NEURAL INFORMATION PROCESSING SYSTEMS

Mitigating Instability in High Residual Adaptive Sampling for PINNs via Langevin Dynamics

Minseok Jeong*, Giup Seo*, and Euiseok Hwang

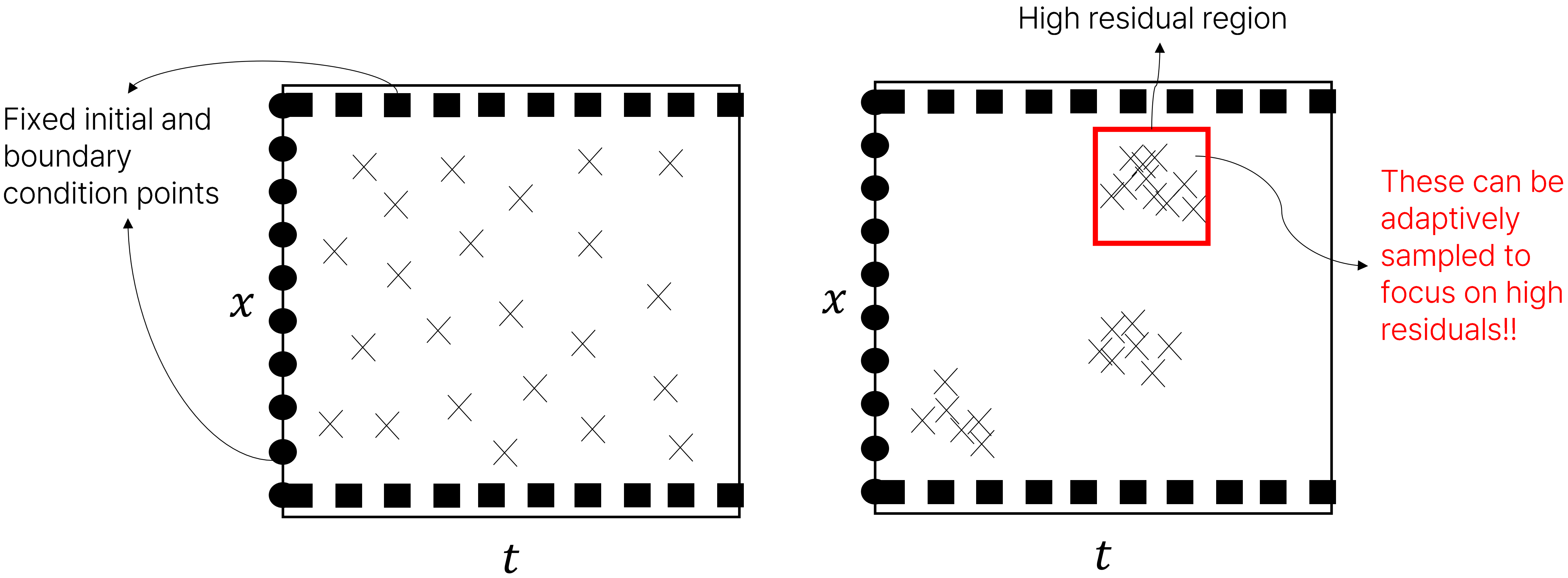
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology (GIST), South Korea

GIST Intelligent Information Systems Laboratory IISL

Motivation & Background

Focusing on more informative samples to enhance the performance

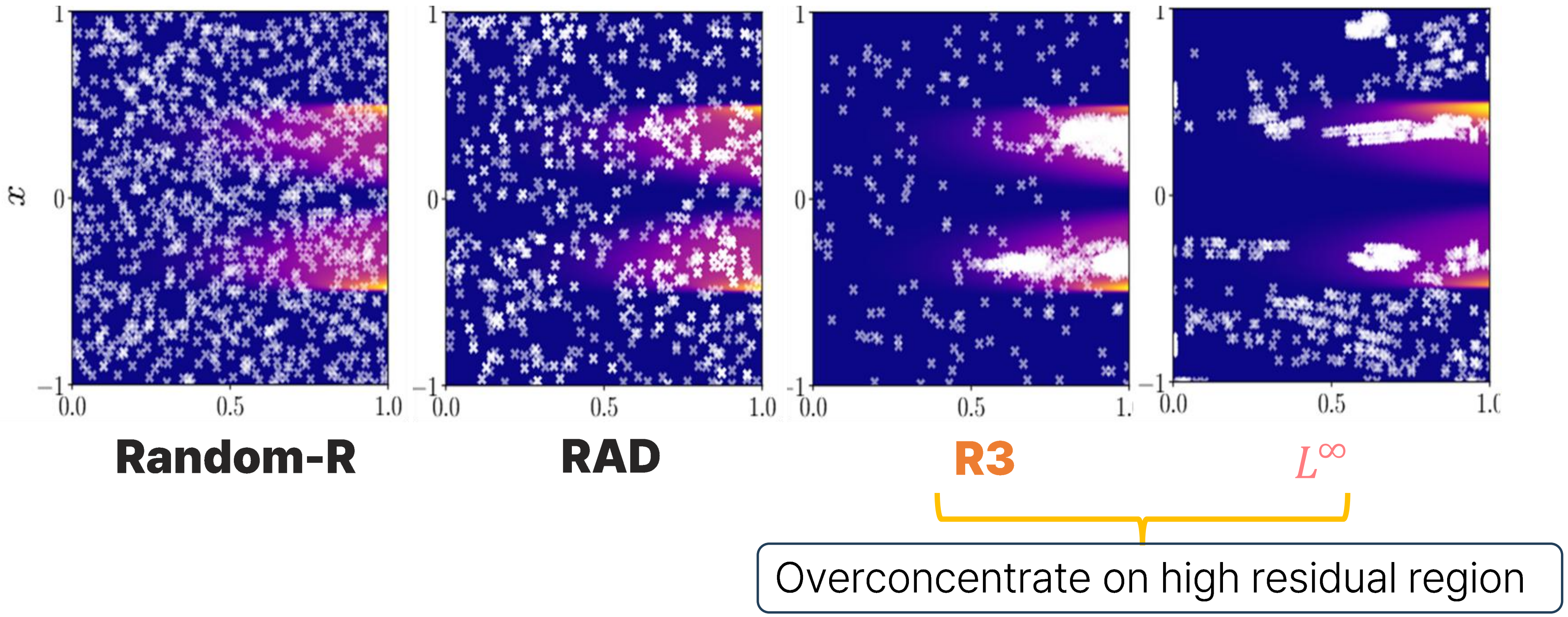
- PINN problem with residual $\mathcal{R}(\mathbf{x}) \triangleq \partial_t f(x,t) - \mathcal{N}[f](x,t)$ where $\mathbf{x} = (x,t)$
- High-residual points are more informative → Adaptively sample and emphasize high-residual regions



Motivation & Background

Adaptive PINN sampling often over-focus on high-residual regions

- Over-focusing algorithms(R3, L^∞) have shown superior performances on STIFF PDEs

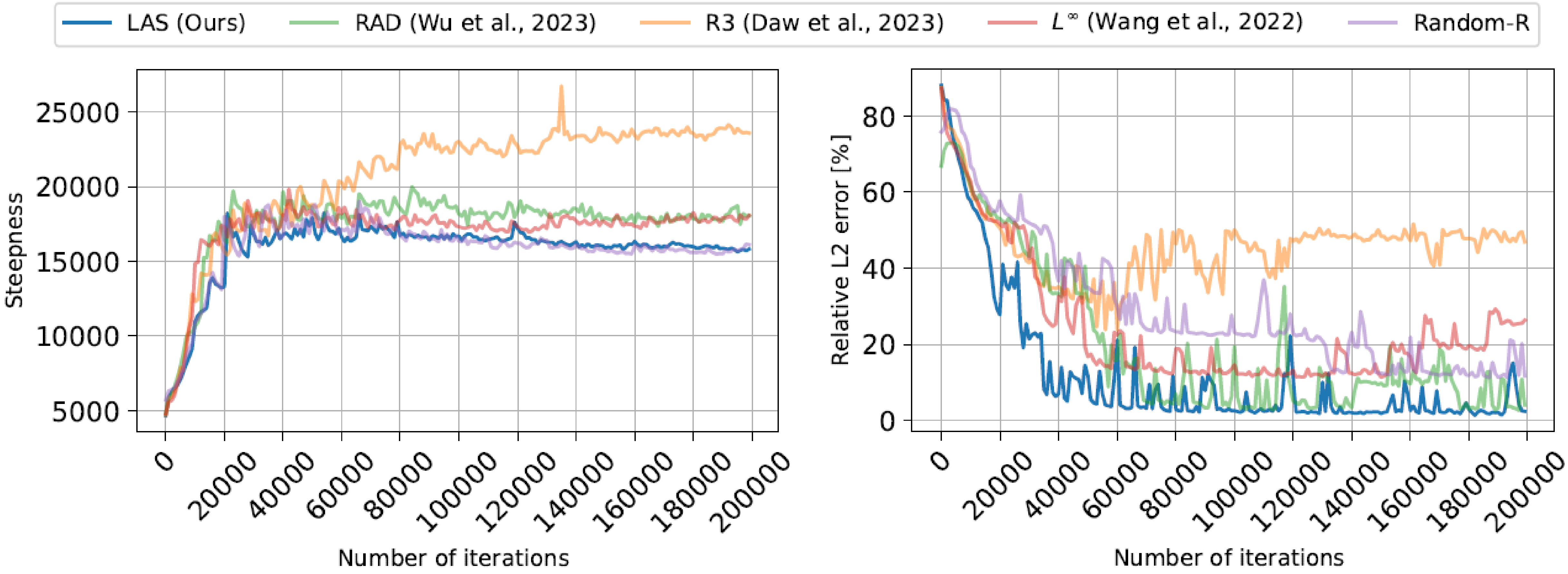


Q. Do we really need to sample only high-residual points?
A. No! theory explains the failure mode.

Analysis

Over-focusing on high residuals can destabilize training.

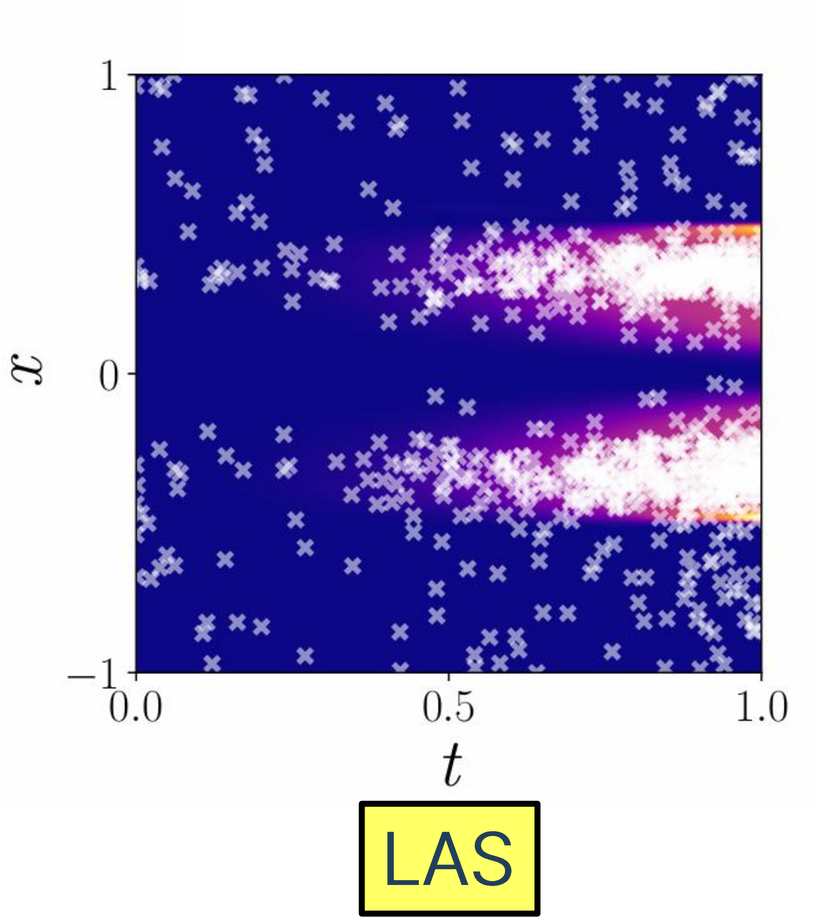
- This sharpens the loss landscape (steep curvature) [Propositions 3.1 and 3.2]
- Then the model needs very small learning rates → training collapses [Theorem 3.1]



Method

LAS: Langevin-based Adaptive Sampling

- We move collocation points using a Langevin update:
    - ➢ Sample high-residual regions in a balanced manner



Langevin Dynamics

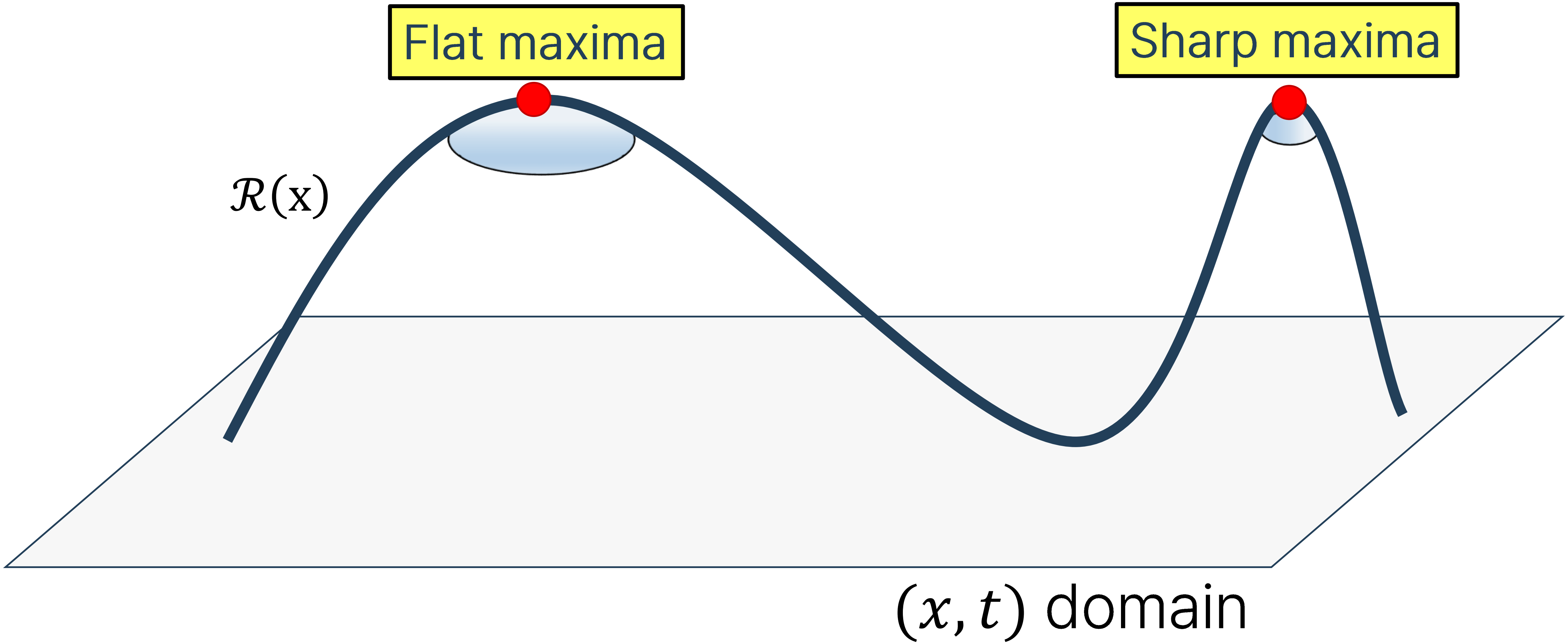
$$\mathrm{x}_{l+1} \leftarrow \mathrm{x}_l + \underbrace{\frac{\tau}{2}\nabla_{\mathrm{x}}|\mathcal{R}|^2}_{\text{Exploitation}} + \underbrace{\beta\sqrt{\tau}z_l}_{\text{Exploration}},$$

$$\lim_{l\to\infty} p_l(\mathrm{x}) \propto \exp\left(\frac{|\mathcal{R}(\mathrm{x})|^2}{\beta}\right)$$

Method

Enhanced stability due to the properties of Langevin dynamics:

- LAS naturally favors flat maxima [Theorem 4.2] → prevent sharp peaks



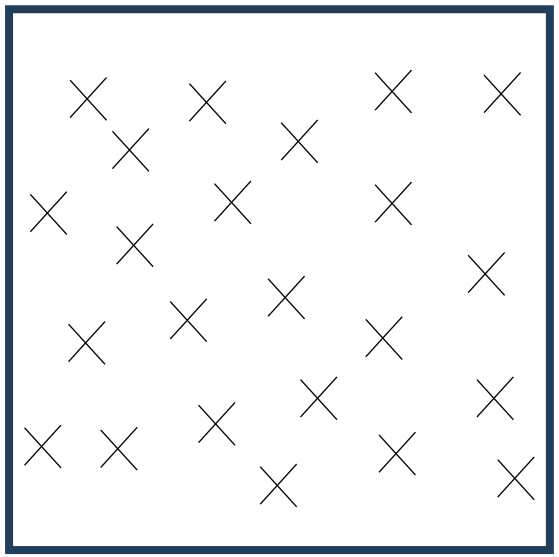
Method

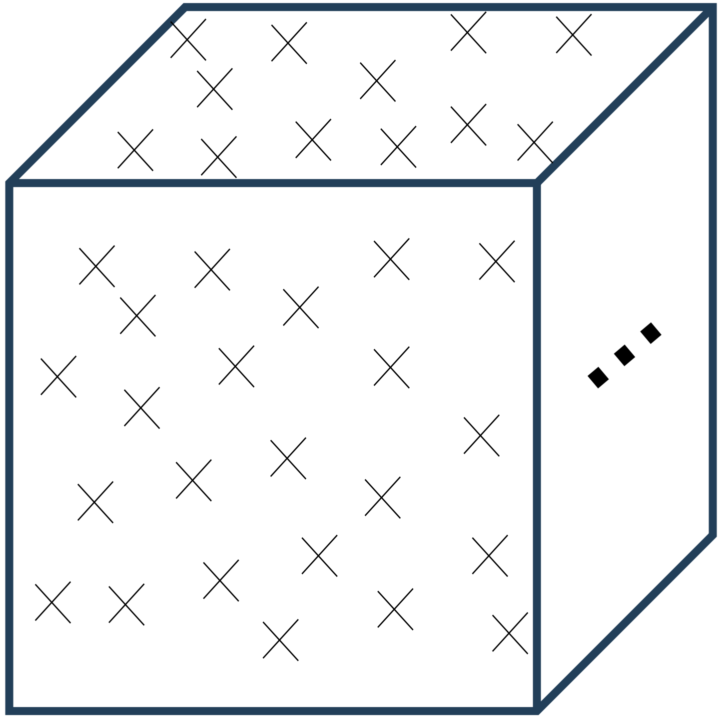
Enhanced stability due to the properties of Langevin dynamics:

- Does not rely solely on Monte-Carlo estimation for sampling
  - Monte-Carlo breaks down in high-dimensional PDEs (curse of dimensionality)

Monte-Carlo estimation

$$\mathbb{E}|\mathcal{R}^2(\mathrm{x})| \approx \frac{1}{n}\sum_{i=1}^{n}|\mathcal{R}(\mathrm{x}_i)|^2$$

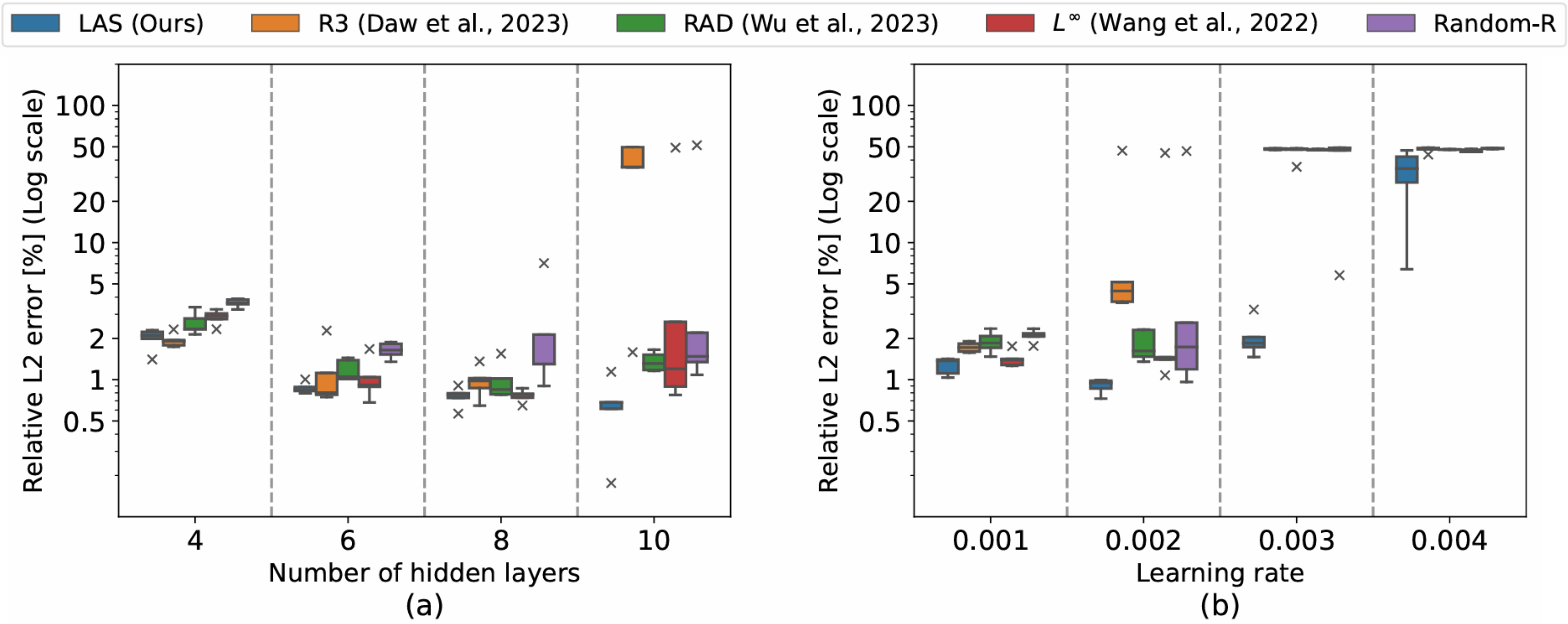




Experimental result

Robustness across various experimental setup

- Scales with **NN depth & learning rate**



Experimental result

Robustness across various experimental setup

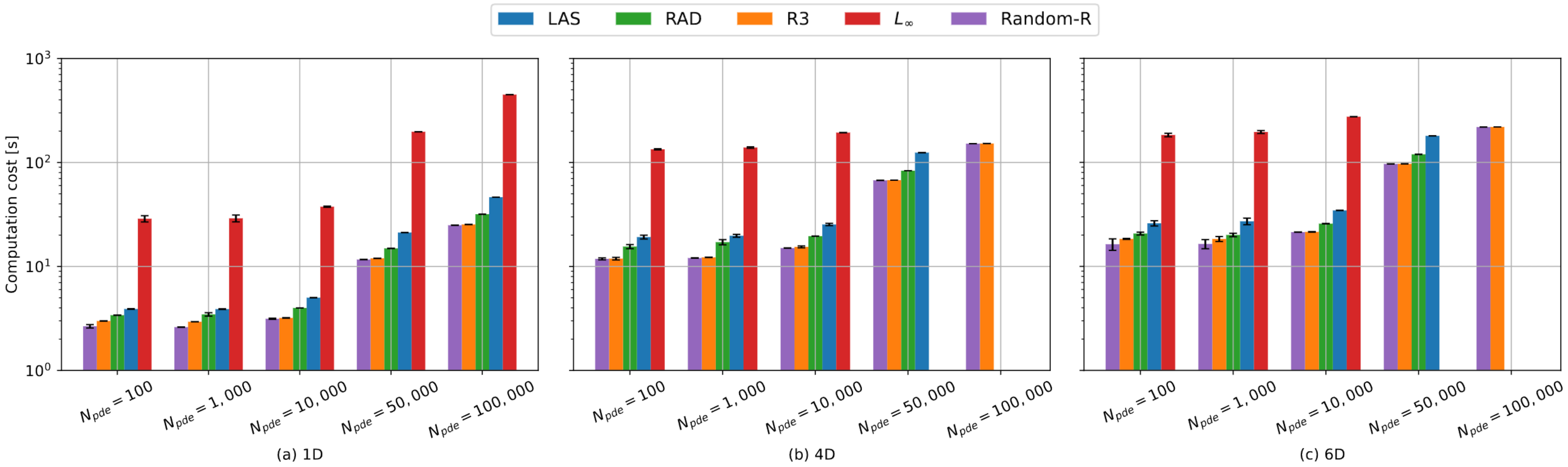
- Stable even in **high-dimensional PDEs** (up to 8D)

Sampling methods		LAS (ours)	Random-R	RAD	R3	L^∞
Number of layers		8	8	8	8	8
1D Allen-Cahn	100,1,1	0.77 ± 0.20	2.54 ± 2.30	0.99 ± 0.29	0.97 ± 0.23	**0.76 ± 0.07**
	200,1,1	0.89 ± 0.03	1.69 ± 0.38	1.15 ± 0.18	10.62 ± 19.56	**0.77 ± 0.05**
	10,1,1	**1.02 ± 0.20**	1.07 ± 0.48	1.25 ± 0.27	75.81 ± 32.72	86.77 ± 7.53
	10,1,5	**1.55 ± 0.80**	2.06 ± 0.40	2.79 ± 0.24	65.51 ± 18.70	98.77 ± 0.79
	1,1,1	13.20 ± 22.52	**3.21 ± 1.19**	14.65 ± 18.34	64.32 ± 7.30	90.81 ± 12.27
4D DF-heat	100,1,1	**1.72 ± 0.23**	9.73 ± 0.44	6.64 ± 2.72	9.53 ± 5.27	2.92 ± 1.95
	200,1,1	**1.82 ± 0.42**	8.87 ± 1.67	5.72 ± 0.42	6.50 ± 1.19	4.06 ± 1.87
	10,1,1	**1.79 ± 0.20**	7.98 ± 37.59	24.98 ± 37.59	16.08 ± 3.25	2.46 ± 0.67
	10,1,5	**2.85 ± 0.34**	25.18 ± 37.84	72.53 ± 32.73	15.57 ± 9.22	42.71 ± 46.79
	1,1,1	**1.86 ± 0.20**	7.73 ± 1.85	12.31 ± 1.24	13.56 ± 5.24	10.95 ± 14.41
6D DF-heat	100,1,1	**3.49 ± 0.20**	6.14 ± 0.42	13.84 ± 2.08	29.91 ± 35.23	4.46 ± 0.42
	200,1,1	**4.18 ± 0.58**	5.21 ± 0.82	10.63 ± 0.93	30.51 ± 34.90	4.39 ± 0.61
	10,1,1	**4.79 ± 0.33**	5.53 ± 0.81	31.81 ± 34.30	49.47 ± 41.28	26.37 ± 37.05
	10,1,5	**7.33 ± 1.37**	70.30 ± 36.59	100.00 ± 0.00	69.37 ± 37.51	90.21 ± 19.57
	1,1,1	**5.20 ± 0.92**	53.18 ± 46.82	100.00 ± 0.00	61.37 ± 39.22	100.00 ± 0.00
8D DF-heat	100,1,1	**7.72 ± 0.59**	52.41 ± 39.33	84.71 ± 30.56	100.00 ± 0.00	100.00 ± 0.00
	200,1,1	**6.92 ± 0.31**	17.63 ± 1.78	83.59 ± 32.80	100.00 ± 0.00	13.21 ± 4.46
	10,1,1	**7.97 ± 0.30**	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
	10,1,5	**11.17 ± 0.65**	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
	1,1,1	**9.08 ± 0.50**	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00

Experimental result

Computational efficiency & robustness to Langevin parameters

- Residual landscape changes slowly → no need for heavy Langevin iteration
  - ➢ Just one step per iteration works
  - ➢ Efficient and scales well as PDE dimension grows



Thank You!