Gaussian Processes & RL

10-403 Recitation 2

Conor Igoe — 02 / 12 / 21

Gaussian Processes & RL

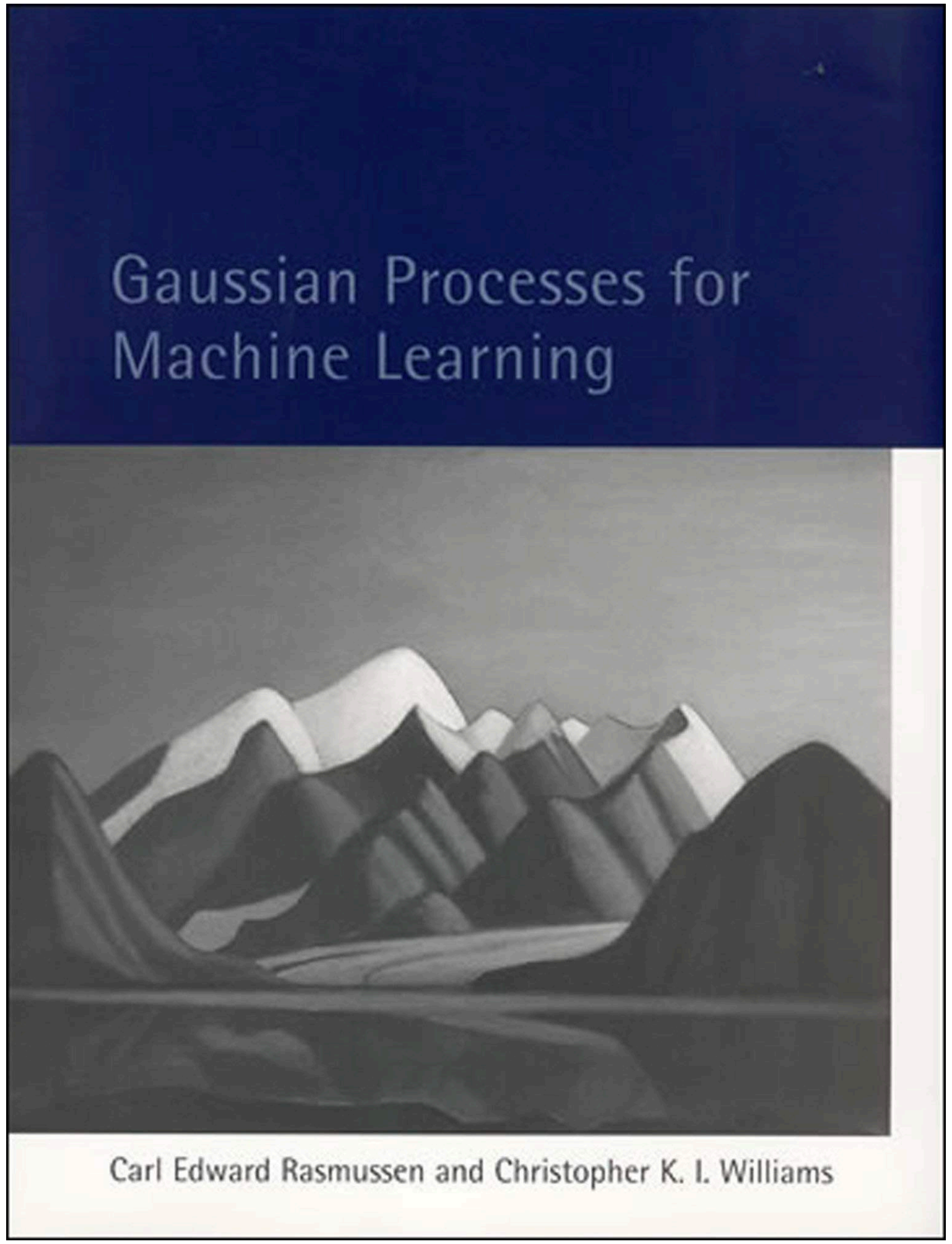
Overview

- Main focus: aleatoric & epistemic uncertainty in RL intuition
  - GP basics
  - Two major types of uncertainty in RL
  - GP-based active exploration algorithms for simple continuous problem
  - GP limitations for RL

Gaussian Processes & RL

GP basics

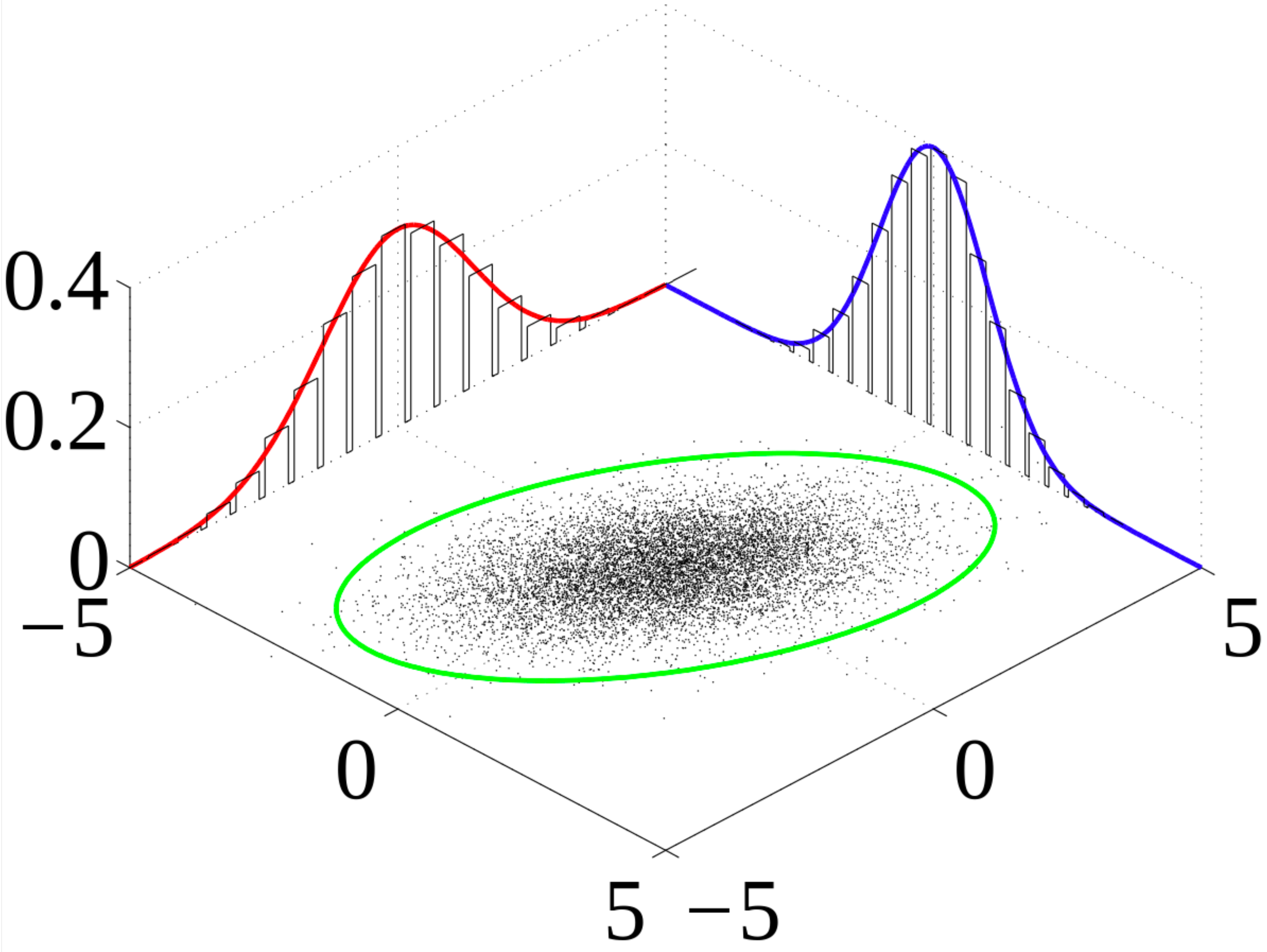
- Excellent reference:
  - C. E. Rasmussen & C. K. I. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006
  - Chapters 1, 2 & 4



Gaussian Processes & RL

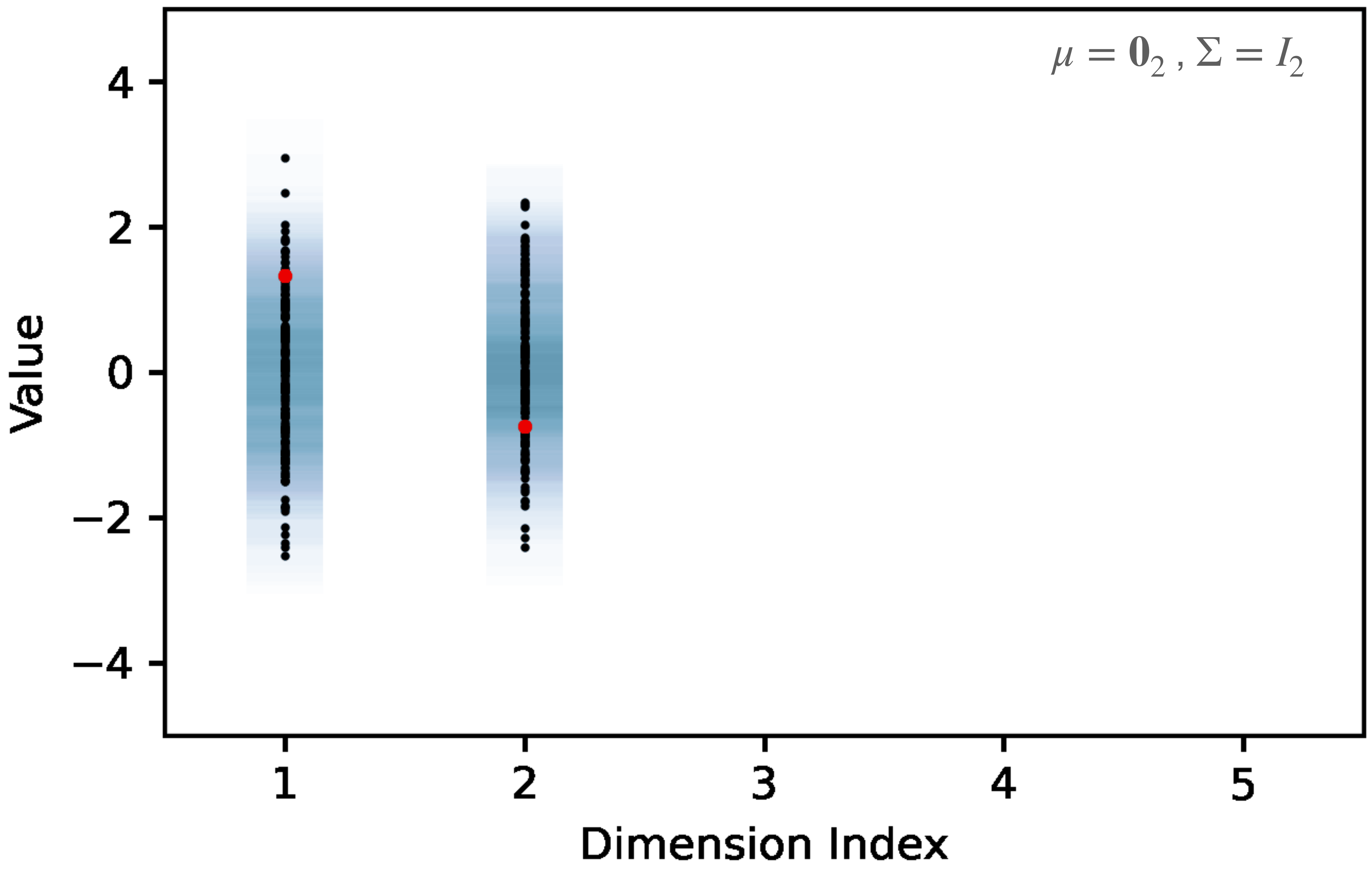
GP basics

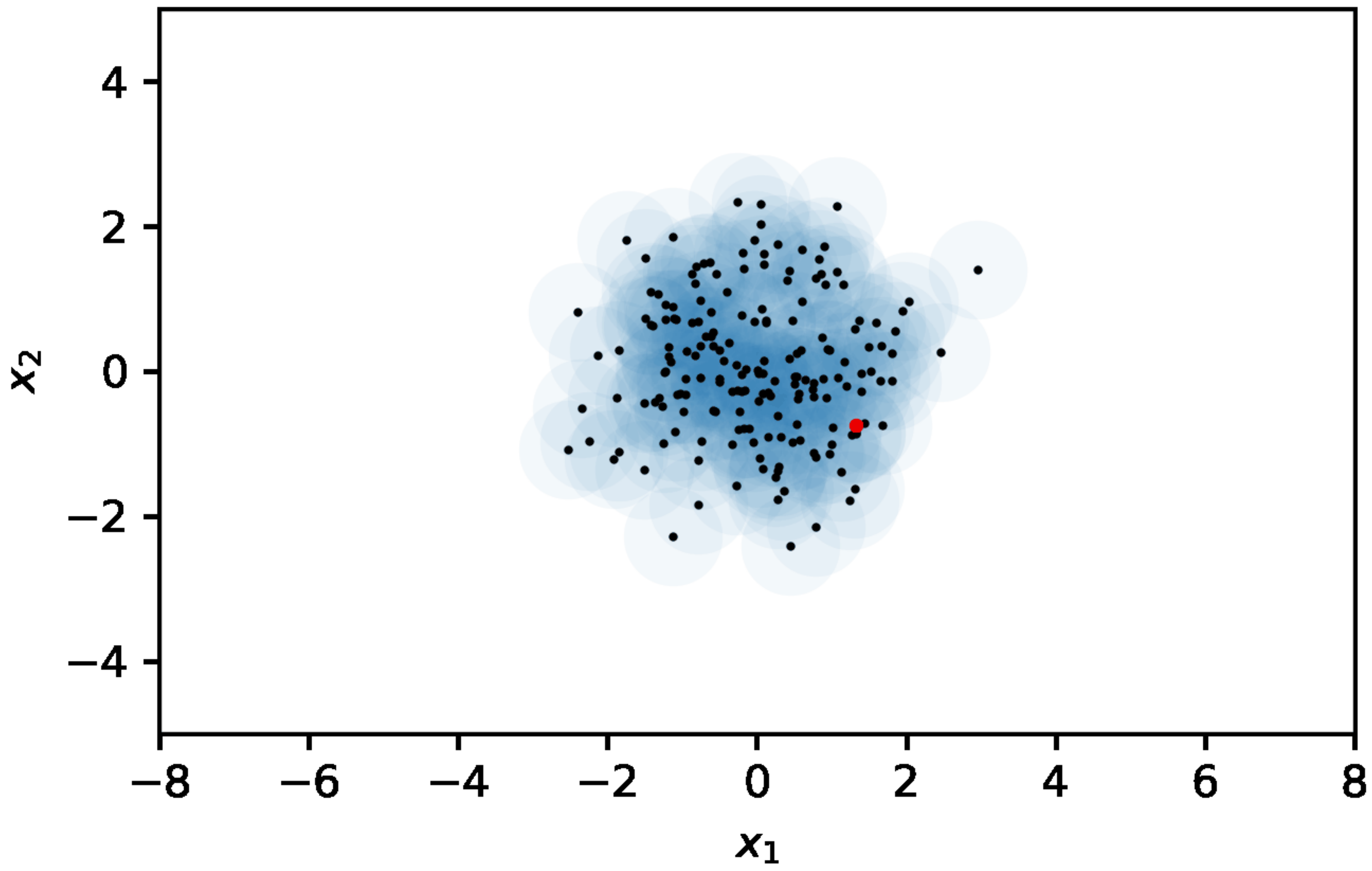
- Multivariate Gaussians
  - $\mathbf{x}$ random k-dimensional vector
  - $x \in \mathbb{R}^k$ realisation of $\mathbf{x}$
  - $\mu \in \mathbb{R}^k$ mean vector
  - $\Sigma \in \mathbb{S}^k_{++}$ covariance matrix
  - $p(\mathbf{x} = x; \mu, \Sigma) = (2\pi)^{\frac{k}{2}} \det(\Sigma) \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$



Gaussian Processes & RL
GP basics

- Multivariate Gaussians

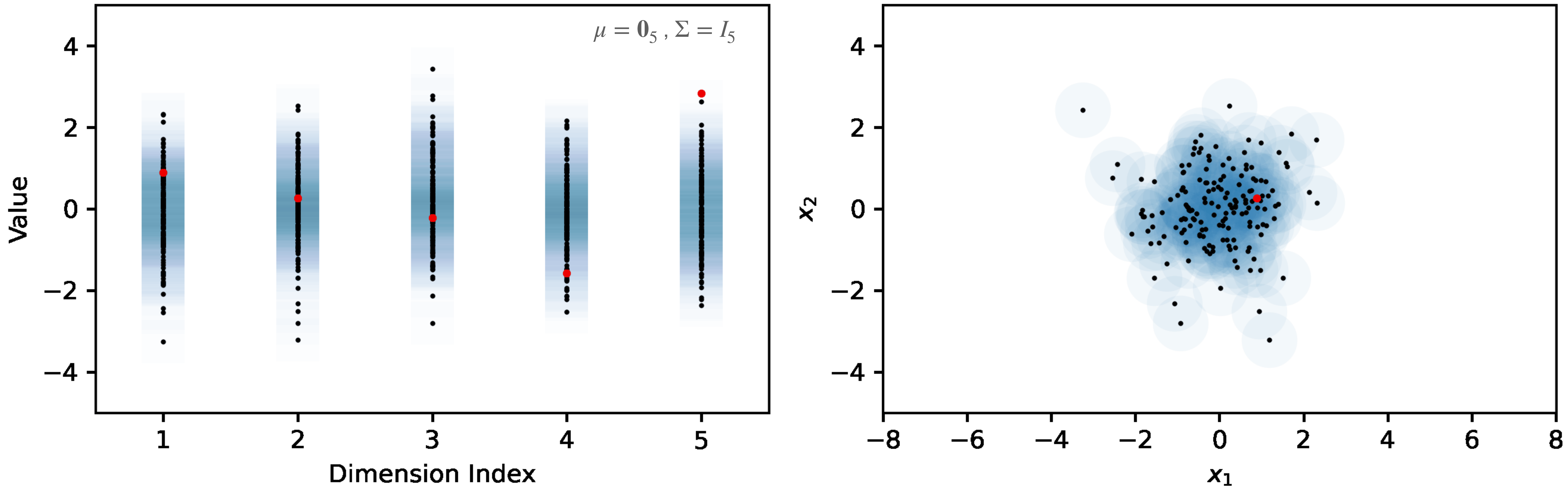




Gaussian Processes & RL

GP basics

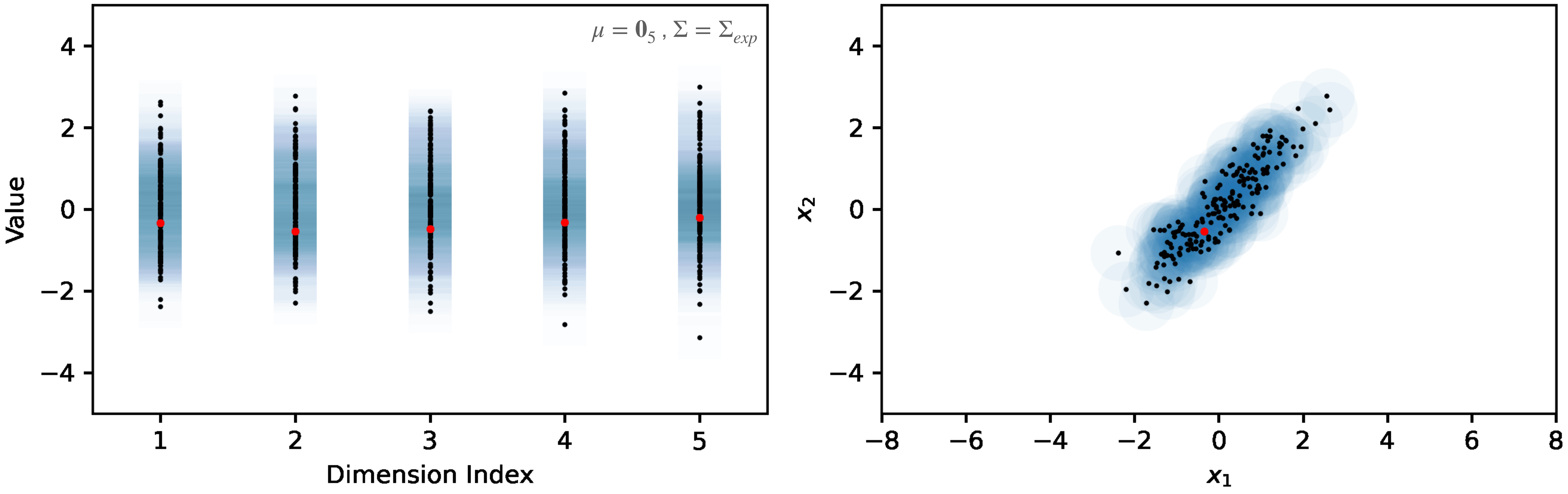
- Multivariate Gaussians



Gaussian Processes & RL

GP basics

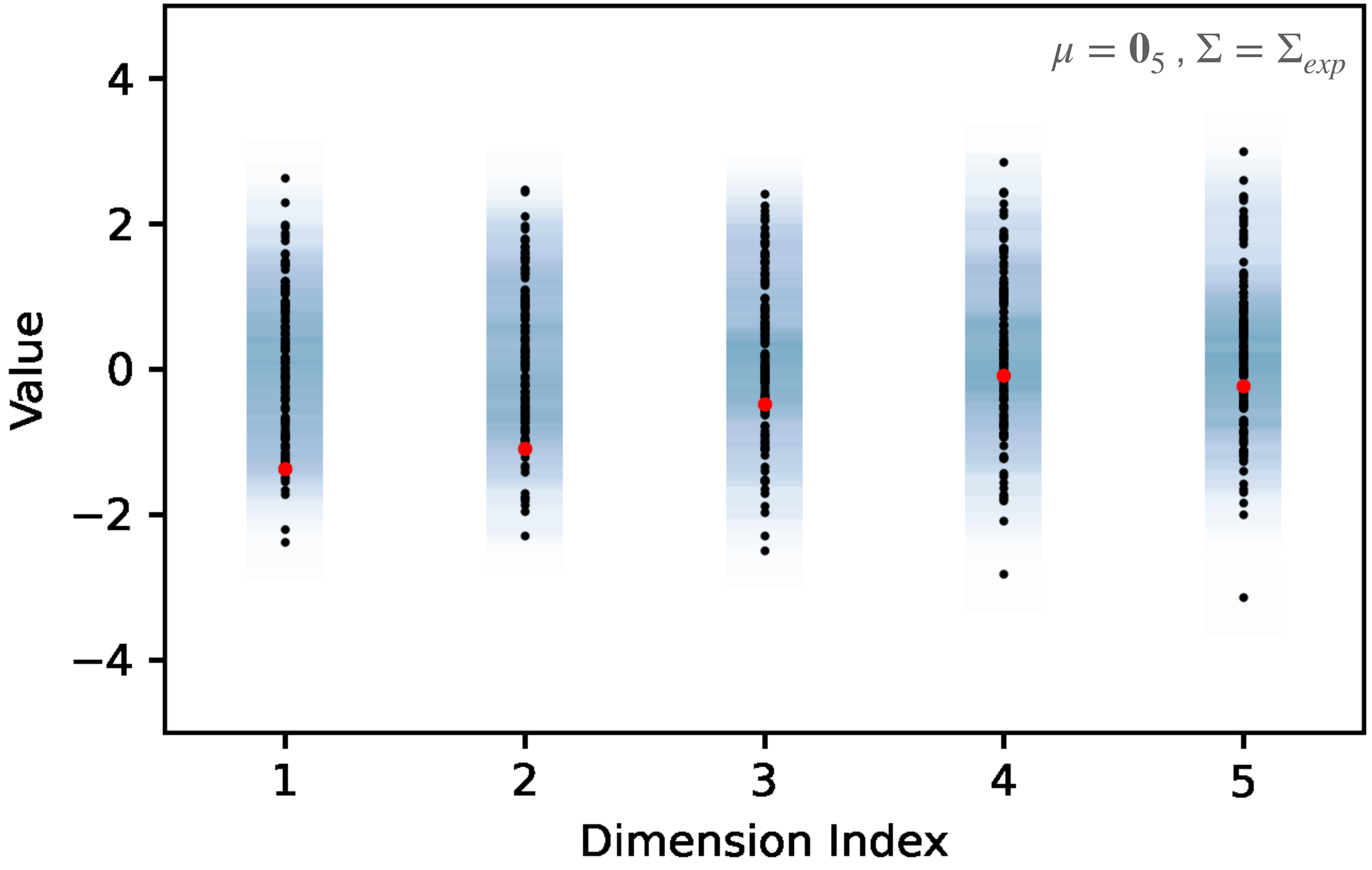
- Multivariate Gaussians



Gaussian Processes & RL

GP basics

- Multivariate Gaussians



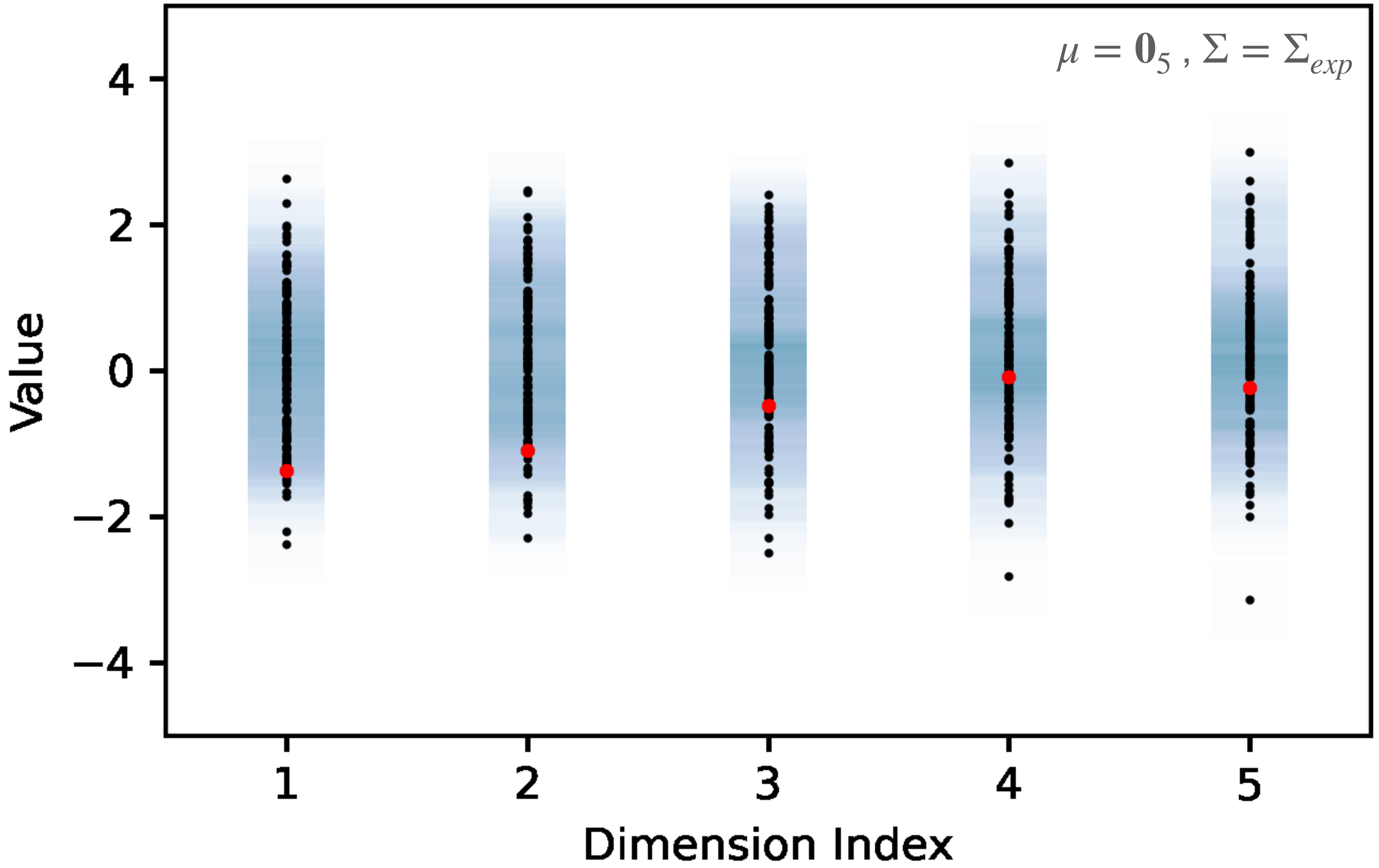
$\Sigma_{exp} =$

	0	1	2	3	4
0	1.000000	0.904837	0.670320	0.406570	0.201897
1	0.904837	1.000000	0.904837	0.670320	0.406570
2	0.670320	0.904837	1.000000	0.904837	0.670320
3	0.406570	0.670320	0.904837	1.000000	0.904837
4	0.201897	0.406570	0.670320	0.904837	1.000000

Gaussian Processes & RL

GP basics

- Multivariate Gaussians

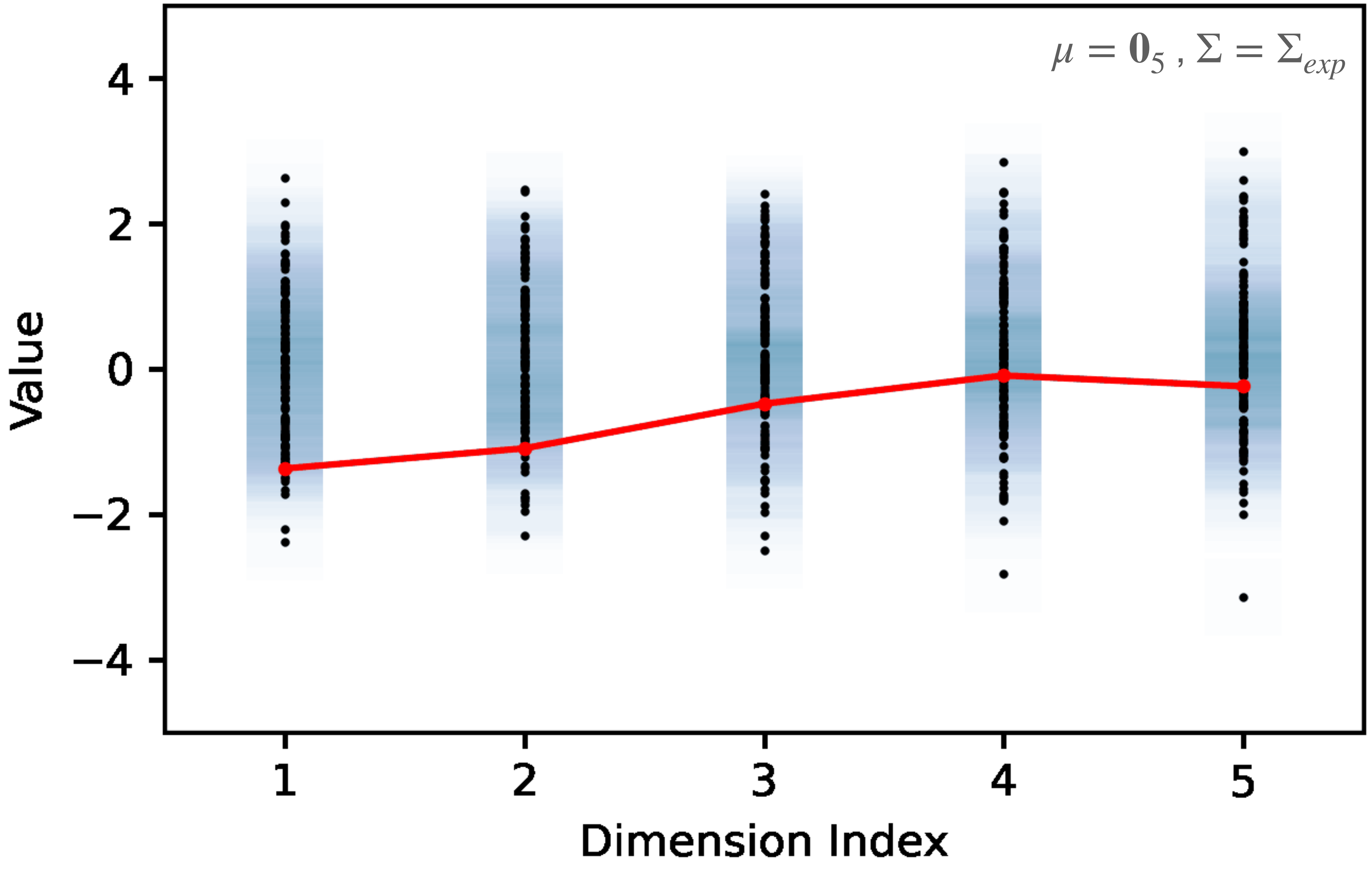


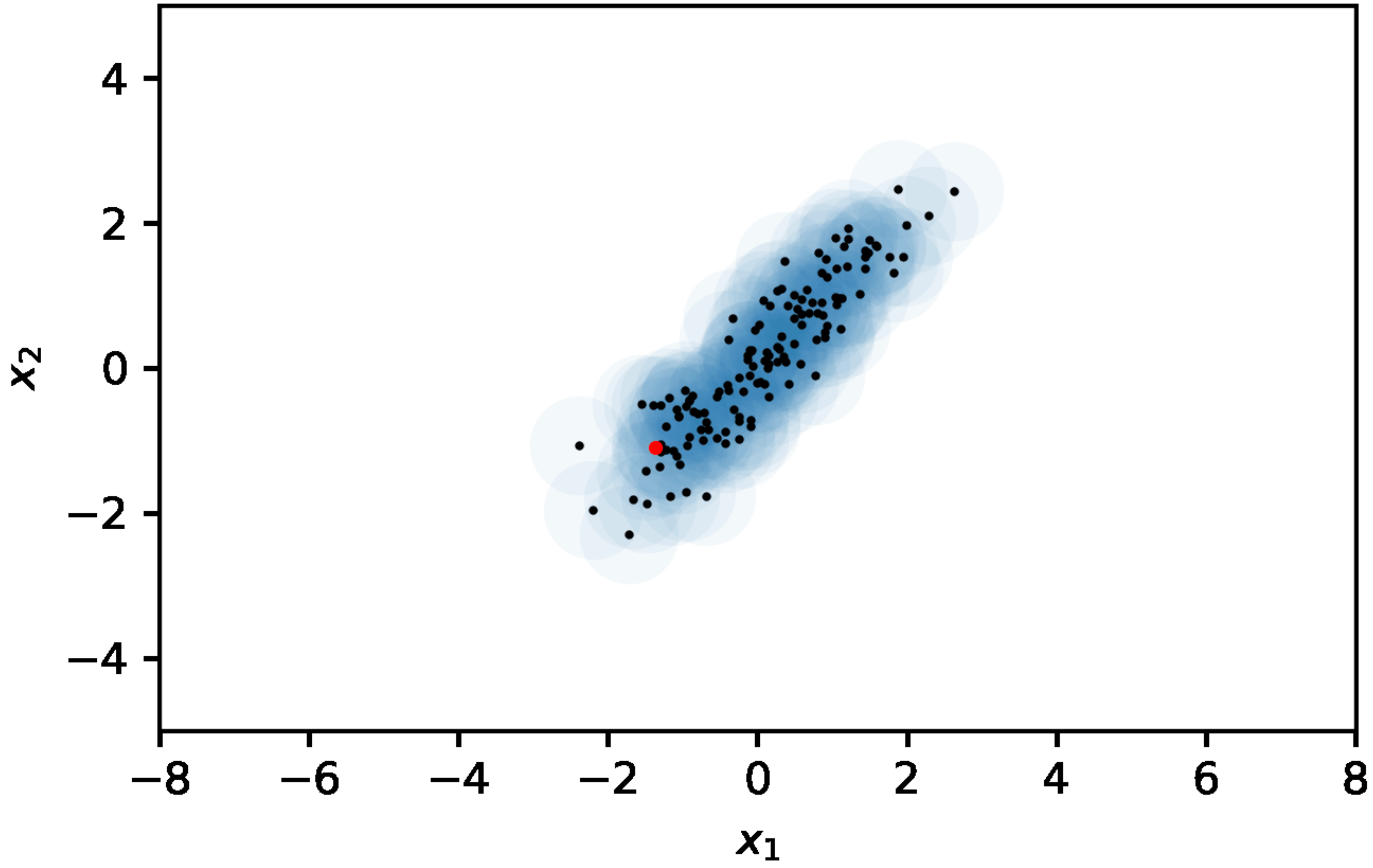
$$\Sigma_{exp}[i, j] = \exp(-(i-j)^2)$$

Gaussian Processes & RL

GP basics

- Multivariate Gaussians

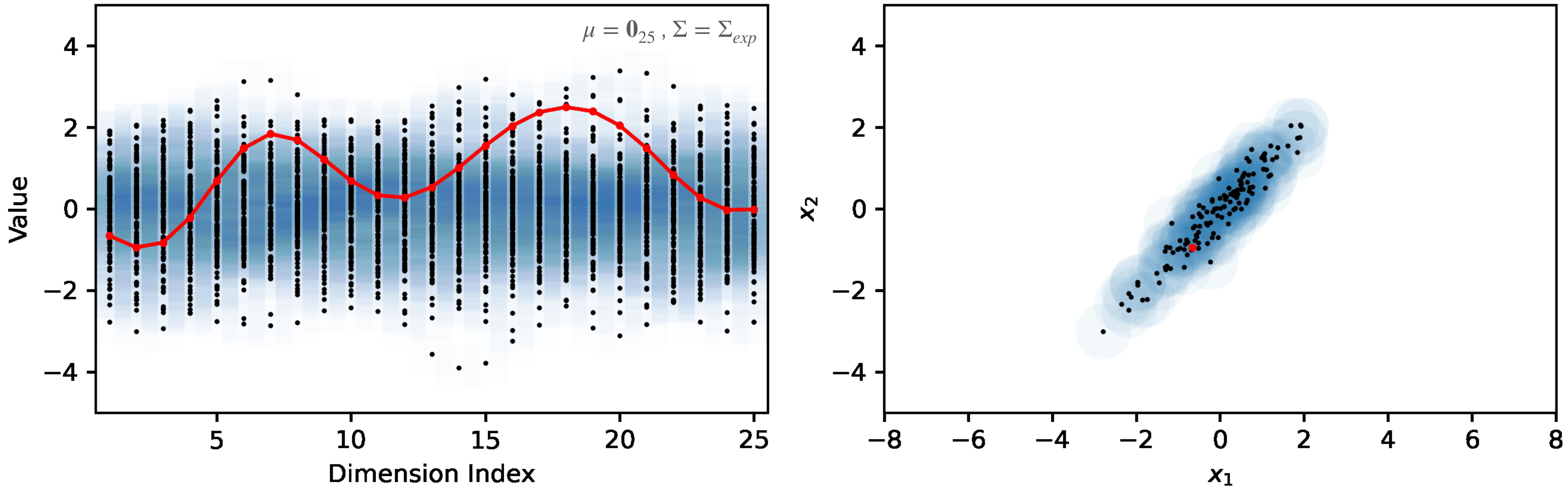




Gaussian Processes & RL

GP basics

- Multivariate Gaussians



Questions?

Gaussian Processes & RL

GP basics

- Multivariate Gaussians
  - 1st important property:
    - If we have $\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1}, \Sigma_{1,2} \\ \Sigma_{1,2}, \Sigma_{2,2} \end{bmatrix} \right)$

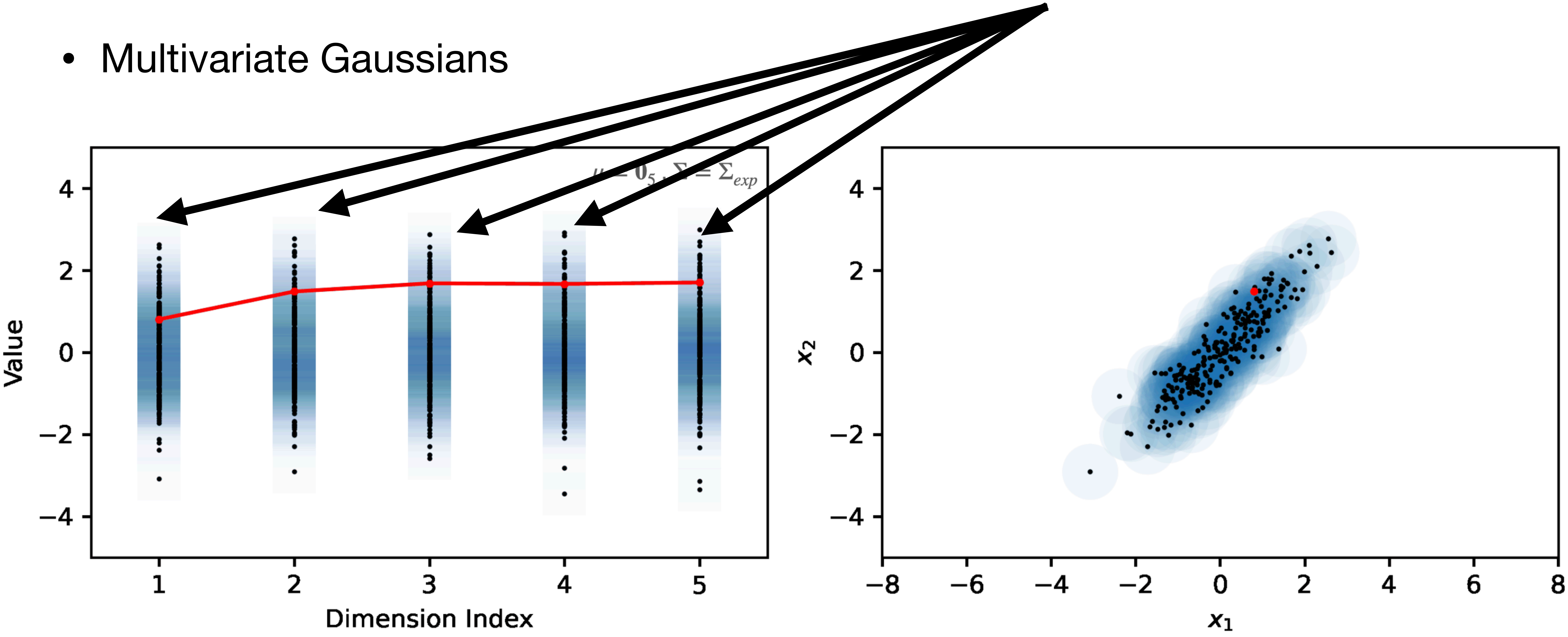
      then $\mathbf{x}_1 \sim \mathcal{N}(\mu_1, \Sigma_{1,1})$

Gaussian Processes & RL

GP basics

Each marginal is also a gaussian

- Multivariate Gaussians



Gaussian Processes & RL

GP basics

- Multivariate Gaussians
  - 2nd important property:
    - If we have $\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1}, \Sigma_{1,2} \\ \Sigma_{1,2}, \Sigma_{2,2} \end{bmatrix} \right)$

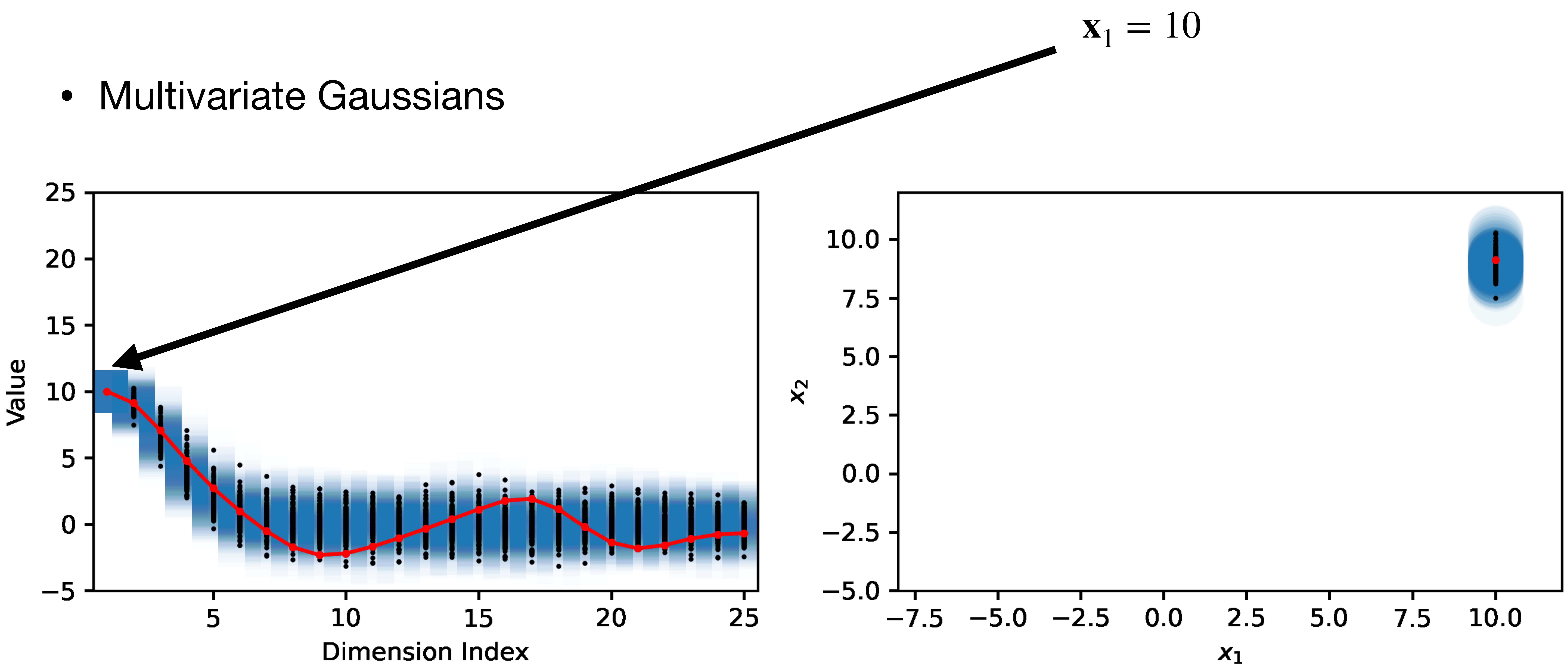
      then $\mathbf{x}_1 \mid \mathbf{x}_2 = x_2 \sim \mathcal{N}(\tilde{\mu}, \tilde{\Sigma})$

      where $\tilde{\mu} = \underbrace{\mu_1}_{\text{prior}} + \underbrace{\Sigma_{1,2}\Sigma_{2,2}^{-1}(x_2 - \mu_2)}_{\text{information from } x_2}$, $\tilde{\Sigma} = \underbrace{\Sigma_{1,1}}_{\text{prior}} - \underbrace{\Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{1,2}^T}_{\text{reduction in uncertainty}}$

Gaussian Processes & RL

GP basics

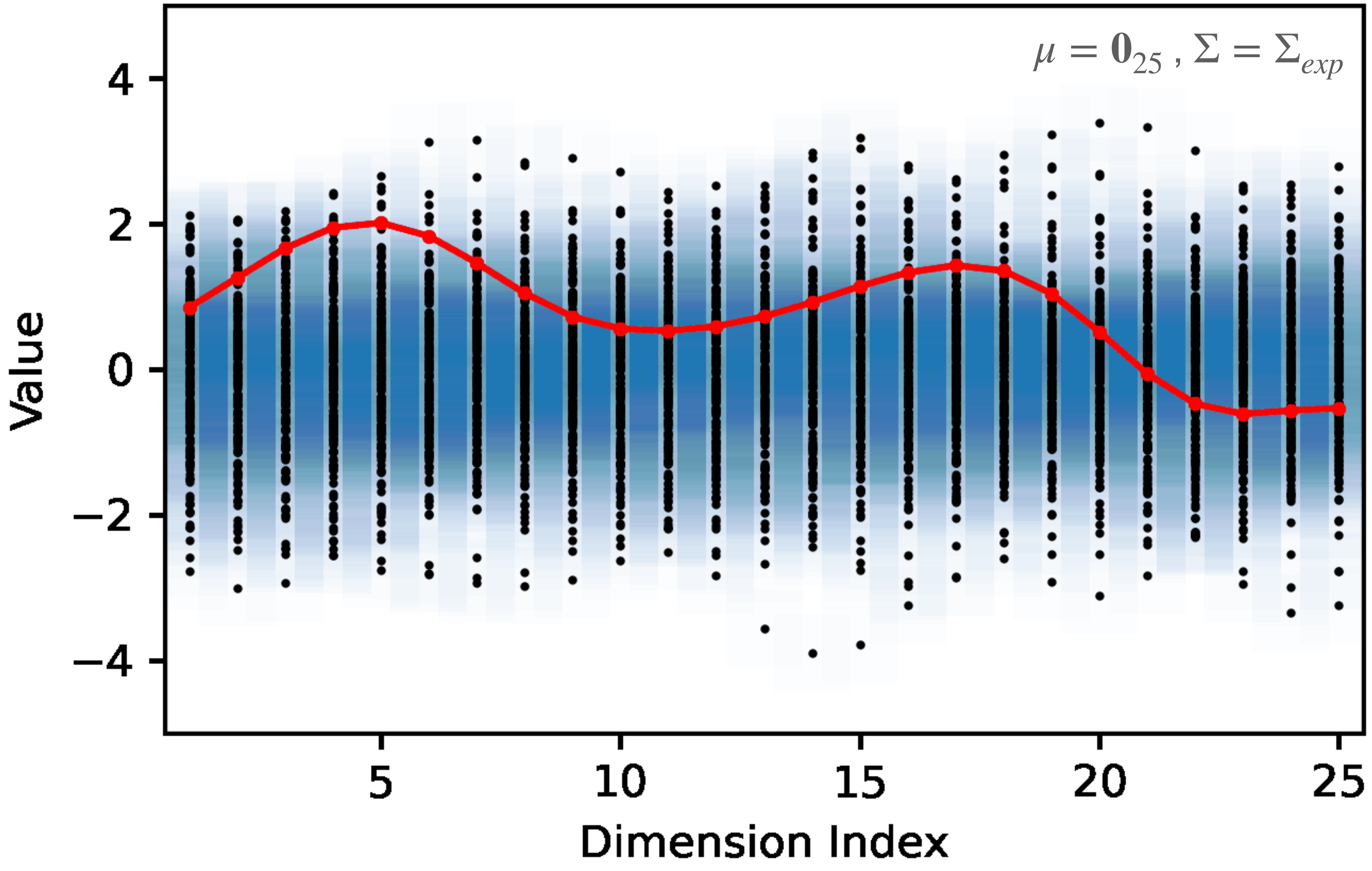
- Multivariate Gaussians



Gaussian Processes & RL

GP basics

- Multivariate Gaussians



- Gaussian processes generalise this to infinitely many indices
- Set of indices (or "inputs") $\mathscr{X}$
  - $\mathscr{X} = \mathbb{N}$
  - $\mathscr{X} = \mathbb{R}$
  - $\mathscr{X} = \mathbb{R}^d$
  - …

(for multivariate Gaussians)

$\mathscr{X} = [k]$

- For multivariate Gaussians, need μ, Σ; for GPs, need:
  - $m : \mathscr{X} \to \mathbb{R}$ — $m(x) = \mu[x]$
  - $k : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ — $k(x, x') = \Sigma[x, x']$

Gaussian Processes & RL

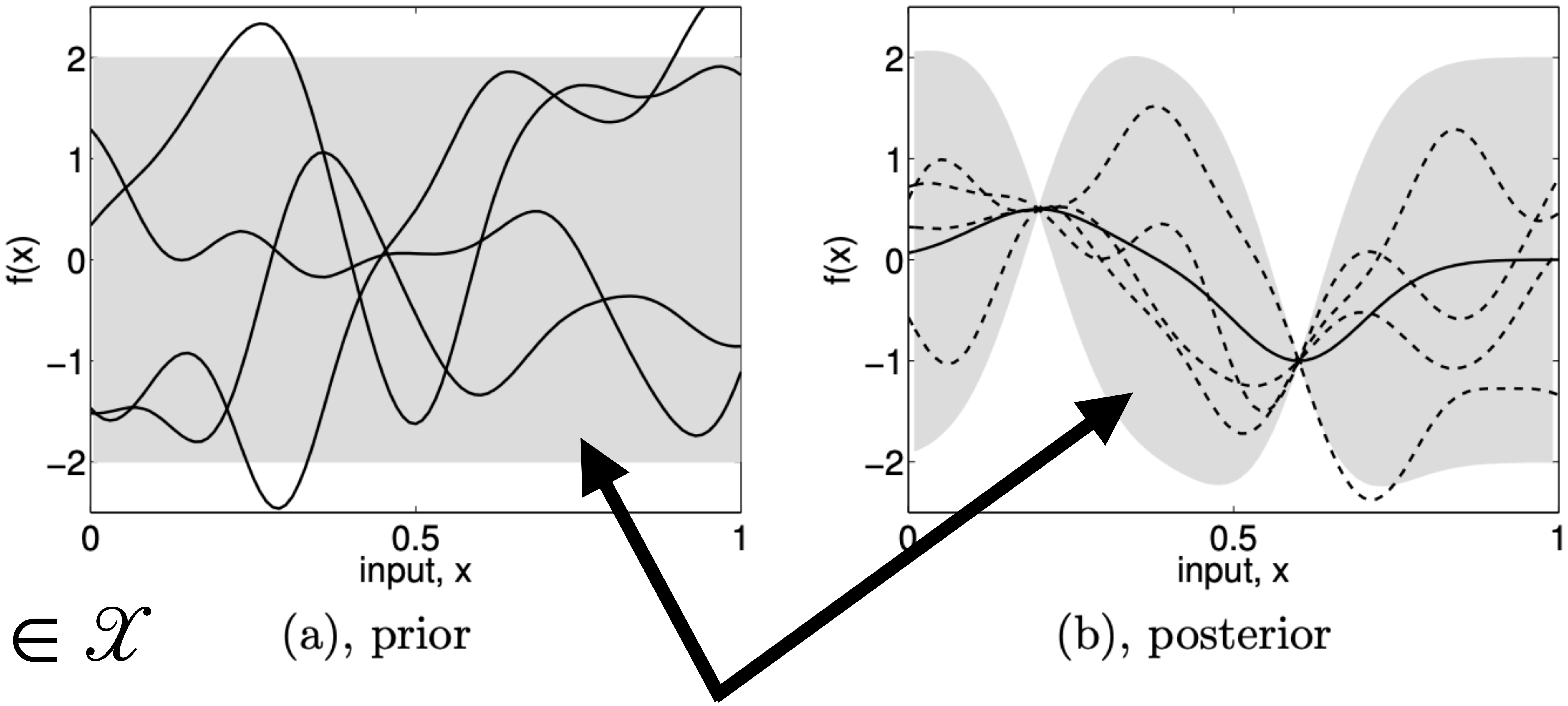
GP basics

- Gaussian Process: Definitions & Examples
  - **Definition: a Gaussian Process is a set of random variables, every finite subset of which are jointly Gaussian.**
  - Individual random variables in the set are denoted $f(x)$, indexed by $x \in \mathcal{X}$
  - Mean function $m(x) := \mathbb{E}[f(x)]$
  - Covariance function (kernel) $k(x, x') := \text{Cov}\big(f(x), f(x')\big)$
    $$= \mathbb{E}\left[\big(f(x) - m(x)\big)\,\big(f(x') - m(x')\big)\right]$$
  - m & k fully define a GP, and we usually write: $f(x) \sim \mathcal{GP}\left(m(x), k\left(x, x'\right)\right)$

Gaussian Processes & RL

GP basics

- Gaussian Process: Definitions & Examples
  - **Definition: a Gaussian Process is a set of random variables, every finite subset of which are jointly Gaussian.**
- Example:
  - $\mathcal{X} = \mathbb{R}$
  - $m(x) = 0 \quad \forall x \in \mathcal{X}$
  - $k(x, x') = \exp\left(-(x - x')^2\right) \quad \forall x, x' \in \mathcal{X}$



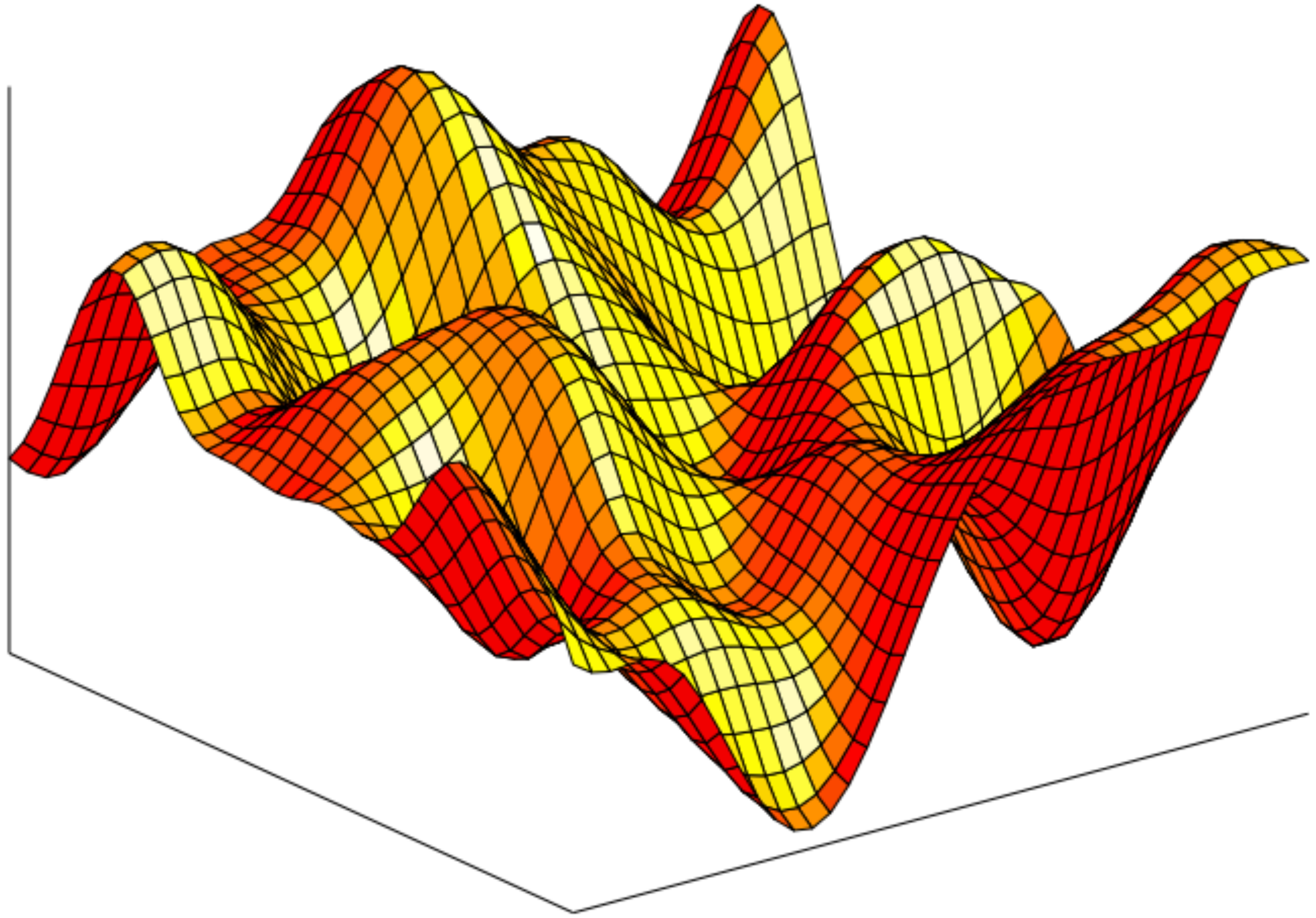
(a), prior (b), posterior

95% confidence region

Gaussian Processes & RL

GP basics

- Gaussian Process: Definitions & Examples
  - **Definition: a Gaussian Process is a set of random variables, every finite subset of which are jointly Gaussian.**



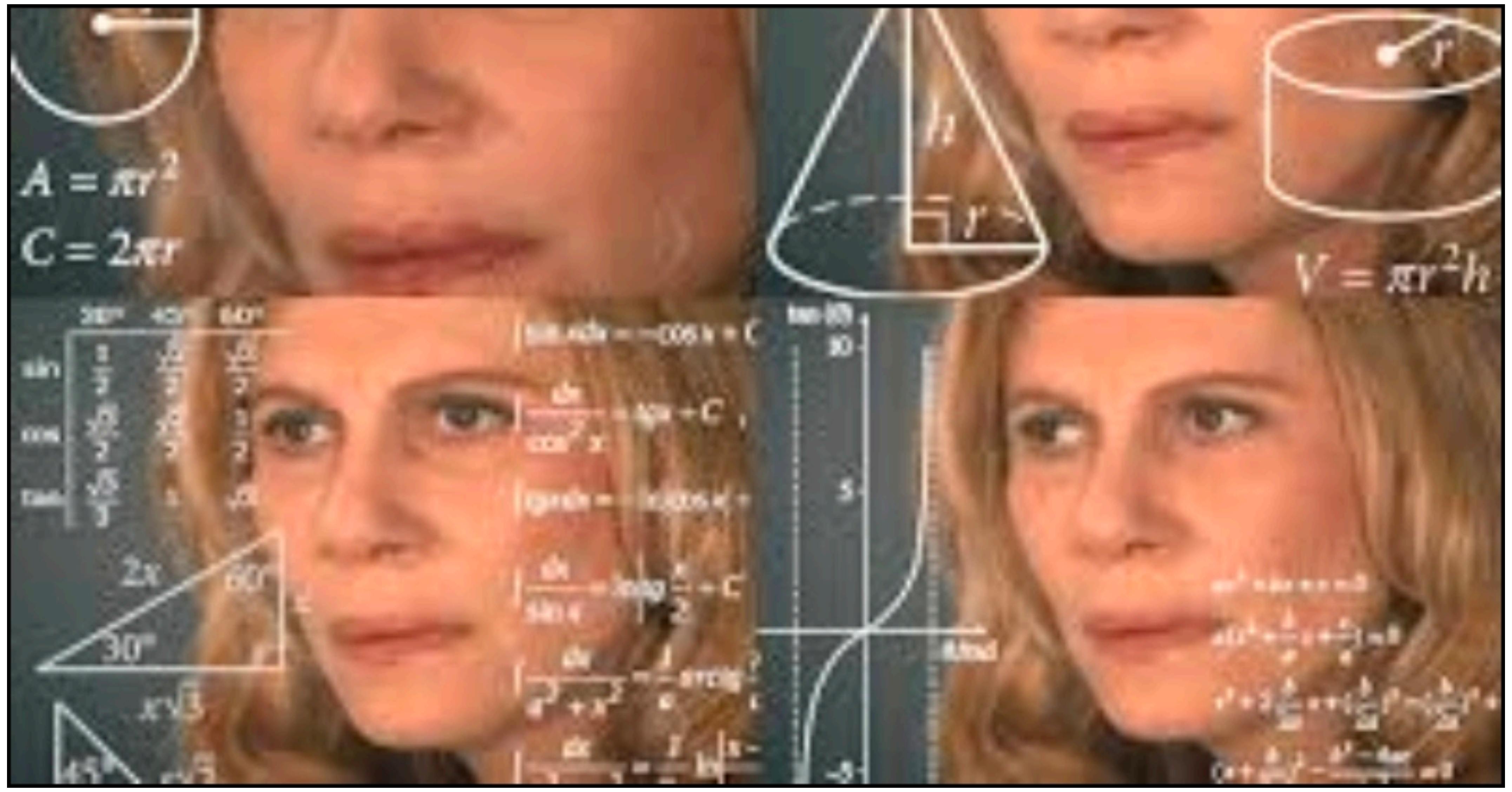
- Example:
  - $\mathcal{X} = \mathbb{R}^2$
  - $m(x) = 0 \quad \forall x \in \mathcal{X}$
  - $k(x, x') = \exp\left(-\|x - x'\|_2^2\right) \quad \forall x, x' \in \mathcal{X}$

Gaussian Processes & RL

GP basics

- Gaussian Process: Definitions & Examples
  - **Definition: a Gaussian Process is a set of random variables, every finite subset of which are jointly Gaussian.**

- Example:
  - $\mathscr{X} = \mathbb{R}^d$
  - $m(x) = 0 \quad \forall x \in \mathscr{X}$
  - $k(x, x') = \exp\left(-\|x - x'\|_2^2\right) \quad \forall x, x' \in \mathscr{X}$



Questions?

Gaussian Processes & RL

GP basics

- Gaussian Processes: Learning / Inference
  - Intuitively, imagine nature draws a function $f \sim \mathcal{GP}(m, k)$
    - i.e. nature jointly draws $\{f(x)\}_{x \in \mathcal{X}}$
  - Then we are given a data set $\mathcal{D} = \{x_i, y_i\}_{i=1}^{N}$, where $y_i = f(x_i)$
  - We might be interested in making inferences for another set of inputs $\{\tilde{x}_i\}_{i=1}^{M}$
  - Thanks to Gaussian conditional & marginalisation properties, inference is straightforward
    - Doesn't matter that there are infinitely many random variables $\{f(x)\}_{x \in \mathcal{X}}$
    - We can just focus on the "training set" RVs $\{f(x_i)\}_{i=1}^{N}$, and the "test set" RVs $\{f(\tilde{x}_i)\}_{i=1}^{M}$

Gaussian Processes & RL

GP basics

- Gaussian Processes: Learning / Inference
  - Focus on the "training set" $\left\{f(x_i)\right\}_{i=1}^{N}$, and the "test set" $\left\{f(\tilde{x}_i)\right\}_{i=1}^{M}$

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \\ f(\tilde{x}_1) \\ \vdots \\ f(\tilde{x}_M) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1}, \Sigma_{1,2} \\ \Sigma_{1,2}, \Sigma_{2,2} \end{bmatrix} \right) = \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_N) \\ m(\tilde{x}_1) \\ \vdots \\ m(\tilde{x}_M) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) & k(x_1, \tilde{x}_1) & \dots & k(x_1, \tilde{x}_M) \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) & k(x_N, \tilde{x}_1) & \dots & k(x_N, \tilde{x}_M) \\ k(\tilde{x}_1, x_1) & \dots & k(\tilde{x}_1, x_N) & k(\tilde{x}_1, \tilde{x}_1) & \dots & k(\tilde{x}_1, \tilde{x}_M) \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ k(\tilde{x}_M, x_1) & \dots & k(\tilde{x}_M, x_N) & k(\tilde{x}_M, \tilde{x}_1) & \dots & k(\tilde{x}_M, \tilde{x}_M) \end{bmatrix} \right)$$

Gaussian Processes & RL

GP basics

- Gaussian Processes: Learning / Inference
  - We can condition the test RVs on the observed training RVs

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \\ f(\tilde{x}_1) \\ \vdots \\ f(\tilde{x}_M) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1}, & \Sigma_{1,2} \\ \Sigma_{1,2}, & \Sigma_{2,2} \end{bmatrix} \right) = \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_N) \\ m(\tilde{x}_1) \\ \vdots \\ m(\tilde{x}_M) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) & k(x_1, \tilde{x}_1) & \dots & k(x_1, \tilde{x}_M) \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) & k(x_N, \tilde{x}_1) & \dots & k(x_N, \tilde{x}_M) \\ k(\tilde{x}_1, x_1) & \dots & k(\tilde{x}_1, x_N) & k(\tilde{x}_1, \tilde{x}_1) & \dots & k(\tilde{x}_1, \tilde{x}_M) \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ k(\tilde{x}_M, x_1) & \dots & k(\tilde{x}_M, x_N) & k(\tilde{x}_M, \tilde{x}_1) & \dots & k(\tilde{x}_M, \tilde{x}_M) \end{bmatrix} \right)$$

Gaussian Processes & RL

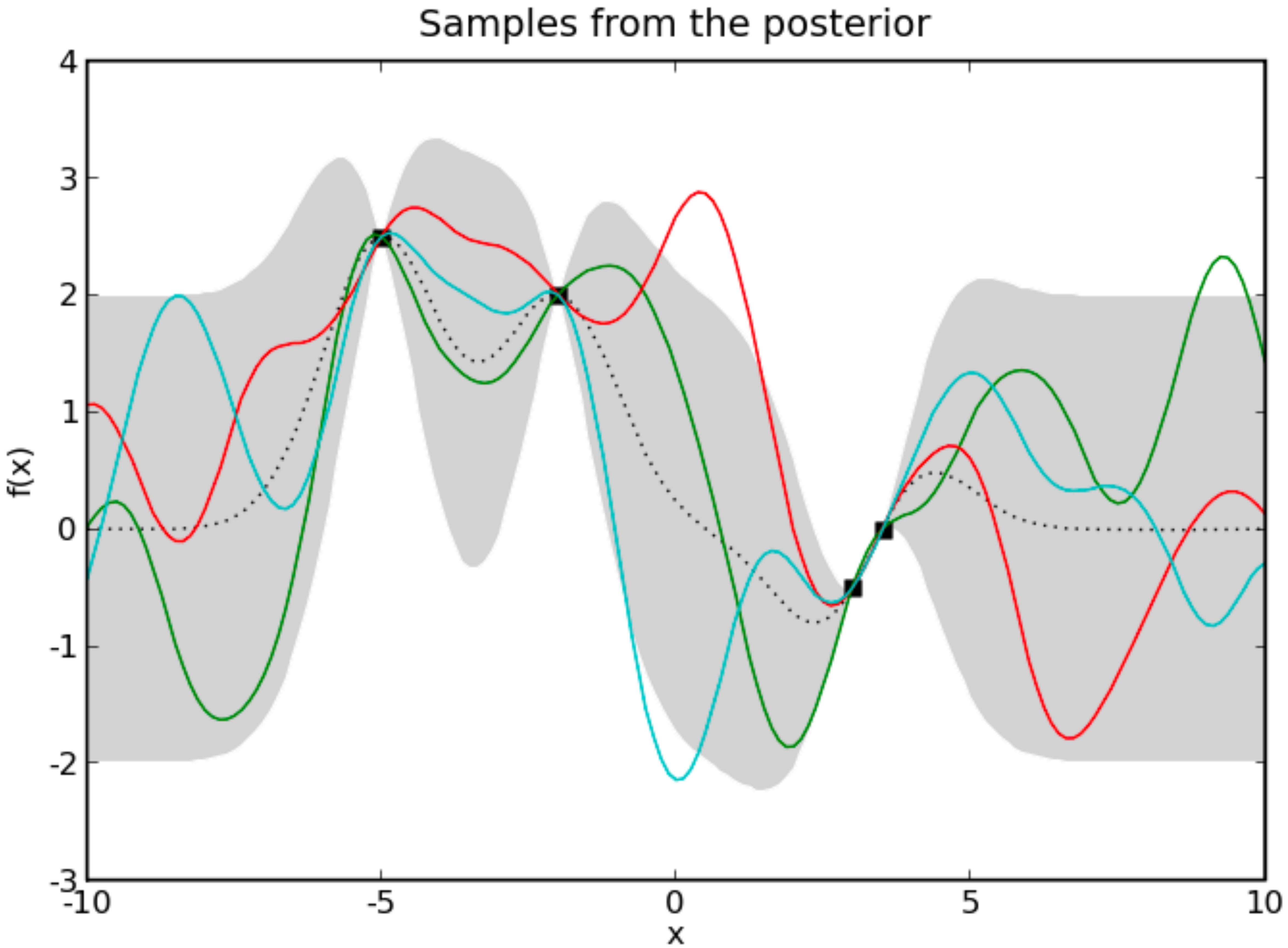
GP basics

- Gaussian Processes: Learning / Inference
  - We can condition the test RVs on the observed training RVs
  - $f(\tilde{x}_1), \ldots, f(\tilde{x}_M) \,|\, f(x_1) = y_1, \ldots, f(x_N) = y_N \sim \mathcal{N}(\tilde{\mu}, \tilde{\Sigma})$
    - where $\tilde{\mu} = \mu_2 + \Sigma_{1,2}\,\Sigma_{1,1}^{-1}\,(y - \mu_1)$, $\tilde{\Sigma} = \Sigma_{2,2} - \Sigma_{1,2}\Sigma_{1,1}^{-1}\Sigma_{1,2}^{T}$

Gaussian Processes & RL

GP basics

- Gaussian Processes: Learning / Inference

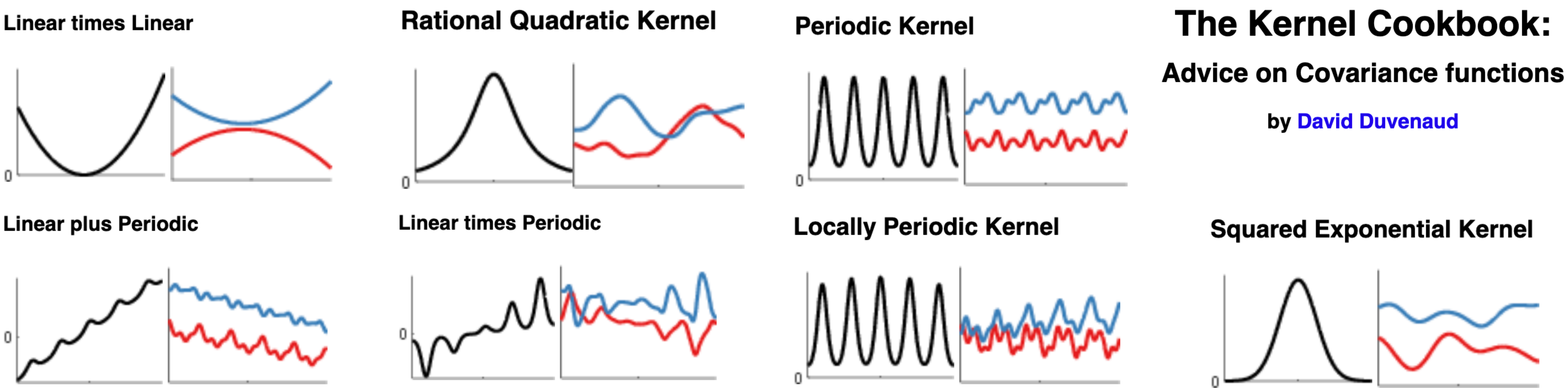


$$\{\tilde{x}_i\}_{i=1}^{M} = \{-10, -9.99, \ldots, 9.99, 10\}$$

Gaussian Processes & RL

GP basics

- Gaussian Processes: Kernels
  - Kernel defines main characteristic behaviour
  - See The Kernel Cookbook



Gaussian Processes & RL

GP basics

- Gaussian Processes
  - See RW text for details on inference for noisy observations, relationship to other statistical models & overview of theoretical results
  - Later, we'll see why Gaussian Processes are not a panacea

Questions?

Gaussian Processes & RL

Two major types of uncertainty in RL

- Aleatoric Uncertainty:
  - Inherent, unavoidable randomness in environment
  - Example:

Bernoulli trial, with success probability = 0.5

Gaussian Processes & RL

Two major types of uncertainty in RL

- Aleatoric Uncertainty:
  - In RL, we typically have two types of Markov state transition functions:
    - $T : \mathscr{S} \times \mathscr{A} \to \mathscr{S}$ (deterministic transition)
    - $T : \mathscr{S} \times \mathscr{A} \to \Delta(\mathscr{S})$ (stochastic transition)
      - Where $\Delta(\mathscr{X})$ is the set of probability distributions defined on a set $\mathscr{X}$
  - We also have two different types of instantaneous reward functions:
    - $r : \mathscr{S} \times \mathscr{A} \times \mathscr{S} \to \mathbb{R}$ (deterministic reward)
    - $r : \mathscr{S} \times \mathscr{A} \times \mathscr{S} \to \Delta(\mathbb{R})$ (stochastic reward)
  - We say the environment has **aleatoric uncertainty** when at least one of these functions is stochastic.

Gaussian Processes & RL

Two major types of uncertainty in RL

- Epistemic Uncertainty:
  - Uncertainty about which environment we're really in
  - Example:

Gaussian Processes & RL

Two major types of uncertainty in RL

- Epistemic Uncertainty:
  - If we are not sure what the true transition function T or reward function r is (or both) then we say we have **epistemic uncertainty**
  - One of the biggest challenges in RL today is coming up with better ways of expressing and leveraging descriptions of epistemic uncertainty
    - Sample efficient exploration & safety
  - In tabular settings, we may have a belief over the true transition matrix
  - In continuous settings, we may have a belief over the true parameters of the state transition function
    - Or we might describe our belief over the state transition function using a Gaussian Process with an appropriate kernel!

Gaussian Processes & RL

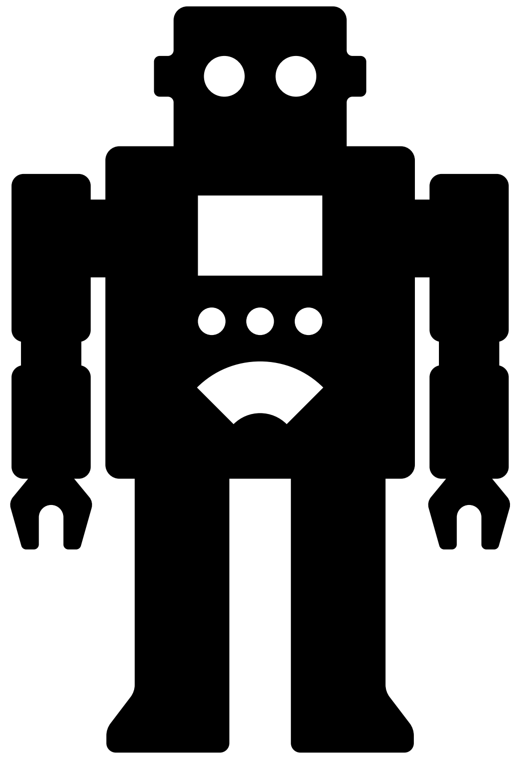
Two major types of uncertainty in RL

- Aleatoric Uncertainty & Epistemic Uncertainty:
  - Many problems have both forms of uncertainty
  - Example:

Questions?

Gaussian Processes & RL

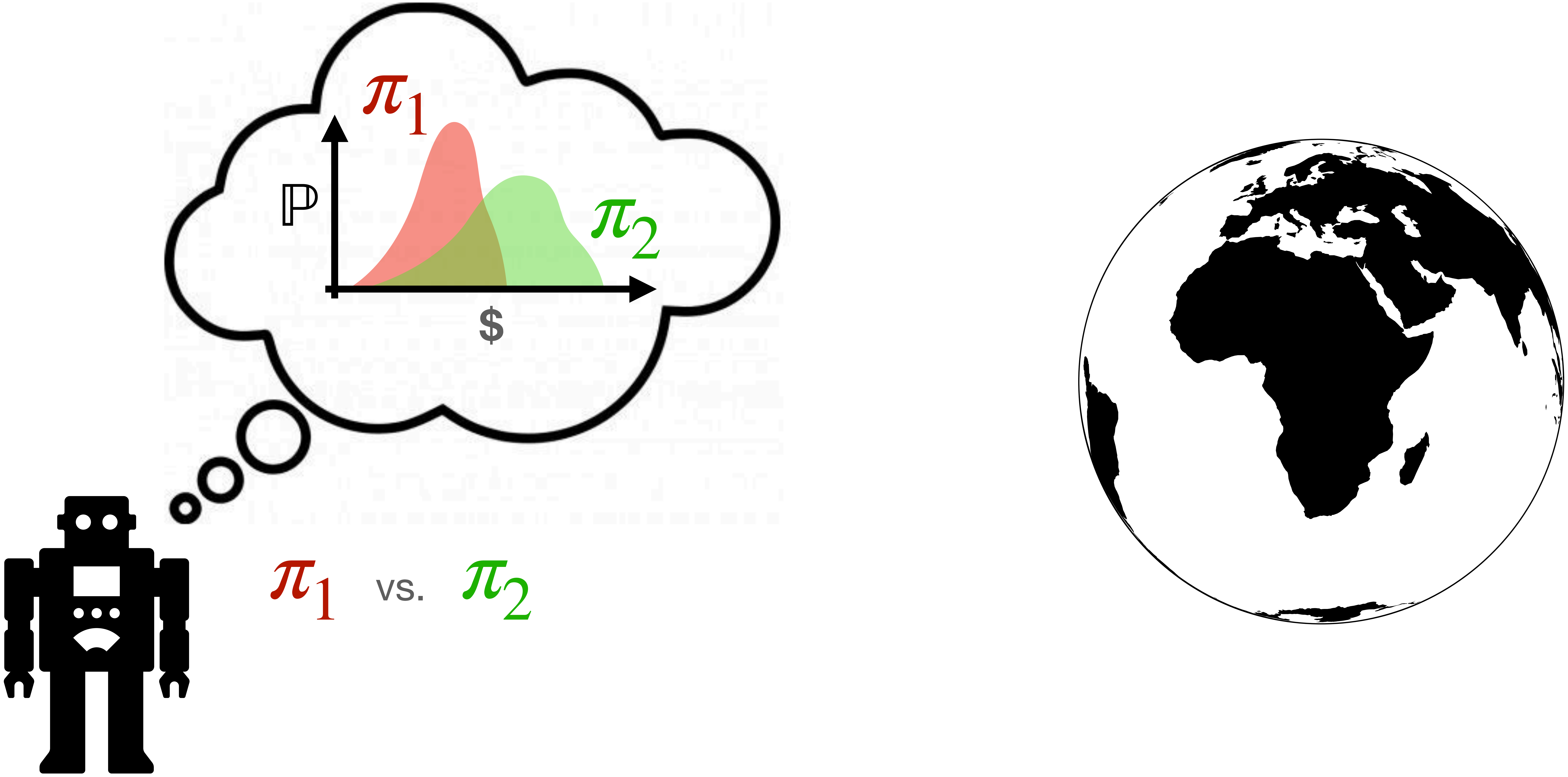
GP-based active exploration algorithms for simple continuous problem



π_1 vs. π_2

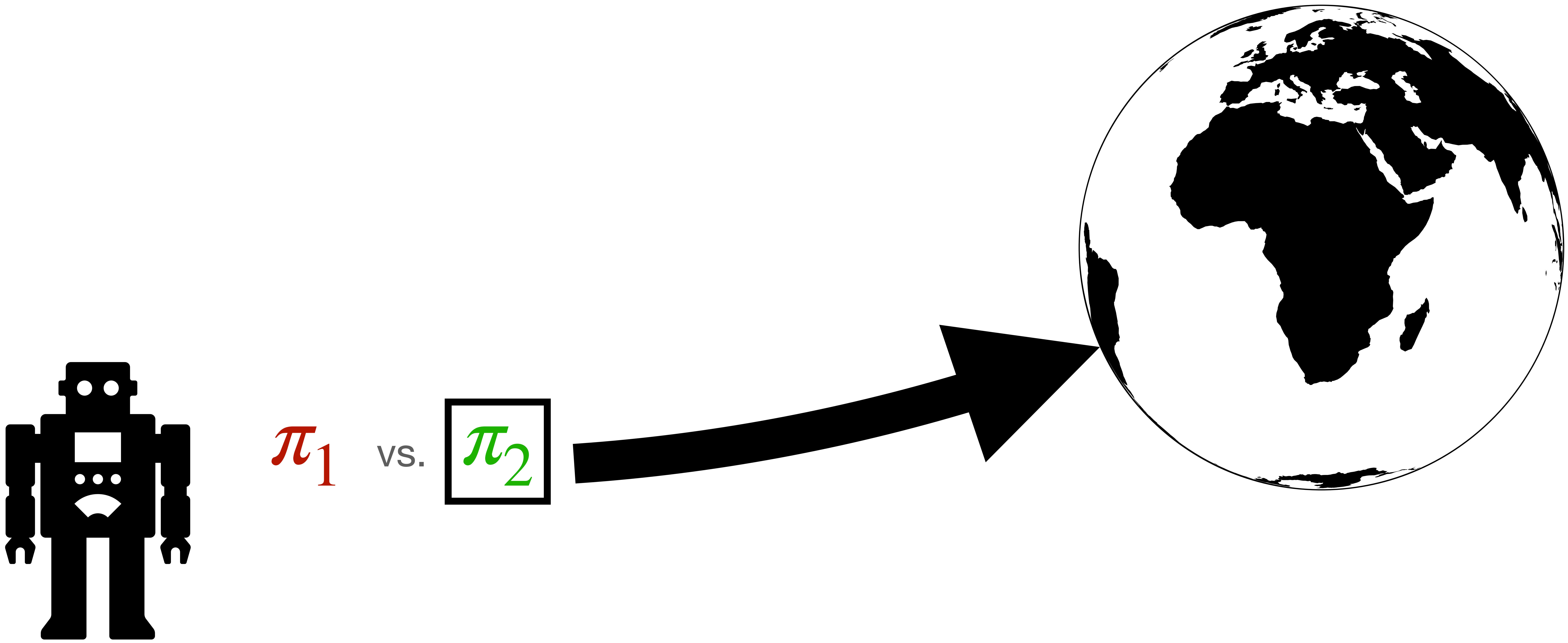
Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem



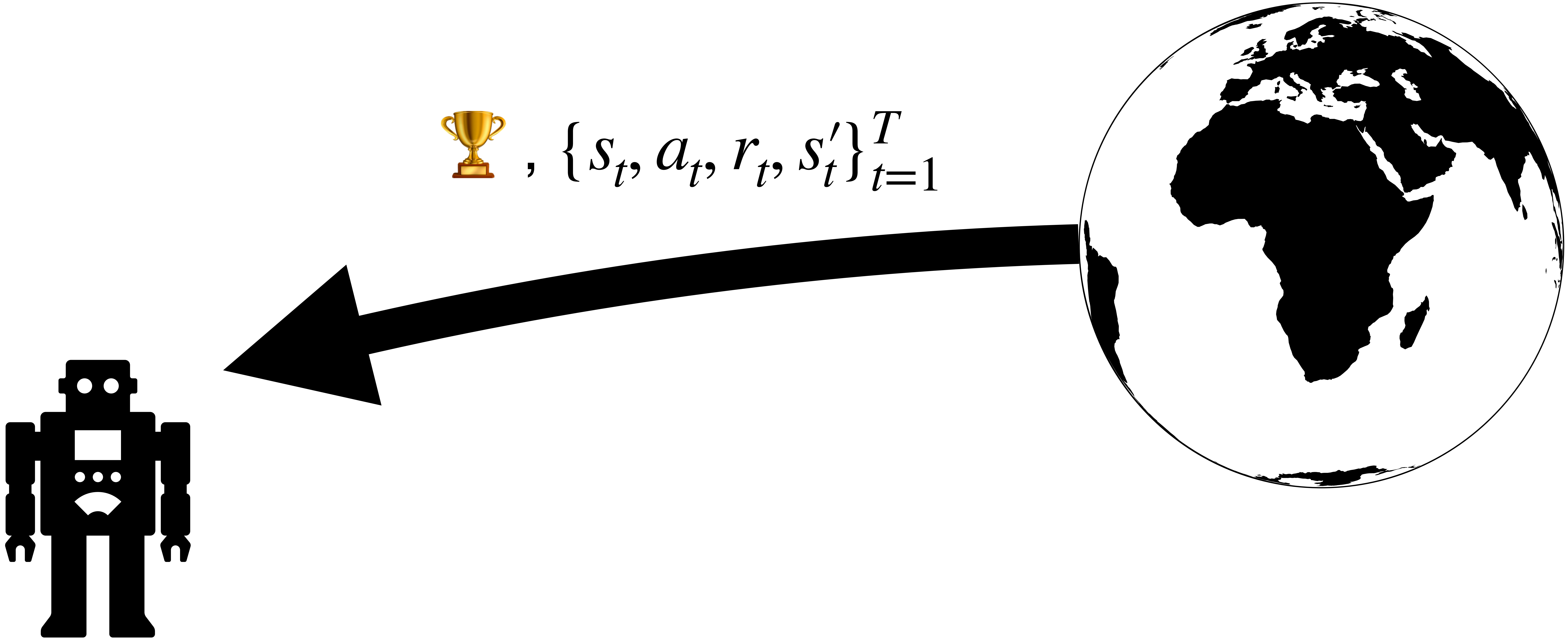
Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem



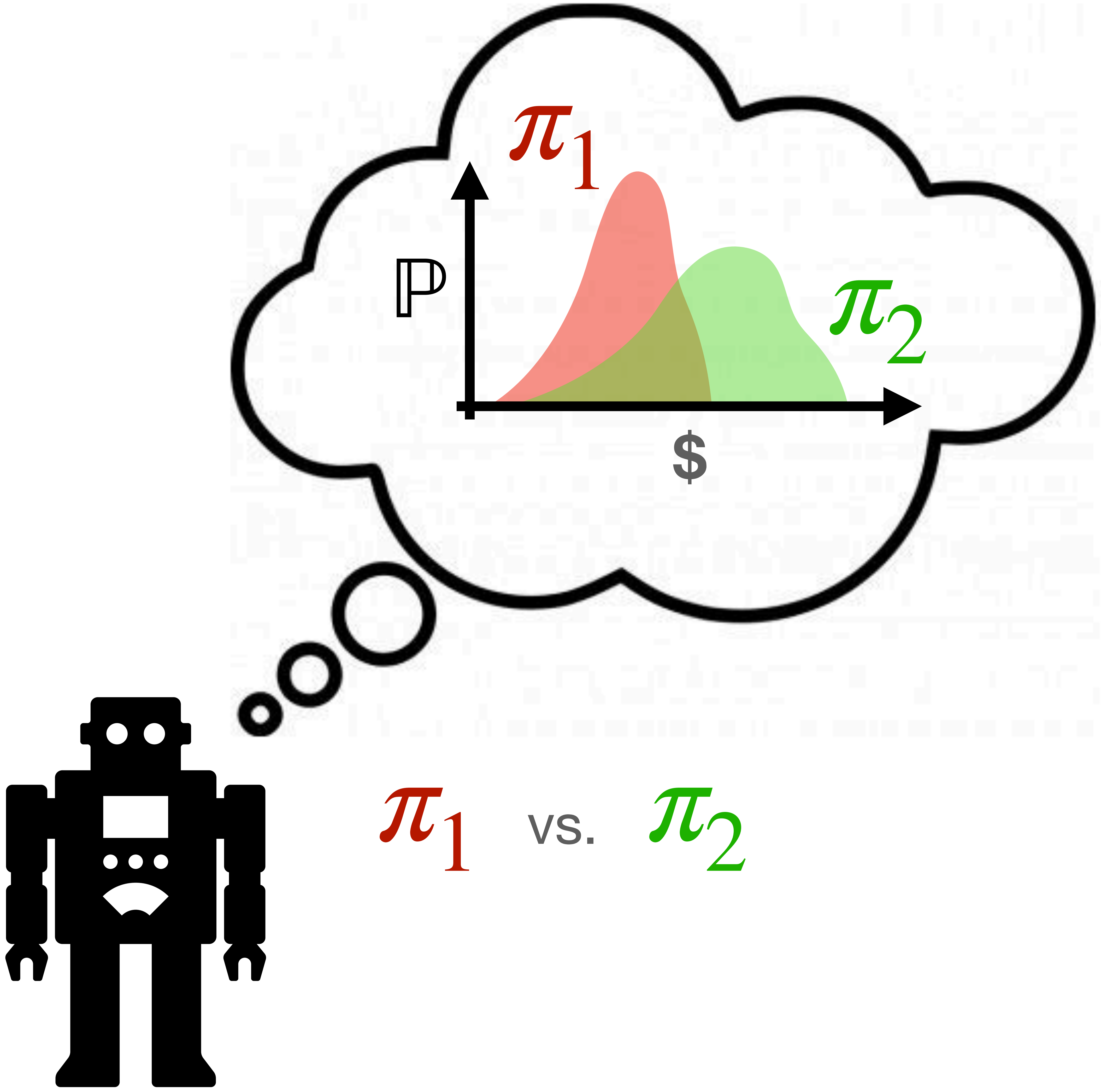
Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem



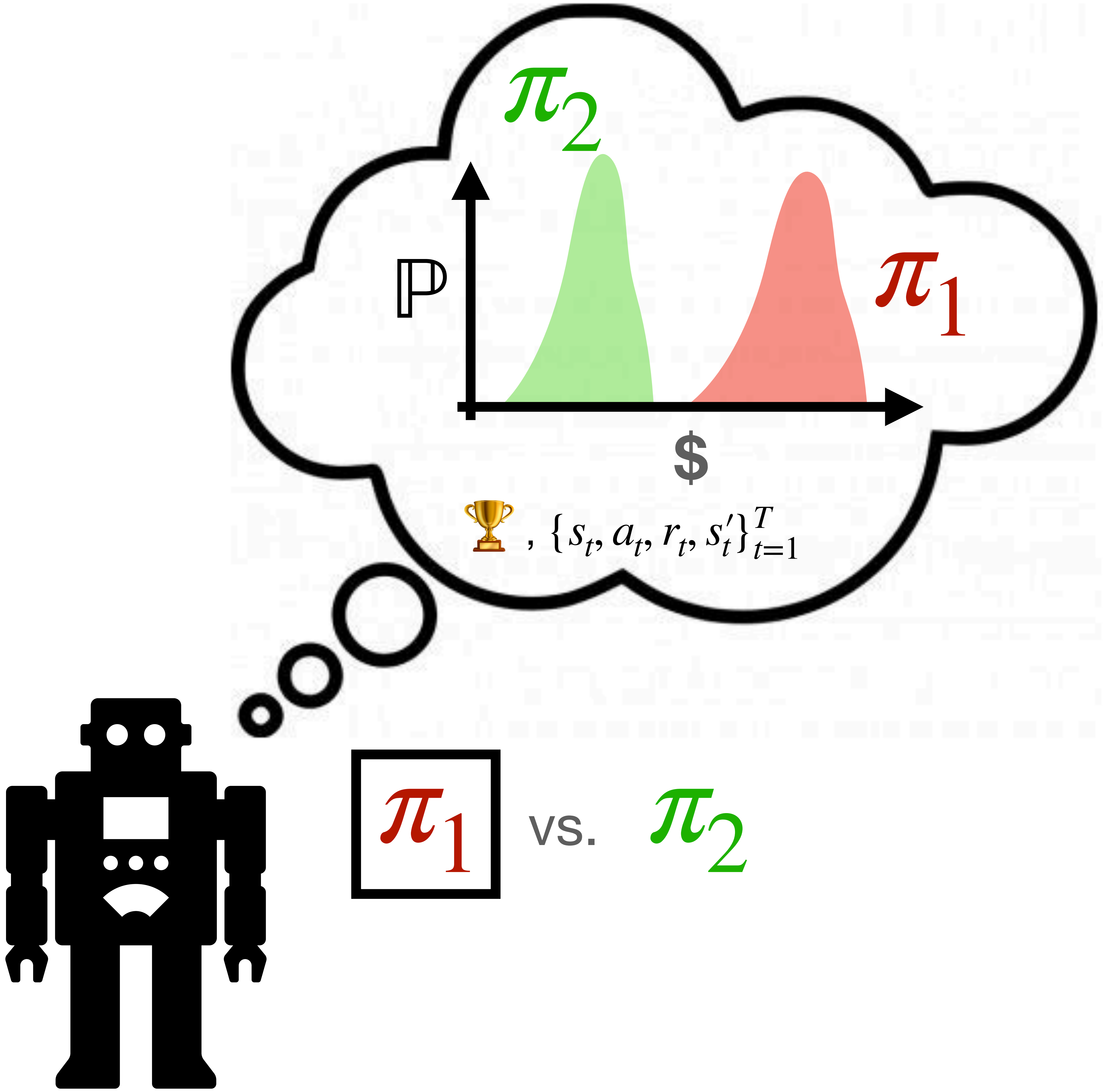
Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem



Gaussian Processes & RL

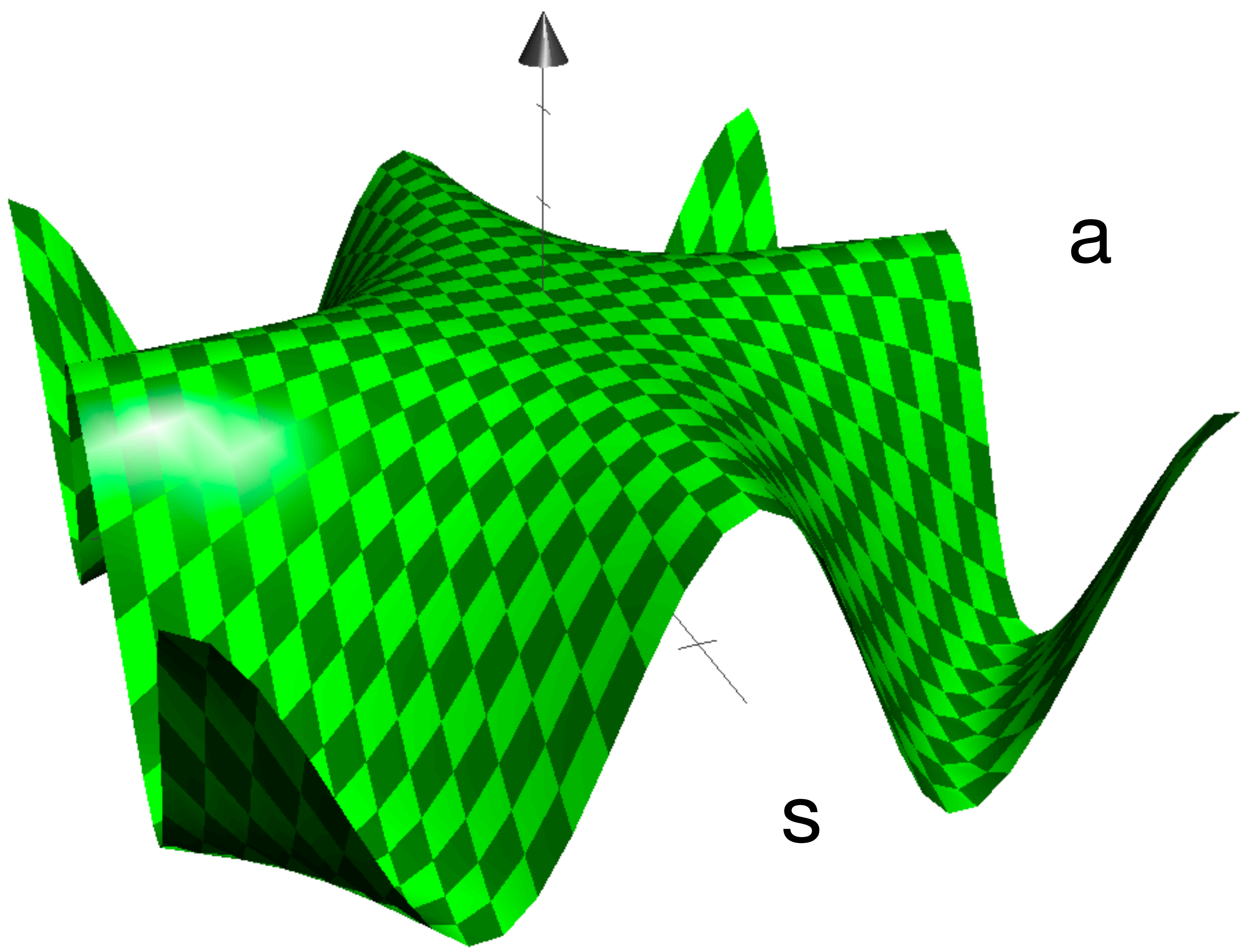
GP-based active exploration algorithms for simple continuous problem



Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

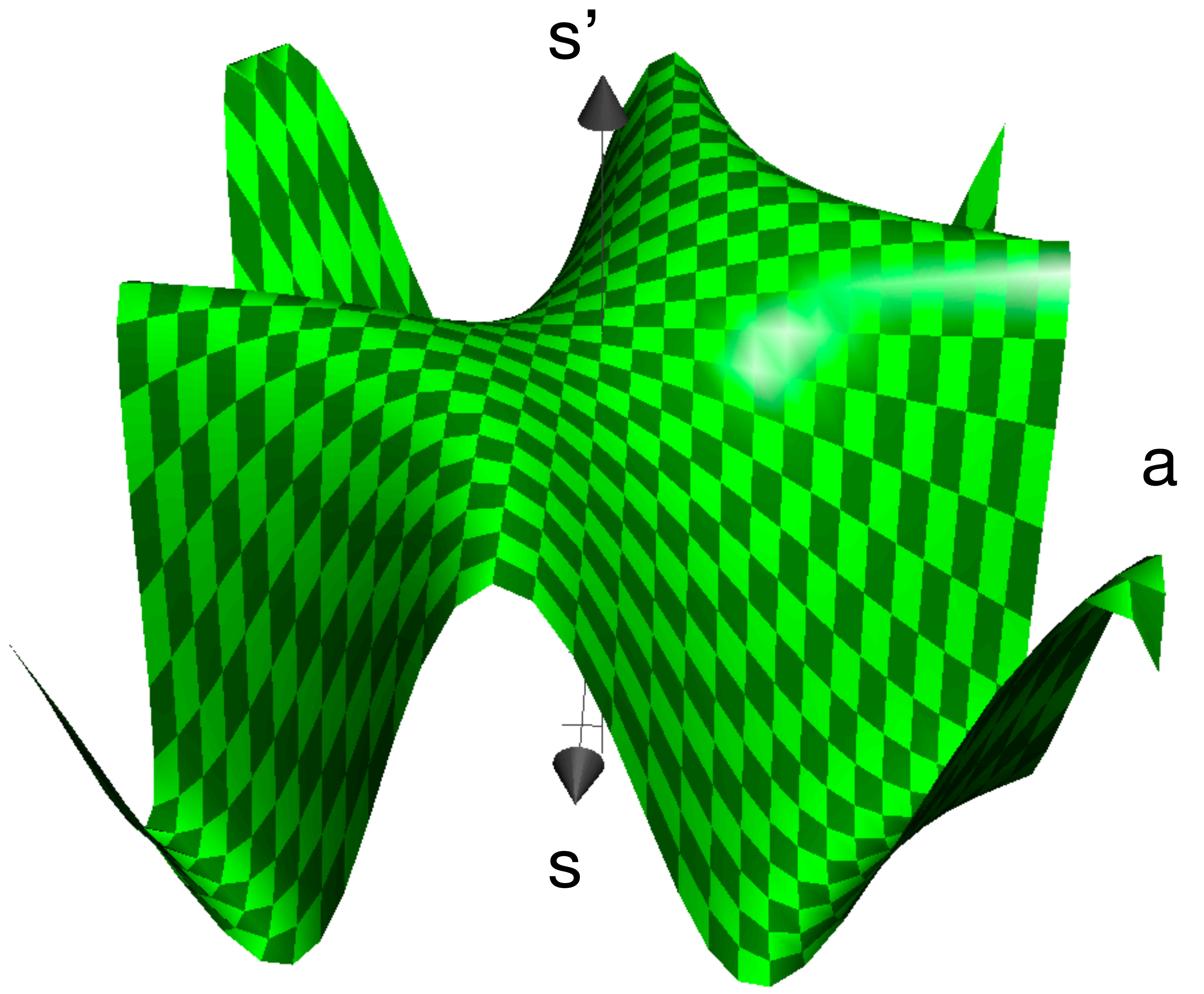
- Using GPs to model epistemic uncertainty
  - $T(s, a) = \cos(sa)$
  - $r(s, a) = -s^2$



Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

- Initialise $\mathscr{D} = \varnothing$
- For i in $[E]$:
  1. Use CMA-ES / CEM / NES to approximately solve
  $$\max_{a_1^i, \ldots, a_T^i} \mathbb{E} \left\{ \sum_{t=1}^{T} r(s_t^i, a_t^i) \middle| \mathscr{D} \right\}$$
  2. Deploy action sequence in true environment, observe R^i and add $\{s_t^i, a_t^i, s_{t+1}^i\}_{t=0}^{T}$ to $\mathscr{D}$

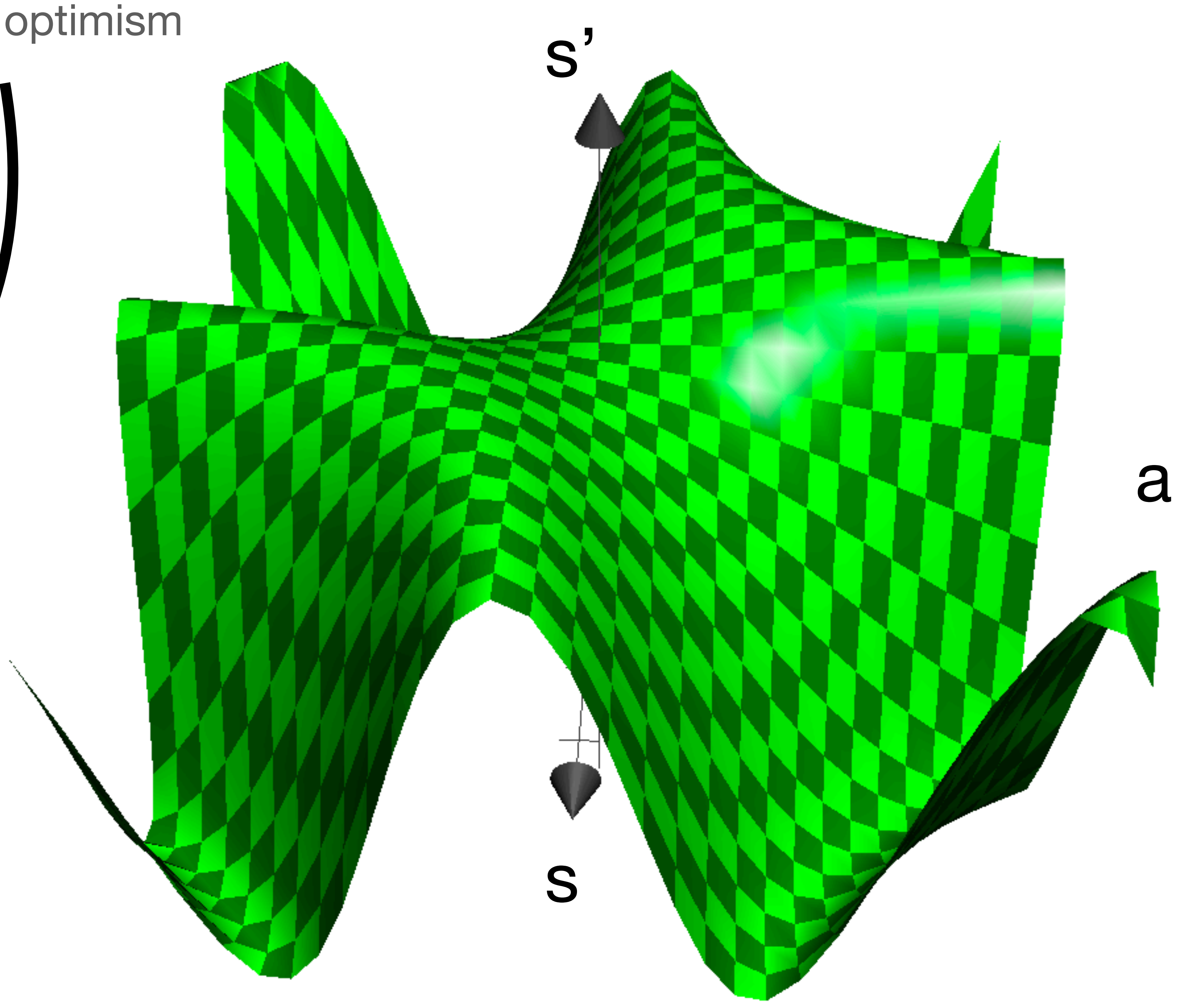


Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

- Initialise $\mathscr{D} = \emptyset$, $R^* = R_{lower}$, $A^* = \emptyset$
- For i in $[E]$:
  1. Use CMA-ES / CEM / NES to approximately solve
  $$\max_{a_1^i,\ldots,a_T^i} \mathbb{E}\left\{ \max\left(\left(\sum_{t=1}^{T} r(s_t^i, a_t^i) \right) - R^* , 0 \right) \middle| \mathscr{D} \right\}$$
  2. Deploy action sequence in true environment, observe R^i and add $\{s_t^i, a_t^i, s_{t+1}^i\}_{t=0}^{T}$ to $\mathscr{D}$
  3. If $R^i > R^*$, set $R^* = R^i$, $A^* = \{a_1^i, \ldots, a_T^i\}$

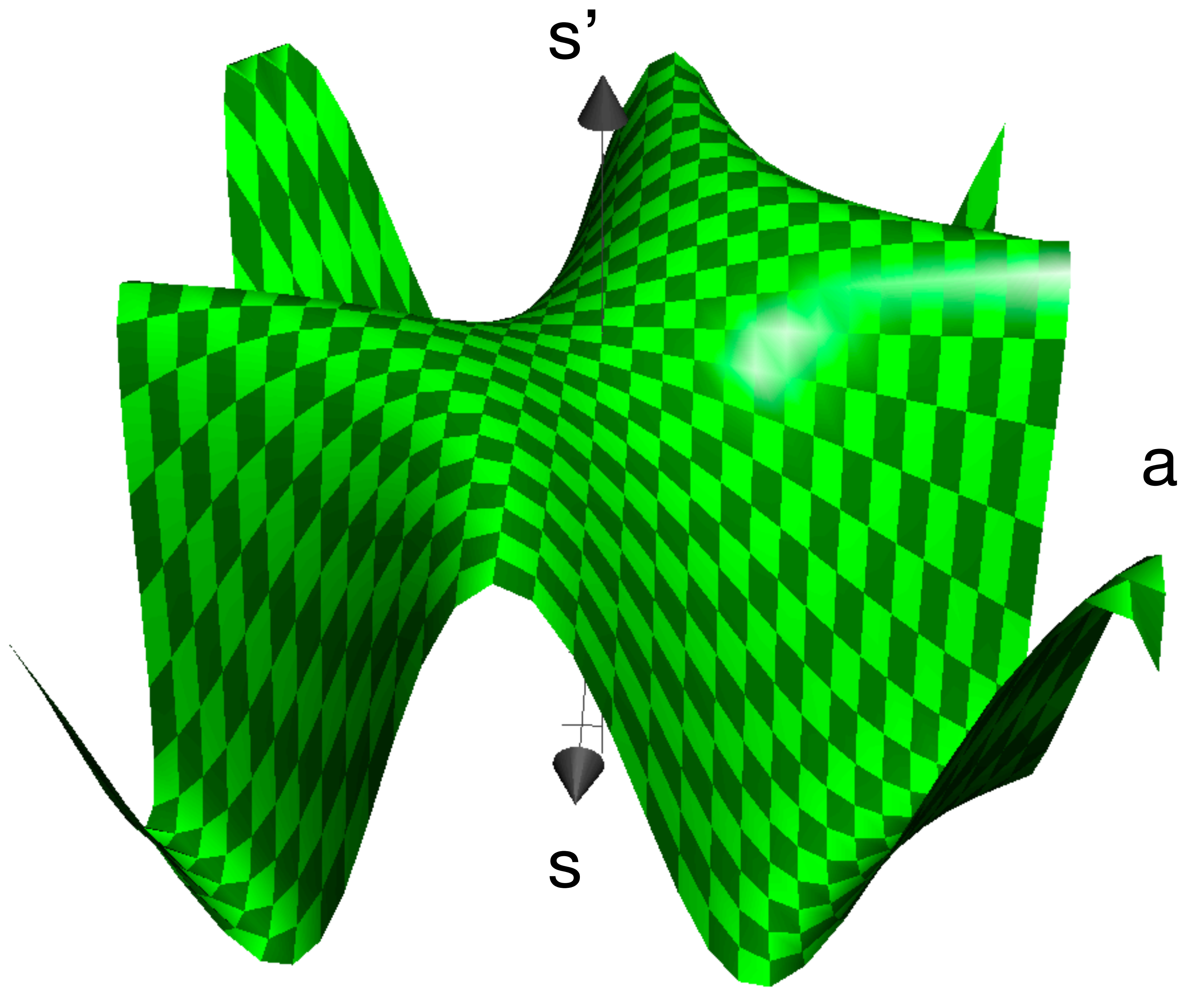
"Expected Improvement", encourages more optimism



Gaussian Processes & RL

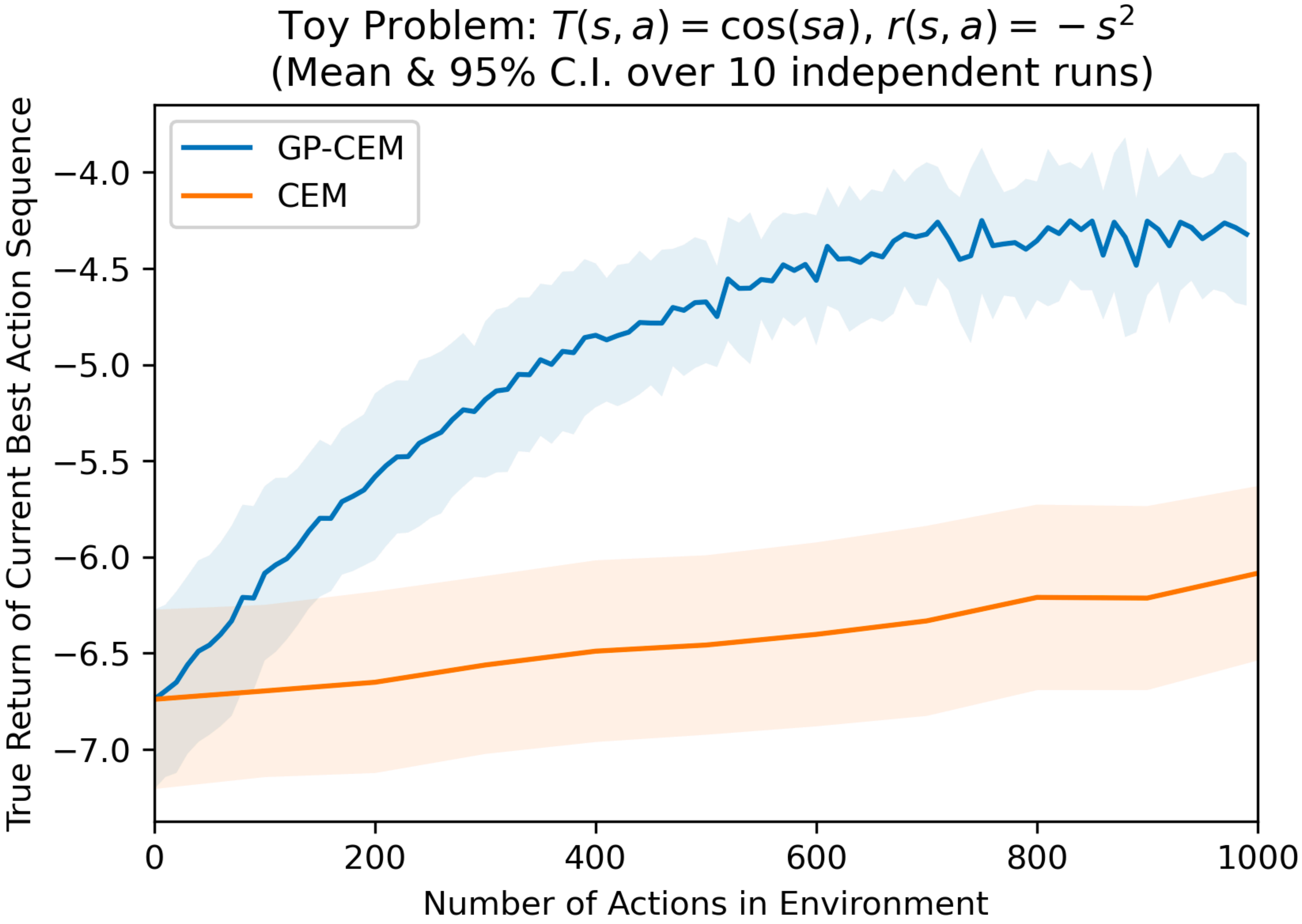
GP-based active exploration algorithms for simple continuous problem

- Classical RL approach would spend most of the early stage of learning by acting randomly
- Previous slides describe an approach that incorporates prior knowledge and explicitly leverages it for more intelligent exploration
  - Possible due to having a full probability distribution over Markov transition functions that permits tractable inference and sampling



Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem



Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

Gaussian Processes & RL

GP-based active exploration algorithms for simple continuous problem

- Many interesting algorithms with theoretical guarantees on sample complexity in the Bayesian Reinforcement Learning Literature
- Most (if not all) are inappropriate for interesting DRL problems
- (Optional reading, not part of the course)

Gaussian Processes & RL

GP limitations for RL

- Some problems make sense to use standard kernels (e.g. PILCO)

PILCO: A Model-Based and Data-Efficient Approach to Policy Search

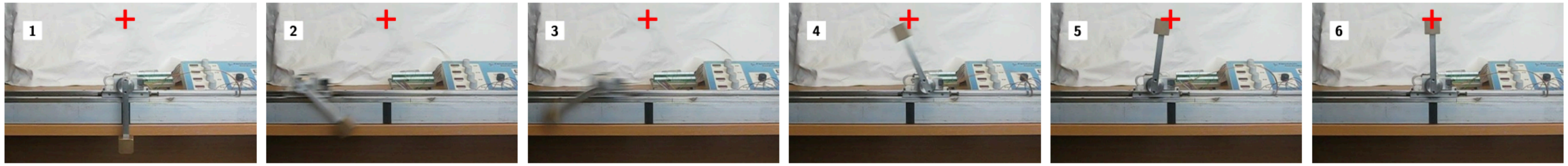


Figure 3. Real cart-pole system. Snapshots of a controlled trajectory of 20 s length after having learned the task. To solve the swing-up plus balancing, PILCO required only 17.5 s of interaction with the physical system.

- But for most problems of interest to DRL, hard to find appropriate kernel
- Computational complexity of inference is $\mathcal{O}(n^3)$ (matrix inversion)
- Very challenging to design differentiable policy / action-sequence optimisation techniques
- Designing multi-variate GPs is a big challenge (co-krigging), but is necessary for most interesting control problems

Gaussian Processes & RL

GP limitations for RL

- So why cover GPs?
  - Serves as a conceptual gold-standard to compare against, a rare setting where we can fully express epistemic uncertainty
    - Different approaches make different sacrifices to full representations of epistemic uncertainty (e.g. by only representing epistemic uncertainty at the marginal state-action level, or by avoiding a Bayesian treatment altogether)
  - Highlights how truly challenging it is to “solve” the full reinforcement learning problem

Questions?