CS-E4075 project work

Spring 2021

Aalto University

Project work outline

Completing the course gives 5 ECTS points.

You can get 2 extra ECTS points by completing an optional small project:

- The work should be done in groups of 1-4 people (expected workload scales)

The project work timeline:

- Kick-off session now
- **Q&A support session** on thursday 4th of March 10:15
- Hand-in a **detailed project report** (one per group) no later than 12th of March (eg. 4-10 pages)
- **Project work seminar** on 18th of March with group presentations (10-20 min)

The project work consist of **at least one of following tasks**

1. **Analyze your favourite dataset** with Gaussian process models of you topic
   - Compare the GP model(s) against baseline methods. Study the inference of the GP model, and study the predictive posteriors of your GP model in your dataset.
2. **Literature survey/comparison** of more advanced Gaussian process models/methods of your topic
   - Read about your topic from scientific literature. Review and discuss the topic.
3. **Implementation of more advanced** Gaussian process models of your topic
   - Choose your favourite programming language and/or library, and implement an advanced GP model of your topic. Describe your implementation and test it.

Topics

1. Iterative kernel learning
2. Bayesian optimization with Gaussian Processes
3. Bayesian quadrature
4. Relationship between Neural networks and GPs
5. Multioutput Gaussian processes & Kronecker structures
6. Gaussian processes for big data
7. Gaussian processes with monotonicity
8. Gaussian process latent variable model (GPLVM)
9. Convolutional Gaussian processes
10. Gaussian process inference (eg. VI, EP, MCMC)
11. Deep Gaussian processes
12. State-space GPs
13. Dynamical GPs
14. Own topic (contact Markus/Arno)

- Please tell us your topic/group: markus.o.heinonen@aalto.fi or arno.solin@aalto.fi

Composite kernel learning

- Learn a composite kernel function form
  - Automatic Statistician (AS)
    - Duvenaud et al 2013. Structure Discovery in Nonparametric Regression through Compositional Kernel Search
    - Kim et al 2018. Scaling up the Automatic Statistician: Scalable Structure Discovery using Gaussian Processes
  - Compositional Kernel Search (CKS) / Neural Kernel Networks (NKN)
    - Sun et al 2018. Differentiable Compositional Kernel Learning for Gaussian Processes
  - + others

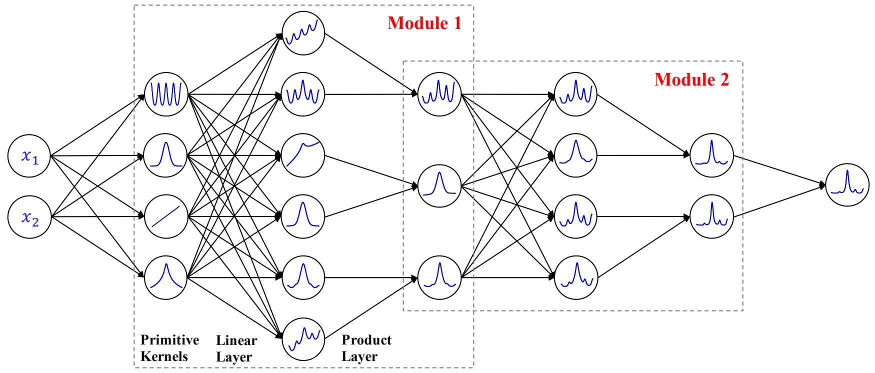
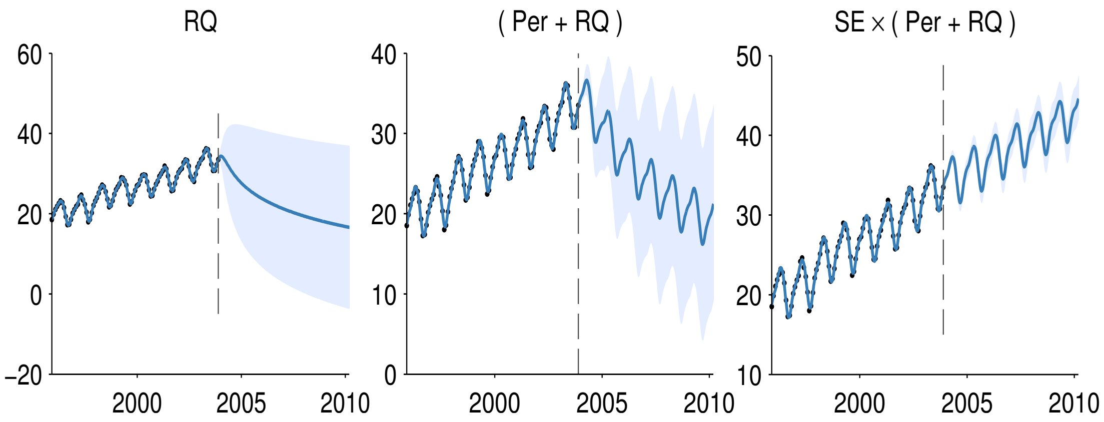
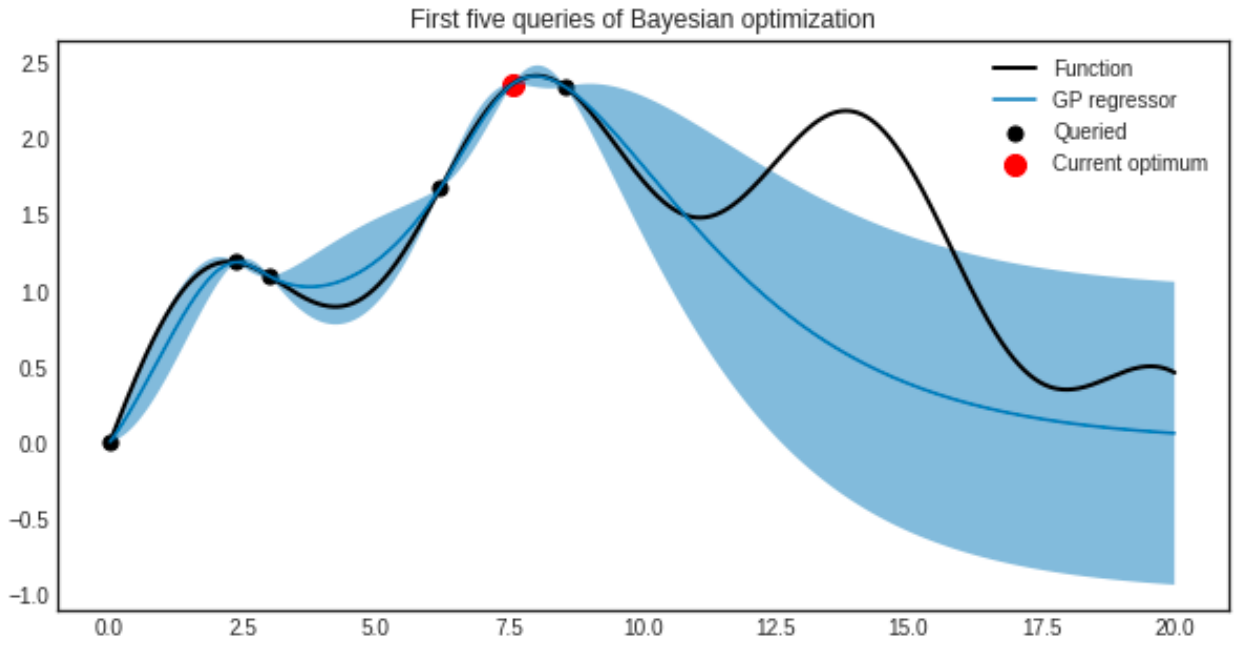


Figure 2. Neural Kernel Network: each module consists of a **Linear** layer and a **Product** layer. NKN is based on compositional rules for kernels, thus every individual unit itself represents a kernel.

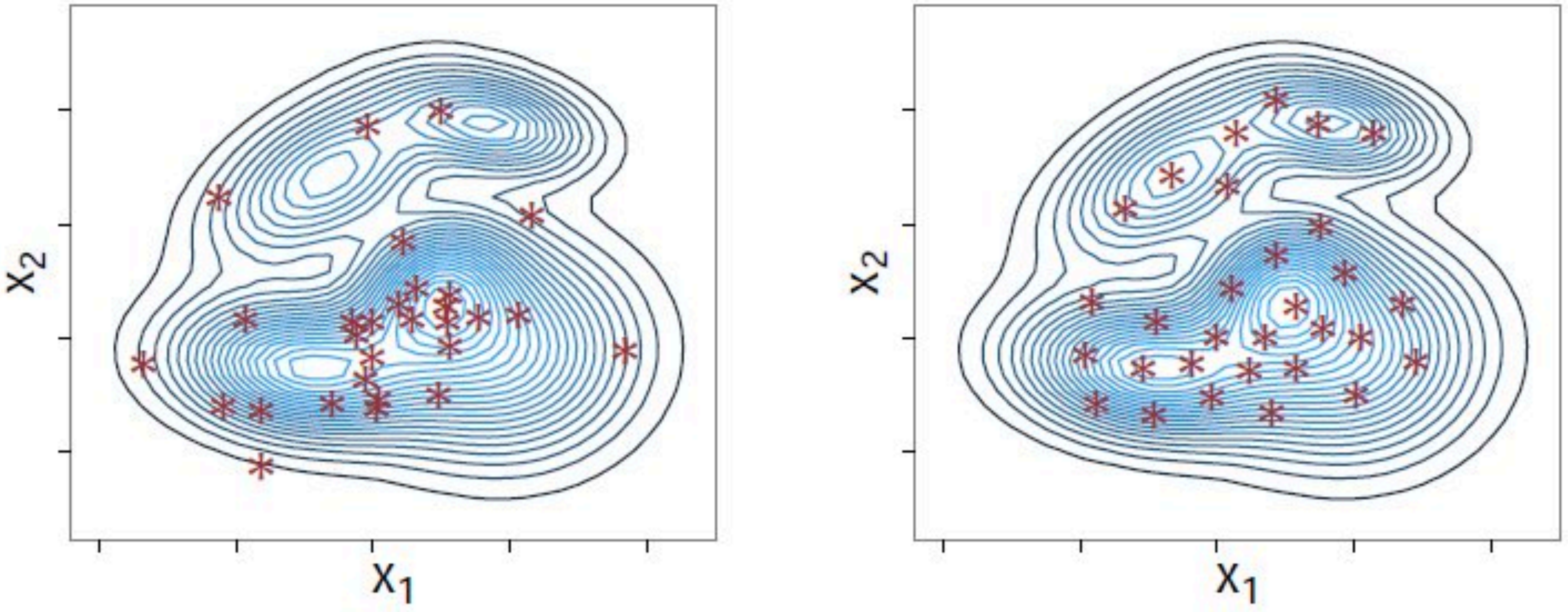


Bayesian optimization



- Find function maximum with least number of tries
  - Shahriari et al: Taking the human out of the loop: A review of bayesian optimization, 2016

Bayesian quadrature



- Estimate an integral numerically with GP assumptions
- Many references

Theoretical neural network / GP connections

- Neural networks are known to converge to Gaussian processes at infinitely wide layers
  - Williams 1997. Computing with infinite networks
  - Lee 2017. Deep neural networks as gaussian processes
  - Matthews 2018. Gaussian process behaviour in wide deep neural networks
- Neural networks induce Neural Tangent Kernel (NTK) behavior
  - Jacot 2017. Neural tangent kernel: Convergence and generalization in neural networks

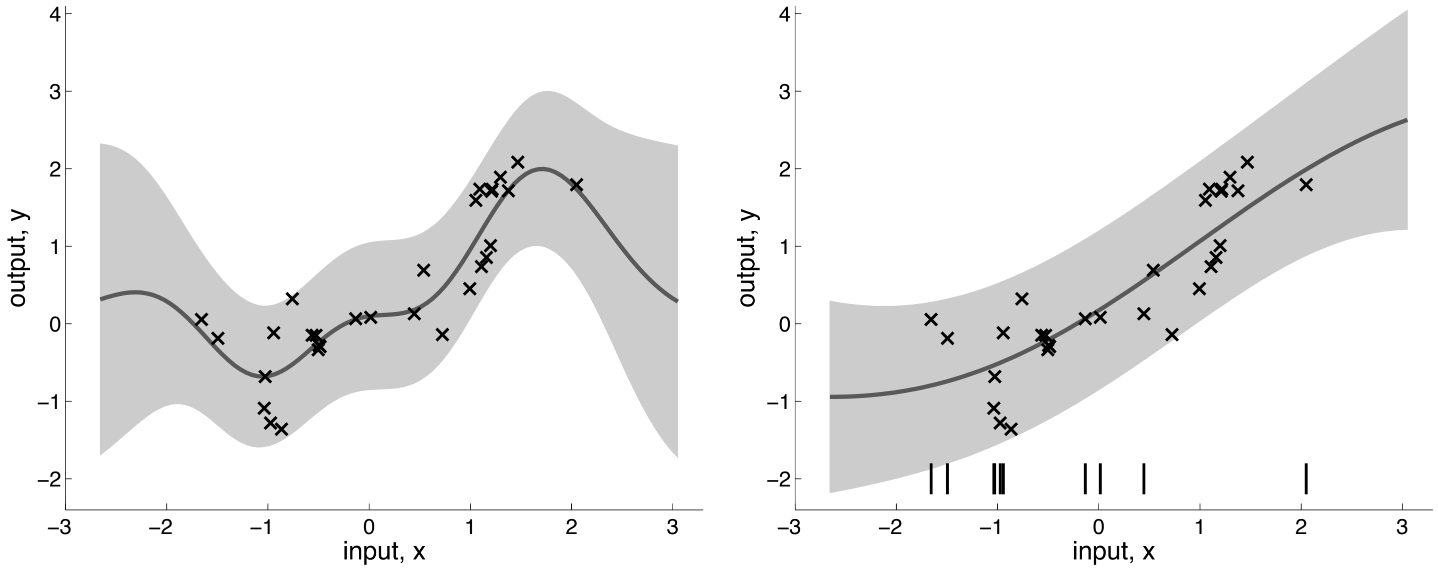
Multioutput GPs

- Bonilla et al 2006. Multi-task Gaussian process prediction
- Stegle et al 2021. Efficient inference in matrix-variate Gaussian models with iid observation noise

GPs for big data

- Scaling GPs to million/billion points
- Hensman et al 2015: Scalable Variational Gaussian process Classification
- Wilson et al 2015: Kernel Interpolation for Scalable Structured Gaussian Processes (KISS-GP)
- Wang 2019. Exact Gaussian Processes on a Million Data Points
- Liu 2019. When Gaussian Process Meets Big Data: A Review of Scalable GPs

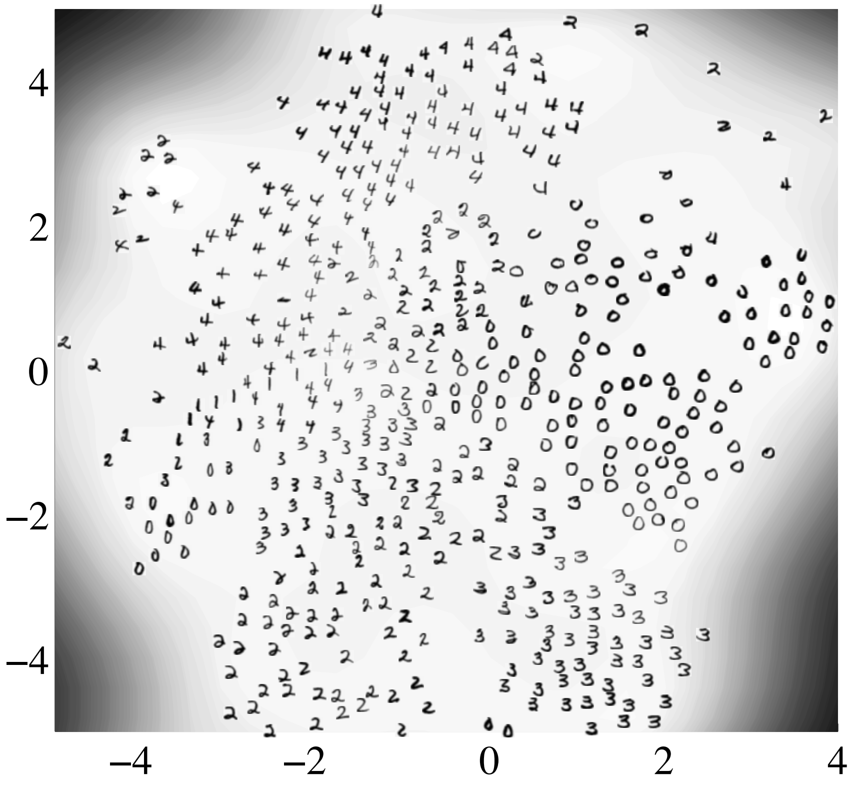
Constrained GPs

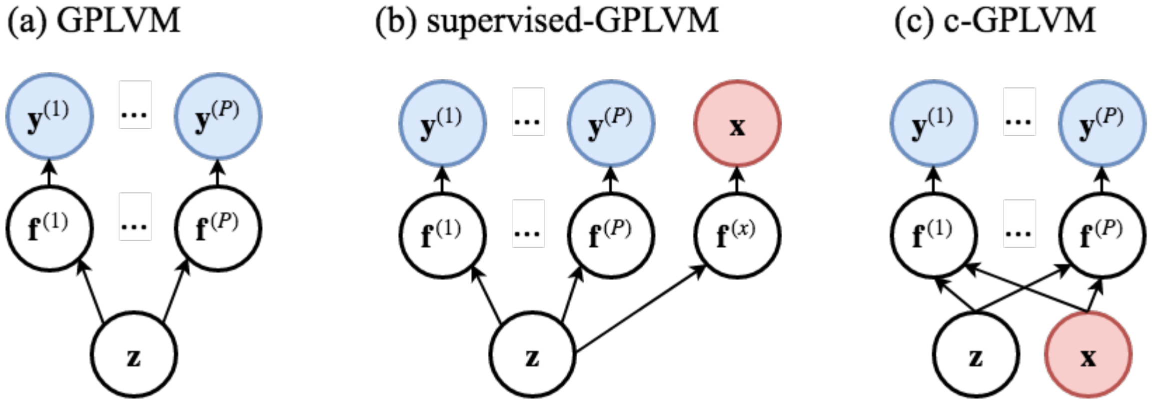


- Riihimäki et al 2011. Gaussian processes with monotonicity information
- Jidling 2017. Linearly constrained Gaussian Processes

GPLVMs

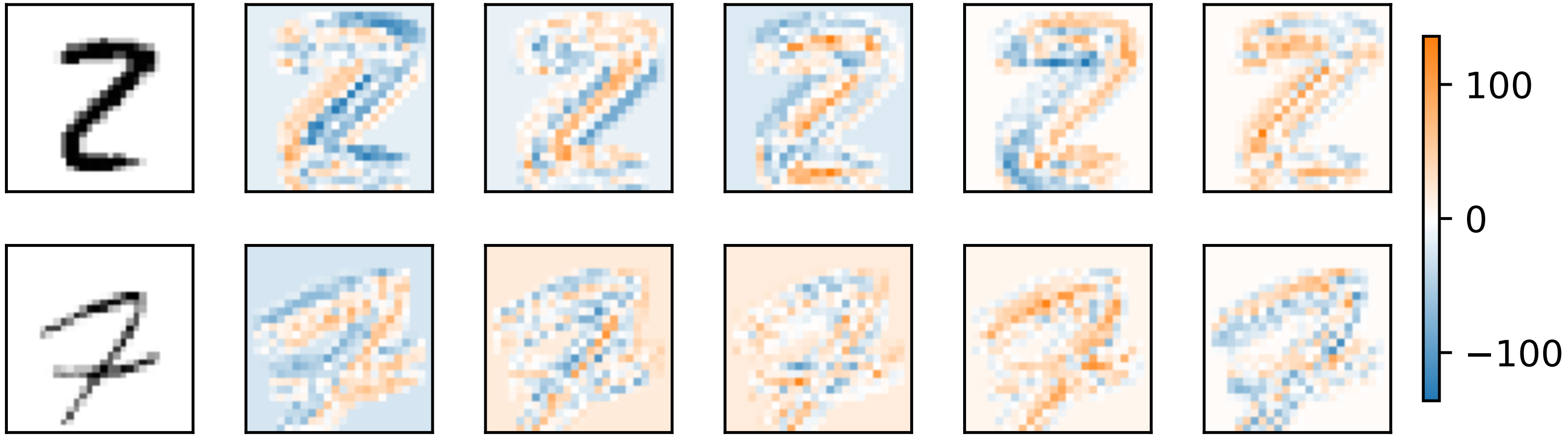
- Titsias 2010. Bayesian Gaussian Process Latent Variable Model
- Märtens 2018. Decomposing feature-level variation with Covariate Gaussian Process Latent Variable Models





Convolutional GPs

- Apply GPs to images
  - Wilk 2017. Convolutional Gaussian Processes
  - Dutordoir 2019. Bayesian Image Classification with Deep Convolutional Gaussian Processes



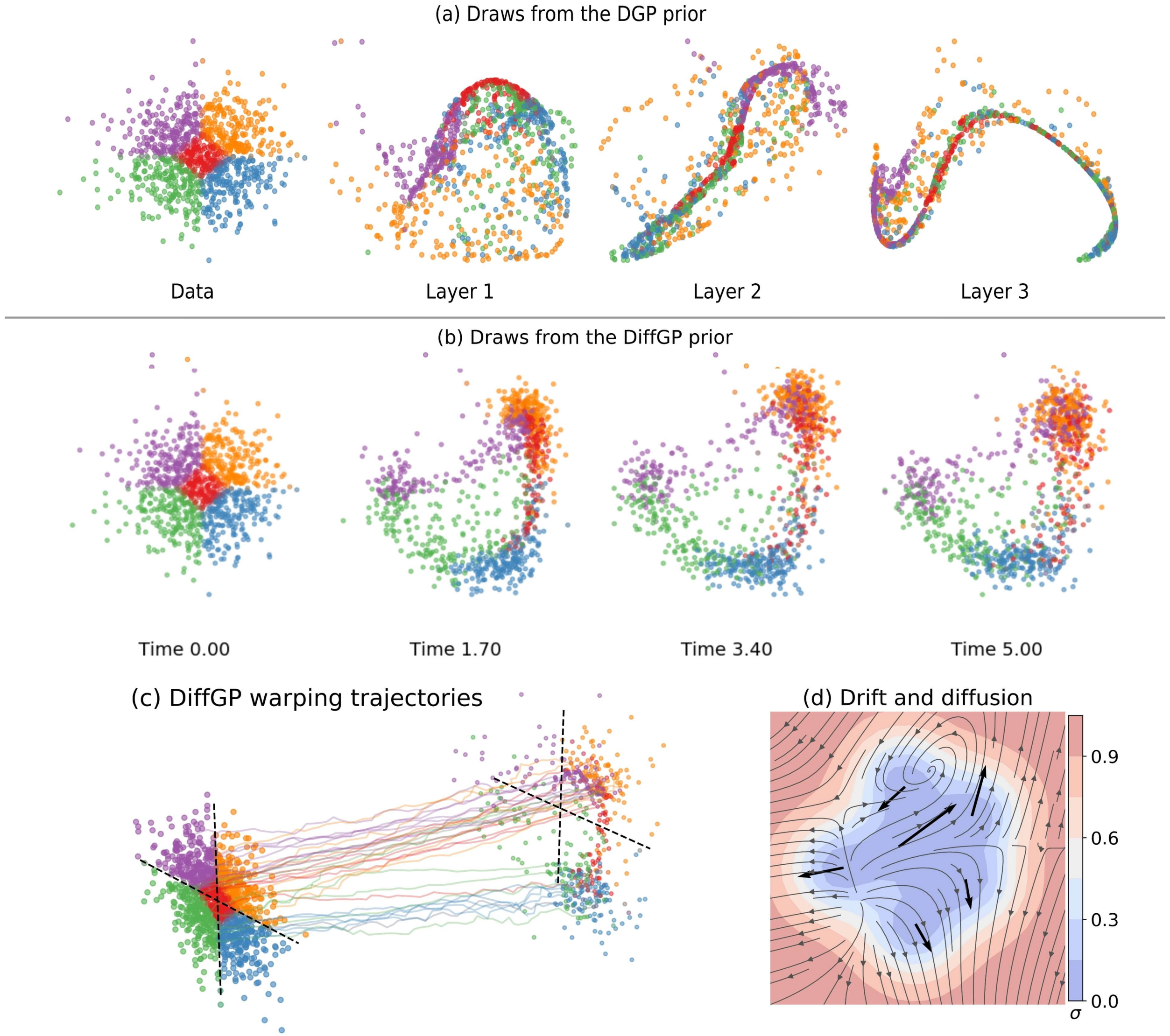
(a) Conv-GP

GP inference (VI, EP, MCMC)

- Inference of GPs is a hot topic, often done in combination with deep GPs
  - Salimbeni 2017. Doubly Stochastic Variational Inference for Deep Gaussian Processes
  - Havasi 2018. Inference in Deep Gaussian Processes using Stochastic Gradient Hamiltonian Monte Carlo
  - Salimbeni 2019. Deep Gaussian Processes with Importance-Weighted Variational Inference

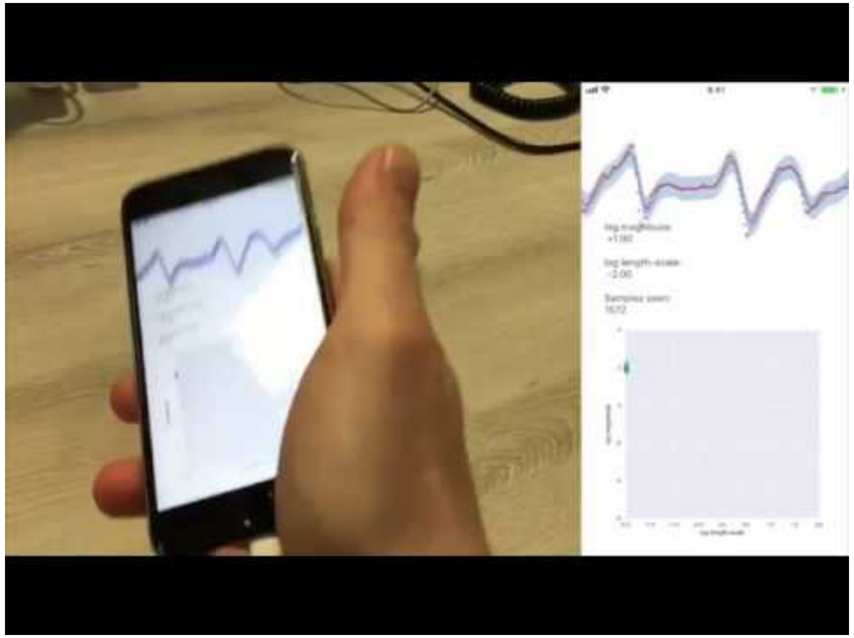
Deep GPs

- The deep GP formulation
  - Salimbeni 2017. Doubly Stochastic Variational Inference for Deep Gaussian Processes
  - Damianou 2013. Deep Gaussian Processes
- Deep GPs are known to suffer from rank collapse
  - Duvenaud 2014. Avoiding pathologies in very deep networks
  - Hegde 2019. Deep learning with differential Gaussian process flows



State-space GPs

- Nickish 2018. State Space Gaussian Processes with Non-Gaussian Likelihood
- Solin 2018. Infinite-Horizon Gaussian Processes



Dynamical GPs

- Learning system dynamics with GPs
    - Wang 2008. Gaussian process dynamical models for human motion
    - Macdonald 2015. Controversy in mechanistic modelling with Gaussian processes
    - Heinonen 2018. Learning unknown ODE models with Gaussian processes
- Applications in RL
    - Deisenroth 2014. Gaussian Processes for Data-Efficient Learning in Robotics and Control [PILCO]
    - Kamthe 2017. Data-Efficient Reinforcement Learning with Probabilistic Model Predictive Control