The process of modelling systems with multiple outputs has a lot of practical value. Example: climate modelling.

Our goal is to incorporate the output relationships into the modelling process to improve predictions [1].

### GPAR Model

Suppose we have a set of functions that have the following relationships:

\[
y_1(x) = f_1(x) \\
y_2(x) = f_2(x,y_1(x)) \\
\vdots \\
y_n(x) = f_n(x,y_1(x),\ldots,y_{n-1}(x))
\]

**Training**

1. Find an appropriate ordering of each of the functions using a greedy approach.
2. Train each GP with inputs composed of available observations and outputs from foregoing GPs.

**Prediction**

\[
y_1 \leftarrow GP_1 + x \\
y_2 \leftarrow GP_2 + x + y_1 \\
\vdots \\
y_n \leftarrow GP_n + x + y_1 + \ldots + y_{n-1}
\]

### Experiments

**GPAR vs independent GPs (IGPs) on synthetic data:**

**GPAR vs IGP using sparse GPs on real data:**

**Noise structure recovery on synthetic data:**

### Freeze-Thaw Bayesian Optimization

Freeze-thaw is an information-theoretic approach that uses Bayesian optimization for hyperparameter tuning [2]. Our goal is to improve this with GPAR.

**Approach**

- Model each loss during the training process using a different set of GPs (using a custom kernel).
- Model the asymptotic cross-validation loss over a set of feasible hyperparameters using a GP.

### Possible Extensions

Deep GPs, parameter tying, optimal conditional ordering, neural architecture search.

### References