# Sensitivity and turnpike results for the optimal control of PDEs and their use for model predictive control

#### Lars Grüne

#### Mathematisches Institut, Universität Bayreuth

based on joint work with Manuel Schaller, Anton Schiela (both Bayreuth), Marleen Stieler (Ludwigshafen)



ICODE Workshop on numerical solutions of HJB equations Paris, 8-10 January 2020

## Outline

- Setting and problem formulation
- What makes model predictive control work?
- Efficient numerical realization for PDEs
- A sensitivity result for general linear evolution equations
- Numerical examples



#### Setting and problem formulation

Consider abstract control systems

$$\dot{y}(t) = f(y(t), u(t)), \quad y(0) = y_0$$

with  $y(t) \in X$ ,  $u(t) \in U$ , X, U suitable spaces



Consider abstract control systems

$$\dot{y}(t) = f(y(t), u(t)), \quad y(0) = y_0$$

with  $y(t) \in X$ ,  $u(t) \in U$ , X, U suitable spaces

Problem: infinite horizon optimal control



Consider abstract control systems

$$\dot{y}(t) = f(y(t), u(t)), \quad y(0) = y_0$$

with  $y(t) \in X$ ,  $u(t) \in U$ , X, U suitable spaces

#### Problem: infinite horizon optimal control Optimality criterion: for a running cost $\ell: X \times U \to \mathbb{R}$ solve

minimize 
$$J_{\infty}(y_0, u) = \int_0^{\infty} \ell(y(t), u(t)) dt$$



Consider abstract control systems

$$\dot{y}(t) = f(y(t), u(t)), \quad y(0) = y_0$$

with  $y(t) \in X$ ,  $u(t) \in U$ , X, U suitable spaces

#### Problem: infinite horizon optimal control Optimality criterion: for a running cost $\ell: X \times U \to \mathbb{R}$ solve

minimize 
$$J_{\infty}(y_0, u) = \int_0^{\infty} \ell(y(t), u(t)) dt$$

subject to state/control constraints  $y(t) \in \mathbb{Y}$ ,  $u(t) \in \mathbb{U}$ 



minimize 
$$J_{\infty}(y_0, u) = \int_0^{\infty} \ell(y(t), u(t)) dt$$

Direct solution of the problem is numerically hard



minimize 
$$J_{\infty}(y_0, u) = \int_0^{\infty} \ell(y(t), u(t)) dt$$

Direct solution of the problem is numerically hard

Alternative method: receding horizon or model predictive control (MPC)



minimize 
$$J_{\infty}(y_0, u) = \int_0^{\infty} \ell(y(t), u(t)) dt$$

Direct solution of the problem is numerically hard

Alternative method: receding horizon or model predictive control (MPC)

Idea: replace the infinite horizon problem by the iterative solution of finite horizon problems

minimize 
$$J_T(y_0, u) = \int_0^T \ell(y(t), u(t)) dt$$

with fixed T>0 and  $y(t)\in\mathbb{Y}\text{, }u(t)\in\mathbb{U}$ 



$$\underset{u(\cdot)}{\operatorname{minimize}} \quad J_{\infty}(y_0,u) = \int_0^{\infty} \ell(y(t),u(t)) dt$$

Direct solution of the problem is numerically hard

Alternative method: receding horizon or model predictive control (MPC)

Idea: replace the infinite horizon problem by the iterative solution of finite horizon problems

$$\underset{u(\cdot)}{\text{minimize}} \quad J_T(y_0, u) = \int_0^T \ell(y(t), u(t)) dt$$

with fixed T>0 and  $y(t)\in\mathbb{Y}\text{, }u(t)\in\mathbb{U}$ 

We obtain a feedback control by a receding horizon technique





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 6/40



black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)





black = predictions (open loop optimization)
red = MPC closed loop





black = predictions (open loop optimization)
red = MPC closed loop





black = predictions (open loop optimization)
red = MPC closed loop





black = predictions (open loop optimization)

red = MPC closed loop  $y_{MPC}(t, y_0)$ 



What is the advantage of MPC over other methods of solving optimal control problems?



What is the advantage of MPC over other methods of solving optimal control problems?

significantly reduced computational complexity

StrobeMediaPlayback.swf



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

StrobeMediaPlayback.swf



What is the advantage of MPC over other methods of solving optimal control problems?

- significantly reduced computational complexity
  - → real time capability
- ability to react to perturbations



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

- ability to react to perturbations
- applicability to real-world industrial applications



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

- ability to react to perturbations
- applicability to real-world industrial applications
- applicability to problems in which data becomes available online



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

- ability to react to perturbations
- applicability to real-world industrial applications
- applicability to problems in which data becomes available online

But: The trajectory delivered by MPC can be far from optimal!



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

- ability to react to perturbations
- applicability to real-world industrial applications
- applicability to problems in which data becomes available online
- But: The trajectory delivered by MPC can be far from optimal!
- → Key questions in this talk:
  - When does MPC yield closed loop trajectories with approximately optimal performance?



What is the advantage of MPC over other methods of solving optimal control problems?

• significantly reduced computational complexity

→ real time capability

- ability to react to perturbations
- applicability to real-world industrial applications
- applicability to problems in which data becomes available online
- But: The trajectory delivered by MPC can be far from optimal!
- → Key questions in this talk:
  - When does MPC yield closed loop trajectories with approximately optimal performance?
  - How can we implement MPC efficiently for PDEs?



#### What makes model predictive control work?

# The turnpike property

The turnpike property demands that there exists a particular trajectory — the turnpike —, such that all optimal trajectories (regardless of initial condition and optimization horizon) stay near this trajectory most of the time [von Neumann '45, Dorfman/Samuelson/Solow '57, McKenzie '83]




# The turnpike property

The turnpike property demands that there exists a particular trajectory — the turnpike —, such that all optimal trajectories (regardless of initial condition and optimization horizon) stay near this trajectory most of the time [von Neumann '45, Dorfman/Samuelson/Solow '57, McKenzie '83]

In the simplest case, this particular trajectory is an equilibrium (but extensions to periodic orbits or more general time varying trajectories exist)



# The turnpike property

The turnpike property demands that there exists a particular trajectory — the turnpike —, such that all optimal trajectories (regardless of initial condition and optimization horizon) stay near this trajectory most of the time [von Neumann '45, Dorfman/Samuelson/Solow '57, McKenzie '83]

In the simplest case, this particular trajectory is an equilibrium (but extensions to periodic orbits or more general time varying trajectories exist)

In this talk we stick to the equilibrium setting



# The turnpike property

The turnpike property demands that there exists a particular trajectory — the turnpike —, such that all optimal trajectories (regardless of initial condition and optimization horizon) stay near this trajectory most of the time [von Neumann '45, Dorfman/Samuelson/Solow '57, McKenzie '83]

In the simplest case, this particular trajectory is an equilibrium (but extensions to periodic orbits or more general time varying trajectories exist)

In this talk we stick to the equilibrium setting

We illustrate it by a simple discrete time example



Minimize the finite horizon objective  $\sum_{k=0}^{N-1}\ell(y(k),u(k))$  with

 $\ell(y,u) = -\ln(Ay^{\alpha} - u), \quad A = 5, \ \alpha = 0.34$  and dynamics y(k+1) = u(k) on  $\mathbb{Y} = \mathbb{U} = [0, 10]$ 



Minimize the finite horizon objective  $\sum_{k=0}^{N-1}\ell(y(k),u(k))$  with

$$\ell(y,u) = -\ln(Ay^{\alpha} - u), \quad A = 5, \ \alpha = 0.34$$
  
nd dynamics  $y(k+1) = u(k)$  on  $\mathbb{Y} = \mathbb{U} = [0, 10]$ 

y =invested capital; u =investment in next time step



а

Minimize the finite horizon objective  $\sum_{k=0}^{N-1}\ell(y(k),u(k))$  with

$$\ell(y,u) = -\ln(Ay^{\alpha} - u), \quad A = 5, \ \alpha = 0.34$$
  
and dynamics  $y(k+1) = u(k)$  on  $\mathbb{Y} = \mathbb{U} = [0, 10]$ 

 $y = \text{invested capital;} \quad u = \text{investment in next time step} \\ Ay^{\alpha} = \text{capital after one time step}$ 



Minimize the finite horizon objective  $\sum_{k=0}^{N-1}\ell(y(k),u(k))$  with

$$\ell(y,u) = -\ln(Ay^{\alpha} - u), \quad A = 5, \ \alpha = 0.34$$
  
and dynamics  $y(k+1) = u(k)$  on  $\mathbb{Y} = \mathbb{U} = [0, 10]$ 

y = invested capital; u = investment in next time step  $Ay^{\alpha} =$  capital after one time step  $Ay^{\alpha} - u =$  consumed capital;  $\ln(\cdot) =$  utility function



Minimize the finite horizon objective  $\sum_{k=0}^{N-1}\ell(y(k),u(k))$  with

 $\ell(y,u)=-\ln(Ay^\alpha-u),\quad A=5,\,\alpha=0.34$  and dynamics  $y(k+1)=u(k)\qquad \text{on}\quad \mathbb{Y}=\mathbb{U}=[0,10]$ 

y = invested capital; u = investment in next time step  $Ay^{\alpha} =$  capital after one time step  $Ay^{\alpha} - u =$  consumed capital;  $\ln(\cdot) =$  utility function

On infinite horizon, it is optimal to stay at the equilibrium  $y^e\approx 2.2344 ~~{\rm with}~ \ell(y^e,u^e)\approx 1.4673$ 







Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 11/40











































The turnpike property makes MPC work...























Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40

















Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40




Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 12/40

In order to formalize how good MPC approximates the infinite horizon problem, we define

- $y_{MPC}(t, y_0) =$  solution generated by MPC starting in  $y_0$
- $u_{MPC}(t) =$ control function generated by MPC

• 
$$J_S^{MPC}(y_0) = \int_0^{\mathbb{Z}} \ell(y_{MPC}(t, y_0), u_{MPC}(t)) dt$$



In order to formalize how good MPC approximates the infinite horizon problem, we define

- $y_{MPC}(t, y_0) =$  solution generated by MPC starting in  $y_0$
- $u_{MPC}(t) = \text{control function generated by MPC}$ •  $J_S^{MPC}(y_0) = \int_0^S \ell(y_{MPC}(t, y_0), u_{MPC}(t)) dt$

#### Furthermore, we define



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 13/40

Theorem: If the turnpike property at an optimal equilibrium  $(y^e, u^e)$  and suitable controllability and regularity conditions hold, then there exist  $\varepsilon_1(T), \varepsilon_2(S) \to 0$  as  $T \to \infty$  and  $S \to \infty$ , such that the following properties hold



Theorem: If the turnpike property at an optimal equilibrium  $(y^e, u^e)$  and suitable controllability and regularity conditions hold, then there exist  $\varepsilon_1(T), \varepsilon_2(S) \to 0$  as  $T \to \infty$  and  $S \to \infty$ , such that the following properties hold

(1) Approximate average optimality:

$$\limsup_{S \to \infty} \frac{1}{S} J_S^{MPC}(y_0) \le \ell(y^e, u^e) + \varepsilon_1(T)$$



Theorem: If the turnpike property at an optimal equilibrium  $(y^e, u^e)$  and suitable controllability and regularity conditions hold, then there exist  $\varepsilon_1(T), \varepsilon_2(S) \to 0$  as  $T \to \infty$  and  $S \to \infty$ , such that the following properties hold

(1) Approximate average optimality:

$$\limsup_{S \to \infty} \frac{1}{S} J_S^{MPC}(y_0) \le \ell(y^e, u^e) + \varepsilon_1(T)$$

(2) Practical asymptotic stability: there is  $\beta \in \mathcal{KL}$ :

 $\|y_{MPC}(t,y_0) - y^e\| \le \beta(\|y_0 - y^e\|, t) + \varepsilon_1(T) \text{ for all } k \in \mathbb{N}$ 



Theorem: If the turnpike property at an optimal equilibrium  $(y^e, u^e)$  and suitable controllability and regularity conditions hold, then there exist  $\varepsilon_1(T), \varepsilon_2(S) \to 0$  as  $T \to \infty$  and  $S \to \infty$ , such that the following properties hold

(1) Approximate average optimality:

$$\limsup_{S \to \infty} \frac{1}{S} J_S^{MPC}(y_0) \le \ell(y^e, u^e) + \varepsilon_1(T)$$

(2) Practical asymptotic stability: there is  $\beta \in \mathcal{KL}$ :

 $\|y_{MPC}(t,y_0) - y^e\| \le \beta(\|y_0 - y^e\|, t) + \varepsilon_1(T) \text{ for all } k \in \mathbb{N}$ 

(3) Approximate transient optimality: for all  $S \in \mathbb{N}$ :

$$J_S^{MPC}(y_0) \le J_S(y_0, \mathbf{u}) + S\varepsilon_1(T) + \varepsilon_2(S)$$

for all admissible  ${\bf u}$  with  $\|y(S,y_0,{\bf u})-y^e\|\leq \beta(\|y_0-y^e\|,S)+\varepsilon_1(T)$ 























Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 15/40







(2):  $y_{MPC}(t)$  converges to the  $\varepsilon_1(T)$ -ball around  $y^e$ 



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 15/40



AYREUTH

(3): cost of all other trajectories reaching the ball at time K is higher than that of  $y_{MPC}(t)$  up to the error  $S\varepsilon_1(T) + \varepsilon_2(S)$ 

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 15/40

## When does the turnpike property hold?

- [Carlson et al. '91, Gr. '13, Gr./Stieler/Pirkelmann '18]: strict dissipativity implies turnpike property for many system classes
- [Gr./Müller '16]: in discrete time the turnpike property is equivalent to strict dissipativity for controllable systems
- [Gr./Guglielmi '18f]: the turnpike property is equivalent to detectability-like characterizations for stabilizable finite dimensional linear quadratic problems
- [Höger/Gr. '19]: Input-output-to-state stability (IOSS) implies strict dissipativity and hence the turnpike property
- [Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f, Gr./Schaller/ Schiela '19f]: turnpike property is implied by stabilizability and detectability for PDE governed linear quadratic problems

[Porretta/Zuazua '13, Gugat/Trélat/Zuazua '16, Zuazua '18,

Gugat/Hante '18]: turnpike property for various hyperbolic PDEs



## When does the turnpike property hold?

- [Carlson et al. '91, Gr. '13, Gr./Stieler/Pirkelmann '18]: strict dissipativity implies turnpike property for many system classes
- [Gr./Müller '16]: in discrete time the turnpike property is equivalent to strict dissipativity for controllable systems
- [Gr./Guglielmi '18f]: the turnpike property is equivalent to detectability-like characterizations for stabilizable finite dimensional linear quadratic problems
- [Höger/Gr. '19]: Input-output-to-state stability (IOSS) implies strict dissipativity and hence the turnpike property

[Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f, Gr./Schaller/ Schiela '19f]: turnpike property is implied by stabilizability and detectability for PDE governed linear quadratic problems

[Porretta/Zuazua '13, Gugat/Trélat/Zuazua '16, Zuazua '18,

Gugat/Hante '18]: turnpike property for various hyperbolic PDEs



## Efficient numerical realization for PDEs

# Idea of efficient numerical approach















What about the global numerical error?

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 18/40

## Sensitivity w.r.t. numerical errors

At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 


At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 

At the second glance, there is a problem



At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 

At the second glance, there is a problem: solving the optimal control problem involves the adjoint equation which is solved backwards



At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 

At the second glance, there is a problem: solving the optimal control problem involves the adjoint equation which is solved backwards

$$\dot{y}(t) = H_{\lambda}(y(t), u(t), \lambda(t)), \qquad y(0) = y_0$$
  
$$\dot{\lambda}(t) = -H_y(y(t), u(t), \lambda(t)), \qquad \lambda(T) = 0$$



At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 

At the second glance, there is a problem: solving the optimal control problem involves the adjoint equation which is solved backwards

$$\dot{y}(t) = H_{\lambda}(y(t), u(t), \lambda(t)) + \varepsilon_{1}(t), \quad y(0) = y_{0}$$
  
$$\dot{\lambda}(t) = -H_{y}(y(t), u(t), \lambda(t)) + \varepsilon_{2}(t), \quad \lambda(T) = 0$$

 $\rightsquigarrow$  large  $\varepsilon_i(t)$  for  $t \approx T$  can propagate backwards to  $t \approx 0$ 



At a first glance, one might conjecture that if the local discretization error  $\varepsilon(t)$  is large only for  $t \approx T$ , then this should not affect the solution at times  $t \approx 0$ 

At the second glance, there is a problem: solving the optimal control problem involves the adjoint equation which is solved backwards

$$\dot{y}(t) = H_{\lambda}(y(t), u(t), \lambda(t)) + \varepsilon_{1}(t), \quad y(0) = y_{0}$$
  
$$\dot{\lambda}(t) = -H_{y}(y(t), u(t), \lambda(t)) + \varepsilon_{2}(t), \quad \lambda(T) = 0$$

 $\rightsquigarrow$  large  $\varepsilon_i(t)$  for  $t \approx T$  can propagate backwards to  $t \approx 0$ Is there a structural property that can save this idea?



A sensitivity result for general linear evolution equations

#### We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 21/40

We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

- $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$
- X Hilbert space



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term

and the optimization objective

$$\min_{u} \frac{1}{2} \int_{0}^{T} \|C(y(t) - y_d)\|_{Y}^{2} + \|R(u(t) - u_d)\|_{U}^{2} dt$$



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term

and the optimization objective

$$\min_{u} \frac{1}{2} \int_{0}^{T} \|C(y(t) - y_{d})\|_{Y}^{2} + \|R(u(t) - u_{d})\|_{U}^{2} dt$$

• U, Y Hilbert spaces



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term

and the optimization objective

$$\min_{u} \frac{1}{2} \int_{0}^{T} \|C(y(t) - y_{d})\|_{Y}^{2} + \|R(u(t) - u_{d})\|_{U}^{2} dt$$

• U, Y Hilbert spaces;  $R \in L(U, U)$  elliptic



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term

and the optimization objective

$$\min_{u} \frac{1}{2} \int_{0}^{T} \|C(y(t) - y_{d})\|_{Y}^{2} + \|R(u(t) - u_{d})\|_{U}^{2} dt$$

- U, Y Hilbert spaces;  $R \in L(U, U)$  elliptic
- B, C admissible, possibly unbounded



We consider general linear evolution equations

$$\frac{d}{dt}y = Ay + Bu + f, \quad y(0) = y_0$$

•  $A: D(A) \subset X \to X$  generates a  $C_0$ -semigroup  $\mathcal{T}(t)$ 

• X Hilbert space;  $f \in L_1(0,T;X)$  source term

and the optimization objective

$$\min_{u} \frac{1}{2} \int_{0}^{T} \|C(y(t) - y_{d})\|_{Y}^{2} + \|R(u(t) - u_{d})\|_{U}^{2} dt$$

- U, Y Hilbert spaces;  $R \in L(U, U)$  elliptic
- B, C admissible, possibly unbounded

Solution concept: mild solution

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 21/40

Optimality condition: Pontryagin's Maximum Principle yields

$$\underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - A^* \\ 0 & E_T \\ \frac{d}{dt} - A & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=: M} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} C^*Cy_d \\ 0 \\ Bu_d + f \\ y_0 \end{pmatrix},$$

where  $E_0 y := y(0)$  and  $E_T \lambda := \lambda(T)$ 



Optimality condition: Pontryagin's Maximum Principle yields

$$\underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - A^* \\ 0 & E_T \\ \frac{d}{dt} - A & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=: M} \begin{pmatrix} \tilde{y} \\ \tilde{\lambda} \end{pmatrix} = \begin{pmatrix} C^*Cy_d \\ 0 \\ Bu_d + f \\ y_0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ 0 \\ \varepsilon_2 \\ 0 \end{pmatrix},$$

where  $E_0 y := y(0)$  and  $E_T \lambda := \lambda(T)$ 



Optimality condition: Pontryagin's Maximum Principle yields

$$\underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - A^* \\ 0 & E_T \\ \frac{d}{dt} - A & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=: M} \begin{pmatrix} \tilde{y} \\ \tilde{\lambda} \end{pmatrix} = \begin{pmatrix} C^*Cy_d \\ 0 \\ Bu_d + f \\ y_0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ 0 \\ \varepsilon_2 \\ 0 \end{pmatrix},$$

where  $E_0 y := y(0)$  and  $E_T \lambda := \lambda(T)$ 

Define  $\delta y = \tilde{y} - y$ ,  $\delta \lambda = \tilde{\lambda} - \lambda$ 



Optimality condition: Pontryagin's Maximum Principle yields

$$\underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - A^* \\ 0 & E_T \\ \frac{d}{dt} - A & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=: M} \begin{pmatrix} \tilde{y} \\ \tilde{\lambda} \end{pmatrix} = \begin{pmatrix} C^*Cy_d \\ 0 \\ Bu_d + f \\ y_0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ 0 \\ \varepsilon_2 \\ 0 \end{pmatrix},$$

where  $E_0 y := y(0)$  and  $E_T \lambda := \lambda(T)$ 

Define  $\delta y = \tilde{y} - y$ ,  $\delta \lambda = \tilde{\lambda} - \lambda$ 

Idea: Use  $\|\delta y\| + \|\delta \lambda\| \le \|M^{-1}\|\|(\varepsilon_1, 0, \varepsilon_2, 0)\|$ 

plus exponential weighting



Theorem: Define  $\rho := \|e^{-\mu t} \varepsilon_1(t)\|_{L_p(X)} + \|e^{-\mu t} \varepsilon_2(t)\|_{L_p(X)}$ for p = 1 or p = 2 and assume the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \end{split}$$

are bounded independently of  $\boldsymbol{T}$ 



Theorem: Define  $\rho := \|e^{-\mu t} \varepsilon_1(t)\|_{L_p(X)} + \|e^{-\mu t} \varepsilon_2(t)\|_{L_p(X)}$ for p = 1 or p = 2 and assume the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \end{split}$$

are bounded independently of T. Then there are  $\mu,c>0$  with

$$\begin{aligned} \|e^{-\mu t} \delta y\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta \lambda\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta u\|_{L_{2}(U)} &\leq c\rho \\ \|e^{-\mu t} \delta y\|_{C(X)} &+ \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho \end{aligned}$$



Theorem: Define  $\rho := \|e^{-\mu t} \varepsilon_1(t)\|_{L_p(X)} + \|e^{-\mu t} \varepsilon_2(t)\|_{L_p(X)}$ for p = 1 or p = 2 and assume the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \end{split}$$

are bounded independently of T. Then there are  $\mu,c>0$  with

$$\begin{aligned} \|e^{-\mu t} \delta y\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta \lambda\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta u\|_{L_{2}(U)} &\leq c\rho \\ \|e^{-\mu t} \delta y\|_{C(X)} &+ \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho \end{aligned}$$

If B is bounded then in addition  $\|e^{-\mu t} \delta u\|_{L_\infty(U)} \leq c\rho$  holds



Theorem: Define  $\rho := \|e^{-\mu t} \varepsilon_1(t)\|_{L_p(X)} + \|e^{-\mu t} \varepsilon_2(t)\|_{L_p(X)}$ for p = 1 or p = 2 and assume the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \end{split}$$

are bounded independently of T. Then there are  $\mu,c>0$  with

$$\begin{aligned} \|e^{-\mu t} \delta y\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta \lambda\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta u\|_{L_{2}(U)} &\leq c\rho \\ \|e^{-\mu t} \delta y\|_{C(X)} &+ \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho \end{aligned}$$

If B is bounded then in addition  $||e^{-\mu t} \delta u||_{L_{\infty}(U)} \leq c\rho$  holds For p = 1 and  $|| \cdot ||_{C(X)}$ -norms this implies

$$||y(t) - \tilde{y}(t)|| \le \int_0^T c e^{\mu(t-s)} (||\varepsilon_1(s)|| + ||\varepsilon_2(s)||) ds$$



Theorem: Define  $\rho := \|e^{-\mu t} \varepsilon_1(t)\|_{L_p(X)} + \|e^{-\mu t} \varepsilon_2(t)\|_{L_p(X)}$ for p = 1 or p = 2 and assume the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \end{split}$$

are bounded independently of T. Then there are  $\mu,c>0$  with

$$\begin{aligned} \|e^{-\mu t} \delta y\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta \lambda\|_{L_{2}(X)} &+ \|e^{-\mu t} \delta u\|_{L_{2}(U)} &\leq c\rho \\ \|e^{-\mu t} \delta y\|_{C(X)} &+ \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho \end{aligned}$$

If B is bounded then in addition  $\|e^{-\mu t} \delta u\|_{L_{\infty}(U)} \leq c\rho$  holds For p = 1 and  $\|\cdot\|_{C(X)}$ -norms this implies

$$||y(t) - \tilde{y}(t)|| \le \int_0^T c e^{\mu(t-s)} (||\varepsilon_1(s)|| + ||\varepsilon_2(s)||) ds$$

 $\rightsquigarrow$  Large errors for  $s \approx T$  are exponentially damped at  $t \approx 0$ 

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 23/40

# Boundedness of $||M^{-1}||$ How do we get a *T*-independent bound for the norms $||M^{-1}||_{(L_1(X) \times X)^2 \to C(X)^2} = ||M^{-1}||_{(L_2(X) \times X)^2 \to C(X)^2} = ||M^{-1}||_{(L_2(X) \times X)^2 \to L_2(X)^2}$ ?



How do we get a T-independent bound for the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \\ \end{split}$$

Definition: (i) We say that (A, B) is exponentially stabilizable, if there is  $K \in L(X, U)$  such that the semigroup generated by A + BK is exponentially stable

(ii) We say that (A,C) is exponentially detectable if  $(A^{\ast},C^{\ast})$  is exponentially stabilizable



How do we get a T-independent bound for the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \\ \end{split}$$

Definition: (i) We say that (A, B) is exponentially stabilizable, if there is  $K \in L(X, U)$  such that the semigroup generated by A + BK is exponentially stable

(ii) We say that (A,C) is exponentially detectable if  $(A^{\ast},C^{\ast})$  is exponentially stabilizable

Theorem: If the control system is exponentially stabilizable and detectable, then the above norms are bounded independently of T



How do we get a T-independent bound for the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \\ \end{split}$$

Definition: (i) We say that (A, B) is exponentially stabilizable, if there is  $K \in L(X, U)$  such that the semigroup generated by A + BK is exponentially stable

(ii) We say that (A,C) is exponentially detectable if  $(A^{\ast},C^{\ast})$  is exponentially stabilizable

Theorem: If the control system is exponentially stabilizable and detectable, then the above norms are bounded independently of  ${\cal T}$ 

This is the hard part of the analysis



How do we get a T-independent bound for the norms

$$\begin{split} \|M^{-1}\|_{(L_1(X)\times X)^2\to C(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to C(X)^2} \\ \|M^{-1}\|_{(L_1(X)\times X)^2\to L_2(X)^2} & \|M^{-1}\|_{(L_2(X)\times X)^2\to L_2(X)^2} \\ \end{split}$$

Definition: (i) We say that (A, B) is exponentially stabilizable, if there is  $K \in L(X, U)$  such that the semigroup generated by A + BK is exponentially stable

(ii) We say that (A,C) is exponentially detectable if  $(A^{\ast},C^{\ast})$  is exponentially stabilizable

Theorem: If the control system is exponentially stabilizable and detectable, then the above norms are bounded independently of T

This is the hard part of the analysis

For details: manuel.schaller@uni-bayreuth.de

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 24/40

• Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup



- Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup
- Particularly, our results apply to hyperbolic PDEs and boundary control and observations



- Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup
- Particularly, our results apply to hyperbolic PDEs and boundary control and observations
- Recall: for PDE governed linear quadratic problems, the turnpike property is implied by stabilizability and detectability [Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f]



- Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup
- Particularly, our results apply to hyperbolic PDEs and boundary control and observations
- Recall: for PDE governed linear quadratic problems, the turnpike property is implied by stabilizability and detectability [Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f]
  - → The same mechanism that generates the turnpike behaviour damps out errors in backward time



- Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup
- Particularly, our results apply to hyperbolic PDEs and boundary control and observations
- Recall: for PDE governed linear quadratic problems, the turnpike property is implied by stabilizability and detectability [Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f]
  - → The same mechanism that generates the turnpike behaviour damps out errors in backward time
- Extension to certain PDEs with nonlinearities (semilinear, quasilinear) possible



- Under the same condition we obtain a turnpike result that generalizes many of the mentioned results in the literature, as we require neither boundedness of *B* and *C* nor that *A* generates an analytic semigroup
- Particularly, our results apply to hyperbolic PDEs and boundary control and observations
- Recall: for PDE governed linear quadratic problems, the turnpike property is implied by stabilizability and detectability [Trélat/Zhang/Zuazua '18, Breiten/Pfeiffer '18f]
  - → The same mechanism that generates the turnpike behaviour damps out errors in backward time
- Extension to certain PDEs with nonlinearities (semilinear, quasilinear) possible work in progress



# Numerical examples
We expect to see the following effect:

Fine grid for small t



We expect to see the following effect:





We expect to see the following effect:





We expect to see the following effect:



However, we do not want to select the grids manually



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 27/40

We expect to see the following effect:



However, we do not want to select the grids manually

→ goal-oriented estimation [Meidner '08, Meidner/Vexler '07ff]



#### Goal oriented error estimation

In goal-oriented error estimation the error of a particular quantity of interest is estimated



#### Goal oriented error estimation

In goal-oriented error estimation the error of a particular quantity of interest is estimated

We use

$$J_T(y, u) := \int_0^T \ell(y(t), u(t)) \, dt \quad \text{and} \quad J_\tau(y, u) := \int_0^\tau \ell(y(t), u(t)) \, dt$$

with time interval  $\tau = 0.5$ 



#### Goal oriented error estimation

In goal-oriented error estimation the error of a particular quantity of interest is estimated

We use

 $J_T(y,u) := \int_0^T \ell(y(t), u(t)) \, dt \quad \text{and} \quad J_\tau(y,u) := \int_0^\tau \ell(y(t), u(t)) \, dt$ 

with time interval  $\tau = 0.5$ 

For the discontinuous Galerkin discretization in time we can prove:

Theorem: Let  $(A,B),~(A^*,C^*)$  be exponentially stabilizable. Then the error indicators  $\eta^\tau$  for  $J_\tau$  satisfy

 $\|\eta^{\tau}(t)\| \sim c(\tau)e^{-\mu t}$ 

with  $c(\tau), \mu > 0$  independent of T

# Test problem

$$\min_{(y,u)} \frac{1}{2} \|y - y_d\|_{L_2([0,30] \times [0,1]^2)}^2 + \frac{\alpha}{2} \|u\|_{L_2([0,30] \times [0,1]^2)}^2$$
$$\frac{d}{dt}y = -d\Delta y + \mu y + u, \quad y(0) = 0, \quad y_d =$$



Open loop optimal solution



t = 0.0



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 30/40

Open loop optimal solution



t = 0.3



Open loop optimal solution



t = 0.6



Open loop optimal solution



 $t = 0.6, \ldots, 2.7$ 



Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 30/40

Open loop optimal solution



t = 3.0







Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40



![](_page_162_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_163_Figure_2.jpeg)

![](_page_163_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_164_Figure_2.jpeg)

![](_page_164_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_165_Figure_2.jpeg)

![](_page_165_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_166_Figure_2.jpeg)

![](_page_166_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_167_Figure_2.jpeg)

![](_page_167_Picture_3.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 31/40

![](_page_168_Figure_2.jpeg)

![](_page_169_Figure_2.jpeg)

![](_page_170_Figure_2.jpeg)

![](_page_171_Figure_2.jpeg)

![](_page_172_Figure_2.jpeg)

![](_page_173_Figure_2.jpeg)

![](_page_174_Figure_2.jpeg)

![](_page_175_Figure_2.jpeg)

![](_page_176_Figure_2.jpeg)

![](_page_177_Figure_2.jpeg)

#### Comparison to standard error estimator

![](_page_178_Figure_2.jpeg)

#### Stability of MPC closed-loop solutions

8 time points

11 time points

![](_page_179_Figure_4.jpeg)

![](_page_179_Picture_5.jpeg)

Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 34/40
#### Adaptive grid in time

#### Cost of MPC closed-loop solutions





# Adaptive grid in space

Goal oriented (bottom) vs. standard error estimator (top)





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 36/40

## Adaptive grid in space





#### Adaptive grid in space and time





Lars Grüne, Sensitivity and turnpike results for optimal control of PDEs and MPC, p. 38/40

 Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems



 Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems ("Model reduction in time")



- Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems ("Model reduction in time")
- The existence of the turnpike property is the key structural property to make this approach work



- Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems ("Model reduction in time")
- The existence of the turnpike property is the key structural property to make this approach work
- Exponential controllability and detectability imply this property for linear quadratic PDE problems



- Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems ("Model reduction in time")
- The existence of the turnpike property is the key structural property to make this approach work
- Exponential controllability and detectability imply this property for linear quadratic PDE problems
- The same mechanism that leads to the turnpike property also causes an exponential damping of numerical errors in backward time



- Model Predictive Control can be seen as a method for splitting up an infinite horizon optimal control problem into the iterative solution of finite horizon problems ("Model reduction in time")
- The existence of the turnpike property is the key structural property to make this approach work
- Exponential controllability and detectability imply this property for linear quadratic PDE problems
- The same mechanism that leads to the turnpike property also causes an exponential damping of numerical errors in backward time
- This can be exploited by adaptive discretization strategies via goal oriented error estimators



#### References

L. Grüne, *Approximation properties of receding horizon optimal control*, DMV Jahresbericht, 118, 3–37, 2016

L. Grüne, *Economic receding horizon control without terminal constraints*, Automatica, 49, 725–734, 2013

L. Grüne, M. Stieler, *Asymptotic stability and transient optimality of economic MPC without terminal conditions*, Journal of Process Control, 24 (Special Issue on Economic MPC), 1187–1196, 2014

L. Grüne, M. Schaller, A. Schiela, *Sensitivity analysis of optimal control for a class of parabolic PDEs motivated by Model Predictive Control*, SIAM J. Control Optim., 2019

L. Grüne, M. Schaller, A. Schiela, *Exponential sensitivity and turnpike* analysis for linear quadratic optimal control of general evolution equations, Journal of Differential Equations, online version appeared 2019

L. Grüne, M. Schaller, A. Schiela, *Space-time adaptivity* for Model Predictive Control of parabolic PDEs, in preparation

