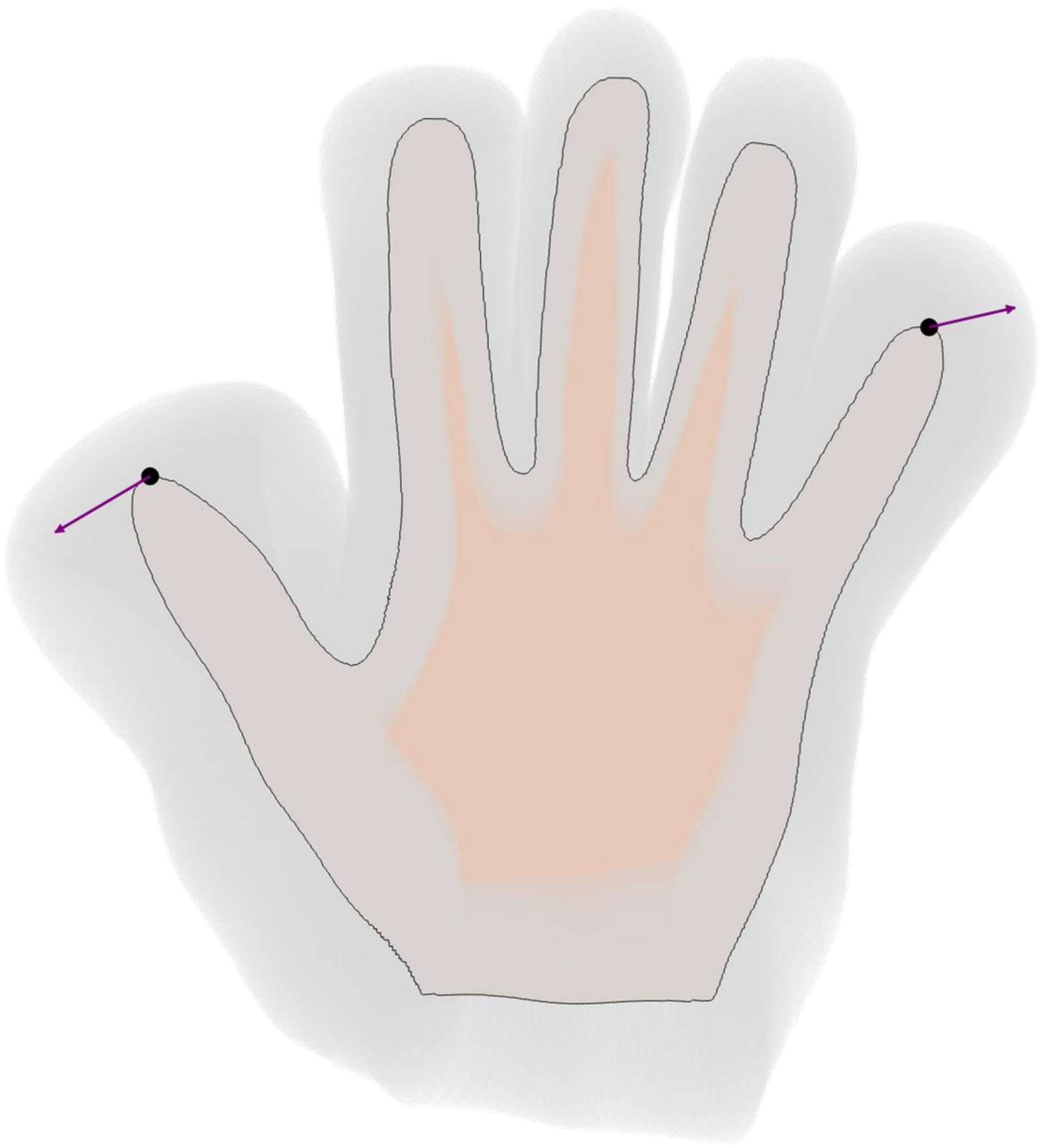
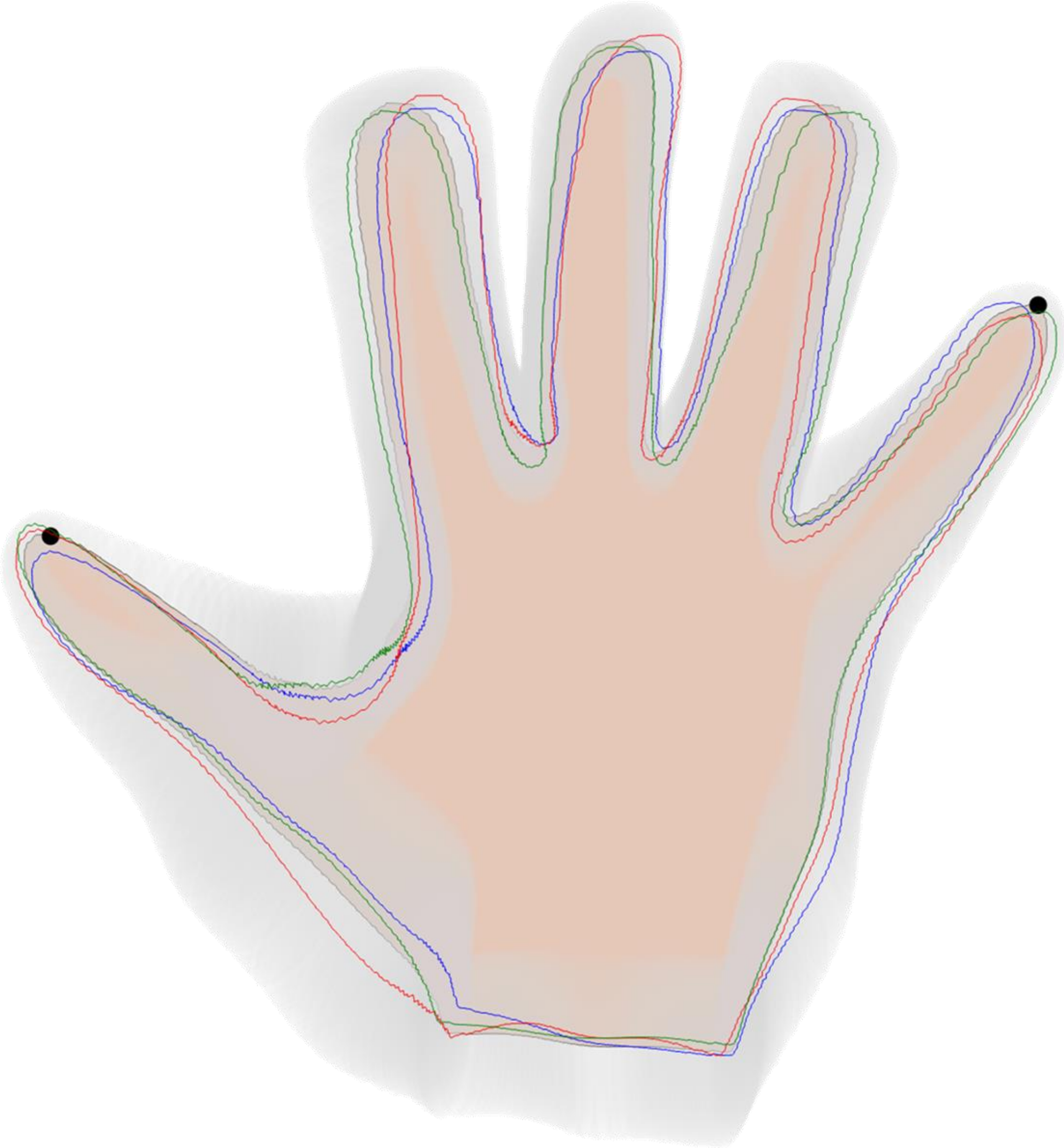
University of Basel

Gaussian process regression

Incorporating known deformations

University of Basel





Conditional distribution

University of Basel

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \\ \mu_{n+1} \\ \vdots \\ \mu_{n+m} \end{pmatrix}, \begin{pmatrix} \Sigma_{1,1} & \dots & \Sigma_{1,n} & \Sigma_{1,n+1} & \dots & \Sigma_{1,m+n} \\ \vdots & & \vdots & & & \vdots \\ \Sigma_{n,1} & & \Sigma_{n,n} & \Sigma_{n,n+1} & & \Sigma_{n,m+n} \\ \Sigma_{n+1,1} & & \Sigma_{n+1,n} & \Sigma_{n+1,n+1} & & \Sigma_{n+1,m+n} \\ \vdots & \dots & \vdots & & \dots & \vdots \\ \Sigma_{m+n,1} & & \Sigma_{m+n,n} & \Sigma_{m+n,n+1} & & \Sigma_{m+n,m+n} \end{pmatrix} \right)$$

Conditional distribution

University of Basel

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \\ \mu_{n+1} \\ \vdots \\ \mu_{n+m} \end{pmatrix}, \begin{pmatrix} \Sigma_{1,1} & \cdots & \Sigma_{1,n} & \Sigma_{1,n+1} & \cdots & \Sigma_{1,m+n} \\ \vdots & & \vdots & & & \vdots \\ \Sigma_{n,1} & & \Sigma_{n,n} & \Sigma_{n,n+1} & & \Sigma_{n,m+n} \\ \Sigma_{n+1,1} & & \Sigma_{n+1,n} & \Sigma_{n+1,n+1} & & \Sigma_{n+1,m+n} \\ \vdots & \cdots & \vdots & & \cdots & \vdots \\ \Sigma_{m+n,1} & & \Sigma_{m+n,n} & \Sigma_{m+n,n+1} & & \Sigma_{m+n,m+n} \end{pmatrix} \right)$$

Simplified notation: $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$

Conditional distribution

University of Basel

For
$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}\right)$$

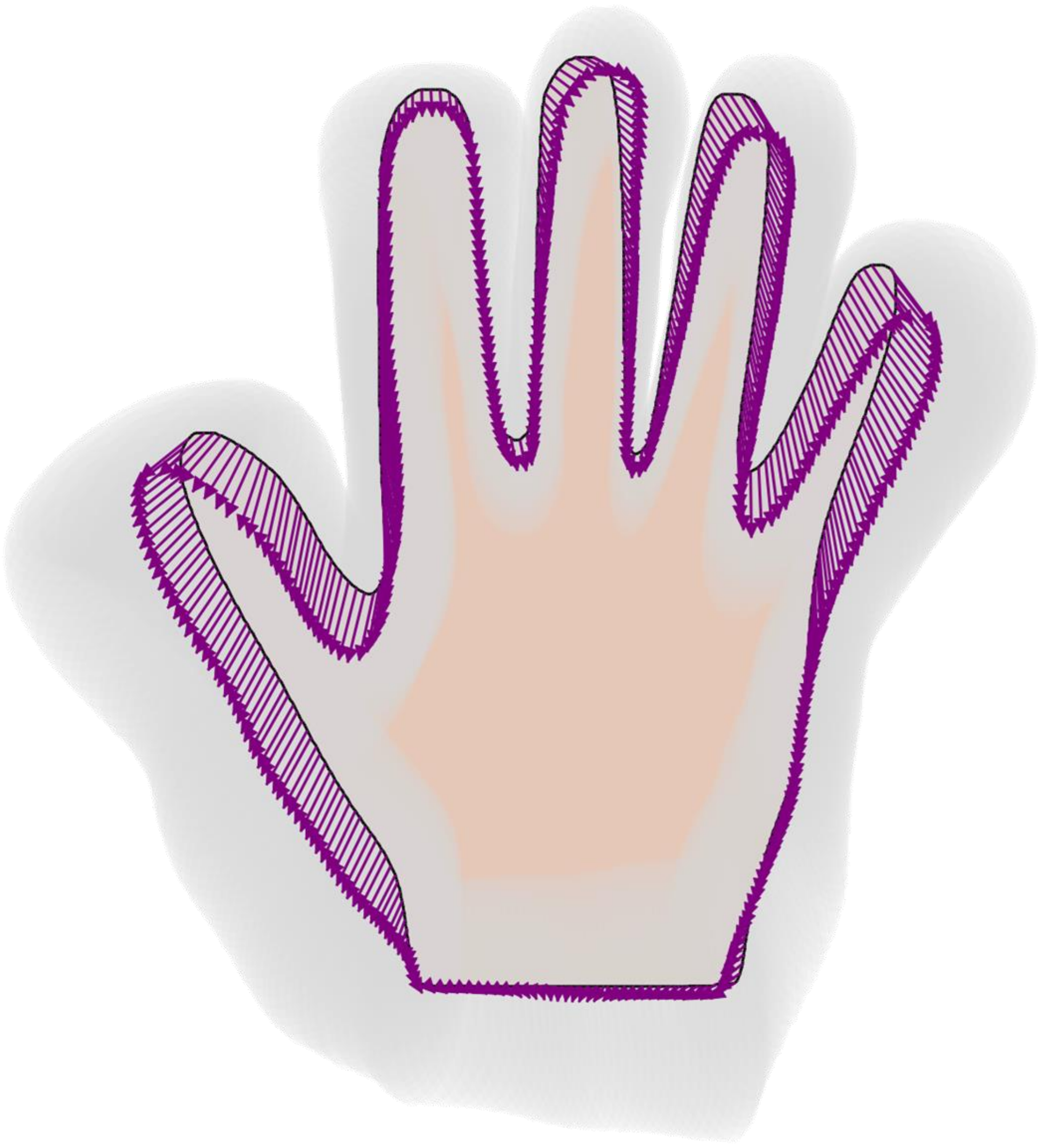
and given observations
$$Y = \tilde{\boldsymbol{a}} = \begin{pmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_m \end{pmatrix}$$

the conditional distribution $X|Y = \tilde{\mathbf{a}}$ is a normal distribution $N(\overline{\mu}, \overline{\Sigma})$.

Its parameters are known in closed form:

$$\overline{\mu} = \mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(\tilde{\mathbf{a}} - \mu_Y)$$
$$\overline{\Sigma} = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$$

Prior model

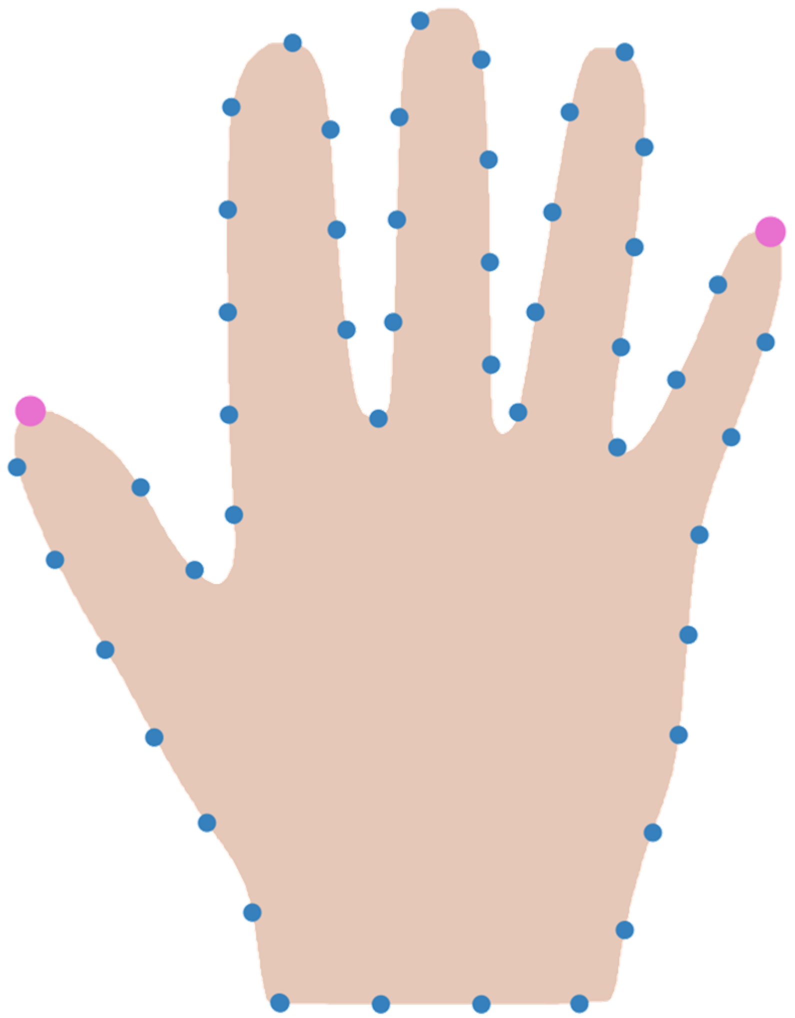


Shape variations modelled using

$$u \sim GP(\mu, k)$$

We call this the prior model.

Prior model



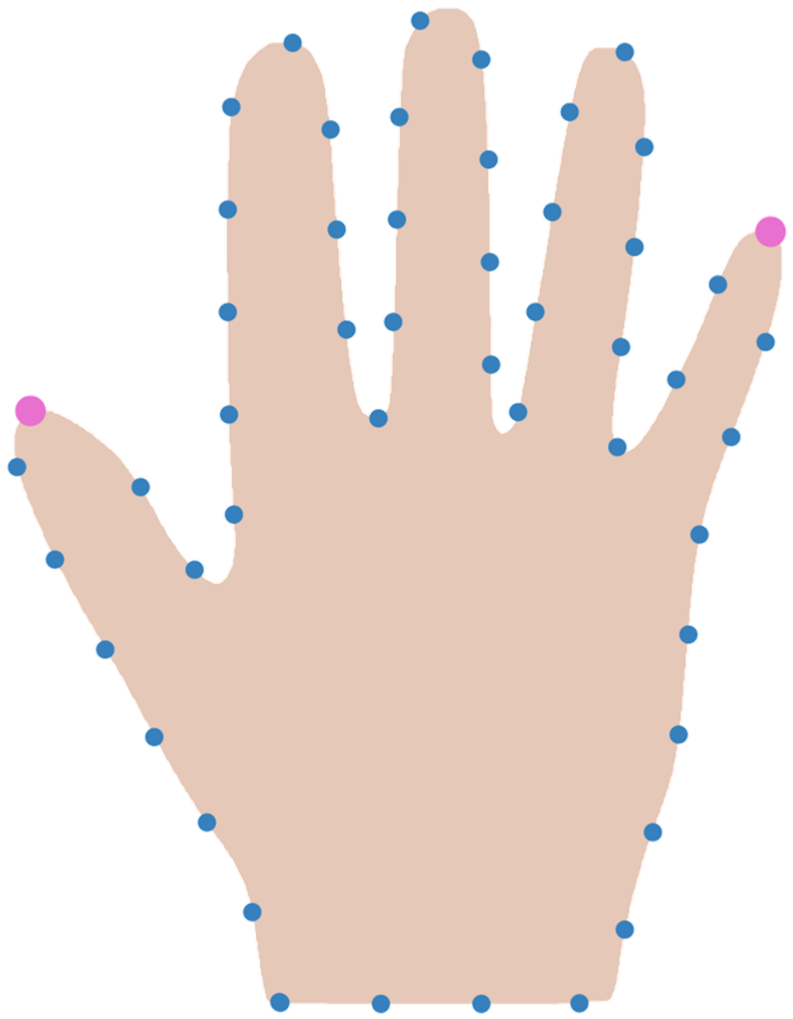
Prior model: $\mathrm{u} \sim GP(\mu, k)$

- For any discretization we can write:

$$\begin{pmatrix} u(x_1) \\ \vdots \\ u(x_n) \\ u(y_1) \\ \vdots \\ u(y_m) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x_1) \\ \vdots \\ \mu(x_n) \\ \mu(y_1) \\ \vdots \\ \mu(y_m) \end{pmatrix}, \begin{pmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) & k(x_1, y_1) & \dots & k(x_1, y_m) \\ \vdots & & & & & \\ k(x_n, x_1) & & k(x_n, x_n) & k(x_n, y_1) & & k(x_n, y_m) \\ k(y_1, x_1) & & k(y_1, x_n) & k(y_1, y_1) & & k(y_1, y_m) \\ \vdots & & & & & \\ k(y_m, x_1) & & k(y_m, x_n) & k(y_m, y_1) & & k(y_m, y_m) \end{pmatrix}\right)$$

Prior model

University of Basel

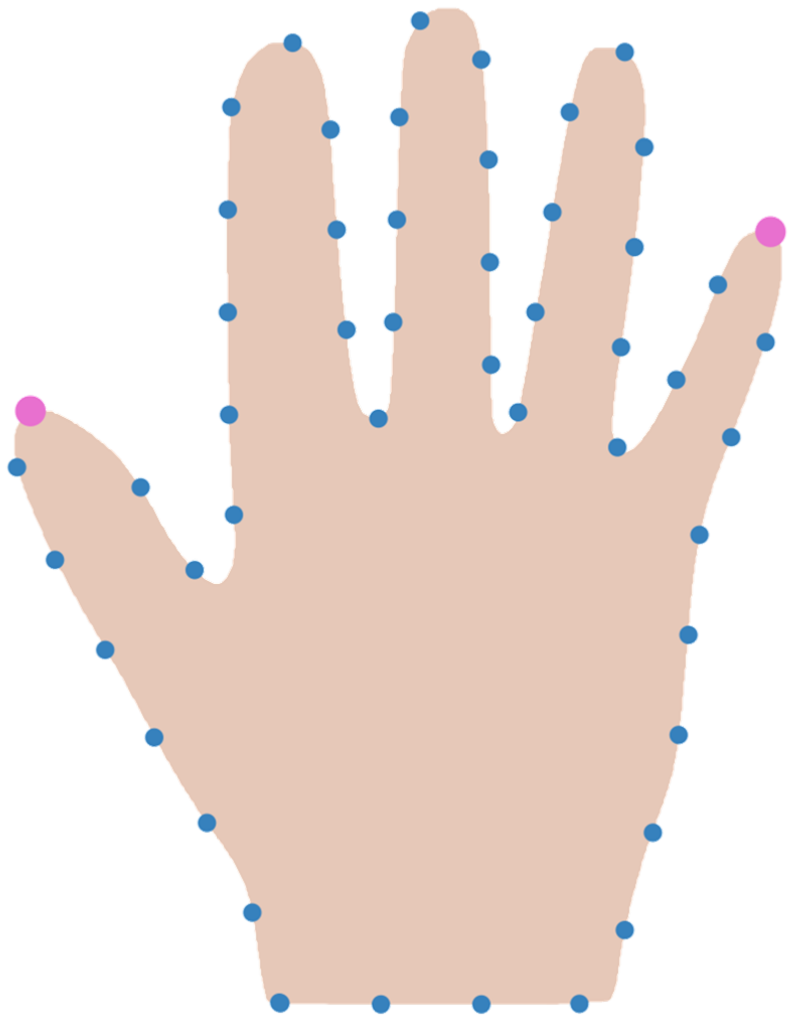


$$\begin{pmatrix} u_1(x_1) \\ u_2(x_1) \end{pmatrix} \qquad \begin{pmatrix} \mu_1(x_1) \\ \mu_2(x_1) \end{pmatrix} \qquad \begin{pmatrix} k_{11}(x_1, x_n) & k_{12}(x_1, x_n) \\ k_{21}(x_1, x_n) & k_{22}(x_1, x_n) \end{pmatrix}$$

$$\begin{pmatrix} u(x_1) \\ \vdots \\ u(x_n) \\ u(y_1) \\ \vdots \\ u(y_m) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x_1) \\ \vdots \\ \mu(x_n) \\ \mu(y_1) \\ \vdots \\ \mu(y_m) \end{pmatrix}, \begin{pmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) & k(x_1, y_1) & \dots & k(x_1, y_m) \\ \vdots & & & & & \\ k(x_n, x_1) & & k(x_n, x_n) & k(x_n, y_1) & & k(x_m, y_m) \\ k(y_1, x_1) & & k(y_1, x_n) & k(y_1, y_1) & & k(y_1, y_m) \\ \vdots & & & & & \\ k(y_m, x_1) & & k(y_m, x_n) & k(y_m, y_1) & & k(y_m, y_m) \end{pmatrix} \right)$$

Prior model

University of Basel

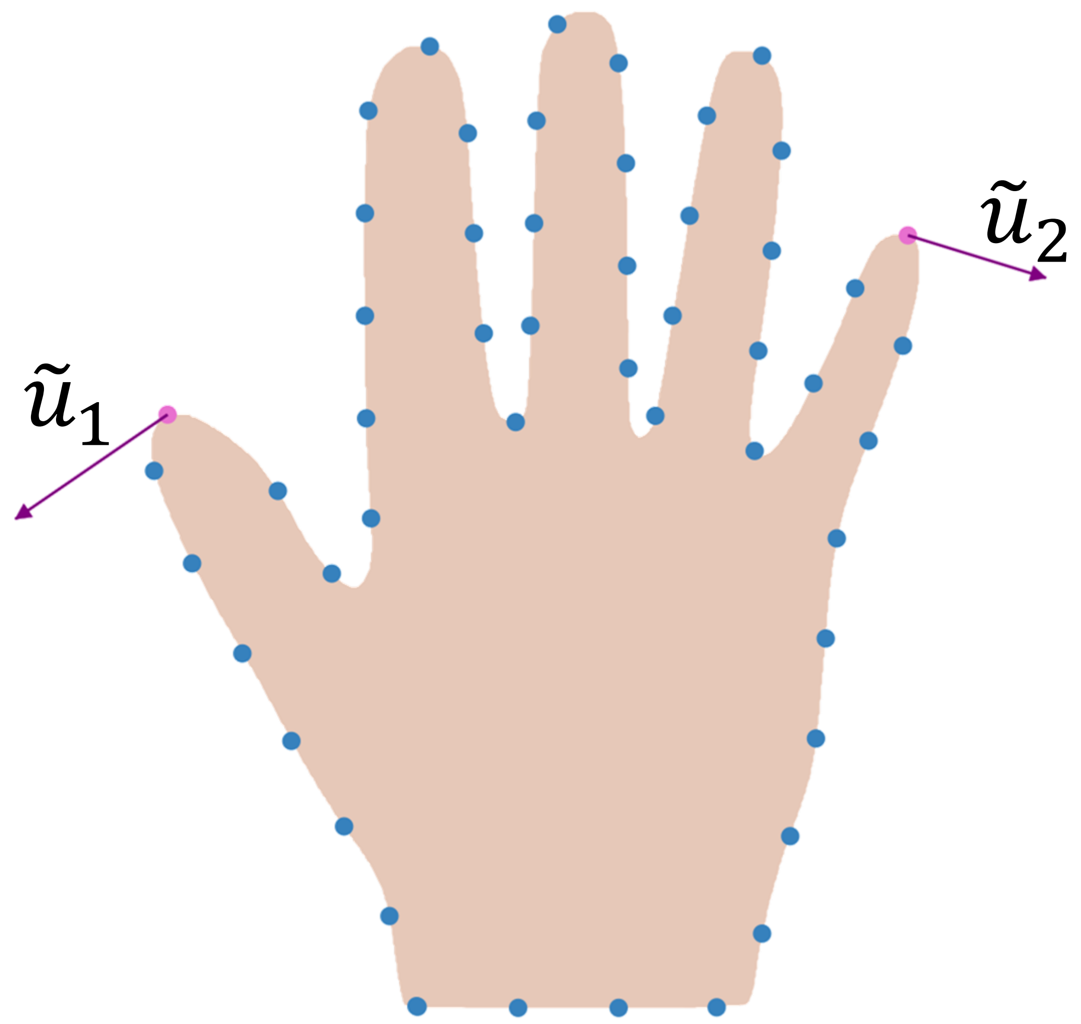


$$\begin{pmatrix} u(x_1) \\ \vdots \\ u(x_n) \\ u(y_1) \\ \vdots \\ u(y_m) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x_1) \\ \vdots \\ \mu(x_n) \\ \mu(y_1) \\ \vdots \\ \mu(y_m) \end{pmatrix}, \begin{pmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) & k(x_1, y_1) & \dots & k(x_1, y_m) \\ \vdots & & & & & \\ k(x_n, x_1) & & k(x_n, x_n) & k(x_n, y_1) & & k(x_m, y_m) \\ k(y_1, x_1) & & k(y_1, x_n) & k(y_1, y_1) & & k(y_1, y_m) \\ \vdots & & & & & \\ k(y_m, x_1) & & k(y_m, x_n) & k(y_m, y_1) & & k(y_m, y_m) \end{pmatrix} \right)$$

Simplified notation:

$$\begin{pmatrix} u(X) \\ u(Y) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(X) \\ \mu(Y) \end{pmatrix}, \begin{pmatrix} K(X, X) & K(X, Y) \\ K(Y, X) & K(Y, Y) \end{pmatrix} \right)$$

The conditional distribution



For given observations:

$$\{(y_i, \tilde{u}_i), i = 1, \dots, m\}$$

we assume that

$$u(y_i) = \tilde{u}_i.$$

Conditional distribution:

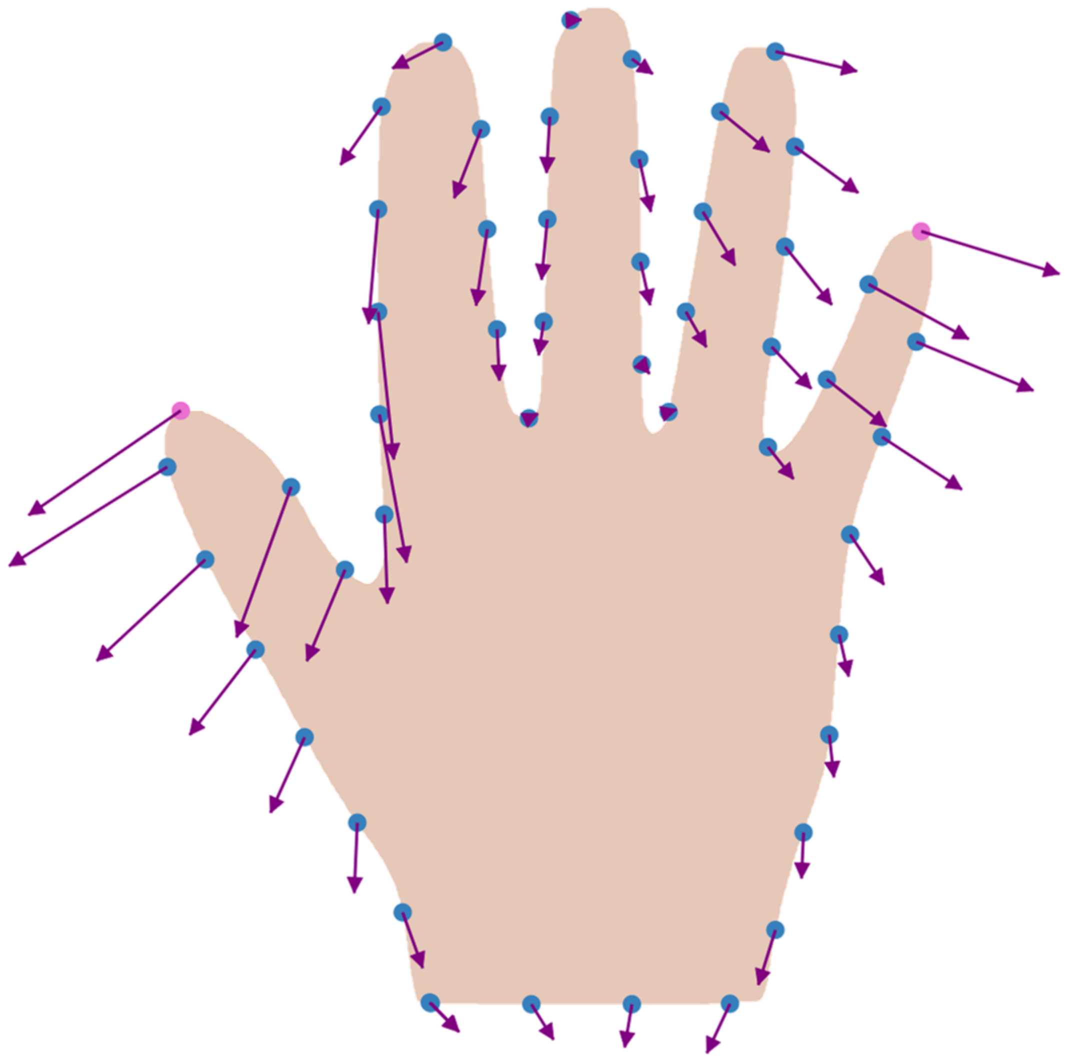
$$p(u(X)|u(Y) = \tilde{\boldsymbol{u}}) = N(\overline{\mu}, \overline{\Sigma})$$

$$\overline{\mu} = \mu(X) + K(X, Y)\mathrm{K}^{-1}(Y, Y)(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$\overline{\Sigma} = K(X, X) - K(X, Y)K(Y, Y)^{-1}K(Y, X)$$

The conditional distribution

University of Basel



For given observations:

$$\{(y_i, \tilde{u}_i), i = 1, \ldots, m\}$$

we assume that

$$u(y_i) = \tilde{u}_i.$$

Conditional distribution:

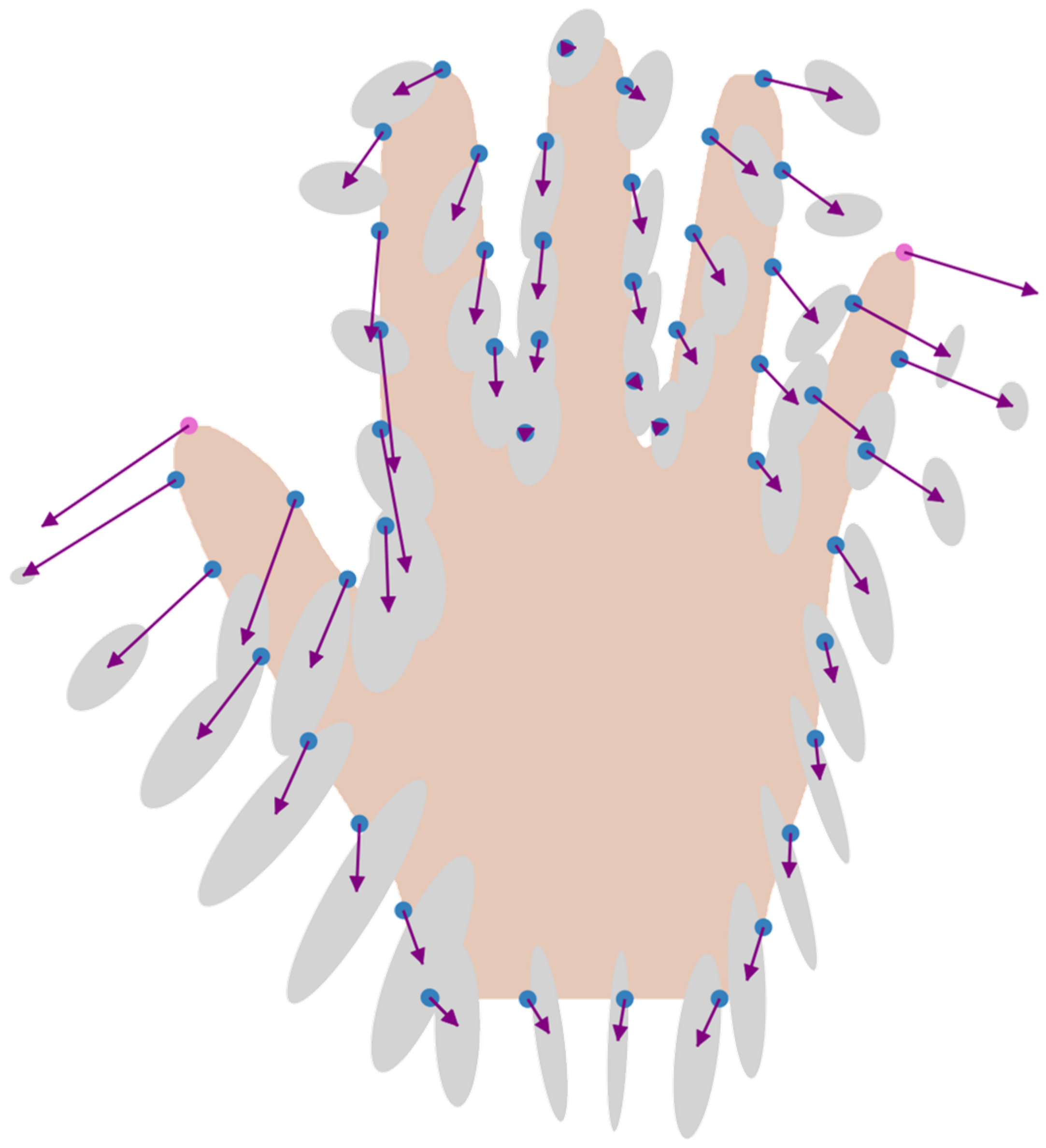
$$p(u(X)|u(Y) = \tilde{\boldsymbol{u}}) = N(\bar{\mu}, \bar{\Sigma})$$

$$\bar{\mu} = \mu(X) + K(X, Y)\mathrm{K}^{-1}(Y, Y)(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$\bar{\Sigma} = K(X, X) - K(X, Y)K(Y, Y)^{-1}K(Y, X)$$

The conditional distribution

University of Basel



For given observations:

$$\{(y_i, \tilde{u}_i), i = 1, \dots, m\}$$

we assume that

$$u(y_i) = \tilde{u}_i.$$

Conditional distribution:

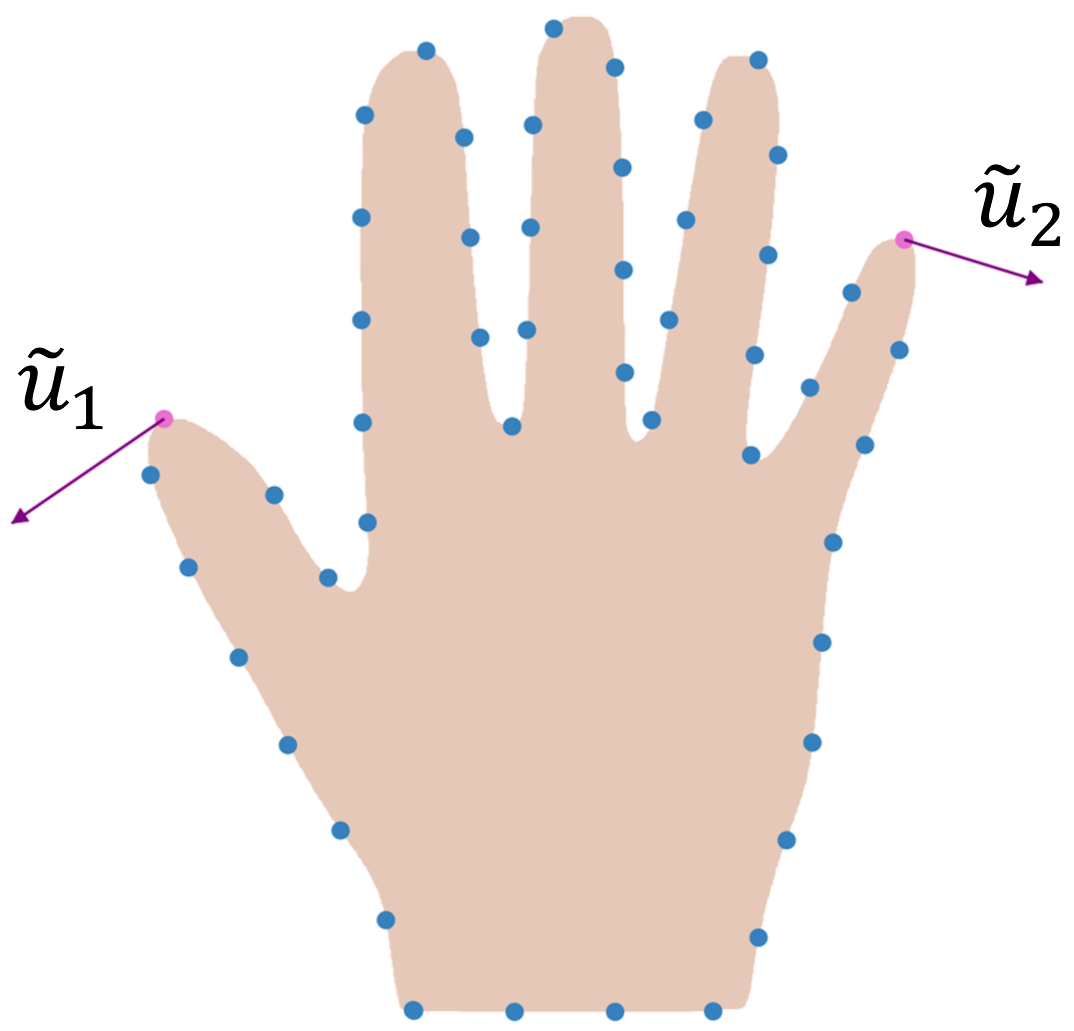
$$p(u(X)|u(Y) = \tilde{\boldsymbol{u}}) = N(\overline{\mu}, \overline{\Sigma})$$

$$\overline{\mu} = \mu(X) + K(X, Y)\mathrm{K}^{-1}(Y, Y)(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$\overline{\Sigma} = K(X, X) - K(X, Y)K(Y, Y)^{-1}K(Y, X)$$

The conditional distribution

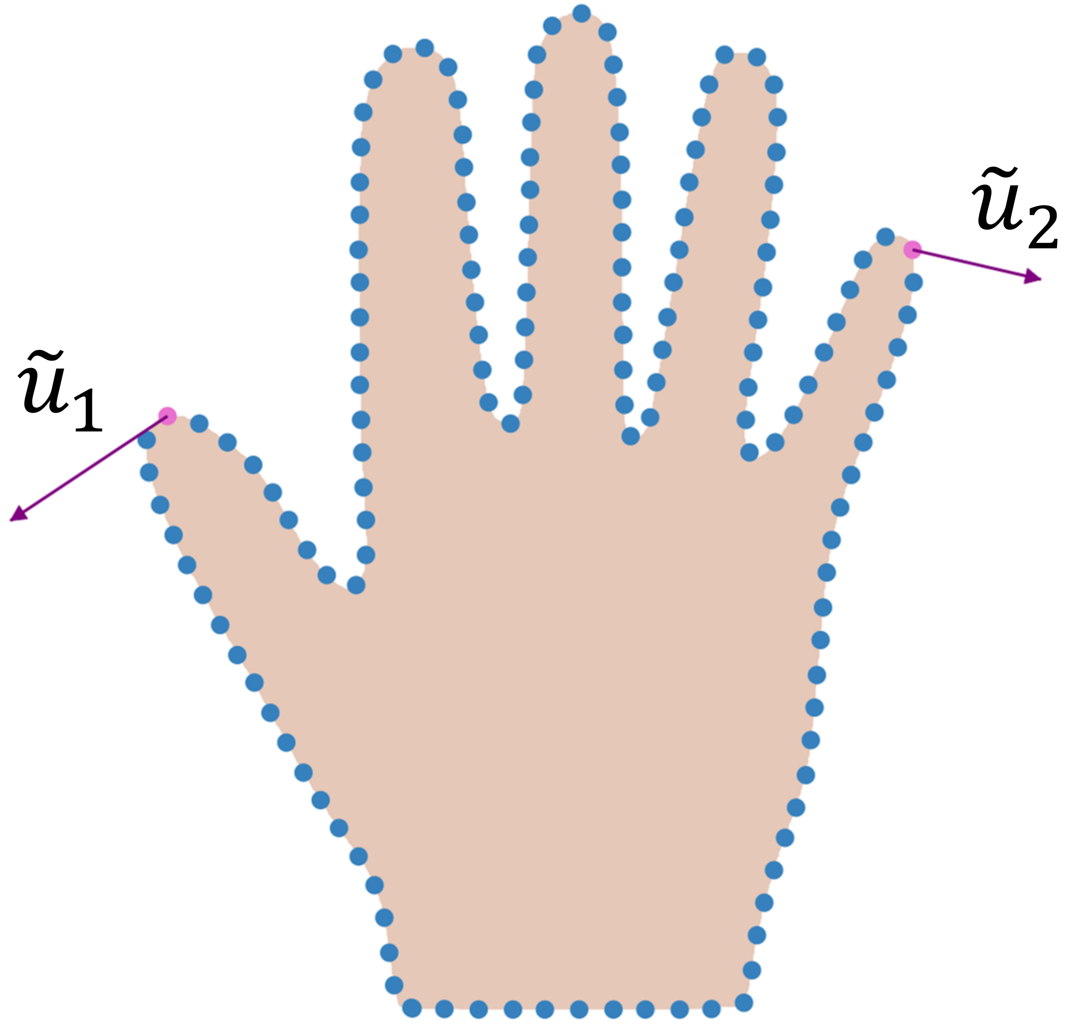
University of Basel



For fixed $u(Y) = \tilde{\boldsymbol{u}}$ the mean and covariance matrix can be computed for an arbitrary number of points $X = \{x_1 \ldots, x_n\}$.

The conditional distribution

University of Basel



For fixed $u(Y) = \tilde{\boldsymbol{u}}$ the mean and covariance matrix can be computed for an arbitrary number of points $X = \{x_1 \ldots, x_n\}$.

Expressions for a single point

$constant \in \mathbb{R}^{2m}$

$$\overline{\mu}(X) = \mu(X) + K(X, Y)K(Y, Y)^{-1}(\tilde{\boldsymbol{u}} - \mu(Y))$$

$constant \in \mathbb{R}^{2m \times 2m}$

Expressions for a single point

$(k(x, y_1) \quad \cdots \quad k(x, y_m)) \in \mathbb{R}^{2\times 2m}$

$constant \ \in \mathbb{R}^{2m}$

$$\mu_p(x) = \mu(x) + K(x, Y)K(Y, Y)^{-1}(\tilde{\boldsymbol{u}} - \mu(Y))$$

$constant \ \in \mathbb{R}^{2m\times 2m}$

Expressions for a single point

University of Basel

$$(k(x,y_1) \quad \cdots \quad k(x,y_m)) \in \mathbb{R}^{2\times 2m}$$

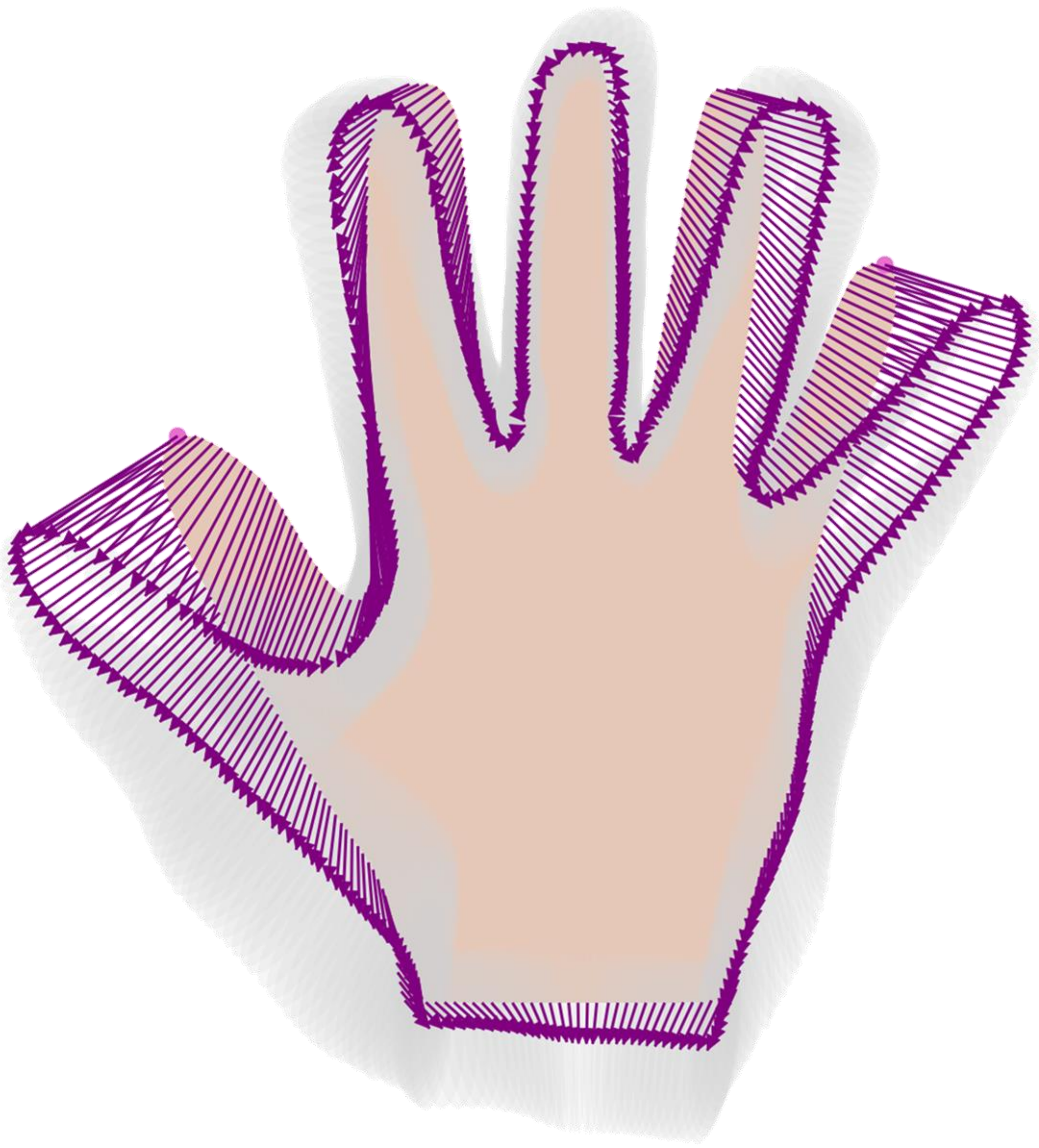
$$\begin{pmatrix} k(y_1,x') \\ \vdots \\ k(y_m,x') \end{pmatrix} \in \mathbb{R}^{2m\times 2}$$

$$k_p(x,x') = k(x,x') - K(x,Y)\,K(Y,Y)^{-1}K(Y,x')$$

$$constant \;\in \mathbb{R}^{2m\times 2m}$$

Posterior model

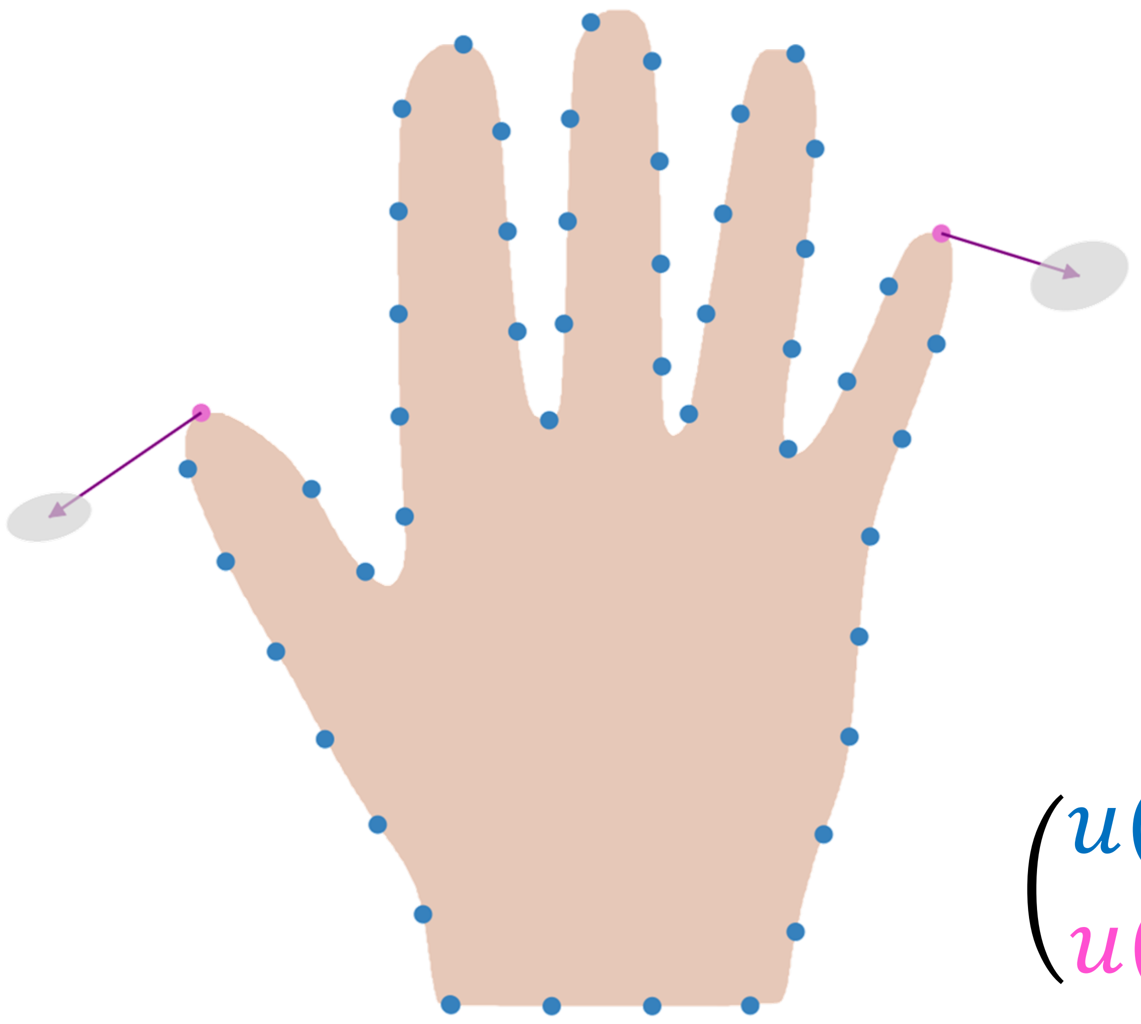
We have defined a Gaussian process $GP(\mu_p, k_p)$



- The process is called the posterior process.
- Defines a distribution of vector-fields, that match the given observations.
- All vector-field match the given observation perfectly.

Noisy observations

University of Basel



For given observations:

$$\{(y_i, \tilde{u}_i), i = 1, \dots, m\}$$

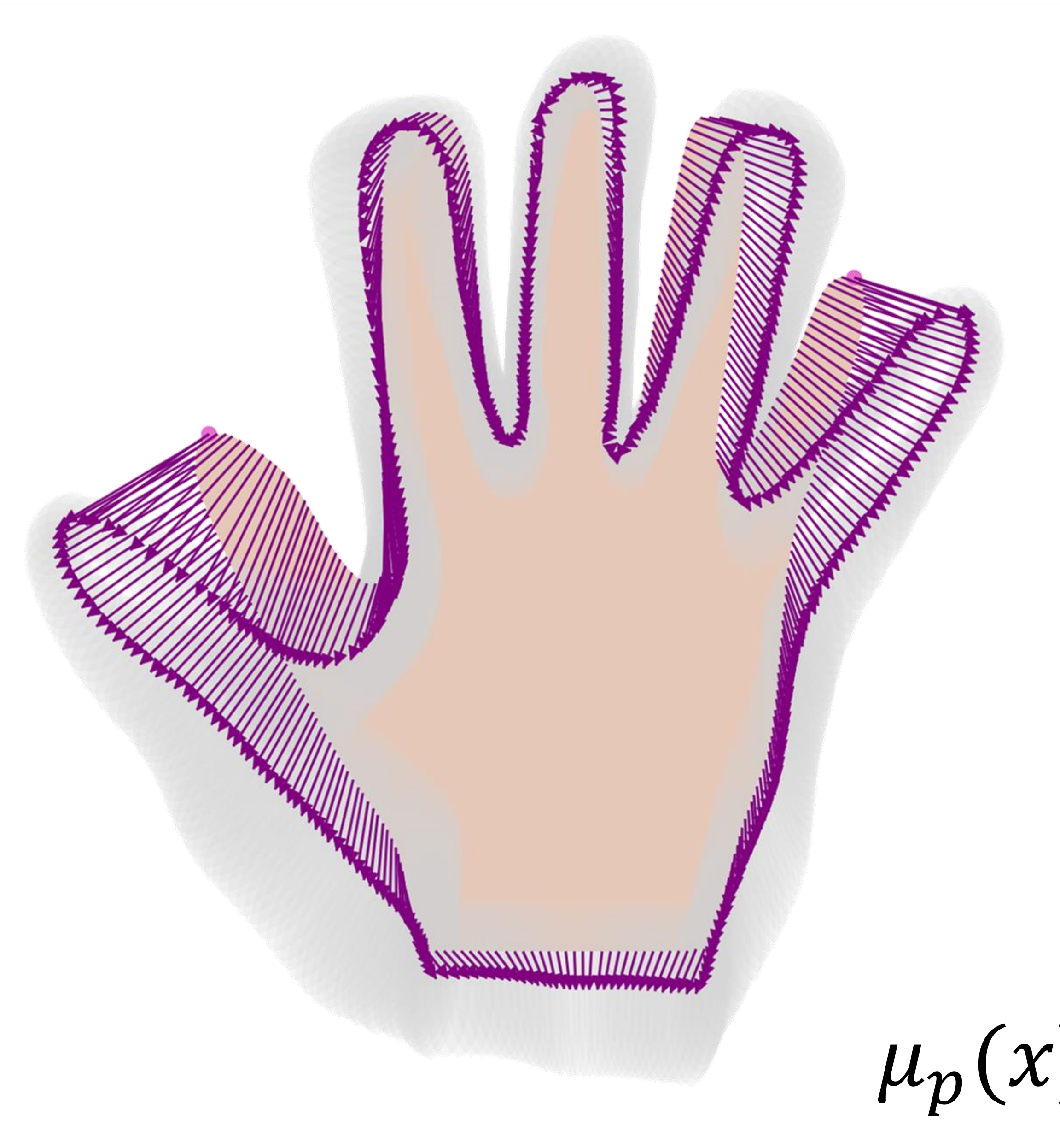
we assume that

$$u(y_i) + \epsilon = \tilde{u}_i.$$

where $\epsilon \sim N(0, \sigma^2 I_{2\times2})$.

$$\begin{pmatrix} u(X) \\ u(Y) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(X) \\ \mu(Y) \end{pmatrix}, \begin{pmatrix} K(X,X) & K(X,Y) \\ K(Y,X) & K(Y,Y) + \sigma^2 I_{2m\times2m} \end{pmatrix} \right)$$

Gaussian process regression



The posterior process $GP(\mu_p, k_p)$ defines a distribution over vector-fields:

- the mean solves the regression problem
- likely deformations agree with observations.
- σ controls accuracy

$$\mu_p(x) = \mu(x) + K(x,Y)(K(Y,Y) + \sigma^2 I_{2m\times 2m})^{-1}(\tilde{\boldsymbol{u}} - \mu(Y))$$

$$k_p(x,x') = k(x,x') - K(x,Y)\,(K(Y,Y) + \sigma^2 I_{2m\times 2m})^{-1} K(Y,x')$$

Posterior shape models