Learning convex bounds for linear quadratic control policy synthesis

Jack Umenberger   Thomas B. Schön

UPPSALA UNIVERSITET

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning →

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning →

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning →

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning

control
(stabilize the upright equilibrium position)

Learning to control

data
(observations of the dynamical system)

learning →

control
(stabilize the upright equilibrium position)

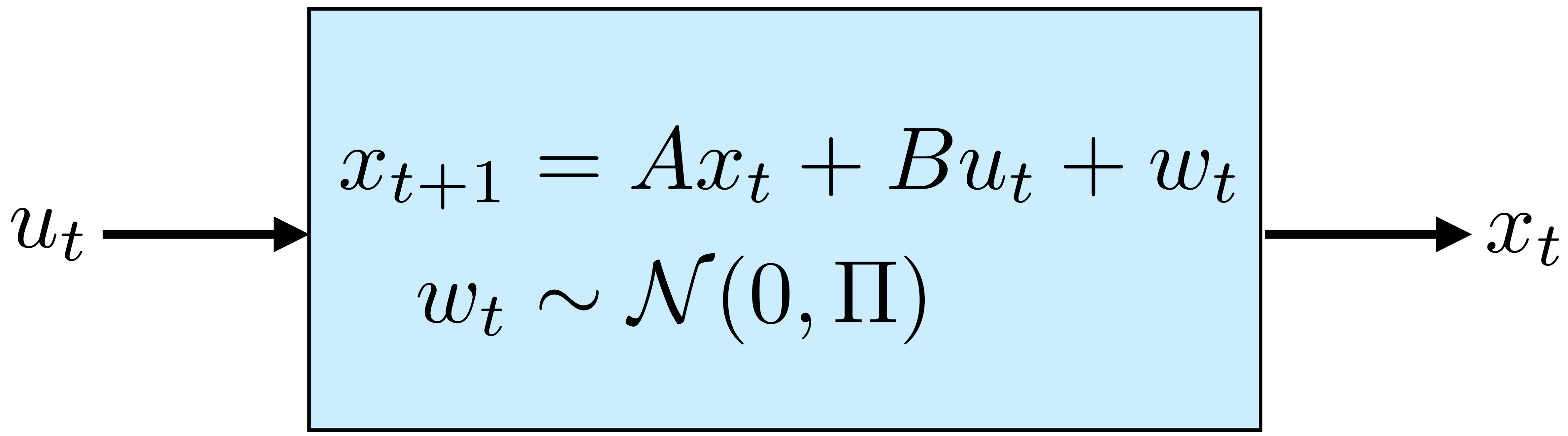
Learning to control

data
(observations of the dynamical system)

learning →

control
(stabilize the upright equilibrium position)

Problem set-up



Problem set-up

u_t →

$$x_{t+1} = Ax_t + Bu_t + w_t$$
$$w_t \sim \mathcal{N}(0, \Pi)$$

→ x_t

Goal: find a static state-feedback controller, $u = Kx$, to minimize

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \mathbb{E}\left[x_t' Q x_t + u_t' R u_t\right],$$

Problem set-up

$$u_t \rightarrow \boxed{\begin{array}{l} x_{t+1} = Ax_t + Bu_t + w_t \\ w_t \sim \mathcal{N}(0, \Pi), \end{array}} \rightarrow x_t$$

Goal: find a static state-feedback controller, *u = Kx,* to minimize

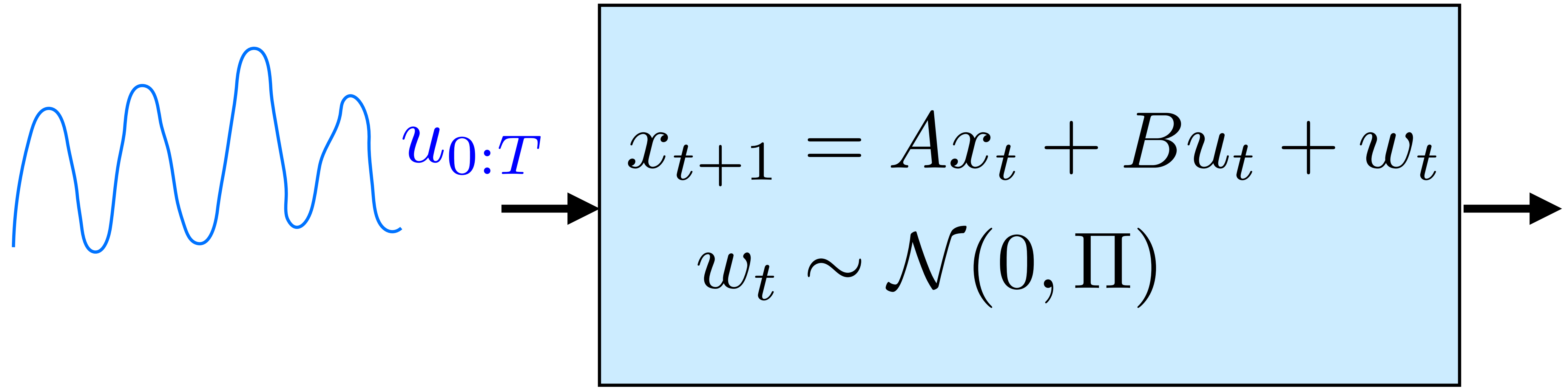
$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \mathbb{E}\left[x_t' Q x_t + u_t' R u_t\right],$$

Challenge: we don't know the system parameters $\theta = \{A, B, \Pi\}$

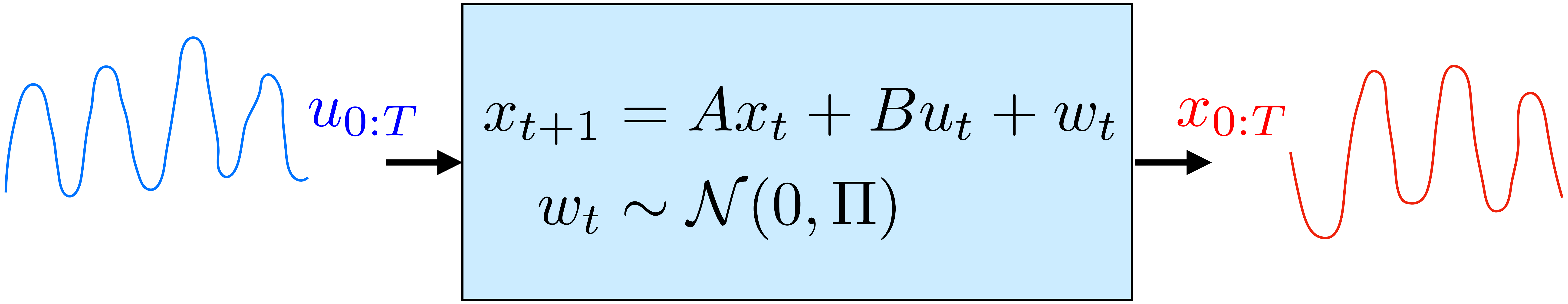
Learning from data

$$x_{t+1} = Ax_t + Bu_t + w_t$$
$$w_t \sim \mathcal{N}(0, \Pi)$$

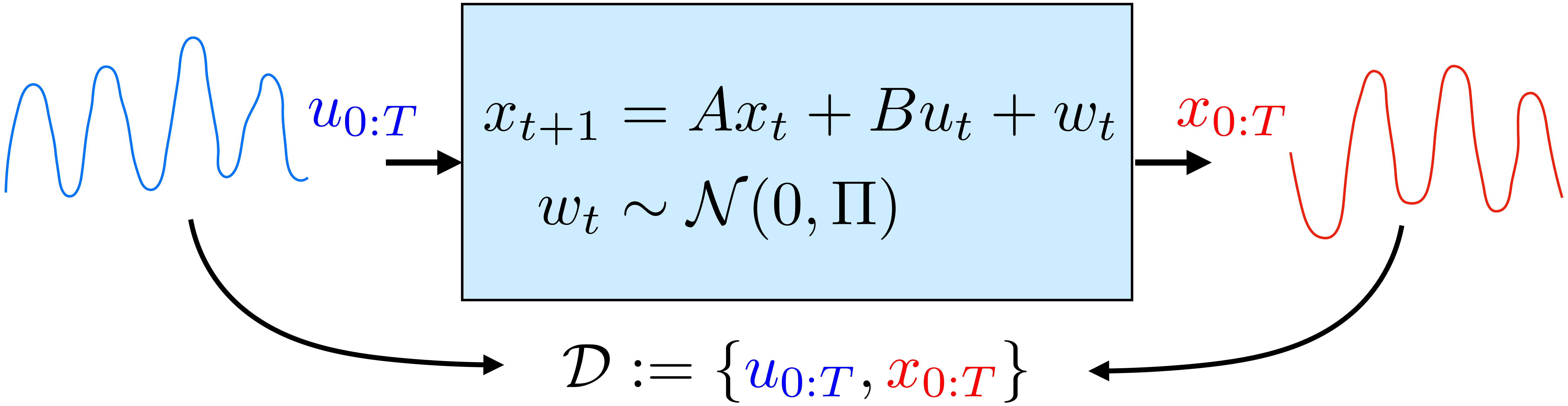
Learning from data



Learning from data

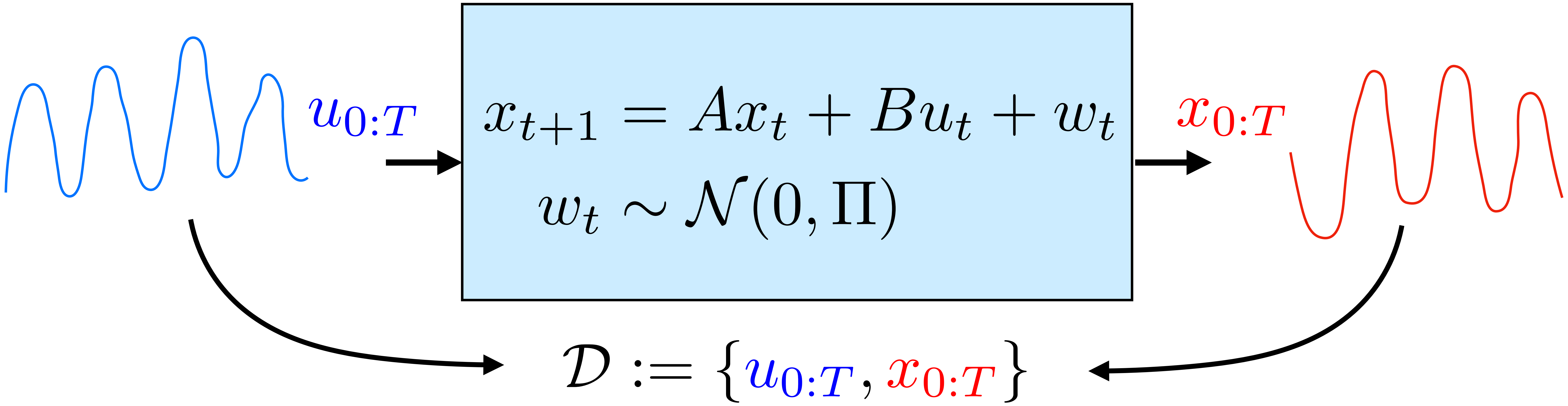


Learning from data



From this data we can form the **posterior** belief over model parameters: $\textbf{posterior}(\theta|\mathcal{D})$

Learning from data

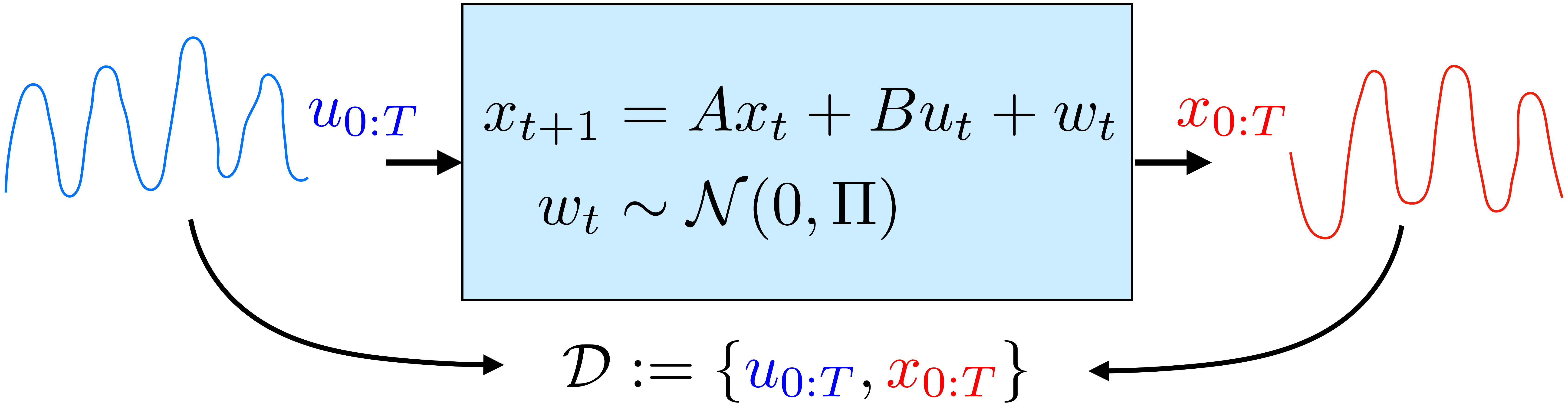


From this data we can form the **posterior** belief over model parameters: $\textbf{posterior}(\theta|\mathcal{D})$

Instead of optimizing the cost for fixed parameters

$$\textbf{cost}(K|\theta)$$

Learning from data



From this data we can form the **posterior** belief over model parameters: $\mathbf{posterior}(\theta|\mathcal{D})$

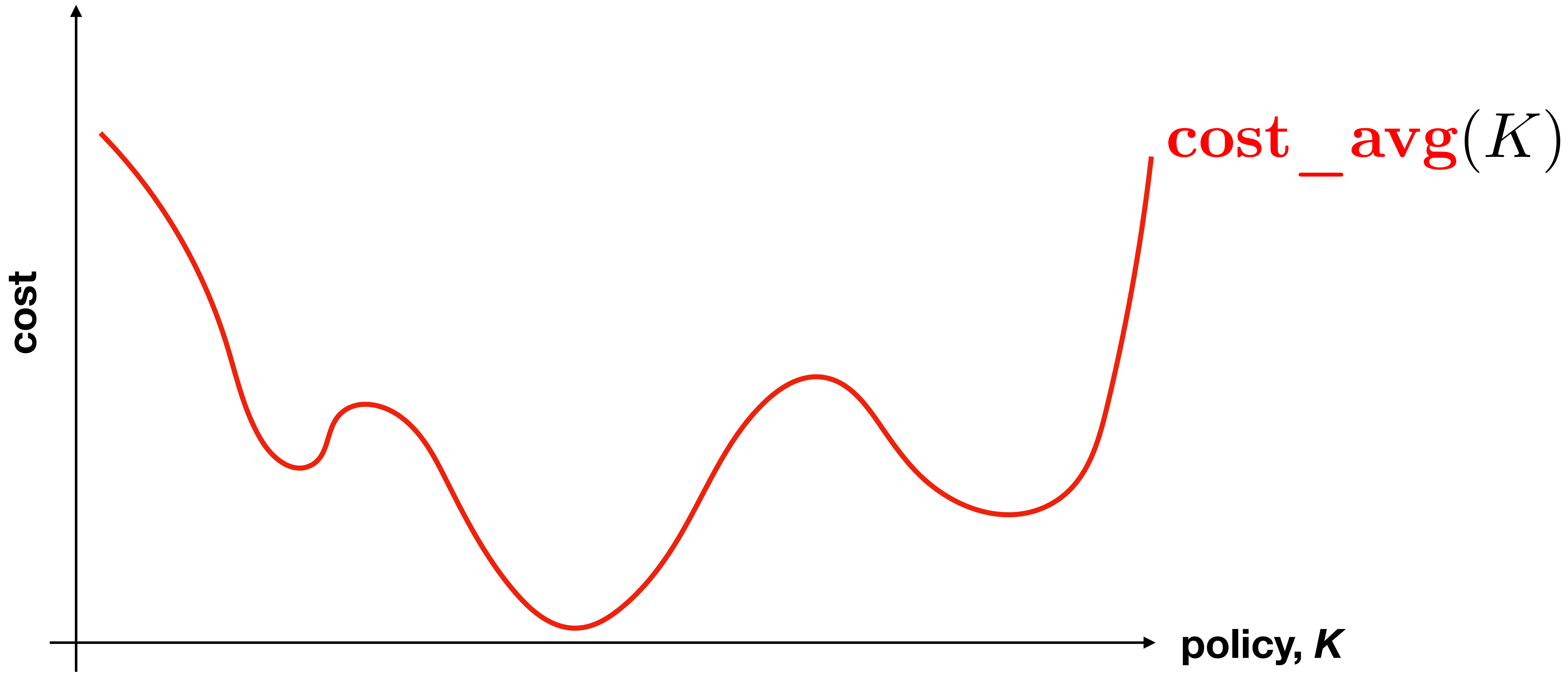
Instead of optimizing the cost for fixed parameters

$$\mathbf{cost}(K|\theta)$$

We can optimize the expected cost over the posterior

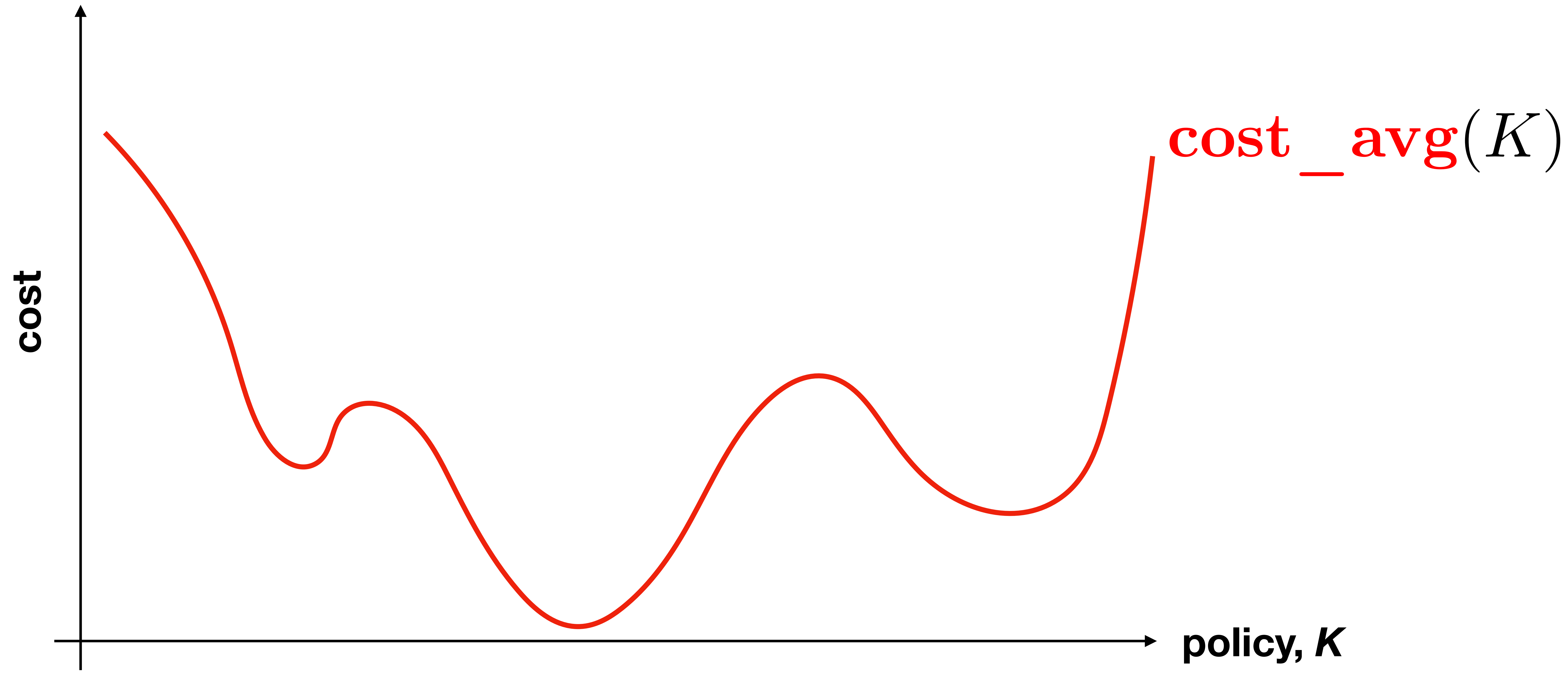
$$\mathbf{cost\_avg}(K) = \int \mathbf{cost}(K|\theta)\mathbf{posterior}(\theta|\mathcal{D})d\theta$$

Convex upper bounds



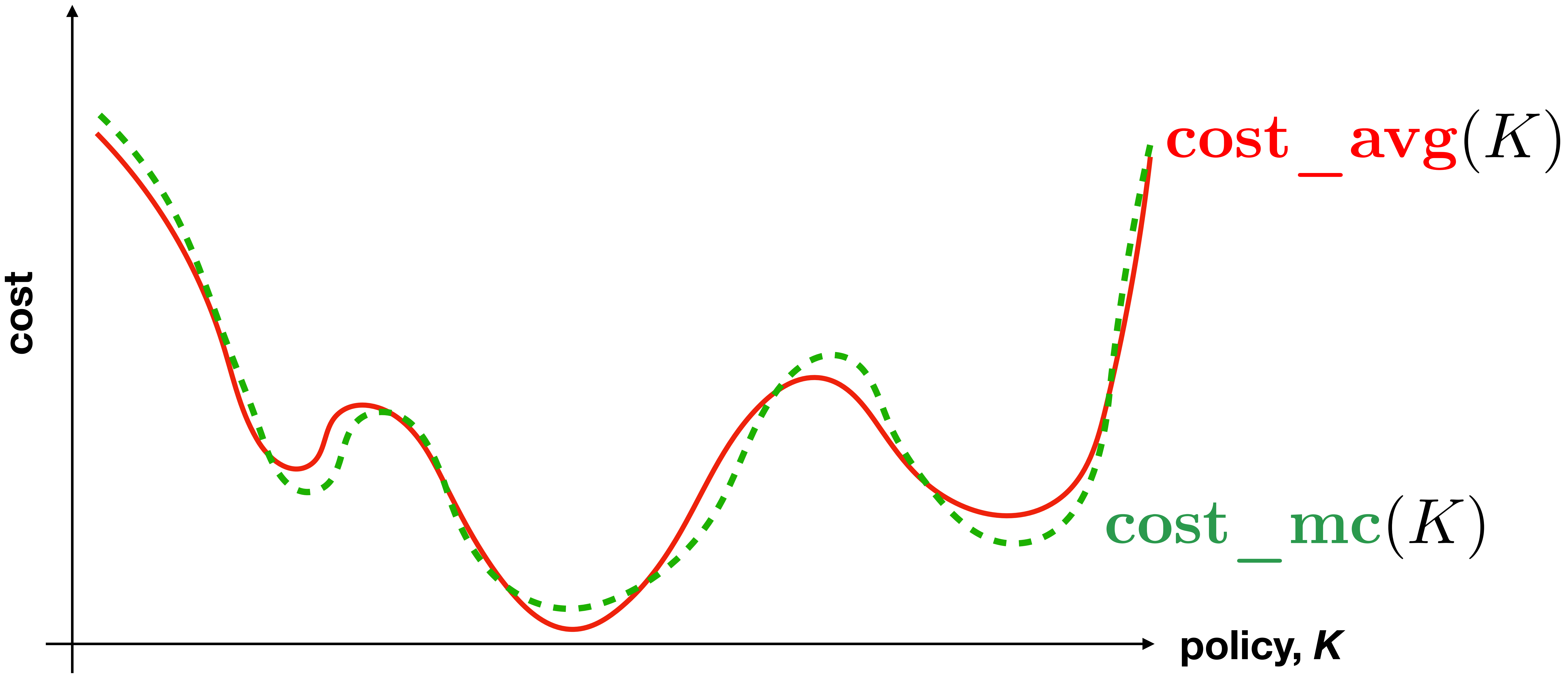
Convex upper bounds

$$\textbf{cost\_avg}(K) \approx \textbf{cost\_mc}(K) := \frac{1}{M}\sum_{i=1}^{M}\textbf{cost}(K|\theta_i) \qquad \theta_i \sim \textbf{posterior}(\theta|\mathcal{D})$$



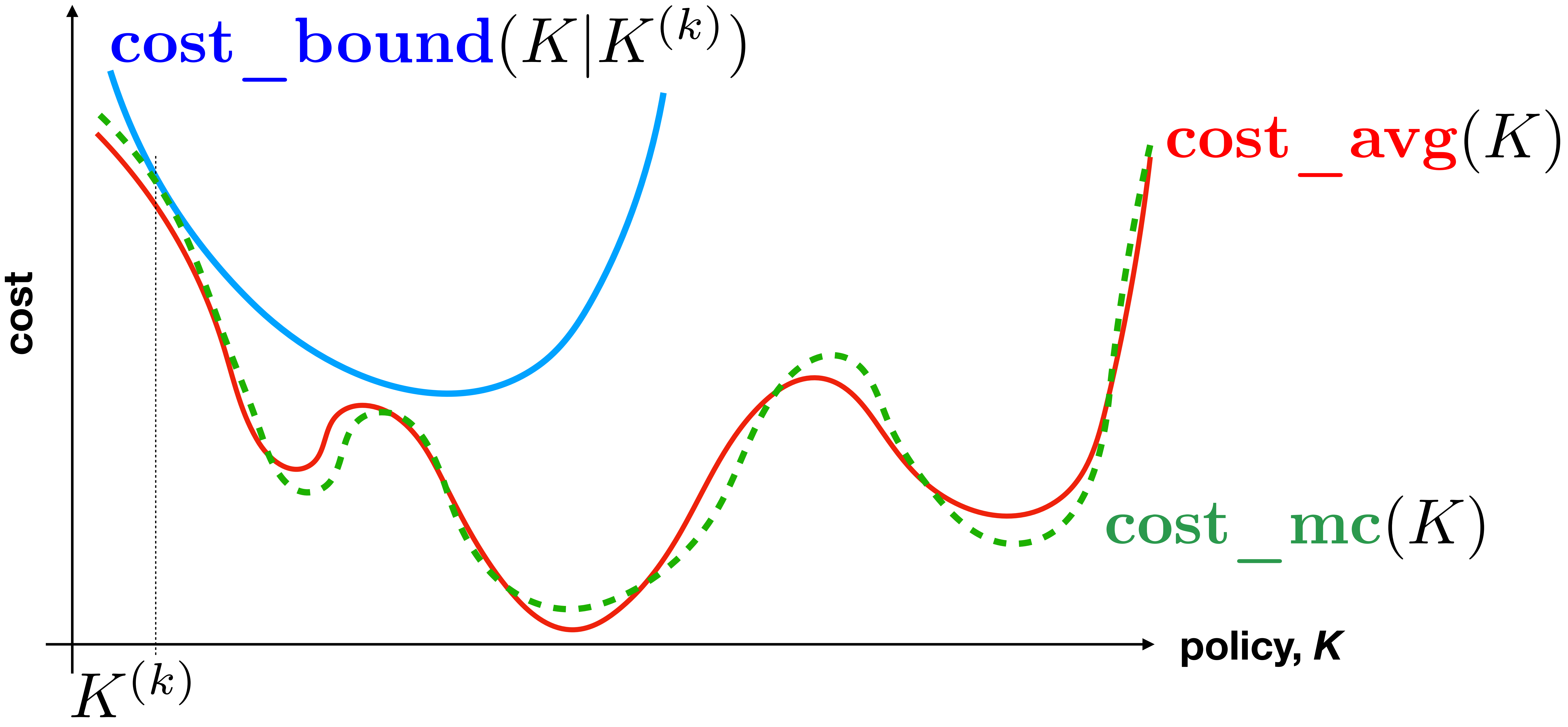
Convex upper bounds

$$\mathbf{cost\_avg}(K) \approx \mathbf{cost\_mc}(K) := \frac{1}{M}\sum_{i=1}^{M}\mathbf{cost}(K|\theta_i) \qquad \theta_i \sim \mathbf{posterior}(\theta|\mathcal{D})$$



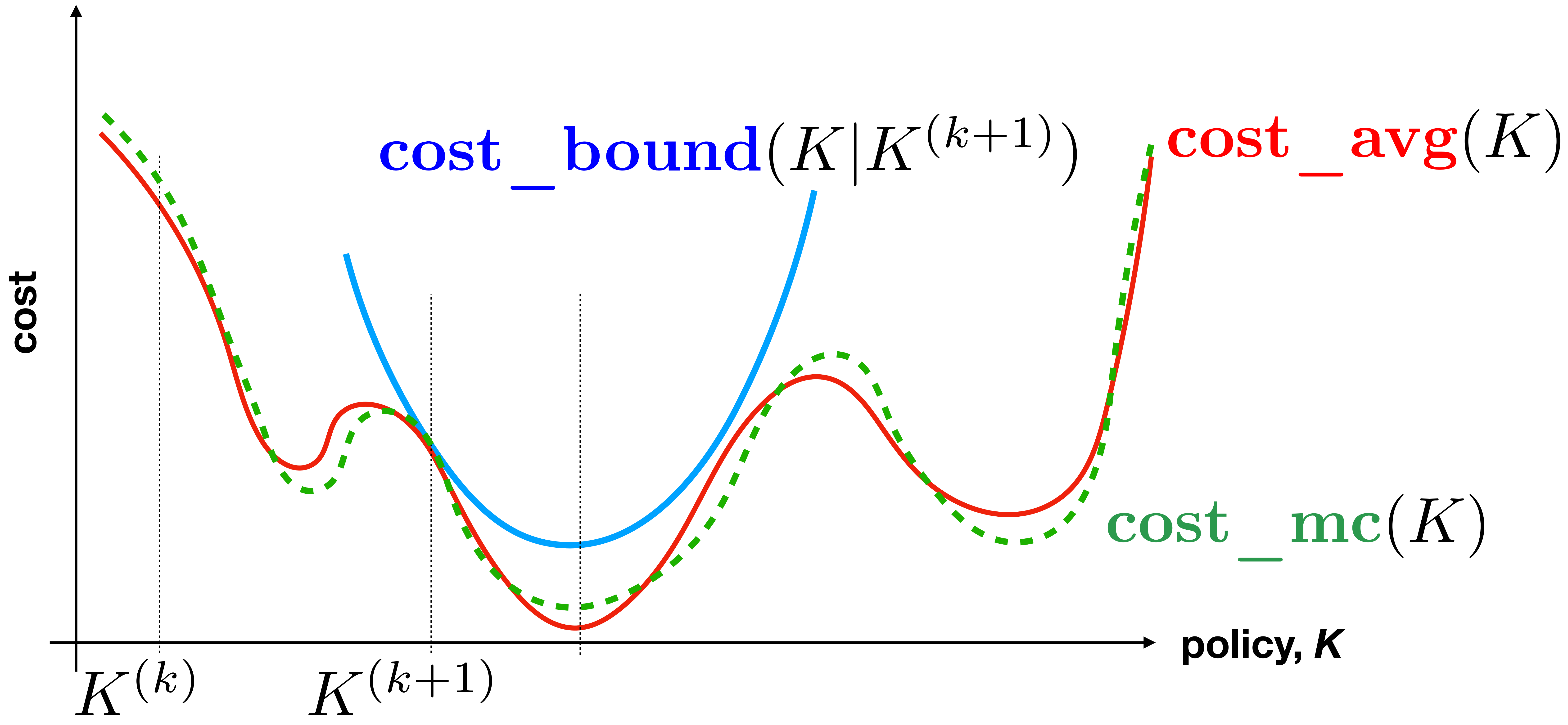
Convex upper bounds

$$\textbf{cost\_avg}(K) \approx \textbf{cost\_mc}(K) := \frac{1}{M}\sum_{i=1}^{M}\textbf{cost}(K|\theta_i) \qquad \theta_i \sim \textbf{posterior}(\theta|\mathcal{D})$$



Convex upper bounds

$$\mathbf{cost\_avg}(K) \approx \mathbf{cost\_mc}(K) := \frac{1}{M}\sum_{i=1}^{M}\mathbf{cost}(K|\theta_i) \qquad \theta_i \sim \mathbf{posterior}(\theta|\mathcal{D})$$



Convexification

The crux of the problem is the matrix inequality

$$\begin{bmatrix} X_i - Q & (A_i + B_i K)' & K' \\ A_i + B_i K & X_i^{-1} & 0 \\ K & 0 & R^{-1} \end{bmatrix} \succeq 0$$

known quantities

decision variables

Convexification

The crux of the problem is the matrix inequality

$$\begin{bmatrix} X_i - Q & (A_i + B_i K)' & K' \\ A_i + B_i K & \boxed{X_i^{-1}} & 0 \\ K & 0 & R^{-1} \end{bmatrix} \succeq 0$$

known quantities

decision variables

Convexification

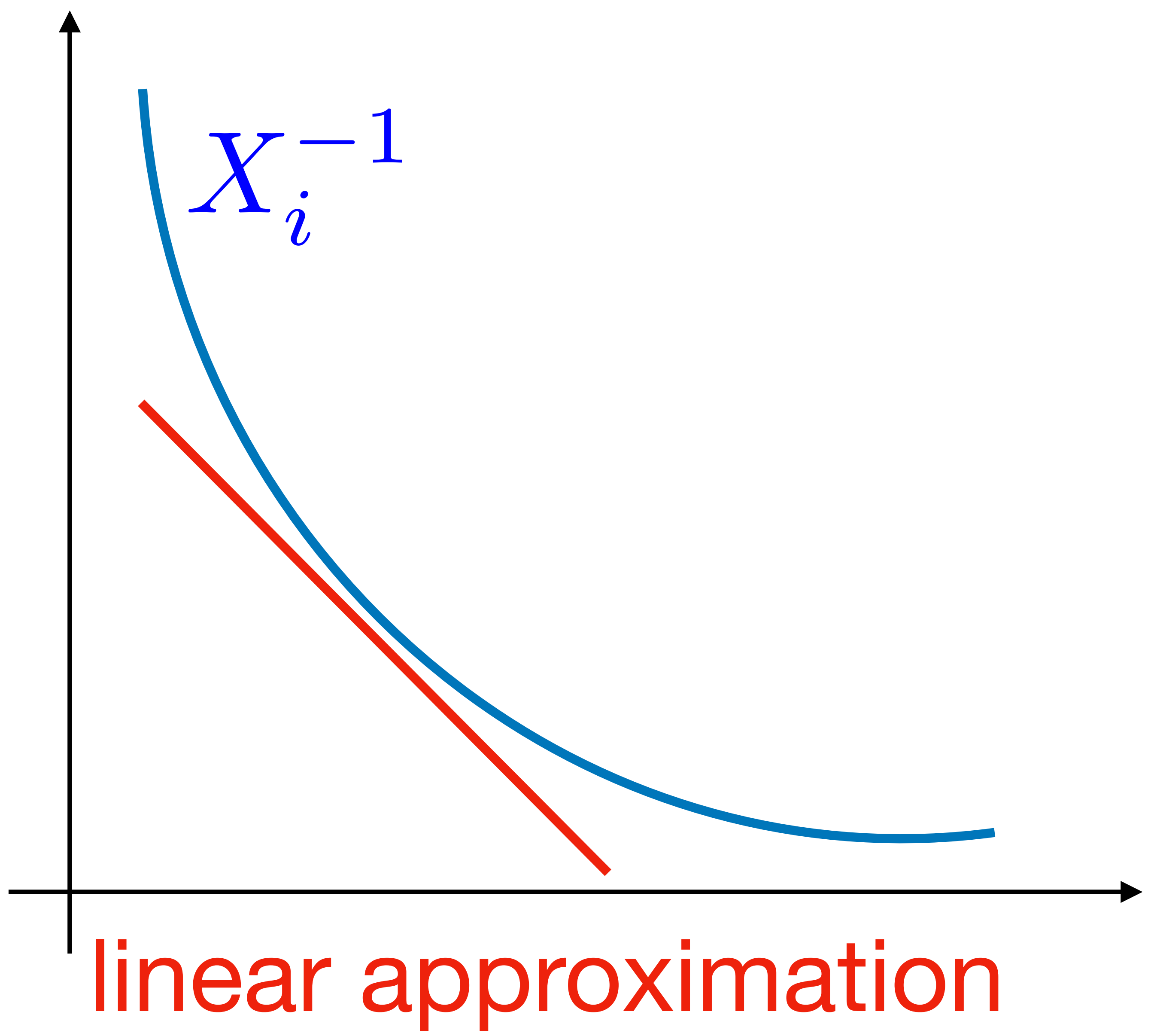
The crux of the problem is the matrix inequality

$$\begin{bmatrix} X_i - Q & (A_i + B_i K)' & K' \\ A_i + B_i K & X_i^{-1} & 0 \\ K & 0 & R^{-1} \end{bmatrix} \succeq 0$$

known quantities

decision variables

- Replace the 'problematic' term with a Taylor series approx.



Convexification

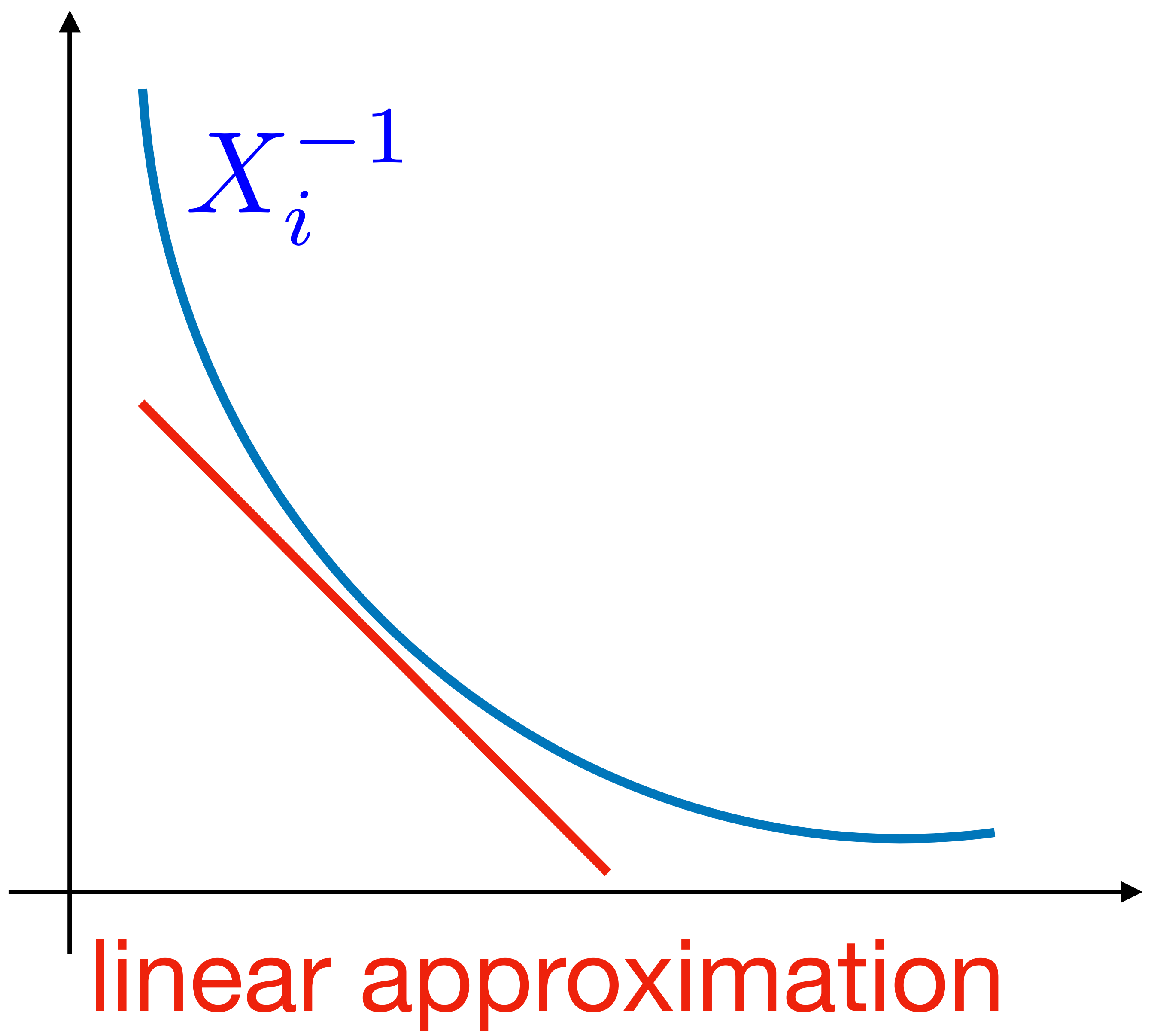
The crux of the problem is the matrix inequality

$$\begin{bmatrix} X_i - Q & (A_i + B_i K)' & K' \\ A_i + B_i K & X_i^{-1} & 0 \\ K & 0 & R^{-1} \end{bmatrix} \succeq 0$$

known quantities

decision variables

- Replace the 'problematic' term with a Taylor series approx.
- Leads to a new linear matrix inequality with a smaller feasible set.



Convexification

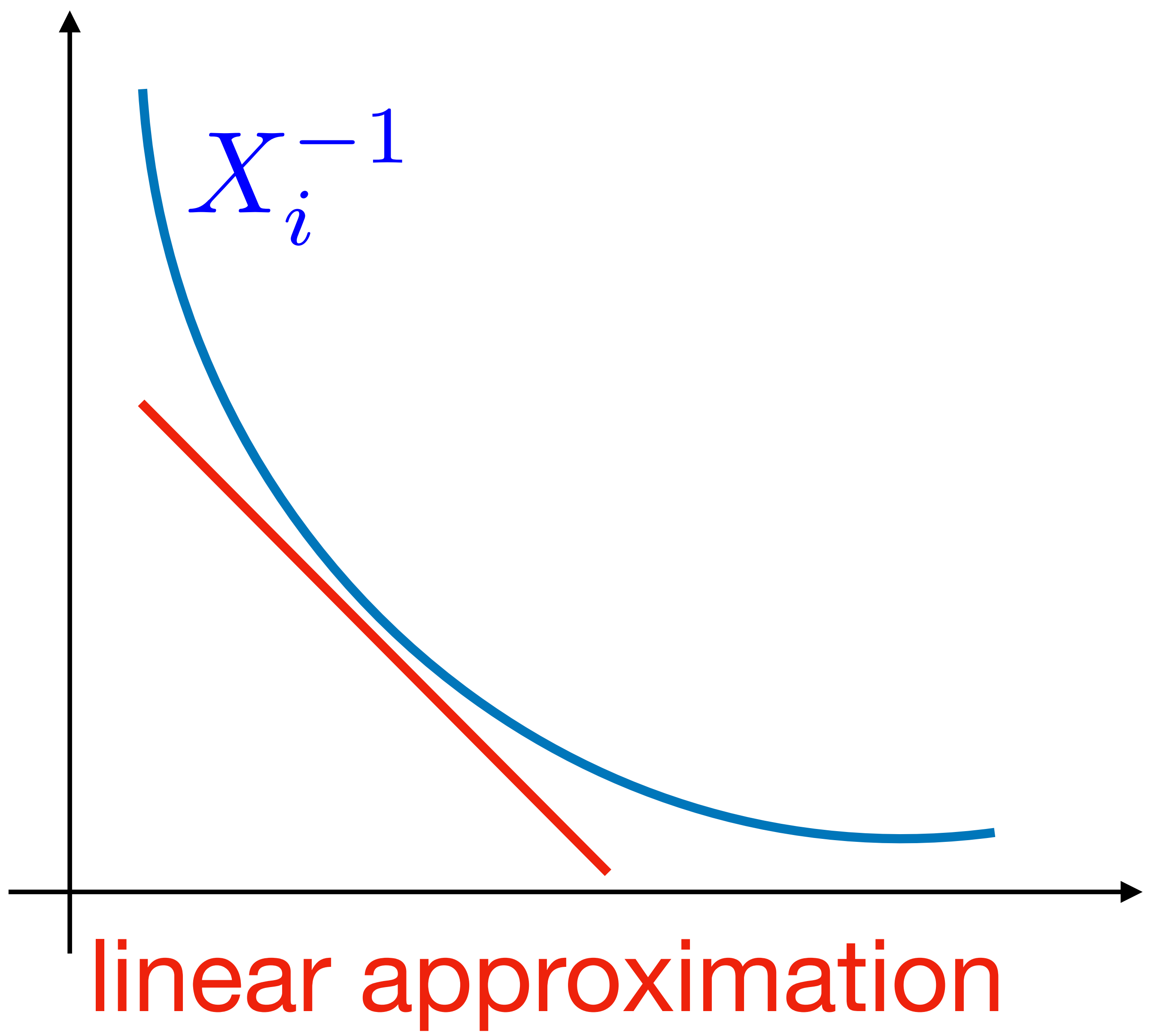
The crux of the problem is the matrix inequality

$$\begin{bmatrix} X_i - Q & (A_i + B_i K)' & K' \\ A_i + B_i K & X_i^{-1} & 0 \\ K & 0 & R^{-1} \end{bmatrix} \succeq 0$$

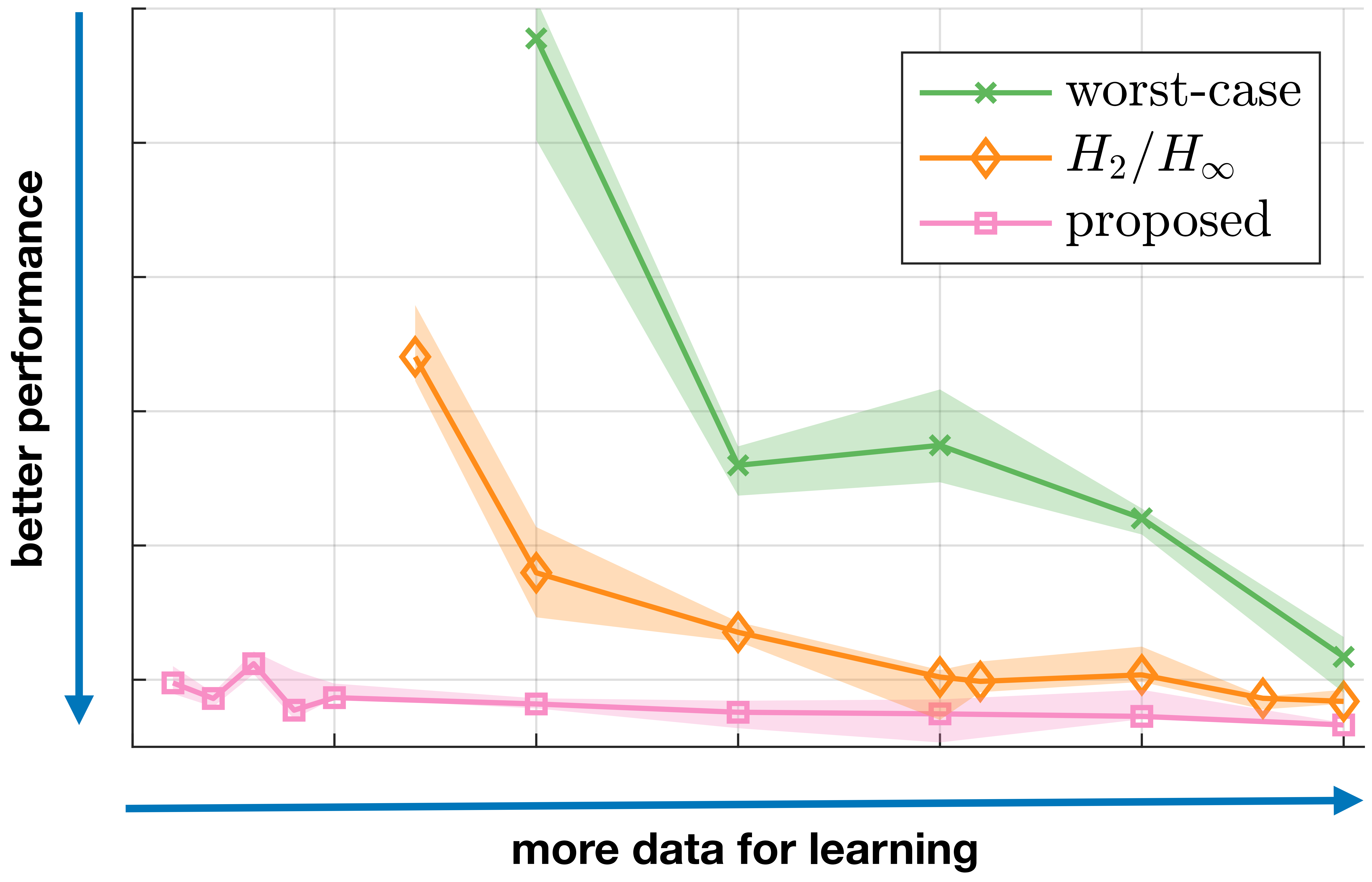
known quantities

decision variables

- Replace the 'problematic' term with a Taylor series approx.
- Leads to a new linear matrix inequality with a smaller feasible set.
- Hence: convex upper bound.



Performance



Poster presentation

Poster #166

Today 05:00 -- 07:00 PM @ Room 210 & 230

UPPSALA UNIVERSITET

SWEDISH FOUNDATION *for* STRATEGIC RESEARCH