Know Your Boundaries:

Constraining Gaussian Processes by Variational Harmonic Features

Arno Solin
Aalto University
`arno.solin@aalto.fi`

Manon Kok
Delft University of Technology
`m.kok-1@tudelft.nl`

Introduction

- **Gaussian processes** (GPs) provide a powerful framework for extrapolation, interpolation, and noise removal in regression and classification
- We constrain GPs to **arbitrarily-shaped domains** with boundary conditions
- **Applications** in, *e.g.*, imaging, spatial analysis, robotics, or general ML tasks
- As a pre-processing step, we solve a **Fourier-like generalised harmonic feature** representation of the GP prior in the domain of interest
- This both constrains the GP and attains a **low-rank representation** that is used for **speeding up inference**
- The method scales as $\mathcal{O}(nm^2)$ in prediction and $\mathcal{O}(m^3)$ in hyperparameter learning (n number of data, m features)
- A Titsias-style variational approach to allow the method to deal with **non-Gaussian likelihoods**

Model

- Consider a **boundary-constrained** GP model of form:

$$\begin{cases} f(\mathbf{x}) \sim \mathrm{GP}(0, \kappa(\mathbf{x}, \mathbf{x}')), & \mathbf{x} \in \Omega \\ \text{s.t. } f(\mathbf{x}) = 0, & \mathbf{x} \in \partial\Omega \end{cases}$$

$$\mathbf{y} \mid \mathbf{f} \sim \prod_{i=1}^{n} p(y_i \mid f(\mathbf{x}_i)) \quad \text{likelihood}$$

where $(\mathbf{x}_i, y_i)$ are the n input–output pairs

- Given a domain $\Omega \subset \mathbb{R}^d$ (d typically 1–3), we follow [3] for projecting the GP onto the eigenbasis of the Laplace operator, ∇^2, that solves the eigenvalue problem:

$$\begin{cases} -\nabla^2 \phi_j(\mathbf{x}) = \lambda_j^2 \phi_j(\mathbf{x}), & \mathbf{x} \in \Omega, \\ \phi_j(\mathbf{x}) = 0, & \mathbf{x} \in \partial\Omega. \end{cases}$$

- The approximate eigenvalues and eigenfunctions of the Laplacian in Ω (s.t. the the boundary conditions) can be solved numerically (see **Fig. 3**).

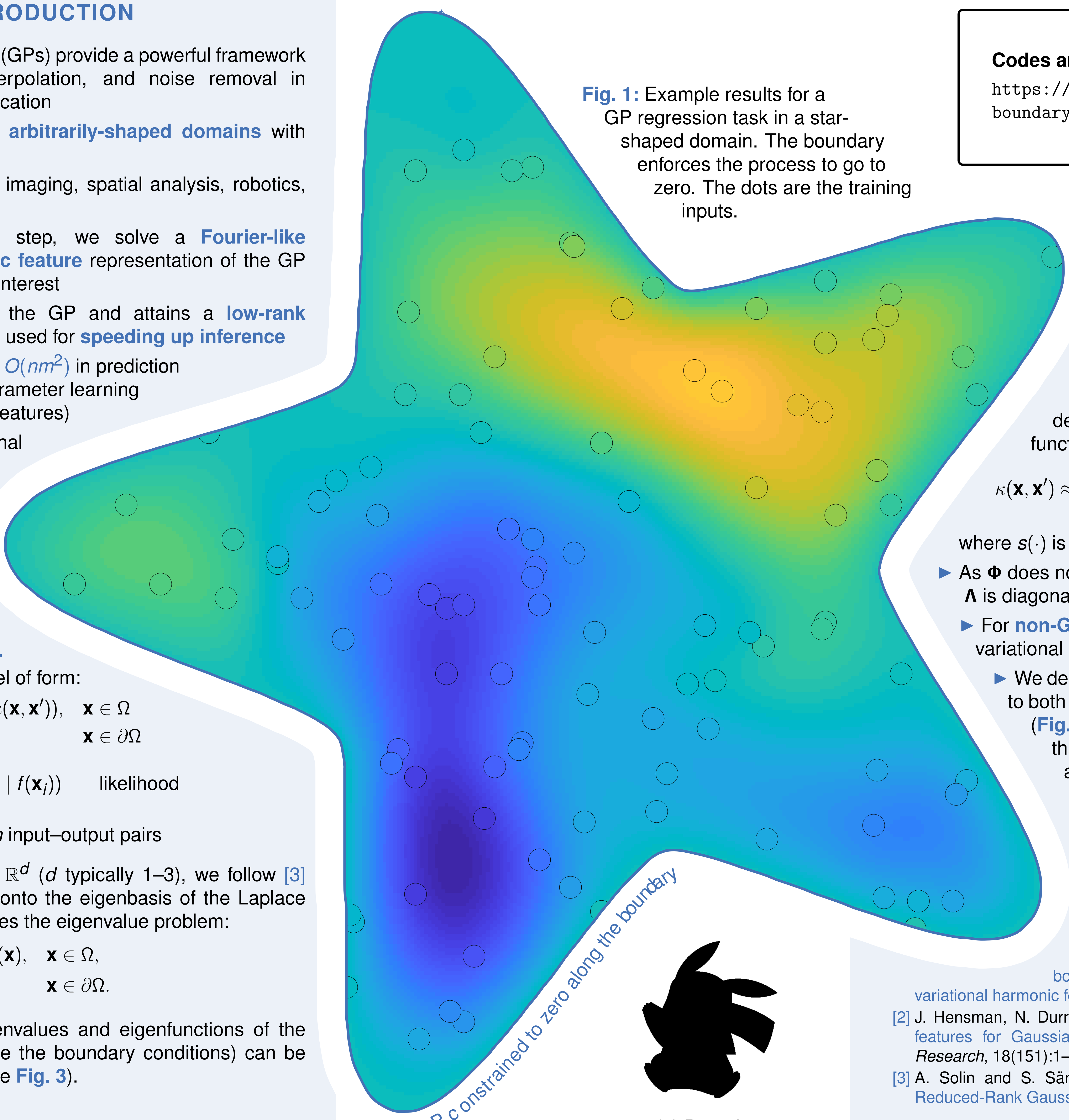


Fig. 1: Example results for a GP regression task in a star-shaped domain. The boundary enforces the process to go to zero. The dots are the training inputs.

Codes and resources available:
`https://github.com/AaltoML/boundary-gp`

GP Inference

- We require the covariance function κ to be **stationary**
- Leverage [3] for **approximating the covariance function** by the eigen-decomposition and the spectral density function:

$$\kappa(\mathbf{x}, \mathbf{x}') \approx \sum_{j=1}^{m} s(\lambda_j)\, \phi_j(\mathbf{x})\, \phi_j(\mathbf{x}') = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^\mathsf{T},$$

where $s(\cdot)$ is the **spectral density function** of $\kappa(\cdot,\cdot)$

- As $\boldsymbol{\Phi}$ does not depend on the hyperparameters and $\boldsymbol{\Lambda}$ is diagonal, we also get a **computational boost**
- For **non-Gaussian likelihoods**, we set up a variational approach and maximize the ELBO [2]
- We demonstrate the applicability of our method to both GP regression (**Fig. 1**) and classification (**Fig. 4**) on simulated data, where we show that encoding boundary information by this approach is beneficial (**Fig. 2**)
- We also consider a **Log-Gaussian Cox process** (Poisson likelihood) study of tick bite count modelling

References

[1] A. Solin and M. Kok (2019). Know your boundaries: Constraining Gaussian processes by variational harmonic features. *Proceedings of AISTATS.*

[2] J. Hensman, N. Durrande, and A. Solin (2018). Variational Fourier features for Gaussian processes. *Journal of Machine Learning Research*, 18(151):1–52.

[3] A. Solin and S. Särkkä (submitted). Hilbert Space Methods for Reduced-Rank Gaussian Process Regression. arXiv:1401.5508

Contact details: arno.solin@aalto.fi, web address: http://arno.solin.fi, AISTATS 2019.

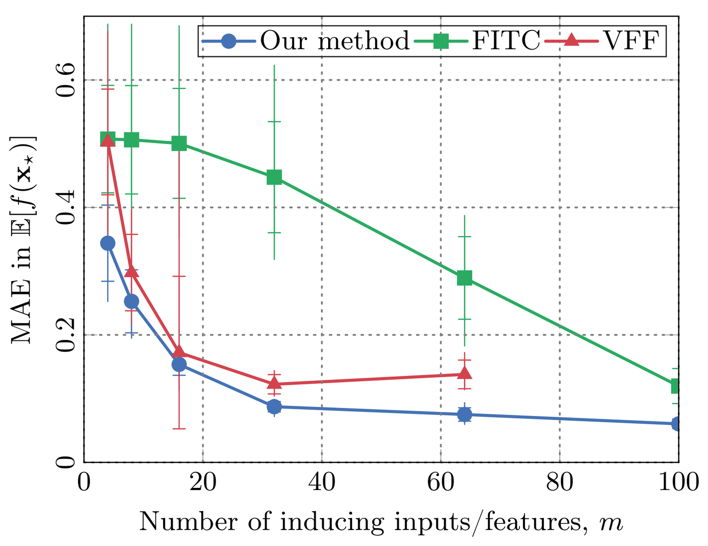
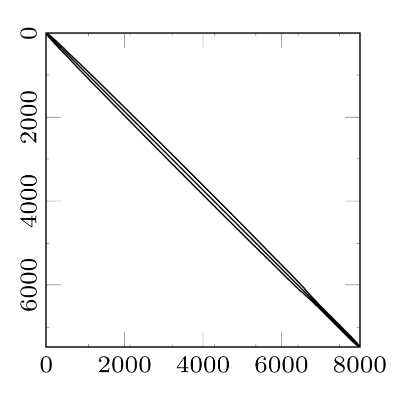
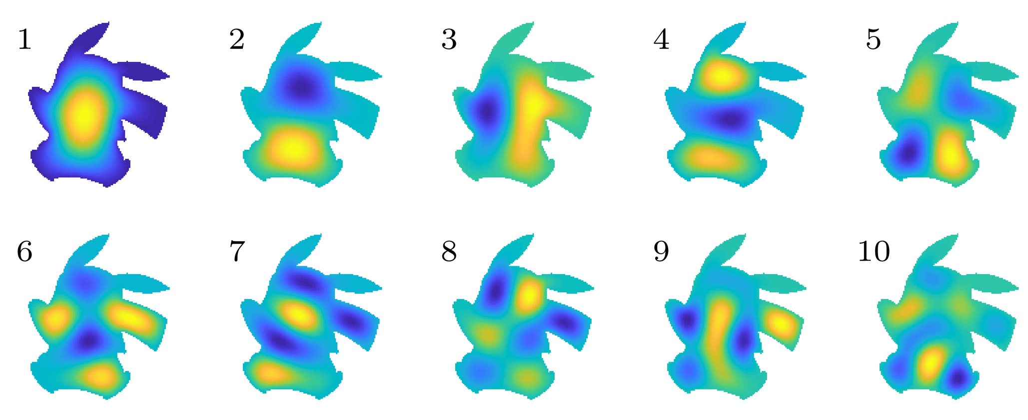


Fig. 2: Effect of increasing the number of inducing inputs/features in **Fig. 1**.

(a) Domain



(b) Sparse stencil matrix



(c) Harmonic basis functions

Fig. 3: Example domain (a) for which we can numerically compute the harmonic basis functions, $\phi_j(\mathbf{x})$, (c) using the sparse stencil matrix (b).

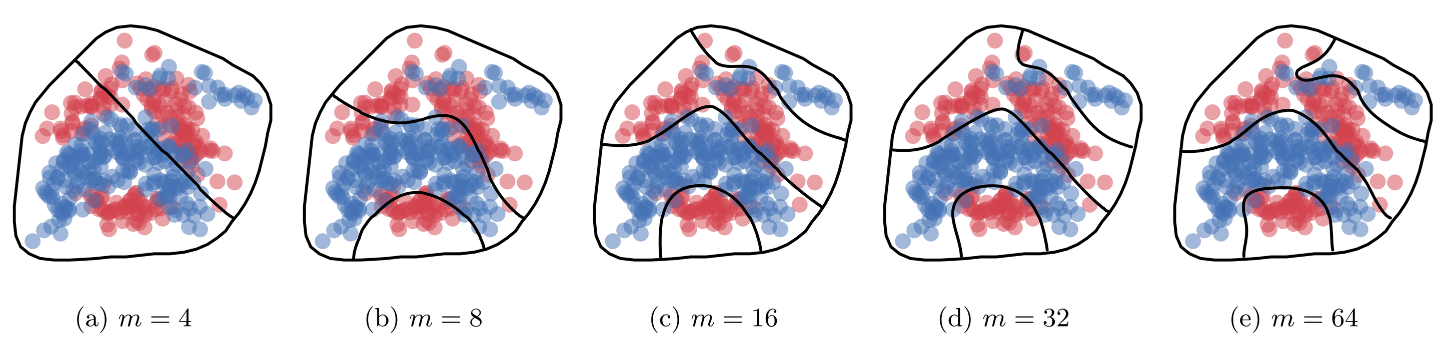


Fig. 4: Increasing the number of inducing features for the **banana classification dataset** with a hard decision boundary. The coloured points represent training data and the decision boundaries are black lines. The outermost line is the pre-defined hard decision boundary.