Overcoming Vanishing Gradients in Inverse Source Localization via Physically-Relaxed PINNs

Dawen Wu [1 2] Amine Ammar [3 1] Francisco Chinesta [3 1]

[1]CNRS@CREATE, 1 Create Way, #08-01 Create Tower, Singapore 138602 [2]School of Computing, National University of Singapore, 13 Computing Drive, Singapore 117417 [3]Arts et Metiers Institute of Technology, Paris, France. Correspondence to: Amine Ammar amine.ammar@ensam.eu.

Abstract

Physics-Informed Neural Networks (PINNs) have established a powerful paradigm for solving inverse problems. However, in inverse source localization within transport equations, standard formulations fail when the initial guess is distant from the ground truth due to vanishing gradients caused by lack of spatiotemporal overlap. To dismantle this barrier, we propose Physically-Relaxed PINNs (PR-PINN). We introduce transient relaxation mechanisms—Laplacian Smoothing, Source Dilatation, and Data Nudging—to artificially broaden the solution support and convexify the optimization landscape. We provide theoretical analysis deriving explicit gradient expressions to demonstrate how these mechanisms restore informative gradients. Empirical validations on advection-diffusion, chaotic reaction-diffusion, and acoustic systems demonstrate that our approach robustly recovers latent source coordinates where baseline methods stagnate.

1. Introduction

The integration of deep learning with physical laws has established a new paradigm in Scientific Machine Learning. Physics-Informed Neural Networks (PINNs) embed partial differential equations (PDEs) directly into the loss function, enabling the solution of inverse problems from sparse data without complex adjoint-state formulations.

This work addresses the inverse problem of source localization in transport-dominated regimes. Standard PINNs frequently encounter convergence failures here, specifically vanishing gradients when the initialization is far from the ground truth. This stems from the localized spatiotemporal support of the physical solution; if the predicted field has no overlap with sensor locations, the loss function becomes insensitive to source parameters.

To overcome this, we propose PR-PINN, a framework grounded in homotopy optimization. We modify the governing PDE to artificially broaden the solution's profile using relaxation mechanisms (smoothing, dilatation, nudging), effectively convexifying the loss landscape to guide the solver toward the global basin of attraction.

2. Problem Definition

We consider the dynamics of a scalar field $u(\mathbf{x}, t)$ governed by:

$$\mathcal{D}_t[u] + \mathcal{N}_{\mathbf{x}}[u] = s(\mathbf{x}, t; \boldsymbol{\theta}_{src}), \quad (\mathbf{x}, t) \in \Omega_T, \tag{1}$$

The objective is to recover unknown source parameters $\boldsymbol{\theta}_{src}$ given sparse measurements $\mathcal{D} = \{(\mathbf{x}_i, t_i, u_i^{obs})\}_{i=1}^{N_{obs}}$.

We employ PINNs to minimize a composite loss:

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\theta}_{src}) = \mathcal{L}_{data}(\mathbf{w}) + \lambda \mathcal{L}_{res}(\mathbf{w}, \boldsymbol{\theta}_{src}). \tag{2}$$

In transport-dominated regimes, if the initial guess is poor, the sensitivity term $\nabla_{\boldsymbol{\theta}_{src}} u(\mathbf{x}_i, t_i)$ vanishes at sensor locations, resulting in a flat loss plateau (vanishing gradient pathology).

3. Methodology: PR-PINN

We propose Physically-Relaxed PINN (PR-PINN) to reshape the loss landscape. We introduce an augmented residual operator $\mathcal{R}_{\epsilon,\eta}$ with relaxation terms:

$$\mathcal{R}_{\epsilon,\eta} := \mathcal{D}_t[u] + \mathcal{N}_{\mathbf{x}}[u] - s(\mathbf{x}, t; \sigma) - \epsilon \Delta u - \eta \sum_{i=1}^{N_{obs}} (u_i^{obs} - u)\delta(\mathbf{x} - \mathbf{x}_i). \tag{3}$$

Laplacian Smoothing ($\epsilon \Delta u$) Injecting transient artificial viscosity to broaden solution support, ensuring "tails" of the distribution reach sensors.

Source Dilatation Initializing with a spatially broad source (large σ) to ensure signal overlap with distant sensors, then annealing σ to the true sharp profile.

Data Nudging A forcing term proportional to observational deviation, acting as an internal source at sensor locations to guide the field toward the data manifold.

We employ a dynamic annealing schedule where $\epsilon(t)$ and $\eta(t)$ decay to zero, recovering the strict physical physics after the optimizer reaches the correct neighborhood.

4. Theoretical Analysis

We analyzed a 1D linear advection-diffusion prototype to explain the pathology. The gradient of the loss w.r.t source location x_{src} involves an term proportional to:

$$\exp\left(-\frac{((x^* - x_{src}) - v(t^* - \tau))^2}{2\Sigma^2(\tau)}\right) \tag{4}$$

When the distance is large, this term decays to zero exponentially. Our analysis proves that increasing diffusivity (D) or source width (σ_s) increases the effective variance $\Sigma^2(\tau)$, preventing the gradient from vanishing and restoring a valid descent direction.

5. Experiments

We evaluate PR-PINN on three systems. The network uses 3 hidden layers (128 units, $\tanh$) and is trained with Adam.

5.1 Advection-Diffusion with Misspecification

We tested a 2D case where the wind field is unknown and misspecified as quiescent. While the Vanilla PINN remained trapped on a zero-gradient plateau, PR-PINN successfully convexified the landscape (via smoothing and nudging), recovering the true source coordinates $\mathbf{x}_{src}$ despite the structural error.

5.2 Chaotic Reaction-Diffusion

In a Fisher-KPP system driven by a chaotic Double Gyre flow, standard methods failed due to complex mixing and lack of overlap. PR-PINN smoothed the energy landscape and compensated for missing advection physics, achieving a mean square error $< 10^{-2}$.

5.3 Acoustic Wave Equation

For hyperbolic wave propagation, finite wave speeds cause severe gradient vanishing. By applying artificial viscosity and source dilatation, PR-PINN allowed the initial wavefronts to interact with sensors. Annealing these parameters enabled the accurate recovery of the source location and wave speed, resolving regimes where baseline methods stagnated.

6. Conclusion

We presented PR-PINN to solve the vanishing gradient problem in inverse source localization. By leveraging homotopy optimization through transient physical relaxations, our method bridges the topological gap between initialization and ground truth. Empirical results across parabolic and hyperbolic systems confirm its robustness, offering a promising direction for complex inverse problems with sparse data.

Acknowledgments

DesCartes: this research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.