
Article:
Gonzalez, O. and Rossiter, J. orcid.org/0000-0002-1336-0633 (2020) A fast hybrid dual mode NMPC for a parallel double inverted pendulum with experimental validation. IET Control Theory and Applications, 14 (16). pp. 2329-2338. ISSN 1751-8644

https://doi.org/10.1049/iet-cta.2020.0130

IET Research Journals

Submission Template for IET Research Journal Papers

IET Journals
The Institution of Engineering and Technology

A Fast Hybrid Dual Mode NMPC for a Parallel Double Inverted Pendulum with Experimental Validation

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Oscar Gonzalez[1]*, *Anthony Rossiter*[1]

[1] *ACSE, The University of Sheffield, Sheffield, UK*
* *E-mail: ojgonzalezvillarreal1@sheffield.ac.uk*

Abstract: This paper presents a novel Fast Nonlinear Model Predictive Control approach for a parallel double inverted pendulum. The approach uses dual mode closed-loop predictions to obtain numerically robust optimal solutions. Moreover, it uses the Real-Time Iteration Scheme to reduce the computational burden and achieve real-time performance. Furthermore, two main modifications are proposed which significantly improve the performance of the RTI Scheme in the presence of large disturbances, namely; additional energy-based costs, and a hybrid switching scheme. Finally, the approach is combined with an Online System Identification scheme to address parameter uncertainty, and with an Extended Kalman Filter for state-estimation. The resulting performance is validated through both simulations and experimental results.

1 Introduction

Nonlinear Model Predictive Control (NMPC) is an advanced optimal control strategy able to handle complex and constrained nonlinear dynamic systems [1]. Although NMPC has a long history of development, its deployment had been restricted to the process industry where relatively slow processes allowed time to compute the required control algorithms [2]. However, recent progress in computing performance has enabled its application in many systems through the use of efficient solutions [1, 3, 4]. One of the most successful and popular approaches for fast NMPC is the Real-Time Iteration (RTI) Scheme, originally developed in [5]. Its efficiency is based on the fact that NMPC is required to succesively solve Optimal Control Problems (OCP) which are closely linked to each other [1]. An excellent tutorial-like paper detailing the main differences between the RTI NMPC and standard NMPC is given in [1]. Moreover, the efficiency of the overall approach depends largely on how the algorithms are programmed, as well as the platforms in which they are deployed, e.g. using embedded hardware such as field-programmable gate-arrays (FPGA) [6]. To address this, several toolkits containing efficient autogeneration routines are available such as the ACADO toolkit [3], VIATOC and CasADi [7], to name a few. Furthermore, the underlying optimisation may be solved using simultaneous or sequential approaches which lead to sparse or condensed OCPs. Authors from [4] concluded condensing based approaches are faster for small to medium systems, whereas sparse solutions give better overall performance for large scale optimisations and deal better with unstable systems [8]. Finally, direct methods are commonly used to discretise the problem, typically by using multiple or single-shooting discretisations [8, 9].

The (double) inverted pendulum is a complex multivariable nonlinear system that presents many challenges such as input-output constraints as well as underactuated, unstable and non-minimum phase dynamics [10]. For this reason, it has been used extensively for benchmarking of NMPC, though mostly for simulation works such as [1, 3, 7], and similar systems such as cranes studied in [6]. Nonetheless, experimental contributions have been achieved in [10–12], and furthermore discussed within. In [10], a triple pendulum swing up was achieved by using a two-degrees of freedom control structure which used offline optimisation to compute a feedforward trajectory, and a feedback controller to stabilise the system along it. In [11], a fast NMPC scheme for a twin parallel pendulum was developed which used a control-parameterisation where the decision variable was able to take only 3 possible values. Finally, authors from [12] presented a NMPC for a single inverted pendulum.

In this paper we propose a novel condensed single-shooting Dual Mode NMPC based on the RTI Scheme for a parallel double inverted pendulum, which differs from all previous works such as [1, 3, 7, 10–12] for the general inverted pendulum problem. The proposed approach cancels the unstable dynamics of the inverted pendulum through the use of closed-loop predictions [13] leading to numerically robust solutions when compared to the standard NMPC. Moreover, we propose two main modifications to improve the performance of the RTI Scheme in the presence of large disturbances, namely; additional energy related costs, and a hybrid switching scheme. Finally, the entire scheme is combined with an Online System Identification (OSI) algorithm based on Recursive Least Squares (RLS) and Delta Modeling approaches, to address parameter uncertainty. The whole approach is validated through both simulations and experimental results. The benefits when compared to [10, 11] were faster performance, and the use of online optimisations, thus allowing large disturbances and model updates. A video of the resulting performance can be found at (https://youtu.be/7E-SXi3YKQo), and the data of the experiments is available in [14] along with a C++ code implementing the approach using the EIGEN library.

The paper is organized as follows: Section 2 introduces the mathematical models of the pendulum along with a discretisation scheme. Section 3 presents a detailed derivation of the proposed dual-mode fast NMPC approach based on the RTI Scheme along with the two aforementioned modifications required to improve the RTI performance, which overall represent the main contribution of the paper. Sections 4 and 5 present Extended Kalman Filter (EKF) and OSI frameworks used for this work without a detailed derivation as both are well known in the literature. Section 6 presents the details of the experimental setup and discusses the experimental results of the proposed approach. Finally, section 7 presents conclusions, summarizes the contribution of the paper and describes future work.

2 Mathematical Modeling

In this section we present the equations of motion for the parallel double inverted pendulum depicted in figure 1 based on the assumption that the pendulums' will have negligible effect on the cart. Moreover, a discretization of the system is presented based on a backward-forward euler scheme which will be used for the NMPC, EKF and OSI, presented in the following sections.

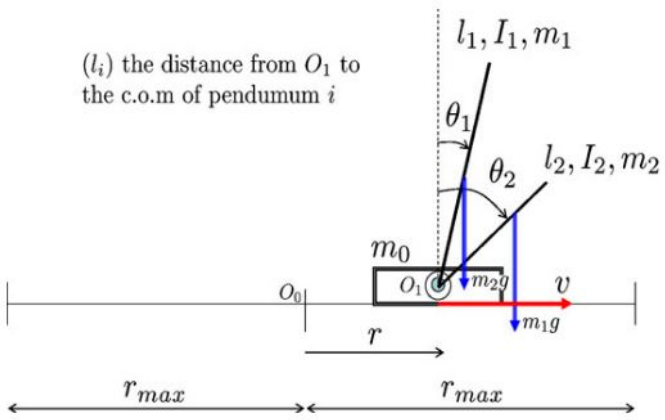


Fig. 1: Diagram from [11]

2.1 Equations of Motion

The equations of motion for the double inverted pendulum can be derived by using Lagrange formalism [10]. As this is well known, this paper uses the model from [11] with additional friction terms, given by (1a) and (1b).

$$\ddot{p} = f_m \dot{p} + ku \tag{1a}$$

$$\ddot{\theta}_i = a_i \dot{\theta}_i + b_i \sin\theta_i + c_i \cos\theta_i \, (f_m \dot{p} + ku) \tag{1b}$$

$$\forall i = [1, 2]$$

where p is the position of the car; u is an input signal to the system which in this case is a Pulse Width Modulated (PWM) signal for the motor driver; k is a constant that relate the PWM with the force and mass of the system; θ_i are the angles of the pendulum; f_m and a_i are viscous friction constants; $b_i = (m_i l_i g)/(m_i l_i^2 + I_i)$ are the pendulums' gravity related terms; $c_i = b_i/g$ are the acceleration-torque related constants; g is the gravity constant; and m_i, l_i, I_i, are the mass, length and moment of inertia of each pendulum, respectively. In this paper the relevant coefficients of the model will be found by the online system identification algorithm presented in section 5. Note that the sign of certain coefficients might be subject to the specific experimental setup depending on orientation, e.g. positive cart motion to the left or positive angle rotation CCW, however, all the viscous friction constants (f_m and a_i) must always be negative. Moreover, it should be noted that thought counter-intuitive, the lengths of the arms should be different to achieve better controllability of the system [11], particularly in the presence of noise which causes a significantly increasing amount of input shattering as the lengths become closer. This was validated through simulations to select appropriate length differences for out particular system.

This model is valid for our particular experimental setup given two conditions are true; both pendulum's masses are much lower than the cart, and the motor driver used has a regenerative breaking feature which further cancels out any possible uncontrolled movement of the cart. In the case where the cart motion is indeed affected by the pendulum's motion, a subordinate controller can be developed to cancel this effects as in [10], or the full nonlinear model can be included in the general NMPC framework as it has been shown in [3].

2.2 Discretization

This paper uses a "direct approach" which requires to "first discretize, then optimise" [6]. Thus, we now look to discretize the equations of motion (1a) and (1b) which will allow us to simulate and linearize the system for both NMPC and EKF frameworks. This can typically be done using some form of integration method such as explicit Euler method or explicit Runge Kutta methods [1].

For this work, a forward Euler method was considered at first, following similar works as in [10, 12], however, based on the observation that only position and angles are measured by the system, this scheme was modified to a forward-backward euler scheme including an extra previous input u_{k-1}, thus augmenting the state to obtain a Non-Minimal State Space (NMSS) [15]. This was motivated by observing that the position dynamics (1a) clearly represent a linear second order model which is known to have an exact ZOH discretization of the form:

$$p_{k+1} = a_1 p_k - a_2 p_{k-1} + b_1 u_k + b_2 u_{k-1} \tag{2}$$

which considering the backward euler approximations:

$$\dot{p}_k = \frac{y_k - y_{k-1}}{T_s} \quad \ddot{p}_{k+1} = \frac{\dot{y}_{k+1} - \dot{y}_k}{T_s} \tag{3}$$

results in the following position acceleration model:

$$\ddot{p}_{k+1} = f_m \dot{p}_k + k_1 u_k + k_2 u_{k-1} \tag{4}$$

where f_m, k_1, and k_2 are some equivalent coefficients after combining equations (1a), (2) and (3).

Similarly, considering the backward euler approximations:

$$\dot{\theta}_{ik} = \frac{\theta_{i_k} - \theta_{i_{k-1}}}{T_s} \quad \ddot{\theta}_{ik+1} = \frac{\dot{\theta}_{ik+1} - \dot{\theta}_{ik}}{T_s} \tag{5}$$

and giving a forward-euler step in equation 1b combined with the position acceleration model (eq. 4) results in:

$$\ddot{\theta}_{i_{k+1}} = a_i \dot{\theta}_{i_k} + b_i \sin\theta_{i_k} + c_i \cos\theta_{i_k} \, (f_m \dot{p}_k + k_1 u_k + k_2 u_{k-1}) \tag{6}$$

$$\forall i = [1, 2]$$

Combining equations (3-6), and considering the state $x_k = [p, \theta_1, \theta_2, \dot{p}, \dot{\theta}_1, \dot{\theta}_2, u_{k-1}]_k$, the simulation step is then given as,

$$x_{k+1} = x_k + T_s f(x_k, u_k) \tag{7a}$$

$$f(x_k, u_k) = \left[f_{up}^T, f_{down}^T, (u_k - u_{k-1})/T_s\right]^T \tag{7b}$$

$$f_{up} = \begin{bmatrix} \dot{p}_k + T_s f_1 \\ \dot{\theta}_{1_k} + T_s f_2 \\ \dot{\theta}_{2_k} + T_s f_3 \end{bmatrix} \quad f_{down} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \tag{7c}$$

$$f_1 = f_m \dot{p}_k + k_1 u_k + k_2 u_{k-1} \tag{7d}$$

$$f_2 = a_1 \dot{\theta}_{1_k} + b_1 \sin\theta_{1_k} + c_1 \cos\theta_{1_k} f_1 \tag{7e}$$

$$f_3 = a_2 \dot{\theta}_{2_k} + b_2 \sin\theta_{2_k} + c_2 \cos\theta_{2_k} f_1 \tag{7f}$$

where T_s is the sampling time. It is noted that the last term of function (7b) was only used to represent the propagation of the input $x_{7_{k+1}} = u_k = u_{k-1} + T_s(u_k - u_{k-1})/T_s$ and doesn't represent, in any way, a "derivative of the input" $\dot{u}_k = (u_k - u_{k-1})/T_s$ which would lead to a completely different integration scheme if more intermediate steps were computed [1].

This discretisation differs slightly from the standard forward euler method in the sense that the latter would compute a forward-euler step on the higher derivative states (p, θ_1, θ_2). However, as stated previously, given only position and angle's measurement were available, the backward euler approximations (3 and 5) were used instead to use the latest information of the system. Moreover, it would only take into account u_k for simulation purposes. Nonetheless, this modification was observed to produce much better predictions in an offline analysis of the system identification process for both, position and angle dynamics, and in fact, increasing the number of previous input terms u_{k-j} was be able to improve them even further, possibly given that it accommodates some unmodeled higher order motor dynamics which are known to be at least $2nd$ order in the angular velocity; $3rd$ order in the angular position; which can be accounted for using convolution/FIR models. However, the system was observed to get good performance whilst only including 2 previous input terms, u_k and u_{k-1}.

3 Nonlinear Model Predictive Control

In this section, a Dual Mode NMPC scheme based on the closed-loop paradigm [16] is proposed to cancel the open-loop unstable dynamics of the double inverted pendulum for numerical robustness of the optimisation. Furthermore, to enable real-time performance, the optimisations are performed within the RTI Scheme which allows the constrained optimisation to be solved within the microsecond range [6]. Finally, a modification to the standard cost used for inverted pendulum control is proposed based on energy considerations along with a hybrid switching scheme which overall significantly improves the convergence of the algorithm, particularly for large disturbances, a situation where the assumptions for local-asymptotic closed-loop stability of the RTI Scheme are lost. Simulations are presented along the section to illustrate the significance of the proposed approach.

3.1 Stable Predictions and Optimization

In this paper we are looking to optimise the system performance along a given prediction horizon N_p by minimizing the cost function (8), defined as,

$$J = \frac{1}{2}(Y_r - \hat{Y})^T Q(Y_r - \hat{Y}) + \frac{1}{2}\hat{U}^T R\hat{U} \quad s.t \tag{8a}$$

$$\hat{x}_k = x_0 \tag{8b}$$

$$\hat{x}_{k+i|k} = f(\hat{x}_{k+i-1|k}, \hat{u}_{k+i-1|k}) \tag{8c}$$

$$\hat{y}_{k+i|k} = g(\hat{x}_{k+i|k}) \tag{8d}$$

$$U_{min} \leq \hat{U} \leq U_{max} \tag{8e}$$

$$Y_{min} \leq \hat{Y} \leq Y_{max} \tag{8f}$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ and $y_k \in \mathbb{R}^{n_y}$ are the states, inputs and outputs of the system, respectively; the notation "$k+1|k$" reads "predicted value at $k+1$ considered at sample time k", and will only be used in full when needed for clarity; $Q > 0 \in \mathbb{R}^{N_p n_y \times N_p n_y}$ and $R > 0 \in \mathbb{R}^{N_p n_u \times N_p n_u}$ are positive-definite matrices for penalising output-errors and inputs, respectively, typically selected as $Q = blkdiag([q_{k+1}, q_{k+2}, \ldots, q_{k+N_p}])$ where q_{k+N_p} is typically referred to as the terminal weight, and $R = r_u I^{N_p n_u \times N_p n_u}$; $Y_r = [r_{k+1}^T, r_{k+2}^T, \ldots, r_{k+N_p}^T]^T \in \mathbb{R}^{N_p n_y}$, $\hat{Y} = [\hat{y}_{k+1}^T, \hat{y}_{k+2}^T, \ldots, \hat{y}_{k+N_p}^T]^T \in \mathbb{R}^{N_p n_y}$, $\hat{U} = [\hat{u}_k^T, \hat{u}_{k+1}^T, \ldots, \hat{u}_{k+N_p-1}^T]^T \in \mathbb{R}^{N_p n_u}$ are references, outputs and inputs column-vectors, respectively; (8b) is the initial condition; (8c) are the state dynamics; (8d) is the function that relates the output with the states; (8e) are the input constraints; and (8f) are the output constraints. For our particular system, the outputs are typically selected as in [1, 10, 12] as $y_{k+i} = [\dot{p}, \dot{\theta}_1 \dot{\theta}_2, p, \theta_1, \theta_2]_{k+i}$ ($n_y = 6$), and the references are selected as $r_{k+i} = [0, 0, 0, p_r, \theta_{1_r}, \theta_{2_r}]_{k+i}$.

Remark 1. *Stability of the resulting closed-loop system can be typically ensured by having long horizons with zero-terminal constraints and/or proper terminal weights [6].*

Cost function (8) for system (7a) represents a non-convex Nonlinear Programming (NLP) problem which is difficult to solve. Sequential Quadratric Programing (SQP) is a popular alternative where the cost is linearised at a given point to formulate a linearized Quadratic Program (QP) and find an optimal search direction, typically based on the Newton method, that eventually drives the solution to the local optimal. Notice in the case of predictive control, future state trajectories $\hat{x}_{k+i}$ required for the linearisation are only defined after a given input trajectory $\hat{u}_{k+i-1}$ has been applied through the state dynamics (8c) with the initial condition (8b). A workaround to this are shooting methods which use an "initially guessed" **nominal input trajectory**, $\bar{U} = [\bar{u}_k^T, \bar{u}_{k+1}^T, \ldots, \bar{u}_{k+N_p-1}^T]^T \in \mathbb{R}^{N_p n_u}$ to generate **nominal state and output trajectories**, $\bar{X} = [\bar{x}_{k+1}^T, \bar{x}_{k+2}^T, \ldots, \bar{x}_{k+N_p}^T]^T \in \mathbb{R}^{N_p n_x}$ and $\bar{Y} = [\bar{y}_{k+1}^T, \bar{y}_{k+2}^T, \ldots, \bar{y}_{k+N_p}^T]^T \in \mathbb{R}^{N_p n_y}$, respectively, by simulating the system with $\bar{U}$ using initial condition (8b) and state dynamics (8c).

The standard NMPC single-shooting approach would linearise the system along this resulting trajectories with a first order Taylor Series on the state dynamics, the output-state function, and the input. However, given the open-loop unstable dynamics of the inverted pendulum in its upright equilibrium, closed-loop dual mode prediction models were motivated [16]. The linearised model at a given time step k is then given by,

$$\hat{x}_{k+1} = \bar{x}_{k+1} + \delta\hat{x}_{k+1} = \bar{x}_{k+1} + A_k\delta\hat{x}_k + B_k\delta\hat{u}_k \tag{9a}$$

$$\hat{y}_k = \bar{y}_k + \delta\hat{y}_k = \bar{y}_k + C_k\delta x_k \tag{9b}$$

$$\hat{u}_k = \bar{u}_k + \delta u_k = \bar{u}_k - K_k\delta x_k + \delta\hat{c}_k \tag{9c}$$

where K_k is a stabilizing gain obtained from solving the Time-Varying Discrete Algebraic Ricatti Equation (DARE) backwards in time along the nominal state trajectory using the same Q and R weights as in [10], and:

$$A_k = \left.\frac{\partial f(\hat{x}_k, \hat{u}_k)}{\partial \hat{x}_k}\right|_{\substack{\hat{x}_k=\bar{x}_k \\ \hat{u}_k=\bar{u}_k}} \quad B_k = \left.\frac{\partial f(\hat{x}_k, \hat{u}_k)}{\partial \hat{u}_k}\right|_{\substack{\hat{x}_k=\bar{x}_k \\ \hat{u}_k=\bar{u}_k}} \tag{10a}$$

$$C_k = \left.\frac{\partial g(\hat{x}_k)}{\partial \hat{x}_k}\right|_{\hat{x}_k=\bar{x}_k} \tag{10b}$$

Notice this dual-mode prediction model differs from common dual-mode schemes in the sense that standard methods would select a gain K, typically a constant one, which stabilizes the system to the origin when the system is in the terminal region [17]. In contrast, our approach uses dual-mode closed loop models to stabilize the predictions and achieve better numerical performance as in [13]. This is because we aim to deal with the situation when the system is not in the terminal region, and finding a stabilizing gain that does the swing up whilst satisfying the constraints is, in general, not a trivial task for our system.

By substituting $\delta\hat{u}_k = -K_k\delta\hat{x}_k + \delta\hat{c}_k$ from (9c) in (9a), a stable state deviation model can be obtained as,

$$\Phi_k = A_k - B_k K_k \tag{11a}$$

$$\delta\hat{x}_{k+1} = \Phi_k\delta\hat{x}_k + B_k\delta\hat{c}_k \tag{11b}$$

Propagating model (11b) N_p steps forward starting from an initial state mismatch δx_0, leads to the following predictions matrices for all future inputs and outputs, $\hat{U}$ and $\hat{Y}$, are condensely represented by,

$$\hat{Y} = \bar{Y} + \delta\hat{Y} = \bar{Y} + G\delta x_0 + H\delta\hat{C} \tag{12a}$$

$$\hat{U} = \bar{U} + \delta\hat{U} = \bar{U} + D\delta x_0 + F\delta\hat{C} \tag{12b}$$

where $\delta x_0 = x_0 - \bar{x}_0$ is the initial condition mismatch which forms a part of the RTI Scheme, $\delta\hat{C}$ are now the decision variables,

$$G = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{N_p} \end{bmatrix} \qquad H = \begin{bmatrix} h_{1,1} & 0 & \cdots & 0 \\ h_{2,1} & h_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{N_p,1} & h_{N_p,2} & \ldots & h_{N_p,N_p} \end{bmatrix} \tag{13a}$$

$$D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N_p} \end{bmatrix} \qquad F = \begin{bmatrix} I^{n_u} & 0 & \cdots & 0 \\ f_{2,1} & I^{n_u} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ f_{N_p,1} & f_{N_p,2} & \ldots & I^{n_u} \end{bmatrix} \tag{13b}$$

where I^{n_u} is a $n_u \times n_u$ identity matrix, $G \in \mathbb{R}^{N_p n_y \times n_x}$, $H \in \mathbb{R}^{N_p n_y \times N_p n_u}$, $D \in \mathbb{R}^{N_p n_u \times n_x}$, $F \in \mathbb{R}^{N_p n_u \times N_p n_u}$, and

$$g_i = C_i \overset{\curvearrowleft}{\prod_{k=0}^{i-1}} \Phi_k \tag{14a}$$

$$h_{i,j} = \begin{cases} C_i B_{j-1}, & i = j \\ C_i \left[\overset{\curvearrowleft}{\prod_{k=j}^{i-1}} \Phi_k \right] B_{j-1}, & i > j \end{cases} \tag{14b}$$

$$d_i = \begin{cases} -K_{i-1}, & i = 1 \\ -K_{i-1} \overset{\curvearrowleft}{\prod_{k=0}^{i-2}} \Phi_k, & i > 1 \end{cases} \tag{14c}$$

$$f_{i,j} = \begin{cases} -K_{i-1} B_{j-1}, & i = j+1 \\ -K_{i-1} \left[\overset{\curvearrowleft}{\prod_{k=j}^{i-2}} \Phi_k \right] B_{j-1}, & i > j+1 \end{cases} \tag{14d}$$

$$\forall i = [1, N_p] \quad \forall j = [1, N_p]$$

Substituting the stable linearized prediction models (12) in (8) and rearranging in terms of the decision variable $\delta\hat{C}$ to obtain the standard QP format gives;

$$J = \frac{1}{2} \delta\hat{C}^T E \delta\hat{C} + \delta\hat{C}^T f + const \quad s.t. \tag{15a}$$

$$E = H^T Q H + F^T R F \tag{15b}$$

$$f = -\left[H^T Q (Y_r - \bar{Y} - G\delta x_0) - F^T R (\bar{U} + D\delta x_0) \right] \tag{15c}$$

$$M \delta\hat{C} \leq \gamma \tag{15d}$$

$$M = \begin{bmatrix} F \\ -F \\ H \\ -H \end{bmatrix} \quad \gamma = \begin{bmatrix} U_{max} - \bar{U} - D\delta x_0 \\ -(U_{min} - \bar{U} - D\delta x_0) \\ Y_{max} - \bar{Y} - G\delta x_0 \\ -(Y_{min} - \bar{Y} - G\delta x_0) \end{bmatrix} \tag{15e}$$

with E known as the Hessian, f typically referred as the linear term, and M and γ are the constraint matrix and vector, respectively. Notice equations (8b), (8c) and (8d) are implicit in the linearisation of $\hat{U}$ and $\hat{Y}$. Moreover, note the not all the outputs may be required to be constrained which can be done by selecting (or computing) only the relevant rows of M and γ. In our particular system, only the position outputs will be constrained.

By derivating (15a) w.r.t. the decision variable $\delta\hat{C}$ and equating to zero ($\frac{\partial J}{\partial \hat{C}} = 0$), the well known unconstrained solution can be found to be $\delta\hat{C} = -E^{-1} f$. For constrained solutions, any QP solver can be used to compute the optimal deviation $\delta\hat{C}$ after having defined E, f, M, γ. For our experiments, an efficient version of the active-set based primal-dual Hildreth's QP found in [15] was used given its simplicity and its ability to be hot-started which is required for achieving fast implementation of the overall scheme. After solving the optimisation, the corrected input $\hat{U}$ can then be recovered by (12b). Only the first input is applied to the system and the process is repeated which is the well known "receding horizon" strategy [11].

3.1.1 Stability and Numerical Robustness: Because this model produces the exact same predictions for a given $\delta\hat{U} = D\delta x_0 + F\delta\hat{C}$ than using the standard model without the stable predictions, and because $\delta\hat{U}$ can always be calculated exactly through the inversion of F which is always invertible, the solution for $\hat{U}$ is exactly the same as the one given by the standard approach using unstable predictions, and therefore presents the exact same stability and convergence properties of the standard single-shooting approach. The benefit of it is that the predictions matrix H is now stable w.r.t the decision variable $\delta\hat{C}$ which leads to a numerically robust Hessian inversion required by the optimisation. This allows the prediction horizon to be increased as much as required without sacrificing numerical robustness. For reference, in our particular system, the condition number (c.n.) of the Hessian at the upward equilibrium when using the proposed approach was around the unit magnitude ($E_{c.n.} = 2.5 \times 10^0$); the c.n. without using the approach was ($E_{c.n.} = 5 \times 10^6$), 6 orders of magnitude larger which indeed shows severe numerical issues given it comes close to singular as the c.n. increases.

Finally, for the interest of the reader, the standard NMPC condensed single-shooting solution can be recovered by enforcing $K_{k+i} = \mathbb{O}, \quad \forall i = [0, N_p - 1]$ which would then use the unstable predictions.

3.2 Real Time Iterations

Ideally, the fully converged NMPC would relinearize cost function (8) until no deviation is required $\delta\hat{C} = \mathbb{O}$ [1]. However, this is not computationally tractable in practice given one must give a solution at every time-step within the available time and avoid solving a problem that is only "getting older" [9]. A very successfull and popular approach to address this is to use the Real-Time Iteration (RTI) Scheme which exploits the fact that NMPC is required to successively solve optimisations which are closely related to each other. The method benefits from the fast contraction rate of Newton-type optimisations and achieves convergence of the solution "on the fly", using the current predictions and measurements rather than through offline reference trajectories [1]. The overall RTI Scheme is based on three well defined strategies.

3.2.1 Initial Value Embedding (IVE): The input trajectory obtained in the previous sampling $\hat{U}_k = [\hat{u}_{k|k}^T, \hat{u}_{k+1|k}^T, \cdots, \hat{u}_{k+N_p-1|k}^T]^T$ is used in a shifted version to hotstart the solution in the next sampling time, typically by duplicating the last value $\hat{U}_{k+1} = [\hat{u}_{k+1|k}^T, \hat{u}_{k+2|k}^T, \cdots, \hat{u}_{k+N_p-1|k}^T, \hat{u}_{k+N_p-1|k}^T]^T$. Moreover, in the case of active-set based QP, the lagrange multipliers λ related to the constraints of the optimisation can also be used for hotstarting the QP in a shifted version.

3.2.2 Single SQP: Only a single linearisation of the QP is performed given the solution is hotstarted from the previous solution which is expected to be close. In the case where the previous solution was indeed close to the optimal solution and no significant disturbances have entered the system, this approach can be proved to have nominal local-asymptotic closed-loop stability [18]. In general, the solution is not given exactly but as an approximation that decreases the sub-optimality of cost J at each iteration. Moreover, one must be satisfied with finding a local minimum, and the solution will be subject to small approximation errors given only one re-linearisation is done.

3.2.3 Computation Separation: In order to avoid the delay related to the computations required by the optimisation, we divide them into a preparation and a feedback phase. A timing diagram that illustrates this is given in [1].

1. Preparation Phase: Uses a predicted nominal state $\hat{\bar{x}}_{k+1|k}$ obtained with the current input and states to compute the matrices E, M and vectors f, γ required by the optimisation assuming $\delta x_0 = 0$.
2. Feedback Phase: As soon as the state x_{k+1} becomes available, the deviation $\delta x_0 = x_{k+1} - \hat{\bar{x}}_{k+1|k}$ is used to complete the calculation of f and γ and the optimal correction $\delta\hat{U}$ to the current trajectory $\hat{U}$.

A slight modification was implemented in our test where, the deviation δx_0 was only applied to the linear term f and the QP was iterated assuming $\delta x_0 = 0$ to find the active set λ before the state x_{k+1} was available. By doing so, the solution is now given by (16), which allows the preparation phase part ($\hat{U}_{pre-computed}$) to be pre-computed prior to the arrival of the measurement, and only the feedback correction $\hat{U} = \bar{U} - \hat{U}_{pre-computed} - K_x \delta x_0$ is required to be computed when the state becomes available, with

$$K_x = FE^{-1}(H^TQG + F^TRD) - D.$$

$$\hat{U} = \bar{U} + \overbrace{D\delta x_0}^{\text{Feedback Phase}} - FE^{-1}\Bigg[\underbrace{\overbrace{-(H^TQ(Y_r - \bar{Y}) - R\bar{U})}^{\text{Unconstrained}} \overbrace{+M^T\lambda}^{\text{Constrained}}}_{\text{Preparation Phase } (\hat{U}_{pre-computed})} \underbrace{+(H^TQG + F^TRD)\delta x_0}_{\text{Feedback Phase}}\Bigg] \tag{16}$$

This modification completely removes the time-delay related to the iterations of the QP required to be done by the standard RTI in the feedback phase. By including this term δx_0 in the linear term f, the stability characteristics of the overall RTI Scheme are preserved, however, given the term is ignored for the calculation of the constrained correction $E^{-1}M^T\lambda$, the system may present small output constraint violations depending on how large is the deviation from the predicted state at a given time. However, this modification is justified considering that feasibility guarantees for output constraints in the presence of disturbances are, in general, difficult to achieve without using robust approaches or slack variables (soft-constriants) which are outside the scope of this paper. Nonetheless, the system presented excellent performance in constraints satisfaction as it will be seen in the experimental results presented in section 6.

3.3 Improving RTI NMPC Performance

A mayor issue with the RTI Scheme is that the solution might give very poor performance whenever an abrupt change is made, e.g., when there is an abrupt change in the reference of the system [1], or a large fault or disturbance enters the system, which may lead to leaving the region of contraction of the Gauss-Newton method and in some cases may even lead to instability of the system [6]. In these cases, the previous solution will not be close to the optimal, and therefore, the method would need to quickly find a suitable correction from the previous solution. This issue may be addressed by adding suitable end weights and other regularity conditions [6]. However, to address this issue, this paper takes a different approach with two main modifications to the standard approach of NMPC of an inverted pendulum such as [1, 3, 7, 12], namely: an additional energy based cost; and a hybrid switching scheme.

3.3.1 Energy Based Costs: Motivated by the fact that a common strategy for the swing up of the pendulum are Energy based control laws [11], along with the fact that standard cost terms defined for inverted pendulum NMPC, e.g. [1, 12], do not actually capture the requirement of "swinging-up" but rather a more restrictive cost requiring the optimisation to drive the angles to a desired reference without considering other upward equilibrium points, the outputs and references used for the cost function (8) were modified to include two extra terms related to the potential energy of both pendulums $E_{\theta_i} = \cos\theta_i$ as;

$$y_k = \left[\dot{p}, \dot{\theta}_1, \dot{\theta}_2, p, \theta_1, \theta_2, \cos\theta_1, \cos\theta_2\right]_k \tag{17a}$$

$$r_k = \left[0, 0, 0, p_r, \theta_{1_r}, \theta_{2_r}, \cos\theta_{1_r}, \cos\theta_{2_r}\right]_k \tag{17b}$$

Remark 2. *With this modification, the optimisation now has $n_y = 8$ outputs.*

Some of the relevant properties of this added term are:

1. Boundedness: The error $e_k = \cos\theta_{1_r} - \cos\theta_1$ is always bounded at $e_k = [-2, 0]$ for the upright position, and $e_k = [0, 2]$ for the downside position. This in general would make the additional term of the linear term (f) bounded.
2. Singularity: The derivative w.r.t. the added term $(C_i = -\sin\theta_i)$ required by (12a) has a singularity at any $N\pi$ multiple given the sensitivity matrix is zero. Thus, if the system is at a steady condition, e.g. all other errors zero, the optimisation would have no sensitivity on it, therefore, not reacting or causing any movement. Although the system can be confined inside an incorrect singularity, if the system is started at any other sufficiently non-singular point, the optimisation will eventually drive the solution to the desired singularity.

By penalising the energy term much higher than the angles directly, the optimisation is more relaxed, essentially aiming to drive the potential energy of the pendulum to the desired state $E_{\theta_i} = \cos\theta_i \rightarrow \cos\theta_{i_r}$ whilst accepting swinging up in either directions. This is because if at a given time the system cannot swing the pendulums up in a given direction, the optimisation would naturally select the other direction which is not the case when the standard costs are used, and the solution was observed to be severely affected when the system reached this infeasibility condition. After a series of simulations it was concluded that imposing higher penalties on this added terms instead of the angles directly, resulted in much better convergence properties than when using the standard cost, which in turn resulted in a larger region of contraction of the Gauss-Newton method. Moreover, notice the stability of the resulting scheme can still be guaranteed by imposing heavy terminal weights in the additional terms to emulate zero-terminal constraints, even though the sensitivity at that condition dissipates, which in turn leads to having the original problem once the system has reached the terminal region.

To visualize the benefits of this approach, a comparison simulation is given in figure 2 where the predicted and closed-loop trajectories are plotted, with and without focusing on the added energy term. For clarity, the predicted responses that presented the erratic behavior are signaled. As it can be seen from (figure 2b), the optimisation penalizing only the angles presented mayor erratic behavior in the input at times $0.5 < t < 1$,and significant differences between predicted and closed-loop responses leading to ill-posed optimisation [16]. In contrast, the optimisation that focused effort on the added energy-cost (figure 2a) presented smooth and much better overall closed-loop performance.

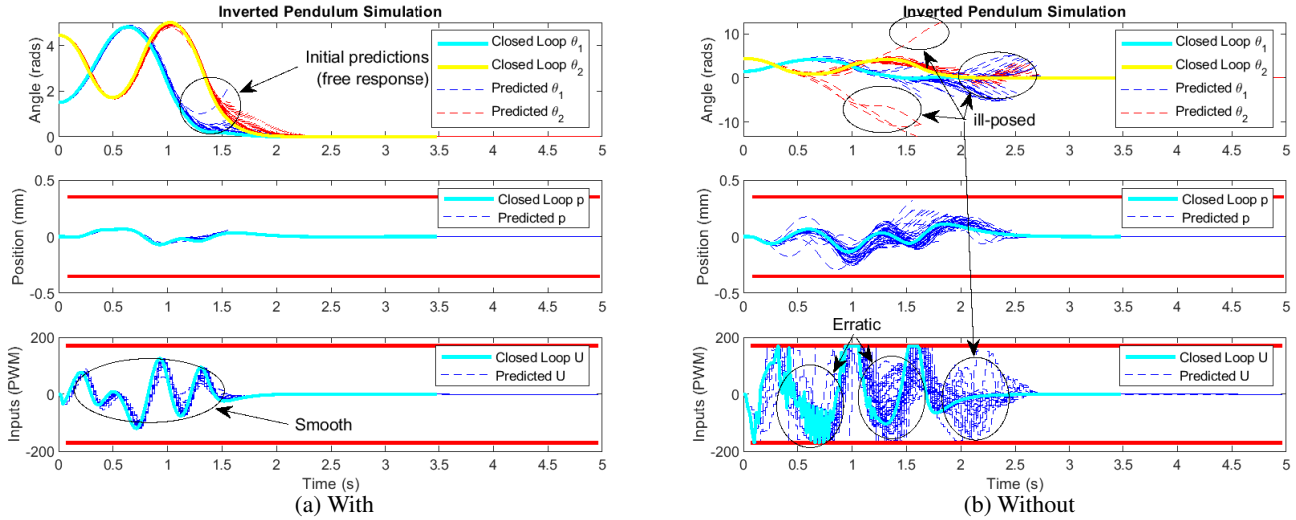
Although this approach might not be immediately generalisable for other control systems and applications, it is very common to find trigonometric terms in robotics systems and mechatronic applications that arise from rotation matrices. This naturally brings the question of whether it is better to target a desired potential energy or an angle directly when dealing with multi-link robots. Indeed, it is well known that the understanding of the inverted pendulum dynamics helped with the development of many robotic applications nowadays, thus its generalisation to multi-link robotic problems, eg. triple inverted pendulum in series [10], could lead to a significant improvement in performance in a broader spectrum of applications, particularly when using the RTI Scheme.

3.3.2 Hybrid Switching Scheme: As discussed previously, penalizing the energy-related terms lead to smoother responses. However, because of aforementioned singularity problem, if only the energy terms are penalized instead of the angles directly, the optimisation would have no sensitivity to potential energy errors at the equilibrium, and would only be sensitive to angular velocity errors. Moreover, if a sufficiently small penalty was imposed on the angles, the optimisation could converge to upright positions were the angle errors were essentially ignored.

To avoid this problems whilst preserving the smoothness of the added energy terms during the swing up phase, a hybrid approach was used where the optimisation would switch between different weightings depending, not only on the region in which the angles where, but also the time that they have been there.

The hybrid switching scheme is given by;

$$q_{\theta_i} = \begin{cases} 1, & t_{lin} < 2 \\ 10, & t_{lin} \geq 2 \end{cases} \tag{18}$$



(a) With (b) Without

Fig. 2: Example comparison of predicted (dashed-lines) and closed-loop (thick lines) responses with (2a) and without (2b) energy costs with parameteres defined in table (1); $T_s = 0.02\ (s)$; $N_p = 75$; $x_0 = [0, 0, 0, 0, 1.51, 4.45, 0]^T$; $\bar{U}_0 = \mathbb{O}$ (free-response); $Q_{with} = diag([1, 0.1, 0.1, 100, 1, 1, 10, 10])$, $Q_{without} = diag([1, 0.1, 0.1, 100, 10, 10, 0, 0])$, and $R = 0.001$.

where q_{θ_i} is the weight of the i_{th} pendulum angle error; and t_{lin} is the time that has elapsed since $\cos\theta_i > 0.9$, ie. the time the system has been in the "linear" zone.

Regarding the stability of this proposed hybrid scheme, it should be noted that both penalization terms of q_{θ_i} were stable for our particular system, and the only reason for this change was to preserve the smoothness of the system during the swing up phase. As the change was implemented when the system was already in the terminal region, the cost of both selected weights dissipated to zero within the available horizon, which in turn made the change between both weights stable. Essentially, the selection of this different terms changes the frequency response of the system to a more "rigid" or fast response for angle perturbations. Indeed, this approach could be used for fault-tolerant applications where the system momentarily has to undergo through a "softer/smoother" set of actions to bring the system back to its target before regaining a more "reactive" state.

4 State Estimation - Extended Kalman Filter

As only position and angle measurements were available in our system, a standard EKF was used for the purpose of state estimation. As this is a well known method in the literature, the details of this are omitted and only the relevant equations and steps are provided.

The EKF uses the following prediction-correction type framework:

1. Prediction Step: The state, output and covariance at time k are estimated based on the previously estimated state and covariance as,

$$\hat{x}_{k|k-1} = \hat{x}_{k-1|k-1} + T_s f(\hat{x}_{k-1|k-1}, u_{k-1}) \tag{19a}$$

$$P_{k|k-1} = A_{k-1} P_{k-1|k-1} A_{k-1}^T + Q_{EKF} \tag{19b}$$

where $Q_{EKF} > 0$ is the process noise; and $P_{k|k-1}$ is the covariance matrix

2. Correction Step: As soon as the outputs of the system become available, the correction step is then given by,

$$\hat{y}_{k|k-1} = g(\hat{x}_{k|k-1}) \tag{20a}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_{EKF}\left[(y_k - \hat{y}_{k|k-1})\right] \tag{20b}$$

$$P_{k|k} = (I - K_{EKF} C_k) P_{k|k-1} \tag{20c}$$

$$K_{EKF} = P_{k|k-1} C_k^T \left(C_k P_{k|k-1} C_k^T + R_{EKF}\right)^{-1} \tag{20d}$$

where $R_{EKF} > 0$ is the output noise covariance matrix, and $C_k = \frac{\partial g(\hat{x}_{k|k-1})}{\partial \hat{x}_{k|k-1}}$.

Remark 3. *As system 7a is a NMSS, the input-related state (x_7) must be included in the measurement for observability. The proof of this is out of the scope of this paper.*

5 Online System Identification

In this section, an OSI scheme based on RLS with forgetting factor combined with a Delta-Modeling approach [19] is presented. The latter was used for the purpose of learning/adapting the parameters of the discrete model (7a), particularly $f_m, k_1, k_2, a_i, b_i, c_i$. As this is a well known method in the literature, the details of this are omitted and only the relevant equations and steps are provided.

The RLS algorithm with forgetting factor (λ) is given by:

$$\tilde{z}_k = z_k - \Psi^T \Theta_{k|k-1} \tag{21a}$$

$$K_{RLS} = P_{k|k-1}\Psi\left(\lambda + \Psi^T P_{k|k-1}\Psi\right)^{-1} \tag{21b}$$

$$\Theta_{k|k} = \Theta_{k|k-1} - K_{RLS}\tilde{z}_k \tag{21c}$$

$$P_{k|k} = \lambda^{-1}\left(P_{k|k-1} - K_{RLS}\Psi^T P_{k|k-1}\right) \tag{21d}$$

where Ψ is known as the regressors vector, $\Theta_{k|k}$ is the parameters vector, $P_{k|k}$ is a covariance matrix of appropriate dimensions, and λ is the forgetting factor, typically selected as $0.98 < \lambda < 1$.

For our system, the definition of Ψ, $\Theta_{k|k}$ and z_k for both, position and pendulum dynamics, are given by;

1. Position Dynamics:

$$\Psi = [\dot{p}_{k-1}, u_{k-1}, u_{k-2}]^T \tag{22a}$$

$$z_k = \ddot{p}_k = \frac{p_k - 2p_{k-1} + p_{k-2}}{T_s^2} \tag{22b}$$

$$\Theta_{k|k} = [f_m, k_1, k_2]^T \tag{22c}$$

2.Pendulum Dynamics:

$$\Psi = [\dot{\theta}_{i_{k-1}}, \sin\theta_{i_{k-1}}, \cos\theta_{i_{k-1}} \ddot{p}_k]^T \tag{23a}$$

$$z_k = \ddot{\theta}_{i_k} = \frac{\theta_k - 2\theta_{k-1} + \theta_{k-2}}{T_s^2} \tag{23b}$$

$$\Theta_{k|k} = [a_i, b_i, c_i]^T \tag{23c}$$

which can be derived from the position and angle acceleration models (4 and 6), and also represent the use of the Delta Δ - modeling approach [19], which is known to numerically perform better than ARX models for system with fast sampling times. The coefficients were then extracted to be used in both NMPC and EKF frameworks presented previously.

On the other hand, several execution rules discussed in [20, 21] were implemented to shut-down the algorithm to protect it from periods of poor excitation which can lead to the rapid grow of covariance matrix $P_{k|k}$.

The implemented shut-down rules were:

1.The trace of the covariance $P_{k|k}$ was limited by

$$P_{k|k} = \frac{k_{lim}}{tr(P_{k|k})} P_{k|k} \quad if \quad tr(P_{k|k}) > k_{lim} \tag{24}$$

to prevent it from becoming ill-conditioned, where for our particular system $k_{lim} = 10$. Additionally, the limitation of the covariance trace allows for better control on the rate of convergence of the parameters.

2.The range of the parameters of both models were limited according to offline analysis, as well as based on the uncertainty of expected models coefficients. In particular, a threshold of $\pm 15\%$ was imposed for the coefficients b_i and c_i calculated from the expressions given in section 2.1, as well as forcing the negative sign of the friction terms f_m and a_i.
If the RLS algorithm moved any of the coefficients outside the available range, they were simply saturated.

3.The RLS algorithms were only run when the system was detected to be moving, in particular, when the angular velocity was greater than a threshold, and kept running for a maximum of 2 seconds after this conditions were satisfied to be able to capture “decaying" dynamics. Specifically, the thresholds used were:

$$||\dot{p}_k|| > 0.4 \left(\frac{m}{s}\right) \qquad ||\dot{\theta} i_k|| > 0.5 \left(\frac{rad}{s}\right) \tag{25}$$

This ensured that the system was properly excited.

4.The model used for the NMPC was only updated if the uncertainty of the coefficients, captured in the covariance matrix $P_{k|k}$ was lower then a threshold. For this system, the uncertainty was simply considered as the trace of the covariance matrix, although a more accurate distribution of the uncertainty could be extracted by using the so called Chi-squared (χ^2) distribution.

Although the rules for shut-down did improve the performance of the online system identification as a potential adaptive controller, a proper excitation signal is required for better model estimation such as PRBS or Frequency Sweep (chirp), as without there is no guarantee that the model estimation would be accurate or even stable, thus no stability guarantee of the combined methodology could be provided, which is generally known. The overall performance of this algorithm can be seen in figure 5 and will be discussed in the results section 6.1.

6 Experimental Results

The test bench used for the experiments is depicted in figure 3. The cart is driven by a brushed 24V DC motor via a toothed-belt and a toothed-pulley of 0.05 (m) diameter. The DC motor is driven by a

Fig. 3: Test Bench Photograph

Motor	Coeffs.	Pend 1	Coeffs.	Pend 2	Coeffs.
f_m	-4.67	a_1	-0.129	a_2	-0.107
k_1	0.0174	b_1	38.4	b_2	49.6
k_2	0.0477	c_1	3.95	c_2	5.11

Table 1 Identified Parameters for the Double Inverted Pendulum (3)

Cytron MD30C Motor Driver operated using sign-magnitude drive with a 8-bit resolution PWM at a frequency of 20 kHz via a Micro Controller Unit (MCU). Three incremental encoders are used to measure both pendulum angles and the DC motor rotation. The resolution of both pendulum encoders and of the motor are 4000 and 2040 counts per revolution, respectively, which are processed by the MCU, leading to angle and position resolutions of $9 \times 10^{-2^\circ}$ and 7.7×10^{-5} (m), respectively. The sampling time of the system is handled by the MCU and kept constant at $T_s = 20$ (ms). Every sampling time, encoders data is streamed via (UART) serial communication to a PC where the calculations related to proposed NMPC approach, OSI and EKF are performed. After the control action is calculated via the RTI Scheme, it is send back to the MCU which generates the motor signals. Due to network communication delays, the motor signal was always implemented exactly 5 ms after the encoders data was streamed to have a constant behavior at least. Figure 4 shows an control diagram detailing the interaction between the different components.

6.1 Online System Identification

To test the online system identification algorithm presented in section 5, the system was excited using a random input $30 < ||u|| < 60$, which reversed every time the system crossed a maximum limit of the position $||x|| > 0.15$ (m) in the current direction. All the parameters were started from completely unknown values $\Theta_0 = \mathbb{O}$ with a forgetting factor of $\lambda = 0.995$ and initial covariance matrices as $P_0 = 1000 I_{3\times 3}$. The resulting performance of the overall OSI algorithm can be seen in figure 5. As it can be seen, the system presented very fast convergence rates, giving settling times for all the parameters of $\tau_s < 2$ (s), indicating that the models are indeed well defined. The resulting parameters after 1 minute of excitation are gathered in table 1, and the input-output data is available in [14]. Notice the theoretical relationship $c_i = b_i/g$ stated in section 2 is very close to the one observed in the resulting parameters. Finally, although the system was only tested for online system identification, it could work using the available adaptation mechanism provided proper rules are used to avoid the periods of poor excitation, as discussed in [20].

6.2 Swing Up, Stabilization and Disturbance Rejection

Regarding the optimisation setup, the cart has a maximum range for the position of $-0.35 < x < 0.35$ and the PWM input was constrained to $-200 < u_{PWM} < 200$ despite the actual maximum being 255 (8-bit) to avoid wearing of the DC motor which defined

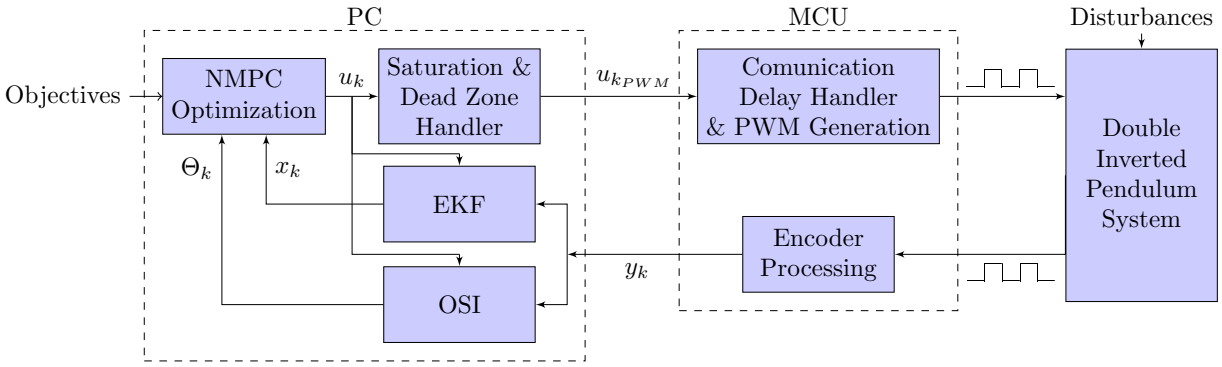


Fig. 4: Control Diagram

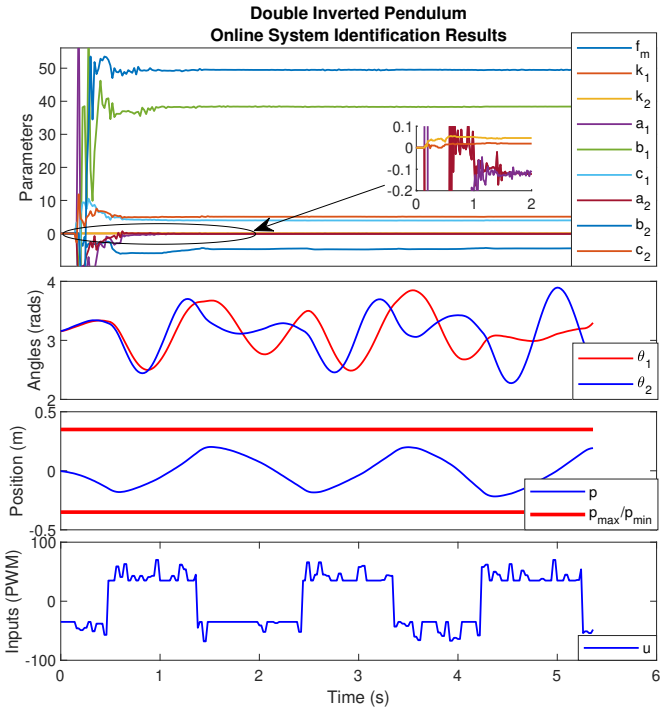


Fig. 5: Online System Identification Example

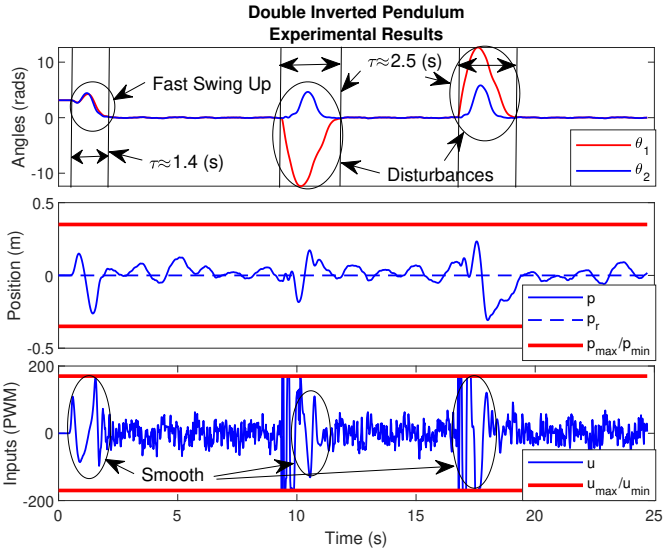


Fig. 6: Example performance with initial condition steady at lower equilibrium and large disturbances at $t \approx 9\ (s)$ and $t \approx 17\ (s)$.

the constraints to be included in the optimisation. Furthermore, the DC motor presented a dead-zone nonlinearity of $u_{dz} \approx 30$ which was removed by implementing the conditional function (26), adjusting constraints to $-200 + u_{dz} < u < 200 - u_{dz}$ and using u in the relevant models to simulate and linearise the system.

$$u_{PWM} = \begin{cases} u + u_{dz}, & u > 0 \\ u - u_{dz}, & u < 0 \end{cases} \tag{26}$$

The prediction horizon for the NMPC was set at $N_p = 75$ ($T_p = 1.5\ (s)$) leading to 600 outputs, 75 decision variables and 300 constraints to be optimised. The output and input weights were selected as $q_{k+i} = diag([1, 0.1, 0.1, 100, q_{\theta_1}, q_{\theta_2}, 10, 10])\ \ \forall i = [1, N_p - 1]$ and $R = 0.003I$ which were observed to give good balanced between fast swing up performance and input chattering due to noise at the steady state. A terminal weight $q_{k+N_p} = 10q_{k+i}$ was selected for the last values of the prediction horizon when $t_{lin} > 2\ (s)$, emulating soft zero terminal constraints to improve the stability characteristics of the optimisation [6]. Moreover, a tailored C++ code available in [14] was developed using EIGEN library following suggestions of [4, 6], and was tested in a laptop running Ubuntu 18.04 with an Intel i7-5700 HQ @ 2.7 GHz giving computation times of $t_{unc} < 800\ \mu s$ for the unconstrained solution and $t_{con} < 2500\ \mu s$ for the constrained one when doing 10 iterations of an efficient version of Hildreth's QP found in [15] . Finally, the EKF weights were set at $Q_{EKF} = diag([0.1, 0.1, 0.1, 0.0001, 0.0001, 0.0001, 1])$ and $R_{EKF} = diag([0.0001, 0.0001, 0.0001, 1])$ based on the variance of the errors observed in an offline analysis of the system identification process.

The resulting performance of the overall scheme can be seen in figure 6 starting from the rest position at the lower equilibrium and introducing large disturbances at $t \approx 9\ (s)$ and $t \approx 17\ (s)$. As it can be seen, the system clearly exhibits much faster performance than [11] giving settling times of $\tau_s \approx 1.4\ (s)$ for the swing up maneuver and of $\tau \approx 2.5\ (s)$ after large disturbances. Moreover, the system presented smooth input shapes during the swing up phases as a result of the added energy costs and the hybrid switching scheme. Furthermore, notice the position constraint is clearly satisfied at $t \approx 17\ (s)$ after the disturbance was given, demonstrating good handling of the rapid active-set changes by the QP. Finally, in some cases the position presented small steady state error and the well known limit cycle, however, this can be removed using standard methods such as integral control or disturbance estimation methods which are not the focus of the paper, and therefore were omitted. For the interest of the reader, an overall video is provided in (https://youtu.be/7E-SXi3YKQo) where the results can be seen, and the input-output data is available in [14].

7 Conclusion

This paper presents a novel NMPC approach based on the RTI Scheme for the swing-up and stabilisation of a parallel double inverted pendulum with experimental validation. The approach uses dual-mode closed loop predictions for the state deviation model to cancel the unstable open-loop dynamics of the double inverted pendulum which improve the numerical robustness of the optimisation. Moreover, two important modifications were introduced for the improvement of the RTI Scheme in the presence of large disturbances, namely; additional energy-related costs and a hybrid switching scheme. The approach was able to compute approximate constrained solutions in $t_c < 2500(\mu s)$, and a C++ code implementing it can be found in [14]. Finally, the approach was combined with an OSI Scheme based on RLS to address parameter uncertainty.

An overall video of the resulting performance is provided in (https://youtu.be/7E-SXi3YKQo), and the data obtained throughout the tests is available in [14].

To the best of the authors knowledge, this is the first contribution presenting numerical and experimental results for the swing-up and stabilization of a parallel double inverted pendulum in the presence of large disturbances based on NMPC using the RTI method. Future work will include the extension to the multiple-shooting scheme along with further analysis of the closed-loop performance with adaptation mechanism active as well as offset-free methods to cancel possible input-output disturbances.

Acknowledgments

This work was funded by CONACyT, Mexico.

8 References

1 S. Gros, M. Zanon, R. Quirynen, A. Bemporad, and M. Diehl, "From linear to nonlinear MPC: bridging the gap via the real-time iteration," *International Journal of Control*, vol. 7179, no. November, pp. 1–19, 2016.

2 H. Seki, S. Ooyama, and M. Ogawa, "Nonlinear Model Predictive Control Using Sucessive Linearization - Application to Chemical Reactors," *Trans. of the Society of Instrument and Control Engineers*, vol. E-3, no. 1, pp. 66–72, 2004.

3 B. Houska, H. J. Ferreau, and M. Diehl, "Autogenerating microsecond solvers for nonlinear MPC: A tutorial using ACADO integrators," *Optimal Control Applications and Methods*, vol. 36, pp. 685–704, 2015.

4 M. Vukov, A. Domahidi, H. J. Ferreau, M. Morari, and M. Diehl, "Auto-generated algorithms for nonlinear model predictive control on long and on short horizons," *52nd IEEE Conference on Decision and Control*, pp. 5113–5118, 2013.

5 M. Diehl, H. G. Bock, and J. P. Schlöder, "A Real-Time Iteration Scheme for Nonlinear Optimization in Optimal Feedback Control," *SIAM Journal on Control and Optimization*, vol. 43, no. 5, pp. 1714–1736, 2005.

6 B. Houska, H. J. Ferreau, and M. Diehl, "An auto-generated real-time iteration algorithm for nonlinear MPC in the microsecond range," *Automatica*, vol. 47, no. 10, pp. 2279–2285, 2011.

7 J. Kalmari, J. Backman, and A. Visala, "A toolkit for nonlinear model predictive control using gradient projection and code generation," *Control Engineering Practice*, vol. 39, pp. 56–66, 2015.

8 S. Gros, R. Quirynen, and M. Diehl, "Aircraft control based on fast non-linear MPC & multiple-shooting," *Proceedings of the IEEE Conference on Decision and Control*, no. 1, pp. 1142–1147, 2012.

9 R. Quirynen, M. Vukov, and M. Diehl, "Multiple Shooting in a Microsecond," in *Multiple Shooting and Time Domain Decomposition Methods*, vol. 9, pp. 183–202, Springer, Cham, 2015.

10 T. Glück, A. Eder, and A. Kugi, "Swing-up control of a triple pendulum on a cart with experimental validation," *Automatica*, vol. 49, no. 3, pp. 801–808, 2013.

11 M. Alamir and A. Murilo, "Swing-up and stabilization of a Twin-Pendulum under state and control constraints by a fast NMPC scheme," *Automatica*, vol. 44, no. 5, pp. 1319–1324, 2008.

12 a. Mills, A. Wills, and B. Ninness, "Nonlinear model predictive control of an inverted pendulum," *American Control Conference*, pp. 2335–2340, 2009.

13 J. A. Rossiter, B. Kouvaritakis, and M. J. Rice, "A numerically robust state-space approach to stable-predictive control strategies," *Automatica*, vol. 34, no. 1, pp. 65–73, 1998.

14 O. J. G. Villarreal, "A Fast Dual Mode NMPC for Parallel Double Inverted Pendulum." Code Ocean, Sept 2019. Available in https://doi.org/10.24433/CO.8048147.v1.

15 L. Wang, *Model Predictive Control System Design and Implementation Using Matlab*. Springer, 2009.

16 J. A. Rossiter, *A first course in predictive control*. Boca Raton: CRC Press, Taylor & Francis, 2nd ed. ed., 2018.

17 D. He, L. Wang, and S. Yu, "A new dual-mode NMPC scheme with terminal control laws of free-parameters," *Proceedings of the 33rd Chinese Control Conference, CCC 2014*, pp. 7667–7672, 2014.

18 M. Diehl, R. Findeisen, F. Allgöwer, H. G. Bock, and J. P. Schlöder, "Nominal stability of real-time iteration scheme for nonlinear model predictive control," *IEE Proceedings-Control Theory and Applications*, vol. 152, no. 3, pp. 296–308, 2005.

19 S. R. Anderson and V. Kadirkamanathan, "Modelling and identification of non-linear deterministic systems in the delta-domain," *Automatica*, vol. 43, no. 11, pp. 1859–1868, 2007.

20 S. F. Campbell, N. T. Nguyen, J. Kaneshige, and K. Krishnakumar, "Parameter Estimation for a Hybrid Adaptive Flight Controller," *AIAA Infotech@ Aerospace Conference, Seattle, WA*, no. April, pp. 1–27, 2009.

21 O. J. Gonzalez Villarreal, J. A. Rossiter, and H. Shin, "Laguerre-Based Adaptive MPC for Attitude Stabilization of Quad-Rotor," *2018 UKACC 12th International Conference on Control, CONTROL 2018*, pp. 360–365, 2018.