ICLR
The International Conference on Learning Representations (ICLR) 2024

The Twelfth International Conference on Learning Representations (ICLR 2024)

Adversarial Adaptive Sampling: Unify PINN and Optimal Transport for the Approximation of PDEs

Kejun Tang*, Jiayu Zhai*, Xiaoliang Wan, Chao Yang *Co-first Author

Problem Setting

The PDE problem considered here is: find $u \in \mathscr{F} : \Omega \mapsto \mathbb{R}$ where $\mathscr{F}$ is a proper function space defined on a computational domain $\Omega \in \mathbb{R}^D$, such that

$$\begin{aligned} \mathcal{L}u(x) &= s(x), \quad \forall x \in \Omega \\ \mathfrak{b}u(x) &= g(x), \quad \forall x \in \partial\Omega, \end{aligned} \tag{1}$$

$\mathcal{L}$: partial differential operator (e.g., the Laplace operator Δ)
$\mathfrak{b}$: boundary operator (e.g., the Dirichlet boundary)
The idea of PINN [2]: NN $u_\theta \to u$. The parameters θ are determined by minimizing the following loss functional

$$J(u_\theta) = J_r(u_\theta) + \gamma J_b(u_\theta) \quad \text{with}$$
$$J_r(u_\theta) = \int_\Omega |r(x;\theta)|^2 dx \ \text{ and } \ J_b(u_\theta) = \int_{\partial\Omega} |b(x;\theta)|^2 dx, \tag{2}$$

where $r(x;\theta) = \mathcal{L}u_\theta(x) - s(x)$, and $b(x;\theta) = \mathfrak{b}u_\theta(x) - g(x)$ are the residuals that measure how well u_θ satisfies the partial differential equations and the boundary conditions, respectively, and $\gamma > 0$ is a penalty parameter to leverage the convergence of the two parts.

Statistical Errors in Neural Network Approximation

Let $S_\Omega = \{x_\Omega^{(i)}\}_{i=1}^{N_r}$ and $S_{\partial\Omega} = \{x_{\partial\Omega}^{(i)}\}_{i=1}^{N_b}$ be two sets of uniformly distributed collocation points on Ω and $\partial\Omega$ respectively. We then minimize the following empirical loss in practice

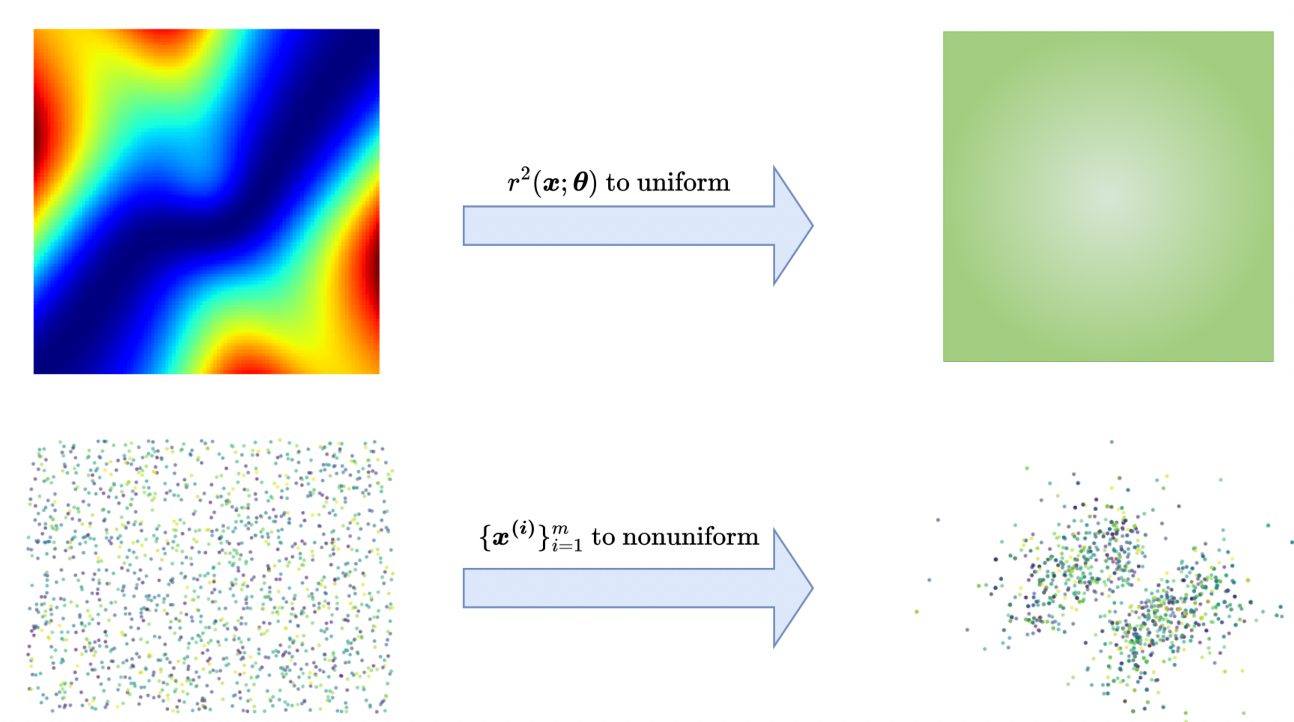
$$J_N(u_\theta) = \frac{1}{N_r}\sum_{i=1}^{N_r} r^2(x_\Omega^{(i)};\theta) + \gamma\frac{1}{N_b}\sum_{i=1}^{N_b} b^2(x_{\partial\Omega}^{(i)};\theta), \tag{3}$$

Monte Carlo (MC) approximation of $J(u_\theta)$ subject to a statistical error of $O(N^{-1/2})$ with N being the sample size. Let $u_{\theta_N^*}$ be the minimizer of the empirical loss $J_N(u_\theta)$ and u_{θ^*} be the minimizer of the original loss functional $J(u_\theta)$. Without taking into the optimization error,

$$\mathbb{E}\left(\left\|u_{\theta_N^*} - u\right\|_\Omega\right) \le \mathbb{E}\left(\left\|u_{\theta_N^*} - u_{\theta^*}\right\|_\Omega\right) + \|u_{\theta^*} - u\|_\Omega, \tag{4}$$

Main Idea

Smaller variance of the Monte Carlo integration for r^2
→ Improved accuracy of the MC approximation(with fixed sample size)
→ More accurate solution of PDEs



Methodology

p_α: sampler (a flow-based deep generative model),
u_θ: approximator (a fully connected neural network).
A minmax formulation

$$\min_\theta \max_{p_\alpha \in V} \mathcal{J}(u_\theta, p_\alpha) = \int_\Omega r^2(x;\theta)p_\alpha(x)dx + \gamma J_b(u_\theta), \tag{5}$$

V: a function space that defines a proper constraint on $p_\alpha(x)$

$$V = \{p(x) \| \|p\|_{\mathrm{Lip}} \le 1, p(x) \ge 0, \int_\Omega p(x)dx = 1\},$$

A practical implementation

$$\min_\theta \max_{\substack{p_\alpha > 0, \\ \int_\Omega p_\alpha(x)dx = 1}} \mathcal{J}(u_\theta, p_\alpha) = \int_\Omega r^2(x;\theta)p_\alpha(x)dx - \beta\int_\Omega |\nabla_x p_\alpha(x)|^2 dx, \tag{6}$$

β: a hyperparameter to be tuned

Main Theorem

Let μ be the Lebesgue measure on $\mathbb{R}^D$, which represents the uniform probability distribution on Ω. Then the optimal value of the min-max problem (5) is 0. Moreover, there is a sequence $\{u_n\}_{n=1}^\infty$ of functions with $r(u_n) \neq 0$ for all n, such that it is an optimization sequence of (5),

$$\lim_{n\to\infty} \mathcal{J}(u_n, p_n) = 0,$$

for some sequence of functions $\{p_n\}_{n=1}^\infty$ satisfying the constraints in (5). Meanwhile, this optimization sequence has the following properties:

1. The residual sequence $\{r(u_n)\}_{n=1}^\infty$ of $\{u_n\}_{n=1}^\infty$ converges to 0 in $L^2(d\mu)$.
2. The renormalized squared residual distributions
$$d\nu_n \triangleq \frac{r^2(u_n)}{\int_\Omega r^2(u_n(x))\,dx} d\mu(x)$$
converge to the uniform distribution μ in the Wasserstein distance d_{W^M}.

Results

Method	one peak	two peaks	high dimensional problem
PINN	9.74e-04	3.22e-02	1.01
RAR [1]	-	-	9.83e-01
DAS-G [3]	3.75e-04	1.51e-03	9.55e-03
DAS-R [3]	1.93e-04	6.21e-03	1.26e-02
AAS (this work)	**2.97e-05**	**1.09e-04**	**1.31e-03**

References

[1] Lu Lu et al. "DeepXDE: A deep learning library for solving differential equations". In: SIAM Review 63.1 (2021), pp. 208–228.
[2] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: Journal of Computational Physics 378 (2019), pp. 686–707.
[3] Kejun Tang, Xiaoliang Wan, and Chao Yang. "DAS-PINNs: A deep adaptive sampling method for solving high-dimensional partial differential equations". In: Journal of Computational Physics 476 (2023), p. 111868.