# Cyber-Physical Systems

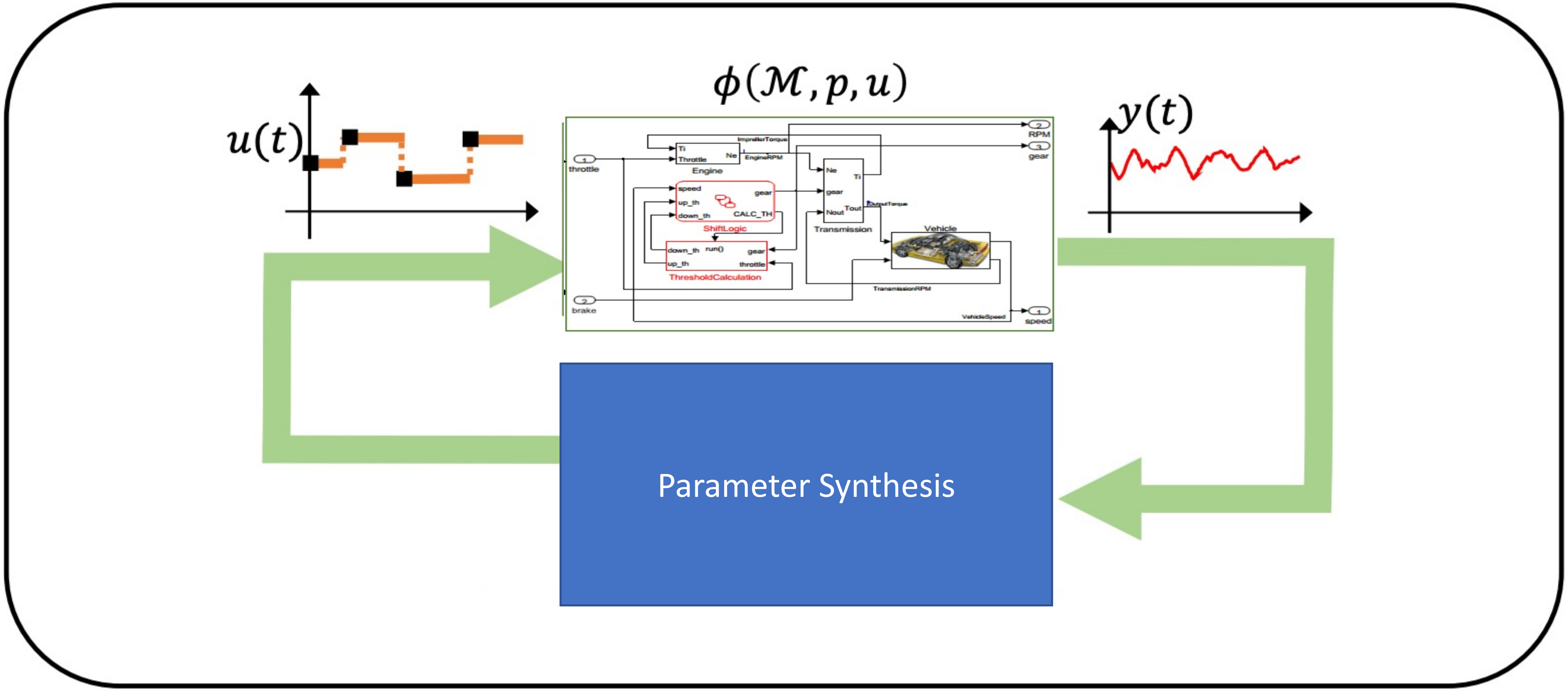
Laura Nenzi

Università degli Studi di Trieste
II Semestre 2020

## Lecture 16: STL applications: parameter synthesis

[Many Slides due to J. Deshmukh, S. Silvetti]

# Parameter Synthesis



# Parameter Synthesis

Problem

Given a model, depending on a set of parameters $\theta \in \Theta$, and a specification $\phi$ (STL formula), find the parameter combination $\theta$ s.t. the system satisfies $\phi$ as more as possible

Solution Strategy

- **rephrase** it as a optimisation problem (maximizing $\rho$)
- **evaluate** the function to optimise
- **solve** the optimisation problem

# Parametric Chemical Reaction Network (PCRN)

Population CTMC models, i.e. CTMC models in the biochemical reactions style.

SET OF SPECIES

$\mathbb{S} = \{S_1, \ldots, S_n\}$, i.e. the different agent states.

STATE SPACE

The state space is described by a vector of $n$ variables

$$\mathbf{X} = (X_{S_1}, \ldots, X_{S_n}) \in \mathbb{N},$$

each counting the number of agents (jobs, molecules, …) of a given kind.

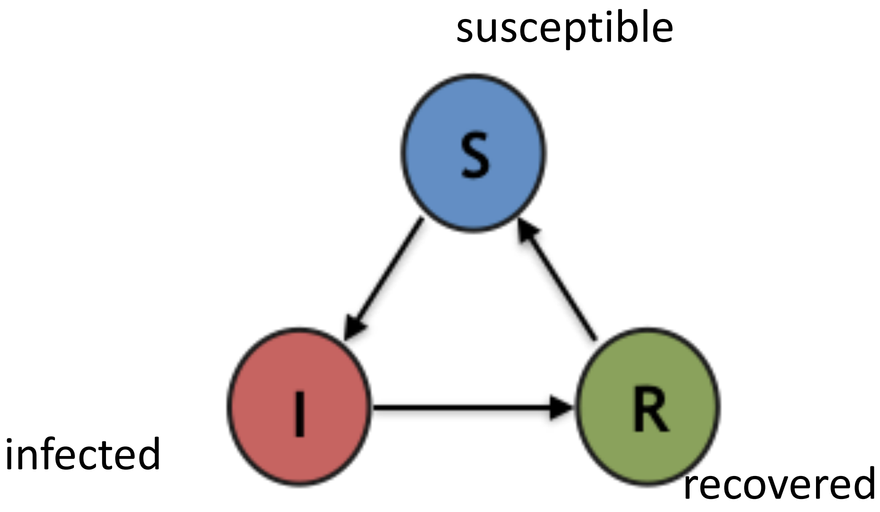
TRANSITIONS

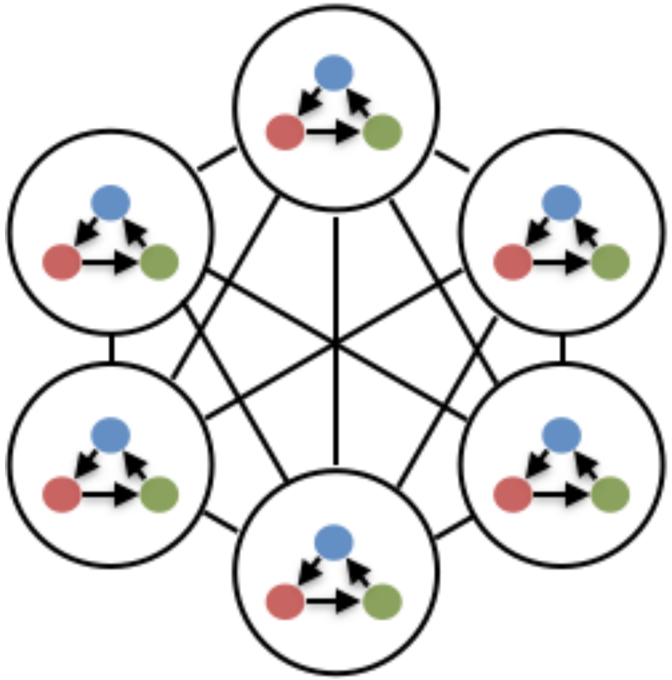
The dynamics is given by a set of chemical reactions:

$$m_1 S_1 + \ldots + m_n S_n \to r_1 S_1 + \ldots + r_n S_n,$$

with a rate given by a function $f(\boldsymbol{X}, \boldsymbol{\theta})$.

# Example: SIR epidemic model



infection: $S + I \rightarrow 2I$ $\qquad f_i(\boldsymbol{X}, \boldsymbol{\theta}) = k_i X_S X_I$

recover: $I \rightarrow R$ $\qquad f_r(\boldsymbol{X}, \boldsymbol{\theta}) = k_r X_I$

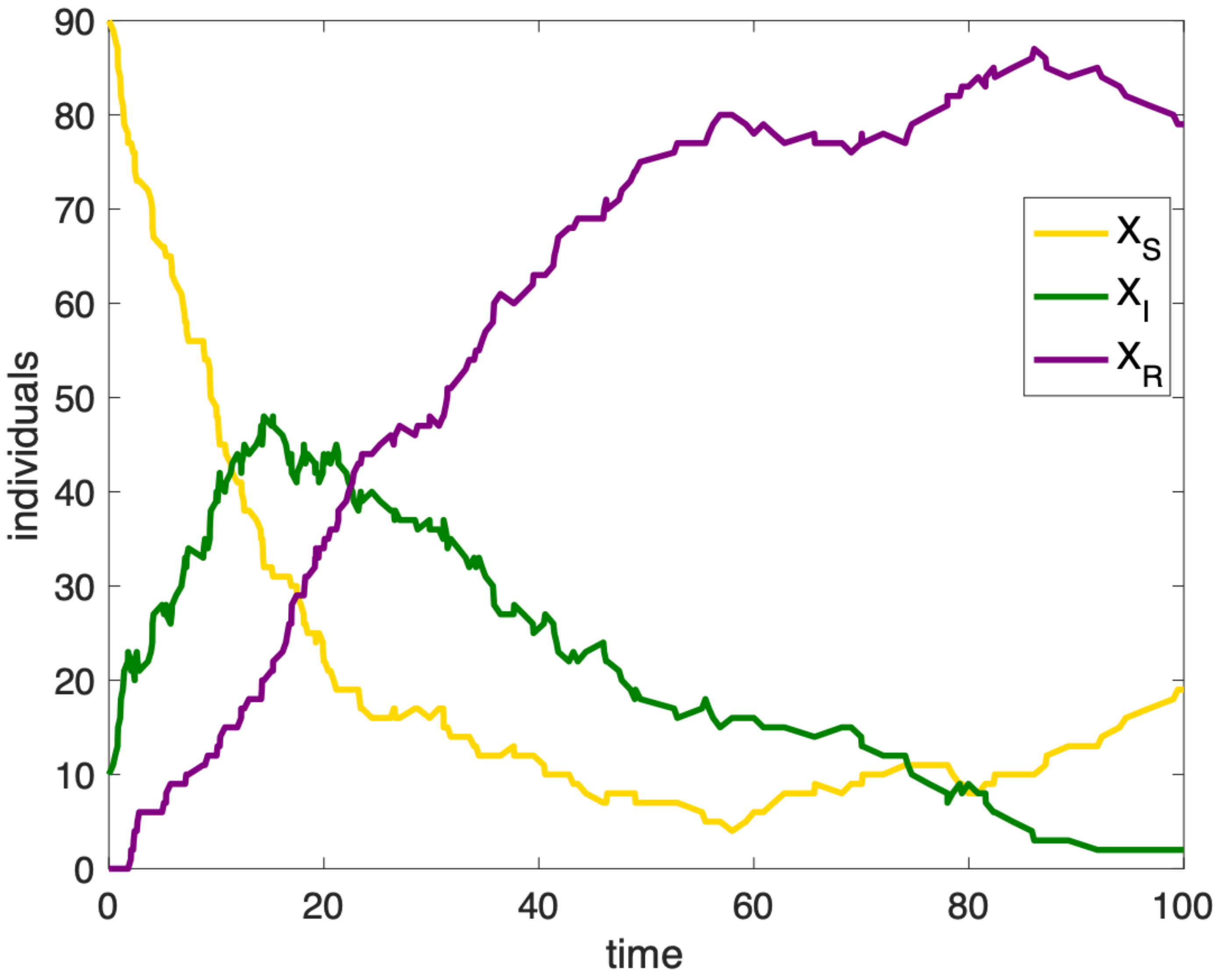
loss of immunity: $R \rightarrow S$ $\qquad f_l(\boldsymbol{X}, \boldsymbol{\theta}) = k_l X_R$

$\mathcal{M}_{\boldsymbol{\theta}}$

State vector: $\mathbf{X} = (X_S, X_I, X_R)$

Vector of parameters: $\boldsymbol{\theta} = (k_i, k_r, k_l)$

# Example: SIRS epidemic model



# Stochastic Semantics

SATISFACTION PROBABILITY (Boolean Semantics)

$$P(\varphi) = \mathbb{P}\{I_\varphi(X) = 1\} := P\{\vec{x} \in Path^{\mathcal{M}} | \mathcal{X}(\vec{x}, 0, \varphi) = 1\}$$

where $I_\varphi(X)$ is a Bernoulli random variable

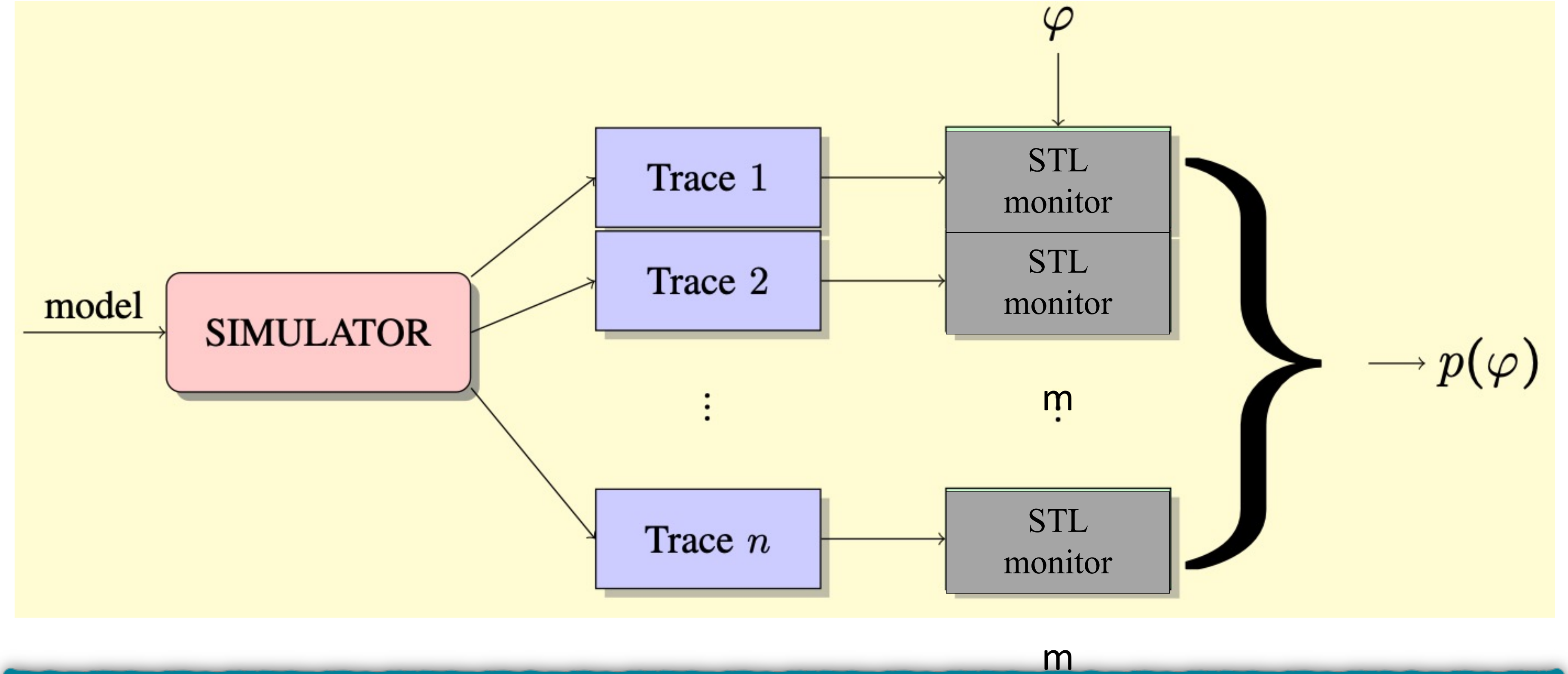
AVERAGE ROBUSTNESS (Quantitative Semantics)

$$\mathbb{P}\{R_\varphi(X) \in [a, b]\} := P\{\vec{x} \in Path^{\mathcal{M}} | \rho(\vec{x}, 0, \varphi) \in [a, b]\}$$

where $R_\varphi(X)$ in a measurable function

# Statistical Model Checking (SMC)

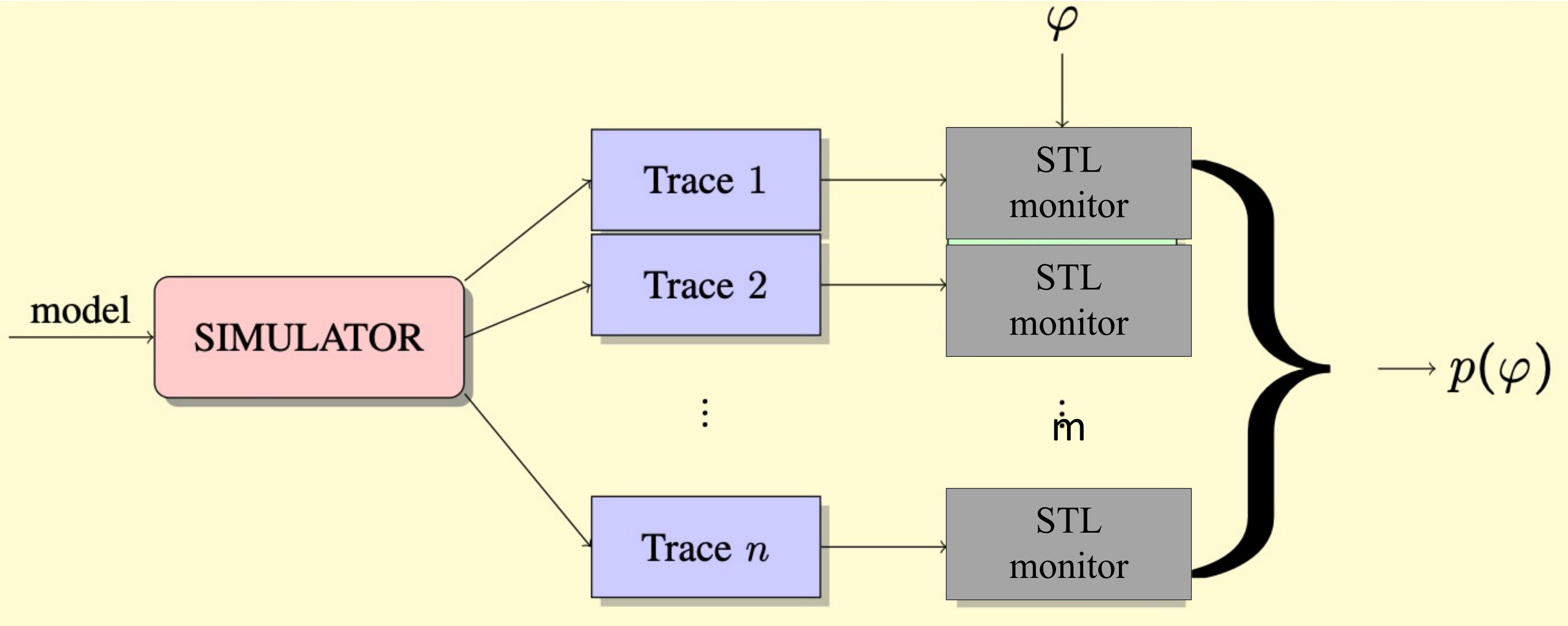
The probability satisfaction can be estimated as an average of the truth values $T_i$ of the formula $\varphi$ over many sample trajectories.




**Bayesian SMC** uses the fact the satisfaction probability of a formula given a model is a number in [0,1], and prior distributions on numbers between [0,1] exist (Beta distribution)}

# Statistical Model Checking

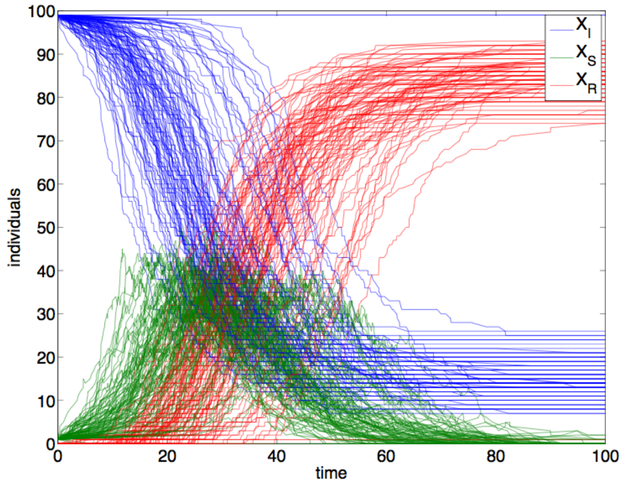
- **Statistical Model Checking:** $p_\phi$ can be estimated as an average of the truth values $T_i$ of the formula $\phi$ over many sample trajectories.
- **Bayesian SMC** specifying (Beta) priors $prob\{p_\phi\}$ and estimating a posteriori $prob\{p_\phi | T_i\}$ using Bayes' theorem and the fact that $prob\{T_i | p_\phi\}$ is Bernoulli.

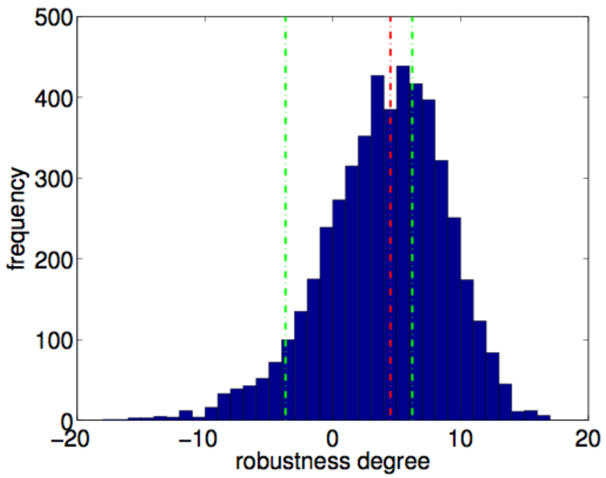


# Parameter Synthesis via Robustness Maximisation

## Robustness Distribution

$$\mathbb{P}\left(R_\varphi(\mathbf{X}) \in [a,b]\right) = \mathbb{P}\left(\mathbf{X} \in \{\mathbf{x} \in \mathcal{D} \mid \rho(\varphi, \mathbf{x}, 0) \in [a,b]\}\right)$$




## Indicators

$\mathbb{E}(R_\varphi)$ (the average robustness degree)

$\mathbb{E}(R_\varphi \mid R_\varphi > 0)$ and $\mathbb{E}(R_\varphi \mid R_\varphi < 0)$ (the conditional averages)

# Parameter Synthesis

Problem

Find the parameter configuration that maximizes $E[R_\phi](\theta)$, of which we have few costly and noisy evaluations.

⬇

Methodology

1. Sample $\{(\theta_{(i)}, y_{(i)}), i = 1,\ldots,n\}$
2. Emulate (**GP Regression**): $E[R_\phi] \sim GP(\mu,k)$
3. Optimize the emulation via **GP-UCB algorithm**, new $\theta_{(n+1)}$

# Gaussian Process Regression

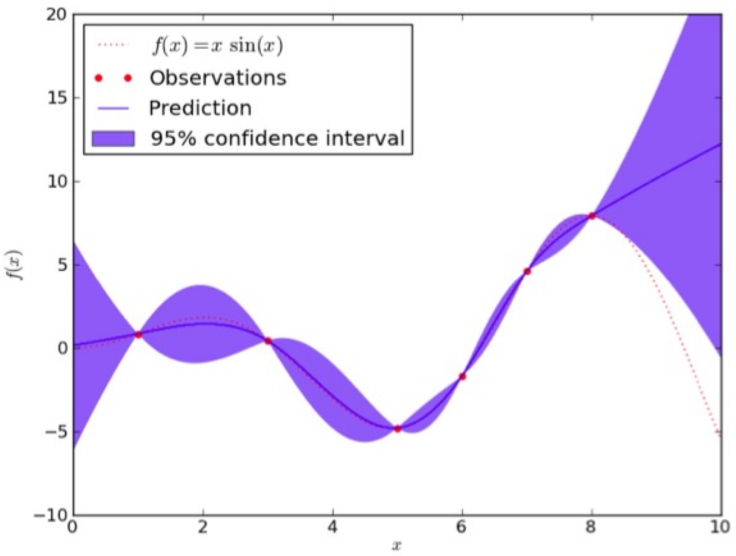
Gaussian Processes can be used for Bayesian prediction and classification tasks.

Idea: put a **GP prior** on functions; condition on **observed data (training set)** $(x_i, y_i)$; we compute a **posterior** distribution on functions; make **predictions**.

**Latent function**: $f$ , GP ; **Noise model**: $p(y_i|f(x_i))$

**Prediction** (latent function $f^*$ at $x^*$)

$$p(f^*|\mathbf{y}) \propto \int df(\mathbf{x})p(f^*, f(\mathbf{x}))p(\mathbf{y}|f(\mathbf{x}))$$




Under Gaussian noise $y(\mathbf{x}) = f(\mathbf{x}) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2)$
predictions have an analytic expression.

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right)$$

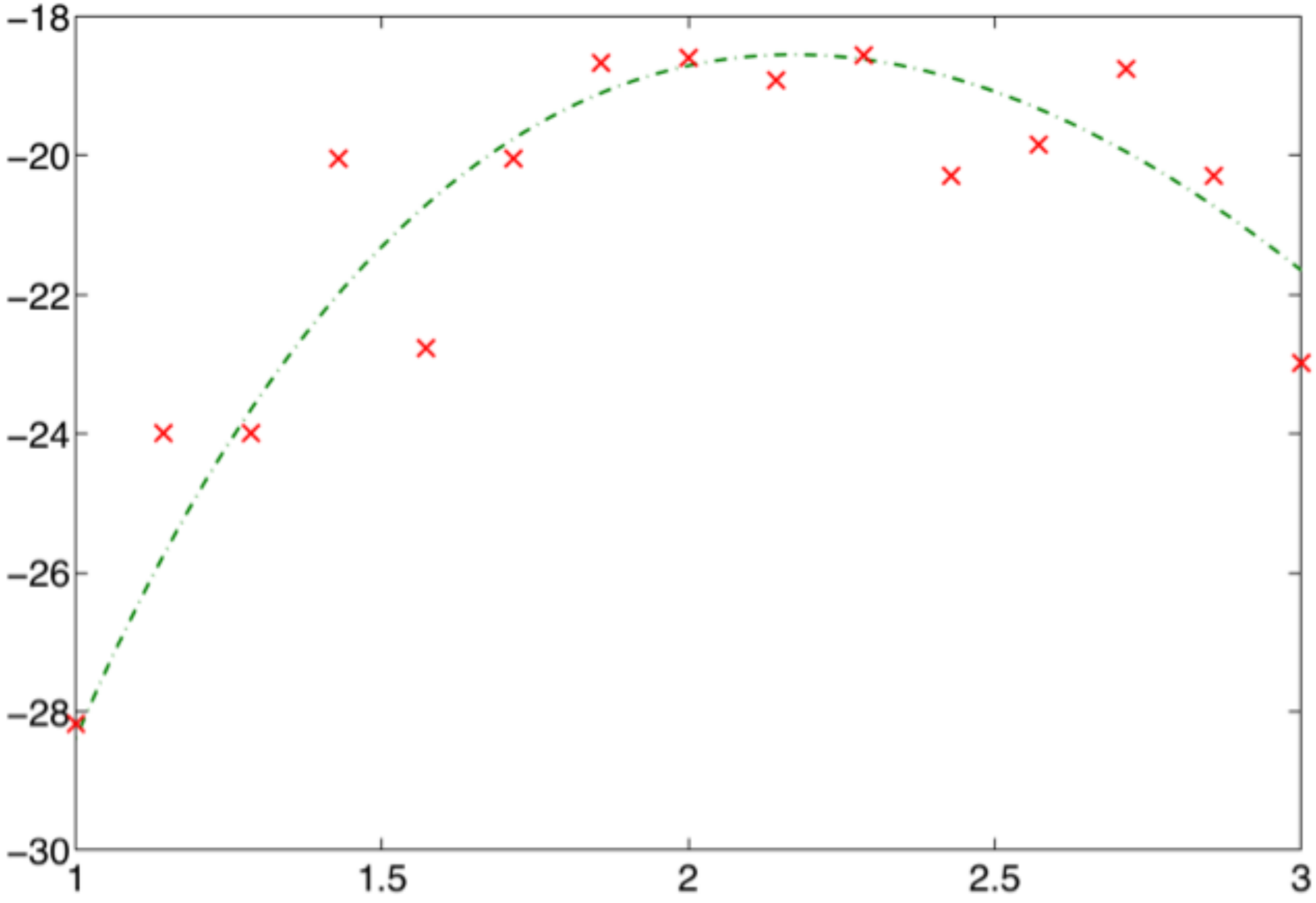
$$\mathbf{f}_*|X, \mathbf{y}, X_* \sim \mathcal{N}(\bar{\mathbf{f}}_*, \operatorname{cov}(\mathbf{f}_*)), \quad \text{where}$$

$$\bar{\mathbf{f}}_* \triangleq \mathbb{E}[\mathbf{f}_*|X, \mathbf{y}, X_*] = K(X_*, X)[K(X,X) + \sigma_n^2 I]^{-1}\mathbf{y},$$

$$\operatorname{cov}(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X,X) + \sigma_n^2 I]^{-1} K(X, X_*)$$

# (1) Sample

Collection of the training set $\{(\theta^{(i)}, y^{(i)}), i = 1,\ldots,m\}$ for parameters values $\theta$.

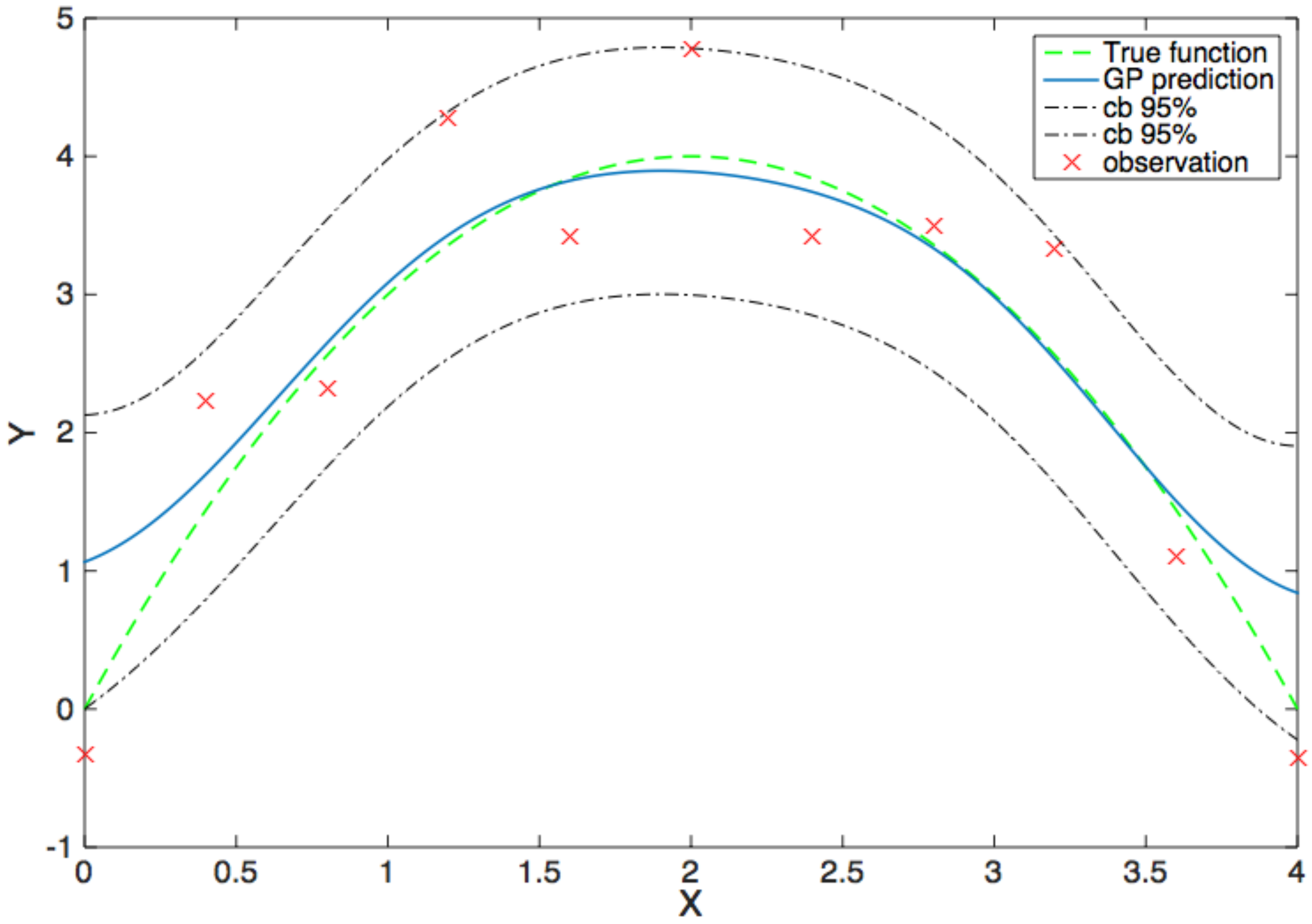


# (2) The GP Regression

We have noisy observations y of the function value distributed around an unknown true value f (θ) with spherical Gaussian noise

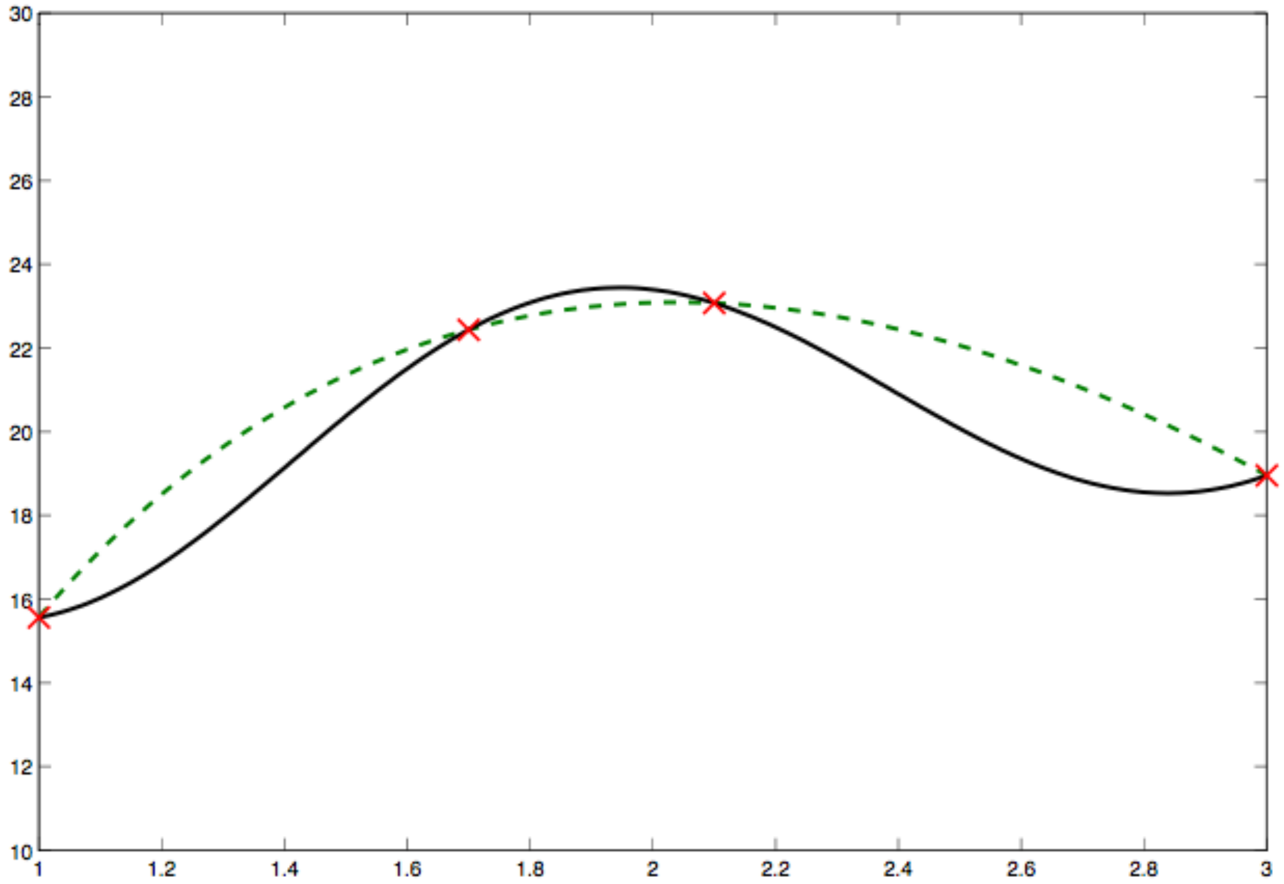
# (2) The GP Regression

We have noisy observations y of the function value distributed around an unknown true value f (θ) with spherical Gaussian noise



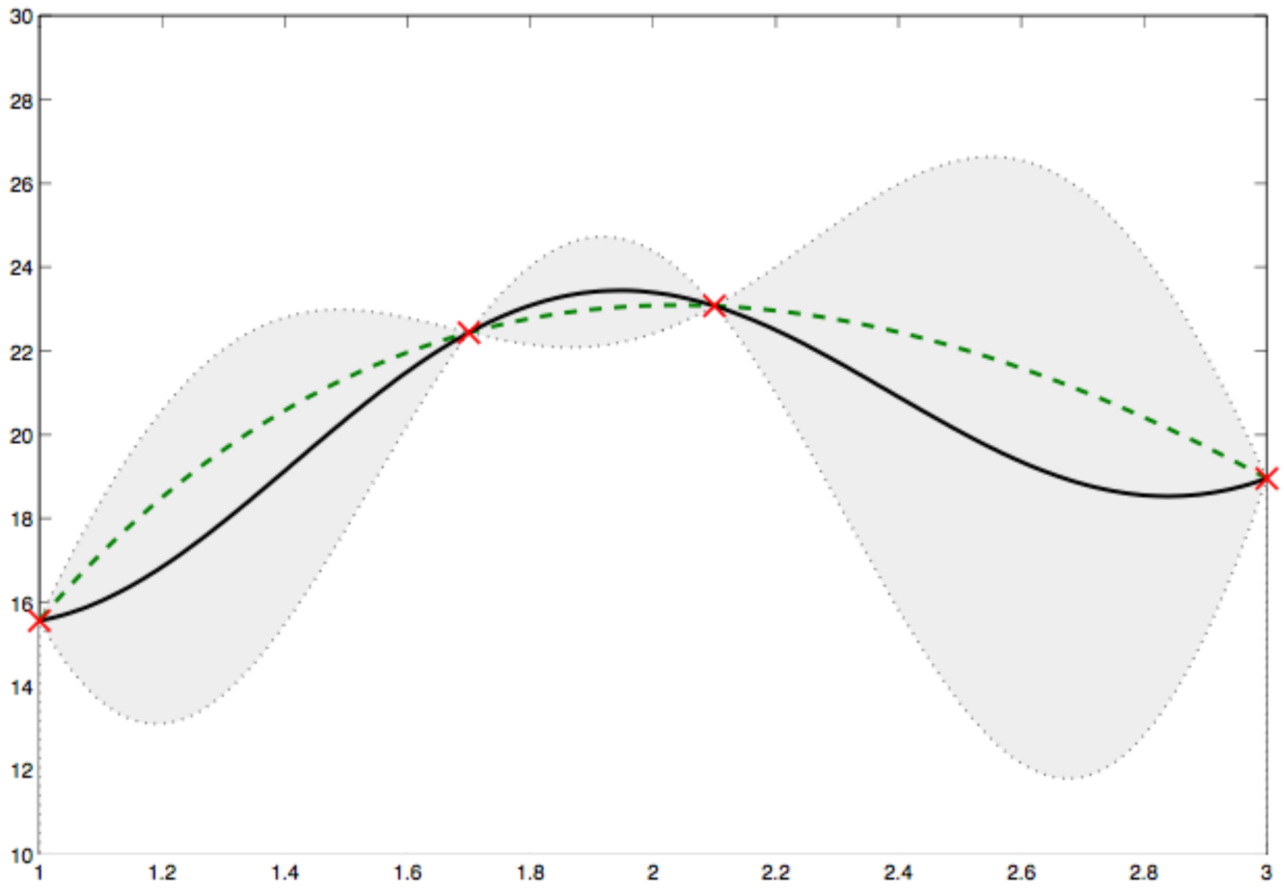
# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_{\theta}[\mu^*(\theta) + \beta_t \sqrt{k^*(\theta,\theta)}]$



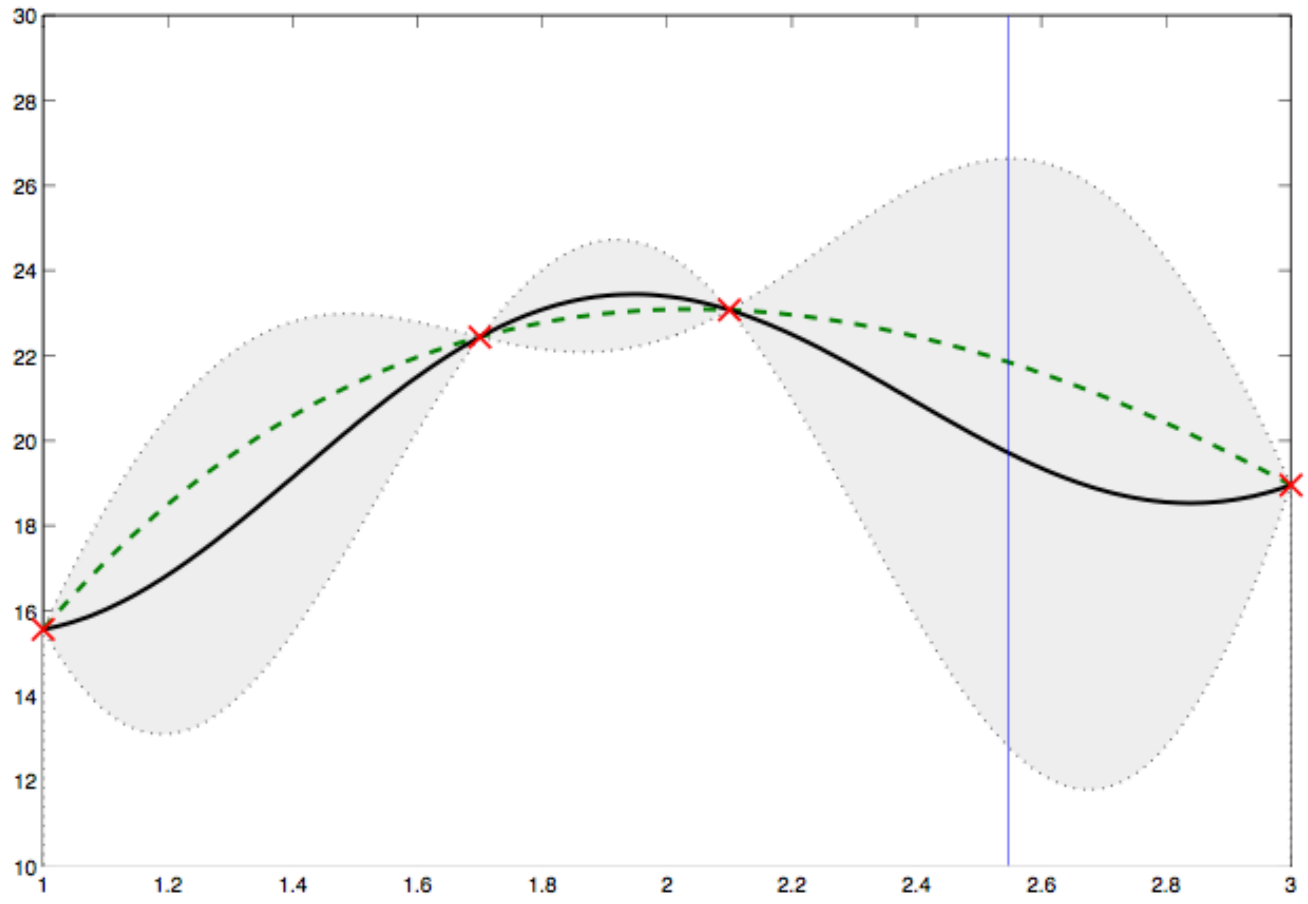
# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_{\theta}[\mu^{*}(\theta) + \beta_t \sqrt{k^{*}(\theta, \theta)}]$



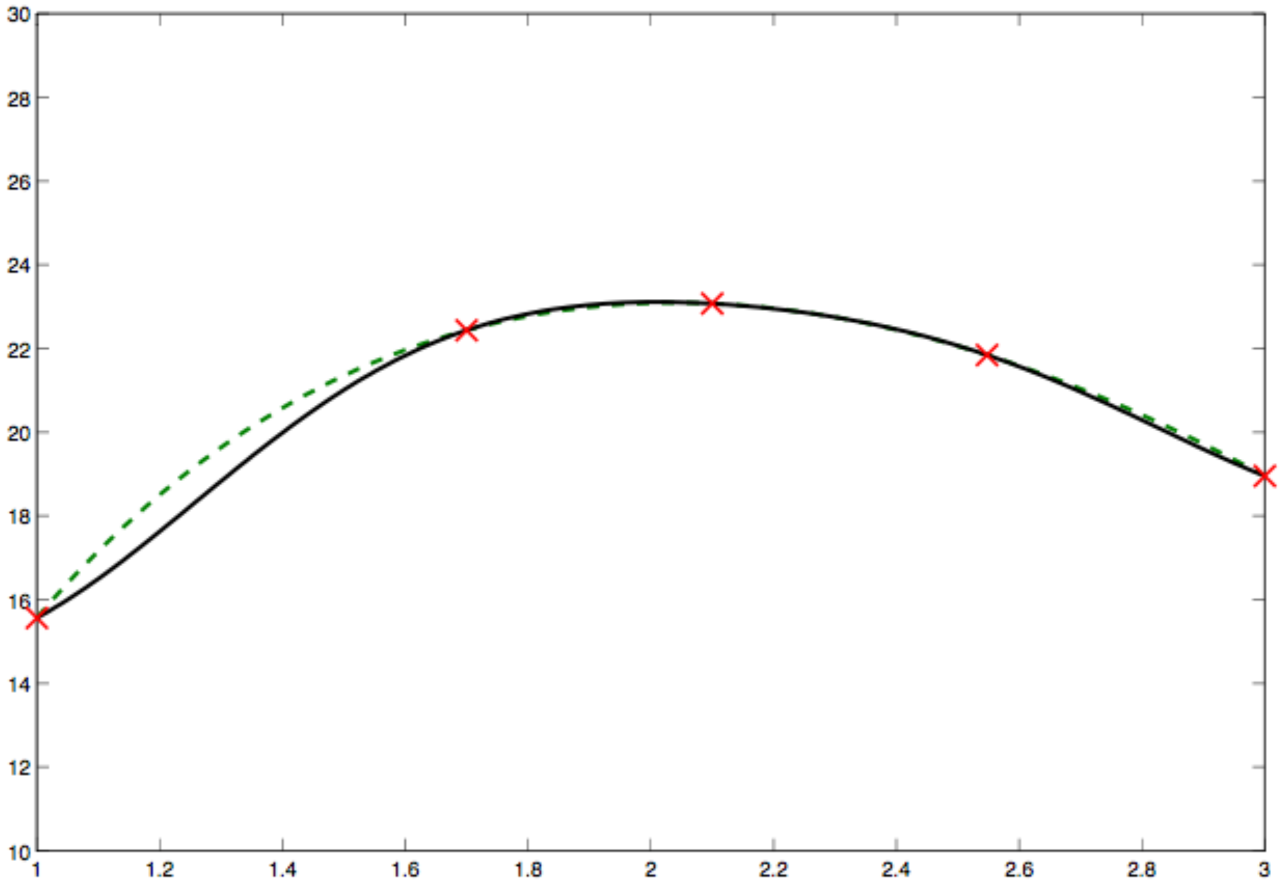
# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_{\theta}[\mu^*(\theta) + \beta_t \sqrt{k^*(\theta, \theta)}]$



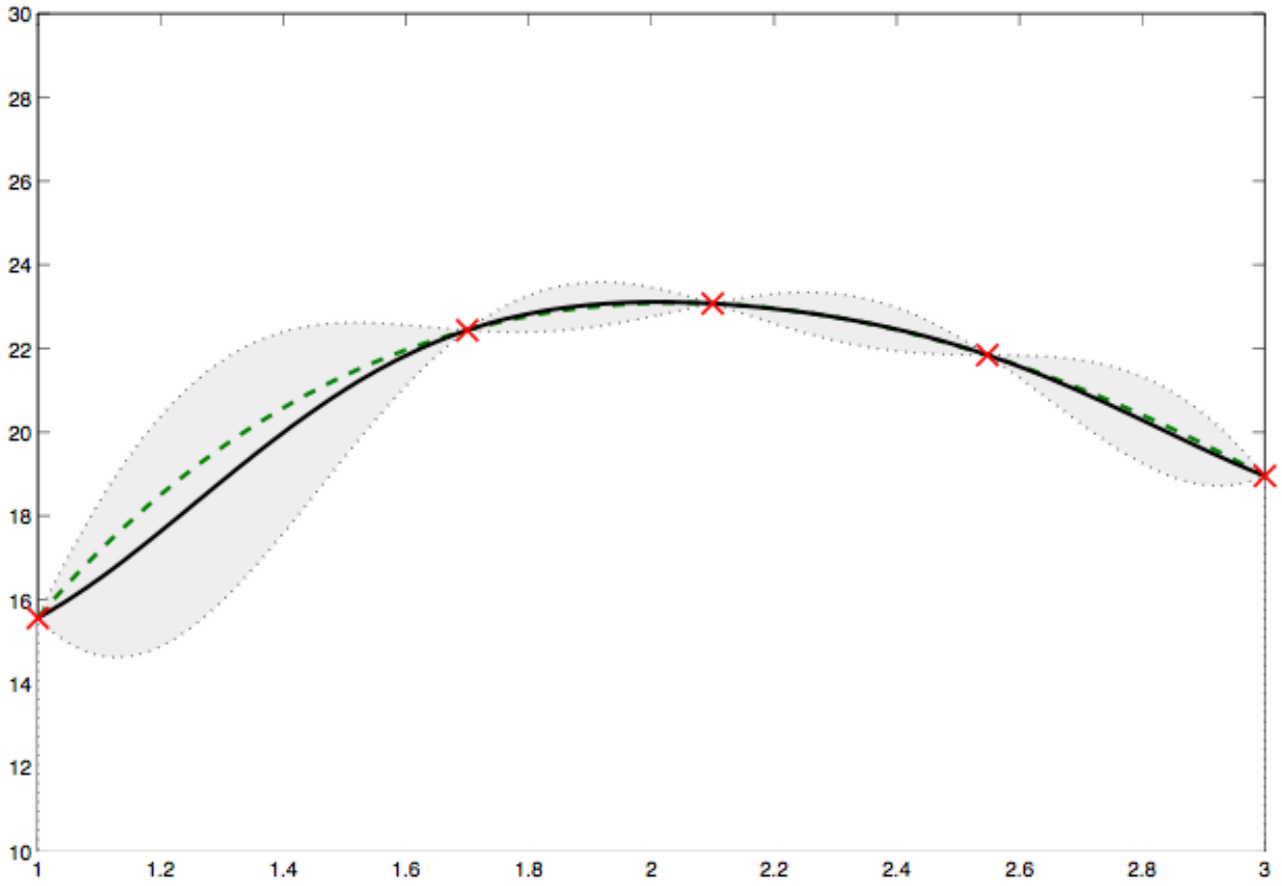
# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_\theta[\mu^*(\theta) + \beta_t\sqrt{k^*(\theta,\theta)}]$



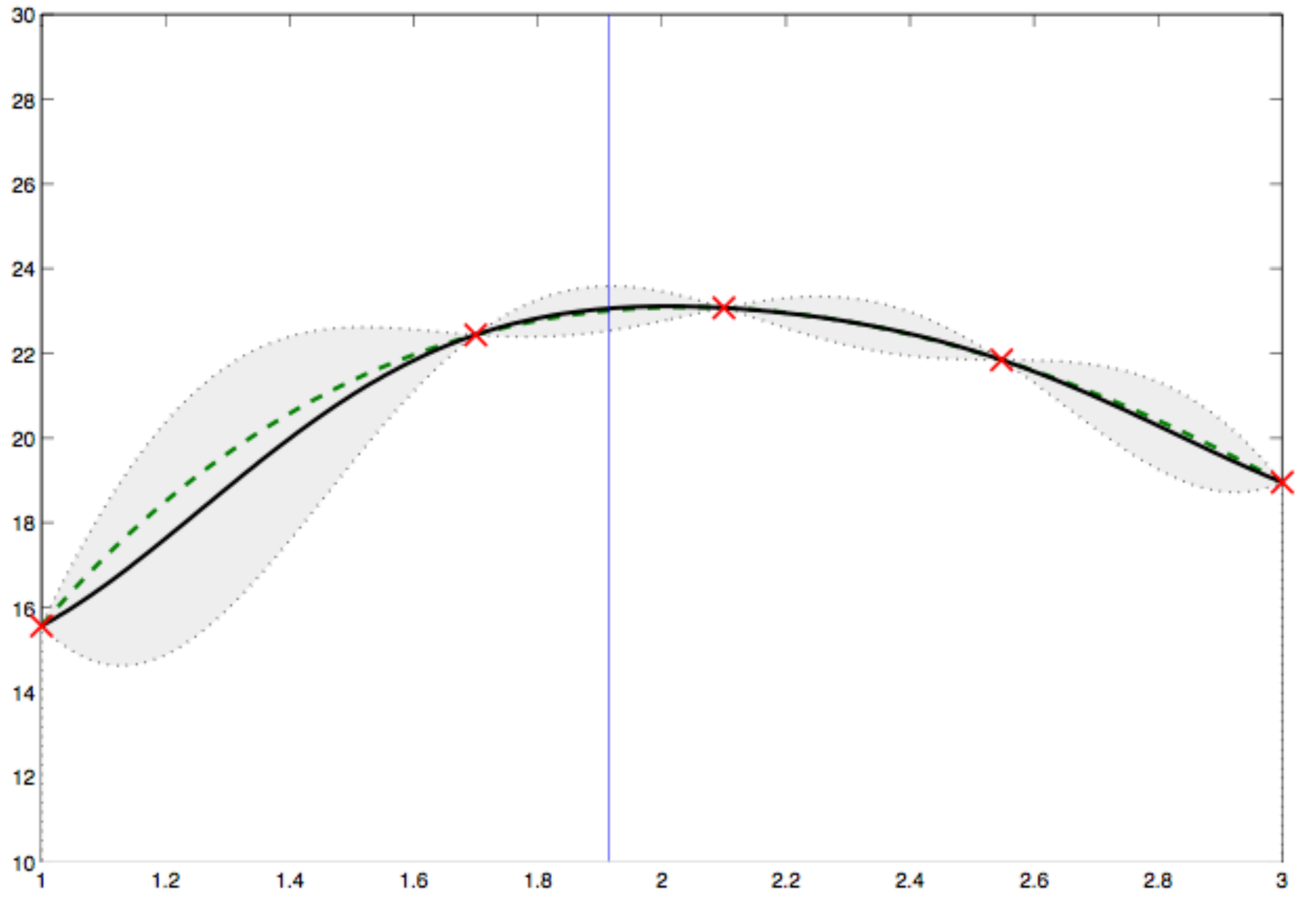
# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_{\theta}[\mu^*(\theta) + \beta_t \sqrt{k^*(\theta,\theta)}]$



# (3) The GP-UCB Algorithm

Balance Exploration and Exploitation: we maximise the **95% upper quantile of the distribution**: $\theta_{t+1} = argmax_\theta[\mu^*(\theta) + \beta_t\sqrt{k^*(\theta,\theta)}]$



# Bibliography

**Parameter Synthesis:**


- Ezio Bartocci, Luca Bortolussi, Laura Nenzi, Guido Sanguinetti, System design of stochastic models using robustness of temporal properties. Theor. Comput. Sci. 587: 3-25 (2015)
- Bortolussi L., Silvetti S. (2018*) Bayesian Statistical Parameter Synthesis for Linear Temporal Properties of Stochastic Models.* TACAS 2018. LNCS, vol 10806. Springer, Cham