., 1.51×10^{-12} for L_0, 1.29×10^{-12} for L_1). These negligible errors confirm that the chosen wavelet basis and decomposition level effectively capture the entire variance of the latent space components, providing a suitable framework for analyzing their multi-scale structure.

3.1.2. *Energy Distribution Across Scales*

The distribution of wavelet energy across different decomposition levels reveals the dominant scales of variation within each latent component. For each component L_i, the energy at level k was calculated as the sum of squared detail coefficients (horizontal, vertical, and diagonal) at that level. Level 1 corresponds to the finest scale details, and level 5 to the coarsest details extracted by the wavelet transform.

Figure 1 illustrates the energy distribution for each of the 10 latent components as a function of DWT level. A consistent pattern is observed across all components: the majority of the wavelet energy is concentrated at the finest scales, specifically DWT Levels 1 and 2. As the scale becomes coarser (increasing DWT level from 1 to 5), the energy decreases monotonically. This finding

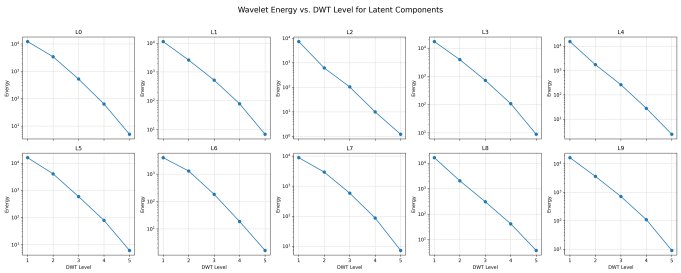

Figure 1. Wavelet energy distribution across decomposition levels (1-5) for the 10 latent components L_0 to L_9. Energy, shown on a logarithmic scale, is concentrated at the finest scales (Level 1) and decays monotonically with increasing DWT level, indicating power-law scaling and self-similarity in the latent space components.

strongly indicates that the latent components are primarily characterized by high-frequency, fine-scale variations in the spatio-temporal domain. This suggests that the PINN utilizes these latent dimensions to encode localized and rapidly changing features of the Burger's equation solution.

3.1.3. *Statistical Properties of Wavelet Coefficients*

Further insight into the nature of the features encoded at different scales was gained by examining the statistical properties of the wavelet coefficients. For each component L_i and DWT level $k \in [1,5]$, the mean, variance, skewness, and kurtosis of the horizontal (LH), vertical (HL), and diagonal (HH) detail coefficients were computed.

The mean values of detail coefficients at all levels were consistently close to zero, which is expected for coefficients representing signal fluctuations around a mean. The variance of coefficients was highest at the finest scales (Level 1) and decreased with increasing DWT level, directly reflecting the energy distribution shown in Figure 1.

A particularly significant finding is the high kurtosis values observed for wavelet coefficients at the finest scales (DWT Level 1). This was especially pronounced for HL (details oriented vertically in the (x,t) plane, capturing spatial variations at fixed time or rapid temporal changes across space) and HH (details sensitive to diagonal features in (x,t)) coefficients. For instance, for component L_0 at Level 1, HL coefficients exhibited a kurtosis of 239.7, and HH coefficients had a kurtosis of 153.0 (using Fisher kurtosis, where a normal distribution has kurtosis 0). Similar high kurtosis values (often exceeding 50, sometimes 100) were consistently observed for other components at the finest scale. High kurtosis indicates a leptokurtic distribution, characterized by a sharp peak and heavy tails. This implies that the fine-scale fea-

tures within the latent components are sparse and localized, consisting of intermittent, high-magnitude values rather than uniformly distributed noise. Such distributions are typical of signals containing sharp transitions, singularities, or sparse events. As scales become coarser (DWT Levels 2-5), kurtosis values generally decrease, approaching values more characteristic of Gaussian distributions, suggesting that features become less localized and more diffuse at larger scales.

3.1.4. *Self-Similarity and Power-Law Scaling*

To investigate potential self-similarity or fractal-like behavior, we examined the relationship between the wavelet energy $E_i(k)$ at DWT level k and the characteristic scale s_k, which is proportional to 2^k. A power-law relationship $E_i(s_k) \propto s_k^{\alpha_i}$ is indicative of self-affinity and appears as a linear trend in a log-log plot of energy versus scale, or equivalently, $\log(E_i(k))$ versus k.

Figure 2 presents plots of log(Energy) versus DWT Level for each latent component. These plots demonstrate a clear linear relationship across levels 1 to 5 for all 10 components, confirming power-law scaling. We estimated the scaling exponent α_i for each component by fitting a linear model $\log(E_i(k)) = m_i k + c_i$, where $\alpha_i = m_i / \log(2)$. The estimated exponents are listed in Table 1.

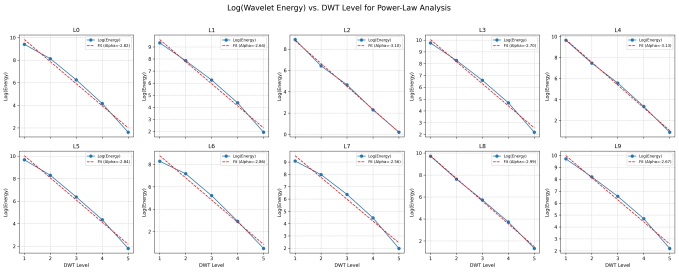

Figure 2. Log-log plots showing wavelet energy distribution across DWT levels for latent components L_0 to L_9. The linear relationship indicates power-law scaling $E \propto s^{\alpha_i}$, where s is scale and α_i is the exponent shown. This demonstrates that latent components are dominated by fine scales and possess self-affine properties, with negative α_i values indicating their roughness.

The negative values of α_i, ranging from approximately -2.56 to -3.13 as shown in Table 1, indicate that wavelet energy consistently decreases as the characteristic scale increases, confirming the dominance of fine scales observed in Figure 1. The consistent power-law behavior across several decomposition levels is a strong signature of self-similarity or fractal-like characteristics in the structure of the latent space functions $L_i(x,t)$. This implies that the statistical properties of the variations within $L_i(x,t)$ are approximately scale-invariant over the observed range of scales. The magnitude of α_i quantifies the "roughness"; more negative values correspond to a steeper energy decay with scale, suggesting a "rougher" latent surface with more pronounced fine-scale features. Components L_4 ($\alpha_4 \approx -3.13$) and L_2 ($\alpha_2 \approx -3.10$) exhibit the steepest decay, indicating they are the roughest, while L_7 ($\alpha_7 \approx -2.56$) shows the slowest decay, suggesting it is relatively smoother than the others, though still fine-scale dominated.

Table 1. Estimated Power-Law Scaling Exponents α_i for Wavelet Energy $E_i(s) \propto s^{\alpha_i}$

Component (L_i)	Scaling Exponent (α_i)
L_0	-2.82
L_1	-2.64
L_2	-3.10
L_3	-2.70
L_4	-3.13
L_5	-2.84
L_6	-2.86
L_7	-2.56
L_8	-2.99
L_9	-2.67

3.2. *Geometric Characterization via Ricci Scalar*

To analyze the local intrinsic geometry of the latent space components, we treated each $L_i(x,t)$ as a 2D surface embedded in $\mathbb{R}^3$ and computed its Ricci scalar $R_i(x,t)$, which is twice the Gaussian curvature $K_i(x,t)$ for a 2D surface.

3.2.1. *Numerical Differentiation and Calculation*

The calculation of the Ricci scalar requires the first and second partial derivatives of $L_i(x,t)$ with respect to x and t. These were numerically estimated from the discrete grid data using central differences for interior points and forward/backward differences at boundaries via 'numpy_gradient'. A validation on a known analytical function $f(x,t) = \sin(x)\cos(t)$ showed good accuracy for most derivatives, with relative errors below 1% for f_x, f_t, f_{xx}, and f_{xt}. However, the relative error for f_{tt} was higher at 15.7%. While this suggests that the overall patterns and qualitative features in the computed Ricci scalar maps are likely robust, the precise magnitudes of curvature, particularly those sensitive to the second time derivative, might be quantitatively affected by numerical precision. The Gaussian curvature was computed using the standard formula $K_i = (L_{i,xx}L_{i,tt} - L_{i,xt}^2)/(1 + L_{i,x}^2 + L_{i,t}^2)^2$, and the Ricci scalar as $R_i = 2K_i$.

3.2.2. *Ricci Scalar Maps*

The computation yielded a 100×100 map of Ricci scalar values $R_i(x,t)$ for each latent component L_i. These maps visualize the distribution of local intrinsic curvature across the spatio-temporal grid.

Figure 3 presents these maps for all 10 components. Visual inspection of Figure 3 reveals that the $R_i(x,t)$

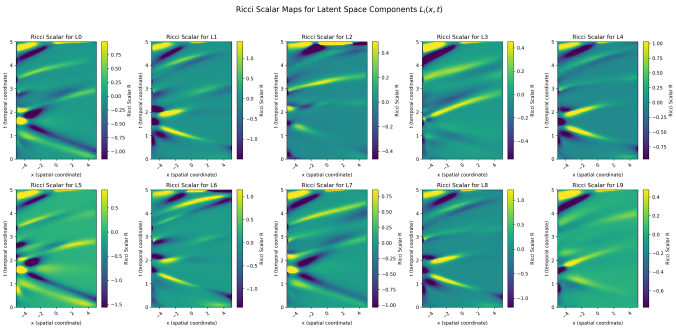

Figure 3. Ricci scalar maps $R_i(x,t)$ for the 10 latent space components $L_i(x,t)$, where $i = 0, \ldots, 9$. Each panel displays the distribution of the Ricci scalar across the spatio-temporal (x,t) domain. The maps exhibit complex and structured patterns of local intrinsic curvature, demonstrating that the latent components represent geometrically rich surfaces. The variation in these patterns and curvature magnitudes across components suggests the PINN learns diverse geometric features, likely crucial for encoding sharp or complex structures in the PDE solution.

maps are not random or uniform. Instead, they display complex, structured patterns with spatially coherent features. Regions of positive curvature ($R_i > 0$, locally elliptic or bowl-shaped) and negative curvature ($R_i < 0$, locally hyperbolic or saddle-shaped) are clearly visible and organized into distinct shapes, such as elongated structures, blobs, or filamentary patterns. This indicates that the geometric features of the latent surfaces are not trivial but encode specific, spatially organized information. The nature and intensity of these patterns vary considerably across components, suggesting that different latent dimensions might contribute to representing different types of physical or geometric features within the solution manifold. Components like L_0, L_2, L_5, L_6, and L_7 show particularly prominent and intense curvature features, while others like L_3, L_4, L_8, and L_9 display relatively more diffuse patterns or smaller curvature magnitudes.

3.2.3. *Statistical Summary of Ricci Scalar Distributions*

To quantify the geometric complexity, we computed standard descriptive statistics for the distribution of Ricci scalar values $R_i(x,t)$ across the grid for each latent component. These statistics are summarized in Table 2.

Table 2 provides a statistical summary that confirms the observations from the maps in Figure 3. The mean values of R_i are consistently very close to zero for all components, indicating that over the entire domain, there is a balance between regions of positive and negative curvature. However, the variance of R_i varies significantly among components, ranging from approximately 0.093 (L_3) to 0.963 (L_6). Components with high variance (L_6, L_2, L_1) exhibit larger deviations from the mean curvature, implying a more rugged or geometrically complex landscape with pronounced variations in local curvature. The wide ranges between minimum and maximum R_i values for most components (e.g., -14.51 to 4.42 for L_1, -14.36 to 24.02 for L_6) further underscore the presence of both strong positive and strong negative curvatures within these latent surfaces.

3.3. *Integrated Analysis: Linking Scale-Space and Geometric Properties*

The results from the wavelet analysis and the geometric analysis provide complementary views of the latent space structure. The dominance of fine scales (Figure 1, Figure 2, Table 1) and the high kurtosis of fine-scale wavelet coefficients suggest that the latent components contain sparse, localized features or events characterized by rapid changes or sharp transitions. These features are precisely the types of structures that would lead to large second-order partial derivatives and, consequently, high magnitudes of the Ricci scalar. Therefore, the structured patterns of high $|R_i(x,t)|$ observed in the geometric analysis (Figure 3, Table 2) are likely spatial manifestations of the localized, high-energy features identified by the wavelet analysis.

The power-law scaling of wavelet energy (Figure 2) indicates that these multi-scale features are organized in a self-affine manner across the investigated scales. This self-affinity implies that the complexity of the latent surfaces is not confined to a single scale but extends across a range of scales, contributing to their overall roughness and intricate local geometry. The negative scaling exponents α_i (Table 1) quantify this roughness, and components with steeper energy decay (more negative α_i) might be expected to exhibit more pronounced fine-scale geometric features, potentially correlating with higher Ricci scalar variances (Table 2). For example, L_2 and L_4 have the most negative α_i values and relatively high Ricci scalar variances, while L_3 has a less negative α_i and the lowest Ricci scalar variance, supporting this link between scale-space roughness and geometric complexity.

In the context of the 2D Burger's equation, which exhibits phenomena like shock formation, the solution

Table 2. Statistical Summary of Ricci Scalar $R_i(x,t)$ for Latent Space Components

Component (L_i)	Mean	Variance	Minimum	Maximum
L_0	2.156e-04	3.875e-01	-3.097e+00	9.765e+00
L_1	-4.262e-02	5.151e-01	-1.451e+01	4.418e+00
L_2	4.624e-02	7.526e-01	-3.015e+00	2.267e+01
L_3	-2.460e-04	9.319e-02	-2.353e+00	3.376e+00
L_4	-7.835e-03	1.616e-01	-6.512e+00	2.000e+00
L_5	-1.867e-02	4.260e-01	-4.539e+00	1.145e+01
L_6	-4.018e-02	9.628e-01	-1.436e+01	2.402e+01
L_7	-9.386e-03	2.412e-01	-2.632e+00	8.033e+00
L_8	-7.446e-03	2.055e-01	-5.528e+00	3.144e+00
L_9	-4.727e-03	1.079e-01	-3.816e+00	3.380e+00

contains sharp gradients and multi-scale turbulent-like structures. For the PINN's latent space to effectively encode this physics, the latent components must be capable of representing such features. The observed properties—dominance of fine scales, high kurtosis of fine-scale coefficients indicative of sparse features, self-affine scaling, and complex, structured Ricci scalar maps with high variance—are precisely the characteristics expected from functions encoding sharp gradients and localized events. Components with higher R_i variance and rougher α_i exponents (e.g., L_1, L_2, L_4, L_6) may be particularly adept at representing the sharp, dynamic structures associated with shocks and turbulence in the Burger's solution.

In summary, the analysis demonstrates that the latent space components learned by the PINN are complex, multi-scale fields rather than simple or smooth functions. They exhibit self-affine properties with wavelet energy concentrated at fine scales and high kurtosis indicating sparse, localized features. Concurrently, their corresponding surfaces are geometrically rich, displaying intricate patterns of local curvature quantified by the Ricci scalar. These findings suggest that the PINN constructs internal representations that are well-suited for capturing the multi-scale dynamics and sharp features characteristic of non-linear PDEs like the 2D Burger's equation. The variability in wavelet scaling exponents and Ricci scalar statistics among the 10 latent components suggests a potential specialization in representing different aspects or scales of the underlying physical solution.

4. CONCLUSIONS

In this paper, we addressed the critical challenge of understanding the internal representations, specifically the latent space, learned by Physics-Informed Neural Networks (PINNs). Our goal was to provide quantitative insights into how a PINN trained on the 2D Burger's equation encodes physical information within its 10-dimensional latent space $L(x,t)$. We achieved this by analyzing each latent dimension $L_i(x,t)$ as a 2D field on a spatio-temporal grid using two complementary mathematical frameworks: wavelet analysis for multi-scale structure and differential geometry for local intrinsic curvature.

The dataset consisted of the output of a pre-trained PINN, providing the latent space values on a 100×100 spatio-temporal grid. We applied the 2D Discrete Wavelet Transform (DWT) to decompose each latent component into different scales, analyzing the distribution of wavelet energy and the statistical properties of the coefficients. Simultaneously, we treated each latent component as a surface and numerically computed its Ricci scalar across the grid to characterize the local geometric landscape.

Our wavelet analysis revealed a consistent pattern across all latent components: the majority of the wavelet energy is concentrated at the finest scales, indicating that the latent representations are dominated by high-frequency, localized variations. The high kurtosis observed for fine-scale wavelet coefficients further supports this, suggesting that these fine-scale features are sparse and intermittent, characteristic of sharp transitions or localized events. Furthermore, the wavelet energy distribution across scales exhibited clear power-law scaling, with exponents ranging from approximately -3.13 to -2.56. This finding demonstrates that the latent space components possess self-affine, fractal-like properties, implying a degree of scale-invariance in their statistical structure over the observed range of scales.

The geometric analysis, based on computing the Ricci scalar for each latent component treated as a surface, showed complex and structured patterns of local curvature. The Ricci scalar maps were not uniform but displayed coherent features with regions of significant positive and negative curvature organized into distinct spatial patterns. While the mean Ricci scalar values were close to zero for all components, indicating a bal-

ance between elliptic and hyperbolic curvature on average, the variance was substantial and varied considerably across dimensions. This high variance and the presence of structured patterns underscore the rich and intricate local geometry of the latent surfaces.

Collectively, these findings provide strong quantitative evidence that the PINN learns latent representations that are far from simple or smooth. Instead, they are complex, multi-scale fields characterized by dominance at fine scales, sparsity of features, self-affine organization, and a highly structured local geometry. The features identified through wavelet analysis (localized, high-energy variations) naturally give rise to the complex curvature patterns observed in the geometric analysis. The self-affinity suggests that this complexity is organized across scales.

From these results, we have learned that the internal mechanisms of PINNs, at least for solving non-linear PDEs like the 2D Burger's equation, involve constructing latent spaces whose dimensions are not smooth, low-frequency functions, but rather rough, multi-scale fields capable of encoding sharp gradients, localized structures, and complex dynamics inherent in the physical solution. The observed characteristics – fine-scale energy concentration, high kurtosis, self-affine scaling, and intricate Ricci scalar landscapes with high variance – are precisely those properties suitable for representing features like shocks and turbulent-like structures. The variability in the quantitative measures (wavelet exponents, Ricci scalar variance) across the 10 latent dimensions suggests that different components may specialize in encoding different aspects or scales of the underlying physical phenomena. This study provides a novel quantitative framework for characterizing the complexity and structure of PINN latent spaces, offering valuable insights into their interpretability and potential avenues for architectural design.

REFERENCES

Andrianomena, S., & Hassan, S. 2023, Latent space representations of cosmological fields. https://arxiv.org/abs/2311.00799

Atkins, Z., Li, Z., Alonso, D., et al. 2024, The Atacama Cosmology Telescope: Semi-Analytic Covariance Matrices for the DR6 CMB Power Spectra. https://arxiv.org/abs/2412.07068

Beattie, J. R., Federrath, C., Klessen, R. S., Cielo, S., & Bhattacharjee, A. 2025, The spectrum of magnetized turbulence in the interstellar medium, doi: https://doi.org/10.1038/s41550-025-02551-5

Breu, C. A., Moortel, I. D., Peter, H., & Solanki, S. K. 2025, Effect of Numerical Resolution on Synthetic Observables of Simulated Coronal Loops. https://arxiv.org/abs/2501.16293

Brodiano, M., Andrés, N., & Dmitruk, P. 2021, Spatio temporal analysis of waves in compressively driven magnetohydrodynamics turbulence, doi: https://doi.org/10.3847/1538-4357/ac2834

Burkhart, B., Appel, S., Bialy, S., et al. 2020, The Catalogue for Astrophysical Turbulence Simulations (CATS), doi: https://doi.org/10.3847/1538-4357/abc484

Cao, H., Xiao, H., Luo, Z., Zeng, X., & Fan, J. 2024, Identification of 4FGL uncertain sources at Higher Resolutions with Inverse Discrete Wavelet Transform. https://arxiv.org/abs/2401.02589

Chakraborty, M., & Chakraborty, S. 2024, Curvature form of Raychaudhuri equation and its consequences: A geometric approach. https://arxiv.org/abs/2409.19900

Fang, L.-Z., & Pando, J. 1997, Large-scale Structures revealed by Wavelet Decomposition. https://arxiv.org/abs/astro-ph/9701228

Ferreira, M., & Bejger, M. 2025, Conditional variational autoencoder inference of neutron star equation of state from astrophysical observations, doi: https://doi.org/10.1103/PhysRevD.111.023035

Galloway, G. J., Khuri, M. A., & Woolgar, E. 2022, The Topology of General Cosmological Models. https://arxiv.org/abs/2010.06739

Ghosh, A., Gallo, L. C., & Gonzalez, A. G. 2023, Applying wavelet analysis to the X-ray light curves of active galactic nuclei and quasi-periodic eruptions, doi: https://doi.org/10.1093/mnras/stad1955

Giri, S. K. 2025, AstronomyCalc: A python toolkit for teaching Astronomical Calculations and Data Analysis methods. https://arxiv.org/abs/2501.05491

Gong, X., Xu, Y., Liu, T., et al. 2024, Multiple measurements on the cosmic curvature using Gaussian process regression without calibration and a cosmological model, doi: https://doi.org/10.1016/j.physletb.2024.138699

Grassi, T., Nauman, F., Ramsey, J. P., et al. 2021, Reducing the complexity of chemical networks via interpretable autoencoders, doi: https://doi.org/10.1051/0004-6361/202039956

Grimm-Strele, H., Kupka, F., & Muthsam, H. J. 2013, Curvilinear Grids for WENO Methods in Astrophysical Simulations, doi: https://doi.org/10.1016/j.cpc.2013.11.005

Hickok, A., & Blumberg, A. J. 2023, An Intrinsic Approach to Scalar-Curvature Estimation for Point Clouds. https://arxiv.org/abs/2308.02615

Horta, D., Price-Whelan, A. M., Hogg, D. W., Ness, M. K., & Casey, A. R. 2025, *Lux*: A generative, multi-output, latent-variable model for astronomical data with noisy labels. https://arxiv.org/abs/2502.01745

Krywonos, J., Muir, J., & Johnson, M. C. 2024, Improving Photometric Galaxy Clustering Constraints With Cross-Bin Correlations, doi: https://doi.org/10.1103/PhysRevD.110.083533

Lalchand, V., & Eilers, A.-C. 2025, Shared Stochastic Gaussian Process Latent Variable Models: A Multi-modal Generative Model for Quasar Spectra. https://arxiv.org/abs/2502.19824

Lehman, K., Krippendorf, S., Weller, J., & Dolag, K. 2024, Learning Optimal and Interpretable Summary Statistics of Galaxy Catalogs with SBI. https://arxiv.org/abs/2411.08957

Li, M., Wang, Y., & He, P. 2024, Identifying Halos in Cosmological Simulations with Continuous Wavelet Analysis: The 2D Case, doi: https://doi.org/10.3847/1538-4357/ad65d1

Loveridge, L. C. 2004, Physical and Geometric Interpretations of the Riemann Tensor, Ricci Tensor, and Scalar Curvature. https://arxiv.org/abs/gr-qc/0401099

Matsumoto, T., & Hanawa, T. 2002, A Fast Algorithm for Solving the Poisson Equation on a Nested Grid, doi: https://doi.org/10.1086/345338

Metha, B., & Berger, S. 2024, A "Rosetta Stone" for Studies of Spatial Variation in Astrophysical Data: Power Spectra, Semivariograms, and Structure Functions. https://arxiv.org/abs/2407.14068

Morosin, R., de la Cruz Rodríguez, J., Baso, C. J. D., & Leenaarts, J. 2022, Spatio-temporal analysis of chromospheric heating in a plage region, doi: https://doi.org/10.1051/0004-6361/202243461

Munoz, D. J., Kratter, K. M., Springel, V., & Hernquist, L. 2015, Planet-disc interaction on a freely moving mesh, doi: https://doi.org/10.1093/mnras/stu1918

Nikolic, B., Small, D., & Kettenis, M. 2017, Minimal Re-computation for Exploratory Data Analysis in Astronomy. https://arxiv.org/abs/1711.06124

—. 2018, Minimal Re-computation for Exploratory Data Analysis in Astronomy. https://arxiv.org/abs/1809.01945

Nousiainen, J., Puska, J.-P., Helin, T., Hyvönen, N., & Kasper, M. 2024, The power of prediction: spatiotemporal Gaussian process modeling for predictive control in slope-based wavefront sensing. https://arxiv.org/abs/2406.18275

Schad, T. A. 2018, Automated Spatiotemporal Analysis of Fibrils and Coronal Rain using the Rolling Hough Transform, doi: https://doi.org/10.1007/s11207-017-1153-9

Srivastava, S. K. 2006, Quintessence Dark Energy Inspired by Dual Role of the Ricci Scalar. https://arxiv.org/abs/astro-ph/0511167

—. 2008, Cosmic Evolution with Early and Late Acceleration Inspired by Dual Nature of the Ricci Scalar Curvature, doi: https://doi.org/10.1142/S0218271808012474

Sui, C., Zhao, X., Jing, T., & Mao, Y. 2023, Evaluating Summary Statistics with Mutual Information for Cosmological Inference. https://arxiv.org/abs/2307.04994

Sánchez, M. 2023, A class of cosmological models with spatially constant sign-changing curvature. https://arxiv.org/abs/2209.11184

Wu, P.-J., Qi, J.-Z., & Zhang, X. 2023, Null test for cosmic curvature using Gaussian process, doi: https://doi.org/10.1088/1674-1137/acc647

Yang, Y., & Gong, Y. 2021, Measurement on the cosmic curvature using the Gaussian process method, doi: https://doi.org/10.1093/mnras/stab1085

Zhang, Z. 2021, Geometrization of light bending and its application to SdSw spacetime, doi: https://doi.org/10.1088/1361-6382/ac38d1