Suboptimality — example

We consider two examples with \( X = \mathbb{R}, U = \mathbb{R} \) for \( N = 2 \)

**Example 1:** \( x^+ = x + u, \quad \ell(x, u) = x^2 + u^2 \)

Terminal constraints \( x_u(N) = x_\star = 0 \)
\( V_\infty(x) \approx 1.618x^2, \quad J^{cl}_\infty(x, \mu_2) = 1.625x^2 \)

**Example 2:** as Example 1, but with \( \ell(x, u) = x^2 + u^4 \)
\( V_\infty(20) \leq 1726, \quad J^{cl}_\infty(x, \mu_2) \approx 11240 \)

General estimates for fixed \( N \) appear difficult to obtain. But we can give an asymptotic result for \( N \to \infty \)

Asymptotic Suboptimality

**Theorem:** For both types of terminal constraints the assumptions of the stability theorems ensure
\( V_N(x) \to V_\infty(x) \)
and thus
\( J^{cl}_N(x, \mu_N) \to V_\infty(x) \)
as \( N \to \infty \) uniformly on compact subsets of the feasible sets, i.e., the MPC performance converges to the optimal one

**Idea of proof:** uses that any approximately optimal trajectory for \( J_\infty \) converges to \( x_\star \) and can thus be modified to meet the constraints with only moderately changing its value

Summary of Section (4)

- \( \mu_N \) is infinite horizon optimal for a suitably altered running cost
- the infinite horizon functional along the \( \mu_N \)-controlled trajectory is bounded by \( V_N \), i.e.,
\[
J^{cl}_\infty(x, \mu_N) \leq V_N(x)
\]
- \( V_N \gg V_\infty \) is possible under terminal constraints
- \( V_N \to V_\infty \) holds for \( N \to \infty \)

(5) Stability and suboptimality without stabilizing constraints
MPC without stabilizing terminal constraints

We return to the basic MPC formulation

$$\text{minimize} \quad J_N(x_0, u) = \sum_{k=0}^{N-1} \ell(x_{u}(k), u(k)), \quad x_u(0) = x_0 = x_{\mu_N}(n)$$

without any stabilizing terminal constraints and costs

In order to motivate why we want to avoid terminal constraints and costs, we consider an example of $P$ double integrators in the plane

A motivating example for avoiding terminal constraints

Example: [Jahn ’10] Consider $P$ 4-dimensional systems

$$\dot{x}_i = f(x_i, u_i) := (x_{i2}, u_{i1}, x_{i4}, u_{i2})^T, \quad i = 1, \ldots, P$$

Interpretation: $(x_{i1}, x_{i3})^T = \text{position}$, $(x_{i2}, x_{i4})^T = \text{velocity}$

Stage cost: $\ell(x, u) = \sum_{i=1}^{P} ||(x_{i1}, x_{i3})^T - x_d|| + ||(x_{i2}, x_{i4})^T||/50$

with $x_d = (0, 0)^T$ until $t = 20s$ and $x_d = (3, 0)^T$ afterwards

Constraints: no collision, obstacles, limited speed and control

The simulation shows MPC for $P = 128$ ($\sim$ system dimension 512) with sampling time $T = 0.02s$ and horizon $N = 6$

Stabilizing NMPC without terminal constraint

(Some) stability and performance results known in the literature:

[Alamir/Bornard ’95] use a controllability condition for all $x \in \mathbb{X}$

[Shamma/Xiong ’97, Primbs/Nevišić ’00] use knowledge of optimal value functions

[Jadbabaie/Hauser ’05] use controllability of linearization in $x_s$

[Grimm/Messina/Tuna/Teel ’05, Tuna/Messina/Teel ’06, Gr./Rantzer ’08, Gr. ’09, Gr./Pannek/Seehafer/Worthmann ’10] use bounds on optimal value functions

Here we explain the last approach

Bounds on the optimal value function

Recall the definition of the optimal value function

$$V_N(x) := \inf_{u \text{ admissible}} \sum_{k=0}^{N-1} \ell(x_{u}(k), u(k))$$

Boundedness assumption: there exists $\gamma > 0$ with

$$V_N(x) \leq \gamma \ell^*(x) \quad \text{for all} \quad x \in \mathbb{X}, N \in \mathbb{N}$$

where $\ell^*(x) := \min_{u \in U} \ell(x, u)$

(sufficient conditions for and relaxations of this bound will be discussed later)
Stability and performance index

We choose $\ell$, such that

$$\alpha_3(\|x-x_\ast\|) \leq \ell^\ast(x) \leq \alpha_4(\|x-x_\ast\|)$$

holds for $\alpha_3, \alpha_4 \in \mathcal{K}_\infty$ (again, $\ell(x,u) = \|x-x_\ast\|^2 + \lambda\|u\|^2$ works)

Then, the only inequality left to prove in order to apply the relaxed dynamic programming theorem is

$$V_N(f(x, \mu_N(x))) \leq V_N(x) - \alpha_N \ell(x, \mu_N(x))$$

for some $\alpha_N \in (0,1)$ and all $x \in \mathcal{X}$

We can compute $\alpha_N$ from the bound $V_N(x) \leq \gamma \ell^\ast(x)$

Decay of the optimal trajectory

We assume $V_N(x) \leq \gamma \ell^\ast(x)$ for all $x \in \mathcal{X}, N \in \mathbb{N}$

We want $\eta_N > 0, k^* \geq 1$ with $\ell^\ast(x^\ast(k^*)) \leq \eta_N \ell^\ast(x^\ast(0))$

Variant 1 [Grimm/Messina/Tuna/Teel ’05]

$V_N(x) \leq \gamma \ell^\ast(x) \Rightarrow \ell(x^\ast(k^*), u^\ast(k^*)) \leq \gamma \ell^\ast(x)/N$ for at least one $k^* \Rightarrow \alpha_N = 1 - \gamma(\gamma - 1)/N$

**Computing $\alpha_N$**

We assume $V_N(x) \leq \gamma \ell^\ast(x)$ for all $x \in \mathcal{X}, N \in \mathbb{N}$

We want $V_N(x^\ast(1)) \leq V_N(x^\ast(0)) - \alpha_N \ell(x^\ast(0), u^\ast(0))$

• use (*) to find $\eta_N > 0, k^* \geq 1$ with $\ell^\ast(x^\ast(k^*)) \leq \eta_N \ell^\ast(x^\ast(0))$

• concatenate $x^\ast(1), \ldots, x^\ast(k^*)$ and the optimal trajectory starting in $x^\ast(k^*): \equiv \tilde{x}(\cdot), \tilde{u}(\cdot)$

$$\Rightarrow V_N(x^\ast(1)) \leq J_N(x^\ast(1), \tilde{u}) \leq V_N(x^\ast(0)) - \frac{(1 - \gamma \eta_N) \ell(x^\ast(0), u^\ast(0))}{\alpha_N}$$

Decay of the optimal trajectory

We assume $V_N(x) \leq \gamma \ell^\ast(x)$ for all $x \in \mathcal{X}, N \in \mathbb{N}$

We want $\eta_N > 0, k^* \geq 1$ with $\ell^\ast(x^\ast(k^*)) \leq \eta_N \ell^\ast(x^\ast(0))$

Variant 2 [Tuna/Messina/Teel ’06, Gr./Rantzer ’08]

$V_N(x) \leq \gamma \ell^\ast(x) \Rightarrow \ell(x^\ast(k), u^\ast(k)) \leq \gamma \left(\frac{2-\gamma}{\gamma}\right)^k \ell^\ast(x)$

$$\Rightarrow k^* = N - 1 \Rightarrow \alpha_N = 1 - (\gamma - 1)^N/\gamma^{N-2}$$
**Decay of the optimal trajectory**

We assume $V_N(x) \leq \gamma \ell^*(x)$ for all $x \in \mathbb{X}, N \in \mathbb{N}$

We want $\eta_N > 0, k^* \geq 1$ with $\ell^*(x^*(k^*)) \leq \eta_N \ell^*(x^*(0))$

**Variant 3** [Gr. ’09, Gr./Pannek/Seehafer/Worthmann ’10]

$V_N(x) \leq \gamma \ell^*(x)$ \Rightarrow formulate all constraints and trajectories

\[ \Rightarrow \text{optimize for } \alpha_N \Rightarrow \alpha_N = 1 - \frac{(\gamma-1)^N}{(\gamma-1)^N - 1} \]

\[
\begin{align*}
V_N(x^*(1)) &\leq \gamma \ell^*(x^*(1)) \\
V_N(x^*(1)) &\leq \ell(x^*(1), u^*(1)) + \gamma \ell^*(x^*(2)) \\
V_N(x^*(1)) &\leq \ell(x^*(1), u^*(1)) + \ell(x^*(2), u^*(2)) + \gamma \ell^*(x^*(3)) \\
& \vdots & \vdots & \vdots
\end{align*}
\]

**Optimization approach to compute $\alpha_N$**

We explain the optimization approach (Variant 3) in more detail. We want $\alpha_N$ such that $V_N(x^*(1)) \leq V_N(x^*(0)) - \alpha_N \ell(x^*(0), u^*(0))$

We assume $\alpha_N^\star$ for all optimal trajectories $x^*(n), u^*(n)$ for $V_N$

The bound and the dynamic programming principle imply:

\[ V_N(x^*(1)) \leq \gamma \ell^*(x^*(1)) \]

\[ V_N(x^*(1)) \leq \ell(x^*(1), u^*(1)) + \gamma \ell^*(x^*(2)) \]

\[ V_N(x^*(1)) \leq \ell(x^*(1), u^*(1)) + \ell(x^*(2), u^*(2)) + \gamma \ell^*(x^*(3)) \]

\[ \vdots \]

**Verifying the relaxed Lyapunov inequality**

Find $\alpha_N$, such that for all optimal trajectories $x^*, u^*$:

\[ V_N(x^*(1)) \leq V_N(x^*(0)) - \alpha_N \ell(x^*(0), u^*(0)) \quad (*) \]

Define $\lambda_n := \ell(x^*(n), u^*(n)), \quad \nu := V_N(x^*(1))$

Then: \((*) \Leftrightarrow \nu \leq \sum_{n=0}^{N-1} \lambda_n - \alpha_N \lambda_0 \)

The inequalities from the last slides translate to

\[ \sum_{n=k}^{N-1} \lambda_n \leq \gamma \lambda_k, \quad k = 0, \ldots, N - 2 \quad (1) \]

\[ \nu \leq \sum_{n=1}^{j} \lambda_n + \gamma \lambda_{j+1}, \quad j = 0, \ldots, N - 2 \quad (2) \]

We call $\lambda_0, \ldots, \lambda_{N-1}, \nu \geq 0$ with \((1), (2)\) admissible
Optimization problem

⇒ if $\alpha_N$ is such that the inequality
\[
\nu \leq \sum_{n=0}^{N-1} \lambda_n - \alpha_N \lambda_0 \iff \alpha_N \leq \frac{\sum_{n=0}^{N-1} \lambda_n - \nu}{\lambda_0}
\]
holds for all admissible $\lambda_n$ and $\nu$, then the desired inequality will hold for all optimal trajectories

The largest $\alpha_N$ satisfying this condition is
\[
\alpha_N := \min_{\lambda_n, \nu \text{ admissible}} \frac{\sum_{n=0}^{N-1} \lambda_n - \nu}{\lambda_0}
\]

This is a linear optimization problem whose solution can be computed explicitly (which is nontrivial) and reads
\[
\alpha_N = 1 - \frac{(\gamma - 1)^N}{\gamma^{N-1} - (\gamma - 1)^{N-1}}
\]

Horizon dependent $\gamma$-values

The theorem remains valid if we replace the bound condition
\[
V_N(x) \leq \gamma \ell^*(x)
\]
by
\[
V_N(x) \leq \gamma_N \ell^*(x)
\]
for horizon-dependent bounded values $\gamma_N \in \mathbb{R}$, $N \in \mathbb{N}$

\[
\alpha_N = 1 - \frac{(\gamma - 1)^N}{\prod_{i=2}^{N} (\gamma_i - 1)}
\]

This allows for tighter bounds and a refined analysis.

Stability and performance theorem

Theorem: [Gr./Pannek/Seehafer/Worthmann '10]: Assume $V_N(x) \leq \gamma \ell^*(x)$ for all $x \in X$, $N \in \mathbb{N}$. If
\[
\alpha_N > 0 \iff N > 2 + \frac{\ln(\gamma - 1)}{\ln \gamma - \ln(\gamma - 1)} \sim \gamma \ln \gamma
\]
then the NMPC closed loop is asymptotically stable with Lyapunov function $V_N$ and we get the performance estimate $J_{\infty}^c(x, \mu_N) \leq V_{\infty}(x)/\alpha_N$ with
\[
\alpha_N = 1 - \frac{(\gamma - 1)^N}{\gamma^{N-1} - (\gamma - 1)^{N-1}} \to 1 \text{ as } N \to \infty
\]

Conversely, if $N < 2 + \frac{\ln(\gamma - 1)}{\ln \gamma - \ln(\gamma - 1)}$, then there exists a system for which $V_N(x) \leq \gamma \ell^*(x)$ holds but the NMPC closed loop is not asymptotically stable.

Controllability condition

A refined analysis can be performed if we compute $\gamma_N$ from a controllability condition, e.g., exponential controllability:

Assume that for each $x_0 \in X$ there exists an admissible control $u$ such that
\[
\ell(x_u(k), u(k)) \leq C \sigma^k \ell^*(x_0), \quad k = 0, 1, 2, \ldots
\]
for given overshoot constant $C > 0$ and decay rate $\sigma \in (0, 1)$

\[
\Rightarrow V_N(x) \leq \gamma_N \ell^*(x) \quad \text{for} \quad \gamma_N = \sum_{k=0}^{N-1} C \sigma^k
\]

This allows to compute the minimal stabilizing horizon
\[
\min\{N \in \mathbb{N} | \alpha_N > 0\}
\]
depending on $C$ and $\sigma$.
Stability chart for $C$ and $\sigma$

(Figure: Harald Voit)

**Conclusion**: for short optimization horizon $N$ it is
more important: small $C$ ("small overshoot")
less important: small $\sigma$ ("fast decay")
(we will see in the next section how to use this information)

Comments and extensions

The "linear" inequality $V_N(x) \leq \gamma \ell^*(x)$ may be too demanding for nonlinear systems under constraints

**Generalization**: $V_N(x) \leq \rho(\ell^*(x)), \quad \rho \in K_\infty$

- there is $\gamma > 0$ with $\rho(r) \leq \gamma r$ for all $r \in [0, \infty]$
  $\Rightarrow$ **global asymptotic stability**

- for each $R > 0$
  there is $\gamma_R > 0$ with $\rho(r) \leq \gamma_R r$ for all $r \in [0, R]$
  $\Rightarrow$ **semiglobal asymptotic stability**

- $\rho \in K_\infty$ arbitrary
  $\Rightarrow$ **semiglobal practical asymptotic stability**

[Grimm/Messina/Tuna/Teel '05, Gr./Pannek '11]

Comments and extensions

- for unconstrained linear quadratic problems:
  existence of $\gamma \iff (A, B)$ stabilizable

- additional weights on the last term can be incorporated into the analysis [Gr./Pannek/Seehafer/Worthmann '10]

- instead of using $\gamma$, $\alpha$ can be estimated numerically online along the closed loop [Pannek et al. '10ff]

- positive definiteness of $\ell$ can be replaced by a detectability condition [Grimm/Messina/Tuna/Teel '05]

- under appropriate uniformity assumptions, the results are easily carried over to tracking time variant references $x_{ref}(n)$ instead of an equilibrium $x_*$ [Gr./Pannek '11]

Summary of Section (5)

- Stability and performance of MPC without terminal constraints can be ensured by suitable bounds on $V_N$

- An optimization approach allows to compute the best possible $\alpha_N$ in the relaxed dynamic programming theorem

- The $\gamma$ or $\gamma_N$ can be computed from controllability properties, e.g., exponential controllability

- The overshoot bound $C > 0$ plays a crucial role or obtaining small stabilizing horizons
(6) Examples for the design of MPC schemes

Design of “good” MPC running costs $\ell$

We want small overshoot $C$ in the estimate

$$\ell(x_u(n), u(n)) \leq C \sigma^\infty \ell^*(x_0)$$

The trajectories $x_u(n)$ are given, but we can use the running cost $\ell$ as design parameter

The car-and-mountains example reloaded

MPC with $\ell(x, u) = \|x - x_s\|^2 + |u|^2$ and $u_{\text{max}} = 0.2$

\(\rightsquigarrow\) asymptotic stability for $N = 11$ but not for $N \leq 10$

Reason: detour around mountains causes large overshoot $C$

Remedy: put larger weight on $x_2$:

$$\ell(x, u) = (x_1 - x_{s,1})^2 + 5(x_2 - x_{s,2})^2 + |u|^2 \quad \rightsquigarrow \text{as. stab. for } N = 2$$

Example: pendulum on a cart

$$\begin{align*}
x_1 &= \theta = \text{angle} \\
x_2 &= \text{angular velocity} \\
x_3 &= \text{cart position} \\
x_4 &= \text{cart velocity} \\
u &= \text{cart acceleration}
\end{align*}$$

\(\rightsquigarrow\) control system

$$\begin{align*}
\dot{x}_1 &= x_2(t) \\
\dot{x}_2 &= -g \sin(x_1) - k x_2 \\
&\quad - u \cos(x_1) \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= u
\end{align*}$$
Example: Inverted Pendulum
Reducing overshoot for swingup of the pendulum on a cart:
\[
\begin{align*}
  \dot{x}_1 &= x_2, & \dot{x}_2 &= g \sin(x_1) - k x_2 + u \cos(x_1) \\
  \dot{x}_3 &= x_4, & \dot{x}_4 &= u
\end{align*}
\]
Let \( \ell(x) = \sqrt{\ell_1(x_1, x_2) + x_3^2 + x_4^2} \) with

\[
\begin{align*}
  \ell_1(x_1, x_2) &= x_2^2 + x_2^2 \\
  4(1 - \cos x_1) + x_2^2 \\
  (\sin x_1, x_2)^T (\sin x_1, x_2)^T + 2(1 - \cos x_1)(1 - \cos x_2)/2
\end{align*}
\]

\( N = 15 \) \( N = 10 \) \( N = 4 \) (swingup only)

A PDE example
We illustrate this with the 1d controlled PDE
\[
y_t = y_x + \nu y_{xx} + \mu y(y + 1)(1 - y) + u
\]
with
\[
\begin{align*}
  \text{domain} \quad \Omega &= [0, 1] \\
  \text{solution} \quad y &= y(t, x) \\
  \text{boundary conditions} \quad y(t, 0) &= y(t, 1) = 0 \\
  \text{parameters} \quad \nu &= 0.1 \text{ and } \mu = 10 \\
  \text{and distributed control} \quad u : \mathbb{R} \times \Omega \to \mathbb{R}
\end{align*}
\]
Discrete time system: \( y(n) = y(nT, \cdot) \), sampling time \( T = 0.025 \)

The uncontrolled PDE
\[
y_t = y_x + \nu y_{xx} + \mu y(y + 1)(1 - y) + u
\]

Goal: stabilize the sampled data system \( y(n) \) at \( y \equiv 0 \)
Usual approach: quadratic \( L^2 \) cost
\[
\ell(y(n), u(n)) = \|y(n)\|_{L^2}^2 + \lambda \|u(n)\|_{L^2}^2
\]
For \( y \approx 0 \) the control \( u \) must compensate for \( y_x \approx u \approx -y_x \)
\sim controllability condition
\[
\ell(y(n), u(n)) \leq C \sigma^n \ell^*(y(0))
\]
\[ \iff \|y(n)\|_{L^2}^2 + \lambda \|u(n)\|_{L^2}^2 \leq C \sigma^n \|y(0)\|_{L^2}^2 \]
\[ \approx \|y(n)\|_{L^2}^2 + \lambda \|y_x(n)\|_{L^2}^2 \leq C \sigma^n \|y(0)\|_{L^2}^2 \]
for \( \|y_x\|_{L^2} \gg \|y\|_{L^2} \) this can only hold if \( C \gg 0 \)

MPC for the PDE example
MPC for the PDE example

Conclusion: because of
\[
\|y(n)\|_{L^2}^2 + \lambda \|y_x(n)\|_{L^2}^2 \leq C\sigma^n \|y(0)\|_{L^2}^2
\]
the controllability condition may only hold for very large \(C\)

Remedy: use \(H^1\) cost
\[
\ell(y(n), u(n)) = \|y(n)\|_{L^2}^2 + \|y_x(n)\|_{L^2}^2 + \lambda \|u(n)\|_{L^2}^2.
\]

Then an analogous computation yields
\[
\|y(n)\|_{L^2}^2 + (1 + \lambda) \|y_x(n)\|_{L^2}^2 \leq C\sigma^n \left(\|y(0)\|_{L^2}^2 + \|y_x(0)\|_{L^2}^2\right)
\]

Boundary Control

Now we change our PDE from distributed to (Dirichlet-) boundary control, i.e.
\[
y_t = y_x + \nu y_{xx} + \mu y(y + 1)(1 - y)
\]
with
- domain \(\Omega = [0, 1]\)
- solution \(y = y(t, x)\)
- boundary conditions \(y(t, 0) = u_0(t), y(t, 1) = u_1(t)\)
- parameters \(\nu = 0.1\) and \(\mu = 10\)

with boundary control, stability can only be achieved via large gradients in the transient phase
\(\sim L^2\) should perform better that \(H^1\)

MPC with \(L^2\) vs. \(H^1\) cost

Boundary control, \(L^2\) vs. \(H^1\), \(N = 20\)

Can be made rigorous for many PDEs [Altmüller et al. ’10ff]
Summary of Section (6)

- Reducing the overshoot constant \( C \) by choosing \( \ell \) appropriately can significantly reduce the horizon \( N \) needed to obtain stability.
- Computing tight estimates for \( C \) is in general a difficult if not impossible task.
- But structural knowledge of the system behavior can be sufficient for choosing a “good” \( \ell \).

Feasibility

Consider the feasible sets

\[ \mathcal{F}_N := \{x \in X | \text{there exists an admissible } u \text{ of length } N\} \]

So far we have assumed

\[ V_N(x) \leq \gamma \ell^*(x) \text{ for all } x \in X \]

which implicitly includes the assumption

\[ \mathcal{F}_N = X \]

because \( V_N(x) = \infty \text{ for } x \in X \setminus \mathcal{F}_N \)

What happens if \( \mathcal{F}_N \neq X \) for some \( N \in \mathbb{N} \)?

(7) Feasibility

The MPC feasibility problem

Even though the open-loop optimal trajectories are forced to satisfy \( x^*(k) \in X \), the closed loop solutions \( x_{\mu_N}(n) \) may violate the state constraints, i.e., \( x_{\mu_N}(n) \notin X \) for some \( n \).

We illustrate this phenomenon by the simple example

\[
\begin{pmatrix}
  x^+_1 \\
  x^+_2
\end{pmatrix}
= \begin{pmatrix}
  x_1 + x_2 + u/2 \\
  x_2 + u
\end{pmatrix}
\]

with \( X = [-1, 1]^2 \) and \( U = [-1/4, 1/4] \). For initial value \( x_0 = (-1, 1)^T \), the system can be controlled to 0 without leaving \( X \).

We use MPC with \( N = 2 \) and \( \ell(x, u) = \|x\|^2 + 5u^2 \).