Buckingham π-Invariant Test-Time Projection for Robust PDE Surrogate Modeling

Seokki Lee[1*], Min-Chul Park[2*], Giyong Hong[1*], Changwook Jeong[1,3]

[1]Graduate School of Semiconductor Materials and Devices Engineering, UNIST,
[2]CSE Team, Samsung Electronics, [3]AI Graduate School, UNIST

UNIST ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

SAMSUNG

ICLR International Conference On Learning Representations

Scale-shift OOD may not be a new physics

- Why many PDE surrogates fail under scale-shift OOD (i.e., different units, magnitudes)?

→ The model implicitly treats ***"numerically different"*** as ***"physically different"***

- However, with classical dimensional analysis, scale changes can leave the underlying physics **unchanged.**

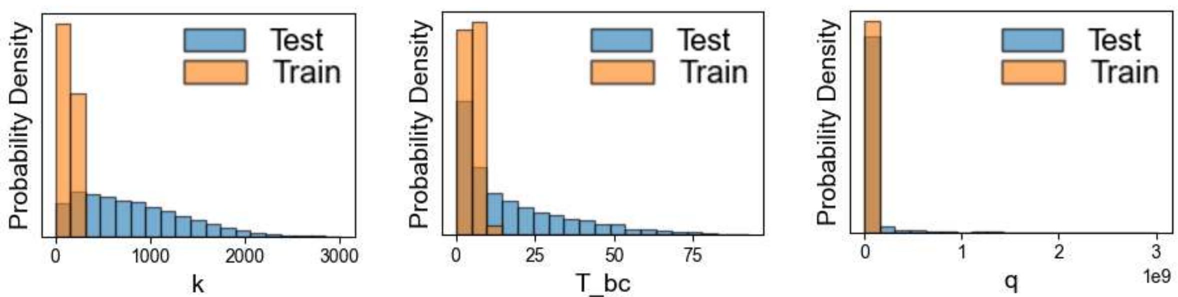
Scale-shift OOD may not be a new physics

- Why many PDE surrogates fail under scale-shift OOD (i.e., different units, magnitudes)?

→ The model implicitly treats ***"numerically different"*** as ***"physically different"***

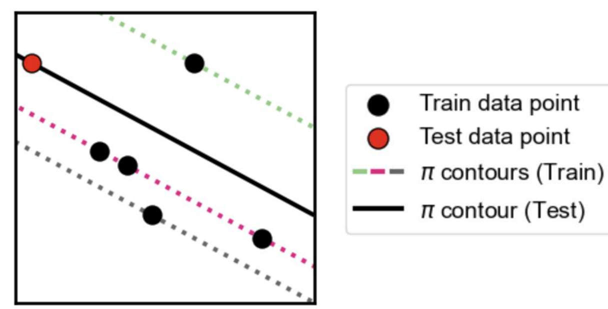
- However, with classical dimensional analysis, scale changes can leave the underlying physics **unchanged.**

Parameter space



Scale-shift OOD

π-space



Physically In-distribution

Buckingham π Theorem

- A theory that **transforms** physical equation with various inputs into **relationships** between **dimensionless** variables (π-groups).

Theorem 1 ((Buckingham π, log form)). *Given B base units, p dimensional variables $x \in \mathbb{R}^p_{>0}$ and a dimension matrix $D \in \mathbb{R}^{|B|\times p}$ (rows = base units, columns = variables), let $r = \text{rank}(D)$. Then there exist $p - r$ independent dimensionless combinations ("π-groups"). If $\Phi \in \mathbb{R}^{p\times(p-r)}$ spans $\ker(D)$, then*

$$\log \Pi(x) = \Phi^\top \log x \iff \Pi(x) = \exp(\Phi^\top \log x)$$

is a complete set of $p - r$ independent π-groups. (See App. A for details.)

Notation. $\Phi = [\phi^{(1)} \cdots \phi^{(p-r)}]$ stacks the null-space basis vectors; each column $\phi^{(\ell)}$ defines one dimensionless monomial (a π-group). The set of log-rescalings that preserve all π-values is $\ker(\Phi^\top) = \{v \in \mathbb{R}^p : \Phi^\top v = 0\}$.

Edgar Buckingham

https://en.wikipedia.org/wiki/Buckingham_pi_theorem

Insights from the Buckingham π theorem: π-equivalence

"A physical system can be expressed using dimensionless π-groups"

Theorem 1 ((Buckingham π, log form)). *Given B base units, p dimensional variables $x \in \mathbb{R}^p_{>0}$ and a dimension matrix $D \in \mathbb{R}^{|B|\times p}$ (rows = base units, columns = variables), let $r = \mathrm{rank}(D)$. Then there exist $p-r$ independent dimensionless combinations ("π-groups"). If $\Phi \in \mathbb{R}^{p\times(p-r)}$ spans $\ker(D)$, then*

$$\log \Pi(x) = \Phi^\top \log x \quad \Longleftrightarrow \quad \Pi(x) = \exp(\Phi^\top \log x)$$

is a complete set of $p-r$ independent π-groups. (See App. A for details.)

Notation. $\Phi = [\phi^{(1)} \cdots \phi^{(p-r)}]$ stacks the null-space basis vectors; each column $\phi^{(\ell)}$ defines one dimensionless monomial (a π-group). The set of log-rescalings that preserve all π-values is $\ker(\Phi^\top) = \{v \in \mathbb{R}^p : \Phi^\top v = 0\}$.

Edgar Buckingham

https://en.wikipedia.org/wiki/Buckingham_pi_theorem

Insights from the Buckingham π theorem: π-equivalence

"A physical system can be expressed using dimensionless π-groups"

Theorem 1 ((Buckingham π, log form)). *Given B base units, p dimensional variables $x \in \mathbb{R}^p_{>0}$ and a dimension matrix $D \in \mathbb{R}^{|B|\times p}$ (rows = base units, columns = variables), let $r = \mathrm{rank}(D)$. Then there exist $p - r$ independent dimensionless combinations ("π-groups"). If $\Phi \in \mathbb{R}^{p\times(p-r)}$ spans $\ker(D)$, then*

$$\log \Pi(x) = \Phi^\top \log x \iff \Pi(x) = \exp(\Phi^\top \log x)$$

is a complete set of $p - r$ independent π-groups. (See App. A for details.)

Notation. $\Phi = [\phi^{(1)} \cdots \phi^{(p-r)}]$ stacks the null-space basis vectors; each column $\phi^{(\ell)}$ defines one dimensionless monomial (a π-group). The set of log-rescalings that preserve all π-values is $\ker(\Phi^\top) = \{v \in \mathbb{R}^p : \Phi^\top v = 0\}$.

Edgar Buckingham

- The solutions depend on inputs via π. Then, the scale-shift can look **OOD** in parameter space, yet be **in-domain** in π-space.

"π-equivalent" = "*physically equivalent*"

https://en.wikipedia.org/wiki/Buckingham_pi_theorem

Calculation of Buckingham π-invariants

- Worked example – thermal conduction

$$-\nabla \cdot (k\nabla T) = q$$

Dimension matrix $D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 \\ -3 & -3 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix}$

Columns: $\{k, q, \Delta T, L\}$ / Rows: $\{M, L, T, \Theta\}$

↑ k [W/mK], q [W/m^3], ΔT [K], L [m]

↑ Base units (kg, m, sec, K)

↓

$D\Phi = 0$ (Dimensionless condition)

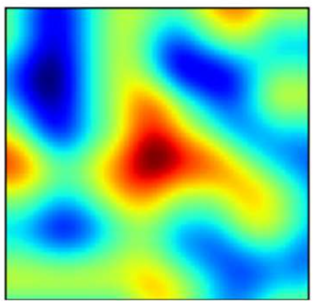
$$\Phi = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 2 \end{bmatrix}$$

→

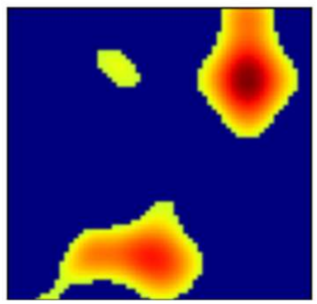
$$\pi_{th} = \frac{qL^2}{k\Delta T} = \frac{(W/m^3)(m^2)}{(W/mK)(K)} = \frac{W/m}{W/m} = 1 \quad \text{(Dimensionless)}$$

Domain-profile Reduction for Stable π-groups

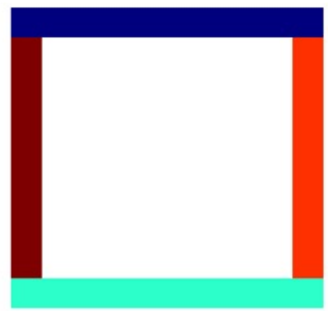
- Using **pixel-wise** π-groups can cause **π-information loss** and **π-degeneracy ($\pi \to 0$).**



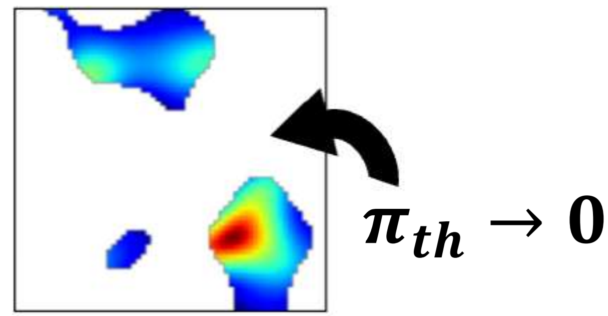
Thermal conductivity (k)



Volumetric heat source (q)



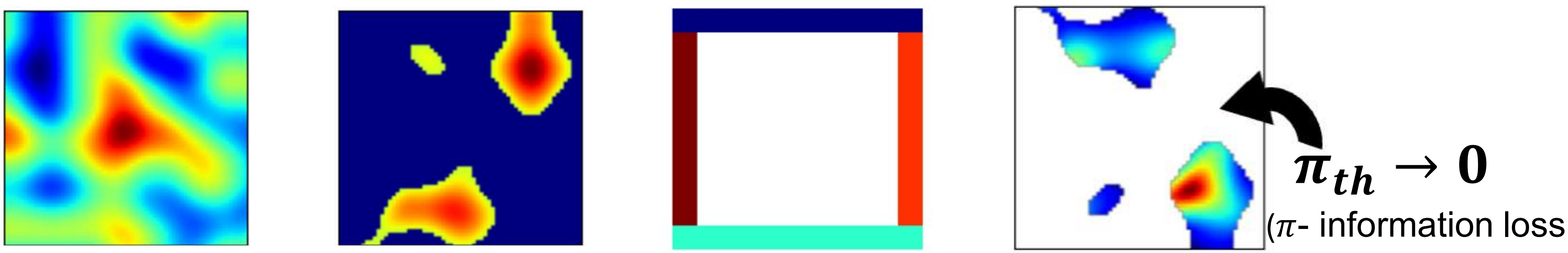
Boundary condition



Pixel-wise π values

Domain-profile Reduction for Stable π-groups

- Using **pixel-wise** π-groups can cause **π-information loss** and **π-degeneracy ($\pi \rightarrow 0$).**



Thermal conductivity (k) | Volumetric heat source (q) | Boundary condition | Pixel-wise π values

→ The issue can be resolved by **representative values** (e.g., **mean**) through domain profile reduction.

$$\bar{\boldsymbol{k}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} k_{ij} \quad \bar{\boldsymbol{q}} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} q_{ij} \quad \Delta T = T_{max} - T_{min}$$

$$\pi_{th} = \frac{\bar{q} L^2}{\bar{k} \Delta T}$$

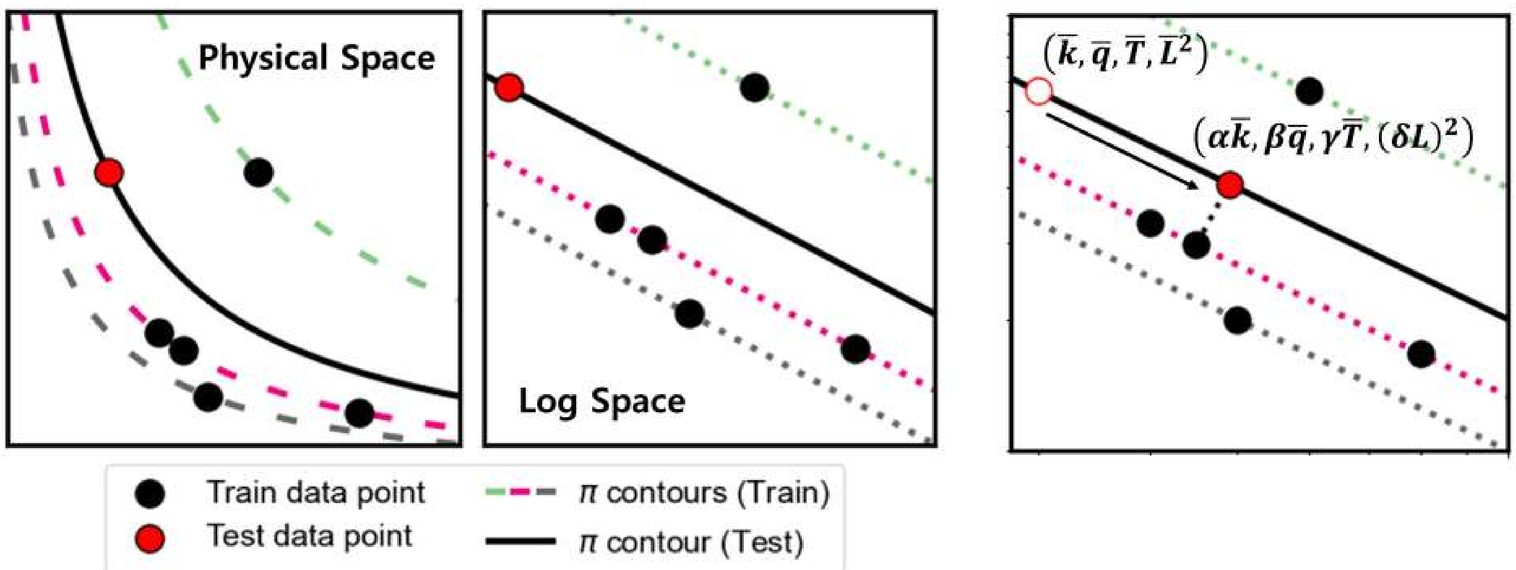
(image-wise π-value)

Definition of π-equivalence class

- A π-equivalence class is a **scale orbit.**

$$[x]_\pi = \{x \odot exp(v) \mid v \in ker(\Phi^{\mathrm{T}})\}$$

,where x is a set of parameters, and v is a set of log-scale scaling vectors that **preserves** π-values.



Decomposition for nearest π–equivalence class search

- **Goal:** Find the **nearest** training π-equivalence class (physics match), then **adjust** scale within the class (projection).

- A π-equivalence class (**log-space**):
$$[z]_\pi = z + \ker(\Phi^T), \text{ where } z = \log x$$

- Unconstrained scaling vector:
$$v_i^t = z_i - \tilde{z}, \text{ where } z_i \text{ for train sample } i, \text{ and } \tilde{z} \text{ for test}$$

- Key idea: **decompose** v_i^t into intra-class (scaling, **π-preserving**) + inter-class (**physics change**).

Decomposition for nearest π–equivalence class search

- Projector $P_{||}$: extracts the **π-preserving (scaling) component** by projecting onto $\ker(\Phi^T)$ and the residual $(I - P_{||})$ measures **physics mismatch.**

$$P_{||} = I - \Phi(\Phi^T\Phi)^{-1}\Phi^T \text{ (or } P_{||} = I - \Sigma^{-1}\Phi(\Phi^T\Sigma^{-1}\Phi)^{-1}\Phi^T \textit{ with a weight } \Sigma)$$

- Decomposition of unconstrained scaling vector:

$$v_i^t = \underbrace{P_{||}v_i^t}_{\text{Intra-class (scaling, } \pi\text{-preserving)}} + \underbrace{(I - P_{||})v_i^t}_{\text{Inter-class (changes } \pi)}$$

Decomposition for nearest π–equivalence class search

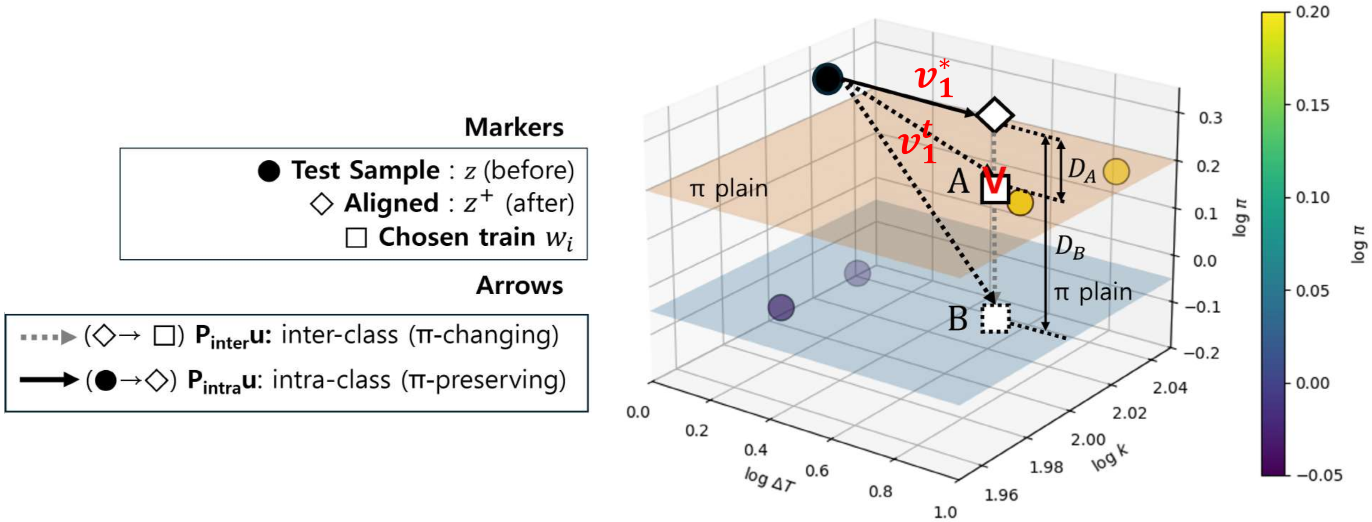
- Optimum train sample to be projected (the **nearest** π-equivalence class)

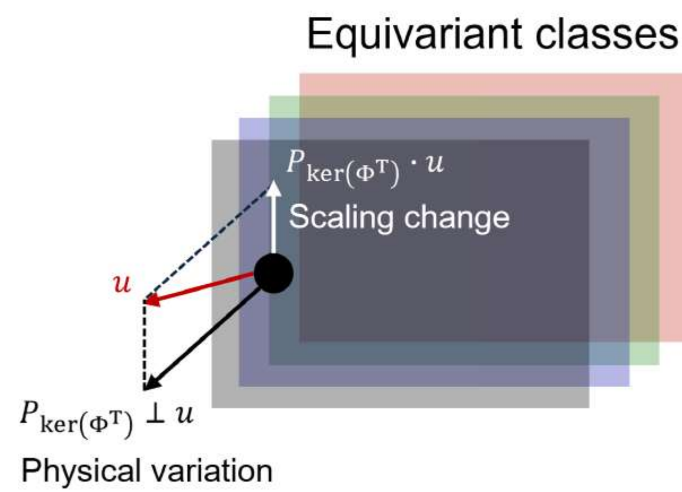
$$i^* = \underset{i=1,2,\dots,M}{\operatorname{argmin}} \left\| v_i^t - v_i^* \right\|_2$$

Decomposition for nearest π–equivalence class search

- Optimum train sample to be projected (the **nearest** π-equivalence class)

$$i^* = \underset{i=1,2,\ldots,M}{\operatorname{argmin}} \left\| v_i^t - v_i^* \right\|_2$$

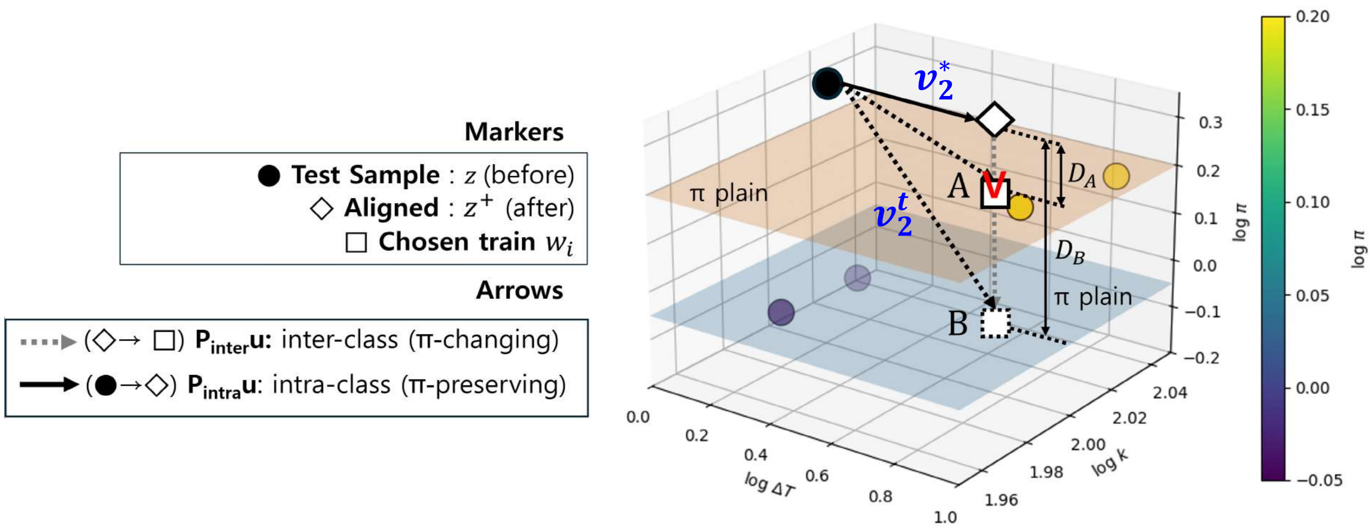


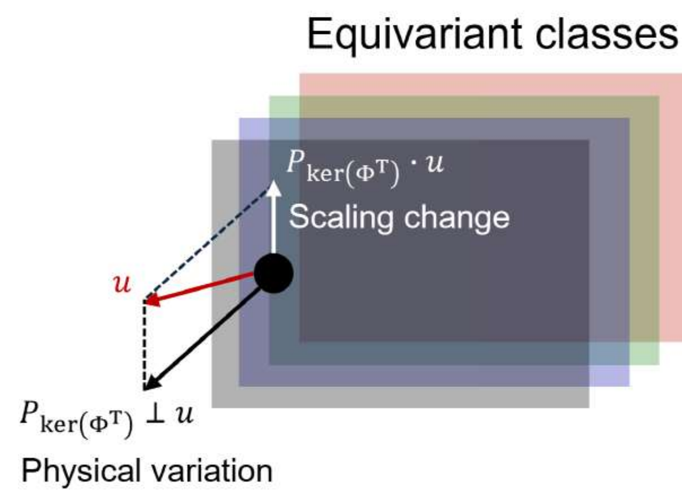


Decomposition for nearest π–equivalence class search

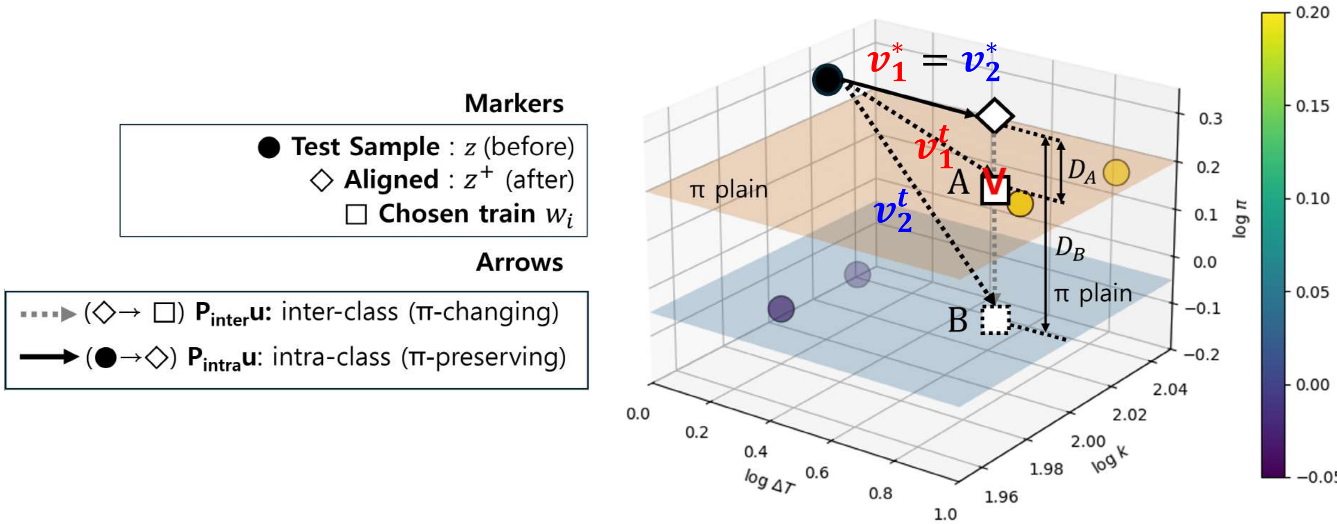
- Optimum train sample to be projected (the **nearest** π-equivalence class)

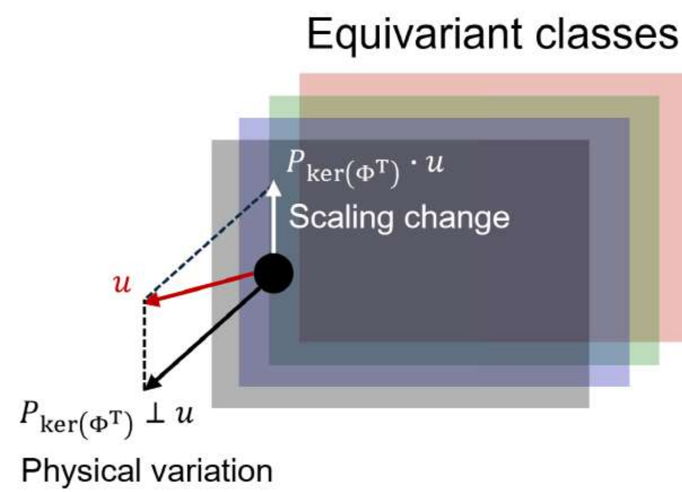
$$i^* = \underset{i=1,2,\ldots,M}{\operatorname{argmin}} \left\| v_i^t - v_i^* \right\|_2$$





Decomposition for nearest π–equivalence class search





- Quotient distance (**class-class**) comparison:

$$D_A = \|v_1^t - v_1^*\|_2 < D_B = \|v_2^t - v_2^*\|_2$$

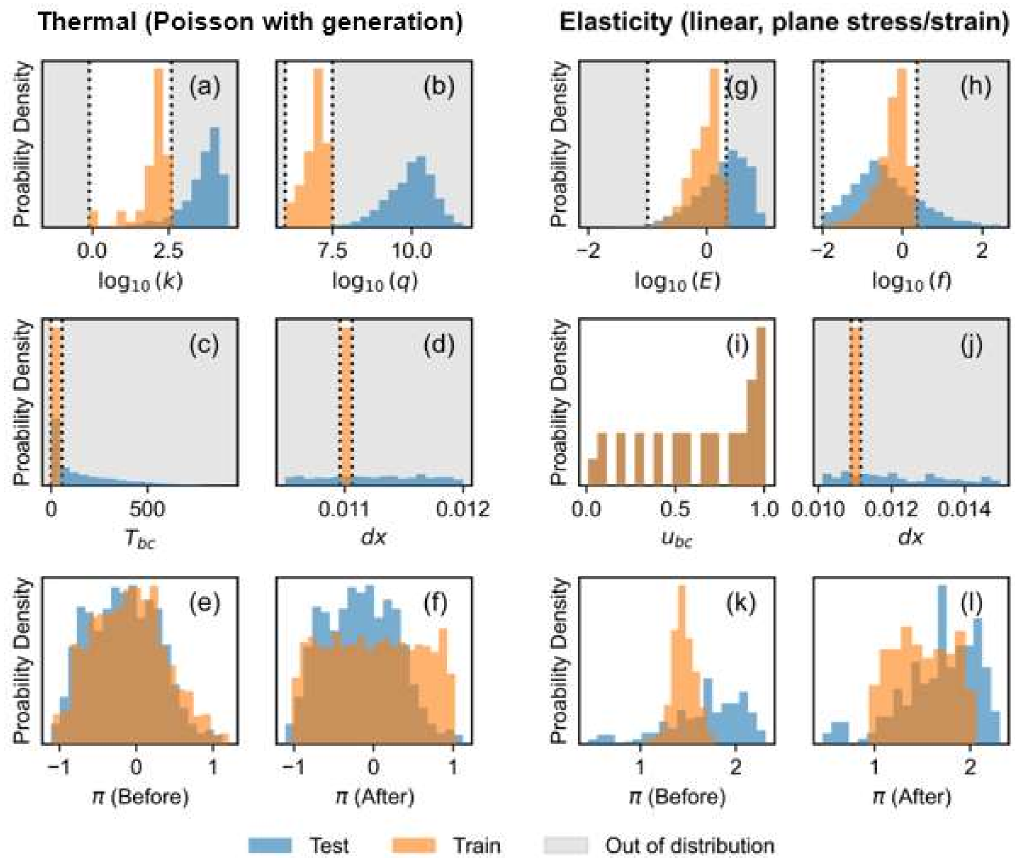
→ The train sample A is **the optimum** sample to be **projected** (the **nearest** π-equivalence class).

- The projected (scaled) test sample is finally given by

$$\tilde{x}^* = \exp(\tilde{z}^*) = \exp(\tilde{z} + v^*) = \tilde{x} \odot \exp(v^*)$$

Buckingham π–Invariant Test−Time Projection

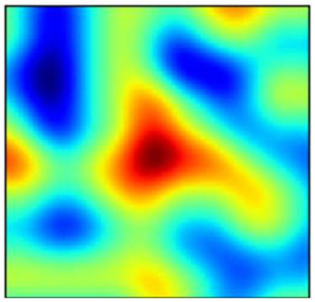
- Given OOD test inputs with different **scales**, predict PDE solutions w/o **retraining.**
- Problem definition:



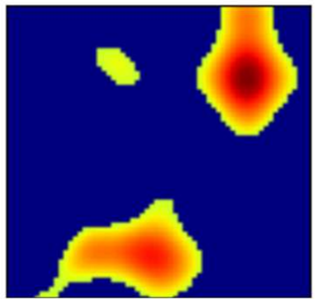
Buckingham π–Invariant Test−Time Projection

- **Step 1.** Domain Profile Reduction

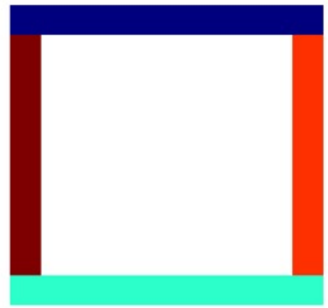
→ extract feature vectors to form **dimensionless $\boldsymbol{\pi}$** using feature extractor ψ.



Thermal conductivity (k)



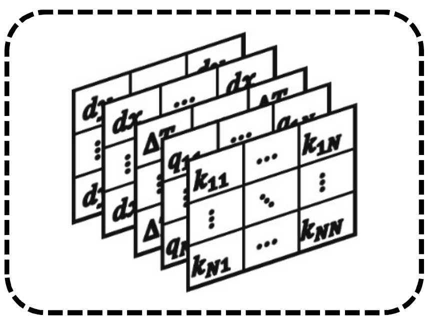
Volumetric heat source (q)



Boundary condition

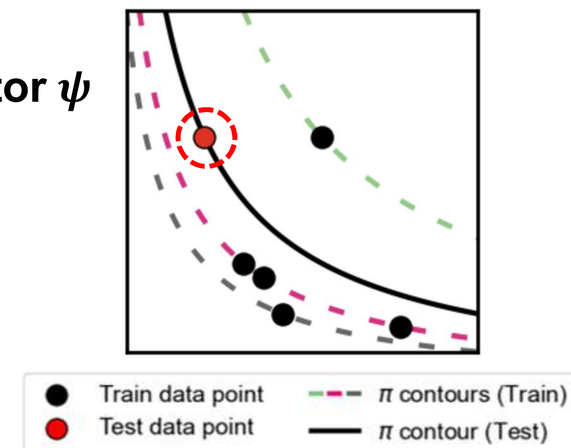
$$\pi_{th} = \frac{\bar{q}L^2}{\bar{k}\Delta T}$$

(image-wise π-value)



Physical input parameters

Feature extractor $\boldsymbol{\psi}$



Buckingham π–Invariant Test−Time Projection

- **Step 2.** π-Invariant Projection

→ **rescale** a test sample $\tilde{\boldsymbol{x}}$ into **the nearest** train sample x^* based on quotient distance calculation.

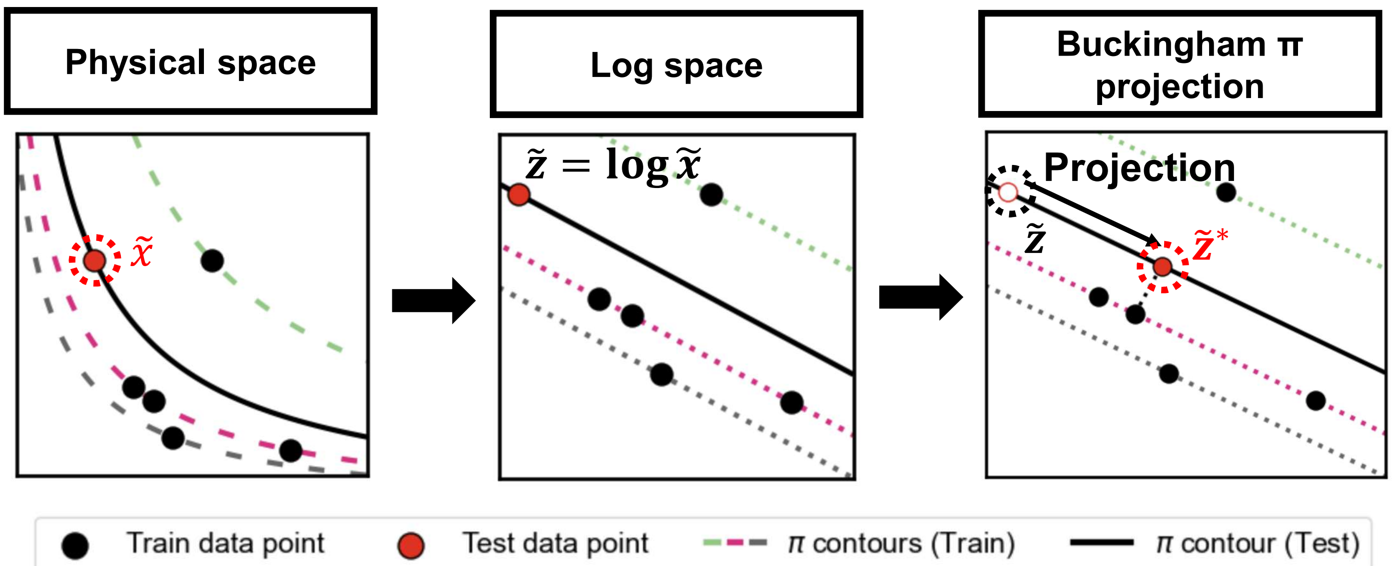
- Leverage **log-space** for efficient **calculation** & preserving π-space **geometry.**

Buckingham π–Invariant Test–Time Projection

- **Step 2.** π-Invariant Projection

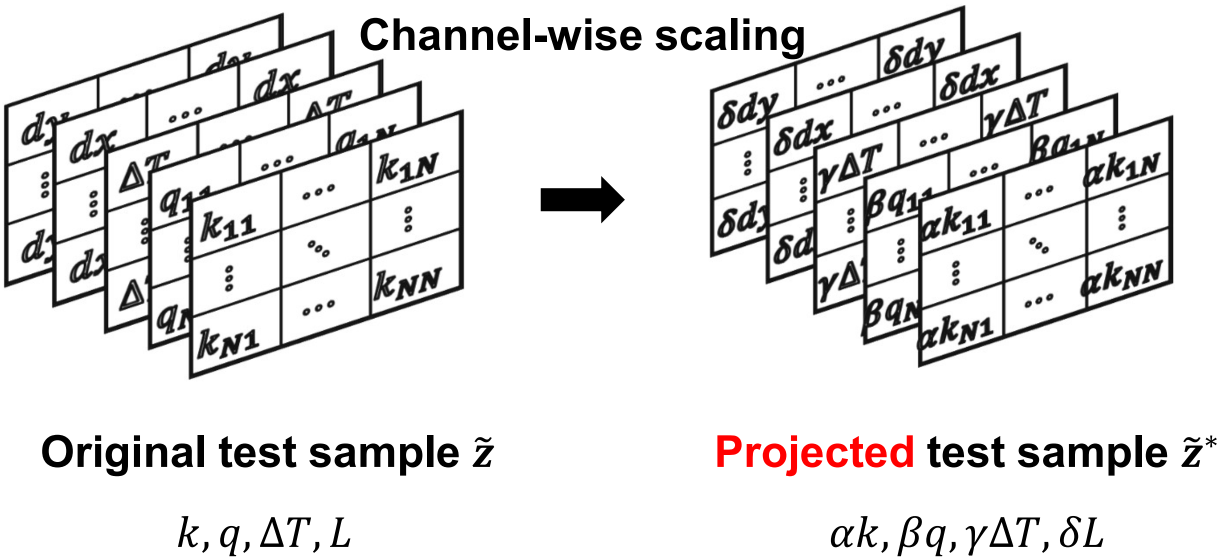
→ **rescale** a test sample $\tilde{x}$ into **the nearest** train sample x^* based on quotient distance calculation.

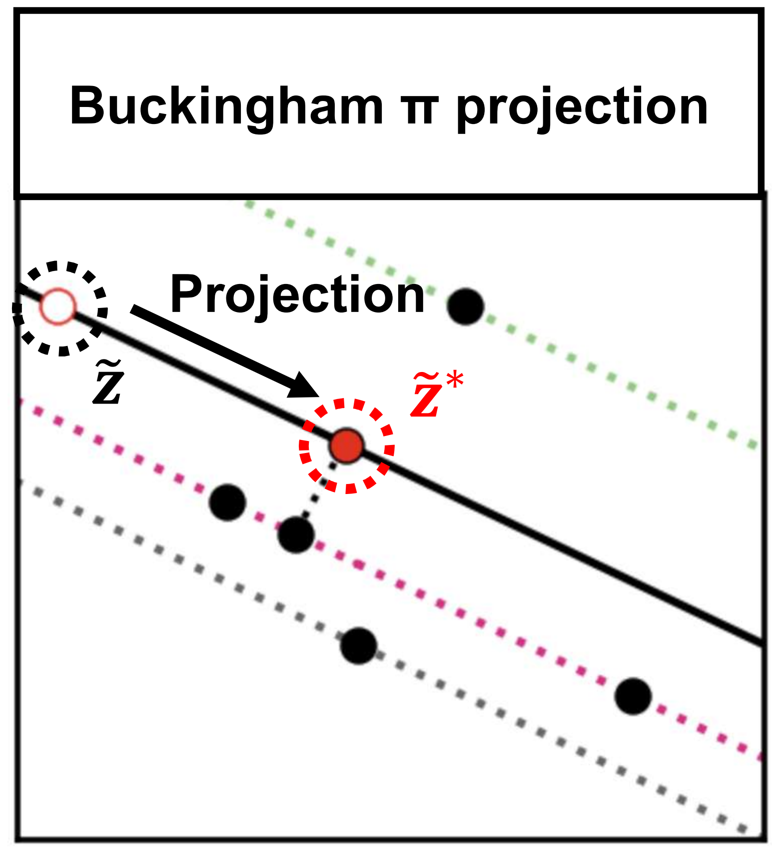
- Leverage **log-space** for efficient **calculation** & preserving π-space **geometry.**



Buckingham π–Invariant Test−Time Projection

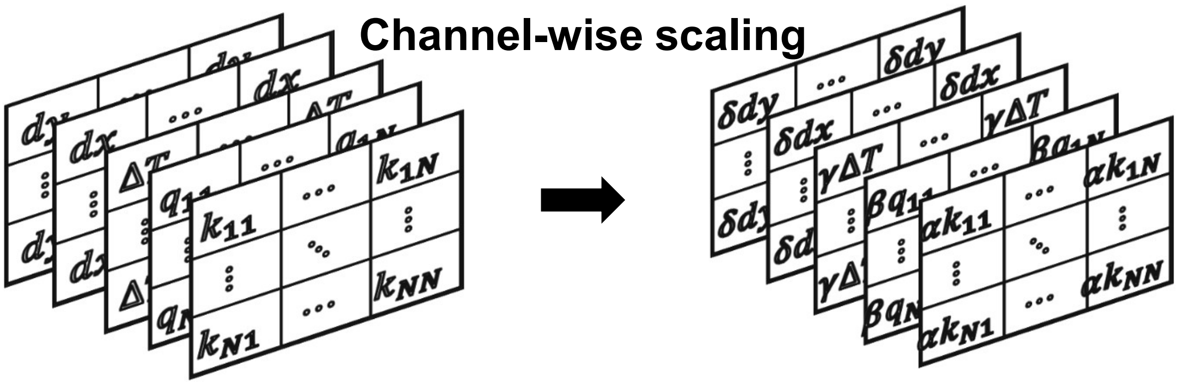
- How projection acts?**:** apply **scaling coefficients** to input parameters **of test** $\tilde{z}$.





Buckingham π–Invariant Test−Time Projection

- How projection acts?**:** apply **scaling coefficients** to input parameters **of test** $\tilde{z}$.

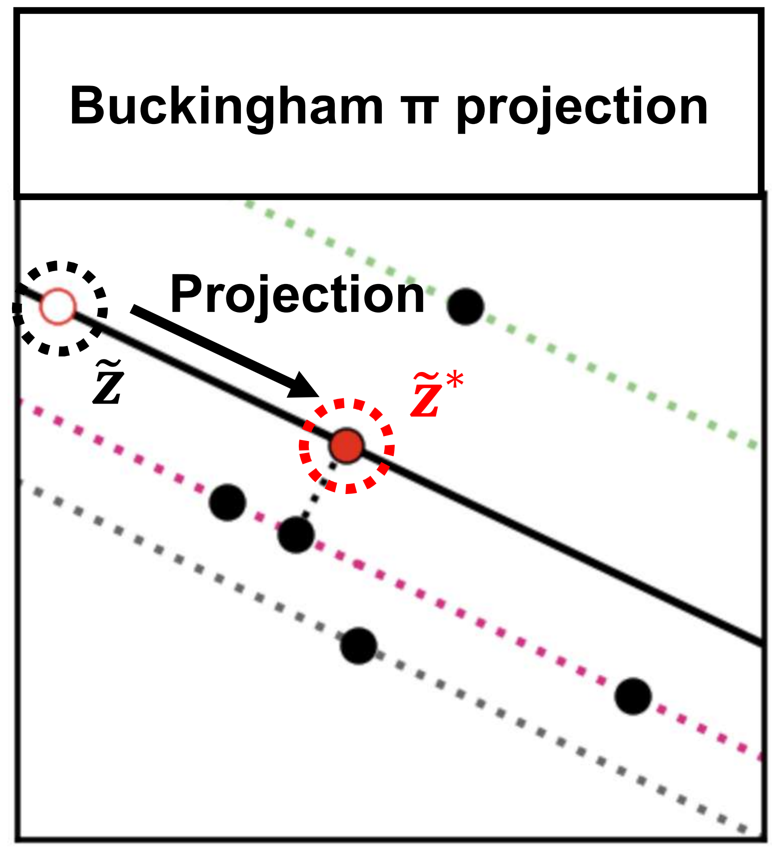


Original test sample $\tilde{z}$

$k, q, \Delta T, L$

Projected test sample $\tilde{z}^*$

$\alpha k, \beta q, \gamma \Delta T, \delta L$



- $\boldsymbol{\pi}$**-preserving** constraint:

$$\frac{qL^2}{k\Delta T} = \frac{(\beta q)(\delta L)^2}{(\alpha k)(\gamma \Delta T)}, \quad where\ \boldsymbol{\beta\delta^2 = \alpha\gamma}$$

$$\mathbf{\log\boldsymbol{\beta} + 2\log\boldsymbol{\delta} - \log\boldsymbol{\alpha} - \log\boldsymbol{\gamma} = 0}$$

Buckingham π–Invariant Test−Time Projection

- **Step 3.** Prediction on projected inputs & Inverse-scaling outputs.
- Prediction on **scaled** inputs with surrogate f_θ.

$$\widehat{\boldsymbol{Y}}_{\boldsymbol{scaled}} = \boldsymbol{f}_{\boldsymbol{\theta}}(\widetilde{\boldsymbol{X}}^{*})$$

- Inverse-scaling: **inverse-scale** the prediction results.

$$\widetilde{\boldsymbol{Y}}_{\boldsymbol{orig}} = \boldsymbol{\rho}(\boldsymbol{g})^{-\mathbf{1}}\, \widetilde{\boldsymbol{Y}}_{\boldsymbol{scaled}}$$

($p(g)$: **output-side** scaling operator)

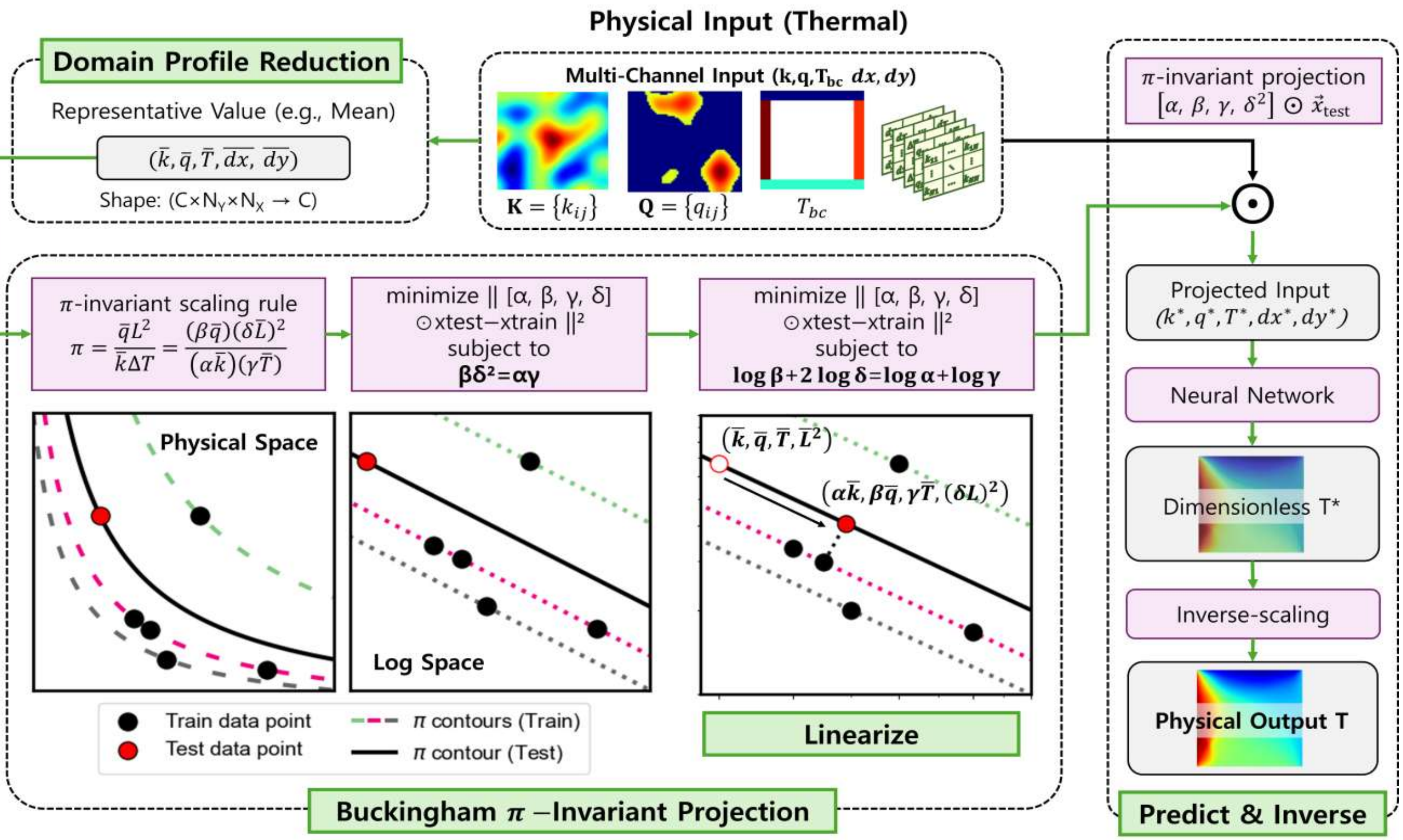
⬇

$$\widehat{\boldsymbol{T}}_{\boldsymbol{orig}} = \boldsymbol{T}_{\boldsymbol{min}} + \boldsymbol{\gamma}^{-\mathbf{1}}(\widehat{\boldsymbol{T}}_{\boldsymbol{scaled}} - \boldsymbol{T}_{\boldsymbol{min}})$$ **(Thermal)**

($T_{\min}$: minimum Dirichlet boundary temperature)

Buckingham π–Invariant Test−Time Projection

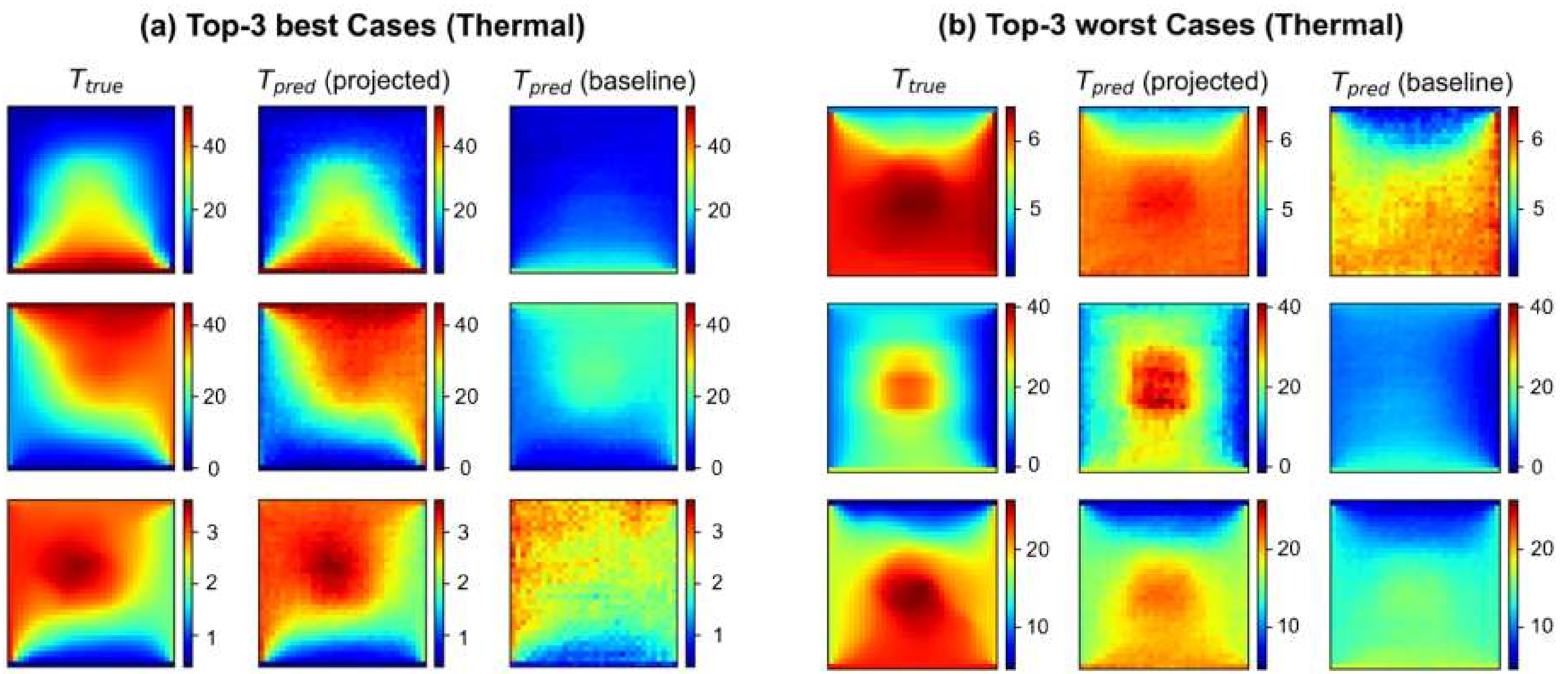
- Flowchart of Buckingham π-Invariant Test-Time Projection



Buckingham π–Invariant Test−Time Projection

- The projection shows **robustness** under scale-shift OOD

→ Projection method can predict the **trend** to some extent even in the **worst-cases.**



Toward Practical OOD-Robust PDE Surrogates

- While projection method shows robustness under scale-shift OOD, the **computational cost** for searching the nearest sample **increases** as the number of train/test samples increases.

→ We adopt π-uniform strategy and clustering projection.

Toward Practical OOD-Robust PDE Surrogates

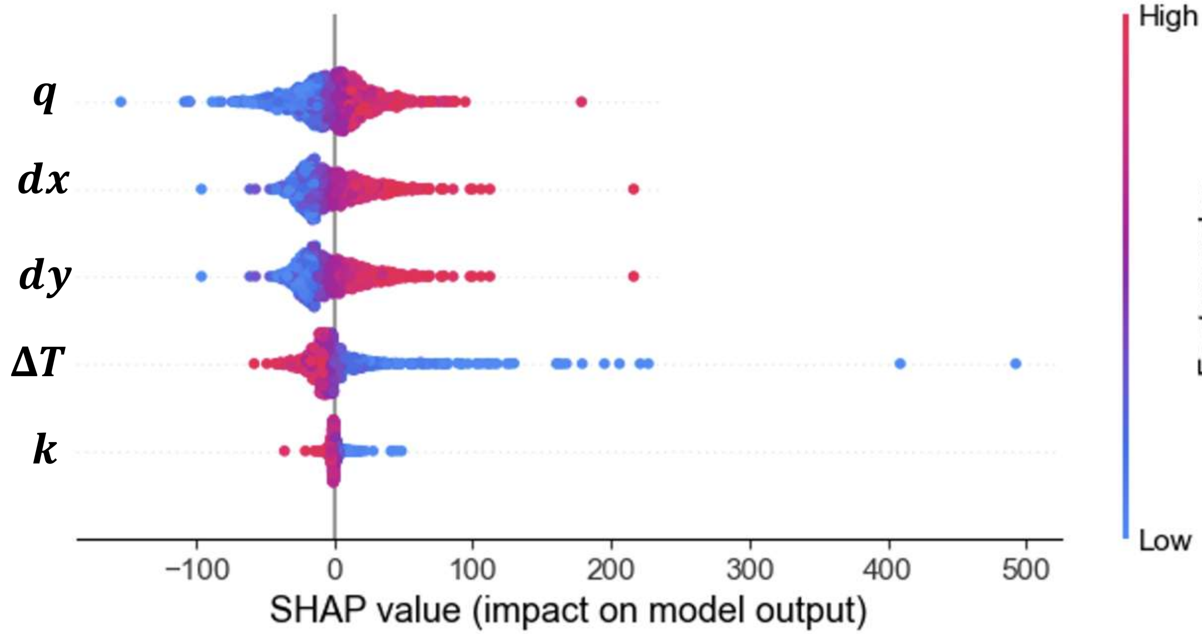
- π-uniform strategy **uniformizes** the sample-wise π **distribution** in train samples by tuning the **dominant scale** input (e.g, $\boldsymbol{q}$ in thermal conduction).

→ Dominant parameter is defined as the input parameter that **directly** controls the Buckingham π-group and tune it for uniform distribution ensures **uniform $\boldsymbol{\pi}$-distribution.**

Toward Practical OOD-Robust PDE Surrogates

- π-uniform strategy **uniformizes** the sample-wise π **distribution** in train samples by tuning the **dominant scale** input (e.g, $\boldsymbol{q}$ in thermal conduction).

→ Dominant parameter is defined as the input parameter that **directly** controls the Buckingham π-group and tune it for uniform distribution ensures **uniform $\boldsymbol{\pi}$-distribution.**

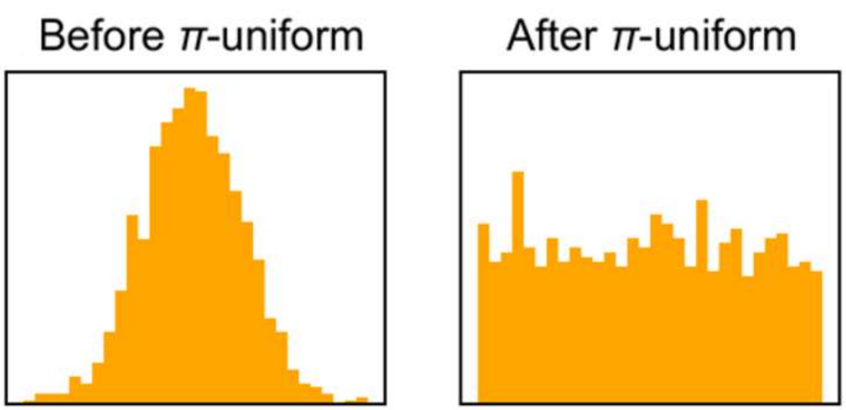


Parameters	SHAP values
q	48.7%
dx	16.4%
dy	16.4%
ΔT	6.0%
k	12.5%

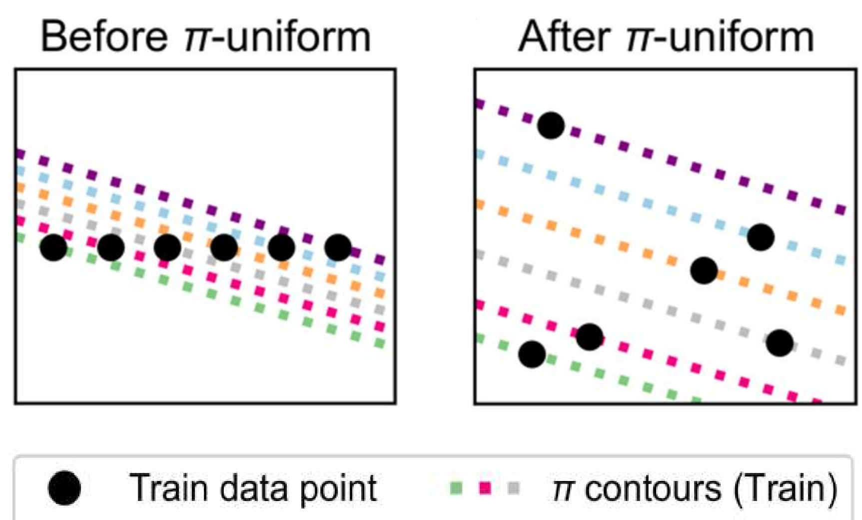
Toward Practical OOD-Robust PDE Surrogates

- π-uniform strategy **uniformizes** the sample-wise π **distribution** in train samples by tuning the **dominant scale** input (e.g, $\boldsymbol{q}$ in thermal conduction).

→ The set of centroids can **cover** the distribution of all train samples.



Log-scale π-distribution

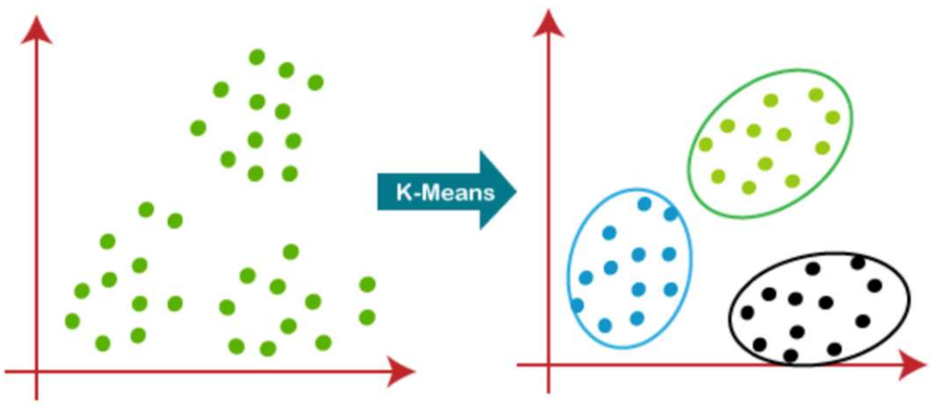


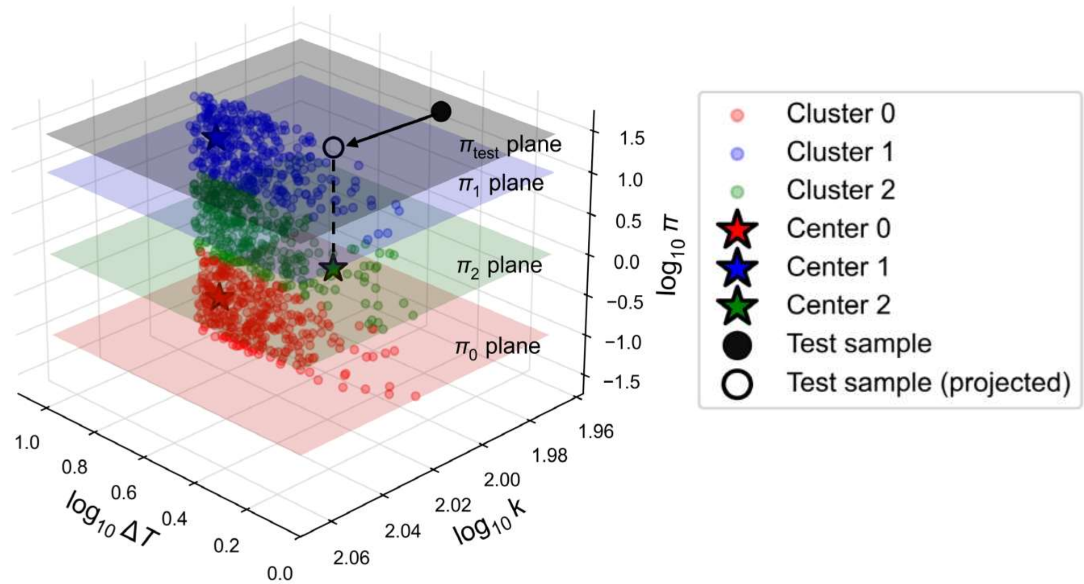
Toward Practical OOD-Robust PDE Surrogates

- Clustering Projection via Centroids:

The vanilla projection method search the entire M train samples for each N test samples, the clustering projection search **only among K centroids** from clusters.

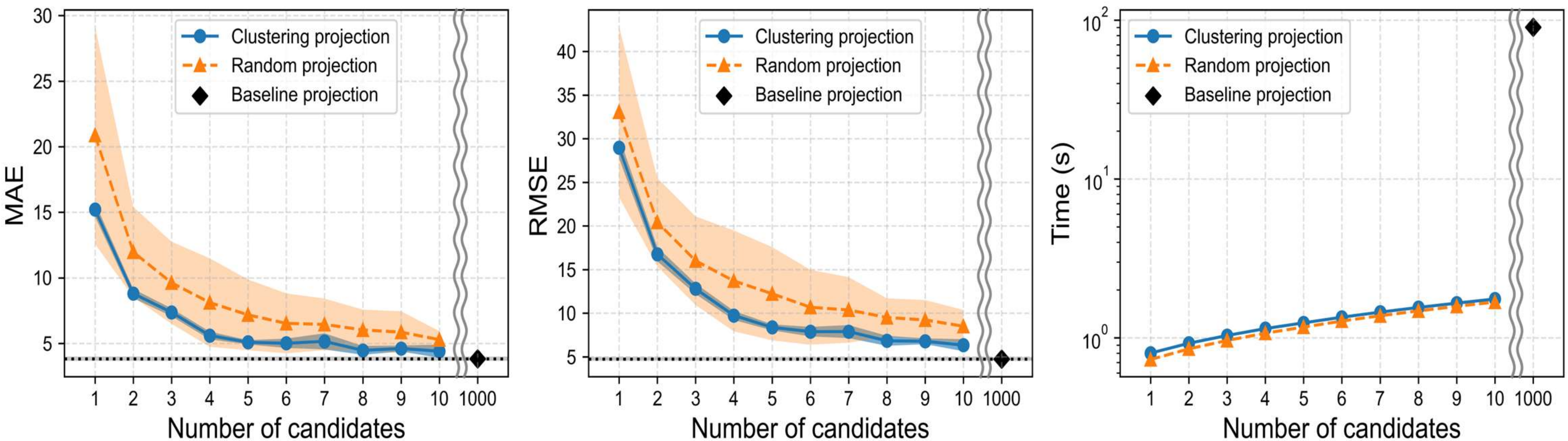
→ Computational cost **reduces** from $\mathcal{O}(MN)$ to $\mathcal{O}(KN)$ $(M > K)$





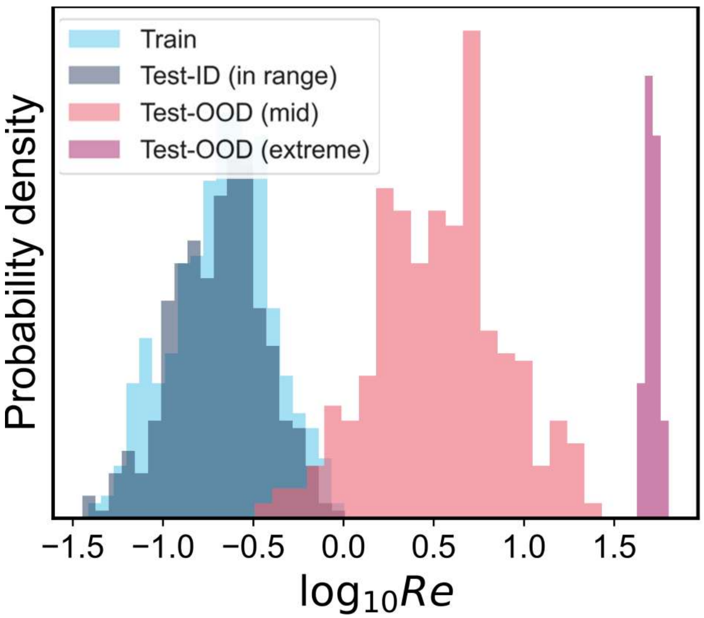
Practical Takeaways

- The clustering projection retains **robustness** on scale-shift OOD with **lower computation cost** for **practical deployment.**



Robustness on OOD in dimensional space

- Even when OOD lies **far outside** the training π**-range** (OOD in physics), the projection **still improves** predictions-though gains shrink as OOD becomes extreme.



Method	Test-ID (in-range)			Test-OOD (mid)			Test-OOD (extreme)		
	MAE	RMSE	R2	MAE	RMSE	R2	MAE	RMSE	R2
CNN	0.02	0.03	0.02	0.20	0.28	0.12	0.38	0.59	-0.04
CNN + Pairwise Projection	**0.005**	**0.009**	**0.82**	**0.08**	**0.13**	**0.76**	**0.34**	**0.52**	**0.18**
U-Net	0.01	0.02	0.50	0.23	0.30	0.10	0.38	0.54	0.18
U-Net + Pairwise Projection	**0.003**	**0.004**	**0.97**	**0.06**	**0.09**	**0.92**	**0.24**	**0.32**	**0.70**
FNO	0.006	0.007	0.95	0.20	0.26	0.39	0.33	0.45	0.14
FNO + Pairwise Projection	**0.002**	**0.002**	**0.99**	**0.04**	**0.05**	**0.97**	**0.21**	**0.27**	**0.78**

For more details

- The **complete description** and **results** can be found in our **paper.**
  - **In-depth dimensional analysis** via matrix null space based on **Buckingham π theorem.**
  - A detailed description of **Buckingham π-Invariant Test-Time Projection** and an **algorithm.**
  - Extensive evaluation on **thermal conduction** and **linear elasticity** across CNN/Unet/FNO.
  - Additional PDE studies on **Navier-Stokes**, including **ideal and incomplete π-groups.**

Paper

Poster