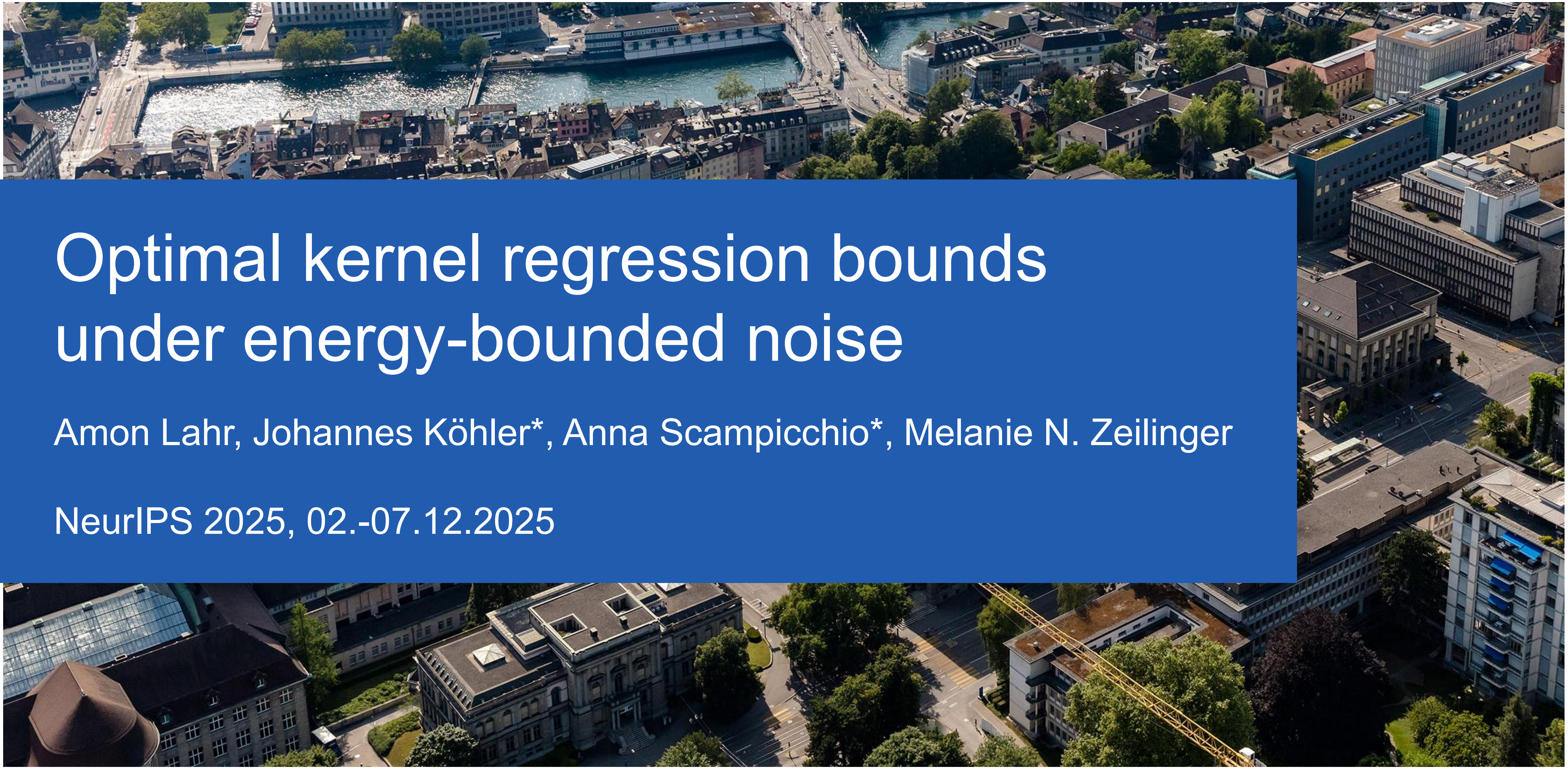
ETH zürich ELO-X

Institute for Dynamic
Systems and Control



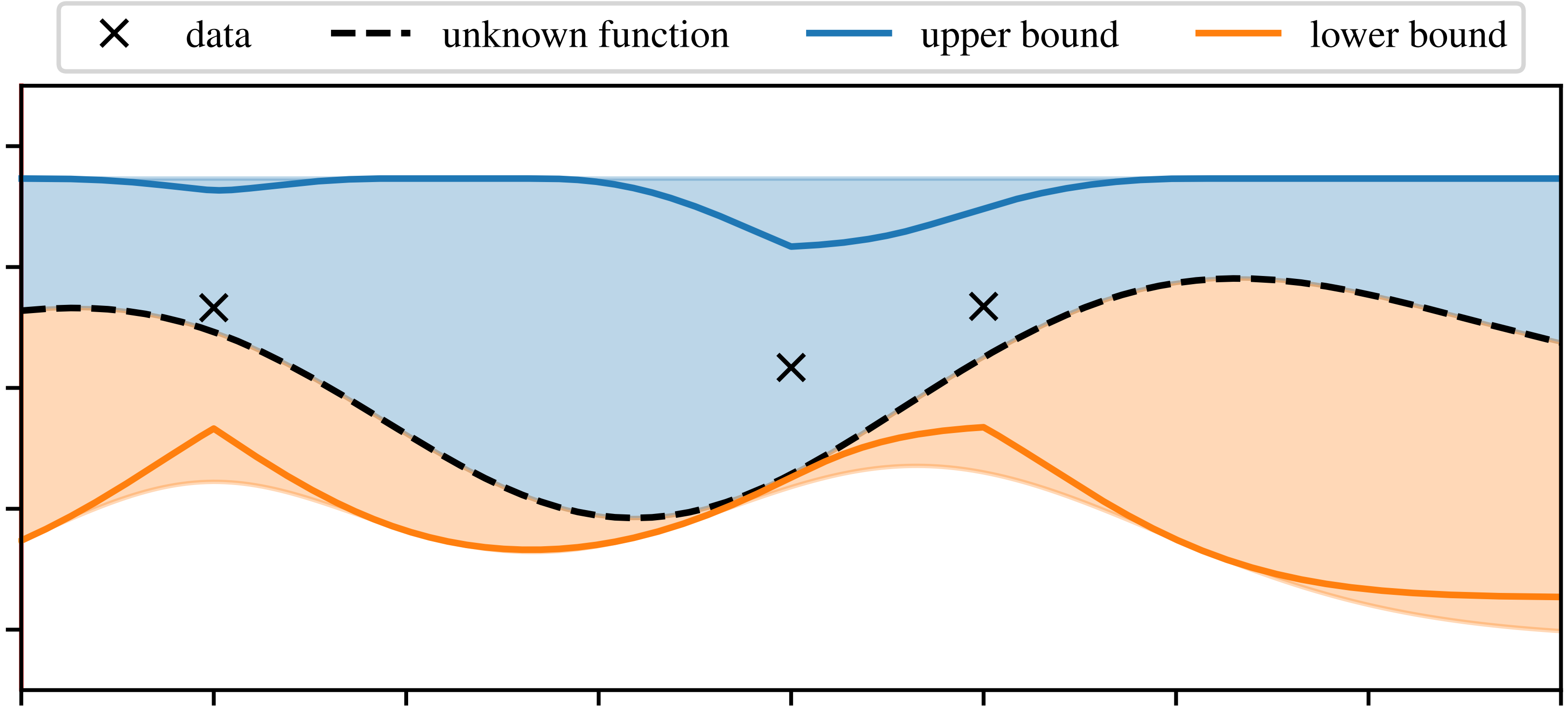
Optimal kernel regression bounds under energy-bounded noise

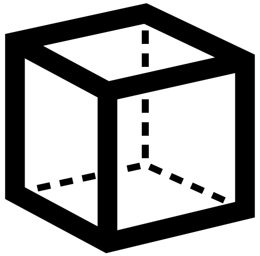
Amon Lahr, Johannes Köhler*, Anna Scampicchio*, Melanie N. Zeilinger

NeurIPS 2025, 02.-07.12.2025

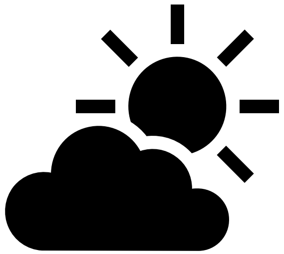
Funded by the
European Union

Learning under uncertainty





Safe exploration & experiment design



Reliable predictions

Autonomous decision-making & control

Data-generating model

$$y_i = f^{\mathrm{tr}}(x_i) + \epsilon_i$$

Uncertainty quantification

$$\underline{f}(x_*) \leq f^{\mathrm{tr}}(x_*) \leq \overline{f}(x_*)$$

Kernel regression bounds in Reproducing Kernel Hilbert Spaces (RKHS)

Regularity assumptions

1. Latent function $f^{\mathrm{tr}} \in \mathcal{H}_{k^f}$ with RKHS norm

$$\|f^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^f}} < \Gamma_f^2$$

2. Noise $\epsilon_i = w^{\mathrm{tr}}(x_i)$ given as realizations of $w^{\mathrm{tr}} \in \mathcal{H}_{k^w}$ with RKHS norm

$$\|w^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^w}} < \Gamma_w^2$$

Optimal upper bound

$$\overline{f}(x_*) = \sup_{\substack{f \in \mathcal{H}_{k^f}, \\ w \in \mathcal{H}_{k^w}}} f(x_*)$$

$$\text{s.t.} \quad f(x_i) + w(x_i) = y_i$$

$$\|f\|^2_{\mathcal{H}_{k^f}} \leq \Gamma_f^2$$

$$\|w\|^2_{\mathcal{H}_{k^w}} \leq \Gamma_w^2$$

Worst-case value of latent function at test point

Assumptions transcribed into optimization problem

Data-generating model

$$y_i = f^{\mathrm{tr}}(x_i) + \epsilon_i$$

Uncertainty quantification

$$\underline{f}(x_*) \leq f^{\mathrm{tr}}(x_*) \leq \overline{f}(x_*)$$

Kernel regression bounds in Reproducing Kernel Hilbert Spaces (RKHS)

Regularity assumptions

1. Latent function $f^{\mathrm{tr}} \in \mathcal{H}_{k^f}$ with RKHS norm

$$\|f^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^f}} < \Gamma_f^2$$

2. Noise $\epsilon_i = w^{\mathrm{tr}}(x_i)$ given as realizations of $w^{\mathrm{tr}} \in \mathcal{H}_{k^w}$ with RKHS norm

$$\|w^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^w}} < \Gamma_w^2$$

Relaxed upper bound (lower bound analogously)

Gaussian process (GP) mean

GP covariance

$$\overline{f}^{\sigma}(x_*) = f_\sigma^\mu(x_*) + \beta_\sigma \sqrt{\Sigma_\sigma(x_*)}$$

Remaining RKHS norm in „data-RKHS“ with minimum-norm data interpolant g_σ^μ:

$$\beta_\sigma^2 = \Gamma_f^2 + \frac{\Gamma_w^2}{\sigma^2} - \|g_\sigma^\mu\|^2_{\mathcal{H}_{k^f + \sigma^2 k^w}}$$

Data-generating model

$$y_i = f^{\mathrm{tr}}(x_i) + \epsilon_i$$

Uncertainty quantification

Valid (conservative) bound for all $\sigma \in (0, \infty)$!

$$\underline{f}^{\sigma}(x_*) \leq \underline{f}(x_*) \leq f^{\mathrm{tr}}(x_*) \leq \overline{f}(x_*) \leq \overline{f}^{\sigma}(x_*)$$

Kernel regression bounds in Reproducing Kernel Hilbert Spaces (RKHS)

Regularity assumptions

1. Latent function $f^{\mathrm{tr}} \in \mathcal{H}_{k^f}$ with RKHS norm
$$\|f^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^f}} < \Gamma_f^2$$
2. Noise $\epsilon_i = w^{\mathrm{tr}}(x_i)$ given as realizations of $w^{\mathrm{tr}} \in \mathcal{H}_{k^w}$ with RKHS norm
$$\|w^{\mathrm{tr}}\|^2_{\mathcal{H}_{k^w}} < \Gamma_w^2$$

Relaxed upper bound (lower bound analogously)

Gaussian process (GP) mean

GP covariance

$$\overline{f}^{\sigma}(x_*) = f^{\mu}_{\sigma}(x_*) + \beta_\sigma \sqrt{\Sigma_\sigma(x_*)}$$

Optimal upper bound

„Tightest relaxed bound“

$$\overline{f}(x_*) = \inf_{\sigma \in (0,\infty)} \overline{f}^{\sigma}(x_*)$$

***Scalar, unconstrained* problem with safe iterates**

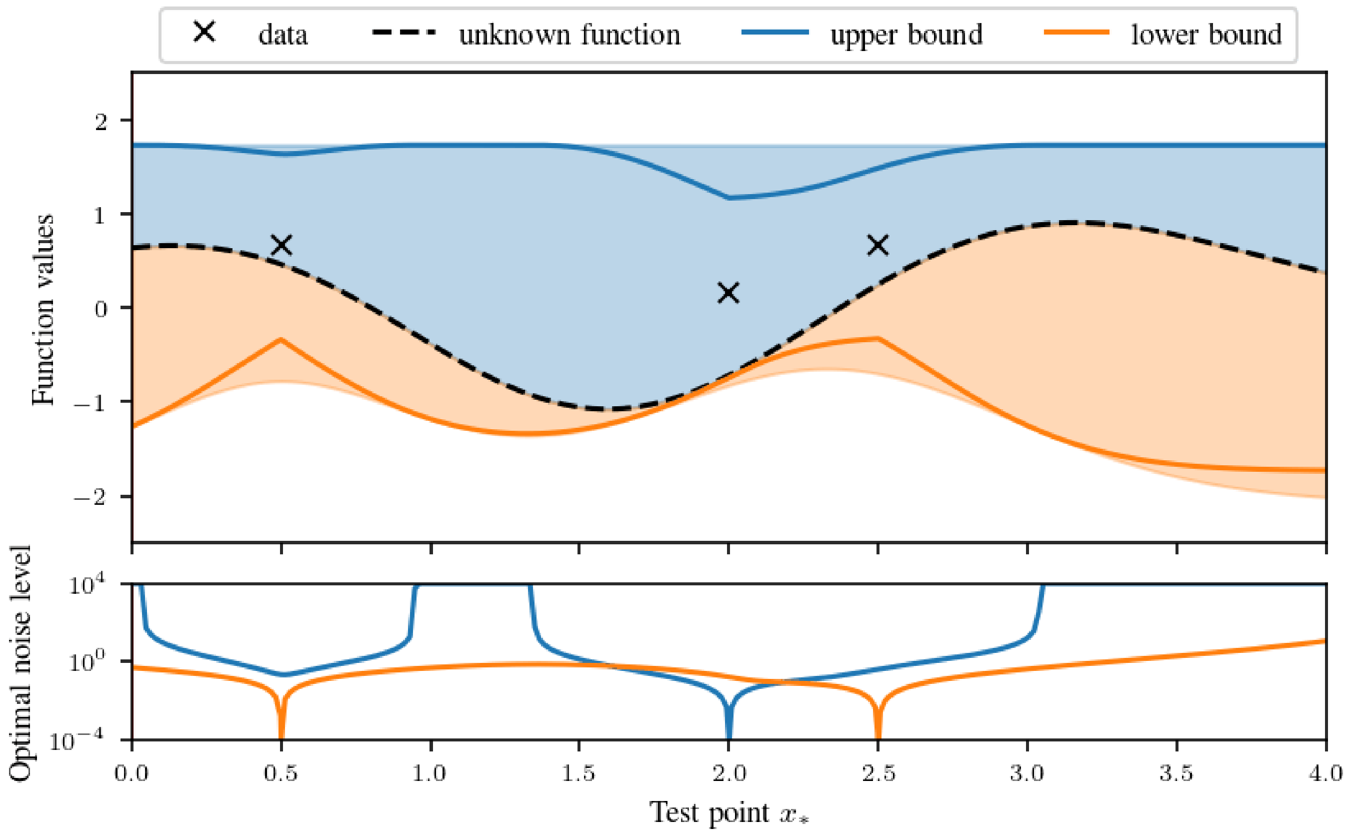
Data-generating model

$$y_i = f^{\mathrm{tr}}(x_i) + \epsilon_i$$

Uncertainty quantification

$$\underline{f}^{\sigma}(x_*) \leq \underline{f}(x_*) \leq f^{\mathrm{tr}}(x_*) \leq \overline{f}(x_*) \leq \overline{f}^{\sigma}(x_*)$$

Kernel regression bounds in Reproducing Kernel Hilbert Spaces (RKHS)



Relaxed bound = optimal bound for test-point-dependent parameter σ

Optimal noise parameter related to importance of RKHS-norm constraints → *More details at the poster session!*

Data-generating model

$$y_i = f^{\mathrm{tr}}(x_i) + \epsilon_i$$

Uncertainty quantification

$$\underline{f}^{\sigma}(x_*) \leq \underline{f}(x_*) \leq f^{\mathrm{tr}}(x_*) \leq \overline{f}(x_*) \leq \overline{f}^{\sigma}(x_*)$$

Numerical comparison: Safe control example

Tighter bound
↕
Higher success rate

Goal: Optimize control input to *satisfy constraints despite model uncertainty*

- Proposed bound *simultaneously* optimizes noise parameter
- Proposed bound enables safe subset-of-data strategy

Compare with **probabilistic bounds** for independent, (sub-)Gaussian noise:

$$\Pr\left[|f^{\mathrm{tr}}(x) - f_\sigma^\mu(x)| \leq \beta\sqrt{\Sigma(x)} \quad \forall x \in \mathcal{X}\right] \geq p$$

Depends on stochastic noise assumption

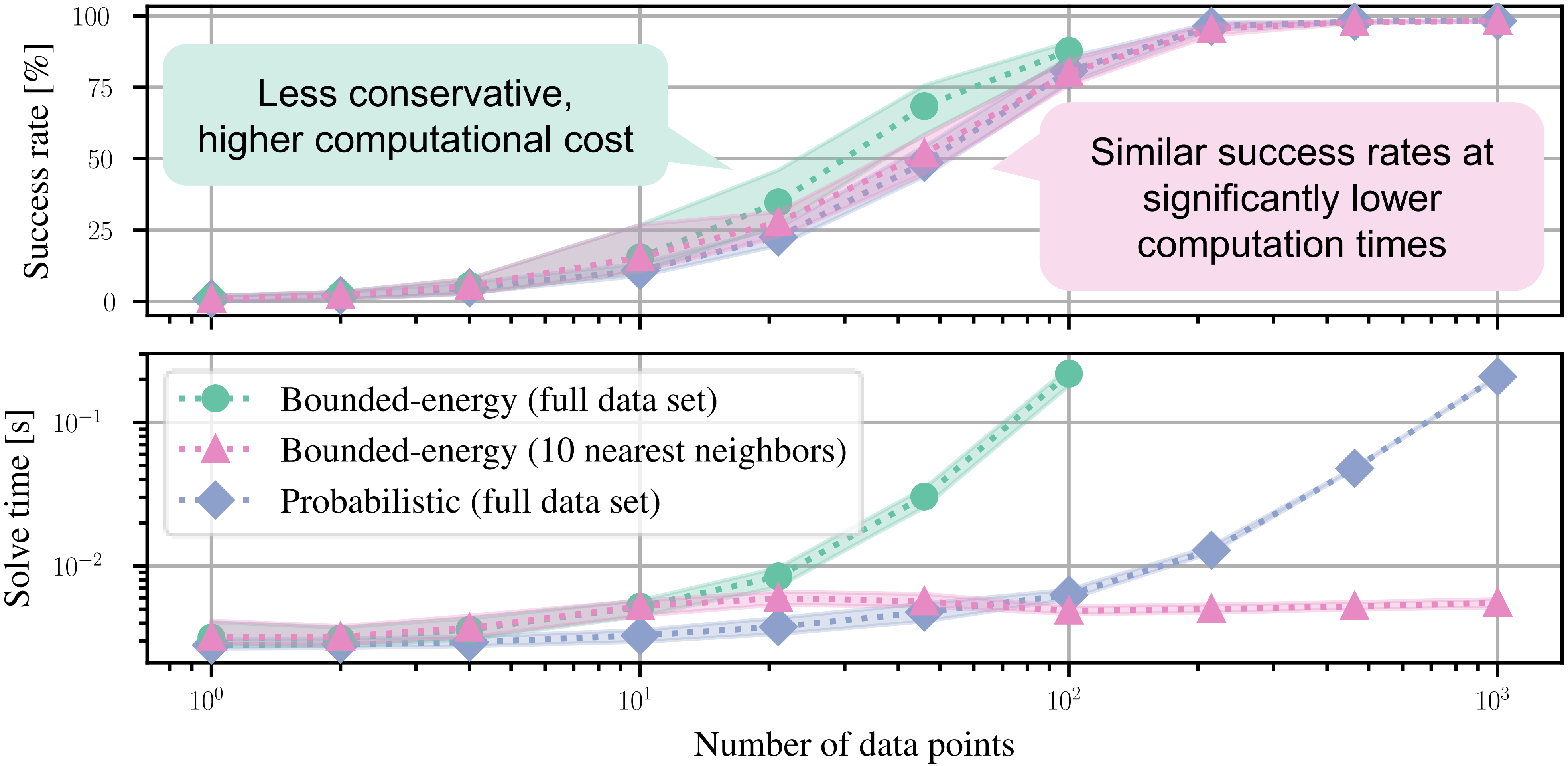
e.g. Srinivas et al., 2012; Abbasi-Yadkori et al., 2013; Chowdhury et al., 2017; Fiedler et al., 2021; Molodchyk et al., 2025

Numerical comparison: Safe control example

Tighter bound
↕
Higher success rate

Goal: Optimize control input to *satisfy constraints despite model uncertainty*

- Proposed bound *simultaneously* optimizes noise parameter
- Proposed bound enables safe subset-of-data strategy



Conclusions

Tight, distribution-free bound for kernel regression under energy-bounded noise:

- computed via scalar, unconstrained optimization,
- suitable for integration into downstream tasks (e.g. Bayesian optimization, safe control),
- enables certification of subset-of-data strategies.

Contact: **amlahr@ethz.ch**

Preprint and code: arXiv:2505.22235