
π -Net: Empirical Latent Representations for Differentiable Physics Heads in Data-Scarce Systems

Akshay Dhalpe*
Tampere University
akshay.dhalpe@tuni.fi

Arnab Chattopadhyay
University of Helsinki
arnab.chattopadhyay@helsinki.fi

Shahriar Bakrani Balani
Tampere University
shahriar.bakranibalani@tuni.fi

Sasu Tarkoma
University of Helsinki
sasutarkoma@helsinki.fi

Eric Cortanéa
Tampere University
eric.cortanea@tuni.fi

Abstract

Developing reliable predictive models for multi-physics systems often forces a compromise between computationally demanding physical simulators and data-hungry, opaque machine learning (ML) algorithms. While incorporating dimensional analysis into directed computation graphs offers a structurally sound alternative to explicitly solving partial differential equations (PDEs), such formulations typically yield underdetermined systems that act as imperfect simulators. We introduce π -Net, a physics-informed latent variable model that resolves these imperfect simulators. The hybrid π -Net architecture consists of an encoder that maps inputs into physically normalized latent space, which are then decoded into final predictions by a differentiable physics head. Evaluated on a real-world, data-scarce Gas Metal Arc Welding (GMAW) case study, the framework achieves an R^2 of 0.864 with only 60 experimental samples. The results demonstrate that π -Net significantly outperforms purely data-driven baselines in data-scarce regimes while providing causal, process-level explainability.

1 Introduction

In complex multi-physics domains like advanced manufacturing, predictive models must be physically consistent, interpretable and effective under severe data constraints. Traditional physical simulators rely on complete governing equations, such as partial and ordinary differential equations (PDEs or ODEs), which are frequently unavailable, partially known, or computationally intensive to solve (Brunton and Kutz [2024], Chen et al. [2018]). Conversely, purely data-driven machine learning (ML) models operate as black boxes. They offer no guarantee that predictions satisfy underlying physical dependencies and suffer sharp performance degradation in data-scarce experimental settings (Lee et al. [2025]). This dichotomy highlights a critical need for hybrid methodologies that can effectively leverage imperfect scientific models.

Physics-informed machine learning (PIML) attempts to bridge this gap, yet most existing approaches still depend on explicit algebraic or differential equations to enforce constraints Karniadakis et al. [2021], Meng et al. [2025]. The Dimensional Analysis and Conceptual Modelling (DACM) framework Coatanéa et al. [2016] provides a powerful alternative by encoding physical phenomena into directed acyclic graphs (DAG) using dimensionless power laws, bypassing the need for explicit PDEs. Despite their flexibility in mapping complex physical interactions, these graph-based representations provide insufficient constraints to uniquely determine all associated parameters Dhalpe et al. [2025]. This renders them imperfect simulators that traditionally require biased assumptions to solve.

To address this, we introduce π -Net, a novel hybrid architecture that reframes these imperfect simulators as physics-informed latent variable models. Specifically, π -Net utilizes a data-driven ML backbone as an **encoder** to map physical inputs into a physically normalized and bounded **latent space** of unresolved dimensionless coefficients (π). Unlike abstract, statistically derived latent representations, typical in deep learning, these variables capture real, bounded physical efficiencies. A DACM-derived DAG then acts as a fixed, **differentiable physics head**, projecting these latent variables to the final target prediction. In this framework, the components mutually resolve their fundamental limitations: the differentiable physics head supplies the causal, domain-grounded structure that pure ML lacks. While the data-driven encoder completes the underdetermined simulator directly from data. Crucially, this integration enables true **end-to-end learning**, where prediction errors backpropagate explicitly through the physical equations in the DAG to guide the encoder.

Our main contributions are as follows:

- **A novel hybrid architecture (π -Net)** that bypasses governing equations, formulating underdetermined simulators as latent variable models by coupling a data-driven encoder to estimate intermediate physical latent variables within a differentiable physics head.
- **End-to-end learning through an imperfect physical simulator.** The DAG directly routes both forward predictions and backward error propagation through fixed physical relationships, learning only the unresolved dimensionless coefficients.
- **Demonstrated high data efficiency**, uniformly outperforming pure data-driven baselines in a data-scarce Gas Metal Arc Welding (GMAW) case study.
- **Process-level explainability** that overcomes the opaque nature of standard ML models by explicitly routing the learned latent representations through a causally physical graph.

2 Related Work

Physics-informed machine learning. Physics-informed machine learning (PIML) traditionally embeds prior knowledge through loss functions or architectural constraints based on explicit governing equations, such as in Physics-Informed Neural Networks (PINN) (Karniadakis et al. [2021], Meng et al. [2025]). While effective, these methods struggle with complex multi-physics systems where the governing equations are incomplete, partially specified, or computationally prohibitive (Brunton and Kutz [2024]). Rather than requiring a complete mathematical description, π -Net incorporates an incomplete scientific model directly into the forward pass as a differentiable physics head. This retains physical structures while the data-driven encoder infers unresolved quantities from data.

Dimensional analysis and machine learning. To bypass explicit PDEs, some ML approaches leverage dimensional analysis and the Buckingham- Π theorem to enforce scale invariance. Recent methods use neural networks, constrained optimization, or symbolic regression to discover these dimensionless groups directly from data (Oppenheimer et al. [2023], Bakarji et al. [2022]). However, these approaches use dimensionless groups as inputs to learn the overall input-output mapping, leaving the interactions among individual variables implicit in the learned model. In contrast, π -Net formulates these unknown coefficients as a physically normalized latent space. By decoding this representation through a differentiable physics head: π -Net backpropagates errors directly through the physical equations, achieving end-to-end learning while preserving causal tracing of variables.

Graph-based physical modelling and imperfect simulators. The Dimensional Analysis and Conceptual Modelling (DACM) framework constructs DAG using dimensionally homogeneous power laws (Coatanéa et al. [2016]). While these directed graphs elegantly capture multiphysics causality, they function as structural templates rather than fully executable simulators, as dimensional homogeneity alone cannot fully constrain the system (Dhalpe et al. [2025]). π -Net resolves this by treating the unresolved dimensionless coefficients as bounded latent variables estimated by an ML encoder. Instead of learning the physics from scratch, it leverages this latent variable formulation to complete an underdetermined, structurally specified simulator.

3 Methodology

The π -Net architecture Given process inputs $\mathbf{x}$, π -Net completes the imperfect simulator as a physics-informed latent variable model. The encoder g_θ maps the inputs into a pre-defined, physically

normalized latent space of unresolved dimensionless coefficients $\hat{\pi}$. These predicted coefficients are constrained to domain-specified physical intervals $[\pi_i^{\min}, \pi_i^{\max}]$ via a scaled sigmoid activation. The physical computation graph $\mathcal{G}$ then operates as a fixed, differentiable physics head that projects this latent space to the final prediction $\hat{y} = \mathcal{G}(\mathbf{x}, \hat{\pi})$. Thus, physics is embedded as a strict structural prior rather than a soft loss penalty, actively participating in forward and backward passes.

The model is trained end-to-end on paired observations $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ with prediction errors backpropagating directly through the differentiable physics head to update the encoder’s parameters θ . As the DAG relies on multiplicative power-law relationships, standard optimization risks numerical instability and gradient explosion. To ensure stable gradient backpropagation through the fixed graph $\mathcal{G}$ to update θ , both the forward pass and the Mean Squared Error (MSE) objective are computed in logarithmic space: $\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N (\log \hat{y}_i - \log y_i)^2$. To validate this framework, the following sections detail the physics head, dataset, and encoder architectures used for evaluation.

The GMAW physics head To demonstrate the π -Net framework, we instantiate it on a complex multi-physics case study: Gas Metal Arc Welding (GMAW). The objective is to map a set of controllable input variables: arc voltage U_a , contact-tip-to-work distance L , wire feed rate v_f , and travel speed v_t to the target: the weld bead cross-sectional area A .

As the PDEs or ODEs required by traditional Physics-Informed Neural Network (PINN) are unavailable for this process, the DACM derived DAG encodes the intermediate physical mechanisms (energy transfer, wire melting, and material deposition) through a sequence of dimensionally homogeneous power-law relationships. The forward computational path is defined by the following sequential transformations:

$$U_a = \pi_1 U_1 \quad (1)$$

$$q_a = \pi_2 I_1 U_a \quad (2)$$

$$q_r = \pi_3 \lambda_r L I_1^2 d_w^{-2} \quad (3)$$

$$q = q_a + q_r \quad (4)$$

$$m_l = \rho_s^{1/2} d_w v_f^{1/2} q^{1/2} \pi_4^{1/2} H_m^{-1/2} \quad (5)$$

$$A = \pi_5 m_l v_t^{-1} \rho_l^{-1} \quad (6)$$

In this formulation, variables such as electrical current I_1 , wire diameter d_w , and enthalpy H_m are known physical constants or measured intermediate quantities. The edges of the computation graph are strictly defined by the physical exponents, representing fixed causal dependencies following the path $\{U_a, L, v_f, v_t\} \rightarrow \{q_a, q_r\} \rightarrow q \rightarrow m_l \rightarrow A$.

However, the dimensionless coefficients $\{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5\}$ capture complex multiphysics interaction within the robotic GMAW process. The dimensionless coefficients are defined through the functional representation and variable clustering, with each coefficient representing either a single function or a cluster of related functions. Thus, dependencies are formed between the dimensionless coefficients of consecutive functions because the structural dependencies remain fixed. This graph operates as an imperfect simulator that acts as a differentiable physics head, waiting to be completed by the data-driven encoder.

Data and encoders To evaluate our framework, we generated a custom experimental dataset specifically for this study using a robotic Gas Metal Arc Welding (GMAW) process. The four controllable input variables ($\mathbf{x}$) were sampled via a Maximum Projection (MAXPRO) design to ensure space-filling diversity, targeting the physical weld cross-sectional area ($y = A$). The prediction target y is the physical weld cross-sectional area A . The resulting dataset contains $N = 60$ experimental runs, partitioned equally into two distinct physical regimes: short-circuit (30 samples) and globular (30 samples) metal-transfer modes. Such data-scarce conditions are inherent to manufacturing use-cases, where each observation requires physical material deposition, thermal stabilization, and costly precision metrology, making large-scale data acquisition time and resource intensive. To isolate the structural benefits of the physics head from representational capacity, we instantiate the encoder g_θ using three distinct architectures (LASSO, MLP, and FT-Transformer) to evaluate on this constrained dataset. Although the encoder predicts sample-dependent coefficient values, their functional identities and interactions are prescribed by the graph.

4 Results

4.1 Evaluating π -Net on Accuracy and Robustness

To establish a baseline, we compared three π -Net variants against their purely data-driven counterparts. Notably, a traditional PINN baseline could not be included, as the complete governing differential equations required for standard PDE-based regularization are unavailable for this complex multi-physics phenomenon. While the baseline models predict the final target directly from the inputs, the π -Net variants utilize the identical architectures strictly as encoders to predict the latent space, which is then routed through the differentiable physics head. As each pair shares identical capacity, any performance gains are strictly attributable to the embedded DAG. We evaluated the models using stratified cross-validation on the 60-sample dataset to isolate the impact of the physics integration. Table 1 shows, the π -Net variants uniformly outperform their data-driven baselines. The π -Net FT-Transformer has the highest accuracy ($\bar{R}^2 = 0.864$, MAPE = 6.93%), demonstrating that the embedded causal structure rather than the backbone complexity is the driver of predictive performance.

To assess robustness, training data was progressively reduced from 80% (5-fold) to 50% (2-fold). As data decreased, purely data-driven models degraded sharply; the baseline FT-Transformer’s R^2 fell to 0.679 at the 50% split (Table 2). In contrast, the π -Net variants maintained strong predictive capability. Notably, the linear π -Net LASSO outperformed the non-linear FT-Transformer baseline across all data splits, confirming that embedded physical constraints provide the necessary structure to succeed in extreme data-scarce environments where data-hungry ML architectures fail to generalize.

Table 1: Performance comparison of backbone architectures and their π -Net variants.

Model	Baseline			π -Net		
	R^2	MAPE	RMSE	R^2	MAPE	RMSE
LASSO	0.742	10.69%	1.919	0.840	7.92%	1.511
MLP	0.809	7.39%	1.652	0.851	7.14%	1.461
FT-Transformer	0.796	8.14%	1.708	0.864	6.93%	1.394

Table 2: Effect of training set size on the cross-validation R^2 performance of backbone architectures and their π -Net variants.

Training split	LASSO		MLP		FT-Transformer	
	Baseline	π -Net	Baseline	π -Net	Baseline	π -Net
80% train; 20% test	0.742	0.840	0.809	0.851	0.796	0.864
75% train; 25% test	0.697	0.830	0.832	0.851	0.782	0.852
67% train; 33% test	0.712	0.833	0.799	0.848	0.770	0.840
50% train; 50% test	0.680	0.821	0.792	0.825	0.679	0.825

4.2 Process level explainability

As illustrated in Figure 1, the π -Net architecture provides structural interpretability by explicitly routing information through a physically meaningful causal graph. For example, the highlighted red pathway demonstrates how a perturbation in the independent variable arc voltage (U_a) propagates downstream through intermediate physical states. The directed graph explicitly maps this causal chain: arc voltage dictates arc heat (q_a), which contributes to total heat (q), driving the wire melting rate (m_l), and ultimately determining the final weld bead cross-sectional area (A). By maintaining this fixed, dimensionally constrained sequence ($U_a \rightarrow q_a \rightarrow q \rightarrow m_l \rightarrow A$), the framework transcends the opacity of standard black-box ML. It enables domain experts to directly trace how input variations influence final performance targets, providing the process-level explainability required to support reliable engineering decision-making.

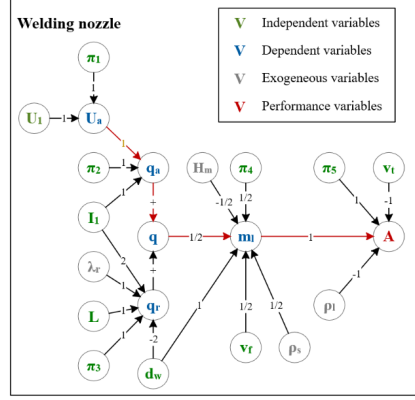


Figure 1: Directed graph of the GMAW process showing the directed propagation path from arc voltage (U_a) to weld bead cross-sectional area (A)

5 Discussion and Future work

Structural regularization via differentiable physics heads. Standard data-driven models must implicitly learn physical symmetries and scaling laws from data, a task that requires massive datasets and frequently fails in complex multi-physics environments. Conversely, traditional PINNs require complete partial differential equations, which are often unavailable or computationally prohibitive for such processes. By using a DACM-derived differentiable physics head as a structural prior, π -Net intrinsically enforces dimensional homogeneity while enabling end-to-end learning. Prediction errors backpropagate explicitly through fixed mathematical relationships to update the encoder. This differentiable prior bounds the hypothesis space to prevent physically impossible predictions and ensure rapid convergence despite severe data scarcity.

Scope and Generalizability. Despite the graph’s regularizing effect, our highly sparse dataset leaves regions of the high-dimensional parameter space underrepresented. Consequently, predictions should be treated with caution when extrapolating out-of-distribution. Furthermore, empirical validation is currently restricted to a single GMAW case study; evaluating the architecture across diverse multi-physics domains remains a critical next step.

Dimensionless latent spaces and transfer learning. In standard machine learning, latent spaces are abstract, statistically derived representations lacking physical meaning. In contrast, the π -Net encoder maps inputs to a latent space of dimensionless coefficients (π) that represent underlying physical phenomena rather than abstract features or absolute magnitude. These coefficients are coupled through the prescribed functional representation and physical graph, so the model learns graph-consistent latent correction factors rather than independently identifiable physical quantities. Furthermore, the inherent scale-invariance of these physically bounded π values might enable zero or few-shot transfer learning across different sized systems.

Partial Explainability. While π -Net resolves the complete opacity of standard ML, it remains a gray-box architecture. It offers structural interpretability: allowing experts to trace exactly how an estimated latent coefficient causally propagates through intermediate variables. However, while the differentiable physics head is fully transparent, the encoder mapping $g_\theta : \mathbf{x} \rightarrow \hat{\pi}$ remains a black box. This decoupled latent architecture represents a pragmatic compromise between the rigid transparency of white-box physical models and the high representational capacity of deep learning.

Future work. π -Net is viewed as a foundational step toward scalable, structurally constrained representation learning. Future work will extend this via a DACM-GNN architecture, where the DAG governs local thermodynamics and phase transitions while a Graph Neural Network handles spatial diffusion across complex topologies. Ultimately, by leveraging scale-invariant physics inherent to these dimensionless latent variables we aim to study ways to implement zero or few-shot transfer learning, scaling seamlessly from experimental geometries to industrial applications such as Wire Arc Additive Manufacturing (WAAM), battery thermal modeling, and silicon processing.

Backbone Model	Architecture Configuration	Key Mechanism & Rationale
LASSO	Single linear layer mapping the 4 input features to 5 dimensionless coefficient outputs ($\pi_2 - \pi_6$).	Employs an L_1 penalty to perform feature selection, identifying which welding parameters exert the greatest influence on each dimensionless group.
MLP	Input: 4 neurons Hidden: 2 layers (16 neurons each) Output: 5 neurons corresponding to the coefficients $\pi_2 - \pi_6$	Compact design deliberately limits trainable parameters to mitigate overfitting risks on the small experimental dataset ($n = 60$). Aligns with standard, robust baselines for GMAW [47,48].
FT-Transformer	Tokenization: Independent linear tokenizer projects each feature to a 64-dimensional embedding with additive learnable positional embeddings. Encoder: A CLS token is prepended (forming a 5-token sequence), processed by a 2-layer pre-norm transformer encoder. Output: Learned CLS representation is decoded to $\pi_2 - \pi_6$.	Attention mechanism explicitly models the complex, strong interdependencies between welding parameters (e.g., arc voltage and wire feed rate).

Table 3: Architectural configuration and structural rationale for the evaluated ML backbones.

References

- Steven L Brunton and J Nathan Kutz. Promising directions of machine learning for partial differential equations. *Nature Computational Science*, 4(7):483–494, 2024.
- Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.
- Hugon Lee, Hyeonbin Moon, Junhyeong Lee, and Seunghwa Ryu. Toward knowledge-guided ai for inverse design in manufacturing: A perspective on domain, physics, and human–ai synergy. *Advanced Intelligent Discovery*, page e202500107, 2025.
- George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.
- Chuizheng Meng, Sam Griesemer, Defu Cao, Sungyong Seo, and Yan Liu. When physics meets machine learning: A survey of physics-informed machine learning. *Machine Learning for Computational Science and Engineering*, 1(1):20, 2025.
- Eric Coatanéa, Ric Roca, Hossein Mokhtarian, Faisal Mokammel, and Kimmo Ikkala. A conceptual modeling and simulation framework for system design. *Computing in Science & Engineering*, 18(4):42–52, 2016.
- Akshay Dhalpe, Hossein Mokhtarian, Suraj Panicker, Di Wu, Joe David, Shahriar Bakrani Balani, Jose Martinez Lastra, and Eric Coatanéa. Integrating dimensional and scaling analyses with functional modelling and graphs: An approach to comprehend mass transfer in welding. *The International Journal of Advanced Manufacturing Technology*, 138(1):61–79, 2025.
- Michael W Oppenheimer, David B Doman, and Justin D Merrick. Multi-scale physics-informed machine learning using the buckingham pi theorem. *Journal of Computational Physics*, 474: 111810, 2023.

Joseph Bakarji, Jared Callahan, Steven L Brunton, and J Nathan Kutz. Dimensionally consistent learning with buckingham pi. *Nature Computational Science*, 2(12):834–844, 2022.