Recent results on numerical dynamic programming for hybrid systems

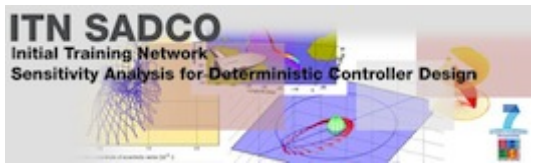
Roberto Ferretti

Department of Mathematics and Physics, Roma Tre University

ferretti@mat.uniroma3.it

Seminario MDN, 28.03.17

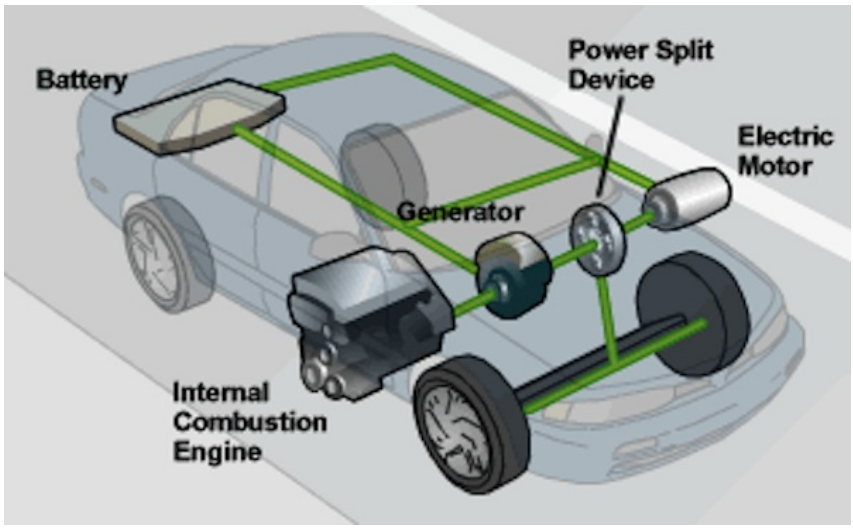
joint works with
A. Festa, A. Sassi and H. Zidani



Outline

1. Introduction
   - A basic example
   - A general framework for hybrid systems
   - The optimal control problem
2. Approximation via monotone schemes
   - Monotone schemes
   - Value iteration
   - Construction of the optimal feedback
   - Acceleration techniques for the solution of the QVI
   - Error estimates for the approximate value function
3. Numerical examples
   - Stabilization of and unstable system
   - Multigear vehicle
   - AC/DC inverter
   - Tacking strategy in a regatta
4. Conclusions

A basic example (1)



A **hybrid vehicle** may use both an **internal combustion engine** and an **electric engine**:

- **ICE**: motion + battery recharging
- **EE**: motion
- **Generator**: battery recharging + braking

A basic example (2)

- It is possible to take the **state of charge of the battery** as the state variable, and the **speed** (or throttle) as the control variable on each of the two engines
- When the battery is fully discharged the system **must turn to the ICE**, while at any intermediate level **the switching may or may not be performed** on the basis of the optimal strategy
- There exists a **switching cost** for changing the active engine

A basic example (2)

- Two **different dynamics**

$$\begin{cases} \dot{X}(t) = f_t(X(t), u(t)) = au(t) > 0 & \text{(thermic engine)} \\ \dot{X}(t) = f_e(X(t), u(t)) = -(b - cX(t))u(t) < 0 & \text{(electric engine)} \end{cases}$$

- A **running cost** related to each of the dynamics, e.g., the fuel consumption

$$\begin{cases} c_t(X(t), u(t)) = u(t)^2 & \text{(thermic engine)} \\ c_e(X(t), u(t)) = 0 & \text{(electric engine)} \end{cases}$$

- An **optimization criterion**, e.g., fuel consumption with a given minimal speed

A basic example (3)

- A set on which **switching is mandatory** + a set on which **switching is optional**

$$\begin{cases} X(t) = 0 & (\text{electric} \to \text{thermic, mandatory}) \\ X(t) > 0 & (\text{any switching is optional}) \end{cases}$$

- A **cost related to switching**

$$\text{switching cost} = \begin{cases} c_{te} = 0 & (\text{thermic} \to \text{electric}) \\ c_{et} > 0 & (\text{electric} \to \text{thermic}) \end{cases}$$

- An infinite number of commutations in finite time **should be avoided**

State equations of a hybrid system (1)

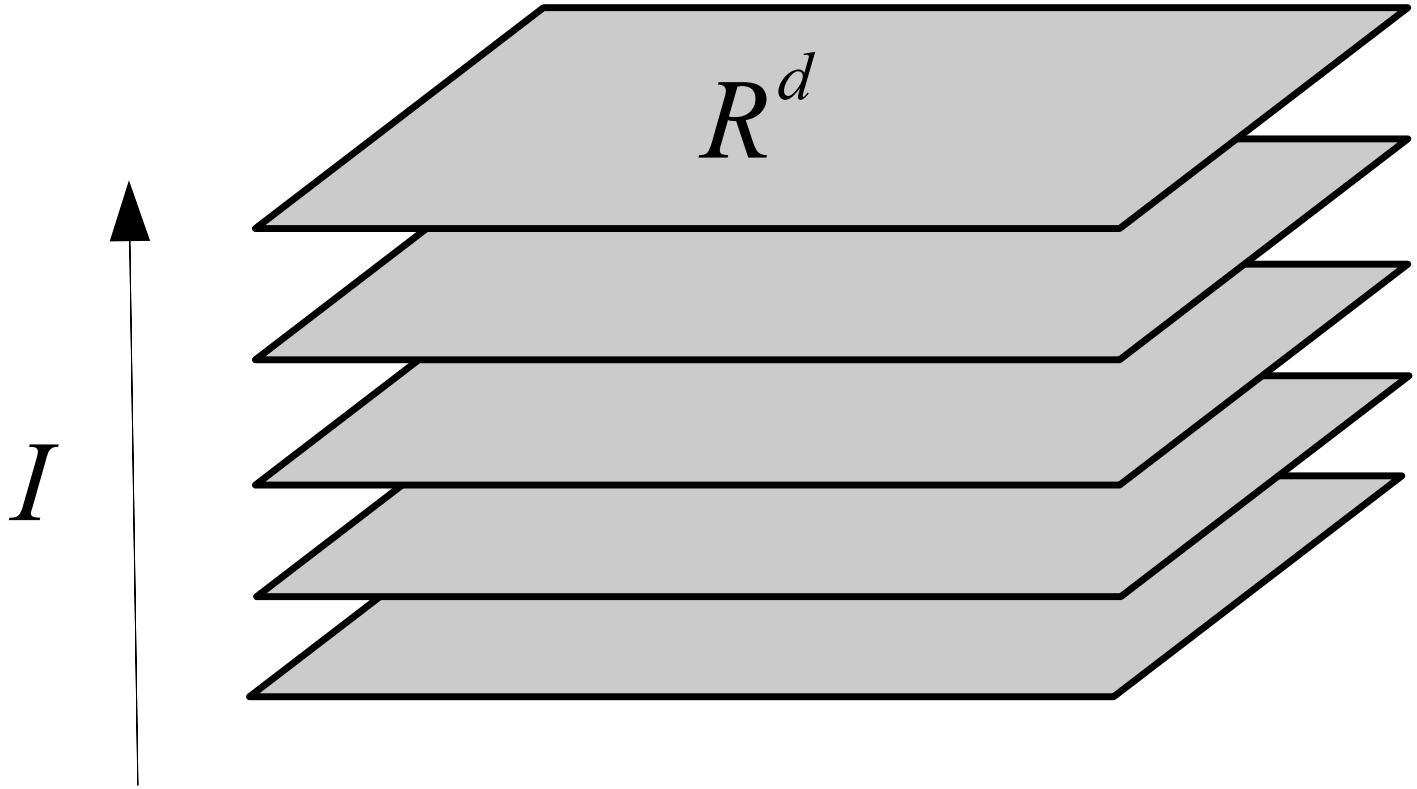
- **State** of the system: $(X(t), Q(t)) \in \Omega \times \mathbb{I}$, with $\Omega \subseteq \mathbb{R}^d$, $\mathbb{I} = \{1, \ldots, Q_m\}$. The discrete variable $Q(t)$ (with initial value $q = Q(0)$) tells **which dynamics is active at time** t
- A measurable **control** $u(t)$ mapping $(0, +\infty)$ into a compact set U

State equation

Evolution for given initial values of X and Q:

$$\begin{cases} \dot{X}(t) = f(X(t), Q(t), u(t)), \\ X(0) = x, \\ Q(0) = q. \end{cases}$$

State equations of a hybrid system (2)



The state space is endowed with the **product topology (metric in x, discrete in q)**

State equations of a hybrid system (3)

Commutations (i.e., changes in $Q(t)$) may occur in **two prescribed sets** A **and** C:

- On hitting the set A at a point (x, q), **the state must jump** to a destination set D. The new state is given by a transition map

$$(x', q') = g(x, q, w)$$

  where $w \in \mathcal{V}$ plays the role of a **control acting on the jump**
- Inside the set C, **the state may jump** from a state (x, q) to a different state $(x', q') \in D$

Control strategy

a **control strategy** for this hybrid system is a quadruple:

$$\theta = \left(u, w, \{\xi_k\}, \left\{(X, Q)\left(\xi_k^+\right)\right\}\right)$$

- u and w are the controls for respectively the **continuous system** f and the **transition map** g
- ξ_k **is a sequence of switching times** for the optional jumps and $(X, Q)(\xi_k^+)$ **are the corresponding states after each jump**

Optimal control problem

Cost functional

In the **discounted infinite horizon case**, the cost functional is defined by

$$J(x,q,\theta) = \int_0^{+\infty} \ell(X(t),Q(t),u(t))e^{-\lambda t}\,dt + \tag{1}$$

$$+\sum_{k=0}^{\infty} C_a(X(\tau_k^-),Q(\tau_k^-),w(\tau_k))e^{-\lambda\tau_k} + \tag{2}$$

$$+\sum_{i=0}^{\infty} C_c(X(\xi_i^-),Q(\xi_i^-),X(\xi_i^+),Q(\xi_i^+))e^{-\lambda\xi_i} \tag{3}$$

- (1) is the cost related to **continuous control**
- (2) is the cost related to **mandatory (autonomous) commutations**
- (3) is the cost related to **optional (controlled) commutations**

Bellman Equation (1)

Once defined the value function

$$V(x,q) = \inf_{\theta} J(x,q,\theta)$$

it can be proved that (in a suitably adapted viscosity sense) V **satisfies the Quasi-Variational Inequality**

QVI

$$\begin{cases} V(x,q) - \mathcal{M}V(x,q) = 0 & x \in A_q, \\ \max(V(x,q) - \mathcal{N}V(x,q), \lambda V(x,q) + H(x,q,D_x V(x,q))) = 0 & x \in C_q, \\ \lambda V(x,q) + H(x, D_x V(x,q)) = 0 & \text{else} \end{cases}$$

Bellman Equation (2)

In QVI:

- $H(x,q,p) = \sup_{u \in U} \{-\ell(x,q,u) - f(x,q,u) \cdot p\}$
  is the term related to continuous control (it **selects the best continuous control strategy**)
- $\mathcal{M}\phi(x,q) = \inf_{w \in \mathcal{V}} \{\phi(g(x,q,w)) + C_a(x,q,w)\}$
  is the term related to autonomuos jumps controlled by w (it **selects the best jump strategy on** A)
- $\mathcal{N}\phi(x,q) = \inf_{(x',q') \in D \times \mathbb{I}} \{\phi(x',q') + C_c(x,q,x',q')\}$
  is the term related to controlled jumps (it **selects the best jump strategy on** C, and the complementarity condition further chooses the **optimal strategy between jumps and continuous control**)

Theoretical results on the QVI

A certain number of **technical assumptions** are to be satisfied to develop the theory. **We skip here a complete overview**, but the numerical examples show that **the framework is robust enough to drop some of them** (at least in some special case).

Main Theoretical results:

- **The value function V is Hölder continuous** (Lipschitz continuous under a slightly less general formulation of the problem)
- A **strong comparison principle** holds

Monotone schemes (1)

A natural approximation approach is to **discretize (in a consistent form) each of the operators appearing in the QVI**:

$$\begin{aligned}
\frac{V^h(x,q) - \Sigma^h(x,q,V^h)}{h} &\leftrightarrow \lambda V(x,q) + H\left(x,q,D_x V(x,q)\right) \\
V^h(x,q) - N^h V^h(x,q) &\leftrightarrow V(x,q) - \mathcal{N} V(x,q) \\
V^h(x,q) - M^h V^h(x,q) &\leftrightarrow V(x,q) - \mathcal{M} V(x,q)
\end{aligned}$$

- The approximation V^h of V is indexed by a **discretization step** h, i.e., space and time steps are related (in this situation, typically h **is to be intended as a time step**)

Monotone schemes (2)

- Σ^h, M^h **and** N^h **are monotone**, i.e., for any $\Phi \geq \Psi$:

$$\begin{aligned} M^h\Phi(x,q) &\geq M^h\Psi(x,q) \\ N^h\Phi(x,q) &\geq N^h\Psi(x,q) \\ \Sigma^h(x,q,\Phi) &\geq \Sigma^h(x,q,\Psi) \end{aligned}$$

- M^h **and** N^h **are invariant w.r.t. the addition of constants**:

$$\begin{aligned} M^h(\Phi + c)(x,q) &= M^h\Phi(x,q) + c \\ N^h(\Phi + c)(x,q) &= N^h\Phi(x,q) + c \end{aligned}$$

- **The operator** Σ^h **is a contraction** for $h < h_0$, that is,

$$\left\| \Sigma^h(\Phi) - \Sigma^h(\Psi) \right\|_\infty \leq (1 - Ch) \left\| \Phi - \Psi \right\|_\infty$$

Monotone schemes (3)

The resulting scheme reads $S^h\left(x, q, V^h(x,q), V^h\right) = 0$, with

$$S^h(x,q,V^h(x,q),V^h) = \begin{cases} V^h(x,q) - M^h V^h(x,q) & \text{if } x \in A_q \\ \max\left\{ V^h(x,q) - N^h V^h(x,q), \dfrac{V^h(x,q) - \Sigma^h(x,q,V^h)}{h} \right\} & \text{if } x \in C_q \\ \dfrac{V^h(x,q) - \Sigma^h(x,q,V^h)}{h} & \text{else.} \end{cases}$$

- Due to the consistency of each building block, S^h **is consistent**
- Due to the monotonicity of each building block, **the scheme is monotone and L^∞ stable**
- Since a **strong comparison principle** holds for the continuous problem, **convergence follows via the Barles–Souganidis theorem** (R. F. – H. Zidani, 2014)

Value iteration

"Classical" approach: value iteration with **monotone** schemes (e.g., upwind, Lax–Friedrichs, Semi-Lagrangian + monotone approximation of the switching operators), for which the **Barles–Souganidis theorem** applies.
Starting from a **time-marching** formulation, the scheme can be put it in the following

Fixed-point form

$$V^h(x,q) = T^h(x,q,V^h) = \begin{cases} M^h V^h(x,q) & \text{if } x \in A_q \\ \min\left\{N^h V^h(x,q), S^h(x,q,V^h)\right\} & \text{if } x \in C_q \\ S^h(x,q,V^h) & \text{else.} \end{cases}$$

This formulation lends itself to computing the solution **via the (convergent) iteration**

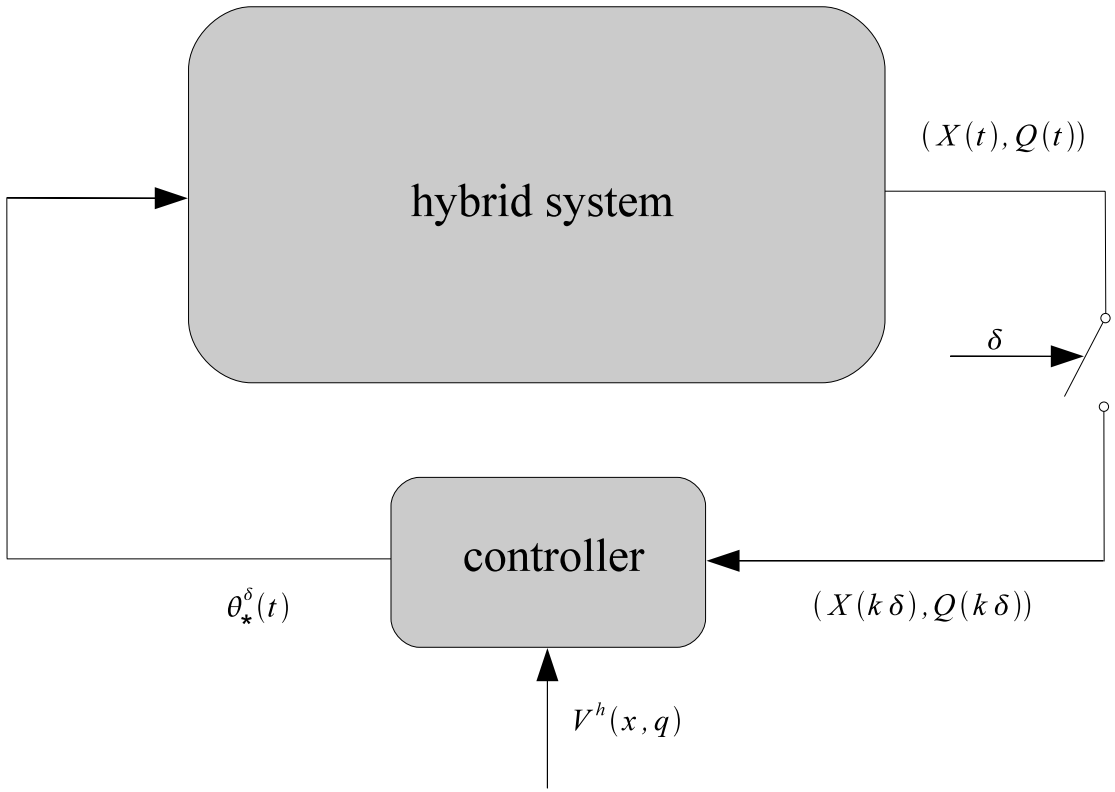
$$V^h_{k+1} = T^h(V^h_k)$$

Basic ideas about the reconstruction of the optimal feedback (1)

The **synthesis of a quasi-optimal control** may be carried out independently of the specific numerical technique used to approximate the HJB equation. **Few convergence results for the feedback law**.

- The state of the system **is sampled with a time step** δ, to be suitably determined
- The controller uses the **information from the value function** to generate either a **piecewise constant control** or a **single switch** in the interval $(k\delta, (k+1)\delta)$

Basic ideas about the reconstruction of the optimal feedback (2)



θ^{δ}_{*} is such that, on any δ-interval, either $u(t)$ is **constant**, or a **single switch** of the state is performed.

Construction of the quasi-optimal control strategy (1)

General idea: set $(x, q) = (X(k\delta), Q(k\delta))$ and construct the optimal strategy as

$$\begin{cases} (x,q) \in A \longrightarrow \text{optimal autonomuos jump} \\ (x,q) \in C \longrightarrow \text{the best between} \begin{cases} \text{optimal continuous control} \\ \text{optimal controlled jump} \end{cases} \\ (x,q) \in (\mathbb{R}^d \times \mathbb{I}) \setminus (A \cup C) \longrightarrow \text{optimal continuous control} \end{cases}$$

using an **approximate discrete dynamic programming** argument on a time step δ:

$$\begin{aligned} V^h(x,q) \quad = \quad & \text{approx cost on the interval } [0,\delta) \text{ starting from } (x,q) + \\ & + \text{discounted value function at the approx arrival point} \end{aligned}$$

Construction of the quasi-optimal control strategy (2)

- **Autonomous switching**:

$$w_*^\delta = \operatorname*{argmin}_{w \in \mathcal{V}} \{ V^h(g(x,q,w)) + C_a(x,q,w) \}$$

- **Controlled switching**:

$$(x_*^\delta, q_*^\delta) = \operatorname*{argmin}_{(x',q') \in D} \left\{ V^h(x',q') + C_c(x,q,x',q') \right\}$$

- **Continuous control**:

$$u_*^\delta = \operatorname*{argmin}_{u \in U} \left\{ \delta \ell(x,q,u) + e^{-\lambda \delta} V^h(x + \delta f(x,q,u), q) \right\}$$

Convergence of the quasi-optimal strategy

- Under the basic assumptions, this procedure turns out to produce **a minimizing sequence of suboptimal controls** (R. F. – H. Zidani, 2014)
- However, **the discretization parameter δ should be chosen so that**

$$\frac{\left\| V - V^h \right\|_\infty}{\delta} \to 0$$

- In practice, **this is a conservative condition** which is required in lack of a result of convergence for gradients

Policy Iteration

Convergence of **value iteration** may be very slow, and **gets slower as** $h \to 0$. A viable acceleration technique is provided by **policy iteration**, also termed as **Howard algorithm**:

- In the step of **policy evaluation**, the control strategy θ_k is frozen, and the solution v_k of the resulting **linear problem** is computed
- In the step of **policy improvement**, a new control strategy θ_{k+1} is computed by performing **a single nonlinear (value) iteration** on v_k

In large, sparse problems a viable solution is the **Modified Policy Iteration**, in which the solution of the linear problem is replaced by a fixed number N_{it} of iterations (or, in other terms, in which the updating of the policy is performed only once every $N_{it} + 1$ iterations)

Convergence of the Policy Iteration

- The linear problem obtained by fixing the control strategy **is not a mere advection equation** (it also involves switching of the state)
- In more conventional cases, it is known that policy iteration provides **superlinear convergence**
- In the hybrid case, a proof of convergence has been given, under the natural assumption of **monotonicity** of the scheme (R. F. – A. Sassi, 2017)

Error estimates

The technical difficulties in obtaining explicit error estimates for the convergence $V^h \to V$ are inherently greater than for the non-hybrid case. The work is **partly in progress**, main technical tools:

- **Modulus of continuity**: the optimal situation is to require Lipschitz continuity for V and V^h. This calls for additional assumptions
- **Regularization**: typically, obtained by enlarging the reachable set of the controlled system
- **Cascade**: we use a sequence of error estimates obtained on control problems with an increasing number of switchings/jumps of the state

The numerical error in the Lipschitz case turns out to be estimated as

$$\|V - V^h\|_\infty \leq Ch^{1/2} \log |h|$$

(R. F., A. Sassi and H. Zidani, preprint)

Hybrid stabilization of an unstable system (1)

An unstable system in $\Omega = [-1, 1]$ (with controls in $U = [-1, 1]$) can be stabilized by means of **two different dynamics**:

- A **weaker, less expensive** controller
- A **stronger, more expensive** controller

The first controller cannot stabilize the system for any $x \in [-1, 1]$, so **it must switch to the second dynamics on hitting the boundary of** Ω, whereas at internal points **switching between the two dynamics is optional**.

Hybrid stabilization of an unstable system (2)

- The two controllers are described by the **dynamics**

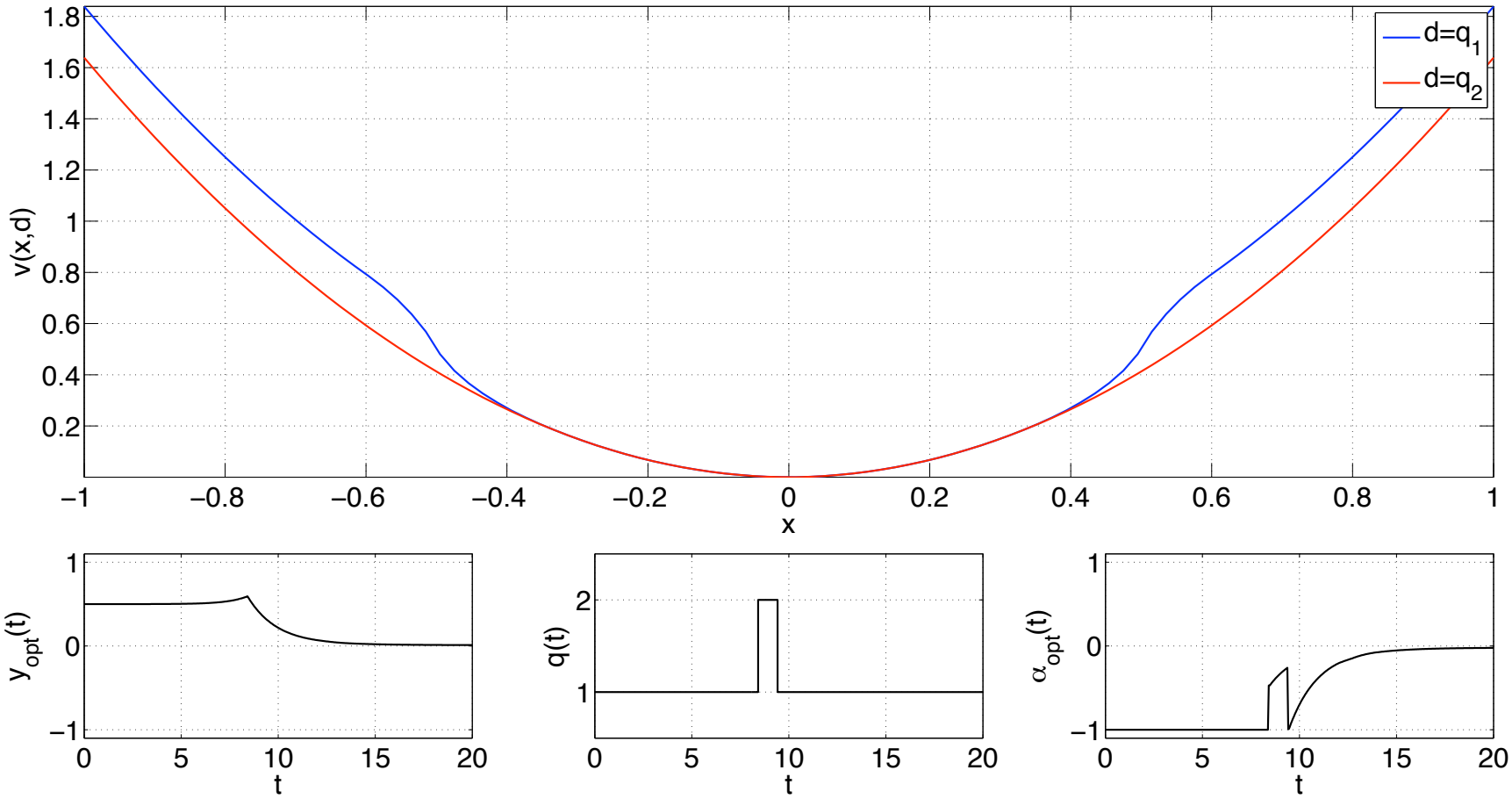
$$f(X(t), 1, u(t)) = X(t) + c_1 u(t)$$
$$f(X(t), 2, u(t)) = X(t) + c_2 u(t),$$

  with $c_2 \gg c_1$, this meaning that the second controller is **"stronger"**;
- Accordingly, the **running costs** associated to the two dynamics are

$$\ell(X(t), 1, u(t)) = X(t)^2 + c_3 u(t)^2,$$
$$\ell(X(t), 2, u(t)) = X(t)^2 + c_4 u(t)^2.$$

  with $c_4 \gg c_3$, this meaning that the second controller is **"more expensive"**;

Hybrid stabilization of an unstable system (3)



Value functions and optimal solution (**unstable initial state**)

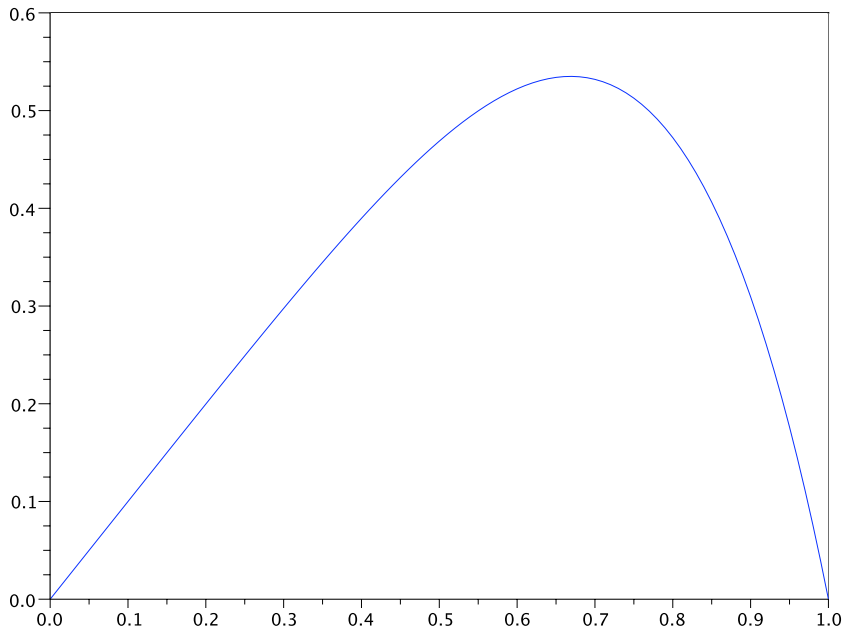
Hybrid stabilization of an unstable system (4)

ϵ	$N_{\text{Value Iteration}}$	$N_{\text{Policy Iteration}}$
10^{-3}	456	8
10^{-6}	1147	10
10^{-12}	2786	12

Convergence of value and policy iteration

Multigear vehicle (1)

A vehicle with **three gears**, and control (throttle) in $U = [0, 1]$). The engine has a given **relationship** $T(\omega)$ **for torque vs angular speed**:



Multigear vehicle (2)

- The **dynamics** associated to the three gears are described by

$$f(X(t), q(t), u(t)) = \frac{3.6}{m}\left(\frac{T\left(\frac{25X(t)}{3r\pi\rho(q(t))}\right)}{r\rho(q(t))}u(t) - c_d\frac{X(t)^2}{3.6}\right)$$

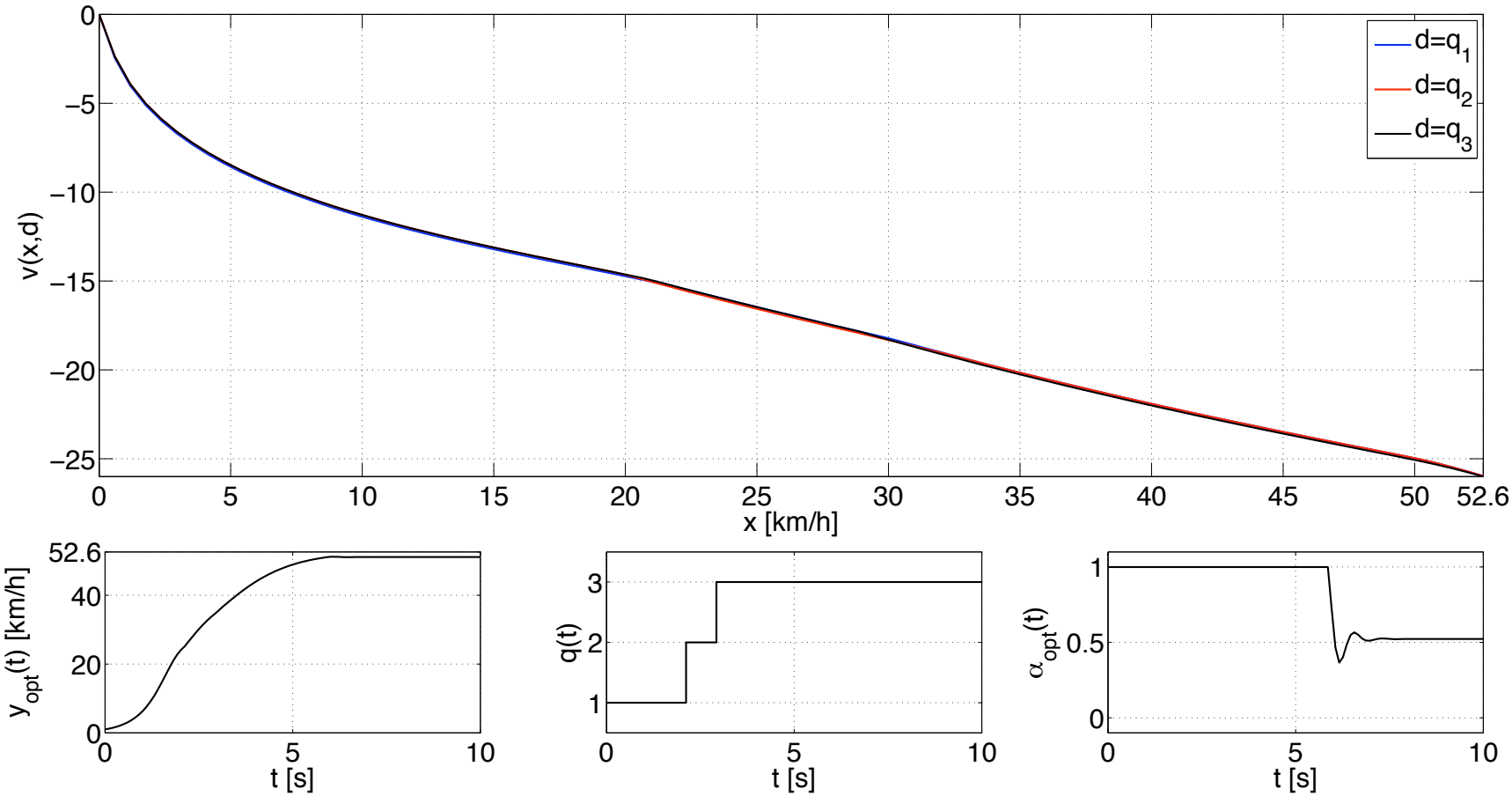
  with $\rho(q)$ **ratio of the gear** q**,** m **mass,** c_d **aerodynamic friction coefficient,** r **radius of the wheels**

- The associated running cost is given by a **combination of travel time and fuel consumption**:

$$\ell(X(t), q(t), u(t)) = -c_1X(t) + c_2u(t)$$

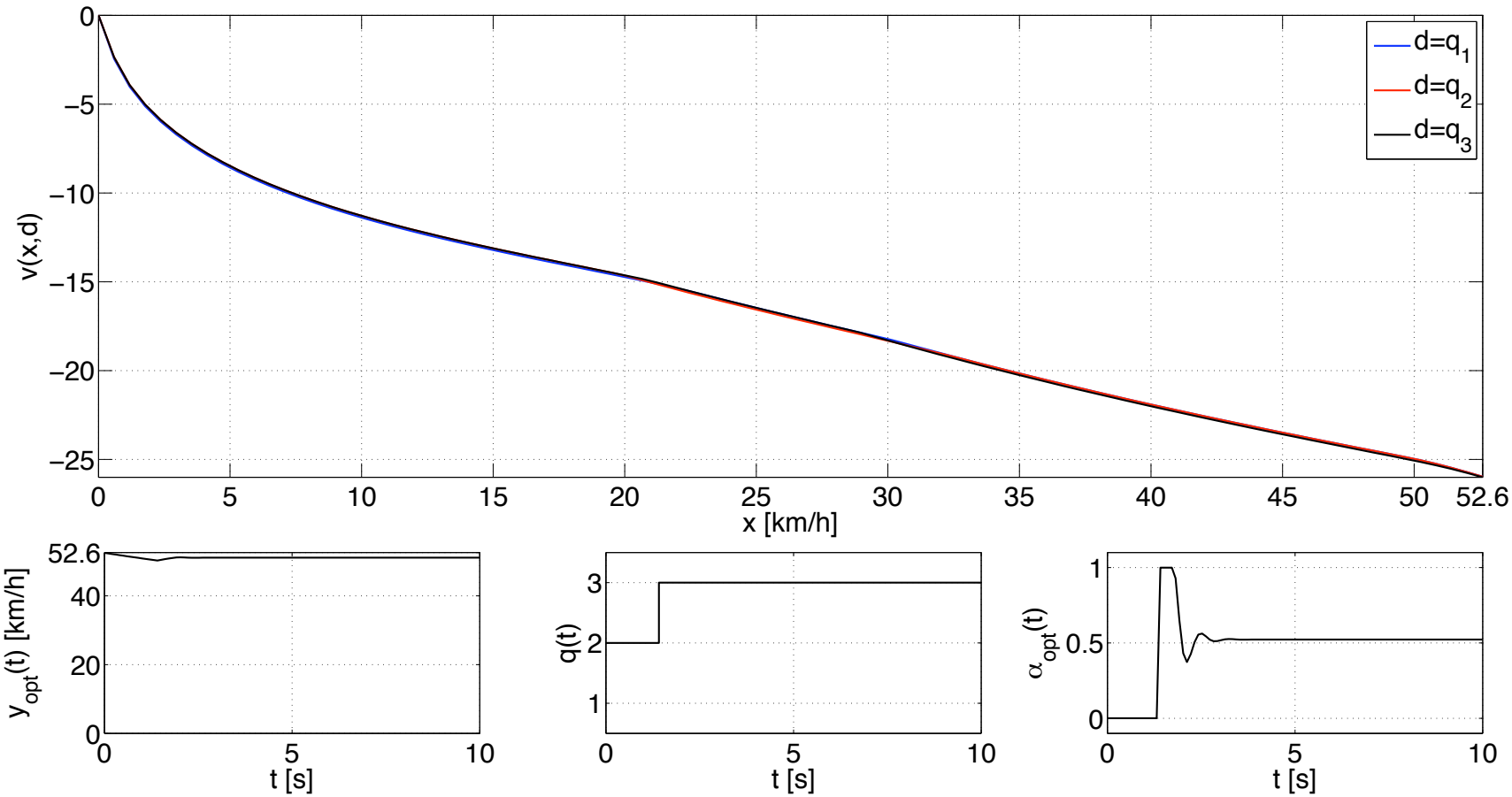
- Constant **switching costs** c_{ij} for changing gear from i to j

Multigear vehicle (3)



First case: starting from a **still vehicle**

Multigear vehicle (4)



Second case: starting from a **moving vehicle**

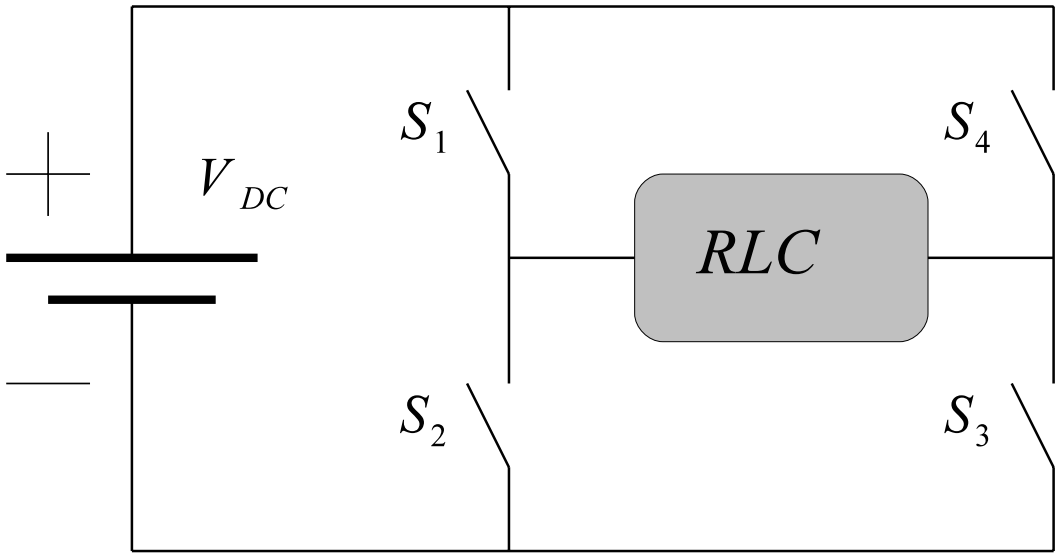
Multigear vehicle (5)

ϵ	$N_{\text{Value Iteration}}$	$N_{\text{Policy Iteration}}$
10^{-3}	337	6
10^{-6}	586	6
10^{-12}	1084	6

Convergence of value and policy iteration

AC/DC inverter (1)

We want to act on the switches $S_1, \ldots, S_4$ in order to obtain a **sinusoidal current** in the RLC load:



Conceptual scheme of a single-phase AC/DC inverter

AC/DC inverter (2)

Following [1], we consider as state variables $X_1 = i_L$ (**the current through the inductor** L, i.e., through the load) and $X_2 = v_C$ (**the voltage across the capacitor** C), the state equations being

$$\begin{cases} \dot{X}_1(t) = \dfrac{V_{DC}}{L}(Q(t) - 2) - \dfrac{R}{L}X_1(t) - \dfrac{1}{L}X_2(t) \\ \dot{X}_2(t) = \dfrac{1}{C}X_1(t). \end{cases} \tag{4}$$

The physical meaning of the discrete state variable **depends on the state of the switches** $S_1, \ldots, S_4$, and more precisely

$$Q(t) = \begin{cases} 1 & \text{if } S_1, S_3 = OFF \text{ and } S_2, S_4 = ON \\ 2 & \text{if } S_1, S_4 = OFF \text{ and } S_2, S_3 = ON \\ 3 & \text{if } S_2, S_4 = OFF \text{ and } S_1, S_3 = ON. \end{cases}$$

AC/DC inverter (3)

The cost functional is defined so as to force the system to evolve (approximately) **along an ellipse of the state space** (see [1]), namely

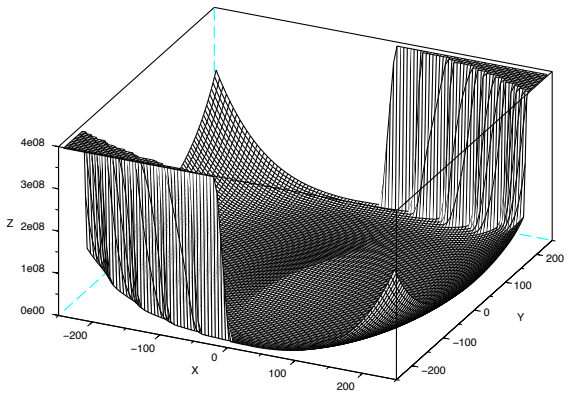
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = c,$$

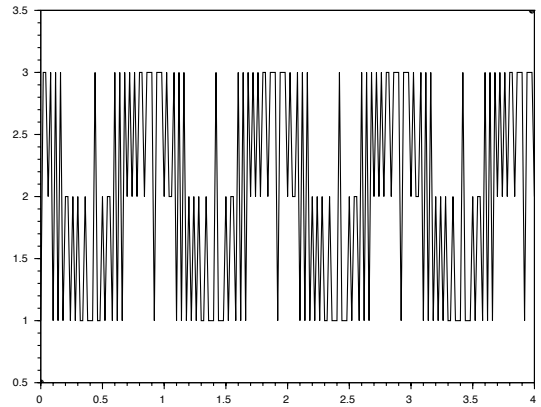
in which the constants a and b are defined in terms of the physical parameters R, L, C and of the desired pulsation ω. This makes it natural to define the **running cost** as

$$\ell(x, q, \alpha) = \left(\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - c \right)^2 . \tag{5}$$

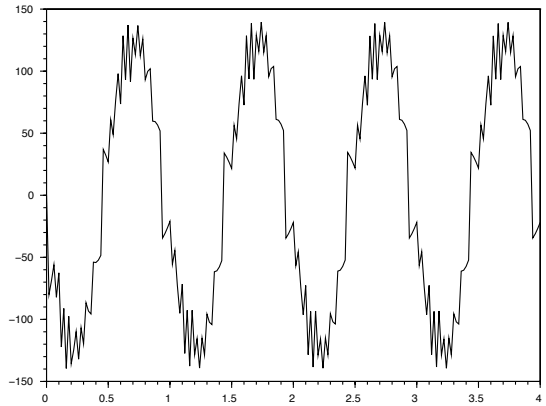
With the parameters chosen, the required output of the system would be given by two sinusoids of amplitude respectively 126 A for X_1 and 200 V for X_2, both at the frequency of 1 Hz.
Discretization: SL, 100×100 grid, Modified Policy Iteration

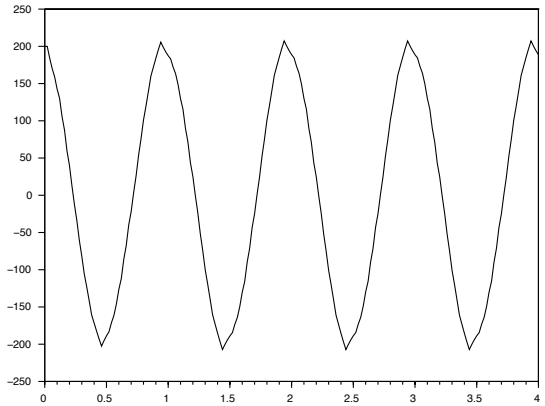
AC/DC inverter (4)





Value function and optimal switching for the DC/AC inverter

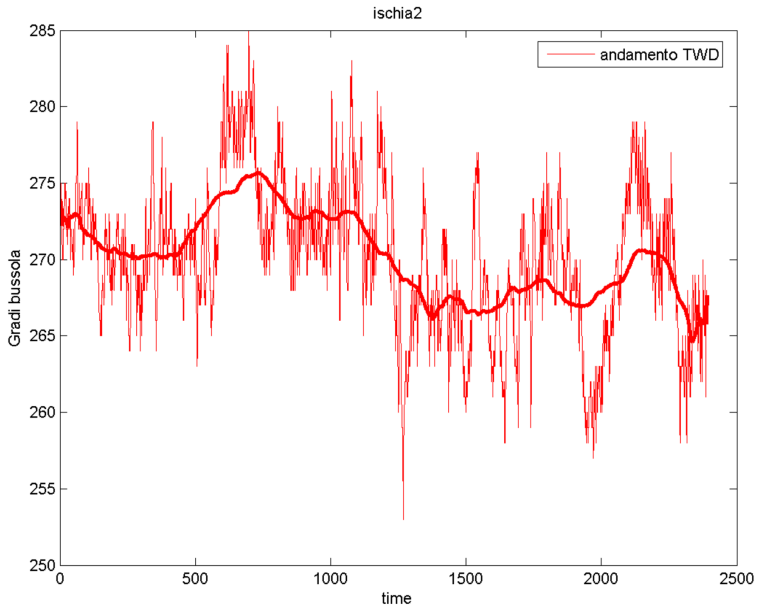




Trajectories $X_1(t)$ and $X_2(t)$ for the DC/AC inverter

Tacking strategy (1)

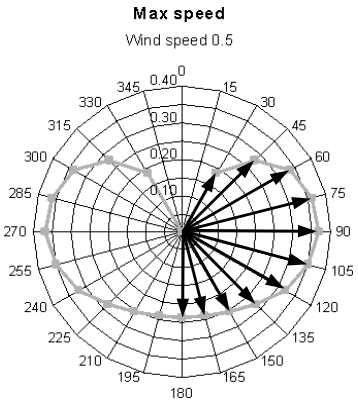
The last example treats a **stochastic hybrid problems**, i.e., the **optimal tacking strategy of a sailing boat** in a windward leg of a regatta.

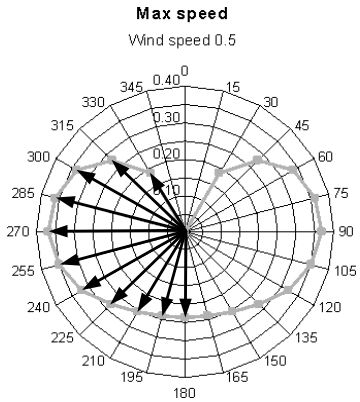


The boat sails at about 45° from the wind direction, which has a **partly stochastic evolution**:

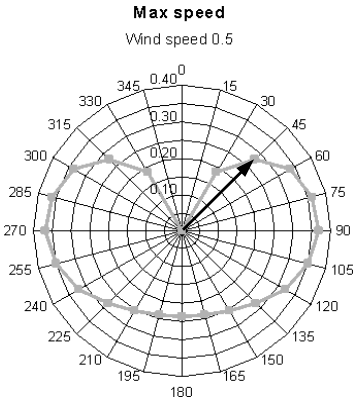
$$d\theta = c_\theta dt + \sigma_\theta dW$$

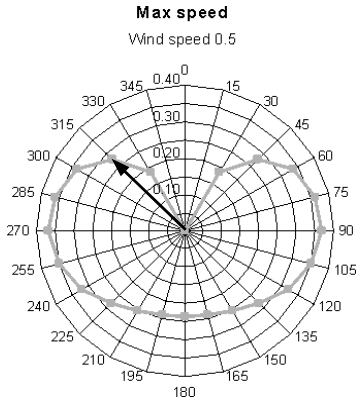
Tacking strategy (2)





original





simplified

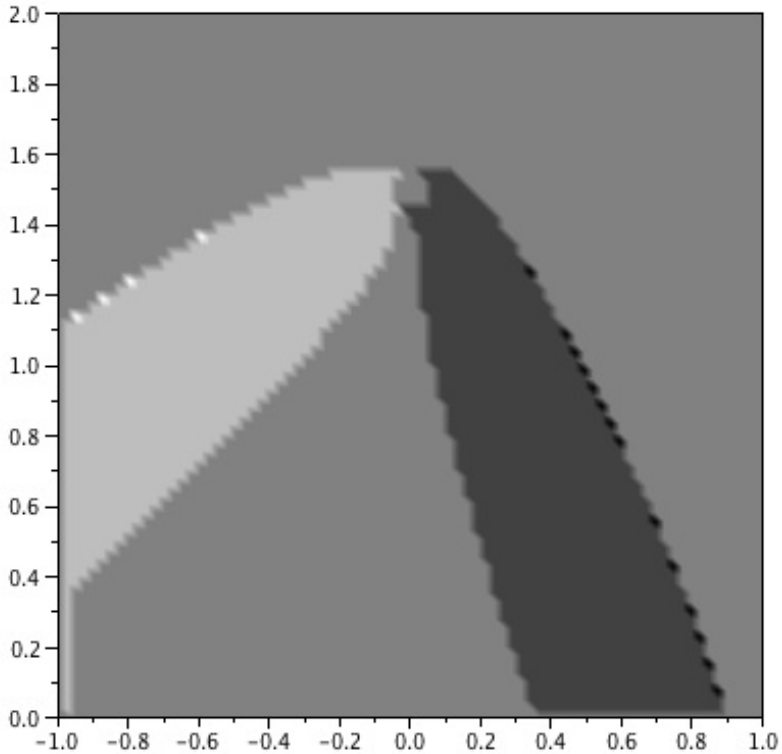
$Q = 1$ $Q = 2$

Original and simplified dynamics of the system

Tacking strategy (3)

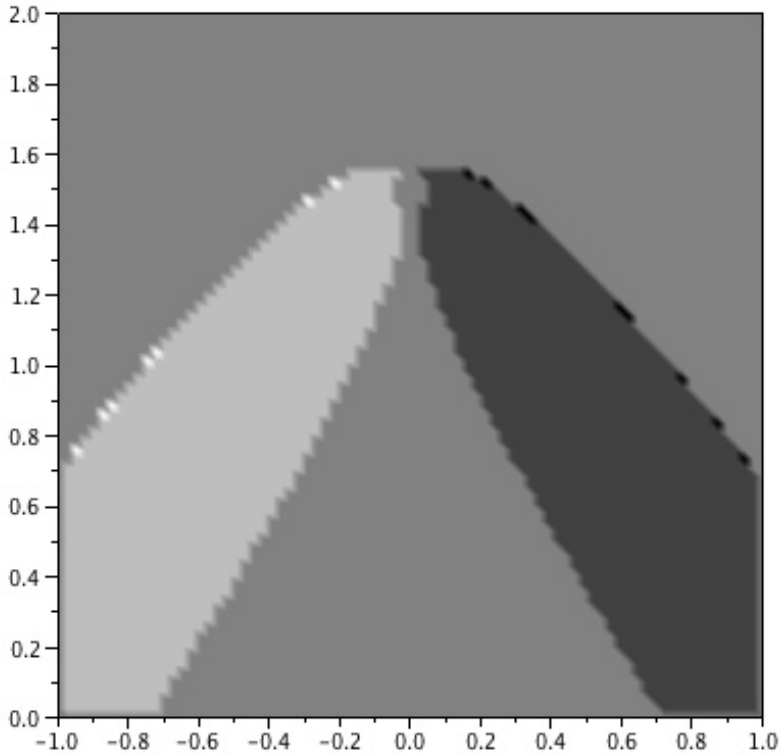
- **State space**: $\mathbb{R}^3 \times \{1, 2\}$ (two space dimensions + wind direction + boat direction (L, R))
- **Target problem**: minimum time + penalized distance from the windward mark as a stopping cost
- **Stochastic problem**: a straightforward extension of the approximation theory can be done for stochastic hybrid problems (R. F. – A. Festa, in preparation)
- **Discretization**: SL, $80 \times 80 \times 80$ grid, Modified Policy Iteration

Tacking strategy (4)



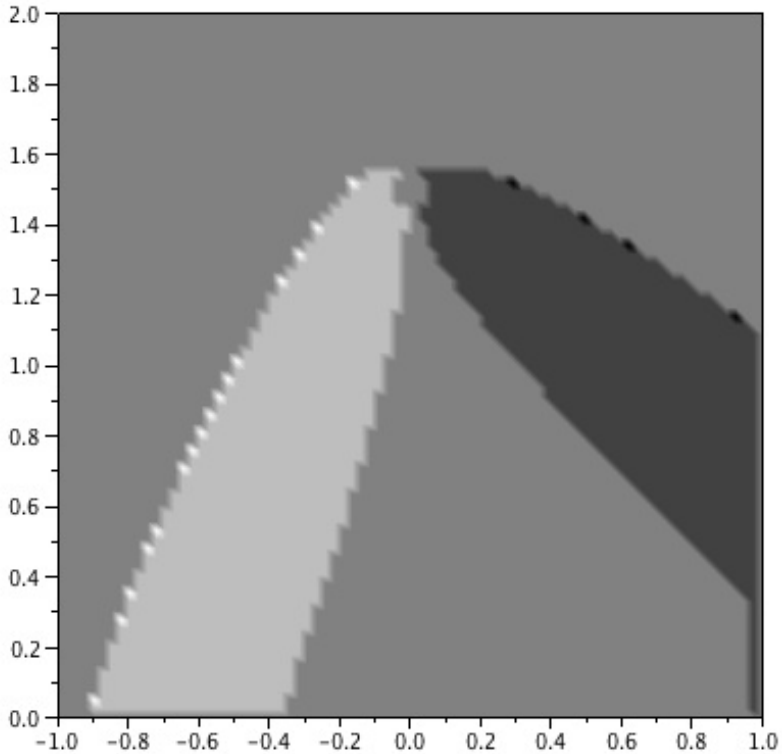
Switching sets, $c_\theta = 0$, wind direction $\theta = -.25$

Tacking strategy (4)



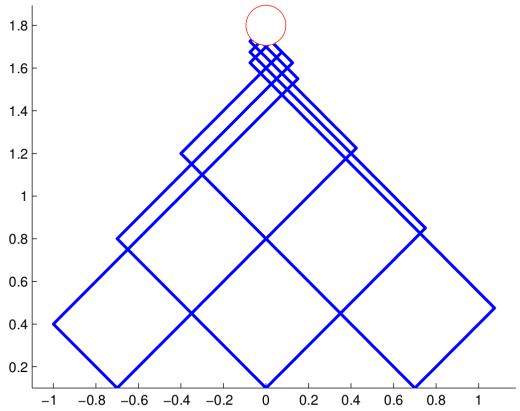
Switching sets, $c_\theta = 0$, wind direction $\theta = 0.0$

Tacking strategy (4)

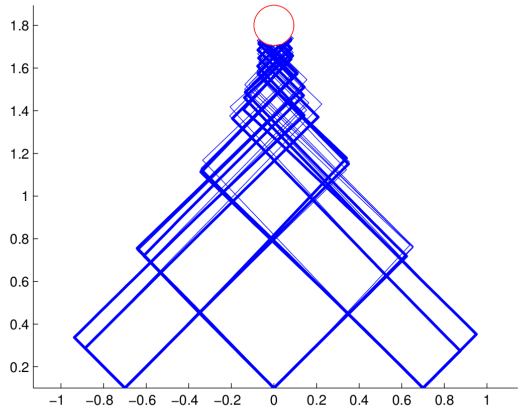


Switching sets, $c_\theta = 0$, wind direction $\theta = 0.25$

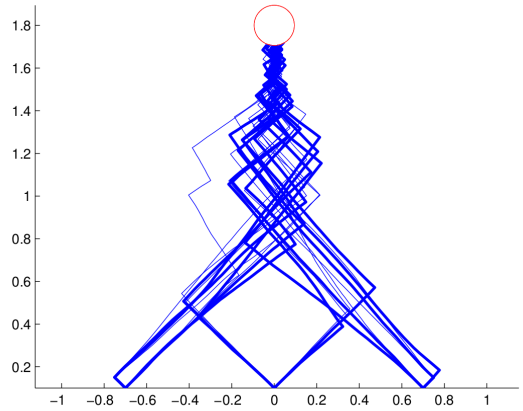
Tacking strategy (5)



$\sigma_\theta = 0$



$\sigma_\theta = 0.01$



$\sigma_\theta = 0.1$

Optimal trajectories for increasing variance of the wind direction

Final remarks

- **Sound theoretical framework**, for both the theoretical and the computational aspects
- **Viable and robust design** of optimization-based controllers in low dimension
- Possibility of using **acceleration techniques** of Policy iteration type
- Application to **stochastic problems** requires minimal adaptations

References

- M.S. Branicky, V. Borkar, and S. Mitter, *A unfied framework for hybrid control problems*, IEEE Trans. Autom. Contr., **43** (1998), 31–45
- J. Chai, R.G. Sanfelice, *Hybrid feedback control methods for robust and global power conversion*, IFAC–PapersOnLine **48-27** (2015), 298–303.
- S. Dharmatti and M. Ramaswamy, *Hybrid control system and viscosity solutions*, SIAM J. Contr. Optim., **34** (2005), 1259–1288
- R. Ferretti and A. Festa, *A hybrid control approach to the route planning problem for sailing boats*, in preparation
- R. Ferretti and A. Sassi, *A semi-Lagrangian algorithm in policy space for hybrid optimal control problems*, ESAIM: COCV, to appear
- R. Ferretti and H. Zidani, *Monotone numerical schemes and feedback construction for hybrid control systems*, J. of Optim. Theory and Applications **165** (2014), 507–531
- A. Sassi, *Tecniche di Programmazione Dinamica nell'ottimizzazione di sisteni di controllo ibridi*, MSc thesis, Università Roma Tre, 2012