2025 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP 2025)
April 06 - 11, 2025 Hyderabad, India

# DAREK

## Distance Aware Error for Kolmogorov Networks

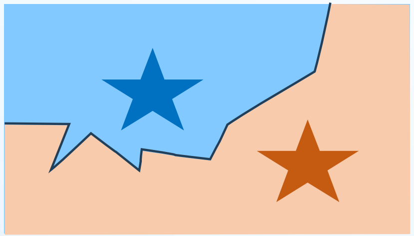
Masoud Ataei

Co-authors: Mohammad Javad Khojasteh and Vikas Dhiman

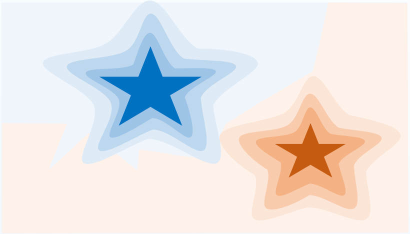
# Motivation

Safe critical applications

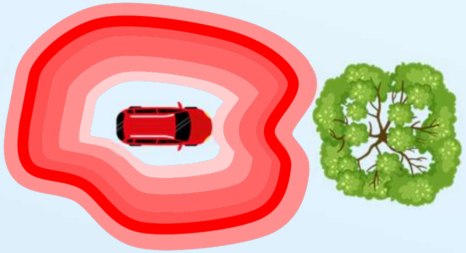
Over confidence model



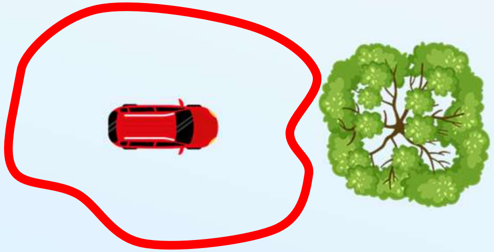
Uncertainty bounded



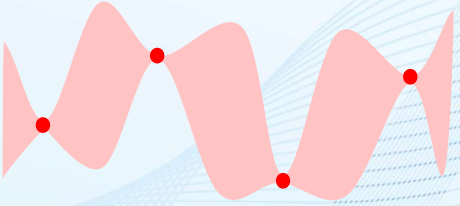
Probabilistic safety



Worst-case bounded



DAREK

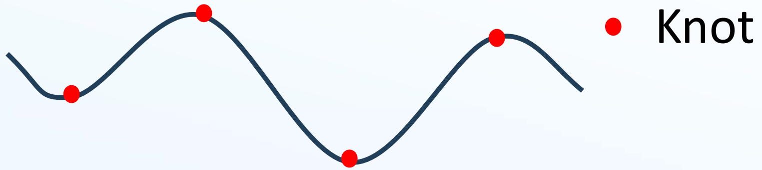


# Introduction

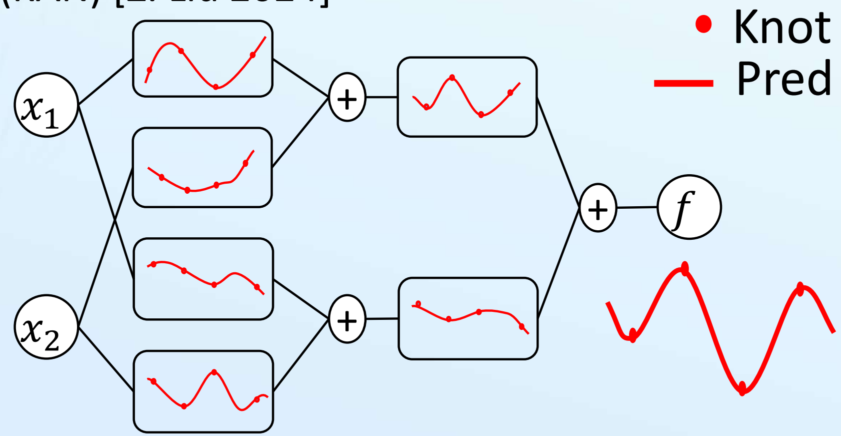
Kolmogorov Arnold Representation Theorem (KAT) [A. N. Kolmogorov 1957]

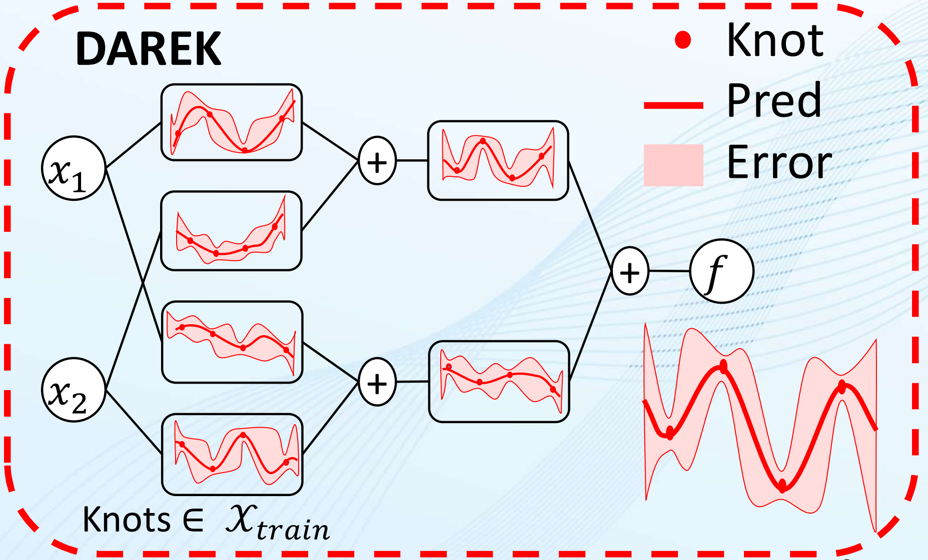
$$f(x_1, x_2, \ldots, x_n) = \sum_{q=1}^{2n+1} \Phi_q\left(\sum_{p=1}^{n} \psi_{p,q}(x_p)\right)$$

Spline

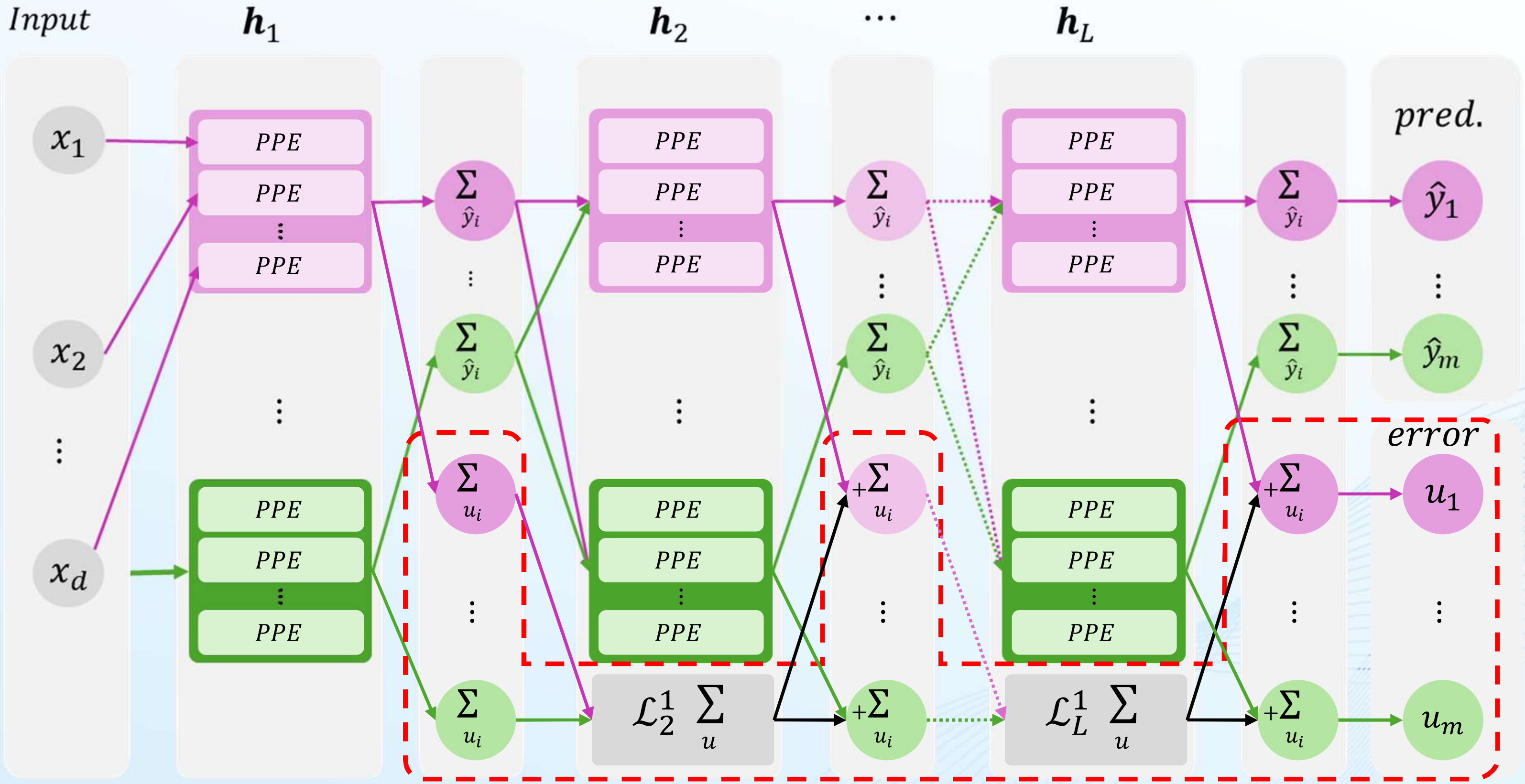



Kolmogorov Arnold Networks (KAN) [Z. Liu 2024]



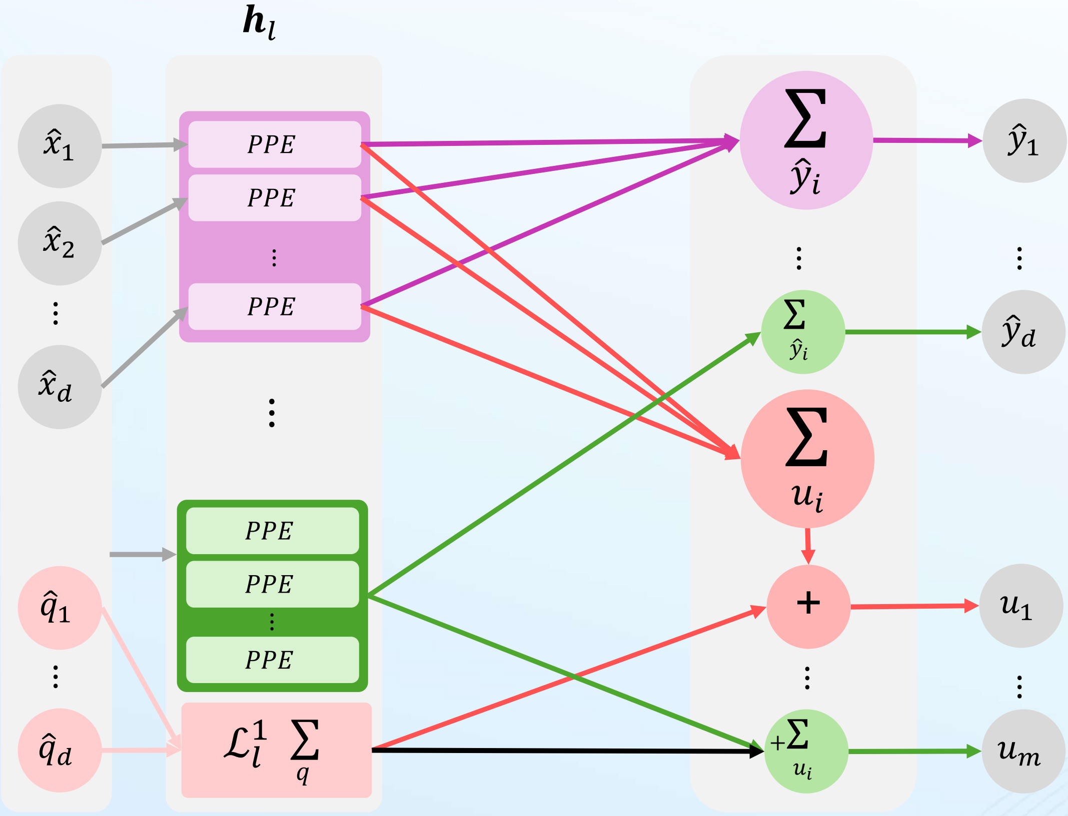



# DAREK




PPE ≡ Piece-wise polynomial error

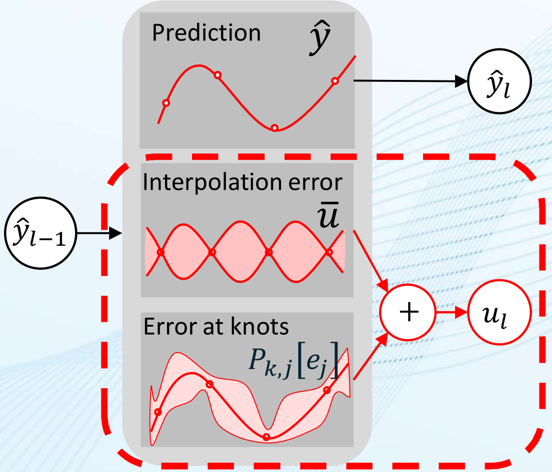
# DAREK




$$|h_l(x) - \hat{h}_l(x)| \leq u_{[l]}(\mathbf{1}_n^\top \hat{\boldsymbol{x}}) + \mathcal{L}_l^1 \mathbf{1}_d^\top \boldsymbol{q}$$

[Thm. 2]

Piecewise Polynomial error $(PPE)$



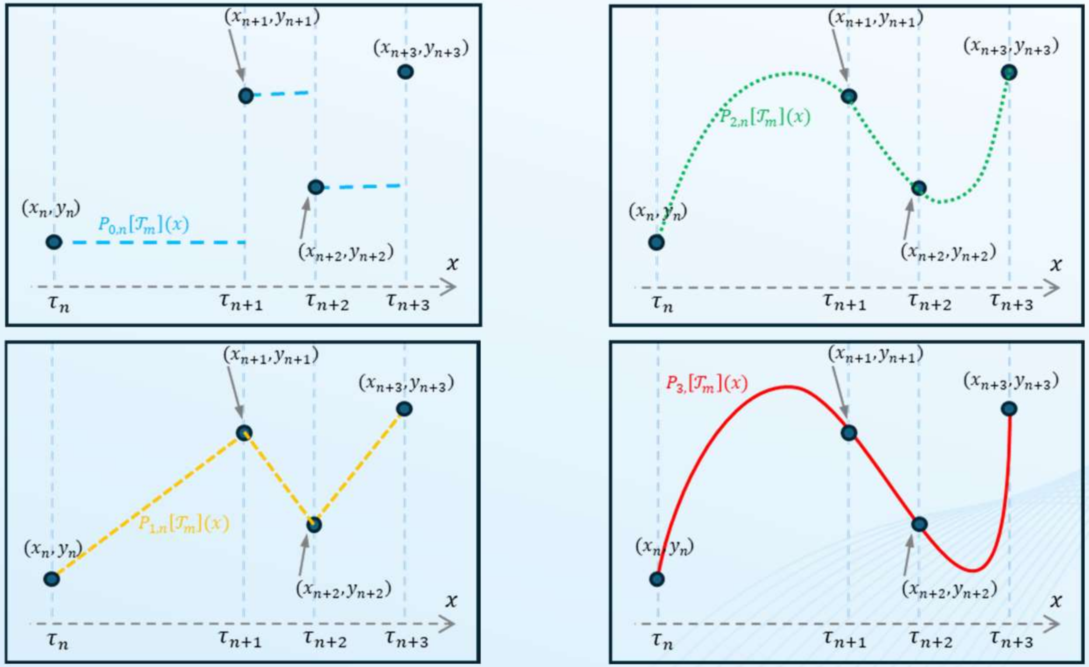

PPE ≡ Piece-wise polynomial error

# Piecewise Newton's Polynomial

$$f(x) \approx \mathcal{P}_{k,n}[f(\tau_{1:m})](x) \coloneqq [\tau_n]f + \sum_{i=1}^{k} [\tau_{n:n+i}]f \prod_{j=n}^{n+i-1} (x - \tau_j)$$

[C. De Boor 1978]

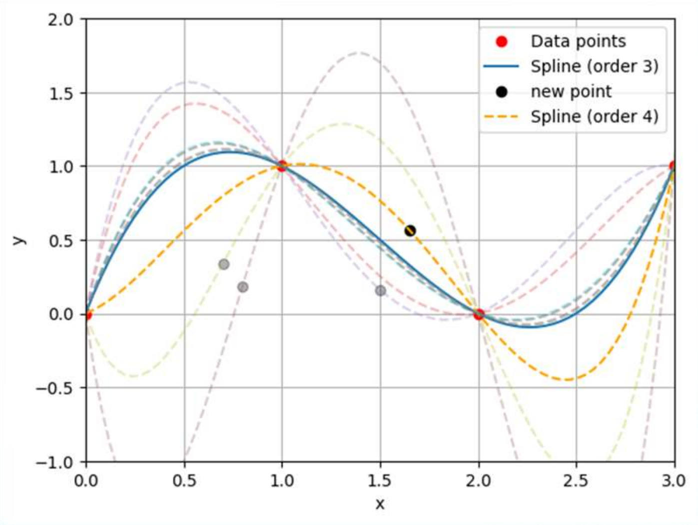
$[.]f \equiv$ Divided differences



# Newton's Polynomial + Reminder

$$f(x) \approx \mathcal{P}_{k,n}[f(\tau_{1:m})](x) \coloneqq [\tau_n]f + \sum_{i=1}^{k} [\tau_{n:n+i}]f \prod_{j=n}^{n+i-1} (x - \tau_j)$$ [C. De Boor 1978]

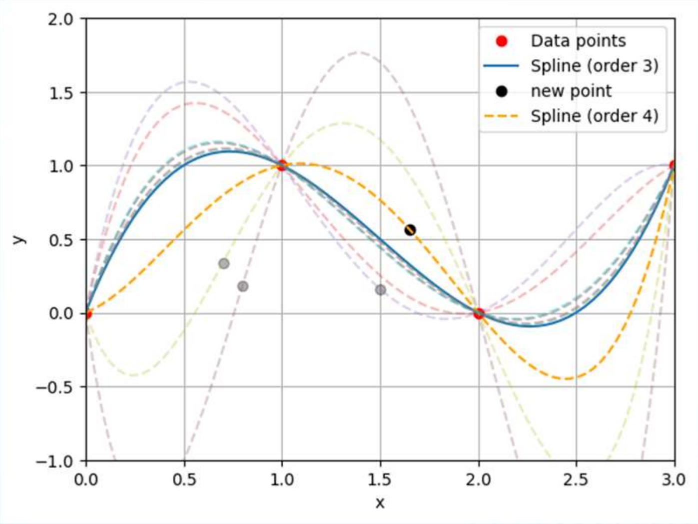
$$f(x) = \mathcal{P}_{k,n}[f(\tau_{1:m})](x) + (x-\tau_n) \dots (x - \tau_{n+k}) [\tau_n, \dots, \tau_{n+k}, x]f$$ [C. De Boor 1978]



# Interpolation error $\bar{u}$

$$f(x) = \mathcal{P}_{k,n}[f(\tau_{1:m})](x) + (x - \tau_n) \dots (x - \tau_{n+k}) [\tau_n, \dots, \tau_{n+k}, x] f$$ [C. De Boor 1978]

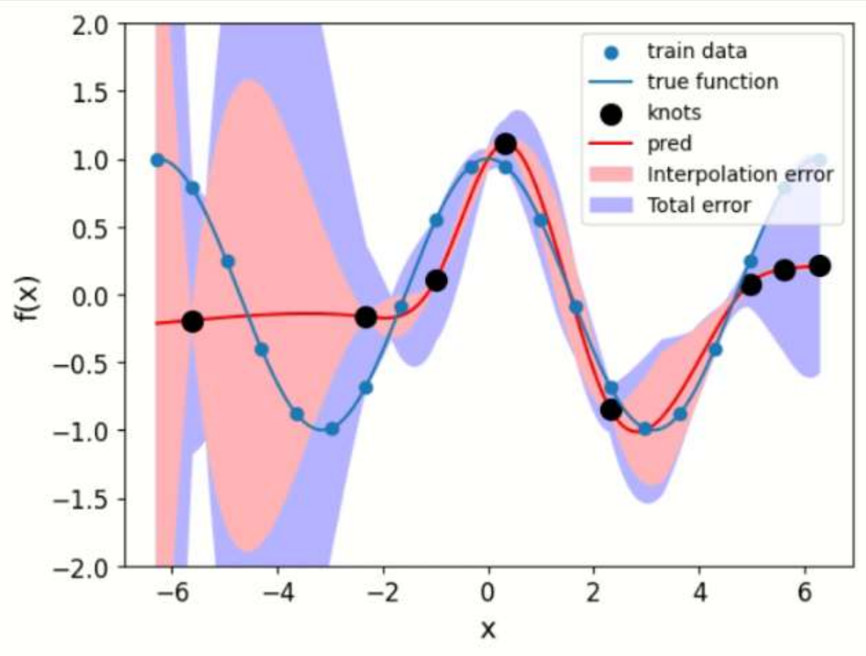
$$\left| f(x) - \mathcal{P}_{k,n}[f(\tau_{1:m})](x) \right| \le \frac{\mathcal{L}_f^{k+1}}{(k+1)!} \left| \prod_{j=n}^{n+i-1} (x - \tau_j) \right| =: \bar{u}_f(x; \tau_{1:m})$$ [Thm. 1]



# Non-zero error at knot

$$e_n^f(\tau_i) \coloneqq f(\tau_i) - \hat{f}_{[n]}(\tau_i)$$

$$\left|f(x) - \hat{f}_{[n]}(x)\right| \le \bar{u}_f(x) + \left|\mathcal{P}_{k,n}\left[e_n^f(\tau_{1:m})\right](x)\right| =: u_f(x;\tau_{1:m})$$ [Lemma 1]



# Lipschitz sharing and error division at knots

- Lipschitz division: $\mathcal{L}_f \Rightarrow \mathcal{L}_h$

$$\mathcal{L}_f = \prod_{l=1}^{L} N_l \, \mathcal{L}_h = (\mathcal{L}_h)^L \prod_{l=1}^{L} N_l \qquad \mathcal{L}_h = \sqrt[L]{\frac{\mathcal{L}_h}{\prod_{l=1}^{L} N_l}}$$

[Inspired by J. Liu 2020]

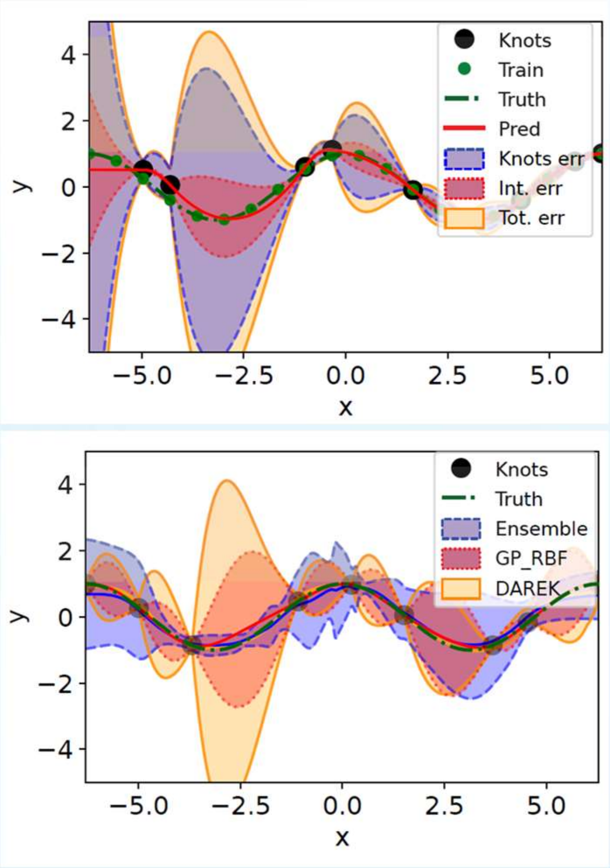
- Error at knot sharing: $e^f \Rightarrow e^h$

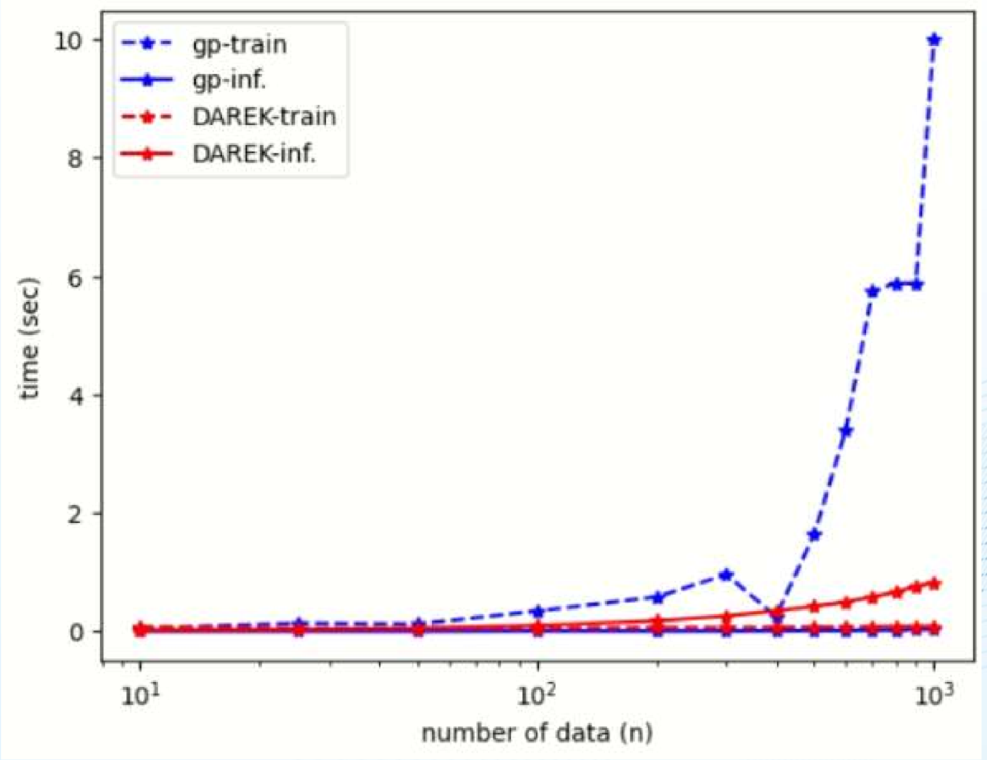
$$e^h(\mathbf{1}^\top \hat{\boldsymbol{g}}(\tau_i)) = e^{g_1}(\tau_i) = \cdots = e^{g_n}(\tau_i) \geq \frac{e^f(\tau_i)}{1 + n\mathcal{L}_h^1}$$

[Assumption 2]

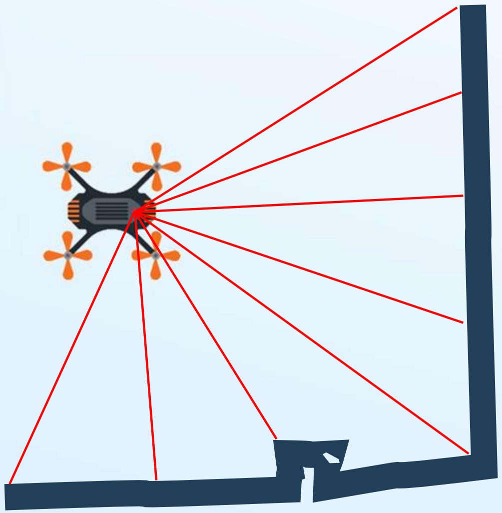
$$e^f(\tau_i) = f(\tau_i) - \hat{f}(\tau_i)$$

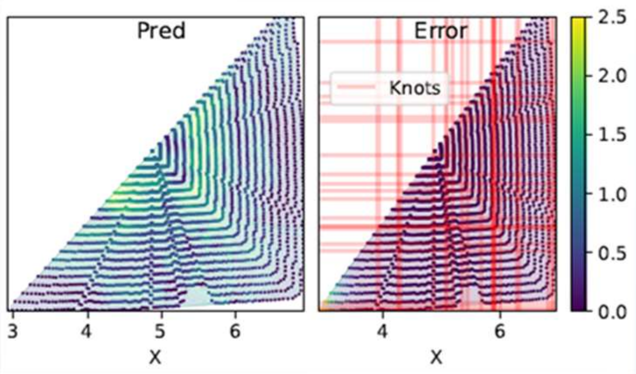
# Experiments

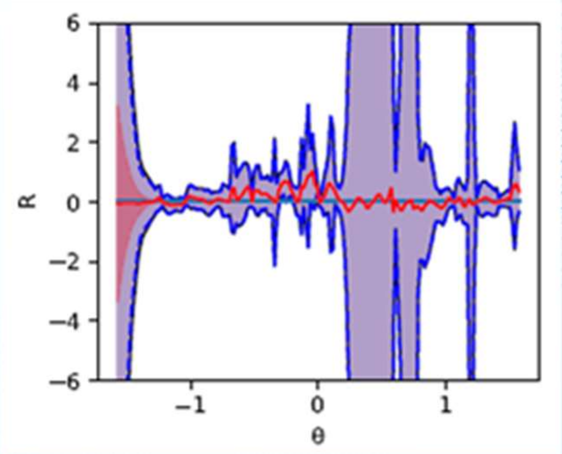


# Experiments





# Conclusion

- **DAREK**, a novel framework for **error estimation** in spline based networks
- Provides **structured, interpretable, and computationally efficient worst-case error bounds**
- Uses **piecewise polynomial error estimation**, ensuring **tight, distance-aware error bounds**

# Future Work

- **Refine Lipschitz division and error propagation** to improve bound tightness
- **Extend DAREK to higher-dimensional problems** for complex applications
- **Integrate DAREK into real-time safe control systems** for autonomous decision-making

# References

1. Z. Liu, Y. Wang et al., “Kan: Kolmogorov-arnold networks,” arXiv preprint, arXiv:2404.19756, 2024.
2. B. Lakshminarayanan et al., “Simple and scalable predictive uncertainty estimation using deep ensembles,” NeuRIPS, vol. 30, 2017.
3. C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning. The MIT Press, 2006, vol. 1, no. 1.
4. L. Jaulin, et al., and ´E. Walter, Interval analysis. Springer, 2001.
5. J. Liu et al., “Simple and principled uncertainty estimation with deterministic deep learning via distance awareness,” Advances in neural information processing systems, vol. 33, pp. 7498–7512, 2020.
6. C. De Boor and C. De Boor, A practical guide to splines. springer New York, 1978, vol. 27.
7. G. Wahba, Spline Functions. John Wiley & Sons, Ltd, 2006. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/0471667196.ess3095.pub2
8. G. M. Phillips, Interpolation and approximation by polynomials. Springer Science & Business Media, 2003, vol. 14.
9. T. M. Apostol, Calculus, Volume 1. John Wiley & Sons, 1967.
10. Z. Shi et al., “Efficiently computing local lipschitz constants of neural networks via bound propagation,” Advances in Neural Information Processing Systems, vol. 35, pp. 2350–2364, 2022.
11. A. Howard, “The robotics data set repository (radish),” http://radish. sourceforge. net/, 2003.
12. A. N. Kolmogorov, “On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition,” in Dok-lady Akademii Nauk, vol. 114, no. 5. Russian Academyof Sciences, 1957, pp. 953–956.

# Thank You