Swing-up and stabilization of a Twin-Pendulum under state and control constraints by a fast NMPC scheme ⋆

Mazen Alamir [a] André Murilo A Pinto [a],

[a] *CNRS/GIPSA-Lab. SYSCO-Team*
Domaine Universitaire
38400 Saint Martin d'Hères, France

Abstract

In this paper, a real-time implementable Nonlinear Model Predictive Control scheme is proposed for the swing-up and the stabilization of a Twin-Pendulum system under control and state constraints. The basic feature lies in a particular parametrization of the set of candidate control profiles leading to a decision variable that may take only 3 admissible values. Simulations are proposed to assess the efficiency of the proposed feedback.

Key words: Twin Pendulum; Swing-up; Constrained Stabilization; Fast NMPC.

1 Introduction

The Twin-Pendulum system is depicted on figure 1. Two pendulums share the same axis of rotation O_1 that is positioned on a cart which is controlled through its acceleration v that must satisfy a saturation constraint. On the other hand, the excursion of the cart is constrained by two boundary stops, namely, $r \in [-r_{max}, +r_{max}]$.

The global stabilization of this system falls in a long tradition of works on swing-up like systems [2,4,9]. An exhaustive study of these works is beyond the scope of this short communication. However, as long as predictive control is concerned and to the best of our knowledge, the only work that used real-time implementable NMPC like approaches [6] to solve such problems has been proposed by [1] where double inverted pendulum on a cart has been stabilized without constraints on the cart excursion.

As long as the specific twin-pendulum system is concerned, there are but few papers [8,3,2] that proposes a

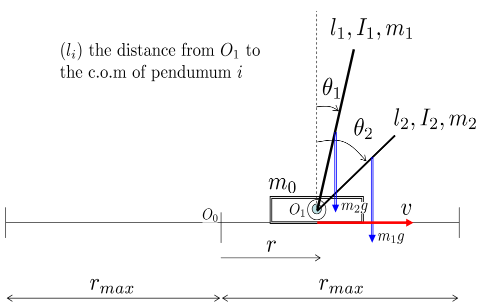

Fig. 1. View of the Twin-Pendulum system

solution to the swing-up and stabilization of this system. In [8,2] Lyapunov approach is used together with stability analysis based on the concept of invariant sets, this enables to show that the two pendulums energies converge to the desired level, the stability issue is not explicitly mentioned. In [3], a rotational twin pendulum is studied where a state feedback linearization approach is used and the stability of the resulting zero dynamics is analyzed.

The approach amounts to stabilize first one of the two pendulums (the smallest) then to stabilize the other using local oscillations of the first around the upward

⋆ This paper was not presented at any IFAC meeting. Corresponding author M. Alamir. Tel. +33476826326. Fax +33476826388.

Email addresses: mazen.alamir@inpg.fr (Mazen Alamir), andre.murilo@lag.ensieg.inpg.fr (André Murilo A Pinto).

desired position. Investigation of the system properties around its vertical upward position is proposed in [5]. All these works tackle the problem in a continuous time setting leading to a necessarily small sampling period. Moreover, state constraints on the position of the cart is never considered.
In this paper, a real time implementable NMPC scheme is proposed that stabilizes the Twin-pendulum system under the control and the state constraints mentioned above.
The paper is organized as follows. First, The system model is given (section 2). Then, the basic idea of the control scheme is explained (section 3) before an explicit expression of the state feedback law is given (section 4). The paper ends with some numerical simulations.

2 The Twin-pendulum model

The equations governing the dynamics of the Twin-Pendulum system are given by :

$$\ddot{r} = v \tag{1}$$

$$\ddot{\theta}_i = -\alpha_i(\cos\theta_i)v + \beta_i \sin\theta_i \quad i \in \{1,2\} \tag{2}$$

where $\alpha_i = \frac{m_i l_i}{m_i l_i^2 + I_i}$ and $\beta_i = g\alpha_i$. Note that the control input here is the acceleration of the cart that is steered via a system of DC-motor driven belt. Note also that for evident controllability reasons, it is assumed that $\alpha_1 \neq \alpha_2$. Finally, while in the real world, friction is unavoidable, it is here neglected to be short and in order to concentrate on the main features. When applied to the real system, the proposed controller has to be coupled with any existing friction compensation scheme (see for instance [7] and the reference therein).

The purpose of the paper is to define a sampled time state feedback that stabilizes the sub-state ($\theta_1 = \dot{\theta}_1 = \theta_2 = \dot{\theta}_2 = \dot{r} = 0$) while continuously maintaining (r, v) in the admissible set $[-r_{max}, +r_{max}] \times [-v_{max}, +v_{max}]$.

Remark 1 *Note that in this paper, saturation on the acceleration $\ddot{r} = v$ are used for simplicity. It goes without saying that constraint on the force can be easily handled through lagrange equations.*

3 Intuitive presentation of the control law

The basic idea is to monitor the energy levels of both pendulums in order to make them simultaneously reach the convenient "*values*". Energy levels may be defined for both pendulum as follows :

$$E_i(x) := \frac{1}{2}\dot{\theta}_i^2 + \beta_i(cos\theta_i - 1) \quad ; \quad i \in \{1,2\}. \tag{3}$$

Indeed, these functions show the following properties :
✓ Since $\dot{E}_i(x) = -\alpha_i\dot{\theta}_i\cos\theta_i v$, it comes that $E_i(x)$ is constant along trajectories such that $v \equiv 0$. In particular, if at some instant t, $E_i(x(t)) = 0$, then the future trajectory under $v = 0$ leads to the state $\theta_i = \dot{\theta}_i = 0$, namely, the desired vertical position.
✓ Whenever $\dot{\theta}_i \cos\theta_i \neq 0$, the level E_i can be "*incrementally*" controlled (to be steered closer to 0) by even very small control values of v.
✓ Over time intervals corresponding to almost vanishing $\dot{\theta}_i \cos\theta_i$'s, even high control levels can be used without sensibly altering the value of E_i.
✓ Combining the two preceding facts together with the fact that $\dot{\theta}_1 \cos\theta_1$ and $\dot{\theta}_2 \cos\theta_2$ have "*no special reasons*" to vanish systematically at the same instants, it is clear that the control may be used over time intervals over which $\dot{\theta}_1 \cos\theta_1 \approx 0$ [resp. $\dot{\theta}_2 \cos\theta_2 \approx 0$] to "*improve* " the value of E_2 [resp. E_1].
In order to put the above discussion in a more formal setting, consider the following definition :

$$E(x) := \max\Big\{\frac{|E_1(x)|}{E_1^{max}}\ ,\ \frac{|E_2(x)|}{E_2^{max}}\Big\},$$

where $E_i^{max} = 2\beta_i$ is the value of $|E_i|$ when the i^{th} pendulum is at the downward vertical position ($\theta_i = \pi, \dot{\theta}_i = 0$). The control objective is therefore to steer the value of the function $E(x)$ to 0 since this means that the two pendulums are on the "*good orbits*". The fact that this is always possible comes from the preceding discussion. However, this would be done by applying control actions v that may lead to the violation of the constraint on the cart excursion.

In order to avoid this constraint violation, a change in the control variable v is made. More precisely, a sampled-time pre-compensation is used on the state of the cart according to :

$$\ddot{r}(k\tau_s + t) = v(k) = -K_{cart}\begin{pmatrix} r(k) - r_d \\ \dot{r}(k) \end{pmatrix}, \tag{4}$$

for all $t \in [0, \tau_s[$ where $\tau_s > 0$ is some sampling period, $K_{cart} \in \mathbb{R}^{1\times 2}$ is a stabilizing feedback gain matrix while $v(k)$ and $r(k)$ stand for $v(k\tau_s)$ and $r(k\tau_s)$ respectively. With this pre-compensation, the new control variable is $r_d \in [-r_{max}, +r_{max}]$ and the constraint on the excursion of r is structurally respected during the swing-up phase. It is the tuning of this pre-compensation loop that enables the control saturation constraint to be respected. This is done by designing a stabilizing gain K_{cart} such that applying the maximal increment $2r_{max}$ at rest position ($\dot{r} = 0$) leads to closed loop trajectory on v that does not violate the constraint.

Consider some instant $k\tau_s$ where the state of the system is $x(k\tau_s)$. Consider the new value of the new control variable $r_d(k)$ to be applied during the next sampling period. The computation of $r_d(k)$ is based on the prediction of the resulting cost function $E(k\tau_s + T_p)$ based on a constant profile $r_d(k\tau_s + t) = r_d(k)$ for all $t \in [0, T_p]$ where T_p is the prediction horizon that is taken greater than the response time of the closed loop (4). More precisely, consider the following cost function :

$$J(x(k), p) := E\Big(F(x(k); T_p; p)\Big), \quad (5)$$

where $F(x(k); T_p; p)$ is the solution of the system (1)-(2) at instant T_p when starting at instant 0 from the initial conditions $x(k)$ under the sampled feedback (4) in which $r_d = p$ is used and where

$$T_p \geq \text{settling of (4)}. \quad (6)$$

This last condition is used in order for the cost function (5) to reflect the value of E one can reach under constant $r_d = p$. Indeed, under the condition (6), v almost vanishes after T_p time units and E becomes invariant (at least over a short time because of friction). The discussion of this section concerning the effect of v on E_1 and E_2 clearly shows that applying the cost function (5) in a receding horizon way leads to the convergence of E to zero. By this way, the resulting receding horizon control achieves the swing-up phase that terminates with the two pendulums in at rest vertical upward positions with zero cart velocity.

The minimization of the cost function (5) has to be a priori done over the admissible set

$$p \in [-r_{max}, +r_{max}].$$

However, the resulting optimization may still be heavy to perform in the available computation time that must be a fraction of the sampling period τ_s. Fortunately, examination of $\dot{E}_i = -\alpha_i \dot{\theta}_i \cos(\theta_i) v$ shows that any neighborhood of $v = 0$ contains improving solutions provided that $\dot{\theta}_i \cos\theta_i \neq 0$ which is almost always the case. That is the reason why, the following over simplified admissible set for p is used at instant $k\tau_s$:

$$\mathbb{P}(k) := \Big\{ r(k) - d_r(E(k)) \ , \ r(k) \ , \ r(k) + d_r(E(k)) \Big\} \bigcap \, [-r_{max}^{RHC} \ , \ +r_{max}^{RHC}] \quad ; \quad r_{max}^{RHC} < r_{max}, \quad (7)$$

where $d_r(E)$ is given by (see figure 2) :

$$d_r(E) = \varepsilon_0 + \delta_r^{max} E. \quad (8)$$

Namely, for maximal value of E ($E = 1$), the allowable increment in $p = r_d$ is maximal and is equal to $\varepsilon_0 + \delta_r^{max}$ while for $E = 0$, the maximal increment of r_d is equal to ε_0.

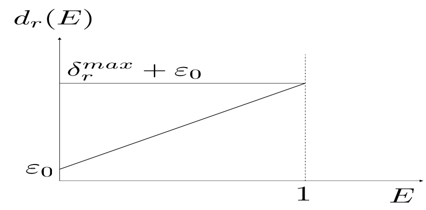

Fig. 2. Definition of the admissible range of the control parameter $p = r_d$ as a function of the present value of the energy function E.

Note also that the resulting value of r during the swing phase must remain in the more restrictive range

$$[-r_{max}^{RHC}, +r_{max}^{RHC}] \subset [-r_{max}, +r_{max}].$$

This leaves a room for manoeuver to be available for the final phase where one switches to a local linear stabilizing controller (see figure 3).

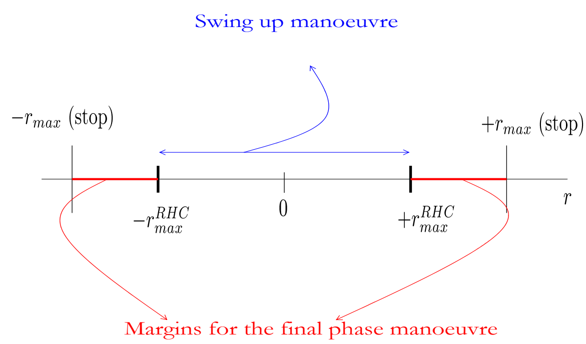

Fig. 3. Definition of the admissible domain for swing up and the margins for the final stabilization manoeuvre.

4 Explicit formulation of the receding-horizon feedback

At each instant $k\tau_s$ during the swing-up phase (the definition of the end of this phase is clearly stated in the forthcoming sequel), the following constrained optimization problem is defined :

$$\hat{p}(x(k)) := arg \min_{p \in \mathbb{P}(k)} \quad J(x(k), p),$$

where the cost function $J(x(k), p)$ is given by (5) and the admissible set $\mathbb{P}(k)$ by (7) and (8). Note that the underlying optimization problem is defined over the discrete set of admissible values that contains 3 elements, namely $r(k)$ and $r(k) \pm d_r(E(k))$. This means that the computation of the optimal solution costs three integration of the differential system (1)-(2) and (4) and takes a rather short and deterministic computation time. More precisely, the simulations on MATLAB are 4-time faster than the real time (on a PENTIUM III 1.2GHz PC). This suggests that using compiled programming languages (C or whatever), the computation would be about 100 time faster than the real-time implementation requirement.

The solution $\hat{p}(x(k))$ is used in (4) to define the sampled time control $v(k)$ during the sampling period $[k\tau_s, (k+1)\tau_s[$:

$$v(k) = K_{RHC}(x(k)) = -K_{cart}\begin{pmatrix} r(k) - \hat{p}(x(k)) \\ \dot{r}(k) \end{pmatrix}. \quad (9)$$

where K_{RHC} is the feedback of the receding horizon controller. This feedback is applied until the two pendulums are almost in at rest upward vertical positions with the cart being at almost vanishing velocity. Before this situation is mathematically characterized, let us define the local feedback law to be used when this configuration is "*fired*".

Since the final position of the cart is not imposed (provided that it belongs to the admissible range $[-r_{max}, +r_{max}]$), the local controller is designed for the five dimensional system with the following state :

$$z := \begin{pmatrix} \dot{r} & \theta_1^m & \dot{\theta}_2 & \theta_2^m & \dot{\theta}_2 \end{pmatrix} \in \mathbb{R}^5,$$

where $\theta_i^m \in [-\pi, +\pi]$ is the angle that is equal to θ_i (modulo 2π). Linearizing the system equations around $z = 0$ gives the following linear system

$$\dot{z} = Az + Bv, \quad (10)$$

with

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\alpha_2 & 0 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} 1 \\ 0 \\ \beta_1 \\ 0 \\ \beta_2 \end{pmatrix}.$$

Using a discrete time version of (10) together with standard linear control design tools (LQR, pole placement or whatever), it is possible to define a feedback gain $K_L \in \mathbb{R}^{1\times 5}$ such that the sampled feedback control law

$$v(k\tau_s + t) = -K_L z(k\tau_s) \quad \forall t \in [0, \tau_s[, \quad (11)$$

locally stabilizes (10) (recall that this is due to the assumption $\alpha_1 \neq \alpha_2$). Moreover, solving the discrete Lyapunov equation for the closed loop system enables a positive definite matrix S to be obtained such that for sufficiently small $\rho_0 > 0$, the neighborhood of the origin given by

$$\mathcal{A}_0 := \left\{ z \in \mathbb{R}^5 \quad | \; z^T S z \leq \rho_0 \right\}, \quad (12)$$

is invariant and contained in the region of attraction of the local linear controller (11).

Classical hybrid control schemes amount then to switch from the receding horizon control law (9) to the linear control law (11) as soon as the state of the system is such that $z(k) \in \mathcal{A}_0$. In our case, this would lead to transient excursion on r that violates the constraint $r(t) \in [-r_{max}, +r_{max}]$ (see figure 4).

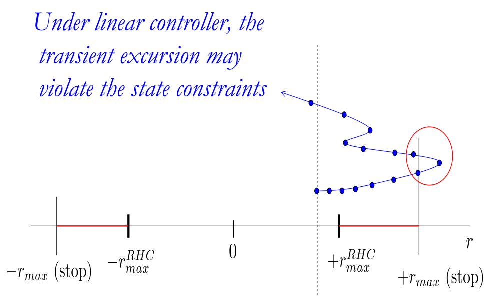

Fig. 4. Switching to the linear controller without any particular care may lead to transient state trajectories that violate the state constraints.

Indeed, while the nonlinear receding horizon controller respects structurally this constraint, the linear controller design does not contain such considerations. Consequently, additional condition has to be checked before switching to the linear controller. This condition must express the fact that transient excursion on the cart position during the stabilization phase under the linear controller remains in the admissible range $[-r_{max}, +r_{max}]$. This can be explicitly written as follows :

$$\varphi(x(k)) \leq r_{max} \quad (13)$$

where :

$$\varphi(x(k)) := \max_{i\in\{1,\dots,N\}} \left| r(k) + \tau_s \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} C(A_d - B_dK_L)^1 z(k) \\ \vdots \\ C(A_d - B_dK_L)^i z(k) \end{pmatrix} \right| \quad (14)$$

where $C = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{1\times 5}$ is the matrix that selects the first element of a 5-dimensional vector (here the cart velocity) while the pair (A_d, B_d) stands for the discrete time linearized system's matrices. Note that the term

$$(A_d - B_dK_L)^i z(k),$$

is nothing but the state z on the closed loop trajectory of the linearized model at instant $(k+i)\tau_s$. Therefore, the term $C(A_d - B_dK_L)^i z(k)$ denotes the velocity at the same instant. Summing up these terms, multiplying the result by τ_s and adding the result to the initial position $r(k)$ gives a tight estimation of the intermediate positions of the cart on the resulting closed loop trajectory. Note that N in (14) is some sufficiently high integer satisfying

$$N\tau_s > \textit{settling time of the local closed loop system.}$$

To summarize, the switch between the receding horizon controller (9) and the local controller (11) is fired when the two following conditions are satisfied at some instant $k\tau_s$ [see (12)-(13)] :

$$z(k) \in \mathcal{A}_0 \quad ; \quad \varphi(z(k)) \le r_{max}.$$

The sampled time hybrid controller can therefore be expressed as follows

$$v(k) := \begin{cases} -K_L z(k) & \text{if } x(k) \in \mathcal{A}_0 \text{ and } \varphi(x(k)) \le r_{max} \\ K_{RHC}(x(k)) & \text{otherwise} \end{cases} \quad (15)$$

5 Simulations

In this section, some simulation results are proposed to assess the efficiency of the proposed feedback in stabilizing the Twin-Pendulum system. In these simulations, the physical parameters given on table 1 are used. The settling time for the cart pre-compensation (4) is taken equal to 0.4 s while the prediction time for the definition of the cost function (5) used in the receding-horizon state feedback is taken equal to $T_p = 1$ s. The other control parameters are depicted on table 2. When several values are given, this refers to different scenarios. The switching parameter ρ_0 appearing in the definition of the subset $\mathcal{A}_0$ [see (12) and (15)] has been defined such that

Parameter	Value	Unity
m_0	2.0	kg
m_1	0.2	kg
m_2	0.1	kg
l_1	1.0	m
l_2	0.5	m
I_1	0.1	$kg \cdot m^2$
I_2	0.0125	$kg \cdot m^2$

Table 1
The physical parameters for the Twin-Pendulum system

Parameter	Value	Unity
ε_0	0.1	m
r_{max}^{RHC}	0.4	m
r_{max}	0.8	m
δ_r^{max}	$\in \{0.1, 0.3, 0.4\}$	m
τ_s	$\in \{0.1, 0.2\}$	s

Table 2
The parameters of the controller for the Twin-Pendulum system. Several values refer to different simulations

$$\rho_0 = z_0^T S z_0 \quad ; \quad z_0 = \begin{pmatrix} 0 & \frac{\pi}{6} & 0 & \frac{\pi}{6} & 0 \end{pmatrix}^T .$$

Figures 5 and 6 show two simulations of the closed-loop behavior when the system starts from the at rest downward position $x(0) = \begin{pmatrix} 0 & 0 & \pi & 0 & \pi & 0 \end{pmatrix}$ for two different values of the maximum allowable increment on the cart position, $\delta_{max}^r = 0.3$ m and 0.1 m respectively. Note that when δ_{max}^r increases, faster achievement of the task may be realized at the expense of higher control $v = \ddot{r}$.

Figure 7 shows the swing-up starting from the same initial state with a different sampling period $\tau_s = 0.2$ s rather than 0.1 s. Note the decrease in the control level despite the fact that the higher value $\delta_{max}^r = 0.4$ m has been used.

Finally, figure 8 shows how the control enables the constraint on the maximum excursion of the cart to be handled. Indeed, the initial state in this scenario is such that the cart is on the boundary of the admissible region $r(0) = 0.8$ m.

6 Conclusion

In this paper, a real-time implementable NMPC scheme is proposed for the swing-up and the stabilization of a Twin-Pendulum system under control and state constraints. Two facts are exploited to derive the fast

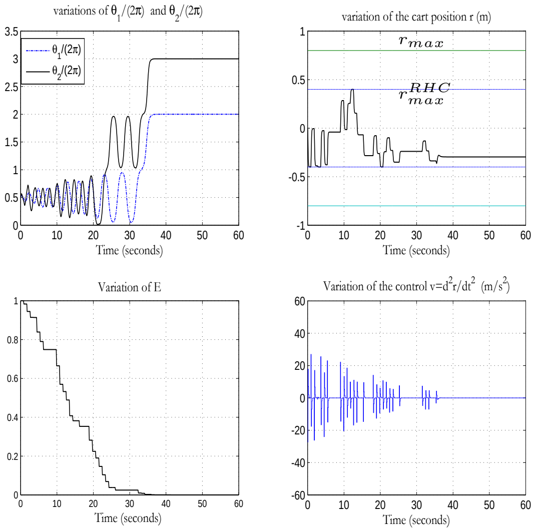

Fig. 5. Closed-loop behavior of the Twin-Pendulum systems. Initial condition $x(0) = (0, 0, \pi, 0, \pi, 0)$. Sampling period $\tau_s = 0.1$ s. Maximum admissible increment on cart position $\delta_r^{max} = 0.3$ m. This scenario is to be compared to the one depicted on figure 6 where $\delta_r^{max} = 0.1$ m is used leading to higher response time but quite lower control level.

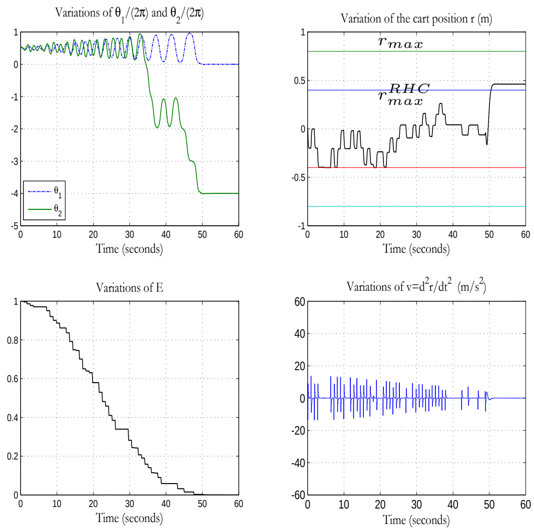

Fig. 6. Closed-loop behavior of the Twin-Pendulum systems. Initial condition $x(0) = (0, 0, \pi, 0, \pi, 0)$. Sampling period $\tau_s = 0.1$ s. Maximum admissible increment on cart position $\delta_r^{max} = 0.1$ m. This scenario is to be compared to the one depicted on figure 5 where $\delta_r^{max} = 0.3$ m is used leading to lower response time but quite higher control level.

NMPC scheme: the first one lies in the extensive use of the energy function while the second is a low dimensional open-loop control parametrization including a change

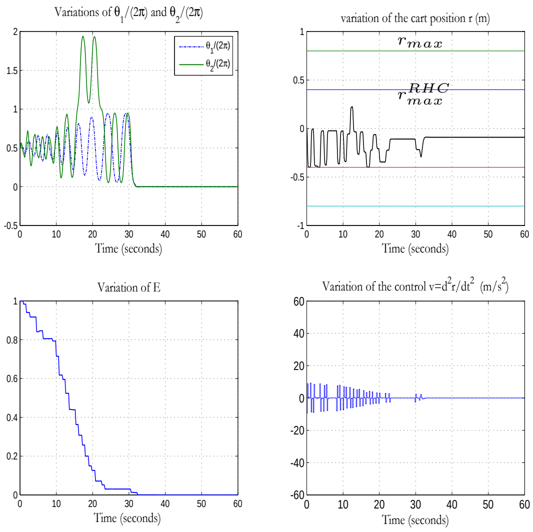

Fig. 7. Closed-loop behavior of the Twin-Pendulum systems. Initial condition $x(0) = (0, 0, \pi, 0, \pi, 0)$. Sampling period $\tau_s = 0.2$ s instead of 0.1 s used in the scenarios depicted on figures 5 and 6. Maximum admissible increment on cart position $\delta_r^{max} = 0.4$ m. Note the very low control value in comparison to figure 5 and 6 despite the fact that the maximum allowable increment is higher. This is due to the high sampling time.

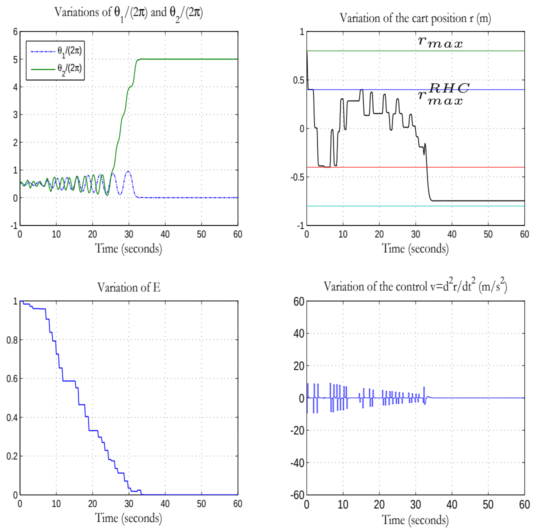

Fig. 8. Closed-loop behavior of the Twin-Pendulum systems. Initial condition $x(0) = (0.8, 0, \pi, 0, \pi, 0)$. Sampling period $\tau_s = 0.2$ s. Note how the controller handles the constraint on the maximum excursion of the cart position that is initially on the boundary of the admissible region ($r(0) = 0.8$ m)

in the manipulated variable in order to structurally meet the state constraint on the cart excursion. This result in a predictive control scheme in which, at each decision instant, the decision variable can only take 3 possible values. The computation time associated to the proposed NMPC scheme is therefore equal to the time necessary to integrate the system equations 3 times.

As it has been mentioned in the introduction, the complete realistic control scheme needs the friction to be compensated as well as constant bias on the position sensors. However, for these problems, there are rather classical solutions that may be used on the experimental setting under construction.

References

[1] M. Alamir. A low dimensional contractive NMPC scheme for nonlinear systems stabilization: Theoretical framework and numerical investigation on relatively fast systems. In *Proceedings of the International Workshop on Assessment and Future Directions of Nonlinear Model Predictive Control. LNCIS Series Springer-Verlag*, 2006.

[2] K. J. Åstrom and K. Furuta. Swing up a pendulum by energy control. *Automatica*, 36(2):287–295, 2000.

[3] S. A. Bortoff. Robust swing-up control for a rotational double pendulum. In *Proceedings of the 13th IFAC World Congress*, pages 413–419, San Fransisco, California, 1996.

[4] D. Chatterjee, A. Patra, and H. K. Joglekar. Swing-up and stabilization of a cartpendulum system under restricted cart track length. *Systems & Control Letters*, 47(4):355–364, 2002.

[5] K. H. Lundberg and J. K. Roberge. Clissical dual-inverted pendulum control. In *Proceedings of the 2003 IEEE Conference on Decision and Control*, pages 4399–4404, 2003.

[6] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert. Constrained model predictive control: stability and optimality. *Automatica*, 36:789–814, 2000.

[7] Jan Swevers, Farid Al-Bender, Chris G. Ganseman, and Tutuko Prajogo. An integrated friction model structure with improved presliding behavior for accurate friction compensation. *IEEE Trans. Autom. Control*, 45(4):675–686, 2000.

[8] X. Xin and M. Kaneda. Analysis of the energy-based control for swinging-up two pendulums. *IEEE Transactions on Automatic Control*, 50(5):679–684, 2005.

[9] W. Zhong and H. Röck. Energy and passivity based control of the double inverted pendulum on a cart. In *Proceedings of the 2001 IEEE International Conference on Control Applications*, pages 896–901, Mexico City, September 2001.