

Data-Driven LQR Control Design

Gustavo R. Gonçalves da Silva, Alexandre S. Bazanella, Charles Lorenzini and Lucíola Campestrini

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

Abstract—This paper presents a data-driven solution to the **discrete-time infinite horizon LQR problem**. The state feedback gain is computed directly from a batch of input and state data collected from the plant. Simulation examples illustrate the convergence of the proposed solution to the optimal LQR gain as the number of **Markov parameters** tends to infinity. Experiments in an uninterruptible power supply are presented, which demonstrate the practical applicability of the design methodology.

Index Terms—Data-driven control, LQR control, Markov parameters, observability matrix.

I. INTRODUCTION

THE Linear Quadratic Regulator (LQR) design is a classical control problem whose analysis and solution can be found in most textbooks on control theory. It consists in computing the state feedback gain that optimizes a quadratic cost function of the plant's state and input. This computation of the gain requires the solution of a Riccati equation and is given as a function of the plant's state-space model. Closed-form (also called *batch-form*) solutions to the Riccati equation have also been provided [1], [2], and these are given as a function of the plant's Markov parameters. Whether applying the classical approach of explicitly solving the Riccati equation, or using a plant's state space description to calculate the Markov parameters and then feed them into the closed-form solution, this is a *model-based* design approach. That is, it is a design approach that is based on the knowledge of a good enough explicit model of the plant and on the use of this model in the control design following the *certainty equivalence* principle.

Data-Driven (DD) optimal control design methods have also been developed based on these closed-form solutions of the Riccati equation [3]–[5]. Although these DD design methods start from the LQR/LQG problem formulation, they do not calculate the state feedback control gain; instead, they directly estimate from data the optimal control input at each time instant. As such, they can not be said to solve the LQR problem in its classical formulation, and can mainly be cast within a predictive control framework.

Motivated by applications in which a state feedback is to be designed but a good enough model is not available and is of no interest *per se*, we present in this paper a DD approach to the solution of the LQR design. Otherwise stated, we provide a

DD solution for the computation of the optimal state feedback gain. In a DD control design, the controller structure is defined *a priori* and the controller's parameters are tuned with the use of a large batch of data, usually after such data are acquired. In most of the DD control literature, the controller structure consists of output feedback with a predefined transfer function with parameters to be tuned – such as a PID controller, for instance. The design itself is based, in most methods, on the Model Reference approach [6]. A control design based on the DD approach leads naturally to the automation of the design process, thus being extremely convenient for auto-tuning and self-tuning, and (all things being equal) also tends to outperform model-based designs, as shown in [7].

The infinite-horizon LQR problem fits this formulation perfectly: one has a fixed controller structure (the state feedback gain) with a few parameters to tune, and the controller must minimize a given quadratic performance criterion. The plant's model is just an intermediate step in the design and often has no interest in itself, the controller being the final and only objective. Thus, in this paper we provide a method to compute the LQR state feedback gain from data without the intermediate step of identifying a model of the system.

The paper is organized as follows. The LQR design problem is presented in Section II, along with its closed-form solution. It is shown that the computation of the LQR state feedback gain by the closed-form solution requires knowledge of two large matrices: an extended observability matrix and a Toeplitz matrix of the plant's Markov parameters. Then, in the ensuing sections III and IV, we provide algorithms to estimate these two matrices directly from data collected from the plant. In Section V we briefly review the Internal Model principle and the formulation of reference tracking as a state feedback problem. Two simulation examples are given in Section VI to illustrate the method's properties. One of our motivating applications – the control of uninterruptible power sources – is explored in Section VII, where we present a practical application of our design methodology. It is seen in the experimental results that our design compares favorably with previously presented model-based solutions to this same practical problem. Finally, concluding remarks are given in Section VIII.

II. PROBLEM STATEMENT

Consider a linear time-invariant discrete-time system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \quad (1)$$

where $x(k)$ is an n -dimension state vector, $u(k)$ is a p -dimension input vector and $y(k)$ a q -dimension output vector.

The authors are with the Department of Automation and Energy – Federal University of Rio Grande do Sul. Av. Osvaldo Aranha, 103 – CEP: 90035-190 – Porto Alegre, RS – Brazil

E-mails {gustavo.rgs, bazanella, charles.lorenzini, luciola}@ufrgs.br.

This work was supported by the National Council for Scientific and Technological Development – CNPq/BR and by the Brazilian Federal Agency for Support and Evaluation of Graduate Education – CAPES/BR.

The *infinite horizon* LQR control problem can be summarized as follows: find the optimal state feedback gain K of the control law

$$u(k) = -Kx(k) \quad (2)$$

such that the quadratic cost function

$$J = \sum_{k=0}^{\infty} (y(k)^T Q y(k) + u(k)^T R u(k)) \quad (3)$$

is minimized subject to system (1), where Q and R are positive definite symmetric weighting matrices. The optimal gain is given by

$$K = (R + B^T P B)^{-1} (B^T P A) \quad (4)$$

where P is the unique positive definite solution to the discrete time algebraic Riccati equation (DARE)

$$P = A^T P A - (A^T P B)(R + B^T P B)^{-1} (B^T P A) + C^T Q C. \quad (5)$$

A closed-form solution to the DARE (5) has been reported in [1], [2]. For a sufficient large N , this solution can be written as

$$P = \mathbf{O}^T (\mathbf{Q}_{N+1}^{-1} + \mathbf{S} \mathbf{R}_{N+1}^{-1} \mathbf{S}^T)^{-1} \mathbf{O}, \quad (6)$$

where

$$\mathbf{O} = [C^T \quad (CA)^T \quad (CA^2)^T \quad \dots \quad (CA^N)^T]^T, \quad (7)$$

$$\mathbf{S} = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ CB & 0 & \ddots & & \vdots \\ CAB & CB & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{N-1}B & CA^{N-2}B & \dots & CB & 0 \end{bmatrix},$$

$$\mathbf{R}_j \triangleq \text{diag}(R, R, \dots, R), \quad \mathbf{Q}_j \triangleq \text{diag}(Q, Q, \dots, Q),$$

with \mathbf{R}_j and \mathbf{Q}_j containing $j = N + 1$ diagonal blocks each.

The matrix \mathbf{O} is an *extended observability matrix* for system (1) and \mathbf{S} is a Toeplitz matrix of its Markov parameters:

$$M^{(i)} = CA^{(i-1)}B, \quad i = 1, 2, \dots, N. \quad (8)$$

As shown in [1], using (6) in (4) and rearranging some terms, the LQR state feedback gain K can be computed as a function of the Markov parameters as

$$K = [R + \mathbf{M}^T (\mathbf{Q}_N^{-1} + \mathbf{S} \mathbf{R}_N^{-1} \mathbf{S}^T)^{-1} \mathbf{M}]^{-1} \times \mathbf{M}^T (\mathbf{Q}_N^{-1} + \mathbf{S} \mathbf{R}_N^{-1} \mathbf{S}^T)^{-1} \mathbf{O}^+, \quad (9)$$

where

$$\mathbf{M} = \begin{bmatrix} M^{(1)} \\ M^{(2)} \\ \vdots \\ M^{(N)} \end{bmatrix} \in \mathbb{R}^{qN \times p}, \quad (10)$$

$$\mathbf{S} = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ M^{(1)} & 0 & \ddots & & \vdots \\ M^{(2)} & M^{(1)} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ M^{(N-1)} & M^{(N-2)} & \dots & M^{(1)} & 0 \end{bmatrix} \in \mathbb{R}^{qN \times pN}, \quad (11)$$

$$\mathbf{O}^+ = [(CA)^T \quad (CA^2)^T \quad \dots \quad (CA^N)^T]^T \in \mathbb{R}^{qN \times n}. \quad (12)$$

Notice that gain K in (4) is a function of A, B , and P , which is also a function of system matrices A, B and C . Now we have an expression that depends basically on \mathbf{S} and \mathbf{O}^+ . If these quantities could be obtained from data, then a data-driven method can be formulated. So, in order to succeed in this data-driven approach, we need to identify the system's Markov parameters – the matrix \mathbf{S} – and an extended observability matrix \mathbf{O}^+ .

Thus, let us pose the problem formally:

Given the data set

$$Z^T = [u(0), u(1), \dots, u(T), y(0), y(1), \dots, y(T), x(0), x(1), \dots, x(T)] \quad (13)$$

find the optimal state feedback gain K as in (9). To do so, a sequence of N Markov parameters and the one-step ahead N extended observability matrix must be estimated from data.

In the sequel we present a procedure to obtain both the Markov parameters and the extended observability matrix without using a model for the system.

III. MARKOV PARAMETERS ESTIMATION

We review next the estimation of the system Markov parameters via the so-called ARMarkov/Toeplitz models¹. We follow the description in [8].

Let the states be repeatedly substituted $N + 1$ times in (1); then

$$x(k + N + 1) = A^{N+1}x(k) + \mathbf{C} \mathbf{u}_m(k)$$

$$\begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+N) \end{bmatrix} = \mathbf{O}x(k) + \mathbf{S} \underbrace{\begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+N) \end{bmatrix}}_{\mathbf{u}_m(k)} \quad (14)$$

where $\mathbf{C} = [A^N B \quad \dots \quad AB \quad B]$.

According to [9], as long as $(N + 1)q \geq n$, it is guaranteed for an observable system that there exists a matrix \mathbf{F} such that $A^{N+1} + \mathbf{F}\mathbf{O} = 0$, which ensures that there exists an expression

¹We remark that estimating the system's Markov parameters is equivalent to identifying an N -th order FIR representation for the system.

where the state is eliminated from (14). This allows to write a predictor for the system's output as follows.

Let the Hankel matrix of a signal $e(k)$ be defined as

$$H(e(k)) \triangleq \begin{bmatrix} e(k) & e(k+1) & \cdots & e(k+L-1) \\ e(k+1) & e(k+2) & \cdots & e(k+L) \\ \vdots & \vdots & \ddots & \vdots \\ e(k+N-1) & e(k+N) & \cdots & e(k+N+L-2) \end{bmatrix}. \quad (15)$$

Define the set of data matrices

$$\begin{aligned} U_p &= H(u(0)) & U_f &= H(u(N)) \\ Y_p &= H(y(0)) & Y_f &= H(y(N)). \end{aligned} \quad (16)$$

Then a predictor of the system output can be written as [8]

$$Y_f = \underbrace{[\mathbf{O}(\mathbf{C} + \mathbf{F}\mathbf{S}) \quad -\mathbf{O}\mathbf{F} \quad \mathbf{S}]}_{\mathbf{W}} \underbrace{\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}}_{\Phi}. \quad (17)$$

Thus, an estimate $\widehat{\mathbf{W}}$ of \mathbf{W} can be obtained by solving the least-squares problem

$$\mathbf{W} = Y_f \Phi^\dagger \quad (18)$$

with $[\cdot]^\dagger$ denoting the Moore-Penrose pseudo-inverse, and an estimate $\widehat{\mathbf{S}}$ as the rightmost pN columns of $\widehat{\mathbf{W}}$. Moreover, \mathbf{M} can be extracted from the first column of $\widehat{\mathbf{S}}$. This estimation has been shown to be consistent for the Markov parameters [10].

For stable systems, the choice of N is closely related to the system's open loop settling time, as the contribution of A^N decreases as N increases. Also, L must satisfy $L \geq 3q(N+1)$ so (18) has a solution.

IV. EXTENDED OBSERVABILITY MATRIX ESTIMATION

Since a state feedback control is to be implemented, we can assume that the states are measurable. Hence, in this section we present two original algorithms to identify an extended observability matrix in the same state coordinates we are measuring and later we discuss their properties. We define the vector of measured states by

$$\mathbf{X} \triangleq [x(0) \ x(1) \ x(2) \ \cdots \ x(L-1)]. \quad (19)$$

A. Algorithm 1

The output equation (14) can also be written with Hankel matrices without eliminating the state vector. It can be put in an extended output matrix equation as [8]

$$Y_p = \mathbf{O}\mathbf{X} + \mathbf{S}U_p. \quad (20)$$

Since an estimate $\widehat{\mathbf{S}}$ can be obtained using the algorithm provided in Section III, and U_p and Y_p can be formed with collected data, we can then solve the system of equations in (20) for \mathbf{O} :

$$\widehat{\mathbf{O}} = (Y_p - \widehat{\mathbf{S}}U_p)\mathbf{X}^\dagger. \quad (21)$$

To obtain \mathbf{O}^+ , one simply removes the first q rows of \mathbf{O} .

B. Algorithm 2

First, define U_{po} as the geometric operator that projects the row space of a matrix onto the orthogonal complement of the row space of the matrix U_p

$$U_{po} \triangleq I_L - U_p^T (U_p U_p^T)^{-1} U_p \quad (22)$$

where I_L is an identity matrix of size L . Then by post-multiplying the extended output matrix (20) by U_{po} , we have

$$\begin{aligned} Y_p U_{po} &= \mathbf{O}\mathbf{X}U_{po} + \mathbf{S}U_p U_{po} \\ Y_p U_{po} &= \mathbf{O}\mathbf{X}U_{po} \end{aligned} \quad (23)$$

Notice that by using the projection U_{po} we eliminate the need to know, or estimate, \mathbf{S} .

An estimator of the extended observability matrix can thus be computed as

$$\widehat{\mathbf{O}} = (Y_p U_{po})(\mathbf{X}U_{po})^\dagger. \quad (24)$$

C. Estimates properties

The estimates just provided allow the computation of the state feedback gain according to (9) and, as shown in [1], the gain thus computed converges asymptotically, as $N \rightarrow \infty$, to the optimal LQR state feedback. A simulation example in Section VI illustrates this property.

On the other hand, when the state measurement is corrupted by noise, one can expect some bias in the estimates of the extended observability matrix, since both solutions presented are of a least squares nature. We now briefly discuss the bias and covariance of these estimates. Consider the system state-space representation (1) with the noise terms

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ev(k) \\ y(k) &= Cx(k) + Fw(k) \end{aligned} \quad (25)$$

with $v(k)$ and $w(k)$ white noise sequences. We can write the extended output equation as

$$Y_p = \mathbf{O}\mathbf{X} + \mathbf{S}U_p + \mathbf{S}_E V_p + F W_p \quad (26)$$

where $V_p = H(v(0))$ and $W_p = H(w(0))$ are Hankel matrices of the noise sequences $v(k)$ and $w(k)$ respectively, and \mathbf{S}_E has the same structure as \mathbf{S} with E in lieu of B . We can also write

$$\mathbf{X} = \mathbf{X}^o + \mathbf{E}\mathbf{V} \quad (27)$$

where \mathbf{V} is a row vector of the noise sequences $v(k)$. We also assume the sequence $u(k)$ uncorrelated with $v(k)$.

1) *Algorithm 1*: Let $\mathcal{E}[\cdot]$ denote the expected value function. The bias of the first algorithm is given by

$$\mathcal{B}(\widehat{\mathbf{O}}) = \mathcal{E}[\widehat{\mathbf{O}}] - \mathbf{O} = \mathcal{E}[(Y_p - \widehat{\mathbf{S}}U_p)\mathbf{X}^\dagger - \mathbf{O}]. \quad (28)$$

Inserting (26) into (28)

$$\begin{aligned} \mathcal{B}(\widehat{\mathbf{O}}) &= \mathcal{E}[(\mathbf{O}\mathbf{X} + \mathbf{S}U_p + \mathbf{S}_E V_p + F W_p - \widehat{\mathbf{S}}U_p)\mathbf{X}^\dagger - \mathbf{O}] \\ &= \mathcal{E}[\mathbf{O}(\mathbf{X}\mathbf{X}^\dagger - I) + (\mathbf{S} - \widehat{\mathbf{S}})U_p\mathbf{X}^\dagger + \mathbf{S}_E V_p\mathbf{X}^\dagger \\ &\quad + F W_p\mathbf{X}^\dagger] \end{aligned} \quad (29)$$

First term of (29) is null since \mathbf{X} has real entries, and the second term is null either if $\widehat{\mathbf{S}} = \mathbf{S}$ or because $u(k)$ is uncorrelated with $v(k)$. We then have

$$\mathcal{B}(\widehat{\mathbf{O}}) = \mathbf{S}_E \mathcal{E}[V_p\mathbf{X}^\dagger] + F \mathcal{E}[W_p\mathbf{X}^\dagger]. \quad (30)$$

If we further assume that $v(k)$ is uncorrelated with $w(k)$, then

$$\mathcal{B}(\hat{\mathbf{O}}) = \mathbf{S}_E \mathcal{E}[V_p \mathbf{X}^\dagger]. \quad (31)$$

The covariance of the first algorithm is given by

$$\begin{aligned} \mathcal{V}(\hat{\mathbf{O}}) &= \mathcal{E}[(\hat{\mathbf{O}} - \mathcal{E}[\hat{\mathbf{O}}])(\hat{\mathbf{O}} - \mathcal{E}[\hat{\mathbf{O}}])^T] \\ &= \mathcal{E}[(\mathbf{S} - \hat{\mathbf{S}})U_p \mathbf{X}^\dagger + \mathbf{S}_E V_p \mathbf{X}^\dagger + F W_p \mathbf{X}^\dagger \\ &\quad - \mathbf{S}_E \mathcal{E}[V_p \mathbf{X}^\dagger] - F \mathcal{E}[W_p \mathbf{X}^\dagger]] \\ &\quad \times ((\mathbf{S} - \hat{\mathbf{S}})U_p \mathbf{X}^\dagger + \mathbf{S}_E V_p \mathbf{X}^\dagger + F W_p \mathbf{X}^\dagger \\ &\quad - \mathbf{S}_E \mathcal{E}[V_p \mathbf{X}^\dagger] - F \mathcal{E}[W_p \mathbf{X}^\dagger])^T] \end{aligned} \quad (32)$$

which after some algebraic manipulation results in

$$\begin{aligned} \mathcal{V}(\hat{\mathbf{O}}) &= \mathcal{E} \left\{ [(\mathbf{S} - \hat{\mathbf{S}})U_p \mathbf{X}^\dagger + \mathbf{S}_E V_p \mathbf{X}^\dagger + F W_p \mathbf{X}^\dagger] \Lambda_1^T \right. \\ &\quad \left. + [(\mathbf{S} - \hat{\mathbf{S}})U_p \mathbf{X}^\dagger + \Lambda_1][(\mathbf{S} - \hat{\mathbf{S}})U_p \mathbf{X}^\dagger]^T \right\} \end{aligned} \quad (33)$$

$$\Lambda_1 = \mathbf{S}_E (V_p \mathbf{X}^\dagger - \mathcal{E}[V_p \mathbf{X}^\dagger]) + F (W_p \mathbf{X}^\dagger - \mathcal{E}[W_p \mathbf{X}^\dagger]) \quad (34)$$

2) *Algorithm 2*: Following the same steps as in Algorithm 1, the following expressions are obtained for the bias and covariance of the estimates given by Algorithm 2.

$$\mathcal{B}(\hat{\mathbf{O}}) = \mathbf{S}_E \mathcal{E}[(V_p U_{po})(\mathbf{X} U_{po})^\dagger]. \quad (35)$$

$$\begin{aligned} \mathcal{V}(\hat{\mathbf{O}}) &= \mathcal{E} \{ \mathbf{B}[\mathbf{B} + \mathbf{S}_E V_p \mathbf{X}^\dagger + F W_p \mathbf{X}^\dagger]^T \\ &\quad + [\mathbf{S}_E V_p \mathbf{X}^\dagger + F W_p \mathbf{X}^\dagger] \mathbf{B}^T \} \end{aligned} \quad (36)$$

$$\begin{aligned} \Lambda_2 &= \mathbf{S}_E [(V_p U_{po})(\mathbf{X} U_{po})^\dagger - \mathcal{E}[(V_p U_{po})(\mathbf{X} U_{po})^\dagger]] \\ &\quad + F [(W_p U_{po})(\mathbf{X} U_{po})^\dagger - \mathcal{E}[(W_p U_{po})(\mathbf{X} U_{po})^\dagger]] (\mathbf{X} U_{po})^\dagger \end{aligned} \quad (37)$$

Notice that, as expected, the bias of both estimates is inversely proportional to the signal to noise ratio and the estimates will be unbiased only if there is no noise in the measurement. We provide an illustrative example in Section VI-B to compare the bias, covariance and – most importantly – the mean square error resulting from the two algorithms.

V. INTERNAL MODEL CONTROLLER AND AUGMENTED STATE SPACE

Most practical control applications consider the reference tracking problem. Thus, in this section we show how to use open-loop data to obtain an augmented state and output vectors in order to adjust the gains also for reference tracking considering a feedback loop with an internal model controller (IMC) [11]. Let

$$\mathbf{C} = (\mathbf{A}_c, \mathbf{B}_c) \quad (38)$$

be a discrete-time realization of the internal model controller state equation, whose n_c states are measurable. Assume that every output of the system is to follow a reference represented

by the internal model controller. Then the augmented open-loop space-state representation of system (1) with controller (38) is given by

$$\begin{aligned} x_a(k+1) &= \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{0}_{n,n_c \times q} \\ -\mathbf{C} \otimes \mathbf{B}_c & \mathbf{I}_q \otimes \mathbf{A}_c \end{bmatrix}}_{\mathbf{A}_a} x_a(k) + \underbrace{\begin{bmatrix} \mathbf{B} \\ \mathbf{0}_{n_c \times q, p} \end{bmatrix}}_{\mathbf{B}_a} u(k) \\ y_a(k) &= \underbrace{\begin{bmatrix} \mathbf{C} & \mathbf{0}_{q,n_c \times q} \\ \mathbf{0}_{n_c \times q, n} & \mathbf{I}_{n_c \times q} \end{bmatrix}}_{\mathbf{C}_a} x_a(k) \end{aligned} \quad (39)$$

where \otimes is the Kronecker product.

From partition (2, 1) of \mathbf{A}_a we see that the open-loop IMC states needed to compute the gain \mathbf{K} can be obtained by simply filtering the plant outputs by $-\mathbf{C}$.

For example, the *state equation* of the integrator $\frac{1}{z-1}$ can be represented by

$$x_c(k+1) = \underbrace{1}_{\mathbf{A}_c} x_c(k) + \underbrace{1}_{\mathbf{B}_c} (r(k) - y(k)) \quad (40)$$

and a resonant controller at frequency ω_n with a pre-warping Tustin representation can be realized as

$$x_c(k+1) \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 2 \cos(\omega_n T_s) \end{bmatrix}}_{\mathbf{A}_c} x_c(k) + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\mathbf{B}_c} (r(k) - y(k)) \quad (41)$$

where T_s is the sampling time.

Let $x_{IMC}(t)$ represent the state of the open-loop IMC², and the augmented output and state vector be given by $y_a(t) = [y(t) \ x_{IMC}(t)]^T$ and $x_a(t) = [x(t) \ x_{IMC}(t)]^T$, respectively. These are the vectors that should be used, along with $u(t)$, to estimate the Markov parameters and the extended observability matrix.

VI. SIMULATION EXAMPLES

A. Regulation control

In this first example we illustrate the convergence of the proposed method. Consider a system as in (1) whose matrices are given by

$$\mathbf{A} = \begin{bmatrix} 1 & 0.15 \\ -0.2 & 0.6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0.04 & 0.01 \\ 0.02 & -0.01 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}. \quad (42)$$

The sampling time is $T_s = 1$ s. Also, let the performance requirements be given by $\mathbf{R} = 0.2 \mathbf{I}_q$ and $\mathbf{Q} = 20 \mathbf{I}_q$, where \mathbf{I}_q is the identity matrix of size q , that is, we are valuing more the evolution of the states than the control effort. The *model-based* optimal LQR controller is given by

$$\mathbf{K} = \begin{bmatrix} 4.6491 & 7.5226 \\ 1.4461 & -1.9886 \end{bmatrix}. \quad (43)$$

In order to apply the proposed methodology we set an open-loop experiment where the input signal is a PRBS with amplitude 1 and length 1022 and the output data was collected, i.e., both $y(k)$ and $x(k)$. We identified two state-feedback gains: one with $N = 10$ and the other with $N = 50$, the latter been

²which can be obtained with MATLAB command `lsim`; for instance, `xIMC=lsim(IMC,y)`.

close to the system's open loop settling time (approximately 45 samples). We obtained $\hat{K}_{10} = \begin{bmatrix} 4.2314 & 7.644 \\ 1.127 & -1.8959 \end{bmatrix}$ and $\hat{K}_{50} = \begin{bmatrix} 4.6491 & 7.5226 \\ 1.4461 & -1.9886 \end{bmatrix}$.

Notice that with $N = 50$, \hat{K} equals the model-based solution (43) up to the fifth significant digit.

B. The noisy case: observability matrix properties

We provide now a simple example to illustrate bias and covariance of the observability matrix estimators. Consider a system with state-space matrices

$$A = 0.14, \quad B = 1.72, \quad C = 1, \quad E = 1, \quad F = 0, \quad (44)$$

and let the LQR performance matrices be $R = 0$ and $Q = 1$, so we are aiming for a dead-beat control, and convergence can be found with $N_m = 3$ (due to row removal, that means the extended observability matrix will be of size 2). The actual extended observability matrix is then $O = [0.14 \ 0.0196]^T$.

We set a Monte Carlo experiment with 5000 runs and with a PRBS input of length 1022 and $e(t)$ as white noise with variance $\sigma^2 = 0.1$. Fig. 1 portrays the results obtained with both estimators (21) and (24).

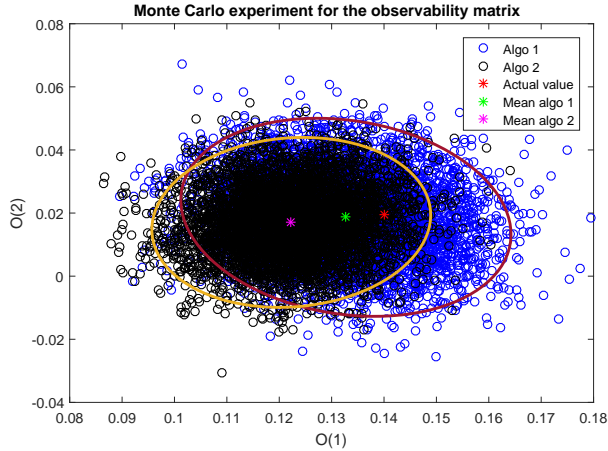


Fig. 1. Estimates of the observability matrix in a Monte Carlo experiment. The ellipses represent the covariance regions around the mean value with 95% confidence.

As mentioned before, the estimate bias with Algorithm 1 is smaller, whereas with Algorithm 2 a smaller covariance is achieved. In fact, we obtained $\mathcal{E}(\hat{O}_1) = [0.1327 \ 0.01865]^T$ and $\mathcal{E}(\hat{O}_2) = [0.1223 \ 0.01704]^T$, and the eigenvalues λ of the covariance matrices $\lambda[\mathcal{E}(\hat{V}_1)] = [1.365 \ 1.938]^T \times 10^{-4}$ and $\lambda[\mathcal{E}(\hat{V}_2)] = [1.093 \ 1.3]^T \times 10^{-4}$. Notice that the largest eigenvalue with Algorithm 2 is approximately the smallest eigenvalue with Algorithm 1.

We also computed the eigenvalues of the MSE matrix of both algorithm and obtained $\lambda(MSE_1) = [1.723 \times 10^{-4} \ 1.812 \times 10^{-2}]^T$ and $\lambda(MSE_2) = [1.182 \times 10^{-4} \ 1.537 \times 10^{-2}]^T$. Note that Algorithm 2 provides much smaller MSE.

VII. CONTROL OF AN UPS

We now consider a practical application of the proposed methodology to an uninterruptible power supply (UPS). This plant has been studied before in [12], [13].

Consider the simplified electrical diagram of the output stage of a single-phase UPS system, as illustrated in Figure 2.

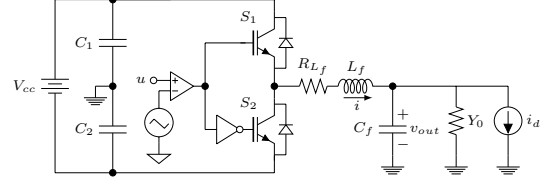


Fig. 2. Schematic representation of the UPS with load.

The load effect on the system output is modeled by a parallel connection of an uncertain admittance $Y_0(t)$ and an unknown periodic disturbance given by the current source $i_d(t)$. The PWM (Pulse Width Modulation) comparator input is modeled by a gain K_{PWM} multiplied by the control input. Also, defining the system state vector as the inductor current and the capacitor voltage, $x(t) = [i(t) \ v(t)]^T$, the *continuous-time* state-space representation for the UPS system is given by:

$$\begin{aligned} \dot{x}(t) &= \underbrace{\begin{bmatrix} \frac{-R_{Lf}}{L_f} & \frac{-1}{L_f} \\ \frac{1}{C_f} & \frac{-Y_0(t)}{C_f} \end{bmatrix}}_A x(t) + \underbrace{\begin{bmatrix} \frac{K_{PWM}}{L_f} \\ 0 \end{bmatrix}}_B u(t) + \underbrace{\begin{bmatrix} 0 \\ \frac{-1}{C_f} \end{bmatrix}}_{B_d} i_d(t) \\ y(t) &= \underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix}}_C x(t) \end{aligned} \quad (45)$$

where $u(t)$ is the PWM control input and $y(t)$ is the output voltage to be controlled.

Closed-loop reference is typically a sinusoid and for our case we have $r(t) = 127\sqrt{2}\sin(120\pi t)$. Since the reference signal is a sinusoid, then the right choice of the IMC is a resonant controller (41). Admittance $Y_0(t)$ can be set as an open circuit (no load), as a nominal resistance $R_0 = 6 \ \Omega$ and as a non-linear load given by a full-bridge circuit.

The control objective can be summarized as: *design a data-driven state feedback controller for sinusoid reference tracking for the UPS operating with non-linear load.*

In order to obtain meaningful data from the system, we set an open-loop experiment as follows. The sampling time was set to $T_s = 1/15000$ s; the input of the PWM was set as a PRBS with amplitude ± 104 V and with length 75000 samples (i.e., a 5 seconds signal); current and voltage were measured as portrayed (zoomed time scale) in Fig. 3 for the UPS operating with its *non-linear load*. The output voltage was filtered by (41) to obtain the IMC states for our proposed algorithm.

The parameters chosen for the LQR were $Q = 200I_3$ and $R = 5000$, that is, we strongly penalized the control signal as to try to achieve closed-loop stability even when there is no load in the UPS and to reduce sensibility due to noise, specially for the current measurement. Prior to any knowledge

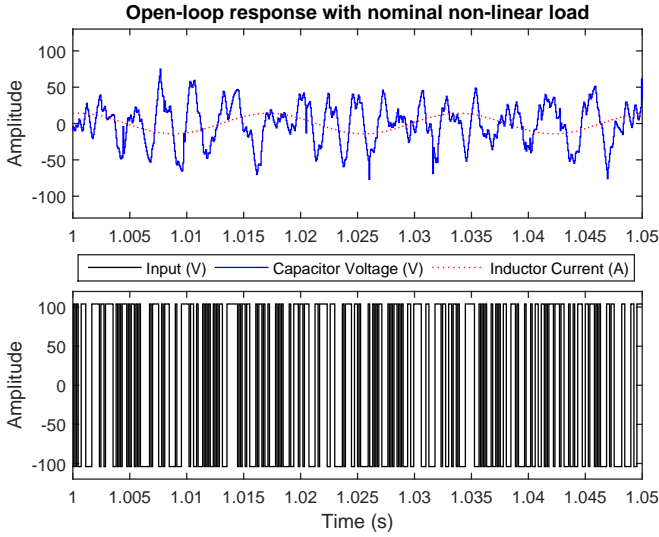


Fig. 3. Open-loop response with nominal non-linear load.

about the system settling time, we also selected $N = 150$. The obtained LQR gain is

$$K = [4.85548 \quad 5.54514 \quad 0.638479 \quad -0.644019]. \quad (46)$$

Fig. 4 shows the closed-loop response for the UPS operating at nominal capacity with non-linear load. Stability and reference tracking were achieved with corresponding Total Harmonic Distortion (THD) of 11.7% – result similar to the one obtained in [12], in which the state feedback gain was designed using a full plant model and a Linear Matrix Inequality approach.

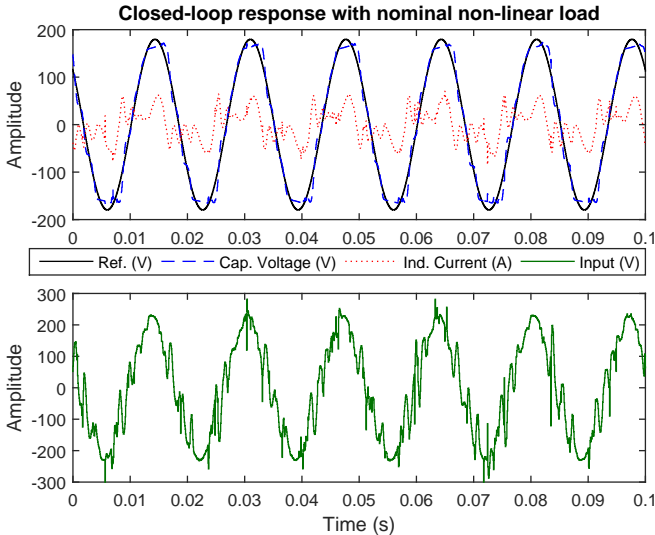


Fig. 4. Closed-loop response with nominal non-linear load.

We also applied the obtained controller to different scenarios: (a) the open circuit (no load) case and (b) the nominal linear load R_0 . Fig. 5 shows the closed-loop responses for cases (a) and (b) respectively. Closed-loop stability was achieved and with very small THD – 1.4% and 1.8% respectively

–, even though the controller was not designed for these scenarios.

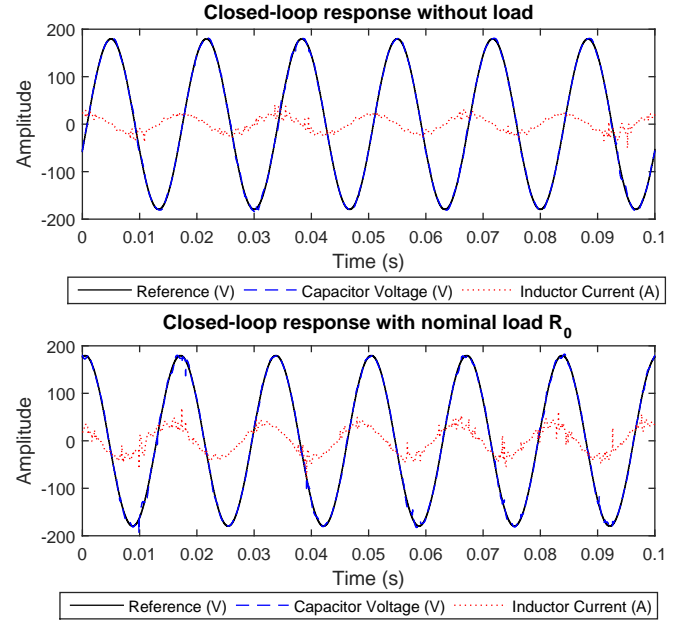


Fig. 5. Closed-loop response with no output load and with nominal linear load.

Notice that with this approach we obtained a linear state-feedback gain with only one experiment on the plant, even though the actual plant has a strong nonlinear behavior and a single linear model would not describe the system with reasonable accuracy. If data were used to identify a plant model, then more than one experiment would be necessary in order to evaluate a plant model and an uncertainty matrix.

VIII. CONCLUSION

In this paper we provided a data-driven method to compute the infinite horizon LQR state feedback gain, without identifying a model of the plant. In our method, the feedback gain is computed from a batch of data and converges to the infinite horizon LQR gain as the amount of data and, by consequence, the number of estimated Markov parameters grow. Simulation examples illustrated the convergence of the method and an experimental application to an UPS showed its practical applicability.

REFERENCES

- [1] F. Lewis, “A generalized inverse solution to the discrete-time singular Riccati equation,” *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 395–398, 1981.
- [2] K. Furuta and M. Wongsaisuwan, “Closed-form solutions to discrete-time LQ optimal control and disturbance attenuation,” *Systems & control letters*, vol. 20, no. 6, pp. 427–437, 1993.
- [3] —, “Discrete-time LQG dynamic controller design using plant Markov parameters,” *Automatica*, vol. 31, no. 9, pp. 1317–1324, 1995.
- [4] R. E. Skelton and G. Shi, “The data-based LQG control problem,” in *Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on*, vol. 2, 1994, pp. 1447–1452.
- [5] W. Aangenent, D. Kostic, B. de Jager, R. van de Molengraft, and M. Steinbuch, “Data-based optimal control,” in *Proceedings of the American Control Conference*, vol. 2, 2005, pp. 1460–1465.

- [6] A. S. Bazanella, L. Campestri, and D. Eckhard, *Data-driven controller design: the H2 approach*. Netherlands: Springer Science & Business Media, 2011.
- [7] L. Campestri, D. Eckhard, A. S. Bazanella, and M. Gevers, "Data-driven model reference control design by prediction error identification," *Journal of the Franklin Institute*, vol. 354, no. 6, pp. 2828–2647, 2017.
- [8] B. De Moor, M. Moonen, L. Vandenberghe, and J. Vandewalle, "Identification of linear state space models with singular value decomposition using canonical correlation concepts," in *International Workshop on SVD and signal processing*. Grenoble, France: North Holland: Elsevier Science Publishers, 1988, pp. 161–169.
- [9] R. K. Lim, M. Q. Phan, and R. W. Longman, "State estimation with AR-Markov models," Department of mechanical and aerospace engineering, Princeton University, Princeton, NJ, Tech. Rep. 3046, October 1998.
- [10] M. Kamrunnagar, B. Huang, and D. Fisher, "Estimation of Markov parameters and time-delay/interactor matrix," *Chemical Engineering Science*, vol. 55, no. 17, pp. 3353–3363, 2000.
- [11] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, no. 5, pp. 457–465, 1976.
- [12] L. F. A. Pereira, J. V. Flores, G. Bonan, D. F. Coutinho, and J. M. G. da Silva, "Multiple resonant controllers for uninterruptible power supplies – a systematic robust control design approach," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 3, pp. 1528–1538, 2014.
- [13] C. Lorenzini, J. V. Flores, L. F. A. Pereira, A. T. Salton, and R. S. Castro, "Repetitive controller with low-pass filter compensation applied to uninterruptible power supplies (UPS)," in *Industrial Electronics Society, IECON 2015-41st Annual Conference of the IEEE*. Yokohama, Japan: New York: IEEE, 2015, pp. 003 551–003 556.