European Journal of Control xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Control

journal homepage: www.elsevier.com/locate/ejcon

Data-driven dynamic relatively optimal control☆

Felice A. Pellegrino[a], Franco Blanchini[b], Gianfranco Fenu[a], Erica Salvato[a,*]

[a] Department of Engineering and Architecture, University of Trieste, Via Alfonso Valerio 6/1, Trieste 34100, Italy
[b] Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Universitá degli Studi di Udine, Via delle Scienze 206, Udine 33100, Italy

ARTICLE INFO

Article history:
Received 8 May 2023
Accepted 8 June 2023
Available online xxx

Recommended by Prof. T Parisini

Keywords:
Stabilizing feedback control
Data-driven control
Linear systems

ABSTRACT

We show how the recent works on data-driven open-loop minimum-energy control for linear systems can be exploited to obtain closed-loop control laws in the form of linear dynamic controllers that are relatively optimal. Besides being stabilizing, they achieve the optimal minimum-energy trajectory when the initial condition is the same as the open-loop optimal control problem. The order of the controller is $N-n$, where N is the length of the optimal open-loop trajectory, and n is the order of the system. The same idea can be used for obtaining a relatively optimal controller, entirely based on data, from open-loop trajectories starting from up to n linearly independent initial conditions.

1. Introduction

Model-based control (MBC) approaches have been widely exploited over the years, exhibiting their ability to provide effective and reliable control laws in a large variety of control tasks. However, their implementation is strictly constrained by the existence of a dynamic model of the system to be controlled, not often easy to derive or identify. To overcome this intrinsic limitations of MBC approaches, the control community is recently focusing on the so-called *data-driven control* (DDC), i.e., the family of model-free control solutions in which the synthesis of controllers is entirely based on input-output data collections.

The key issue in data-driven control is how to replace process models with data. A first solution in this regard was postulated, in the case of linear systems, by Willems et al. [18], who states that a linear system can be dynamically represented by a finite set of system trajectories, provided that these trajectories come from sufficiently excited dynamics. This lemma has been more or less explicitly exploited for the design of data-driven controls.

In [3], for example, authors propose an off-line approach leading to optimal open-loop input sequences from data-batch collected in preliminary experiments. Here, explicit formulas for the open-loop minimum energy control problem, based entirely on experimental data, are derived for linear, unconstrained discrete-time systems. A less restricted experimental framework is presented in Baggio and Pasqualetti [4], while some applications on complex systems, such as power-grid networks and brain networks, are reported in Baggio et al. [2].

Data-driven closed-loop solutions are instead proposed in De Persis and Tesi [13], Rotulo et al. [17] where linear quadratic regulator (LQR) problems are faced respectively with infinite and finite optimization time horizons. Further results on the data-driven LRQ approach applied on nonlinear systems, and in case of data corrupted by noise, can be found in De Persis and Tesi [14], De Persis and Tesi [15]. Data-driven solutions to address the model predictive control (MPC) problem are instead proposed in Berberich et al. [5], Coulson et al. [11,12], Yang and Li [19]. A *one-shot* robust controller synthesis, based on an expert operator's trajectory is instead proposed in Blanchini et al. [7].

In the present work, we deal with open-loop optimal control sequences obtained by the sole experimental data, and specifically, by sequences of inputs and the corresponding states. The aim is to exploit such open-loop sequences to get a closed-loop control law. In a previous work [16], we employed the state-space partitioning technique described in Blanchini and Pellegrino [10], to get a static, nonlinear, state-feedback controller having the property (besides being stabilizing), of guaranteeing the achievement of the optimal trajectory when starting from the same initial condition of the optimal open loop trajectory. Here, based on the dynamic relatively optimal control (ROC) [8], we show how to synthesize a *linear dy-*

☆ This work has been partially supported by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant no. 2017YKXYXJ. It was also carried out within the PNRR research activities of the consortium iNEST (Interconnected North-Est Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043). This manuscript reflects only the Authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

* Corresponding author.
E-mail addresses: fapellegrino@units.it (F.A. Pellegrino), blanchini@uniud.it (F. Blanchini), fenu@units.it (G. Fenu), erica.salvato@dia.units.it (E. Salvato).

https://doi.org/10.1016/j.ejcon.2023.100839
0947-3580/© 2023 European Control Association. Published by Elsevier Ltd. All rights reserved.

namic controller having the same properties. As a result, we get a data-driven dynamic ROC. The remaining of the paper is organized as follows: Section 2 recalls how the dynamic ROC can be used for synthesizing a closed-loop controller starting from optimal open-loop input and state sequences; Section 3, based on [16], provides an explicit, data-driven, formula for the minimum energy control sequence leading the state to zero from an arbitrary initial state and shows how to recover the corresponding state trajectory; Section 4 provides a generalization to the case of multiple initial conditions; two numerical examples are provided in Section 5, and conclusions are drawn in Section 6.

2. Dynamic relatively optimal control

The Relatively Optimal Control (ROC) [8] is a kind of control that, besides being stabilizing, guarantees the optimality of certain trajectories, specifically, those starting from a given (or a set of given) *nominal initial conditions*. Both linear dynamic [8,9], and non-linear static [10] implementations of ROC for discrete-time linear systems have been proposed, as well as a continuous-time solution based on the Youla–Kučera parameterization [6]. For the sake of completeness, we report next the essentials of the ROC, from Blanchini and Pellegrino [8].

Consider the time-invariant discrete-time linear system

$$x(k+1) = Ax(k) + Bu(k), \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, while $x(k) \in \mathbb{R}^n$ and $u(k) \in \mathbb{R}^m$ denote respectively the state and the input at time $k \in \mathbb{N}$. For a given horizon K, and initial state x_{ini}, the following open-loop control problem can be formulated:

$$J_{\text{opt}}(x_{\text{ini}}) = \min \sum_{k=0}^{K-1} l(x(k), u(k))$$
$$\text{subject to:} \quad (1)\ k = 0 \dots K-1$$
$$x(0) = x_{\text{ini}},\ x(K) = 0, \tag{2}$$

where $l(\cdot,\cdot)$ is a convex function of its arguments, and the decision variables are the control actions $u(0), \dots, u(K-1)$.

Let the optimal state and input sequences obtained by solving the above problem, be arranged in the matrices:

$$\mathcal{X} = [x(0)\ \ x(1)\ \ \dots\ \ x(K-1)]$$
$$\mathcal{U} = [u(0)\ \ u(1)\ \ \dots\ \ u(K-1)], \tag{3}$$

and choose a $(K-n)\times K$ matrix

$$\mathcal{Z} = [0\ \ z(1)\ \ z(2)\ \ \dots\ \ z(K-1)]$$

such that

$$M = \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix} = \begin{bmatrix} 0 & z(1) & z(2) & \dots & z(K-1) \\ x(0) & x(1) & x(2) & \dots & x(K-1) \end{bmatrix}, \tag{4}$$

is invertible. Such a choice is possible provided that $\mathcal{X}$ is full row rank.

Now, consider the dynamic compensator

$$z(k+1) = Qz(k) + Rx(k) \tag{5}$$

$$u(k) = Sz(k) + Tx(k) \tag{6}$$

where Q, R, S, T are achieved as the unique solution of the linear equation

$$\begin{bmatrix} \mathcal{Z}\mathcal{P} \\ \mathcal{U} \end{bmatrix} = \begin{bmatrix} Q & R \\ S & T \end{bmatrix} \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix}, \tag{7}$$

where the square matrix $\mathcal{P}$ is the K–Jordan block associated with the 0 eigenvalue:

$$\mathcal{P} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \tag{8}$$

The main result of Blanchini and Pellegrino [8] states that the above compensator (whose order is $K-n$) is relatively optimal, namely, it is stabilizing and results in the optimal trajectory $\mathcal{X}$ when the initial states of the system and the compensator are, respectively, $x(0) = x_{\text{ini}}$ and $z(0) = 0$. Notice that from (7) and $\mathcal{X}\mathcal{P} = A\mathcal{X} + B\mathcal{U}$ (which holds true by construction), it follows that

$$\begin{bmatrix} Q & R \\ BS & A+BT \end{bmatrix} \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix} = \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix} \mathcal{P}, \tag{9}$$

implying that the closed-loop matrix is similar to the nilpotent matrix $\mathcal{P}$. In other words, an asymptotically stable, in fact dead-beat, control is achieved. Regarding the choice of K, the only constraint is $K \geq n$, which in the case of $K = n$ leads to a static compensator of limited significance. Therefore, we are more interested in considering the case $K > n$. Indeed, since the design parameter K is the number of steps allowed to reach the origin, taking it small may cause excessive control exploitation we avoid taking K not too small. Clearly, the cost of the problem (2) is monotonic decreasing with respect to K, so we have a potential benefit.

In the following, we apply the dynamic ROC to obtain a closed-loop control law from data-driven, open-loop, optimal trajectories.

3. Data-driven minimum energy control

The present section reports some results on data-driven minimum energy control from Pellegrino et al. [16] and, specifically, provides explicit formulas for getting the optimal input and state trajectory directly from data. From that trajectories, then, the dynamic ROC can be synthesized.

For a given horizon K, the minimum-energy control problem to zero is that of finding, among the input sequences that drive the state from $x(0) = x_{\text{ini}}$ to $x(K) = 0$, the one of minimum energy, i.e., the one minimizing $\sum_{k=0}^{K-1} \|u(k)\|_2^2$. Clearly, the problem is a special case of (2), corresponding to $l(x(k), u(k)) = \|u(k)\|_2^2$. Let us denote (with slight abuse) by u the sequence $\left[u(K-1)^\top, \dots, u(0)^\top\right]^\top$, and by u^* the optimal one. With the same notation, the optimal input sequence u^* can be expressed as the minimum 2-norm solution of the following equation:

$$0 = A^K x_{\text{ini}} + \underbrace{\begin{bmatrix} B & AB & \dots & A^{K-1}B \end{bmatrix}}_{R_K} u, \tag{10}$$

where R_K is the K-steps reachability matrix. For A and B (and thus R_K) known, the solution to the above problem is well-known to be [1]

$$u^* = -R_K^\dagger \left(A^K x_{\text{ini}}\right), \tag{11}$$

where † denotes the Moore–Penrose pseudo-inverse.

Here, we are interested in solving Eq. (10) relying on experimental data only. In addition, since the ROC technique described in the previous section requires the optimal open loop *state* trajectory, besides the optimal input sequence, we need to compute the optimal state sequence from data as well. The mentioned issues are dealt with in the following subsections. When the optimal input and state sequences have been computed based on data, the ROC technique can be applied, leading to a closed-loop, stabilizing, data-driven control law. An example is reported in Section 5.

3.1. Optimal input sequence from data

The experimental data employed is similar to that of Baggio et al. [3], in which a set of $N \geq n$ experiments is available, each starting from $x_0 = 0$ and lasting K steps. Denoting by u_i the ith (arbitrary) input *sequence*, and by x_i the state reached at time K of the ith experiment, the matrices

$$U = [u_1 \quad \ldots \quad u_N], \quad \text{and} \quad X = [x_1 \quad \ldots \quad x_N], \tag{12}$$

are constructed. We remark that matrix U is $(mK) \times N$ (since u_i denotes an input sequence of length K), while X is $n \times N$. Here, according to Baggio et al. [3], we assume that U is a full rank matrix. Clearly, we have:

$$x_i = R_K u_i, \quad i = 1, \ldots, N. \tag{13}$$

The previous can be written as $X = R_K U$, and the solution of problem $\min \|X - R_K U\|_F^2$, i.e.,

$$R_K^* = XU^\dagger, \tag{14}$$

is an estimate of the K-steps reachability matrix. Under the assumption of $N = Km$ [3], and $\mathrm{rank}\begin{bmatrix} X^\top & U^\top \end{bmatrix} = \mathrm{rank}\begin{bmatrix} U^\top \end{bmatrix}$, R_K^* exactly matches the reachability matrix. Note that the full rank property of the U matrix is a sufficient condition to ensure this match.

In the following, we will employ such an estimate in place of the unknown reachability matrix, and we will denote it by R_K.

Due to the term $A^K x_{\mathrm{ini}}$, substituting Eq. (14) in Eq. (11) is not sufficient to get a solution based on data only. A possibility would be to use the results of Baggio and Pasqualetti [4], which extends [3] to more general problems and less restrictive experimental setups. As a simpler alternative, we propose to collect N sequences of length $2K$:

$$\underbrace{u(2K-1)^\top, \ldots, u(K)^\top}_{\hat{u}_i^\top}, \underbrace{u(K-1)^\top, \ldots, u(0)^\top}_{u_i^\top}$$

$$x(1), \ldots, \underbrace{x(K)}_{x_i}, x(K+1), \ldots, \underbrace{x(2K)}_{\hat{x}_i}$$

and construct the following matrices:

$$U = [u_1, \ldots, u_i, \ldots, u_N], \qquad \hat{U} = \begin{bmatrix} \hat{u}_1, \ldots, \hat{u}_i, \ldots, \hat{u}_N \end{bmatrix}$$

$$X = [x_1, \ldots, x_i, \ldots, x_N], \qquad \hat{X} = \begin{bmatrix} \hat{x}_1, \ldots, \hat{x}_i, \ldots, \hat{x}_N \end{bmatrix},$$

where U and X are the same of Eq. (12), while $\hat{U}$ and $\hat{X}$ correspond to trajectories of length K, starting (in general) from nonzero states. Then, by construction, $\forall i = 1 \ldots N$ we have:

$$\hat{x}_i = A^K x_i + R_K \hat{u}_i,$$

which can be written in compact form as

$$A^K X = \hat{X} - R_K \hat{U}, \tag{15}$$

and, in view of Eq. (14), as

$$A^K X = \hat{X} - XU^\dagger \hat{U}. \tag{16}$$

Thus, the right-hand side of Eq. (16) can be used to compute the term $A^K x_{\mathrm{ini}}$ for any x_{ini} in the column space of X. Specifically, let $\alpha \in \mathbb{R}^N$ be such that

$$x_{\mathrm{ini}} = X\alpha. \tag{17}$$

Then, we have

$$A^K x_{\mathrm{ini}} = A^K X \alpha = \left(\hat{X} - XU^\dagger \hat{U}\right)\alpha.$$

Using the least-norm solution for α in Eq. (17), i.e., $\alpha = X^\dagger x_{\mathrm{ini}}$, we get

$$A^K x_{\mathrm{ini}} = \left(\hat{X} - XU^\dagger \hat{U}\right) X^\dagger x_{\mathrm{ini}}. \tag{18}$$

Finally, by substituting in Eq. (11), namely $u^* = -R_K^\dagger\left(A^K x_{\mathrm{ini}}\right)$, and recalling that $R_K = XU^\dagger$, we get:

$$u^* = \left(XU^\dagger\right)^\dagger \left(XU^\dagger \hat{U} - \hat{X}\right) X^\dagger x_{\mathrm{ini}}, \tag{19}$$

which gives a data-driven open-loop minimum energy control sequence leading the state to zero in K steps from x_{ini}. The formula provides the optimal sequence when x_{ini} belongs to the column space of X. In particular, if X is rank n, then x_{ini} can be arbitrary.

3.2. Optimal state trajectory from data

To get the closed-loop control law by means of the technique described in Section 2, besides the optimal input sequence u^*, given by Eq. (19), the corresponding optimal state trajectory is needed. A such trajectory can be recovered from u^* and the data obtained from the same N sequences collected before.

It is sufficient to define the matrices U_k, $\hat{U}_k$, X_k, and $\hat{X}_k$, similarly as before, but based on *subsequences* of length $2k$, for $k = 1 \ldots K-1$. Let δ_k denote the starting index of the subsequences of length $2k$, and define

$$U_k = \begin{bmatrix} {}^k u_1, \ldots, {}^k u_i, \ldots, {}^k u_N \end{bmatrix},$$
$$\hat{U}_k = \begin{bmatrix} {}^k \hat{u}_1, \ldots, {}^k \hat{u}_i, \ldots, {}^k \hat{u}_N \end{bmatrix},$$
$$X_k = \begin{bmatrix} {}^k x_1, \ldots, {}^k x_i, \ldots, {}^k x_N \end{bmatrix},$$
$$\hat{X}_k = \begin{bmatrix} {}^k \hat{x}_1, \ldots, {}^k \hat{x}_i, \ldots, {}^k \hat{x}_N \end{bmatrix},$$

where ${}^k u_i = \begin{bmatrix} u(\delta_k + k - 1)^\top, \ldots, u(\delta_k)^\top \end{bmatrix}^\top$, ${}^k \hat{u}_i = \begin{bmatrix} u(\delta_k + 2k - 1)^\top, \ldots, u(\delta_k + k)^\top \end{bmatrix}^\top$, ${}^k x_i = x(\delta_k + k)$, and ${}^k \hat{x}_i = x(\delta_k + 2k)$.

Hence, by letting $U_K = U$, $X_K = X$, $\hat{U}_K = \hat{U}$, and $\hat{X}_K = \hat{X}$ we can write

$$A^k X_k = \hat{X}_k - R_k \hat{U}_k, \quad k = 1, \ldots, K, \tag{20}$$

where R_k is the k-step reachability matrix, corresponding to the first k columns of R_K:

$$R_k = R_K \begin{bmatrix} I_k \\ 0 \end{bmatrix} = XU^\dagger \begin{bmatrix} I_k \\ 0 \end{bmatrix},$$

where I_k denotes the identity matrix of dimension k. Eq. (20) holds irrespective of the choice of the subsequences (i.e., of indices δ_k). However, it is convenient to choose the subsequences in such a way that $\mathrm{rank}[X]_k = n$, $\forall k$. This is always possible, provided that $\mathrm{rank}[X] = n$, and can be achieved by taking $\delta_k = K - k$, leading to $X_k = X$, $\forall k$. The full-rank condition on the X_k guarantees that any initial state x_{ini} can be written as a linear combination of the columns of any of the X_k. As a consequence, the optimal state trajectory, in terms of data, is given by:

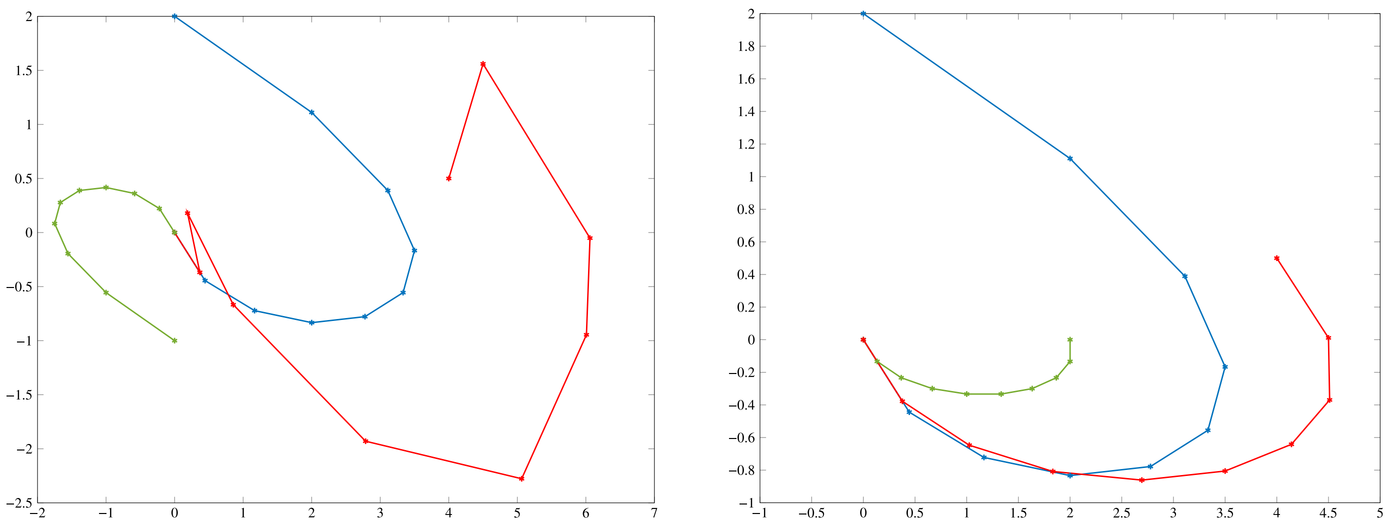
$$x(k) = \left(\hat{X}_k - R_k \hat{U}_k\right) X_k^\dagger x_{\mathrm{ini}} + R_k u_k^* \qquad k = 1, \ldots, K, \tag{21}$$

where u_k^* is the vector composed by the first k steps of the optimal input sequence: $u_k^* = \begin{bmatrix} u^*(k-1)^\top, \ldots, u^*(0)^\top \end{bmatrix}^\top$. Finally note that from (19) and (21), we can derive a data-driven relatively optimal control for the initial condition x_{ini} by completing the matrix X as in (4) and adopting (7).

Notice that for a given initial state x_{ini} and a K-steps input sequence u, the arrival to zero condition can be expressed as

$$XU^\dagger u = -\left(\hat{X} - XU^\dagger \hat{U}\right) X^\dagger x_{\mathrm{ini}}.$$

As a consequence, the minimum energy problem already discussed can be stated equivalently in the following quadratic programming (QP) form:



(a) (b)

Fig. 1. (a) Optimal trajectory corresponding to the optimal input sequence in (28) in blue. Trajectory obtained applying the resulting controller starting from $x(0) = -x_{\text{ini}}/2 = [0 \;\; 1]^\top$ in green. Trajectory obtained applying the resulting controller starting from $x(0) = [4 \;\; 0.5]^\top$ in red. (b) Optimal trajectory corresponding to the optimal input sequences in (29) and (30) respectively in blue and green. Trajectory obtained applying the resulting controller starting from $x(0) = [4 \;\; 0.5]^\top$ in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$\min J = \|u\|_2^2$$

subject to:

$$XU^\dagger u = -\big(\hat{X} - XU^\dagger \hat{U}\big)X^\dagger x_{\text{ini}}. \tag{22}$$

This formulation allows the use of different cost functions, thus, for example, the LQR-type cost $J = \sum x^\top(k)\mathcal{Q}x(k) + u^\top(k)\mathcal{R}u(k)$. Furthermore, it admits the introduction of state constraints given that the state at time $k \le K$ can be written as

$$x(k) = \big(\hat{X}_k - R_k\hat{U}_k\big)X_k^\dagger x_{\text{ini}} + R_k u_k, \tag{23}$$

where u_k is the vector composed by the first k steps of the input sequence: $u_k = \big[u(k-1)^\top, \ldots, u(0)^\top\big]^\top$.

4. Relative optimality from multiple initial conditions

The linear implementation of the relatively optimal control [8], allows to obtain a single dynamic controller which is relatively optimal from several, linearly independent, initial states $x_{\text{ini}}^{(i)} = x^{(i)}(0)$, $i = 1 \ldots r \le n$. Indeed, denoting by

$$\begin{aligned}\mathcal{X}^{(i)} &= \big[x^{(i)}(0) \;\; x^{(i)}(1) \;\; \ldots \;\; x^{(i)}(K_i - 1)\big]\\ \mathcal{U}^{(i)} &= \big[u^{(i)}(0) \;\; u^{(i)}(1) \;\; \ldots \;\; u^{(i)}(K_i - 1)\big],\end{aligned} \tag{24}$$

the optimal state and input trajectory (of length K_i), define the matrices:

$$\begin{aligned}\mathcal{X} &= \big[\mathcal{X}^{(1)} \;\; \mathcal{X}^{(2)} \;\; \ldots \;\; \mathcal{X}^{(r)}\big]\\ \mathcal{U} &= \big[\mathcal{U}^{(1)} \;\; \mathcal{U}^{(2)} \;\; \ldots \;\; \mathcal{U}^{(r)}\big].\end{aligned} \tag{25}$$

Then, it is sufficient to complete $\mathcal{X}$ as:

$$M = \begin{bmatrix}\mathcal{Z}\\ \mathcal{X}\end{bmatrix} = \begin{bmatrix}\mathcal{Z}^{(1)} & \mathcal{Z}^{(2)} & \mathcal{Z}^{(3)} & \ldots & \mathcal{Z}^{(r)}\\ \mathcal{X}^{(1)} & \mathcal{X}^{(2)} & \mathcal{X}^{(3)} & \ldots & \mathcal{X}^{(r)}\end{bmatrix}, \tag{26}$$

where each matrix $\mathcal{Z}^{(i)}$ has the form

$$\mathcal{Z}^{(i)} = \big[0 \;\; z^{(i)}(1) \;\; z^{(i)}(2) \;\; \ldots \;\; z^{(i)}(K_i - 1)\big],$$

and in such a way that M is invertible. The relatively optimal compensator, guaranteeing optimality from the set of initial conditions $x_{\text{ini}}^{(i)}$, $i = 1 \ldots r$, is given by the unique solution of (7), where

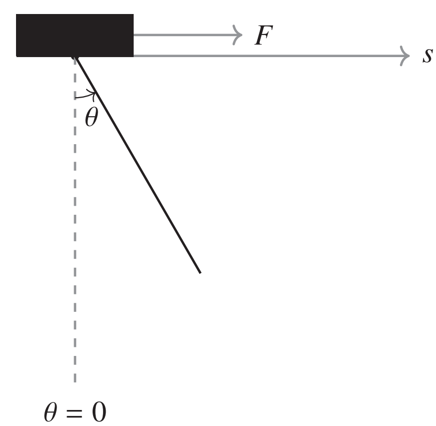
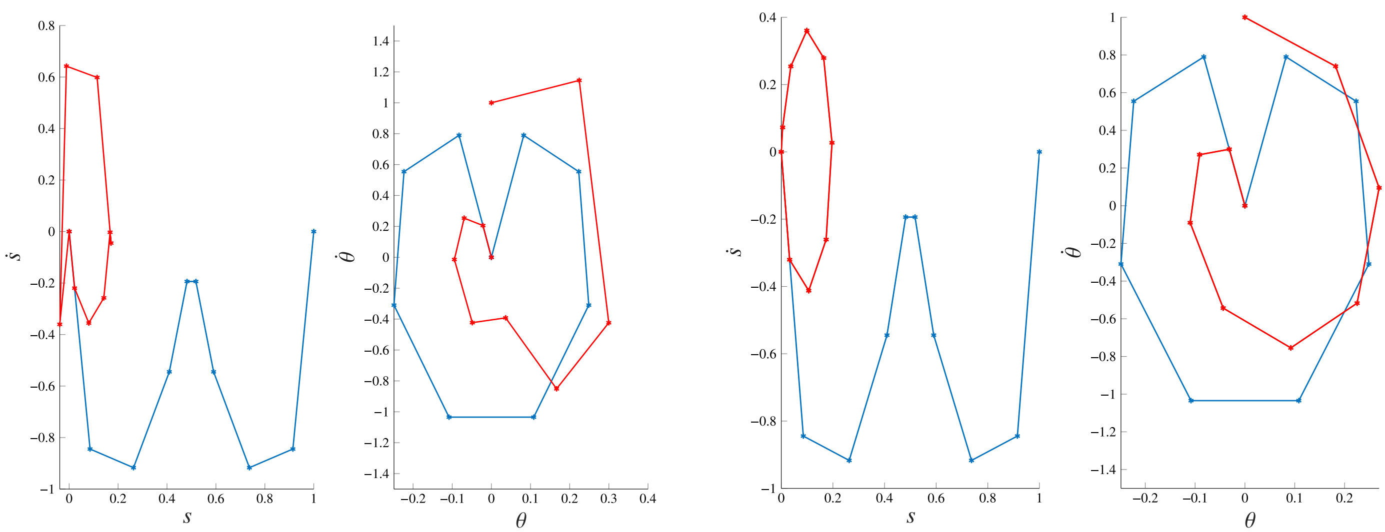


Fig. 2. Cart and pole system.

$\mathcal{P} = \text{diag}\left\{\mathcal{P}^{(i)}\right\}$ is a diagonal block matrix whose blocks have the form (8).

The data-driven methodology shown in the previous section can be readily applied to the case of multiple initial conditions, leading to a dynamic compensator that guarantees minimum energy trajectories (possibly of different lengths) from each of the initial conditions. As far as the energy of the control sequence from a generic initial state $x(0)$ is concerned, the following considerations apply. Consistently with the relatively control framework, the closed-loop control is optimal, and thus, minimum energy, only from $x_{\text{ini}}^{(i)}$, $i = 1 \ldots r$ (and for the null initial state of the compensator). However, since the closed-loop system is linear, the cost (the energy) is a convex function of the initial state. Thus, if the generic initial state $x(0)$ can be written as $x(0) = \sum_i^r \alpha_i x^{(i)}$, it follows that the cost from $x(0)$ is bounded by $\sum_{i=1}^r |\alpha_i| J^{(i)}$, where $J^{(i)}$ is the (optimal) cost from $x_{\text{ini}}^{(i)}$. In particular, for $x(0)$ belonging to the convex hull of the $x_{\text{ini}}^{(i)}$, $i = 1 \ldots r$, the cost is bounded by the largest cost associated with all the vertices (Fig. 1).

Finally, we point out that, as shown in the numerical example, resorting to a set of initial conditions spanning the whole state



(a) (b)

Fig. 3. (a) Optimal trajectory obtained solving the data-driven ROC problem on the cart and pole system (Fig. 2), starting from $x_{\text{ini}} = [1 \;\; 0 \;\; 0 \;\; 0]^\top$ (blue). Trajectory obtained by applying the same compensator starting from $x_{\text{ini}} = [0 \;\; 0 \;\; 0 \;\; 1]^\top$ (red). (b) Optimal trajectories obtained solving the data-driven ROC problem on the cart and pole system (Fig. 2) starting from the two linearly independent initial conditions $x_{\text{ini}}^{(1)} = [1 \;\; 0 \;\; 0 \;\; 0]^\top$ (blue), and $x_{\text{ini}}^{(2)} = [0 \;\; 0 \;\; 0 \;\; 1]^\top$ (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

space can lead to compensators being less sensitive to the initial conditions.

5. Numerical examples

We consider the double integrator:

$$x(k+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k). \tag{27}$$

We set $K = 9$ and perform $N = 20$ experiments lasting $2K = 18$ steps, starting from $x_0 = \begin{bmatrix} 0 & 0 \end{bmatrix}^\top$ and applying randomly chosen inputs. The collected states and inputs are then used to construct the U, $\hat{U}$, X and $\hat{X}$ matrices of Eq. (19). We choose $x_{\text{ini}} = \begin{bmatrix} 0 & 2 \end{bmatrix}^\top$, and we solve the minimum-energy control problem by applying Eq. (19), thus obtaining the minimum energy input sequence that leads the system to $x(9) = 0$ starting from $x(0) = x_{\text{ini}}$. The obtained optimal (open-loop) control sequence is:

$$\begin{aligned} u^* = [&-0.889 \;\; -0.722 \;\; -0.556 \;\; -0.389 \\ &-0.222 \;\; -0.056 \;\; 0.111 \;\; 0.278 \;\; 0.444], \end{aligned} \tag{28}$$

while the resulting optimal trajectory, say x^*, obtained by applying (21), is shown in blue in Figure 1. We then synthesize the dynamic compensator by applying (7), where $\mathcal{X} = [x^*(0) \; \ldots \; x^*(K-1)]$ and $\mathcal{Z} = [0 \; z(1) \; \ldots \; z(K-1)]$ is randomly chosen and such that $M = \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix}$ is invertible[1]. The compensator is of order $K - n$, and, as expected, results in the optimal trajectory x^* of Fig. 1. The closed-loop trajectories starting from any initial state αx_{ini}, aligned with x_{ini}, are optimal as well, e.g., the one starting from $-x_{\text{ini}}/2$ (represented in green in Fig. 1). However, closed-loop trajectories starting from non-nominal initial conditions may exhibit undesirable behaviour, e.g., the one represented in red which is obtained by applying the same dynamic controller starting from the initial condition $x(0) = \begin{bmatrix} 4 & 0.5 \end{bmatrix}^\top$. This is not surprising, since the ROC, albeit stabilizing, does not guarantee global optimality. Such a sensitivity on the initial conditions is undesirable and may vary according to the random choice of $\mathcal{Z}$. However, the problem can be avoided by employing a set of initial conditions, as stated in Section 4.

We choose a set of two linearly independent initial states $x_{\text{ini}}^{(1)} = \begin{bmatrix} 0 & 2 \end{bmatrix}^\top$, $x_{\text{ini}}^{(2)} = \begin{bmatrix} 2 & 0 \end{bmatrix}^\top$ (thus, the set of initial states spans the whole state-space $\mathbb{R}^2$), we set $K_1 = K_2 = K = 9$, and we solve the same minimum-energy control problem from each of the initial states. The resulting optimal open-loop control sequences are, respectively:

$$\begin{aligned} \mathcal{U}^{(1)} = [&-0.889 \;\; -0.722 \;\; -0.556 \;\; -0.389 \\ &-0.222 \;\; -0.056 \;\; 0.111 \;\; 0.278 \;\; 0.444] \end{aligned} \tag{29}$$

$$\begin{aligned} \mathcal{U}^{(2)} = [&-0.133 \;\; -0.1 \;\; -0.067 \;\; -0.033 \\ &0 \;\; 0.033 \;\; 0.067 \;\; 0.1 \;\; 0.133], \end{aligned} \tag{30}$$

thus leading to the optimal trajectories $\mathcal{X}^{(1)}$, and $\mathcal{X}^{(2)}$ shown respectively in blue and in green in Fig. 1. We therefore synthesize the dynamic compensator by applying (7), with $\mathcal{X} = \begin{bmatrix} \mathcal{X}^{(1)} & \mathcal{X}^{(2)} \end{bmatrix}$, and $\mathcal{Z} = \begin{bmatrix} \mathcal{Z}^{(1)} & \mathcal{Z}^{(2)} \end{bmatrix}$ randomly chosen such that $M = \begin{bmatrix} \mathcal{Z} \\ \mathcal{X} \end{bmatrix}$ is invertible. The obtained compensator is of order $2K - n$ and, besides resulting in the optimal trajectories from $x_{\text{ini}}^{(1)}$ and $x_{\text{ini}}^{(2)}$, it exhibits a better behaviour from non-nominal initial conditions, e.g., the red trajectory in Fig. 1 obtained starting from $x(0) = \begin{bmatrix} 4 & 0.5 \end{bmatrix}^\top$ (the same non-nominal initial condition of the red trajectory of Fig. 1).

Finally, we report results obtained by applying the proposed data-driven ROC approach to a cart and pole system (Fig. 2). We denote by $x = \begin{bmatrix} s & \dot{s} & \theta & \dot{\theta} \end{bmatrix}^\top$ the state vector, where s, $\dot{s}$, θ, $\dot{\theta}$ denote, respectively, the position and the speed of the cart, and the angular position and the angular speed of the pole.

The zero-order-hold sampling of the linearized system leads to the following state-space representation (the parameters are: mass of the cart 0.3 kg, mass of the pole 0.1 kg, length of the pole 1 m, gravity acceleration 9.81 m s^{-2}, friction neglected, sampling time

[1] To avoid numerical issues, we set a threshold on the condition number of the resulting matrix M.

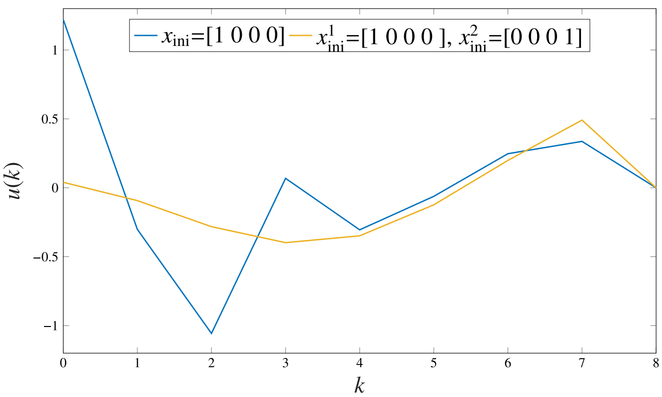


Fig. 4. Comparison of the control input trajectories obtained applying respectively the compensator of order $K-n$ (blue), and the compensator of order $2K-n$ (yellow) starting from the same $x_{ini} = [0 \quad 0 \quad 0 \quad 1]^\top$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.2 s):

$$x(k+1) = Ax(k) + Bu(k), \tag{31}$$

where

$$A = \begin{bmatrix} 1 & 0.2 & 0.06259 & 0.00425 \\ 0 & 1 & 0.59844 & 0.06259 \\ 0 & 0 & 0.74961 & 0.18301 \\ 0 & 0 & -2.3938 & 0.74961 \end{bmatrix}, \quad B = \begin{bmatrix} 0.065953 \\ 0.65251 \\ -0.06381 \\ -0.61004 \end{bmatrix}, \tag{32}$$

and u is the force F applied to the cart.

We set $K = 9$, performing $N = 20$ experiments lasting $2K = 18$ steps, starting from $x_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^\top$, and applying randomly chosen inputs. We compute the U, $\hat{U}$, X and $\hat{X}$ matrices of Eq. (19), and we repeat the same procedure used in the double integrator example. The closed-loop trajectory obtained by applying the compensator of order $K-n$, solution of Eq. (19), starting from $x_{ini} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^\top$ is reported in blue in Fig. 3.

Fig. 3 also shows, in red, the trajectory obtained from the non-nominal initial condition $x_{ini} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top$.

We then consider the two linearly independent initial states $x^{(1)}_{ini} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^\top$, $x^{(2)}_{ini} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top$ and obtain a compensator of order $2K-n$. The resulting closed-loop trajectories can be observed in Fig. 3. The red trajectory exhibits a better behaviour compared with the one of Fig. 3, obtained starting from the same initial state. The control actions are different as well, as it is clear from Fig. 4. Indeed, as expected, the input trajectory of the compensator of order $2K-n$ (yellow), starting from $x_{ini} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\top$, is characterized by less intense control actions than the one of the compensator of order $K-n$ (blue), starting from the same x_{ini}.

6. Conclusions

In this paper, we derived a novel data-driven approach to obtain closed-loop control laws from open-loop data-driven optimal control sequences. The approach is based on the dynamic ROC, which leads to a linear, dynamic, and globally stabilizing controller, starting from optimal state and control sequences. It can be applied whenever an open-loop, optimal control sequence is available that leads the system to zero from a given initial state. The approach can handle more than a single optimal control sequence and, differently, from the static approach [16], is not based on the partition of the state space, thus it is suitable for high-order systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P.J. Antsaklis, A.N. Michel, Linear Systems, Springer Science & Business Media, 2006.
[2] G. Baggio, D.S. Bassett, F. Pasqualetti, Data-driven control of complex networks, Nat. Commun. 12 (1) (2021) 1429, doi:10.1038/s41467-021-21554-0.
[3] G. Baggio, V. Katewa, F. Pasqualetti, Data-driven minimum-energy controls for linear systems, IEEE Control Syst. Lett. 3 (3) (2019) 589–594.
[4] G. Baggio, F. Pasqualetti, Learning minimum-energy controls from heterogeneous data, in: 2020 American Control Conference (ACC), IEEE, 2020, pp. 3991–3996.
[5] J. Berberich, J. Köhler, M.A. Müller, F. Allgöwer, Data-driven model predictive control with stability and robustness guarantees, IEEE Trans. Autom. Control 66 (4) (2020) 1702–1717.
[6] F. Blanchini, P. Colaneri, Y. Fujisaki, S. Miani, F.A. Pellegrino, A Youla–Kučera parameterization approach to output feedback relatively optimal control, Syst. Control Lett. 81 (2015) 14–23, doi:10.1016/j.sysconle.2015.04.006.
[7] F. Blanchini, F. Dabbene, G. Fenu, F.A. Pellegrino, E. Salvato, Model-free feedback control synthesis from expert demonstration, IEEE Control Syst. Lett. 7 (2023) 1604–1609, doi:10.1109/LCSYS.2023.3251575.
[8] F. Blanchini, F.A. Pellegrino, Relatively optimal control and its linear implementation, IEEE Trans. Autom. Control 48 (12) (2003) 2151–2162, doi:10.1109/TAC.2003.820070.
[9] F. Blanchini, F.A. Pellegrino, Relatively optimal control with characteristic polynomial assignment and output feedback, IEEE Trans. Autom. Control 51 (2) (2006) 183–191, doi:10.1109/TAC.2005.863493.
[10] F. Blanchini, F.A. Pellegrino, Relatively optimal control: a static piecewise-affine solution, SIAM J. Control Optim. 46 (2) (2007) 585–603, doi:10.1137/050643180.
[11] J. Coulson, J. Lygeros, F. Dörfler, Data-enabled predictive control: In the shallows of the DeePC, in: 2019 18th European Control Conference (ECC), IEEE, 2019a, pp. 307–312.
[12] J. Coulson, J. Lygeros, F. Dörfler, Regularized and distributionally robust data-enabled predictive control, in: 2019 IEEE 58th Conference on Decision and Control (CDC), IEEE, 2019b, pp. 2696–2701.
[13] C. De Persis, P. Tesi, Formulas for data-driven control: stabilization, optimality, and robustness, IEEE Trans. Autom. Control 65 (3) (2019) 909–924.
[14] C. De Persis, P. Tesi, Designing experiments for data-driven control of nonlinear systems, arXiv:2103.16509 (2021).
[15] C. De Persis, P. Tesi, Low-complexity learning of linear quadratic regulators from noisy data, Automatica 128 (2021) 109548, doi:10.1016/j.automatica.2021.109548.
[16] F.A. Pellegrino, F. Blanchini, G. Fenu, E. Salvato, Closed-loop control from data-driven open-loop optimal control trajectories, in: 2022 European Control Conference (ECC), IEEE, 2022, pp. 1379–1384.
[17] M. Rotulo, C. De Persis, P. Tesi, Data-driven linear quadratic regulation via semidefinite programming, IFAC-PapersOnLine 53 (2) (2020) 3995–4000.
[18] J.C. Willems, P. Rapisarda, I. Markovsky, B.L. De Moor, A note on persistency of excitation, Syst. Control Lett. 54 (4) (2005) 325–329.
[19] H. Yang, S. Li, A data-driven predictive controller design based on reduced Hankel matrix, in: 2015 10th Asian Control Conference (ASCC), IEEE, 2015, pp. 1–7.