

FCAI

fcai.fi

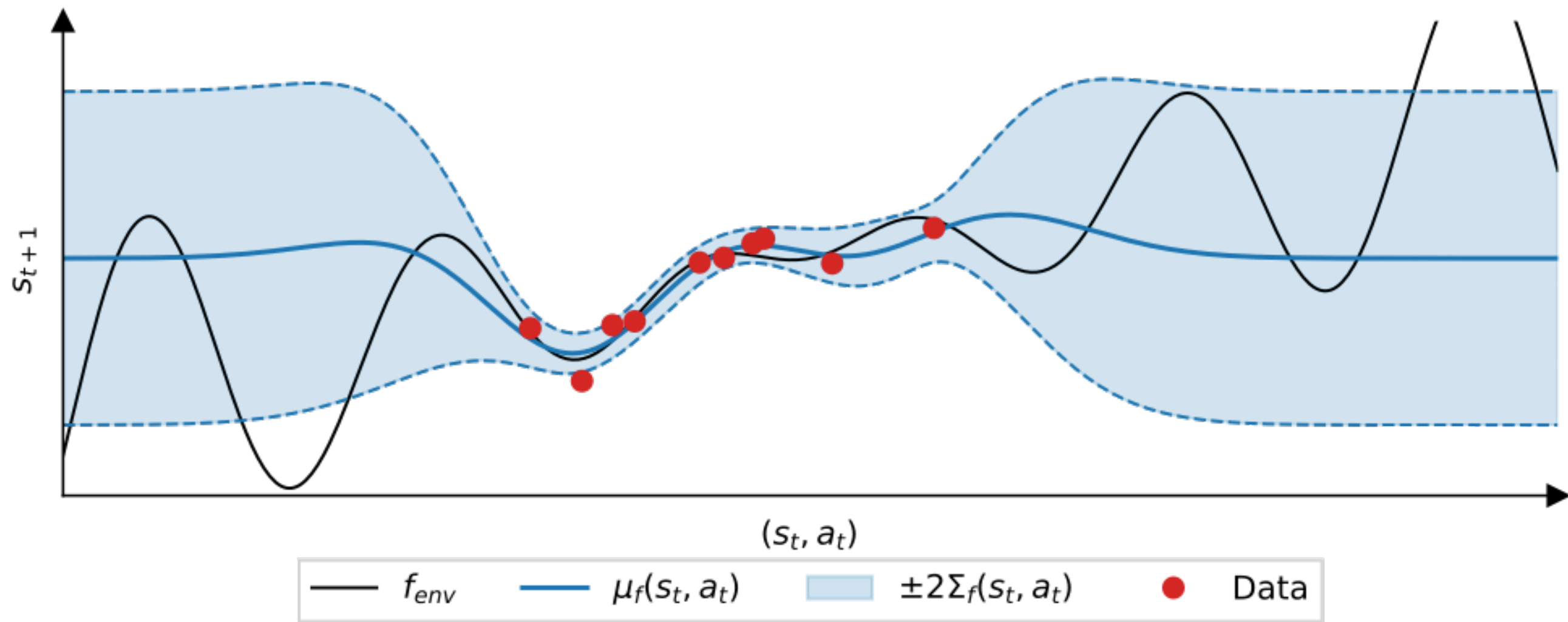
Exploration via Upper Confidence Bound

4

0

$$\pi^{UCB} = \arg \max_{\pi} \max_{f \in \mathcal{M}} J(\pi; f)$$

$$\mathcal{U} = \{f \mid \|f(s, a) - \mu_f(s, a)\| \leq \beta \Sigma_f(s, a)\}$$







$$f(s,a) \equiv \mu_f(s,a) - \beta \Sigma_f(s,a)$$

$$f(s,a) \equiv \mu_f(s,a) + \beta \Sigma_f(s,a)$$

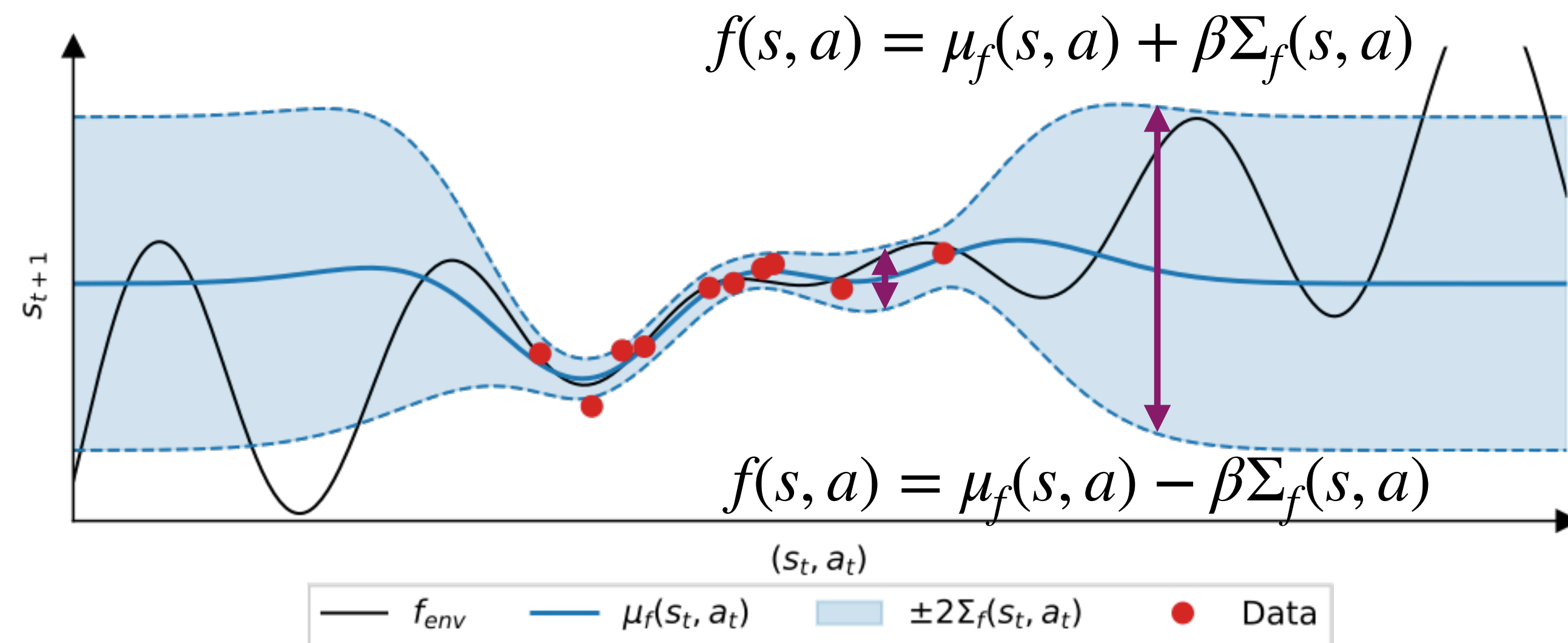
Curral (2020) Efficient Model-Based Reinforcement Learning through Optinistic Policy Search and Planning. NeurIPS

Extrahypersparameter β

Exploration via Upper Confidence Bound

$$\pi^{UCB} = \arg \max_{\pi} \max_{f \in \mathcal{M}} J(\pi; f)$$

$$\mathcal{M} = \{f \mid \|f(s, a) - \mu_f(s, a)\| \leq \beta \Sigma_f(s, a)\}$$



Extra hyperparameter β

How to Quantify Uncertainty in Dynamics?

