



HAL
open science

A Simple Machine Learning Technique for Model Predictive Control

Didier Georges

► **To cite this version:**

Didier Georges. A Simple Machine Learning Technique for Model Predictive Control. MED 2019 - 27th Mediterranean Conference on Control and Automation, Jul 2019, Akko, Israel. <10.1109/MED.2019.8798512>. <hal-02179706>

HAL Id: hal-02179706

<https://hal.science/hal-02179706v1>

Submitted on 11 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

A Simple Machine Learning Technique for Model Predictive Control

Didier Georges¹

Abstract—This paper is devoted to a simple approach for the offline computation of closed-loop optimal control for dynamical systems with imposed terminal state arising in Model Predictive Control Scheme (MPC). The here-proposed approach simply relies on some integrations of the characteristic equations associated to the optimal control problem, together with the classical supervised learning of a one-hidden-layer neuron network, to get a closed-loop MPC completely computed offline. Some examples are provided in the paper, which demonstrate the ability of this approach to tackle some quite large problems, with state dimensions reaching 50, without encountering limitations due to the so-called curse of dimensionality.

I. INTRODUCTION

The computation of nonlinear model predictive control in closed-loop form still remains a challenge due to the so-called curse of dimensionality of the associated optimal control problem. An alternative solution is to solve the associated optimal control problem online. However the computational cost may be incompatible with real-time for fast systems. Except for linear-quadratic optimal control problems, the computation of closed-loop solutions remains largely challenging when the state dimension is typically greater than 5. Several attempts have been made to compute the offline closed-loop solution to nonlinear optimal control problems by using a polynomial approximation of the solution of the associated Hamilton-Jacobi-Bellman equation [?] or some more general functional approximations of the optimal control thanks to some Galerkin approaches [?], [?]. Several approaches based on reinforcement learning and adaptive dynamic programming have been proposed [?]. However these latter approaches remain computationally expensive for medium or large scale nonlinear systems. In practice, all the above-mentioned approaches fail to offer practical solutions to problems whose state dimension is greater than 3 or 4, and are not appropriate to include terminal constraints. Other approaches can be derived by using model reduction to deal with the optimal control of a "small" system (see [?] for a recent paper in the linear case). However the extension of such an approach is to be made for the closed-loop optimal control of nonlinear systems. In this paper, the combination of a supervised learning technique with the integration of the characteristics of the associated Hamilton-Jacobi equation of the optimal control problem associated to a MPC scheme with a terminal state constraints and possibly some input bounds is proposed and experimented on several case studies.

The paper is organized as follows. In section 2, some background is provided on nonlinear MPC and the necessary

conditions for optimality of the associated optimal control problem. Section 3 describes the control design methodology proposed in this paper. In section 4, four illustrative examples demonstrate the effectiveness and easiness of the approach. Some conclusions and perspectives are given in section 5.

II. SOME BACKGROUND ON MODEL PREDICTIVE CONTROL OF NONLINEAR SYSTEMS

We consider a class of nonlinear systems defined by

$$\dot{x}(t) = F(x(t)) + G(x(t))u(t) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$, with $F(0) = 0$ (0 is an equilibrium state of the system). F and G are also assumed to be at least continuously differentiable.

In this paper we consider the continuous-time model predictive control of such systems around the origin with terminal state constraint, which consist in

- 1) Solving at time instant t , knowing the current state $x(t)$, an open-loop optimal control problem with finite control horizon $t + T$ defined by:

$$\min_u \int_t^{t+T} L(x(\tau), u(\tau)) d\tau \quad (2)$$

subject to $\dot{x}(\tau) = F(x(\tau)) + G(x(\tau))u(\tau)$, $\tau \in [t, t + T]$, with $x(t)$ known, $x(t + T) = 0$, and where

$$L(x, u) = l(x) + \frac{1}{2}u^T R(x)u - g(x)^T u \quad (3)$$

$$l(x) \geq 0, \forall x \neq 0, l(0) = 0, \quad (4)$$

$$R(x) = R^T(x) > 0, \forall x, \quad (5)$$

where $R(x)$ is a $m \times m$ matrix, and $g(x)$ is a vector of \mathbb{R}^m .

- 2) Applying optimal control solution $u^*(t)$ obtained at time t . At time $t + \epsilon$, the system reaches a new state $x(t + \epsilon) = x(t) + \int_t^{t+\epsilon} (F(x(\tau)) + G(x(\tau))u^*(\tau)) d\tau \approx x(t) + \epsilon(F(x(t)) + G(x(t))u^*(t))$.
- 3) Repeating the above sequence with $t + \epsilon \rightarrow t$.

Under mainly system controllability and zero-state observability of $(L(x, 0), F(x))$ assumptions, it can be shown that optimal cost function $V(t, x(t)) = \min_u \int_t^{t+T} L(x(\tau), u(\tau)) d\tau$ with $x(t + T) = 0$ is a Lyapunov function of the closed-loop system [?] under control law $u^*(t, x(t))$, optimal control solution of (??) at time t . Therefore, this model predictive control scheme described above ensures asymptotic stability around the origin.

¹Didier Georges is with Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, 38000 Grenoble, France *Institute of Engineering Univ. Grenoble Alpes, didier.georges@grenoble-inp.fr

In what follows, we will consider a nonsingular optimal control problem defined from 0 to T , since both F and G do not explicitly depend on t (time-invariant systems).

If $H(x, p) = L(x, u) + p^T(F(x) + G(x)u)$ defines the so-called Hamiltonian associated to the problem, Pontryagin's principle [?] for optimality provides the following necessary conditions for optimality

$$\begin{aligned} \nabla_u H &= 0 \\ \Leftrightarrow u^*(t) &= g(x(t)) - R^{-1}(x(t))G^T(x(t))p(t), \end{aligned} \quad (6)$$

$$\nabla_p H = 0 \Leftrightarrow \dot{x} = F(x) + G(x)u^*, \quad (7)$$

$$x(0) \text{ known}, \quad x(T) = 0, \quad (8)$$

$$\begin{aligned} \nabla_x H &= 0 \Leftrightarrow \dot{p} = -\nabla_x L(x, u^*) \\ &\quad - \frac{\partial}{\partial x}[F(x) + G(x)u^*]^T p. \end{aligned} \quad (9)$$

where $p(t)$ is the adjoint state of the system.

This defines a two-point boundary value problem (TP-BVP) which can be solved, for instance by using a shooting method which consists in finding $p(0, x(0))$ such that $x(T) = 0$, in order to get $u^*(0)$ thanks to (??). A basic shooting method can be defined as the solution of the following nonlinear least-square problem:

$$\begin{aligned} \min_{p(0)} & \frac{1}{2} \|x(T)\|^2 \\ \text{s.t.} & \end{aligned} \quad (10)$$

$$\begin{aligned} \dot{x} &= F(x) + G(x)g(x) - G(x)R^{-1}(x)G^T(x)p, \\ x(0) & \text{ known} \end{aligned} \quad (11)$$

$$\begin{aligned} \dot{p} &= -\nabla_x L(x, g(x) - R^{-1}(x)G^T(x)p) \\ &\quad - \frac{\partial}{\partial x}[F(x) + G(x)g(x) - G(x)R^{-1}(x)G^T(x)p]^T p. \end{aligned} \quad (12)$$

Since $p(0)$ is obtained, optimal control at time $t = 0$ $u^*(0, x(0))$ can be easily derived from (??).

Multi-shooting methods ([?]) are recommended for the computation of TPBVPs with large horizon T to avoid ill-conditioning.

In practice, the solution of TPBVP (??)-(??) can be very time consuming, especially for large systems, and the approach is then not appropriate for fast systems, since the TPBVP has to be solved at each time instant of control.

Rather than trying to solve TPBVP (??)-(??) at every time instant of control, we will consider an approach which will generate offline a closed-loop solution of the model predictive control defined above. **This approach relies on a sequence of simple integrations of differential equations (??) and (??) performed backward in time and starting always from $x(T) = 0$.** It can be noticed that equations (??) and (??) are nothing else than the equations of the characteristics of Hamilton-Jacobi-Bellman equation

$$\frac{\partial V}{\partial t}(t, x) + \min_{u(\cdot)} H(x, \frac{\partial V}{\partial x}) = 0, \quad (13)$$

associated to optimal problem (??), where $p(t) = \frac{\partial V}{\partial x}$.

Constrained input case: According to Pontryagin principle, if u is constrained to belong to a compact set U (for

instance a hypercube of \mathbb{R}^m), necessary condition (??) has to be replaced by

$$\begin{aligned} u^* &= \underset{u \in U}{\text{Arg min}} H \\ \Leftrightarrow u^* &= \text{Proj}_U(g(x) - R^{-1}(x)G^T(x))p \end{aligned} \quad (14)$$

where Proj_U denotes the projection operator onto U .

The approach remains unchanged by replacing (??) by (??) in characteristic equations (??) and (??), that however become discontinuous, what potentially makes integration much more tricky.

III. CLOSED-LOOP SOLUTION TO THE MODEL PREDICTIVE CONTROL PROBLEM

The offline derivation of a closed-loop solution of the model predictive control scheme relies on two ingredients:

- 1) The definition of a **one-hidden-layer neural network**, with N neurons in the hidden layer with n input neurons and m output neurons, used to approximate $u^*(x)$, which represents the model predictive control given by (??) at time $t = 0$, as a function of any initial state $x(0) = x$ in a domain large enough:

$$u^*(x) \approx \sum_{i=1}^N w_i \sigma(\alpha_i^T x + b_i) = \Phi_\theta(x), \quad (15)$$

where $\sigma(x) = \frac{1}{1 + e^{-x}}$ or \tanh represents the neuron activation function, and $w_i \in \mathbb{R}^m$, $\alpha_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, and $\theta = (w_i, \alpha_i, b_i)_{i=1, \dots, N}$.

It has been shown in [?] that such networks can be used as multi-dimensional approximants. In this paper, best results have been obtained with \tanh activation function. The use of multi-layer networks seems not to provide significant improvements in the present case according to preliminary experiments. However this point should be more deeply investigated.

- 2) The use of **low-discrepancy sequences** such as the ones proposed by Halton, Sobol, Faure (see [?] for instance) to generate learning sequences.

Such sequences have been proposed to solve the problem of optimally choosing M samples x_i in a hypercube $C = [0, 1]^n$ to "minimize holes" in the sense of the best possible approximation of integrals:

$$\left| \frac{1}{M} \sum_{i=1}^M f(x_i) - \int_C f(x) dx \right| \leq W(f) \frac{\log(M)^n}{M} \quad (16)$$

$W(f)$ is the variation of f in the sense of Hardy & Krause.

This approach usually provides better approximation results than other approaches based on random sequences for $n \leq 20$.

Low-discrepancy sequences could be used to generate initial state sequences by "filling" a hypercube of \mathbb{R}^n and then compute the corresponding sequence of TPBVPs (??)-(??). Even though this approach certainly can provide an effective way to get closed-loop

approximation to the nonlinear optimal solution thanks to supervised learning, it appears to be very computationally expensive for medium/large-scale systems.

In order to overcome this drawback, a low-discrepancy sequence of M samples $\{p^i(T)\}_{i=1,M}$, each of them belonging to a domain V_d of \mathbb{R}^n , is generated.

Then a **standard backward-in-time numerical integration of characteristic equations (??) and (??)** is performed starting from $t = T$ with each of the M samples $\{p^i(T)\}_{i=1,M}$ and $x(T) = 0$ to initial time $t = 0$ in order to get a sequence of M samples $x^i(0, p^i(T))$, representing the initial state of the optimal trajectory satisfying $x(T) = 0$, together with the set of M samples of control $u_i^*(x^i(0, p^i(T)))$ obtained by using control equation (??) or (??).

A **classical supervised learning technique** is finally used to tune parameters θ of the neural network, as solution of the following nonlinear regression problem, using the M pairs of samples $(x^i(0, p^i(T)), u_i^*(x^i(0, p^i(T))))$, $i = 1, \dots, M$ generated above:

$$\min_{\theta} \frac{1}{2} \sum_{i=1}^M \|\Phi_{\theta}(x^i(0, p^i(T)) - u_i^*(x^i(0, p^i(T)))\|^2. \quad (17)$$

Many approaches can be used to solve this problem (stochastic gradient, quasi-Newton ...) (see [?] for instance). In this paper, the Levenberg-Marquardt algorithm [?] implemented in MATLAB neural network toolbox has been used. According to my experience, this approach appears to be more efficient than the stochastic gradient approach for this class of problems. Due to nonconvexity, several learning trials are needed to retain the best solution in the least-square sense.

Remarks.

- Adjoint state domain V_d should be adequately chosen to ensure a well defined domain V of the $x^i(0, p^i(T))$'s including the origin (for instance a hypercube of the form $V = \prod_{i=1}^n [a_i, b_i]$). For that purpose, the computation of a finite number of TPBVPs (??)-(??) with initial state $x(0)$ defined in V , can be helpful to infer V_d as an hypercube in \mathbb{R}^n of the form $V_d = \prod_{i=1}^n [a_i^d, b_i^d]$.
- Control horizon T should be chosen not too large to avoid ill-conditioning and/or discontinuities which occur with nonlinear problems when the trajectories of characteristic equations (??) or (??) intersect then generating shocks and leading to integration failures and singularities.

IV. SOME ILLUSTRATIVE EXAMPLES

A. Constrained nonlinear model predictive control of a boost converter

In this section, the model predictive control of a boost converter is considered.

DC-to-DC boost converters are a class of power electronics devices used to elevate input voltage ([?]). This system exhibits non minimum phase behavior, which renders it quite difficult to control.

The average model of a boost converter is given by

$$\begin{aligned} \frac{di}{dt} &= -(1-u)v/L + E/L \\ \frac{dv}{dt} &= (1-u)i/C - v/(RC) \end{aligned} \quad (18)$$

where E is the input voltage, i is the inductance current, v is the output voltage, and u is the duty cycle of the converter acting as control input. The duty cycle is constrained to belong to interval $[0, 1]$.

In order to avoid ill-conditioning, a classical normalization procedure is performed using following changes of variables:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{E}\sqrt{L/C} & 0 \\ 0 & \frac{1}{E} \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix}, \quad \tau = \frac{t}{\sqrt{LC}} \quad (19)$$

The normalized average model is then given by

$$\begin{aligned} \frac{dz_1}{d\tau} &= -(1-u)z_2 + 1 \\ \frac{dz_2}{d\tau} &= (1-u)z_1 - \frac{1}{(R + \Delta R)\sqrt{C/L}}z_2, \end{aligned} \quad (20)$$

Here we consider the nonlinear MPC around the equilibrium defined by (z_1^e, z_2^e, u_e) of the boost converter described by Table ??.

The optimal control problem associated to the MPC scheme is defined by

$$\min_{u \in [0,1]} \frac{1}{2} \int_0^2 (\|z - z_e\|^2 + (u - u_e)^2) dt \quad (21)$$

where $z = (z_1, z_2)^T$, and $z_e = (z_1^e, z_2^e)^T$, subject to average model dynamics (??).

TABLE I

BOOST CONVERTER PARAMETERS

| L | C | R | E | i^e | v^e | u_e |
|------|------|----------------|-----|--------|-------|-------|
| 15mH | 50μF | $R = 50\Omega$ | 12V | 2.165A | 2.5V | 0.6 |

Fig. ?? shows the closed-loop dynamics of the boost converter under the constrained MPC. Fig. ?? shows the constrained closed-loop control input. Clearly, the bounds of the duty cycle are satisfied. The number of neurons in the hidden layer is equal to 40. The number of Sobol sequence samples of final time adjoint state $p(T) \in [-0.4, 0.4]^2$ is equal to 2000. Although the characteristics equations are unsmooth due to the input constraints, MATLAB function *ode45* was successfully used for the numerical integration. The retained solution was the best among 10 learning trials.

B. A small-scale linear MPC problem

The goal here is to compare the approximate solution of the following MPC with an controllable and unstable linear MPC system (with 3 unstable eigenvalues) with control horizon $T = 4$:

$$\min_u \frac{1}{2} \int_t^{t+4} (\|x\|^2 + \|u\|^2) d\tau \quad (22)$$

subject to

$$\dot{x} = Ax + Bu, \quad x \in \mathbb{R}^6, \quad u \in \mathbb{R}^2, \quad x(t+4) = 0, \quad (23)$$

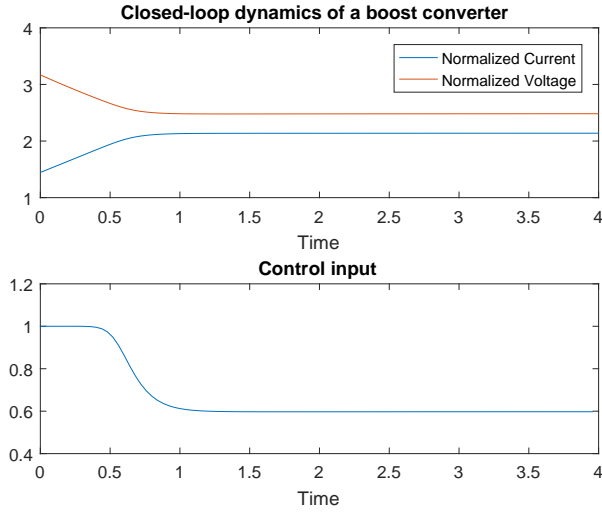


Fig. 1. Constrained MPC of the boost converter.

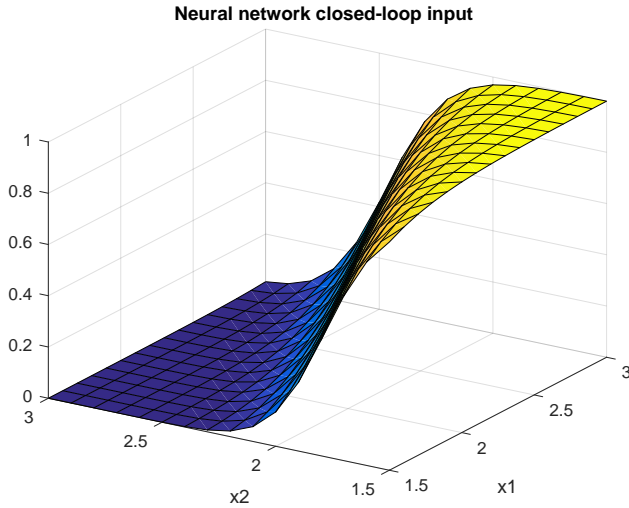


Fig. 2. Constrained MPC duty cycle as a function of the two states.

where coefficients of both A and B were randomly generated by using a uniform distribution on $[-1, 1]$, with the reference solution derived from the explicit solution obtained from the Hamiltonian system given by

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = H \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} A & -BB^T p \\ -I_6 & -A^T \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix}, \quad (24)$$

$x(t)$ known, $x(t+T) = 0$.

The closed-loop control is explicitly given by

$$u(t, x(t)) = -B^T p(t) \quad (25)$$

$$p(t) = -E_2(t+T)^{-1} E_1(t+T)x(t) \quad (26)$$

where E_1 and E_2 are given by

$$E(t) = \begin{pmatrix} E_1(t) & E_2(t) \\ E_3(t) & E_4(t) \end{pmatrix} = \exp(Ht). \quad (27)$$

Supervised learning was performed from a Sobol sequence of 3000 samples $p_i(t+T)$ defined in hypercube $[-0.5, 0.5]^6$. The hidden layer of the neural network used 25 neurons. Fig ?? shows the closed-loop dynamics with the best approximate control in terms of least-square cost function value obtained after solving 10 learning problems (??) for different initializations of network parameter θ .

Figs. ?? and ?? show very similar behaviors. Relative error between Hamiltonian solution u and approximate solution u_a $\frac{\int_t^{t+T} \|u(\tau) - u_a(\tau)\|^2 dt\tau}{\int_t^{t+T} \|u(\tau)\|^2 dt\tau}$ is equal to $7.6e^{-2}$.

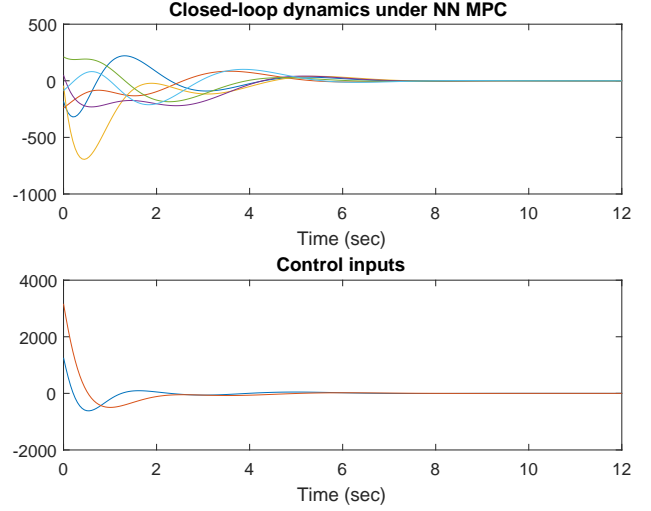


Fig. 3. Approximate MPC solution.

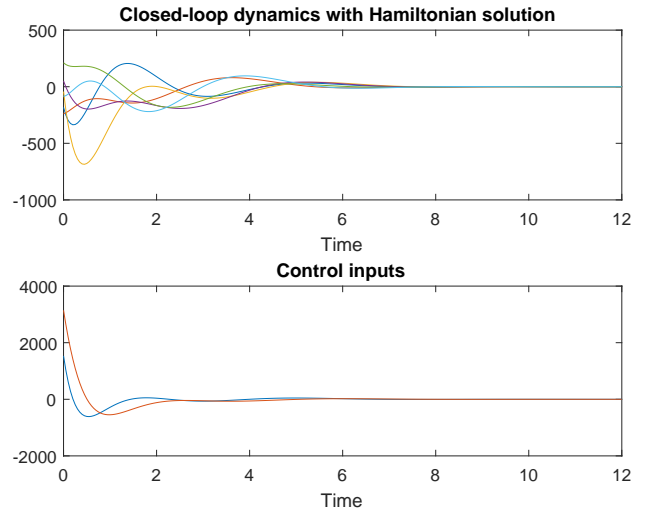


Fig. 4. Explicit linear MPC solution.

C. A medium-scale nonlinear MPC problem

In this section, the ability of the approach to deal with a medium-scale MPC problem is investigated.

Again a randomly generated controllable and unstable linear system of dimension 50 with 20 control inputs, and 22 unstable eigenvalues is tested, where coefficients of both A and B were randomly generated by using a uniform distribution on $[-1, 1]$.

Here the MPC problem is defined by the following convex but nonlinear optimal problem with control horizon $T = 0.25$ defined as what follows:

$$\min_u \frac{1}{2} \int_t^{t+0.25} (\log(1 + \frac{1}{2}\|x(\tau)\|^2) + \|u(\tau)\|^2) d\tau. \quad (28)$$

Supervised learning was performed from a Sobol sequence of only 1000 samples $p_i(t+T)$ defined in hypercube $[-0.15, 0.15]^{50}$. The hidden layer of the neural network used 25 neurons. Pairs $(x_i(t), u_i(t))$ were obtained from the backwards-in-time integration of characteristic equations

$$\dot{x} = Ax - BB^T p, \quad (29)$$

$$\dot{p} = -\frac{x}{1 + \frac{1}{2}\|x\|^2} - A^T p, \quad (30)$$

$$x(t+T) = 0, p_i(t+T), i = 1, \dots, 1000. \quad (31)$$

Fig. ?? shows the closed-loop dynamics with the best approximate control in terms of least-square cost function value obtained after solving 10 learning problems (??) for different initializations of network parameter θ . Fig. ?? demonstrates the effectiveness of the proposed approach to deal with medium-scale MPC problems at least when the dynamics is linear and the cost function is convex and nonlinear.

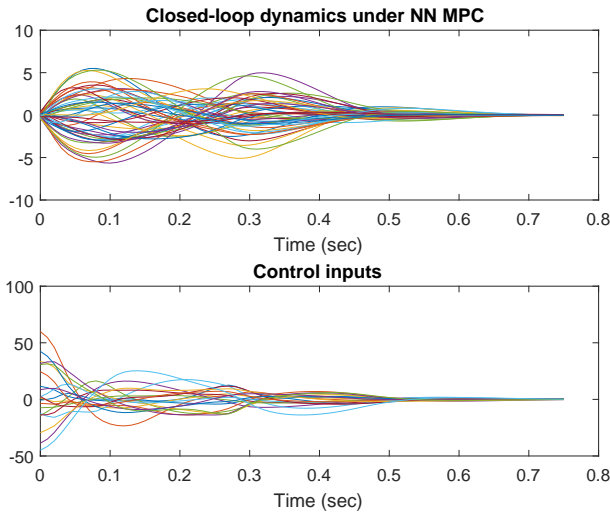


Fig. 5. MPC of a medium-scale linear system with a nonlinear convex cost function.

D. MPC of a nonlinear unstable exothermic process

In this section, the approximate solution of a MPC applied to an unstable 1D diffusion-reaction exothermic process (see [?]) is studied.

The process defined in a rod is fed with species A and an exothermic catalytic reaction of the form $A \rightarrow B$ occurs on the rod. The temperature profile is controlled with an actuator situated in the middle of the rod.

The exothermic distributed dynamics is modeled on Domain $[0, \pi]$ by the following nonlinear PDE

$$\frac{\partial z}{\partial t}(x, t) = \frac{\partial^2 z}{\partial x^2}(x, t) + \beta_T e^{-\frac{\gamma}{1+z(x, t)}} - \beta_T e^{-\gamma} + \beta_U(b(x)u(t) - z(x, t)) \quad (32)$$

where $z(x, t)$ denotes the dimensionless temperature, with Dirichlet boundary conditions $z(0, t) = 0$ and $z(\pi, t) = 0$.

The same coefficients as the ones in [?] have been used: Dimensionless heat of reaction $\beta_T = 50$ (implying instability), heat transfer coefficient $\beta_U = 2$, and activation energy $\gamma = 4$. Input operator $b(x)$ is taken to be nonzero and equal to 1 in finite spatial interval $[0.5\pi - \epsilon, 0.5\pi + \epsilon]$.

A classical method of lines consisting in the approximation of the Laplacian operator by finite differences has been used to get a finite-dimensional ODE model of the PDE defined on a spatial grid of $[0, \pi]$:

$$\frac{\partial^2 z_i}{\partial x^2} \approx \frac{z_{i+1} - 2z_i + z_{i-1}}{dx^2} \quad (33)$$

where dx is the spatial step and z_i denotes $z(x = idx)$.

The MPC problem (with $T = 0.27$) is defined by

$$\min_u \frac{1}{2} \int_t^{t+0.27} (\|Z(\tau)\|^2 + 5\|u(\tau)\|^2) d\tau \quad (34)$$

subject to

$$\dot{Z} = F(Z, u), Z(t) \in \mathbb{R}^7, \quad (35)$$

where F is the vector field defined by the method of lines on \mathbb{R}^7 (the temperatures evaluated at 7 spatial samples inside of the domain are used as state variables of the process dynamics), and 0 is the reference dimensionless temperature to be controlled.

Supervised learning was performed from a Sobol sequence of 2000 samples $p_i(t+T)$ defined in hypercube $[-0.1, 0.1]^7$. The hidden layer of the neural network used 40 neurons. Fig ?? shows the initial state distribution resulting from the integration of the characteristics.

Fig. ?? shows the closed-loop dynamics with the best approximate control in terms of least-square cost function value obtained after solving 10 learning problems (??) for different initializations of network parameter θ . The approximate closed-loop MPC is applied to an approximate dynamical model of dimension 20. Fig. ?? shows the related temperature profile in both time and space with $z_0(x) = 0.4 \sin(\pi x)$. These results demonstrate the effectiveness of the proposed approach with a limited number of 7 measurements corresponding to the 7 state components of reduced model (??).

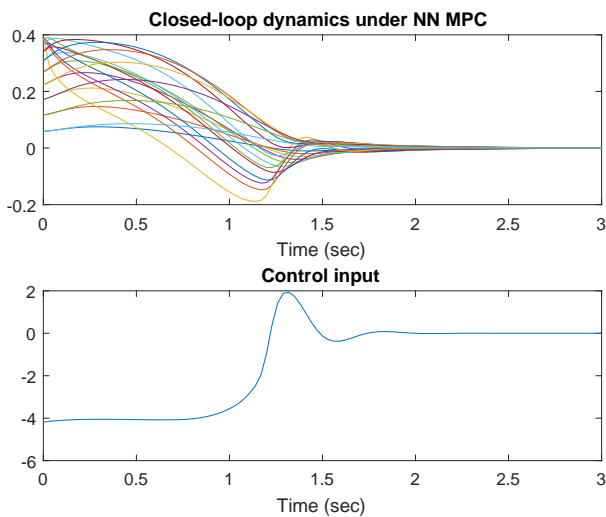


Fig. 6. Closed-loop response of the exothermic process.

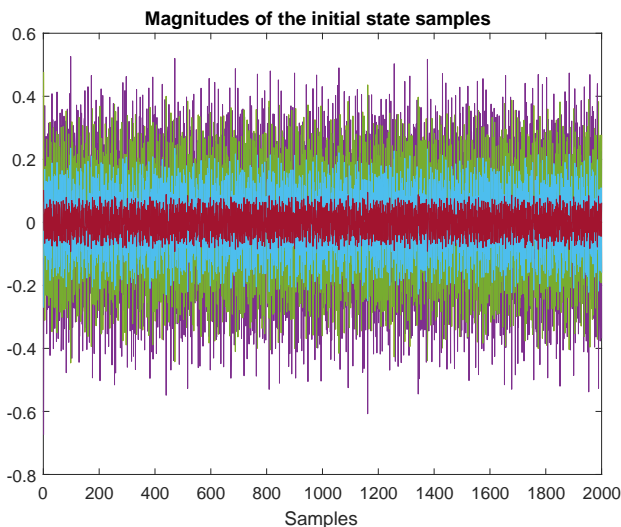


Fig. 7. Initial states distribution obtained after integration of the characteristics.

V. CONCLUSIONS AND PERSPECTIVES

Some preliminary results on the application of a neural-network-based machine learning method using a simple backward-in-time integration of the characteristic equations associated to the optimal control problem of a MPC scheme with a terminal state constraint have been provided. This approach can be easily extended to the case of discrete-time nonlinear MPC, thanks to the discrete-time optimality principle. The here-proposed approach seems to be able to cope with some quite large dynamical systems (state dimension up to 50). Since the method intrinsically solves an optimal control problem with a final state constraint, it can be used to solve other regular optimal control problems (with a final time cost for instance). A potential limitation

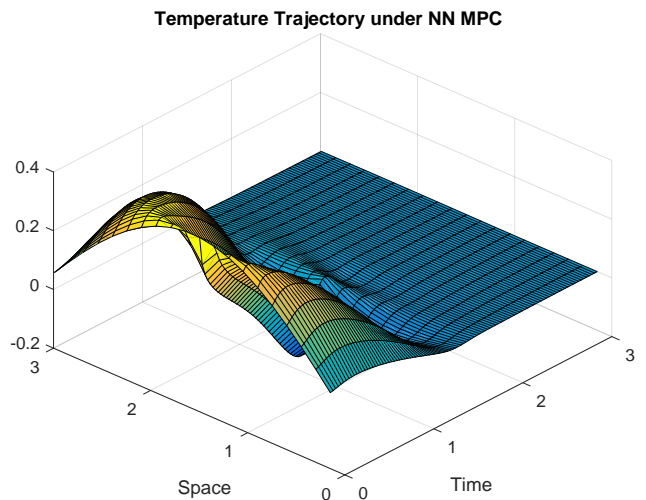


Fig. 8. Temperature profile in closed-loop (the boundary conditions ($= 0$) are not displayed).

to this approach is ill-conditioning and loss of continuities of the characteristic equations when the value of the control horizon is too large. Small control horizons potentially can lead to large optimal controls which could be sometimes incompatible with physical constraints. This limitation can be partially circumvented by introducing control bounds, using a solver suitable for discontinuous ODEs. Further studies are still needed to evaluate medium-scale nonlinear problems and investigate optimal state estimation.

REFERENCES

- [1] Beard, R.W., Sardis, G.N., and Wen, J.T. (1998). Approximate solutions to the time-invariant HJB equation, *Journal of Optimal Theory and Applications*, 96, 589-626.
- [2] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Systems*, 2 (1989), 303-314.
- [3] Daraghmeha A., Qatanani N., and Hartmann C. (2018). Optimal control of linear systems with balanced reduced-order models: Perturbation approximations, *Applied Mathematics and Computation* Volume 337, 15 November 2018, Pages 119-136.
- [4] El-Farra N.H., and Christofides P.D. (2004). Coordinating feedback and switching for control of spatially distributed processes, *Computers and Chemical Engineering* 28 (2004) 111-128.
- [5] Erickson, R. W., and Maksimovic, D. (2000). *Fundamentals of power electronics*, Second edition, Kluwer Academic Publishers, ISBN 0-7923-7270-0.
- [6] Fletcher C.A.J. (1984). *Computational Galerkin methods*, Springer Series in Computational Physics. Springer-Verlag, 1984.
- [7] Gander, M.J.; Vandewalle, S. (2007). Analysis of the Parareal TimeParallel TimeIntegration Method. *SIAM Journal on Scientific Computing*, SIAM. 29 (2): 556-578.
- [8] Georges, D. (1996). Solutions of nonlinear optimal regulator and H_∞ control problems via Galerkin methods, *European Journal of Control* (1996) 2:211-226.
- [9] Kelley C.T. (1999). *Iterative Methods for Optimization*, SIAM Frontiers in Applied Mathematics, no 18, 1999, ISBN 0-89871-433-8.
- [10] Lewis F.L., and Vrabie D. (2009). *Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control*, IEEE Circuits and System Magazine, 2009.
- [11] Mayne D.Q., Rawlings J.B., Rao C.V., and Sckaert P.O.M. (2000). Constrained model predictive control: Stability and optimality, *Automatica* 36 (2000) 789-814.

- [12] Mohri M., Rostamizadeh A., and Talawalkar A. (2012). Foundations of Machine Learning, MIT Press.
- [13] Niederreiter H. (1988). Low-Discrepancy and Low-Dispersion Sequences, *Journal of Number Theory* 30: 5170.
- [14] Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., and Mishchenko, E.F. (1962), *The Mathematical Theory of Optimal Processes*. English translation. Interscience. ISBN 2-88124-077-1.