What are we going to discuss?

- Learning for MPC A closed-loop performance view
- Safety & stability in Learning for MPC
- MPC and Markov Decision Processes When is learning beneficial?

 $Q_{+}(\mathbf{x},\mathbf{u}) \leftarrow L(\mathbf{x},\mathbf{u}) + \gamma \mathbb{E}\left[V\left(\mathbf{x}_{+}
ight) \mid \mathbf{x},\mathbf{u}
ight]$



samples = 1000000

What are we going to discuss today?

MPC for MDPs

- MPC: purpose and usage
- MPC as a practical solution to tackle MDPs
- Planning vs. Policing
- Repeated planning as a policy
- MPC as an MDP model
- Optimal MPC model?

Model Predictive Control for

Markov Decision Processes

Sébastien Gros

Cybernetic, NTNU

Freiburg PhD School

Model Predictive Control (MPC) Tuning for Performance

$$\begin{split} \text{MPC: at current state s solve} \\ \min_{x,u} \quad \mathcal{T}(x_N) + \sum_{k=0}^{N-1} \mathcal{L}(x_k, u_k) \\ \text{s.t.} \quad x_{k+1} = f_{\theta}(x_k, u_k) \\ \quad \mathbf{h}(x_k, u_k) \leq 0 \\ x_0 = s \\ \text{gives policy $\pi_{\theta}^{\text{MPC}}$(s) = u_0^{\star}} \end{split}$$

MPC: at current state s solve
$$\begin{split} \min_{\mathbf{x},\mathbf{u}} \quad & \mathcal{T}_{\theta}\left(\mathbf{x}_{N}\right) + \sum_{k=0}^{N-1} \mathcal{L}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ \text{s.t.} \quad & \mathbf{x}_{k+1} = \mathbf{f}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ & \mathbf{h}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \leq \mathbf{0} \\ & \mathbf{x}_{0} = \mathbf{s} \\ & \text{gives policy } \boldsymbol{\pi}_{\theta}^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \end{split}$$

- "Classic" view
- MPC built around the model f_{θ}
- θ fits \mathbf{f}_{θ} to data
- Model fitting→optimality is tricky

- "Holistic" view
- MPC is a model of Q^*
- heta fits Q^{MPC} to Q^{\star}
- $RL \rightarrow optimality$, also BO btw!!

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

Model Predictive Control (MPC) Tuning for Performance

$$\begin{split} \text{MPC: at current state s solve} \\ \min_{x,u} \quad \mathcal{T}(x_N) + \sum_{k=0}^{N-1} \mathcal{L}(x_k, u_k) \\ \text{s.t.} \quad x_{k+1} = f_{\theta}(x_k, u_k) \\ \quad h(x_k, u_k) \leq 0 \\ \quad x_0 = s \\ \text{gives policy $\pi_{\theta}^{\mathrm{MPC}}$(s) = u_0^{\star}} \end{split}$$

MPC: at current state s solve $\begin{array}{ll} \min_{\mathbf{x},\mathbf{u}} & T_{\theta}\left(\mathbf{x}_{N}\right) + \sum_{k=0}^{N-1} \mathcal{L}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ \text{s.t.} & \mathbf{x}_{k+1} = \mathbf{f}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ & \mathbf{h}_{\theta}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \leq 0 \\ & \mathbf{x}_{0} = \mathbf{s} \\ & \text{gives policy } \boldsymbol{\pi}_{\theta}^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \end{array}$

- "Classic" view
- MPC built around the model f_{θ}
- θ fits \mathbf{f}_{θ} to data
- Model fitting→optimality is tricky

- "Holistic" view
- MPC is a model of Q^*
- heta fits Q^{MPC} to Q^{\star}
- $\bullet \ \mathsf{RL} \to \mathsf{optimality, also BO btw!!}$

イロト 不得 トイラト イラト 一日

Can we get $Q^{\rm MPC} = Q^*$ from tuning the MPC model alone?? Can we get $Q^{\rm MPC} = Q^*$ from fitting the model to data??

We are unpacking the maths

Outline

1 MPC & MDP: Let's rehearse the background

2 MPC Model for Performance

- 3 Optimal MPC models
- 4 Stochastic MPC models



Infinite horizon & discounted

 $oldsymbol{\pi}^{\star}_{\infty} = rgmin_{oldsymbol{\pi}} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k L(\mathbf{x}_k, \mathbf{u}_k)
ight]$

Policy π : state \rightarrow action belongs to a function space

イロト イポト イヨト イヨト

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Policy π : state ightarrow action belongs to a function space

イロト イポト イヨト イヨト

Finite-horizon equivalent:

$$\pi_{0,...,N-1}^{\star} = \underset{\pi_{0,...,N-1}}{\arg \min} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$$
If $T = V_{\star}$, then $\pi_{0,...,N-1}^{\star} = \pi_{\infty}^{\star}$

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Finite-horizon equivalent:

$$\pi_{0,\ldots,N-1}^{\star} = \underset{\pi_{0,\ldots,N-1}}{\arg\min} \mathbb{E}\left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k})\right]$$

If $T = V_{\star}$, then $\pi_{0,\ldots,N-1}^{\star} = \pi_{\infty}^{\star}$

Planning instead of policing:

$$\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$$

i.e. restrict policies to fixed $\mathbf{u}_{0,...,N-1}$

Policy π : state ightarrow action belongs to a function space

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Finite-horizon equivalent:

$$\pi_{0,\dots,N-1}^{\star} = \underset{\pi_{0,\dots,N-1}}{\arg\min} \mathbb{E}\left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k})\right]$$

If $T = V_{\star}$, then $\pi_{0,\dots,N-1}^{\star} = \pi_{\infty}^{\star}$

Planning instead of policing:

$$\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$$

i.e. restrict policies to fixed $\mathbf{u}_{0,...,N-1}$

Policy π : state ightarrow action belongs to a function space

Deterministic approximation: $\min_{\substack{\pi_{0,...,N-1}}} \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})$ s.t $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k}, \mathbf{u}_{k})$ $\mathbf{x}_{0} = \mathbf{s}$ i.e. adopt deterministic model

S. Gros (NTNU)



Finite-horizon equivalent:

$$\pi_{0,\dots,N-1}^{\star} = \underset{\pi_{0},\dots,N-1}{\arg\min} \mathbb{E}\left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k})\right]$$

If $T = V_{\star}$, then $\pi_{0,\dots,N-1}^{\star} = \pi_{\infty}^{\star}$

Why attacking the problem in these ways?

Planning instead of policing:

$$\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$$

i.e. restrict policies to fixed $\mathbf{u}_{0,...,N-1}$

Deterministic approximation: $\min_{\mathbf{u}_{0,...,N-1}} \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})$ s.t $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k}, \mathbf{u}_{k})$ $\mathbf{x}_{0} = \mathbf{s}$ i.e. adopt deterministic model

Planning instead of policing:

$$\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$

A (10) F (10) F (10)

Planning instead of policing:

$$\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$$

i.e. restrict policies to fixed $\mathbf{u}_{0,...,N-1}$



Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \qquad \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$





MDP & SDP

→ Ξ →

Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0,...,N-1}} \mathbb{E}\left[\left. \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right| \, \mathbf{x}_{0} = \mathbf{s} \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,...,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \middle| \mathbf{x}_{0} = \mathbf{s} \right] \min_{\pi_{0},\ldots,N-1} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Planning instead of policing: $\min_{\mathbf{u}_{0,\ldots,N-1}} \mathbb{E} \left[\left| T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right| \mathbf{x}_{0} = \mathbf{s} \right] = \min_{\pi_{0,\ldots,N-1}} \mathbb{E} \left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \right]$

i.e. restrict policies to fixed $\mathbf{u}_{0,\dots,N-1}$







Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Policing: $\min_{\boldsymbol{\pi}_{0,...,N-1}} \mathbb{E}\left[T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$ i.e. optimize over policies











Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$



S. Gros (NTNU)

MDP & SDP

Fall, 2023 8 / 29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

Deterministic approximation:

$$egin{aligned} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & ext{ s.t } & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & \mathbf{x}_{0} = \mathbf{s} \end{aligned}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$${{{\pi }^{{
m{MPC}}}}\left({
m{s}}
ight)} = {{
m{u}}_0^\star }$$

from repeated planning



S. Gros (NTNU)

MDP & SDP

Fall, 2023 8 / 29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$${{{\pi }^{{
m{MPC}}}}\left({
m{s}}
ight)} = {{
m{u}}_{0}^{\star }}$$

from repeated planning



Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$${{{\pi }^{{
m{MPC}}}}\left({
m{s}}
ight)} = {{
m{u}}_0^\star }$$

from repeated planning



Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$${{{\pi }^{{
m{MPC}}}}\left({
m{s}}
ight)} = {{
m{u}}_{0}^{\star }}$$

from repeated planning



Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$${{{\pi }^{{
m{MPC}}}}\left({
m{s}}
ight)} = {{
m{u}}_{0}^{\star }}$$

from repeated planning



Infinite horizon & discounted $\pi_{\infty}^{\star} = \operatorname*{arg\,min}_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit ...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$





Infinite horizon & discounted $\pi_{\infty}^{\star} = \operatorname*{arg\,min}_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit ...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$





Infinite horizon & discounted $\pi_{\infty}^{\star} = \operatorname*{arg\,min}_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit ...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$

from repeated planning


Infinite horizon & discounted $\pi_{\infty}^{\star} = \operatorname*{arg\,min}_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit ...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$

from repeated planning



MDP & SDP

Fall, 2023 8/29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$

from repeated planning



MDP & SDP

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$

from repeated planning



MDP & SDP

Fall, 2023 8 / 29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

Deterministic approximation:

$$\begin{array}{ll} \min_{\mathbf{u}_{0,\ldots,N-1}} & T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} & \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Remarks:

- MPC predictions from f are a simplified representation of the real dynamics
- Finer models can be built, e.g. scenario trees, more on this in a bit...
- MPC defines a policy:

$$\pi^{\mathrm{MPC}}\left(\mathbf{s}
ight)=\mathbf{u}_{0}^{\star}$$

from repeated planning



MDP & SDP

MPC as a policy

Deterministic MPC: $\min_{\mathbf{u}_{0,...,N-1}} \quad T(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} L(\mathbf{x}_{k}, \mathbf{u}_{k})$ s.t $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k}, \mathbf{u}_{k})$ $\mathbf{x}_{0} = \mathbf{s}$

Defines policy:

$$\pi^{ ext{MPC}}\left(ext{s}
ight)= ext{u}_{0}^{\star}$$

How does π^{MPC} relate to π^* ?

No reason to match:

- Planning rather than policing
- Plan ignores stochasticity



Can we clarify the relationship?

・ロト ・ 母 ト ・ ヨ ト ・ ヨ

Historically MPC focuses on **constraints satisfaction & stability**. Cost is for reference tracking, not representative of the system performance.

< ∃ > <

Historically MPC focuses on **constraints satisfaction & stability**. Cost is for reference tracking, not representative of the system performance.

- "Tracking MPC"
- Classic stability theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
- "MPC is for constraints satisfaction" (public statement)

Historically MPC focuses on **constraints satisfaction & stability**. Cost is for reference tracking, not representative of the system performance. More recently, focus on **closed-loop performance**, e.g. energy, time, money. Cost is generic, representative of the system performance.

- "Tracking MPC"
- Classic stability theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
- "MPC is for constraints satisfaction" (public statement)

Historically MPC focuses on **constraints satisfaction & stability**. Cost is for reference tracking, not representative of the system performance.

- "Tracking MPC"
- Classic stability theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
- "MPC is for constraints satisfaction" (public statement)

More recently, focus on **closed-loop performance**, e.g. energy, time, money. Cost is generic, representative of the system performance.

- "Economic MPC"
- Dissipativity theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
 - ... debatable approach for performance

< ∃ > < ∃

Historically MPC focuses on **constraints satisfaction & stability**. Cost is for reference tracking, not representative of the system performance.

- "Tracking MPC"
- Classic stability theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
- "MPC is for constraints satisfaction" (public statement)

More recently, focus on **closed-loop performance**, e.g. energy, time, money. Cost is generic, representative of the system performance.

- "Economic MPC"
- Dissipativity theory
- Uncertainty via
 - Robust MPC
 - Stochastic MPC
 - ... debatable approach for performance

MPC for closed-loop performance

- is not a very old topic
- unclear in the presence of stochasticity
- partially clarified by recent results

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} L(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

$$\begin{array}{l} \text{MPC policy } \boldsymbol{\pi}^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$$





イロト イポト イヨト イヨト

S. Gros (NTNU)

MDP & SDP

◆ ■ シ へ ○
 Fall, 2023 11 / 29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

$$\begin{array}{l} \text{MPC policy } \boldsymbol{\pi}^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$$

MPC as a model of the MDP

$$egin{aligned} &\mathcal{V}^{ ext{MPC}}\left(\mathbf{s}
ight) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ } \mathbf{x}_0 = \mathbf{s} \end{aligned}$$





イロト イボト イヨト イヨト

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

$$\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$$

MPC as a model of the MDP

$$egin{aligned} V^{ ext{MPC}}\left(\mathbf{s}
ight) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ } \mathbf{x}_0 = \mathbf{s} \end{aligned}$$

$$egin{aligned} Q^{ ext{MPC}}\left(\mathbf{s},\mathbf{a}
ight) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & ext{ s.t } \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & ext{ } \mathbf{x}_{0} = \mathbf{s}, \quad \mathbf{u}_{0} = \mathbf{a} \end{aligned}$$





S. Gros (NTNU)

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

$\begin{array}{l} \textbf{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$

MPC as a model of the MDP

$$egin{aligned} \mathcal{V}^{ ext{MPC}}\left(\mathbf{s}
ight) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \ & ext{ } \mathbf{x}_0 = \mathbf{s} \end{aligned}$$

MPC is consistent:

$$egin{aligned} V^{\mathrm{MPC}}\left(s
ight) &= \min_{a} \ Q^{\mathrm{MPC}}\left(s,a
ight) \ \pi^{\mathrm{MPC}}\left(s
ight) &= rgmin \ Q^{\mathrm{MPC}}\left(s,a
ight) \end{aligned}$$

$$\begin{split} Q^{\text{MPC}}\left(\mathbf{s},\mathbf{a}\right) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \mathbf{x}_{0} = \mathbf{s}, \quad \mathbf{u}_{0} = \mathbf{a} \end{split}$$

◆□▶ ◆□▶ ◆三▶ ◆三▶ ● ○○○

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

$$egin{aligned} \mathcal{V}^{ ext{MPC}}\left(\mathbf{s}
ight) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & ext{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \ & ext{ } \mathbf{x}_{0} = \mathbf{s} \end{aligned}$$

$$\begin{split} Q^{\text{MPC}}\left(\mathbf{s},\mathbf{a}\right) &:= \min_{\mathbf{u}} \ \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \mathbf{x}_{0} &= \mathbf{s}, \quad \mathbf{u}_{0} = \mathbf{a} \end{split}$$

$$\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(\mathbf{s}\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ & \underset{\mathbf{u}_{0},\ldots,N-1}{\min} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$$

MPC is consistent:

$$egin{aligned} V^{\mathrm{MPC}}\left(\mathrm{s}
ight) &= \min_{\mathrm{a}} \ Q^{\mathrm{MPC}}\left(\mathrm{s},\mathrm{a}
ight) \ \pi^{\mathrm{MPC}}\left(\mathrm{s}
ight) &= rgmin \ Q^{\mathrm{MPC}}\left(\mathrm{s},\mathrm{a}
ight) \end{aligned}$$

MPC is a complete model of MDP if:

$$Q^{\mathrm{MPC}}\left(\mathbf{s},\mathbf{a}
ight)=Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)$$

for all s, a. Then **optimality** holds:

$$\pi^{\mathrm{MPC}}\left(\mathrm{s}
ight)=\pi^{\star}\left(\mathrm{s}
ight)$$

S. Gros (NTNU)

MDP & SDP

Fall, 2023 11 / 29

Outline

MPC & MDP: Let's rehearse the background

2 MPC Model for Performance

- 3 Optimal MPC models
- 4 Stochastic MPC models



Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

MPC policy $\mathbf{\pi}^{ ext{MPC}}\left(ext{s} ight)= ext{u}_{0}^{\star}$ from			
$\min_{\mathbf{u}_{0,\ldots,N-1}}$	$\mathcal{T}(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k)$		
s.t	$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)$		
	$\mathbf{x}_0 = \mathbf{s}$		

When does

$$Q^{\mathrm{MPC}}\left(\mathbf{s},\mathbf{a}
ight)=Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)$$
 hold?

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

MPC policy
$$\pi^{\text{MPC}}(\mathbf{s}) = \mathbf{u}_0^{\star}$$
 from

$$\min_{\mathbf{u}_{0,...,N-1}} \quad \mathcal{T}(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k)$$
s.t $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)$
 $\mathbf{x}_0 = \mathbf{s}$

Bellman equation for all s, a

$$Q^{\mathrm{MPC}}\left(\mathrm{s},\mathrm{a}
ight) =Q^{\star}\left(\mathrm{s},\mathrm{a}
ight)$$
 hold?

$$\mathcal{Q}^{\star}\left(\mathbf{s},\mathbf{a}
ight)=\mathcal{L}\left(\mathbf{s},\mathbf{a}
ight)+\gamma\mathbb{E}\left[\mathcal{V}^{\star}\left(\mathbf{s}_{+}
ight)\mid\mathbf{s},\mathbf{a}
ight]$$

イロト イポト イヨト イヨト

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k}) \right]$

$$\begin{array}{l} \textbf{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \underset{\mathbf{u}_{0,\ldots,N-1}}{\min} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Bellman equation for all \mathbf{s}, \mathbf{a}

$$Q^{\star}\left(\mathbf{s},\mathbf{a}
ight) =L\left(\mathbf{s},\mathbf{a}
ight) +\gamma\mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}
ight) \left| \,\mathbf{s},\mathbf{a}
ight]$$

When does

$$Q^{\mathrm{MPC}}\left(\mathbf{s},\mathbf{a}
ight)=Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)$$
 hold?

Theorem: if $T = \gamma^N V^*$, MPC is a **complete** MDP solution[†] if for some *c*

$$Q^{\star}(\mathbf{s}, \mathbf{a}) = L(\mathbf{s}, \mathbf{a}) + \gamma V^{\star}(\mathbf{f}(\mathbf{s}, \mathbf{a})) + \boldsymbol{c}$$

イロト 不得 トイラト イラト 一日

holds for all $s,a\ ({\sf with\ technical\ assumption})$

[†]up to a constant

Infinite horizon & discounted $\pi_{\infty}^{\star} = \arg\min_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

$$\begin{array}{l} \textbf{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0,...,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$$

When does

$$Q^{\mathrm{MPC}}\left(\mathrm{s},\mathrm{a}
ight) =Q^{\star}\left(\mathrm{s},\mathrm{a}
ight)$$
 hold?

[†]up to a constant

Bellman equation for all \mathbf{s}, \mathbf{a}

$$Q^{\star}\left(\mathbf{s},\mathbf{a}
ight) =L\left(\mathbf{s},\mathbf{a}
ight) +\gamma\mathbb{E}\left[\mathcal{V}^{\star}\left(\mathbf{s}_{+}
ight) \left| \mathbf{s},\mathbf{a}
ight]$$

Theorem: if $T = \gamma^N V^*$, MPC is a **complete** MDP solution[†] if for some *c*

$$Q^{\star}(\mathbf{s}, \mathbf{a}) = L(\mathbf{s}, \mathbf{a}) + \gamma V^{\star}(\mathbf{f}(\mathbf{s}, \mathbf{a})) + \boldsymbol{c}$$

holds for all \mathbf{s}, \mathbf{a} (with technical assumption)

Proof: telescopic sums, Bellman identities, properties of the advantage function, some measure theory, devil is in the details, approximation is a some set of the set of the

S. Gros (NTNU)

MDP & SDP

Fall, 2023 13 / 29

Infinite horizon & discounted $\pi_{\infty}^{\star} = \operatorname*{arg\,min}_{\pi} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right]$

$$\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$$

When does

$$Q^{\mathrm{MPC}}\left(\mathbf{s},\mathbf{a}
ight)=Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)$$
 hold?

What does it mean?

[†]up to a constant

Bellman equation for all \mathbf{s}, \mathbf{a}

$$Q^{\star}\left(\mathrm{s,a}
ight) =L\left(\mathrm{s,a}
ight) +\gamma\mathbb{E}\left[V^{\star}\left(\mathrm{s_{+}}
ight) \left| \mathrm{s,a}
ight]$$

Theorem: if $T = \gamma^N V^*$, MPC is a **complete** MDP solution[†] if for some **c**

$$Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)=L\left(\mathbf{s},\mathbf{a}
ight)+\gamma V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight)+\boldsymbol{c}$$

holds for all \mathbf{s}, \mathbf{a} (with technical assumption)

Proof: telescopic sums, Bellman identities, properties of the advantage function, some measure theory, devil is in the details, a set the set the set of the set of

S. Gros (NTNU)

MDP & SDP

Fall, 2023 13 / 29

 $\begin{array}{l} \text{MPC policy } \pi^{\mathrm{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \underset{\mathbf{u}_{0,\ldots,N-1}}{\min} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Theorem: if $T = \gamma^N V^*$, MPC is complete

$$Q^{\star}(\mathbf{s}, \mathbf{a}) = L(\mathbf{s}, \mathbf{a}) + \gamma V^{\star}(\mathbf{f}(\mathbf{s}, \mathbf{a})) + c$$

Theorem: if $T = \gamma^N V^*$, MPC is complete

$$Q^{\star}\left(\mathbf{s},\mathbf{a}
ight)=L\left(\mathbf{s},\mathbf{a}
ight)+\gamma V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight)+oldsymbol{c}$$

Equivalent statement

$$\mathbb{E}\left[{{V}^{\star }\left({{{\rm{s}}_{+}}} \right)|\,{\rm{s}},{\rm{a}}} \right] = {V}^{\star }\left({{\rm{f}}\left({{\rm{s}},{\rm{a}}} \right)} \right) + c$$

requirement on model f !!!

◆□▶ ◆□▶ ◆三▶ ◆三▶ ● ○○○

$$\begin{split} \text{MPC policy } \pi^{\mathrm{MPC}}\left(s\right) &= \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{u}_{0,...,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{ s.t } \quad \mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \quad \mathbf{x}_{0} &= \mathbf{s} \end{split}$$

Theorem: if $T = \gamma^N V^*$, MPC is complete $Q^*(\mathbf{s}, \mathbf{a}) = L(\mathbf{s}, \mathbf{a}) + \gamma V^*(\mathbf{f}(\mathbf{s}, \mathbf{a})) + c$

Equivalent statement

$$\mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s},\mathbf{a}\right] =V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}\right) \right) +\textbf{\textit{c}}$$

requirement on model f !!!

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

E.g. c = 0, V^{\star} quadratic



 $\begin{array}{l} \text{MPC policy } \pi^{\mathrm{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Theorem: if $T = \gamma^{N} V^{\star}$, MPC is complete $Q^{\star}(\mathbf{s}, \mathbf{a}) = L(\mathbf{s}, \mathbf{a}) + \gamma V^{\star}(\mathbf{f}(\mathbf{s}, \mathbf{a})) + c$

Equivalent statement

$$\mathbb{E}\left[{{V}^{\star }\left({{\mathbf{s}}_{+}} \right)\left| {\,\mathbf{s}},{\mathbf{a}} \right]} = {V}^{\star }\left({\mathbf{f}\left({{\mathbf{s}},{\mathbf{a}}} \right)} \right) + {\textit{c}}$$

requirement on model f !!!

E.g. c = 0, V^* quadratic



Remark

Even for a simple V^* neither

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\right|\mathbf{s},\mathbf{a}\right]$$

nor

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right) = \max \varrho \left[\left. \mathbf{s}_{+} \right| \mathbf{s},\mathbf{a} \right]$$

<ロト <回ト < 回ト < 回ト

make the MPC complete / optimal

S. Gros (NTNU)

MDP & SDP

Fall, 2023 14 / 29

MPC is **complete** if model **f** satisfies $\mathbb{E}\left[V^{\star}\left(s_{+}\right) \mid s, \mathbf{a}\right] = V^{\star}\left(\mathbf{f}\left(s, \mathbf{a}\right)\right) + c$

 $\begin{array}{ll} \text{MPC policy } \pi^{\mathrm{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ & \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mathbf{x}_{0} = \mathbf{s} \end{array}$

Model f built via Least-Squares fit

$$\min_{\boldsymbol{\theta}} \sum_{k=0}^{N} \|\mathbf{f}_{\boldsymbol{\theta}}\left(\mathbf{s}_{k}, \mathbf{a}_{k}\right) - \mathbf{s}_{k+1}\|^{2}$$

Classic approach in SYSID & ML

MPC is **complete** if model **f** satisfies $\mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s}, \mathbf{a}\right] = V^{\star}\left(\mathbf{f}\left(\mathbf{s}, \mathbf{a}\right)\right) + c$

くぼう くほう くほう しほ

 $\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \underset{\mathbf{u}_{0},\ldots,N-1}{\text{min}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$

Model f built via Least-Squares fit

$$\min_{\boldsymbol{\theta}} \sum_{k=0}^{N} \left\| \mathbf{f}_{\boldsymbol{\theta}} \left(\mathbf{s}_{k}, \mathbf{a}_{k} \right) - \mathbf{s}_{k+1} \right\|^{2}$$

Classic approach in SYSID & ML

In ideal conditions:

$$\mathrm{f}_{{m heta}^{\star}}\left(\mathrm{s},\mathrm{a}
ight) \,
ightarrow \, \mathbb{E}\left[\left.\mathrm{s}_{+} \,|\,\mathrm{s},\mathrm{a}\,
ight]$$

i.e. one-step ahead expected transition

MPC is **complete** if model **f** satisfies $\mathbb{E}\left[V^{\star}(\mathbf{s}_{+}) \mid \mathbf{s}, \mathbf{a}\right] = V^{\star}\left(\mathbf{f}(\mathbf{s}, \mathbf{a})\right) + \boldsymbol{c}$ ž k

Trajectory from simulating $f_{\theta^{\star}}$

<u> </u>	\sim	(N 1 T N 1 I I)
<u> </u>	-roc I	
J.	0103 1	

Fall, 2023 15 / 29

 $\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \\ \underset{\mathbf{u}_{0},\ldots,N-1}{\text{min}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \text{s.t} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \\ \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$

Model f built via Least-Squares fit

$$\min_{\boldsymbol{\theta}} \sum_{k=0}^{N} \left\| \mathbf{f}_{\boldsymbol{\theta}} \left(\mathbf{s}_{k}, \mathbf{a}_{k} \right) - \mathbf{s}_{k+1} \right\|^{2}$$

Classic approach in SYSID & ML

In ideal conditions:

$$\mathbf{f}_{\boldsymbol{\theta}^{\star}}\left(\mathbf{s},\mathbf{a}\right)\,\rightarrow\,\mathbb{E}\left[\left.\mathbf{s}_{+}\right|\mathbf{s},\mathbf{a}\left.\right]$$

i.e. one-step ahead expected transition

MPC is **complete** if model **f** satisfies $\mathbb{E}\left[V^{\star}(\mathbf{s}_{+}) \mid \mathbf{s}, \mathbf{a}\right] = V^{\star}\left(\mathbf{f}(\mathbf{s}, \mathbf{a})\right) + \boldsymbol{c}$ ž k

Trajectory from simulating $\mathbf{f}_{\boldsymbol{\theta}^\star}$

<u> </u>	<u> </u>	(N 1 T N 11 1	t
<u> </u>	-roc I		
J.	0103 1		
		`	

Fall, 2023 15 / 29

Model f built via Least-Squares fit

$$\min_{\boldsymbol{\theta}} \sum_{k=0}^{N} \|\mathbf{f}_{\boldsymbol{\theta}}\left(\mathbf{s}_{k}, \mathbf{a}_{k}\right) - \mathbf{s}_{k+1}\|^{2}$$

Classic approach in SYSID & ML

In ideal conditions:

$$\mathbf{f}_{\boldsymbol{\theta}^{\star}}\left(\mathbf{s},\mathbf{a}\right)\,\rightarrow\,\mathbb{E}\left[\left.\mathbf{s}_{+}\right|\mathbf{s},\mathbf{a}\right]$$

i.e. one-step ahead expected transition

MPC is complete if model f satisfies $\mathbb{E}\left[V^{\star}(\mathbf{s}_{+}) \mid \mathbf{s}, \mathbf{a}\right] = V^{\star}\left(\mathbf{f}(\mathbf{s}, \mathbf{a})\right) + \boldsymbol{c}$ XX

k

Trajectory from simulating $f_{\theta^{\star}}$ Expected trajectory

MDP & SDP

 $\begin{array}{ll} \mbox{MPC policy } \pi^{\rm MPC}\left(s\right) = \mathbf{u}_{0}^{\star} \mbox{ from} \\ & \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mbox{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

MPC is **complete** if model **f** satisfies $\mathbb{E}\left[V^{*}\left(s_{+}\right) \mid s, \mathbf{a}\right] = V^{*}\left(\mathbf{f}\left(s, \mathbf{a}\right)\right) + \mathbf{c}$

The gap in

$$Q^{ ext{MPC}}\left(ext{s}, ext{a}
ight)pprox Q^{\star}\left(ext{s}, ext{a}
ight), \qquad \pi^{ ext{MPC}}\left(ext{s}
ight)pprox \pi^{\star}\left(ext{s}
ight)$$

when using MPC models based on one-step ahead Least-Squares fitting (PEM) comes from the lack of commutativity (up to a constant) between V^* and $\mathbb{E}[.]$, i.e.

$$\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \left| \, \mathrm{s},\mathrm{a}
ight] -V^{\star}\left(\mathbb{E}\left[\, \mathrm{s}_{+} \left| \, \mathrm{s},\mathrm{a}
ight]
ight)
eq c$$

To our best knowledge other methods (sim error, max likelihood) do not fix that

S. Gros (NTNU)

◆□▶ ◆□▶ ◆三▶ ◆三▶ ○三 のへ⊙

An Important Exception

 $\begin{array}{l} \mbox{MPC policy } \pi^{\rm MPC}\left(s\right) = \mathbf{u}_{0}^{\star} \mbox{ from} \\ & \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mbox{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

イロト イポト イヨト イヨト 二日

An Important Exception - The LQR Case

$$\begin{array}{l} \mbox{MPC policy } \pi^{\rm MPC}\left(s\right) = \mathbf{u}^{\star}_{0} \mbox{ from} \\ & \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mbox{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$$

Consider

- *L* is **quadratic**, no constraints
- Real dynamics: for some density φ

$$arrho\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]=arphi\left(\mathbf{s}_{+}-oldsymbol{\mu}\left(\mathbf{s},\mathbf{a}
ight)
ight)$$

where μ is affine

MPC model selected as:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\right|\mathbf{s},\mathbf{a}\right]$$

An Important Exception - The LQR Case

Then...

- V^* is quadratic
- f is affine
- There is c such that

 $\mathbb{E}\left[{{V}^{\star }\left({{\rm{s}}_{+}} \right)|\,{\rm{s}},{\rm{a}}} \right] = {V}^{\star }\left({{\rm{f}}\left({{\rm{s}},{\rm{a}}} \right)} \right) + \textit{c}$

• MPC is **complete** for some *T*, i.e.

$$egin{aligned} & Q^{ ext{MPC}}\left(ext{s}, ext{a}
ight) = Q^{\star}\left(ext{s}, ext{a}
ight), \ & \pi^{ ext{MPC}}\left(ext{s}
ight) = \pi^{\star}\left(ext{s}
ight) \end{aligned}$$

Consider

- *L* is **quadratic**, no constraints
- Real dynamics: for some density φ

$$arrho\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]=arphi\left(\mathbf{s}_{+}-oldsymbol{\mu}\left(\mathbf{s},\mathbf{a}
ight)
ight)$$

where μ is affine

MPC model selected as:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\right.\mathbf{s}_{+}\left.\right|\,\mathbf{s},\mathbf{a}\left.\right]$$

• • = • • = •

An Important Exception - The LQR Case

Then...

- V^* is quadratic
- f is affine
- There is c such that

 $\mathbb{E}\left[{{V}^{\star }\left({{s_ + }} \right)\left| {\,{s_ ,a}} \right] = {V^{\star }}\left({f\left({{s_ ,a}} \right)} \right) + c$

• MPC is **complete** for some *T*, i.e.

$$egin{aligned} & Q^{\mathrm{MPC}}\left(\mathbf{s},\mathbf{a}
ight) = Q^{\star}\left(\mathbf{s},\mathbf{a}
ight), \ & \pi^{\mathrm{MPC}}\left(\mathbf{s}
ight) = \pi^{\star}\left(\mathbf{s}
ight) \end{aligned}$$

Consider

- *L* is **quadratic**, no constraints
- Real dynamics: for some density φ

$$arrho\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]=arphi\left(\mathbf{s}_{+}-oldsymbol{\mu}\left(\mathbf{s},\mathbf{a}
ight)
ight)$$

where μ is affine

• MPC model selected as:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\right.\mathbf{s}_{+}\left.\right|\,\mathbf{s},\mathbf{a}\left.\right]$$

This is LQR + i.i.d state noise!

Why is this relevant for MPC?

・ 同 ト ・ ヨ ト ・ ヨ ト

Local Optimality of Classic MPC

Assume:

- $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \right]$ is smooth in \mathbf{s}, \mathbf{a} , for all \mathbf{s}_{+}
- L is smooth
- π^{*} is such that system dynamics converge to steady state density ρ_{*} (.) (dissipative) which is "off-constraints"

▲ □ ▶ ▲ □ ▶ ▲ □ ▶
$\begin{array}{ll} \mbox{MPC policy } \pi^{\rm MPC}\left(s\right) = \mathbf{u}_{0}^{\star} \mbox{ from} \\ & \min_{\mathbf{u}_{0},\ldots,N-1} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \mbox{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ & \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Assume:

- $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \right]$ is smooth in \mathbf{s}, \mathbf{a} , for all \mathbf{s}_{+}
- L is smooth
- π^{*} is such that system dynamics converge to steady state density ρ_{*} (.) (dissipative) which is "off-constraints"

Then MPC based on "expected-value" model

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)=\mathbb{E}\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]$$

yields a **locally optimal policy**, optimality loss in the order of the moments of $\rho_{\star}(.)$

 $\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{*} \text{ from} \\ \min_{\mathbf{u}_{0,\ldots,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Assume:

- $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \right]$ is smooth in \mathbf{s}, \mathbf{a} , for all \mathbf{s}_{+}
- L is smooth
- π^{*} is such that system dynamics converge to steady state density ρ_{*} (.) (dissipative) which is "off-constraints"

Then MPC based on "expected-value" model

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)=\mathbb{E}\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]$$

yields a **locally optimal policy**, optimality loss in the order of the moments of $\rho_{\star}(.)$

If problem is smooth & optimal policy drives and keeps the system "tightly" to its optimal steady state, then one can expect the MPC based on an "expected-value" model to perform well

・ロト ・ 母 ト ・ ヨ ト ・ ヨ ト … ヨ

 $\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{*} \text{ from} \\ \min_{\mathbf{u}_{0,...,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Assume:

- $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \right]$ is smooth in \mathbf{s}, \mathbf{a} , for all \mathbf{s}_{+}
- L is smooth
- π^{*} is such that system dynamics converge to steady state density ρ_{*} (.) (dissipative) which is "off-constraints"

Then MPC based on "expected-value" model

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)=\mathbb{E}\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]$$

yields a **locally optimal policy**, optimality loss in the order of the moments of $\rho_{\star}(.)$

If problem is smooth & optimal policy drives and keeps the system "tightly" to its optimal steady state, then one can expect the MPC based on an "expected-value" model to perform well Classic MPC paradigm works well under these conditions

イロト 不得下 イヨト イヨト 二日

 $\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{u}_{0,...,N-1}} \quad \mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \text{ s.t } \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_{k},\mathbf{u}_{k}) \\ \quad \mathbf{x}_{0} = \mathbf{s} \end{array}$

Assume:

- $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \right]$ is smooth in \mathbf{s}, \mathbf{a} , for all \mathbf{s}_{+}
- L is smooth
- π^{*} is such that system dynamics converge to steady state density ρ_{*} (.) (dissipative) which is "off-constraints"

Then MPC based on "expected-value" model

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)=\mathbb{E}\left[\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}
ight]$$

yields a **locally optimal policy**, optimality loss in the order of the moments of $\rho_{\star}(.)$

If problem is smooth & optimal policy drives and keeps the system "tightly" to its optimal steady state, then one can expect the MPC based on an "expected-value" model to perform well

Classic MPC paradigm works well under these conditions

Not "classic"?

- Economic / non-smooth cost
- No dissipativity / "disturbances"
- Non-smooth problem

S. Gros (NTNU)

MDP & SDP

Fall, 2023 17 / 29

Consider the dynamics:

 $s_+=s+a+w$ with $s,a,\,w\in\mathbb{R},$ and $w\sim\mathcal{N}\left(0,\sigma\right)$ i.i.d.

on a restricted interval.

Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

Constraints:

$$\mathbf{a} \in [-0.25, 0.25]$$

Consider the dynamics:

 $s_+=s+a+w$ with $s,a,\,w\in\mathbb{R},$ and $w\sim\mathcal{N}\left(0,\sigma\right)$ i.i.d.

on a restricted interval.

Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

Constraints:

$$\mathbf{a} \in [-0.25, 0.25]$$

MPC model:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\,\right|\,\mathbf{s},\mathbf{a}\,\right]=\mathbf{s}+\mathbf{a}$$

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ i.i.d. on a restricted interval. **Constraints:**

 $\mathbf{a}\,\in\,[-0.25,0.25]$

MPC model:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\,\right|\,\mathbf{s},\mathbf{a}\,\right]=\mathbf{s}+\mathbf{a}$$

If V^* is quadratic i.e. $V^*(s) = s^\top W s + d^\top s + V_0$ then

Consider the dynamics:

 $\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$

Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ i.i.d. on a restricted interval. **Constraints:**

$$\mathbf{a} \in [-0.25, 0.25]$$

MPC model:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\,\right|\,\mathbf{s},\mathbf{a}\,\right]=\mathbf{s}+\mathbf{a}$$

If V^* is quadratic i.e. $V^*(s) = s^\top W s + d^\top s + V_0$ then

$$\begin{split} \mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s}, \mathbf{a} \right] &= \mathbb{E}\left[\left. \mathbf{s}_{+}^{\top} W \, \mathbf{s}_{+} + \mathbf{d}^{\top} \mathbf{s}_{+} + V_{0} \right| \, \mathbf{s}, \mathbf{a} \right] \\ &= \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right)^{\top} \, W \, \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right) + \mathbf{d}^{\top} \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right) + V_{0} + \mathbb{E}\left[\mathbf{w}^{\top} \, W \, \mathbf{w} \right] \\ &= V^{\star}\left(\mathbf{f}\left(\mathbf{s}, \mathbf{a}\right)\right) + \underbrace{\mathbb{E}\left[\mathbf{w}^{\top} \, W \, \mathbf{w} \right]}_{=c} \end{split}$$

Consider the dynamics:

 $\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$

Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ i.i.d. on a restricted interval. **Constraints:**

$$\mathbf{a} \, \in \, [-0.25, 0.25]$$

MPC model:

$$\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)=\mathbb{E}\left[\left.\mathbf{s}_{+}\,\right|\,\mathbf{s},\mathbf{a}\,\right]=\mathbf{s}+\mathbf{a}$$

If V^* is quadratic i.e. $V^*(s) = s^\top W s + d^\top s + V_0$ then

$$\begin{split} \mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s}, \mathbf{a} \right] &= \mathbb{E}\left[\left. \mathbf{s}_{+}^{\top} W \, \mathbf{s}_{+} + \mathbf{d}^{\top} \mathbf{s}_{+} + V_{0} \right| \, \mathbf{s}, \mathbf{a} \right] \\ &= \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right)^{\top} W \, \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right) + \mathbf{d}^{\top} \mathbf{f}\left(\mathbf{s}, \mathbf{a}\right) + V_{0} + \mathbb{E}\left[\mathbf{w}^{\top} W \, \mathbf{w} \right] \\ &= V^{\star}\left(\mathbf{f}\left(\mathbf{s}, \mathbf{a}\right)\right) + \underbrace{\mathbb{E}\left[\mathbf{w}^{\top} W \, \mathbf{w} \right]}_{= \mathbf{c}} \end{split}$$

Hence theory predicts MPC produces optimal policy!

S. Gros (NTNU)

MDP & SDP

Consider the dynamics:

 $s_+=s+a+w$ with $s,a,\,w\in\mathbb{R},$ and $w\sim\mathcal{N}\left(0,\sigma\right)$ i.i.d.

on a restricted interval.

Cost:

$$L\left(\mathbf{s},\mathbf{a}\right)=\mathbf{a}^{2}+\left(\mathbf{s}-0.5\right)^{2}$$

Constraints:

$$\mathbf{a} \in [-0.25, 0.25]$$



Consider the dynamics:

 $\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ i.i.d. on a restricted interval. Cost:

$$L(s, a) = a^{2} + (s - 0.5)^{2}$$

Constraints:

$$a \in [-0.25, 0.25]$$



S. Gros (NTNU)

MDP & SDP

< 回 > < 三 > < 三

Illustrations - Dissipative / Non-Smooth Problem

Consider the dynamics:

$$\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval. Cost:

$$L(s, a) = |a| + |s - 0.5|$$

Constraints:

 $\mathbf{a}\,\in\,[-0.25,0.25]$

Illustrations - Dissipative / Non-Smooth Problem

Consider the dynamics:

$$\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval.

Cost:

$$L(s, a) = |a| + |s - 0.5|$$

Constraints:

$$\mathbf{a}\,\in\,[-0.25,0.25]$$



Illustrations - Dissipative / Non-Smooth Problem

Consider the dynamics:

 $\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval. Cost:

$$L(s, a) = |a| + |s - 0.5|$$

Constraints:

$$\mathbf{a}\,\in\,[-0.25,0.25]$$



S. Gros (NTNU)

MDP & SDP

< 回 > < 三 > < 三 >

Illustrations - Non-Dissipative / Non-smooth Problem

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with s, a, $\mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval.

What is this problem?

Cost:

$$L(\mathbf{s}, \mathbf{a}) = \begin{cases} \mathbf{a} & \text{if } \mathbf{a} \leq 0\\ 2\mathbf{a} & \text{if } \mathbf{a} > 0 \end{cases}$$

Constraints:

$${\bf s}\,\in\,[0,1],\quad {\bf a}\,\in\,[-0.25,0.25]$$

Illustrations - Non-Dissipative / Non-smooth Problem

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with s, a, $\mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval.

What is this problem?

$$L(\mathbf{s}, \mathbf{a}) = \begin{cases} \mathbf{a} & \text{if } \mathbf{a} \le \mathbf{0} \\ 2\mathbf{a} & \text{if } \mathbf{a} > \mathbf{0} \end{cases}$$

Constraints:

$${\bf s}\,\in\,[0,1],\quad {\bf a}\,\in\,[-0.25,0.25]$$



Illustrations - Non-Dissipative / Non-smooth Problem

Consider the dynamics:

$$\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$$

What is this problem?

with $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$, and $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval.

Cost:

$$L(\mathbf{s}, \mathbf{a}) = \begin{cases} \mathbf{a} & \text{if } \mathbf{a} \leq 0\\ 2\mathbf{a} & \text{if } \mathbf{a} > 0 \end{cases}$$

Constraints:

$${\bf s}\,\in\,[0,1],\quad {\bf a}\,\in\,[-0.25,0.25]$$



S. Gros (NTNU)

MDP & SDP

Fall, 2023

20 / 29

Outline

MPC & MDP: Let's rehearse the background

2 MPC Model for Performance

- 3 Optimal MPC models
 - 4 Stochastic MPC models





Stochastic state transition $\varrho \left[\left. \mathbf{s}_{+} \right| \mathbf{s}, \mathbf{a} \right]$

Model likelihood:

 $\varrho\left[\mathbf{f}\left(\mathbf{s},\mathbf{a}\right) \mid \mathbf{s},\mathbf{a}\right] > \mathbf{0}$

<ロト < 四ト < 回ト < 回ト

is desired, ideally maximal

◆ ■ ● ○ < ○
 Fall, 2023 22 / 29



Stochastic state transition $\varrho \left[\left. \mathbf{s}_{+} \right| \mathbf{s}, \mathbf{a} \right]$

Model likelihood:

 $\varrho\left[\mathbf{f}\left(\mathbf{s},\mathbf{a}\right) \mid \mathbf{s},\mathbf{a}\right] > \mathbf{0}$

<ロト < 四ト < 回ト < 回ト

is desired, ideally maximal

◆ ■ ● ○ < ○
 Fall, 2023 22 / 29



Stochastic state transition $\varrho\left[\,\mathbf{s}_{+}\,|\,\mathbf{s},\mathbf{a}\,\right]$

Model likelihood:

 $\varrho \left[\mathbf{f} \left(\mathbf{s}, \mathbf{a} \right) \mid \mathbf{s}, \mathbf{a} \right] > \mathbf{0}$

is desired, ideally maximal

Existence yes... but not fully clear yet For c = 0 and V^* continuous and ϱ of convex support, there is a f(s, a) with $\varrho[f(s, a) | s, a] > 0$ and (1) for all s, a.

< □ > < □ > < □ > < □ > < □ >



Stochastic state transition $\varrho \left[\mathbf{s}_{+} \, | \, \mathbf{s}, \mathbf{a} \, \right]$

Model likelihood:

 $\varrho\left[{\,{\bf f}\left({{\bf s},{\bf a}} \right)\,\left| \,{{\bf s},{\bf a}} \,\right]} > 0 \right.$

is desired, ideally maximal

Uniqueness no... but max likelihood

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

 \mathbf{s}_+

Fall, 2023 22 / 29



S. Gros (NTNU)

Fall, 2023 22 / 29

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with

• $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$

• $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval

Cost: $L(\mathbf{s}, \mathbf{a}) = |\mathbf{a}| + |\mathbf{s} - 0.5|$ Constraints: $\mathbf{s} \in [0, 1], \quad \mathbf{a} \in [-0.25, 0.25]$

イロト 不得 トイラト イラト 一日

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with

• $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$

• $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval

Cost: $L(\mathbf{s}, \mathbf{a}) = |\mathbf{a}| + |\mathbf{s} - 0.5|$ Constraints: $\mathbf{s} \in [0, 1], \quad \mathbf{a} \in [-0.25, 0.25]$

Expected value model:

$$\mathbb{E}[\mathbf{s}_+] = \mathbf{s} + \mathbf{a}$$

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with

- $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$
- $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval

Cost: $\mathcal{L}(s, \mathbf{a}) = |\mathbf{a}| + |s - 0.5|$ Constraints: $s \in [0, 1], \quad \mathbf{a} \in [-0.25, 0.25]$

Expected value model:

$$\mathbb{E}[\mathbf{s}_+] = \mathbf{s} + \mathbf{\epsilon}$$

Max Likelihood optimal model(s)

$$\begin{split} \mathcal{C}_{c}\left(\mathbf{s},\mathbf{a}\right) &= \\ \underset{\hat{\mathbf{s}}_{+}}{\operatorname{arg\,max}} \quad \varrho\left[\hat{\mathbf{s}}_{+} \mid \mathbf{s},\mathbf{a}\right] \\ \text{s.t.} \quad V^{\star}\left(\hat{\mathbf{s}}_{+}\right) &= \mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s},\mathbf{a}\right] - c \end{split}$$

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

Consider the dynamics: Cost: L(s, a) = |a| + |s - 0.5| $\mathbf{s}_{+} = \mathbf{s} + \mathbf{a} + \mathbf{w}$ with **Constraints:** • $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$ $s \in [0, 1], a \in [-0.25, 0.25]$ • $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval c = 0.1Expected value model: c = 0.13c = 0.15 $\mathbb{E}[\mathbf{s}_+] = \mathbf{s} + \mathbf{a}$ 0.8 0.6 Max Likelihood optimal model(s) (s, a)0.4 $\mathbf{f}_{c}(\mathbf{s},\mathbf{a}) =$ arg max $\varrho[\hat{\mathbf{s}}_+ | \mathbf{s}, \mathbf{a}]$ 0.2 ŝ₊ s.t. $V^{\star}(\hat{\mathbf{s}}_{+}) = \mathbb{E}[V^{\star}(\mathbf{s}_{+}) | \mathbf{s}, \mathbf{a}] - c$ 0.0 0.2 0.4 0.6 0.8 0.0 1.0 s + aS. Gros (NTNU) MDP & SDP Fall, 2023 23 / 29

Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with

- $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$
- $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval

Cost: L(s, a) = |a| + |s - 0.5|Constraints:

 ${\bf s}\,\in\,[0,1],\quad {\bf a}\,\in\,[-0.25,0.25]$

Observations

- Existence: not for all c
- Continuity: not for all c
- Both: specific c(?)
- Linearity is lost



Consider the dynamics:

 $\mathbf{s}_+ = \mathbf{s} + \mathbf{a} + \mathbf{w}$

with

• $\mathbf{s}, \mathbf{a}, \mathbf{w} \in \mathbb{R}$

• $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma)$ on a restricted interval

Cost: L(s, a) = |a| + |s - 0.5|Constraints: $s \in [0, 1], a \in [-0.25, 0.25]$

Observations

- Existence: not for all c
- Continuity: not for all c
- Both: specific c(?)
- Linearity is lost

This is to be further investigated



S. Gros (NTNU)

Model
$$\mathbf{f}(\mathbf{s}, \mathbf{a})$$
 such that
 $V^{\star}(\mathbf{f}(\mathbf{s}, \mathbf{a})) = \mathbb{E}\left[V^{\star}(\mathbf{s}_{+}) \mid \mathbf{s}, \mathbf{a}
ight] - c$
holds for some c ???

What if $\mathbf{s}, \mathbf{a} \rightarrow$ possible infeasibility? I.e.

$$\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \mid \mathrm{s},\mathrm{a}
ight] = \infty$$
 (1)

Â

イロト イポト イヨト イヨ

holds.

What should the model do then?







What if $s, a \rightarrow possible infeasibility?$ I.e.

$$\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \mid \mathrm{s},\mathrm{a}
ight] = \infty$$
 (1)

holds.

What should the model do then?

• (1) implies that:

$$\mathbb{P}\left[\left.V^{\star}\left(\mathbf{s}_{+}\right)=\infty\left.\right|\mathbf{s},\mathbf{a}\right.\right]>0$$

i.e. s_+ may land where V^\star is ∞

Â

- 4 回 ト 4 ヨ ト 4 ヨ ト





What if $\mathbf{s}, \mathbf{a} \rightarrow$ possible infeasibility? I.e.

$$\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \mid \mathrm{s},\mathrm{a}
ight] = \infty$$
 (1)

holds.

What should the model do then?

• (1) implies that:

 $\mathbb{P}\left[\left.V^{\star}\left(s_{+}\right)=\infty\left|\left.s,a\right.\right]>0\right.$

i.e. s_+ may land where V^\star is ∞

Model must reproduce that, i.e.

$$V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight) =\infty$$

i.e. $f\left({{\mathbf{s}},{\mathbf{a}}} \right)$ picks a point among the infeasible onesÂ

イロト イポト イヨト イヨト

 \mathbf{s}_+



What if $\mathbf{s}, \mathbf{a} \rightarrow$ possible infeasibility? I.e.

$$\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \mid \mathrm{s},\mathrm{a}
ight] = \infty$$
 (1)

holds.

What should the model do then?

• (1) implies that:

$$\mathbb{P}\left[\left.V^{\star}\left(\mathbf{s}_{+}\right)=\infty\left|\left.\mathbf{s},\mathbf{a}\right.\right]>0\right.$$

i.e. s_+ may land where V^\star is ∞

• Model must reproduce that, i.e.

$$V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight)=\infty$$

i.e. $\mathbf{f}\left(\mathbf{s},\mathbf{a}\right)$ picks a point among the infeasible onesÂ

Theory requires conservative model for constraints violations

S. Gros (NTNU)

MDP & SDP

Outline

MPC & MDP: Let's rehearse the background

Stochastic MPC models 4



MPC with stochastic models?

$$\begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \underset{x,u}{\min} \quad \mathcal{T}\left(x_{N}\right) + \sum_{k=0}^{N-1} \mathcal{L}\left(x_{k}, \mathbf{u}_{k}\right) \\ \text{s.t.} \quad x_{k+1} = \mathbf{f}\left(x_{k}, \mathbf{u}_{k}\right) \\ x_{0} = \mathbf{s} \end{array} \qquad \begin{array}{l} \text{MPC policy } \pi^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \underset{u_{0}, \pi_{1, \dots, N-1}}{\min} \quad \mathbb{E}\left[\mathcal{T}(x_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(x_{k}, \mathbf{u}_{k})\right] \\ \text{s.t.} \quad x_{k+1} \sim \hat{\varrho}[\, . \, | \, x_{k}, \mathbf{u}_{k} \,], \\ x_{0} = \mathbf{s}, \quad \mathbf{u}_{k} = \pi_{k}\left(x_{k}\right), \, k > 0 \end{array}$$

 ▲ ■
 ■
 ∽ ۹

 Fall, 2023
 26 / 29

<ロト < 四ト < 回ト < 回ト

MPC with stochastic models?

 $\begin{array}{l} \text{MPC policy } \pi^{\mathrm{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{x},\mathbf{u}} \quad \mathcal{T}\left(\mathbf{x}_{N}\right) + \sum_{k=0}^{N-1} L\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ \text{s.t.} \quad \mathbf{x}_{k+1} = \mathbf{f}\left(\mathbf{x}_{k},\mathbf{u}_{k}\right) \\ \mathbf{x}_{0} = \mathbf{s} \end{array}$



MPC policy $\pi^{\mathrm{MPC}}(\mathbf{s}) = \mathbf{u}_0^{\star}$ from $\min_{\mathbf{u}_0, \boldsymbol{\pi}_1, \dots, N-1} \mathbb{E}\left[T(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k L(\mathbf{x}_k, \mathbf{u}_k)\right]$ s.t $\mathbf{x}_{k+1} \sim \hat{\rho}[. | \mathbf{x}_k, \mathbf{u}_k],$ $\mathbf{x}_0 = \mathbf{s}, \quad \mathbf{u}_k = \boldsymbol{\pi}_k(\mathbf{x}_k), \ k > 0$



Policy from planning

Proper policing, but difficult

<ロト <部ト <きト <きト = 1

S. Gros (NTNU)

MDP & SDP

Fall, 2023 26 / 29
MPC with stochastic models - Scenario trees

MPC policy
$$\pi^{\text{MPC}}(\mathbf{s}) = \mathbf{u}_0^{\star}$$
 from

$$\min_{\mathbf{u}_0, \pi_1, \dots, N-1} \mathbb{E}\left[T(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k L(\mathbf{x}_k, \mathbf{u}_k)\right]$$
s.t $\mathbf{x}_{k+1} \sim \hat{\varrho}[. | \mathbf{x}_k, \mathbf{u}_k],$
 $\mathbf{x}_0 = \mathbf{s}, \quad \mathbf{u}_k = \pi_k(\mathbf{x}_k), \ k > 0$

Scenario tree MPC

- $\hat{\varrho}$ is a discrete probability distribution
- Tree of scenarios
- Implicitly produces decision policies
- Exploding complexity over horizon



MDP & SDP

MPC with stochastic models - Other methods

$$\begin{aligned} & \text{MPC policy } \boldsymbol{\pi}^{\text{MPC}}\left(s\right) = \mathbf{u}_{0}^{\star} \text{ from} \\ & \min_{\mathbf{u}_{0}, \boldsymbol{\pi}_{1}, \dots, N-1} \quad \mathbb{E}\left[\mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right] \\ & \text{ s.t } \quad \mathbf{x}_{k+1} \sim \hat{\varrho}[\,.\,|\,\mathbf{x}_{k}, \mathbf{u}_{k}\,], \\ & \mathbf{x}_{0} = \mathbf{s}, \quad \mathbf{u}_{k} = \boldsymbol{\pi}_{k}\left(\mathbf{x}_{k}\right), \, k > 0 \end{aligned} \end{aligned}$$

"Spectral" representations of $\hat{\varrho}$

- Gaussian Processes
- Polynomial Chaos Expansion
- RKHS

Representations of $\pi_{1,...,N-1}$

Linear feedback

$$oldsymbol{\pi}_{k}\left(\mathbf{s}_{k}
ight)=\mathbf{ar{u}}_{k}-\mathcal{K}_{k}\left(\mathbf{s}_{k}-\mathbf{ar{s}}_{k}
ight)$$

More advance forms...

Model $f\left(s,a\right)$ such that

 $V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight)=\mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}
ight)\mid\mathbf{s},\mathbf{a}
ight]-c$

A (10) A (10) A (10)

Model $f\left(s,a\right)$ such that

 $V^{\star}\left(\mathbf{f}\left(\mathbf{s},\mathbf{a}
ight)
ight)=\mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}
ight)\mid\mathbf{s},\mathbf{a}
ight]-c$

Generalization?

Model $\hat{\mathbf{s}}_{+}\sim\hat{\varrho}[\,.\,|\,\mathbf{s},\mathbf{a}\,]$ such that

 $\mathbb{E}\left[\mathcal{V}^{\star}\left(\hat{\mathrm{s}}_{+}
ight) \left| \, \mathrm{s},\mathrm{a}
ight] =\mathbb{E}\left[\mathcal{V}^{\star}\left(\mathrm{s}_{+}
ight) \left| \, \mathrm{s},\mathrm{a}
ight] -c$

イロト イポト イヨト イヨト

Model $f\left(s,a\right)$ such that

 $V^{\star}\left(\mathrm{f}\left(\mathrm{s},\mathrm{a}
ight)
ight) =\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \left| \mathrm{s},\mathrm{a}
ight] -c
ight.$

Generalization?

$$\begin{split} \text{Model } \hat{s}_{+} &\sim \hat{\varrho}[\,.\,|\,s,a] \text{ such that} \\ &\mathbb{E}\left[\mathcal{V}^{\star}\left(\hat{s}_{+}\right)\,|\,s,a\right] = \mathbb{E}\left[\mathcal{V}^{\star}\left(s_{+}\right)\,|\,s,a\right] - c \end{split}$$

$$\begin{split} \text{MPC policy } \pi^{\text{MPC}}\left(\mathbf{s}\right) &= \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{u}_{0}, \pi_{1, \dots, N-1}} & \mathbb{E}\left[\mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right] \\ \text{ s.t } & \mathbf{x}_{k+1} \sim \hat{\varrho}[\, . \, | \, \mathbf{x}_{k}, \mathbf{u}_{k} \,], \\ & \mathbf{x}_{0} = \mathbf{s}, \quad \mathbf{u}_{k} = \pi_{k}\left(\mathbf{x}_{k}\right) \\ \textbf{ yields } \pi^{\text{MPC}}\left(\mathbf{s}\right) &= \pi^{\star}\left(\mathbf{s}\right) \text{ (with correct } T) \end{split}$$

イロト イボト イヨト イヨト

Model $f\left(s,a\right)$ such that

 $V^{\star}\left(\mathrm{f}\left(\mathrm{s},\mathrm{a}
ight)
ight) =\mathbb{E}\left[V^{\star}\left(\mathrm{s}_{+}
ight) \left| \mathrm{s},\mathrm{a}
ight] -c
ight.$

Generalization?

$$\begin{split} \text{Model } \hat{s}_{+} &\sim \hat{\varrho}[\,.\,|\,s,a] \text{ such that} \\ &\mathbb{E}\left[\mathcal{V}^{\star}\left(\hat{s}_{+}\right)\,|\,s,a\right] = \mathbb{E}\left[\mathcal{V}^{\star}\left(s_{+}\right)\,|\,s,a\right] - c \end{split}$$

$$\begin{split} \text{MPC policy } \pi^{\text{MPC}}\left(\mathbf{s}\right) &= \mathbf{u}_{0}^{\star} \text{ from} \\ \min_{\mathbf{u}_{0}, \pi_{1, \dots, N-1}} & \mathbb{E}\left[\mathcal{T}(\mathbf{x}_{N}) + \sum_{k=0}^{N-1} \gamma^{k} \mathcal{L}(\mathbf{x}_{k}, \mathbf{u}_{k})\right] \\ &\text{ s.t } \mathbf{x}_{k+1} \sim \hat{\varrho}[. \mid \mathbf{x}_{k}, \mathbf{u}_{k}], \\ &\mathbf{x}_{0} = \mathbf{s}, \quad \mathbf{u}_{k} = \pi_{k}\left(\mathbf{x}_{k}\right) \\ & \text{ yields } \pi^{\text{MPC}}\left(\mathbf{s}\right) = \pi^{\star}\left(\mathbf{s}\right) \text{ (with correct T)} \end{split}$$

To be further explored!

S. Gros (NTNU)

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > <

Model f(s, a) such that $V^{\star}(\mathbf{f}(\mathbf{s},\mathbf{a})) = \mathbb{E}\left[V^{\star}(\mathbf{s}_{+}) \mid \mathbf{s},\mathbf{a}\right] - c$ **Generalization?** Model $\hat{\mathbf{s}}_+ \sim \hat{\varrho}[\,.\,|\,\mathbf{s},\mathbf{a}\,]$ such that $\mathbb{E}\left[V^{\star}\left(\hat{\mathbf{s}}_{+}\right) \mid \mathbf{s}, \mathbf{a}\right] = \mathbb{E}\left[V^{\star}\left(\mathbf{s}_{+}\right) \mid \mathbf{s}, \mathbf{a}\right] - c$ **MPC policy** $\pi^{\text{MPC}}(s) = \mathbf{u}_0^{\star}$ from $\min_{\mathbf{u}_0, \boldsymbol{\pi}_1, \dots, N-1} \mathbb{E} \left[T(\mathbf{x}_N) + \sum_{k=0}^{N-1} \gamma^k \mathcal{L}(\mathbf{x}_k, \mathbf{u}_k) \right]$ s.t $\mathbf{x}_{k+1} \sim \hat{\rho}[.|\mathbf{x}_k, \mathbf{u}_k],$ $\mathbf{x}_0 = \mathbf{s}, \quad \mathbf{u}_k = \boldsymbol{\pi}_k (\mathbf{x}_k)$ yields $\pi^{\mathrm{MPC}}\left(\mathrm{s}
ight)=\pi^{\star}\left(\mathrm{s}
ight)$ (with correct T)



To be further explored!

S. Gros (NTNU)

Fall, 2023 29 / 29

Model f(s, a) such that $V^*(f(s, a)) = \mathbb{E}[V^*(s_+) | s, a] - c$ Discrete Stochastic Model Models $f_{1,...,m}(s, a)$ such that $\sum_{i=1}^{m} \omega^i V^*(f_i(s, a)) = \mathbb{E}[V^*(s_+) | s, a] - c$

this describes a scenario tree !!



MDP & SDP