# Probabilistic Shape Modelling
# - Part 2. Fitting probabilistic models -

14. April 2020

Marcel Lüthi

Graphics and Vision Research Group
Department of Mathematics and Computer Science
University of Basel

# Probabilistic Shape Modelling

Online Course / Futurelearn

Shape Modelling

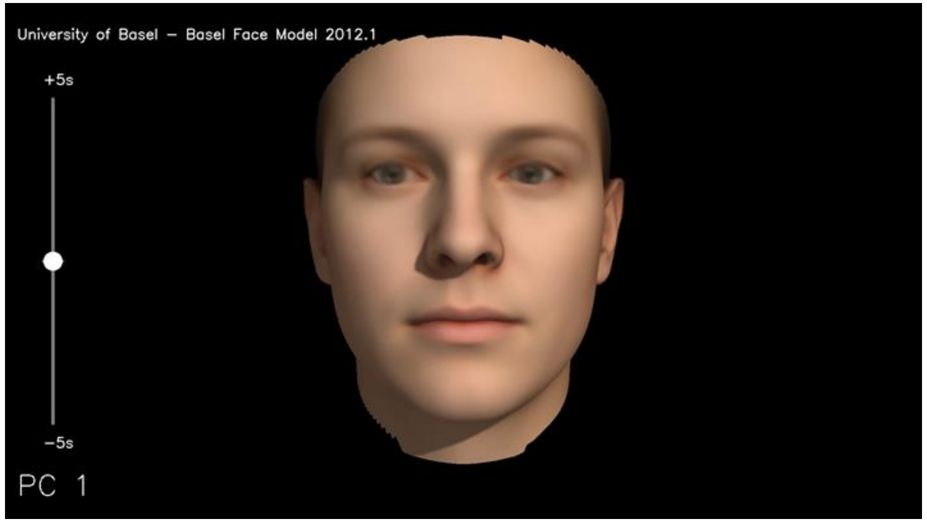
Next lectures

Model fitting

Scalismo

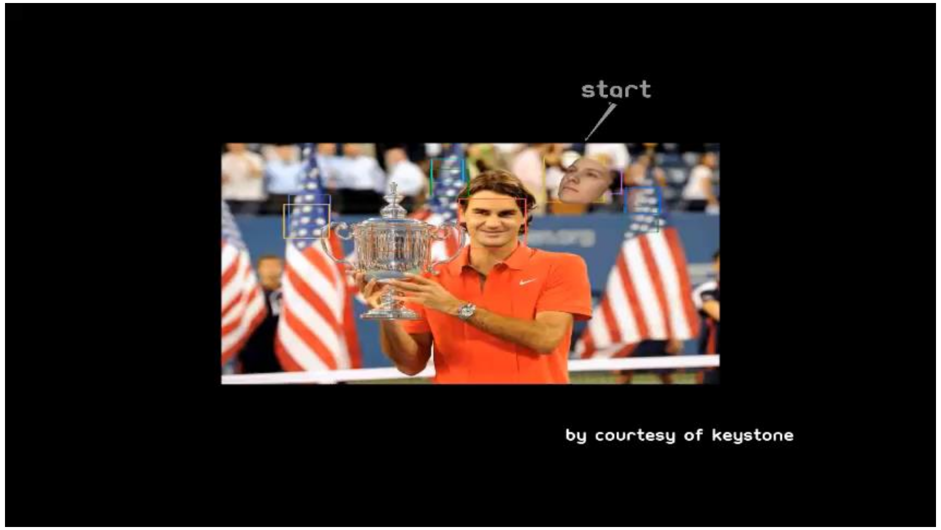
# Probabilistic Shape Modelling

Online Course / Futurelearn




Shape Modelling

Next lectures




Model fitting

Scalismo

# Programme

| | Lecture (14.15 – 16.00) | |
|---|---|---|
| 14. April | • Analysis by Synthesis<br>• Introduction to Bayesian modelling | |
| 21. April | • Markov Chain Monte Carlo – Concepts and main ideas<br>Applications to Shape modelling | • Introduction to exercise 3 and project 2<br>• Working on exercise sheet 3 |
| 28. April | • MCMC: Filtering, diagnostics and logging<br>• Likelihood Functions for shape and image analysis | • Working on exercise sheet 3 |
| 5. Mai | • Metropolis – Hastings. Why does it work? | • Discussion: Exercise sheet 3 |
| 12. Mai | • Face Image Analysis | • Working on Project 2 |
| 19. Mai | • Gaussian processes<br>More insights / connections to other methods | • Working on Project 2 |
| 26. Mai | • Summary | |

# Administrative issues

## Exam

- Will be changed to oral exam due to Covid-19
- Date remains the same (2. Juli 2020)

## Project 2

- You may regroup if you ended up alone or unhappy in a group
- Project introduction: 21. April

## Lectures

- Lectures on Tuesdays, 14:15 – 16:00
- Exercises, questions and discussions, Tuesday’s 16:15-18:00

# Outline

## Analysis by synthesis – Main ideas

- The conceptual framework we follow in this course

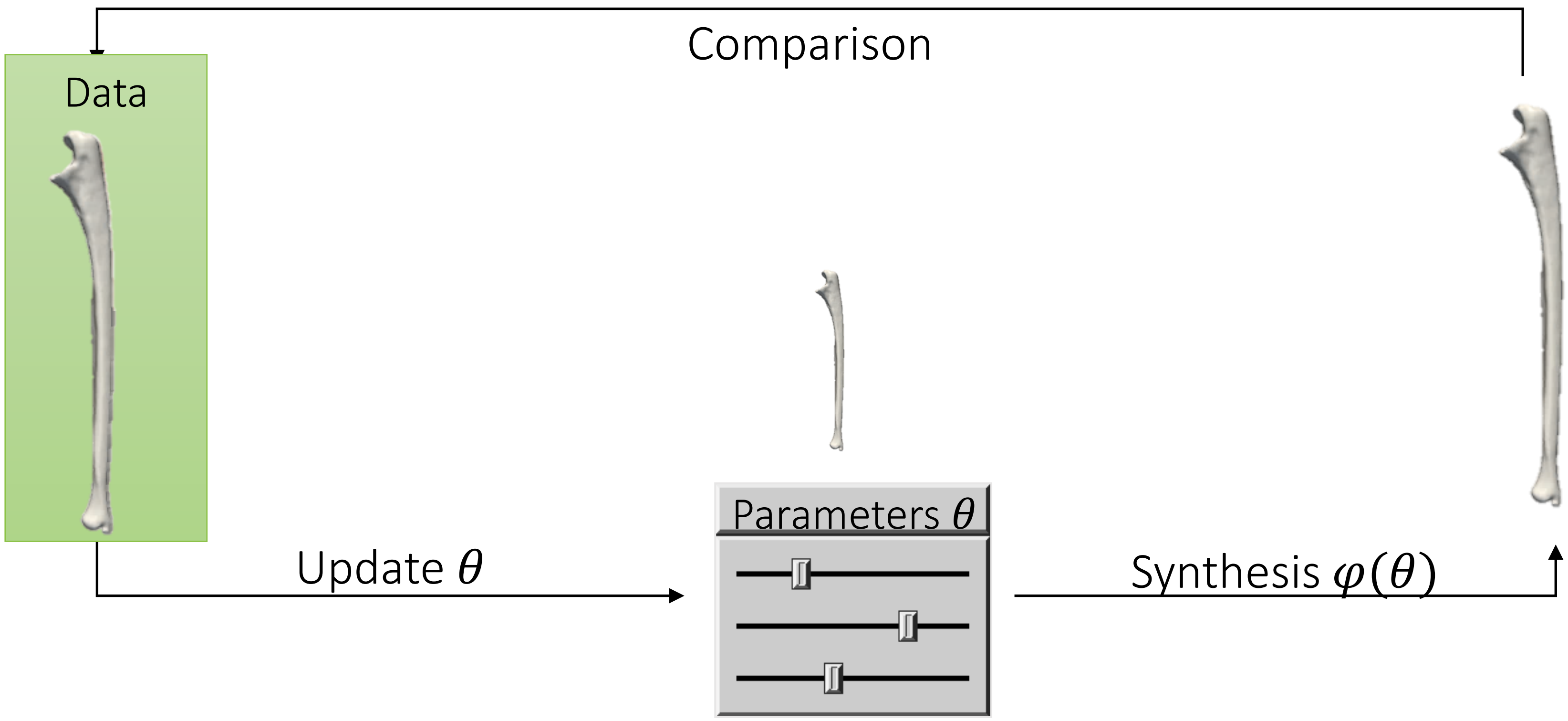
## Bayesian inference

- How we reason in this course

## Analysis by Synthesis in 5 (simple) steps

- A step by step guide to image analysis

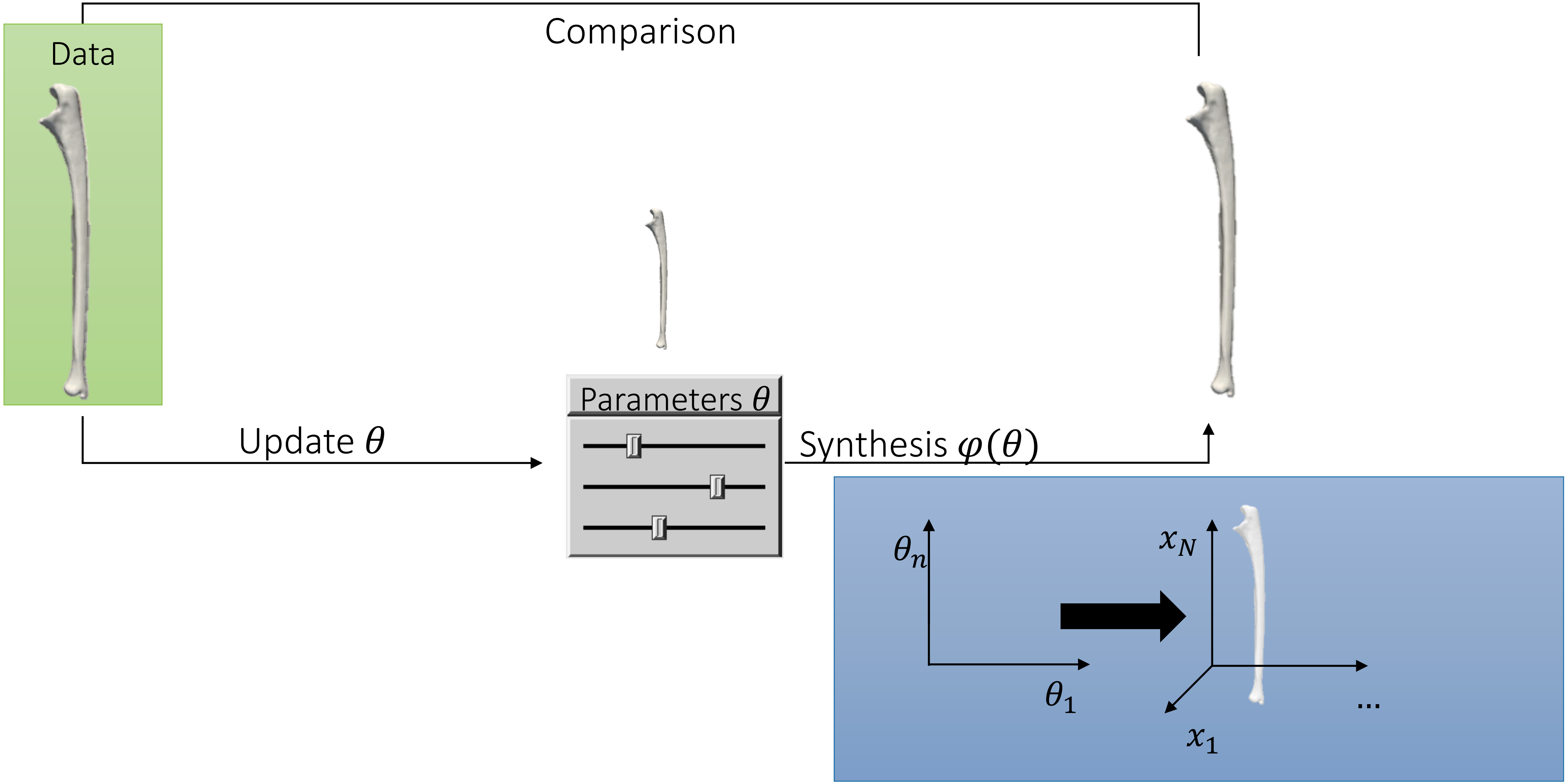
# Conceptual Basis: Analysis by synthesis



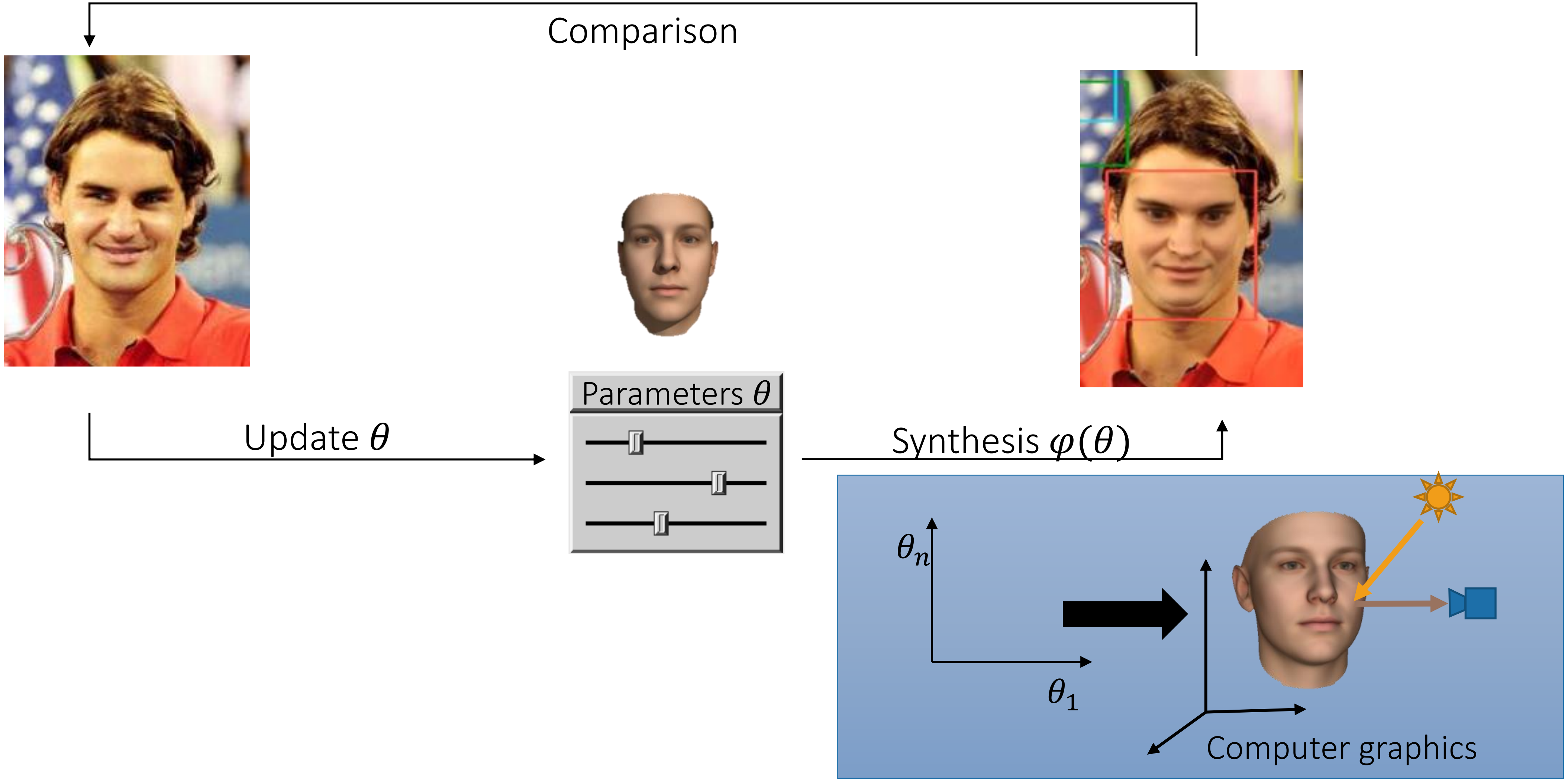

We analyze *our* world by synthesizing relevant aspects of it using *our* model

- Once synthesis produces observed data, we have an explanation of the data
- Allows reasoning about unseen parts

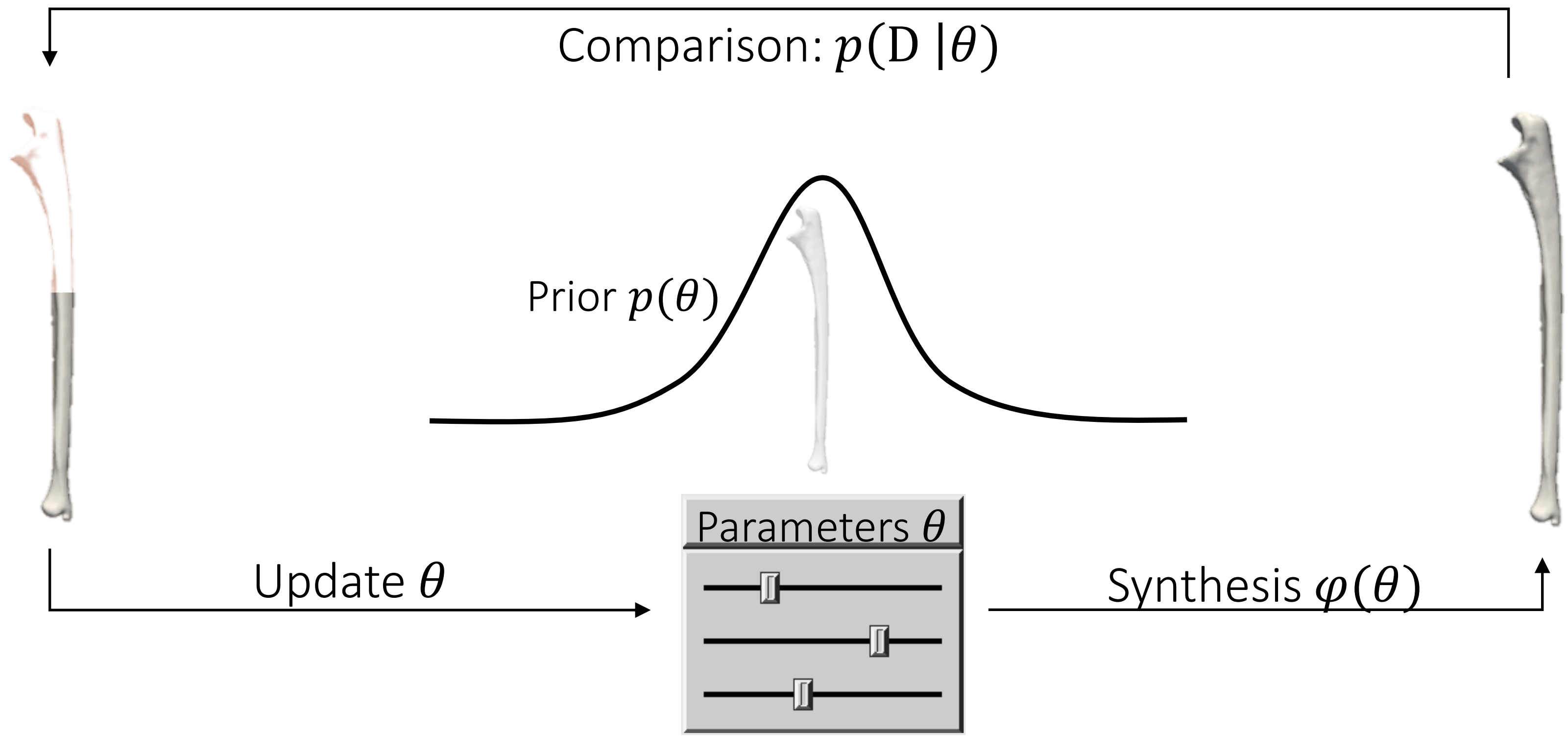
# Conceptual Basis: Analysis by synthesis



# Conceptual Basis: Analysis by synthesis

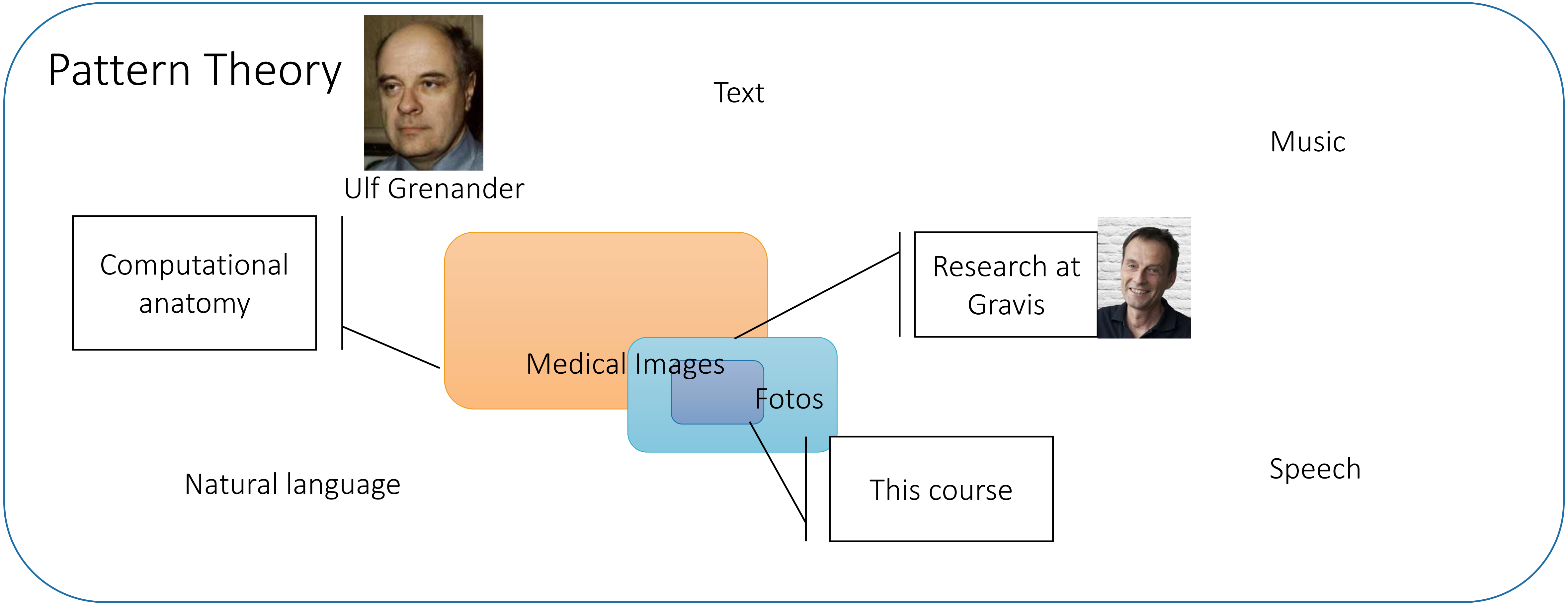


# Mathematical Framework: Bayesian inference

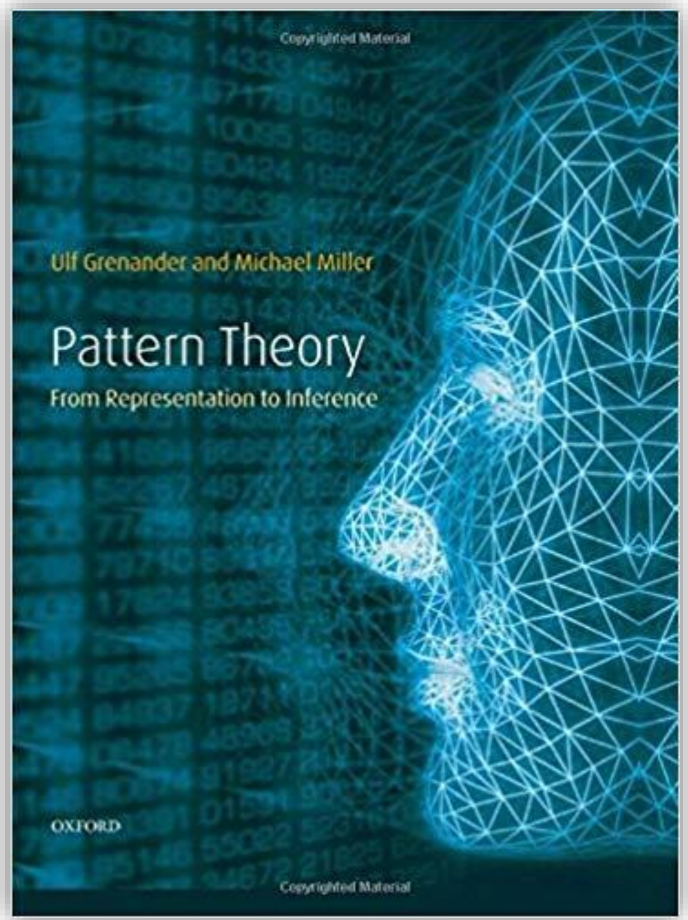


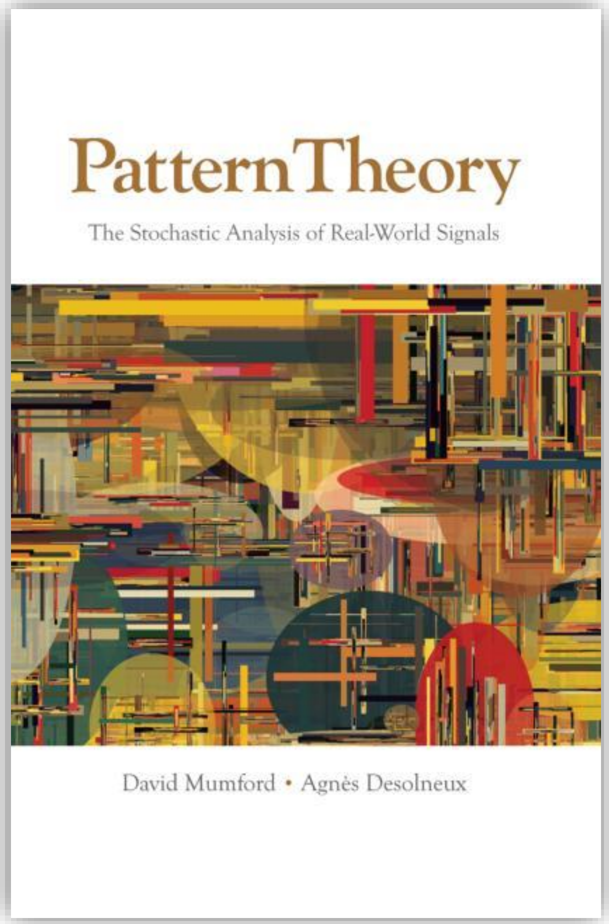

Principled way of dealing with uncertainty.

# The course in context



# Pattern theory – The mathematics



# Bayesian inference

# Probabilities: What are they?

Four possible interpretations:

1. Long-term frequencies
   - Relative frequency of an event over time
2. Physical tendencies (propensities)
   - Arguments about a physical situation (causes of relative frequencies)
3. Degree of belief (Bayesian probabilities)
   - Subjective beliefs about events/hypothesis/facts
4. Logic
   - Degree of logical support for a particular hypothesis

# Degree of belief: An example

Does a dentist's patient have a cavity?

$$P(\text{cavity}) = 0.1$$

$$P(\text{cavity}|\text{toothache}) = 0.8$$

$$P(\text{cavity}|\text{toothache}, \text{gum problems}) = 0.4$$

**Observation: Patient either has a cavity or does not!**

- There is no 80% cavity
- Having a cavity does not depend on whether the patient has a toothache or gum problems
- Does not depend on what the dentist believes

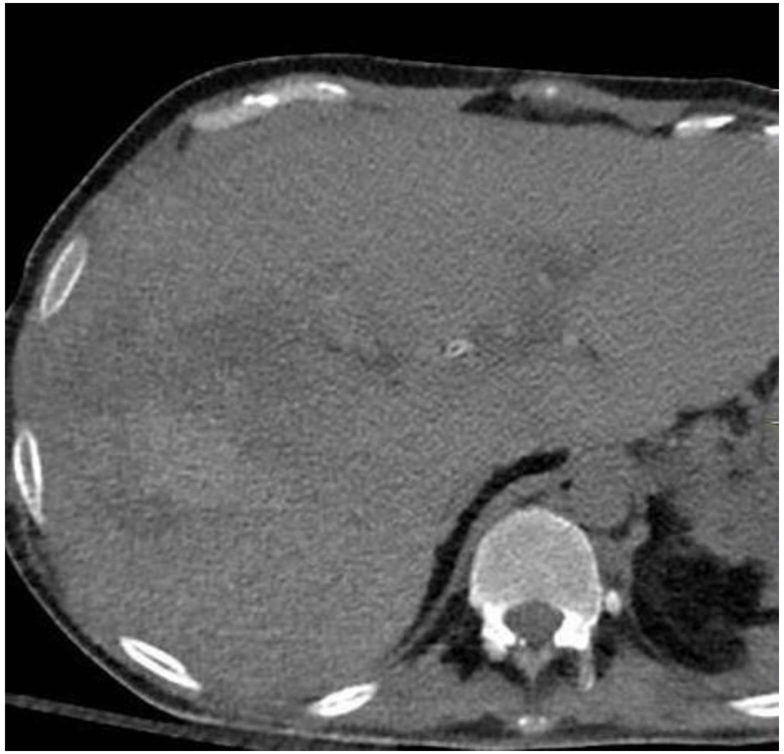
*Statements summarize* ***the dentist's knowledge (model)*** *about the patient*

# Bayesian probabilities for image analysis

*Bayesian probabilities make sense where frequentists interpretations are not applicable!*

- No amount of repetition makes organ boundaries sharper
  - Uncertainty is not due to random effect

- Still possible to use Bayesian inference.
  - Build model of situation
    - Our believe how image was generated
  - Add uncertainty where we are ignorant



# Subjectivity

- Bayesian probabilities rely on a *subjective* perspective:
  - Probabilities express our *current knowledge*.
  - Can *change* when we learn or see more
  - More data -> more *certain* about our result.

*Subjectivity*: There is no single, real underlying distribution. A probability distribution expresses our knowledge – It is different in different situations and for different observers since they have different knowledge.

# Rules for updating beliefs

Given: Joint distribution

$$p_{x,y}(x,y)$$

## Marginal

Distribution of certain points only

$$p_x(x) = \int_y p_{x,y}(x,y)dy$$

## Conditional

Distribution of points conditioned on *known* values of others

$$p_{x|y}(x|y) = \frac{p_{x,y}(x,y)}{p_y(y)}$$

Product rule:

$$p_{x,y}(x,y) = p_{x|y}(x|y)p_y(y)$$

# Bayes rule

From the product rule:

$$p_y(y)p_{x|y}(x|y) = p_{x,y}(x,y) = p_x(x)p_{y|x}(y|x)$$

Bayes rule follows by dividing by $p_y(y)$

$$p_{x|y}(x|y) = \frac{p_x(x)p_{y|x}(y|x)}{p_y(y)}$$

Since $p_y(x) = \int p_{x,y}(x,y)dy = \int p_x(x)p_{y|x}(y|x)dy$ we get

$$p_{x|y}(x|y) = \frac{p_x(x)p_{y|x}(y|x)}{\int p_x(x)p_{y|x}(y|x)dx}$$

# Bayes inference - Terminology

Posterior

Prior

Likelihood

$$p(x|y) = \frac{p(x)p(y|x)}{\int p(x)p(y|x)dx}$$

Marginal Likelihood

# Updating beliefs

Given

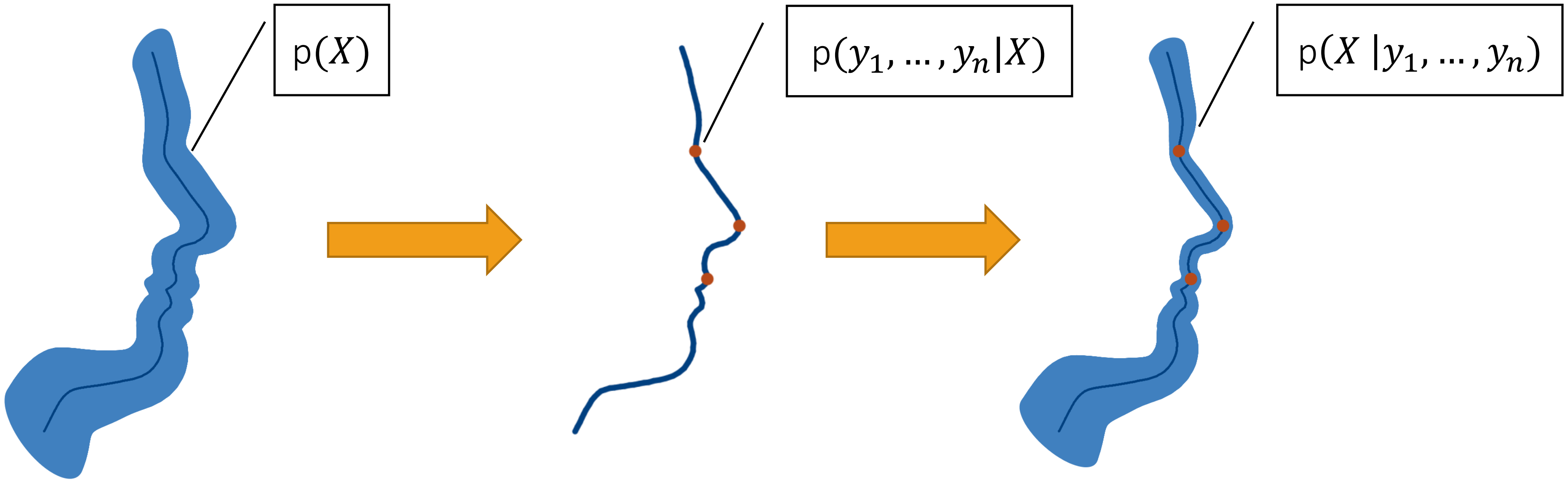
- prior knowledge $p(x)$ - (dentists knowledge about cavities)
- Observation $p(y|x)$ - (probability of toothache given cavity)

We can compute posterior probability: (probability of cavity given toothache)

- $p(x,y) = \frac{p(x)p(y|x)}{\int p(x)p(y|x)dx}$

*Once distributions are fixed, updating beliefs follows laws of probability and is not subjective!*

# Modelling example




**Model**
Shape distribution

**Observation**
Point location (uncertain)

**Posterior**
Shape distribution
*consistent with observation*

Prior belief

More knowledge

Posterior belief

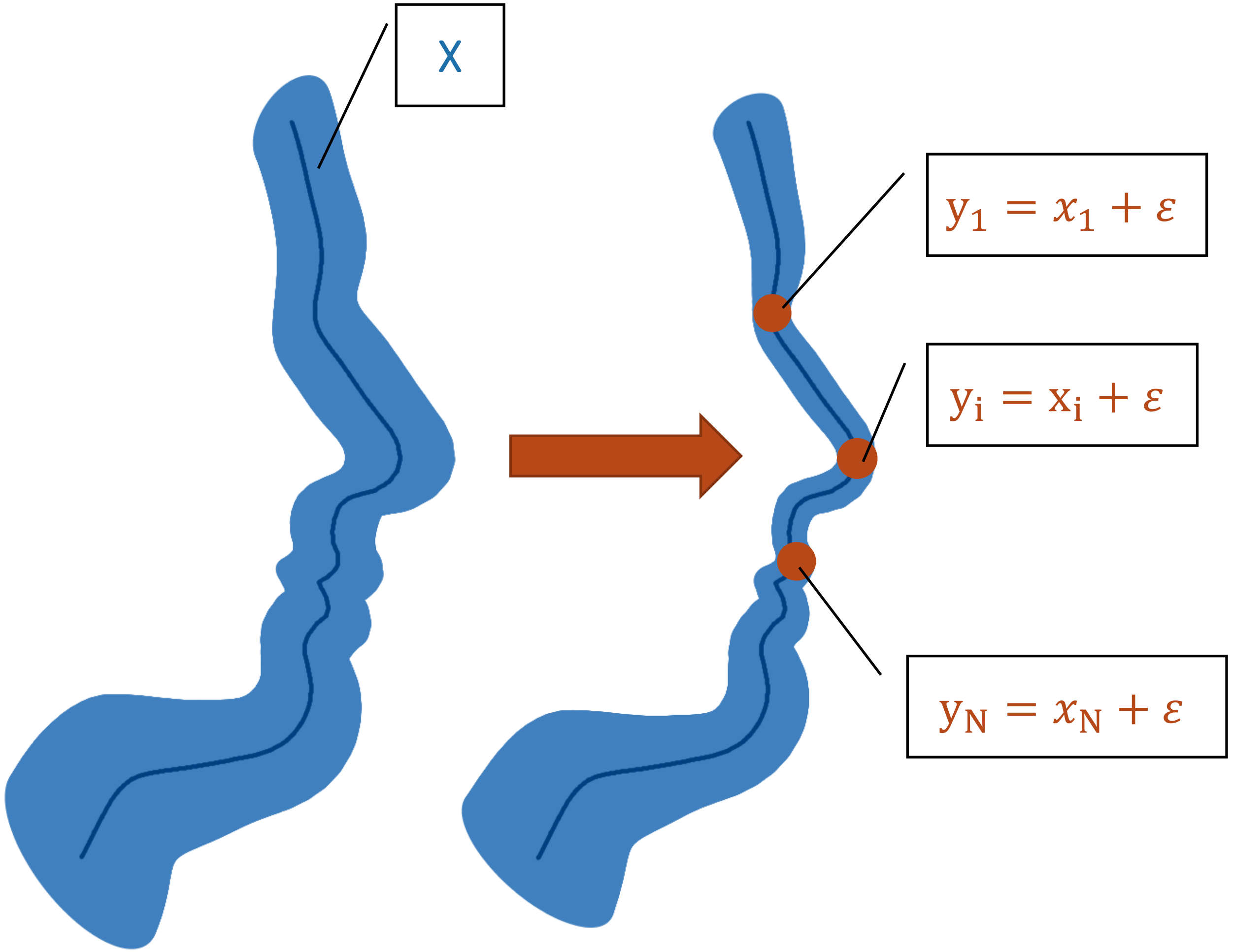
# Belief update

- Observation $y_i$ is noisy measurements of (unobserved) surface point: $y_i = x_i + \epsilon$
- Distribution of $X$ after *observing* $y_1, \dots, y_N$:

$$P(X|y_1 \dots y_N)$$

- Posterior

$$P(X|y_1 \dots y_N) = \frac{P(y_1, \dots, y_N|X)P(X)}{P(y_1, \dots, y_N)}$$



# Belief update (II)

- Each update changes our belief
- Data can be processed sequentially
  - Posterior becomes prior in next step

$$p(X)$$

$$\rightarrow p(X|y_1) = \frac{p(X)p(y_1|X)}{p(y_1)}$$

$$\rightarrow p(X|y_1, y_2) = \frac{p(X)p(y_1|X)p(y_2|y_1, X)}{p(y_1)p(y_2)} = \frac{p(X|y_1)p(X|y_1, y_2)}{p(y_2)}$$

$$\rightarrow \cdots$$

# Joint-Factorisation in Bayesian Inference

Joint Likelihood Prior

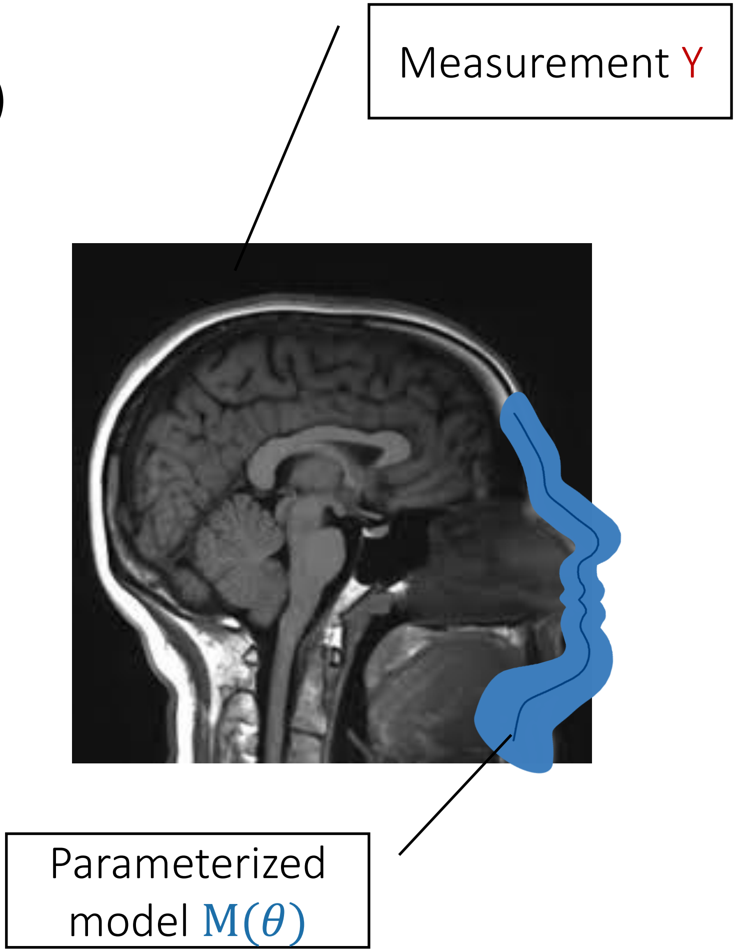
$$P(X,Y) = P(Y|X)P(X)$$

- *Likelihood x prior:* factorization is more flexible than full joint
  - Prior: distribution of core model *without observation*
  - Likelihood: describes how observations are distributed
    - May be related to model variables in very complicated ways

# General Bayesian Inference

- Observation of *additional* variables
  - Common case, e.g. image intensities, surrogate measures (size, sex, …)
  - Coupled to core model via likelihood factorization
- General Bayesian inference case:
  - Distribution of data $Y$
  - Parameters $\theta$

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{P(Y)} = \frac{P(Y|\theta)P(\theta)}{\int P(Y|\theta)P(\theta)d\theta}$$



# Summary: Bayesian Inference

- *Belief*: formal expression of an *observer's knowledge*
  - Subjective state of knowledge about the world
- Beliefs are expressed as *probability* distributions
  - Formally not arbitrary: Consistency requires laws of probability
- *Observations* change knowledge and thus beliefs
- Bayesian inference formally updates *prior beliefs* to *posteriors*
  - Conditional Probability
  - Integration of observation via *likelihood x prior* factorization

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{\int P(\theta)P(Y|\theta)}$$

# Analysis by Synthesis in 5 (simple) steps

# Analysis by synthesis in 5 simple steps

1. Decide which parameters you would like to model
   - Parameters are your representation of the world
   - state of the world is determined by parameters $\theta = (\theta_1, \dots, \theta_n)$

*Everything that is not represented by the parameters cannot be explained by the model*

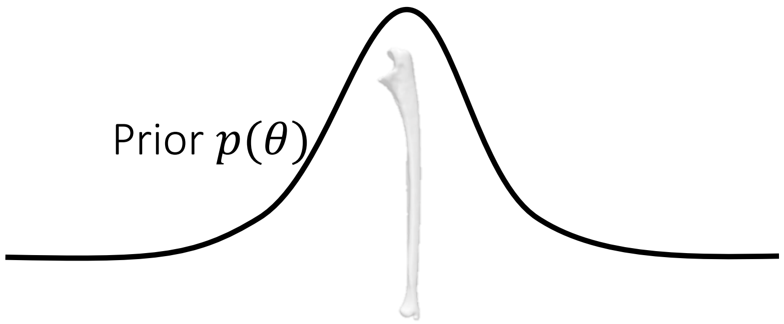
**Shape reconstruction example:**

Parameters: Shape parameters (KL-Expansion coefficients) of GP

# Analysis by synthesis in 5 simple steps

2. Define prior distribution: $p(\theta) = p(\theta_1, \ldots, \theta_n)$

- Our believe about the “state of the world”



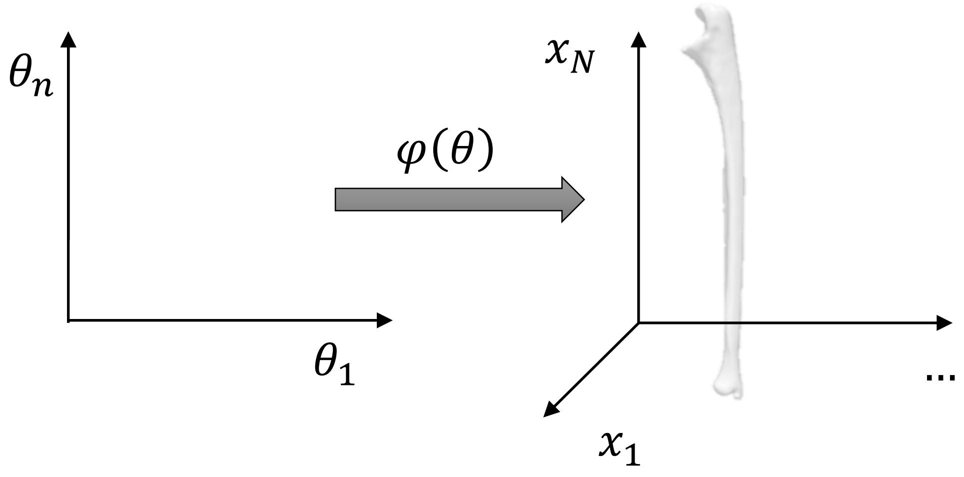

*Subjective and part of our modelling*

**Shape reconstruction example:**

Prior Distribution: Multivariate normal $\theta \sim N(0, I)$

# Analysis by synthesis in 5 simple steps

3. Define a synthesis function $\varphi(\theta)$
   - generates/synthesize the data given the "state of the world"
   - $\varphi$ can be deterministic or stochastic




**Shape reconstruction example:**

Synthesis function:
- Warp of reference surface with deformation vector field $u$ where $\mathrm{u}[\theta](x) = \sum_i \theta_i \lambda_i \phi_i(x)$

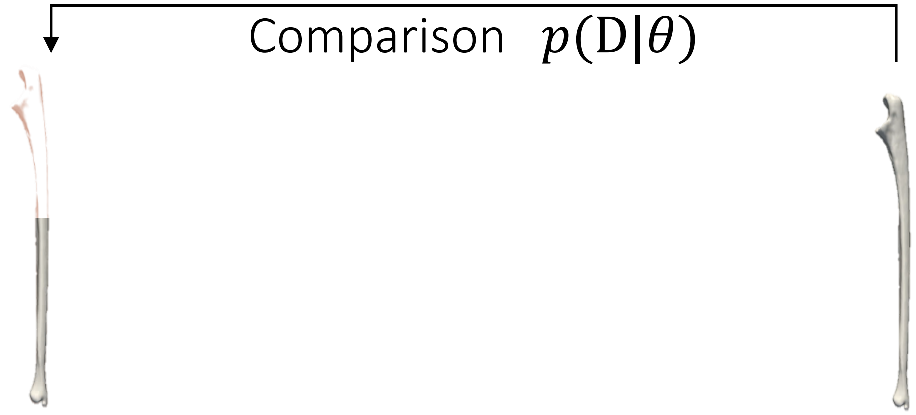
# Analysis by synthesis in 5 simple steps

4. Define likelihood function:

- Define a probabilistic model

$$p(\mathrm{D}|\theta) = \mathrm{p}(\mathrm{D}|\varphi(\theta))$$

  - How likely is $D$ given our synthesized $\varphi(\theta)$
- Includes stochastic factors on the data, such as noise
- Needs to include limitations of model and synthesis function




**Shape reconstruction example:**

Likelihood function for target point position $y(x) \in \Gamma_T \subset \mathbb{R}^3$:

$$p(y(x)|\theta, x) = N(x + u[\theta](x), \sigma^2)$$

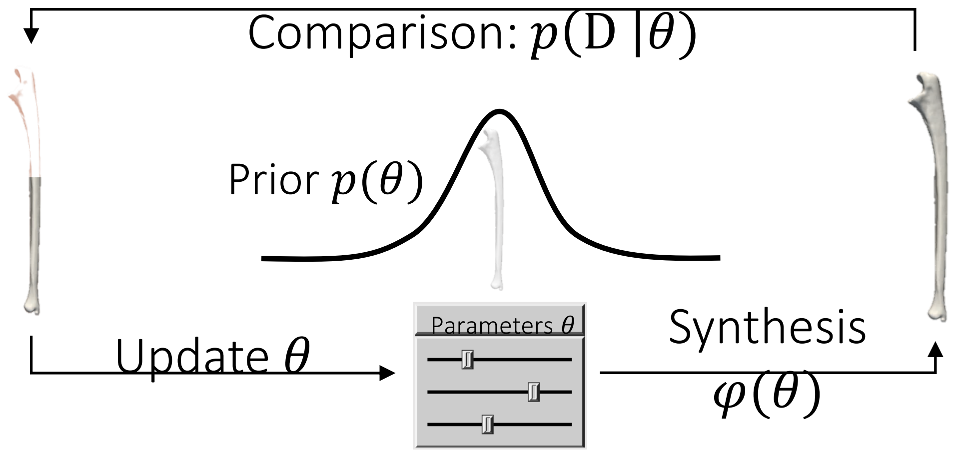
# Analysis by synthesis in 5 simple steps

5. Observe data and update the posterior

$$p(\theta|\mathrm{D}) = \frac{p(\theta)p(\mathrm{D}|\theta)}{\int p(\theta)p(D|\theta)d\theta}$$

**Purely conceptual:**

- Independent of algorithmic implementation

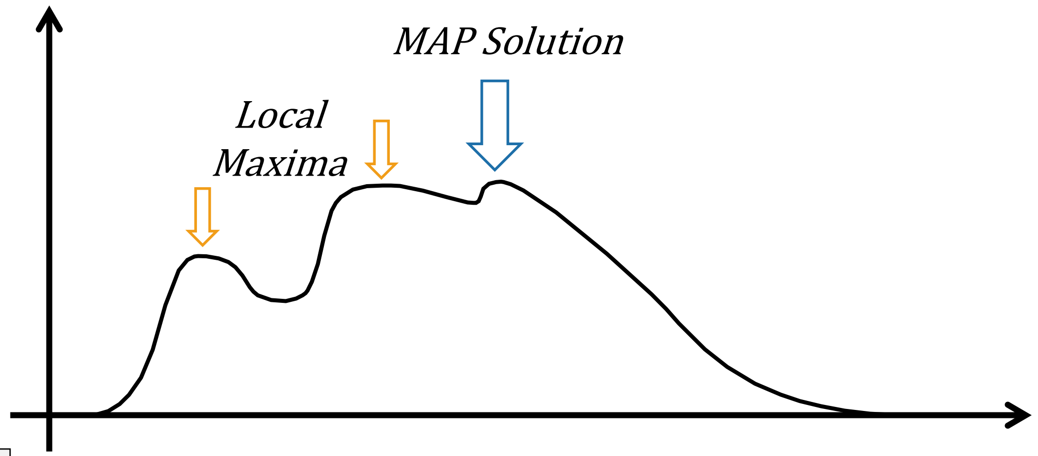


# Analysis by synthesis in 5 simple steps

5a. Implement numerical procedure to do actual inference

**Possibilities**

1. Computing MAP solution
   - No uncertainty – leaves out information
2. Analytic Solution
   - Often not practical
3. Posterior approximation
   - Core of this course




**Shape reconstruction example:**

GP Regression (Analytic posterior)
MAP – Solution (ICP)