Physics-Guided Discovery of Highly Nonlinear Parametric Partial Differential Equations

Yingtao Luo, Qiang Liu, Yuntian Chen, Wenbo Hu, Tian Tian, Jun Zhu

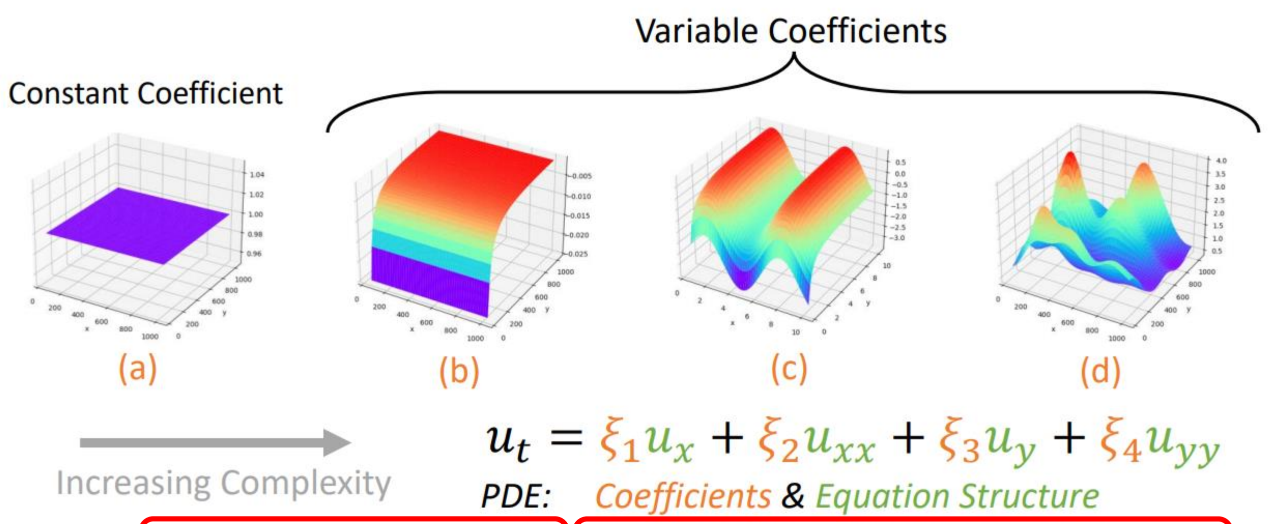
Background

Partial Differential Equations (PDEs) are ubiquitous.

Based on observations, scientists explain electromagnetism by:

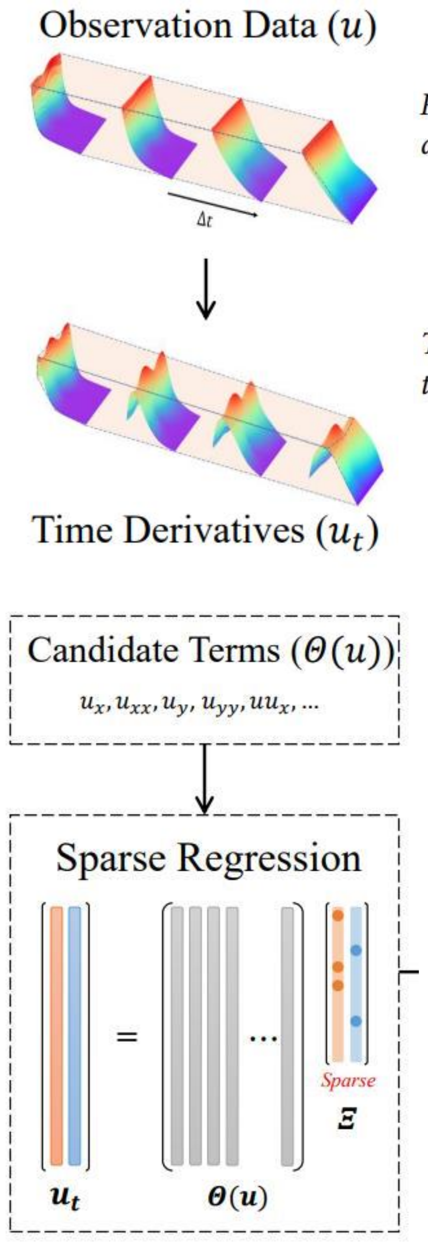
$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}$$
$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$
$$\nabla \cdot \boldsymbol{B} = 0$$
$$\nabla \cdot \boldsymbol{D} = \rho$$

Problem & Challenges



(a) 1, **(b)** $-1/x$, **(c)** $-1/x + \sin y$, **(d)** Highly Nonlinear (HIN) coefficient

PDE-Net/SGTRidge/A-DLGA — Ours



DEFINITION 1 (PDES WITH CONSTANT COEFFICIENTS, PDES-CC). *PDEs-CC are the simplest PDEs, whose coefficients ξ_i are fixed along all coordinates:*

$$u_t = \sum_{i=1}^{p} \Theta(u)_i \xi_i,\ \Theta(u)_i \in [1, u, u_x, u_y, u_{xx}, \ldots, uu_x, \ldots]. \quad (1)$$

DEFINITION 2 (PDES WITH VARIABLE COEFFICIENTS, PDES-VC). *The coefficients of PDEs-VC are changing in some dimensions, e.g., the spatial dimensions:*

$$u_t = \sum_{i=1}^{p} \Theta(u)_i \xi_i(x, y), \Theta(u)_i \in [1, u, u_x, u_y, u_{xx}, \ldots, uu_x, \ldots]. \quad (2)$$

Accordingly, the goal of PDE discovery is to determine:

- **Terms**: which coefficient ξ_i is nonzero so that the term $\Theta(u)_i$ exists in the PDE structure;
- **Coefficients**: the exact values of all nonzero coefficients at each spatial coordinate.

Sparse Regression (filter irrelevant terms):

$$Y = XW + \epsilon, \quad \epsilon \sim \eta \mathcal{N}(0, \sigma^2) \in \mathbb{R}^h,$$

$$\widehat{W} = \underset{W}{\operatorname{argmin}} \|Y - XW\|_2^2 + \lambda \|W\|_0,$$

Independent coefficient estimation across space

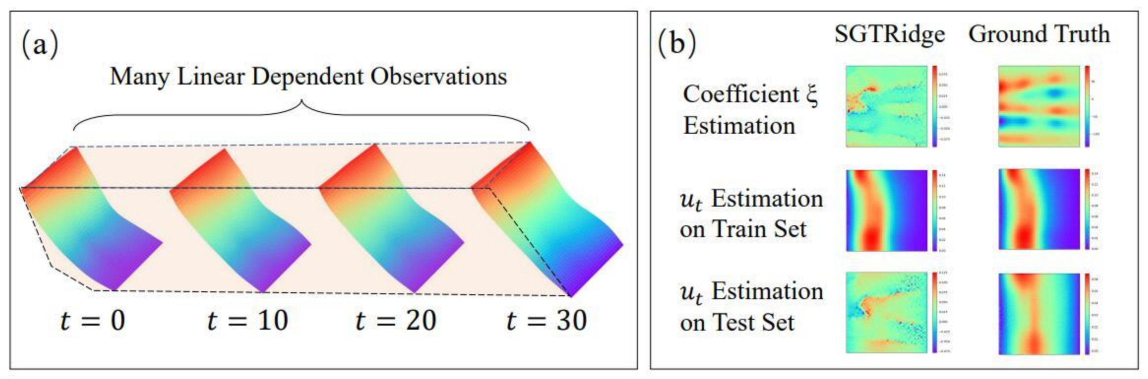


Figure 2: The linear dependent observations and data quality issues cause the overfitting of baselines such as SGTRidge [44]. The estimated PDE coefficients are fairly irregular, and cannot match the ground truth. Although it can fit the training data well, it fails to generalize to the test data.

Motivations

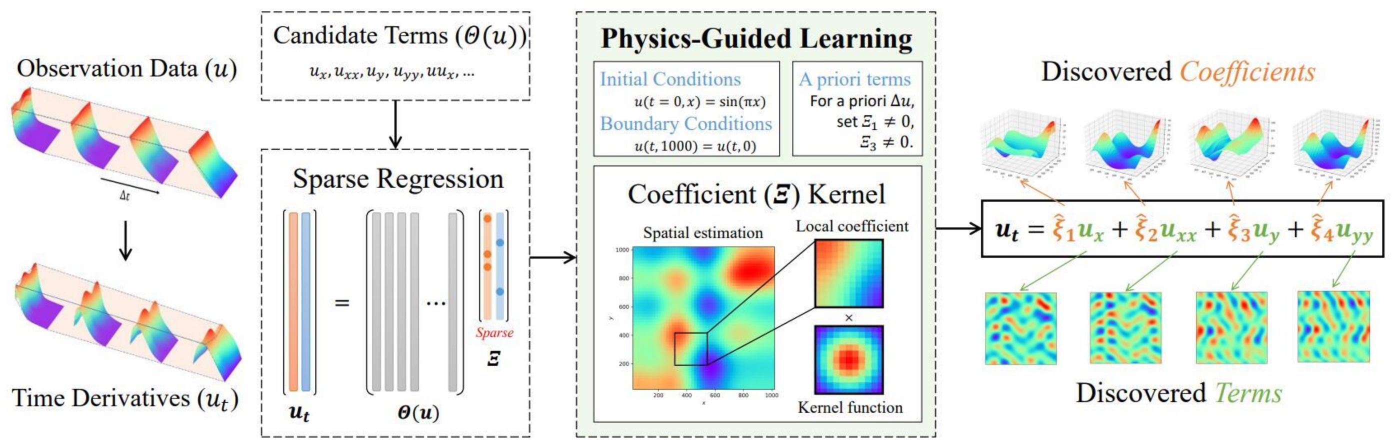
Physics Knowledge as Constraints:
Initial and boundary conditions, a priori terms.

Physics Principles as Constraints:
Conservation laws, smoothness.

DEFINITION 3 (LOCAL SMOOTHNESS). *Given coefficient $\xi(x, y)$, the coefficients within a local area (with radius r) can be considered as a k-Lipschitz continuous function. Given the spatial distance of any two adjacent coordinates $Dist = \|S(x, y) - S(x', y')\| \le r$ where $S(x, y)$ is the spatial coordinate vector, the slope of the coefficient function is bounded by $\alpha \ge 0$ as $\frac{|\xi(x,y) - \xi(x',y')|}{\|S(x,y) - S(x',y')\|} \le \alpha$.*

Recall: Differentiation smoothness is a first principle.
Scope: Most physical fields. For those with mutations or discontinuities, need shock/discontinuity-capturing methods).

Methodology



We denote the spatial coordinate vector as $S(x, y)$ and denote the distance between two spatial coordinates (x, y) and (x', y') as $\|S(x, j) - S(x', y')\|$. For each (x, y), the proposed model considers all (x', y') that $\|S(x, y) - S(x', y')\| < r$ to compute

$$\boxed{\widehat{W}} = \underset{W}{\operatorname{argmin}} \|Y - X\boxed{\Xi}\|_2^2, \quad (5)$$

Parameters — Coefficients

where

$$\Xi_i^{[x,y]} = \frac{\sum K_i^{[x',y']} W_i^{[x',y']}}{\sum K_i^{[x',y']}}, \quad (6)$$

$$K_i^{[x',y']} = \exp\left(-\frac{D^{[x',y']}}{2\gamma}\right), \quad (7)$$

$$D^{[x',y']} = \|S(x, y) - S(x', y')\|_2^2, \quad (8)$$

$$\mathcal{L}(W; u) = \|Y - X\Xi\|_2^2 + \boxed{\beta \|\hat{u} - u\|_2^2} + \lambda \|W\|_0, \quad (9)$$

Knowledge Constraint

Theoretical Properties

Kernel Complexity: Linearly proportional to the size of the dataset.
(The data used for estimating coefficient is "local")

Estimation Error: Strictly lower than works like Adaptive-DLGA.
(Far-away coefficients have low density in kernel)

Robustness: Strictly lower than independent coefficient estimation.
For I.I.D. Gaussian noises, $|\widehat{\Xi}^{[x,y]} - \xi^{[x,y]}| \sim \eta N(0, \sum_{(x',y')} \frac{K^{[x',y']}\sigma^2}{K^{[x',y']}})$
which is lower than $|\widehat{W}^{[x,y]} - \xi^{[x,y]}| \sim \eta N(0, \sigma^2)$.

Mesh-Free: (Kernel only calculates spatial distance. No grid is needed.)

Results – Constant Coefficient

Burgers' Equation

$$u_t = 0.005u_{xx} + 0.005u_{yy} - uu_x - vu_y,$$
$$v_t = 0.005v_{xx} + 0.005v_{yy} - uv_x - vv_y.$$

KdV Equation

$$u_t = -uu_x - 0.0025u_{xxx}.$$

C-I Equation

$$u_t = u_{xx} - u + u^3.$$

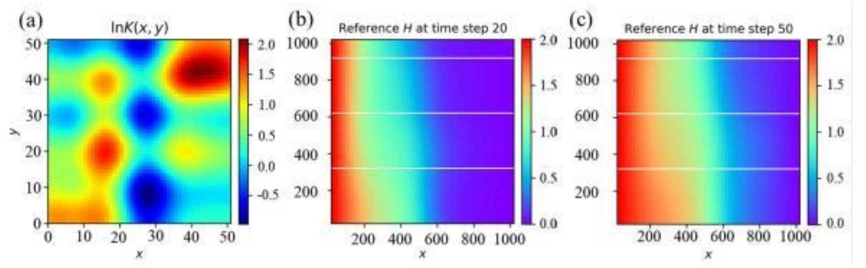
Metrics	Recall (%)			Coef. Error ($\times 10^{-3}$)			Fitting Error ($\times 10^{-3}$)		
Noise Level	0%	10%	20%	0%	10%	20%	0%	10%	20%
Burgers' Equation	100	100	100	2.603	6.124	6.946	0.205	0.356	1.004
KdV Equation	100	100	100	1.417	7.385	14.36	3.729	187.8	375.5
C-I Equation	100	100	100	3.623	12.69	25.38	1.691	11.85	23.71

Results – HIN Coefficient

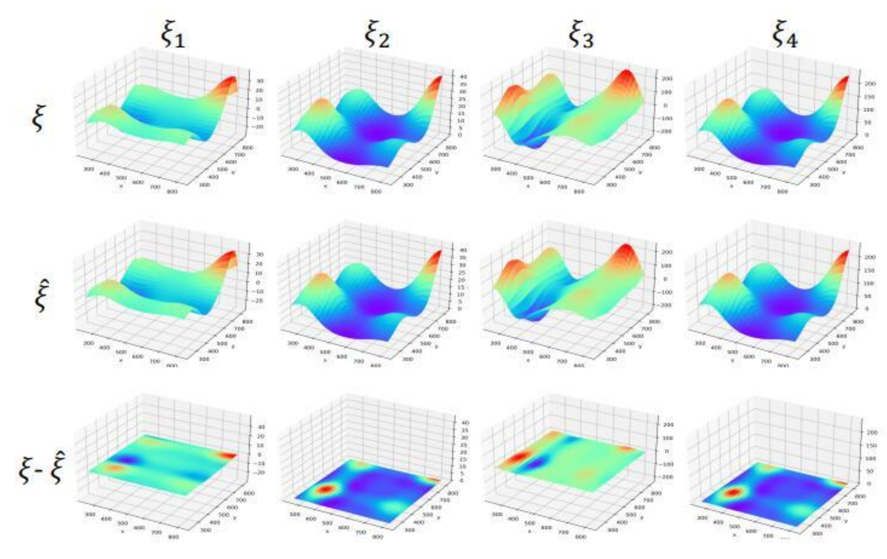
Five governing equation cases of underground seepage (a form of Convection–diffusion equations)

$$S_s \frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(K(x, y)\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(K(x, y)\frac{\partial u}{\partial y}\right).$$

where Ss denotes the specific storage; $K(x, y)$ denotes the hydraulic conductivity field; u denotes the hydraulic head. u is the physical field. K is the coefficient field as shown below. The hydraulic conductivity field $K(x, y)$ is set to be heterogeneous to simulate real situations, which is random fields with higher complexity.



Dataset	Metric	SGTRidge 5%	PDE-Net 5%	A-DLGA 5%	HIN-PDE 5%	HIN-PDE 10%	HIN-PDE 20%
1-HNC	Train Err ($\times 10^{-3}$)	1.148	2.190	16.70	3.314	6.283	11.88
	Dev Err ($\times 10^{-3}$)	3.907	10.85	47.39	3.919	6.373	12.10
	Test Err ($\times 10^{-3}$)	28.25	31.80	136.8	**3.686**	6.367	12.29
	Recall (%)	50	50	25	**100**	100	75
2-HNC	Train Err ($\times 10^{-3}$)	2.283	2.697	19.83	3.794	5.676	8.329
	Dev Err ($\times 10^{-3}$)	26.87	42.23	43.04	3.794	5.674	8.332
	Test Err ($\times 10^{-3}$)	106.1	169.2	123.8	**3.585**	5.499	8.318
	Recall (%)	25	25	25	**100**	100	75
3-HNC	Train Err ($\times 10^{-3}$)	0.134	0.129	1.033	0.331	0.583	0.969
	Dev Err ($\times 10^{-3}$)	1.588	1.563	2.661	0.342	0.588	0.997
	Test Err ($\times 10^{-3}$)	9.095	9.005	7.872	**0.343**	0.589	1.018
	Recall (%)	25	25	50	**100**	100	75
4-HNC	Train Err ($\times 10^{-3}$)	0.336	0.301	21.50	1.733	2.235	3.662
	Dev Err ($\times 10^{-3}$)	11.52	10.13	37.61	1.729	2.302	4.037
	Test Err ($\times 10^{-3}$)	94.47	85.60	150.0	**1.703**	2.521	7.505
	Recall (%)	0	0	0	**100**	75	25
5-HNC	Train Err ($\times 10^{-3}$)	1.218	1.506	20.66	1.940	3.139	4.438
	Dev Err ($\times 10^{-3}$)	6.624	7.508	42.72	1.984	2.984	4.254
	Test Err ($\times 10^{-3}$)	13.63	15.17	109.6	**1.733**	3.049	4.345
	Recall (%)	50	50	50	**100**	100	100



State-of-the-arts in all three metrics:

(1) Recall of the discovered PDE terms compared to ground truth;
(2) The mean absolute error of coefficient ξ estimation;
(3) The mean absolute error of target u_t fitting.