Proceedings of Machine Learning Research vol 168:1–12, 2022 4th Annual Conference on Learning for Dynamics and Control

Noise handling in data-driven predictive control: a strategy based on dynamic mode decomposition

Andrea Sassella ANDREA.SASSELLA@POLIMI.IT
Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milano, Italy

Valentina Breschi VALENTINA.BRESCHI@POLIMI.IT
Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milano, Italy

Simone Formentin SIMONE.FORMENTIN@POLIMI.IT
Dipartimento di Elettronica e Informazione, Politecnico di Milano, Milano, Italy

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

A major issue when exploiting data for direct control design is noise handling, since overlooking or improperly treating noise might have a catastrophic impact on closed-loop performance. Nonetheless, standard approaches to mitigate its effect might not be easily applicable for data-driven control design, since they often require tuning a set of hyper-parameters via potentially unsafe closed-loop experiments. By focusing on data-driven predictive control, we propose a noise handling approach based on truncated dynamic mode decomposition, along with an automatic tuning strategy for its hyper-parameters. By leveraging on pre-processing only, the proposed approach allows one to avoid dangerous closed-loop calibrations, while being effective in coping with noise, as illustrated on a benchmark simulation example.

Keywords: Data-driven predictive control; dynamic mode decomposition; noise handling.

1. Introduction

The possibility to cut corners by avoiding the modeling/identification of the plant made *data-driven control* an appealing alternative to traditional model-based design approaches (see Formentin et al. (2014) for a comparison between these two paradigms). In particular, translating the ideas of *model predictive control* (MPC) into a purely data-based framework, *data-driven predictive control* (DDPC) allows to combine the advantages of constrained control with the ones of direct design strategies. Indeed, as standard MPC, DDPC allows one to find the optimal control action in a receding horizon fashion, while accounting for possible constraints on the control input and the system output. Nonetheless, DDPC does not require a parametric model for the system under control to be known, being instead grounded on a behavioral description of the system Willems et al. (2005), directly obtainable by a set of matrices constructed from experimental data. The first examples of DDPC for Linear Time Invariant (LTI) systems appeared in Coulson et al. (2019a); Berberich et al. (2020). Since then, several extensions have been proposed, including the one to nonlinear systems Berberich et al. (2021) and the derivation of explicit data-driven predictive controllers in Sassella et al. (2021); Breschi et al. (2021).

A key requirement for these techniques to result in satisfactory performance is a proper handling of noisy measurements. Indeed, real data are always corrupted by noise and, consequently, this affects the data-based predictor which DDPC strategies rely on. Noise treatment in data-driven control has been studied in Berberich et al. (2020); Coulson et al. (2019a,b), with Berberich et al. (2020); Coulson et al. (2019a) focusing on the case of bounded noise, while Coulson et al. (2019b) tackling noise handling from a probabilistic perspective. Ultimately, all these strategies result in the introduction of regularization terms in the performance-oriented cost of DDPC, as comprehensively summarized in Dörfler et al. (2021). Even though these regularization-based approaches have shown to be effective, they require a careful selection of the regularization penalties to properly weight the additional terms in the cost, whose choice is crucial to obtain satisfactory closed-loop performance. However, little to no guidelines are provided on how to tune these important parameters, with their choice generally performed in closed-loop, according to a *cross-validation* rationale. Such a calibration procedure can be quite critical, possibly endangering the system and/or requiring a lengthy iterative tuning period, which might be unfeasible when high performance are required. By relying on the data-based predictor introduced in De Persis and Tesi (2021), an alternative approach to handle noise is presented in Sassella et al. (2021), which relies on performing several replicas of the same experiment to average out noise from data. Even if this approach has proven its effectiveness, repeated experiments are generally costly, with the imposed constraints on the data collection campaign being likely limiting in practice.

To overcome the problems related to hyper-parameter tuning characterizing regularization-based strategies and the experimental limitations imposed by the averaging approach in Sassella et al. (2021), in this paper we propose to handle measurement noise via a pre-processing phase relying on truncated *dynamic mode decomposition* (DMD) Kutz et al. (2016). Firstly derived to decompose complex flows into simpler spatio-temporal data structures Schmid (2010), DMD has also proven to be effective in extracting the relevant dynamics and input/output features of a dynamic system from data Proctor et al. (2016). This approach allows us to avoid the introduction of additional terms in the cost, with the management of noise moved prior to the actual control design phase, along the same line of Sassella et al. (2021). Nonetheless, differently from the averaging strategy, this approach for noise handling does not severely constraint experiment design. At the core of DMD lies *singular value decomposition* (SVD), which allows one to detect the relevant dynamics by truncation. Since this procedure requires the choice of the truncation thresholds (see *e.g.*, Gavish and Donoho (2014)), we introduce an automatic tuning strategy for these hyper-parameters. This allows us to overcome one of the major limitation of regularization-based strategies, with hyper-parameters tuning performed in the data pre-processing phase, *without requiring additional experiments in closed-loop*.

The paper is organized as follows. Section 2 introduces the considered framework for data-driven predictive control, while Section 3 formalizes the noise handling problem we aim at tackling. The proposed strategy is then presented in Section 4, along with an heuristic deputed to the automatic tuning of the hyper-parameters. The results obtained on a benchmark example are discussed in Section 5 and compared with the averaging strategy. Finally, Section 6 concludes the work by summarizing the contribution and highlighting possible future research directions.

2. Background

Before presenting the strategy we propose for noise handling in *data-driven predictive control* (DDPC), it is crucial to introduce the framework under which the control problem is solved. To this end, let us consider a *linear time invariant* (LTI), discrete-time system, described by the *un-*

known difference equations

$$\begin{cases} x(t+1) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \tag{1}$$

where $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are the inputs and outputs of the plant at time $t \in \mathbb{N}$, respectively, while $x(t) \in \mathbb{R}^n$ is the corresponding state vector, here assumed to be only partially accessible. Suppose that the order n of the system is not known a-priori, whilst we dispose of an upper bound $\rho \geq n$ on the system's order. Accordingly, as in (De Persis and Tesi, 2021, Section VI) let us introduce a non-minimal state $z(t) \in \mathbb{R}^{(m+p)\rho}$ made of past inputs/outputs of the system, namely

$$z(t) = \begin{bmatrix} u(t-\rho)^T & \cdots & u(t-1)^T & y(t-\rho)^T & \cdots & y(t-1)^T \end{bmatrix}^T. \tag{2}$$

To gather information on the plant, let us assume that we can inject an input sequence $\{u(t)\}_{t=0}^{T-1}$ of length $T \in \mathbb{N}$, which is persistently exciting of order $\zeta + 1$ according to the following definition, with $\zeta = (m+p)\rho$.

Definition 1 (Persistence of excitation) *An input sequence* $\{u(t)\}_{t=0}^{T-1}$ *is said to be* persistently exciting *of order* ℓ *if the Hankel matrix*

$$U_{0,\ell,T-1} = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-\ell) \\ u(1) & u(2) & \cdots & u(T-\ell+1) \\ \vdots & \vdots & \ddots & \vdots \\ u(\ell-1) & u(\ell) & \cdots & u(T-1) \end{bmatrix}, \tag{3}$$

is full row rank, i.e., *rank*$(U_{0,\ell,T-1}) = m\ell$.

Note that, when $\ell = 1$ $U_{0,1,T-1}$ is compactly denoted as $U_{0,T-1}$. Given $\{u(t)\}_{t=0}^{T-1}$, suppose that we can gather the corresponding outputs $\{y(t)\}_{t=0}^{T-1}$, so as to construct the Hankel matrices

$$Z_{0,T-1} = \begin{bmatrix} z(0) & z(1) & \ldots & z(T-1) \end{bmatrix} \tag{4a}$$

$$Z_{1,T} = \begin{bmatrix} z(1) & z(2) & \ldots & z(T) \end{bmatrix}, \tag{4b}$$

where $z(t)$ is obtained from the measured data according to (2). For (De Persis and Tesi, 2021, Lemma 3), the features of the input sequence guarantee that the following rank condition holds:

$$\text{rank} \begin{bmatrix} U_{0,T-1} \\ Z_{0,T-1} \end{bmatrix} = \zeta + m, \tag{5}$$

whenever the dataset is sufficiently long, *i.e.,* $T \geq (m+1)\zeta + m$. In this case, within a noiseless setting, the dynamics of the non-minimal realization of the unknown system dictated by $z(t)$ in (2) can be equivalently described via the data-driven representation outlined in the following theorem (De Persis and Tesi, 2021, Theorem 7).

Theorem 2 (Data-driven dynamics) *Consider the non-minimal state-space realization in the reachability form* (1) *associated with the extended state* (2)*. Let condition* (5) *hold. Then, the non-minimal realization of the unknown system in reachability form has the equivalent data-driven representation:*

$$z(t+1) = Z_{1,T}\Omega^\dagger \begin{bmatrix} u(t) \\ z(t) \end{bmatrix}, \tag{6}$$

where $\Omega = \begin{bmatrix} U_{0,T-1} \\ Z_{0,T-1} \end{bmatrix}$ *and* $\Omega^\dagger$ *denotes its Moore-Penrose right inverse.*

It is worth stressing that this representation is more general than the one introduced in Sassella et al. (2021), since it does not require the state to be directly measured nor the order of the system to be exactly known. These advantages are counterbalanced by the increased dataset size required to obtain an equivalent description of the system. Nonetheless, while the state is generally difficult to be accessed in practice, designing an experiment that meets the required specifics in terms of dataset length is likely to be less of an issue.

2.1. An input-output DDPC formulation

Let N, N_u and N_c denote the prediction, input and constraint horizons respectively, with $N_u \leq N$. In this work, we leverage on the non-minimal state defined in (2) and on Theorem 2 to introduce the following DDPC problem:

$$\min_{\bar{u}} \sum_{k=0}^{N-1} \left[\bar{z}(k)^T Q \bar{z}(k) + \bar{u}(k)^T R \bar{u}(k)\right] + \bar{z}(N)^T P \bar{z}(N) \tag{7a}$$

$$\text{s.t. } \bar{z}(k+1) = Z_{1,T}\Omega^\dagger \begin{bmatrix} \bar{u}(k) \\ \bar{x}(k) \end{bmatrix}, \quad k = 0, \dots, N-1, \tag{7b}$$

$$\bar{z}(0) = z(t), \tag{7c}$$

$$\bar{u}(k) \in \mathcal{U}, \quad k = 0, \dots, N_c - 1, \tag{7d}$$

$$\bar{y}(k) \in \mathcal{Y}, \quad k = 0, \dots, N_c - 1, \tag{7e}$$

$$\bar{u}(k) = K\bar{z}(k), \quad N_u \leq k < N, \tag{7f}$$

whose target is to steer both the predicted extended states $\{\bar{z}(k)\}_{k=1}^{L}$ and the inputs $\{u(k)\}_{k=0}^{L-1}$ to zero starting from the initial condition $\bar{z}(0) = z(t)$, which is dictated by the current configuration of the system. The trade-off between the regulation to zero of the extended state and the control effort is governed by the positive semi-definite tunable matrix $Q \in \mathbb{R}^{\zeta \times \zeta}$, the positive definite weight $R \in \mathbb{R}^{m \times m}$ and the terminal non-negative definite penalty $P \in \mathbb{R}^{\zeta \times \zeta}$. Note that, when the system is known to be open-loop stable, $K \in \mathbb{R}^{m \times \zeta}$ can be set as a matrix of zeros and P can be designed from data as described in (Sassella et al., 2021, Section 5). Otherwise, according to Bemporad et al. (2002b); Sassella et al. (2021), both P and K can be retrieved by solving a linear quadratic regulation problem on the non-minimal system, based on the approach proposed in De Persis and Tesi (2021). Differently from the DDPC formulation proposed in Sassella et al. (2021), here we explicitly constrain the inputs and outputs of the system only (see (7d)-(7e)), as the state is not directly accessible.

3. Problem statement

Suppose that the available dataset $\mathcal{D}_T = \{u(t), y(t)\}_{t=0}^{T-1}$ comprises noisy output measurements only, namely

$$y(t) = y^{\text{o}}(t) + v(t), \tag{8}$$

with $y^{\text{o}}(t) \in \mathbb{R}^p$ being the noiseless (not measured) output and $v(t) \in \mathbb{R}^p$ being the realization of a zero mean white noise. Our goal is to design a data-driven predictive controller by iteratively solving (7), while handling the noise on the data in the construction of the data-based matrices in (7b). We aim at fulfilling this objective without (i) constraining the data collection phase, *i.e.,* by not demanding for repeated experiments with the same inputs as in Sassella et al. (2021), and without (ii) requiring potentially unsafe closed-loop tests to calibrate the hyper-parameters needed to handle noise, as one might be asked to do using the approaches in Berberich et al. (2020); Dörfler et al. (2021).

4. Dynamic mode decomposition for noise handling in DDPC

To attain the goal highlighted in Section 3, in this work we propose to pre-process the data matrices in (7b) via *dynamic mode decomposition* (DMD) Kutz et al. (2016), since this tool is known to allow for the extraction of meaningful information from available data, by braking them down into a set of dynamic modes. Prior to the control design phase, we thus propose to perform *singular value decomposition* (SVD) on both $Z_{1,T}$ and Ω, thus translating them into the following products:

$$\Omega = U_\Omega \Sigma_\Omega V_\Omega^*, \quad Z_{1,T} = U_Z \Sigma_Z V_Z^*, \tag{9}$$

where $U_\Omega \in \mathbb{C}^{(m+\zeta)\times(m+\zeta)}$, $U_Z \in \mathbb{C}^{\zeta\times\zeta}$, $V_\Omega \in \mathbb{C}^{T\times T}$ and $V_Z \in \mathbb{C}^{T\times T}$ are unitary matrices[1], V_Ω^* and V_Z^* are the conjugate transposes of V_Ω and V_Z, respectively, and $\Sigma_\Omega \in \mathbb{C}^{(m+\zeta)\times T}$ and $\Sigma_Z \in \mathbb{C}^{\zeta\times T}$ are diagonal matrices, whose elements correspond to the singular values of Ω and $Z_{1,T}$. Since our objective is to have a predictor (7b) that is *sufficiently accurate* to perform the control task at hand, but that will never be used standalone, we have decided to handle noise by truncation of the matrices in (9), in line with the strategy presented in Proctor et al. (2016). As such, we try to extract insights on the *noiseless* dynamics by introducing two truncation tunable thresholds r_Ω and r_Z, dictating the number of singular values of Ω and $Z_{1,T}$ to be preserved. The DDPC scheme with DMD is summarized in Algorithm 1, where the distinction between the pre-processing stage needed for noise handling and the actual control design phase is highlighted.

4.1. A heuristic strategy for hyper-parameter tuning

The choice of the tunable thresholds r_Ω and r_Z to be fed to Algorithm 1 is of paramount importance, to guarantee that the relevant dynamics of the system are preserved, while noise is sufficiently mitigated. In this work, we introduce a heuristic for the *automatic selection* of these hyper-parameters, that allows us to avoid any closed-loop calibration phase. To devise such strategy, we rely on the inkling that the distinction between modes representative of the system and of noise is dictated by a knee in the distribution of the singular values of both Ω and $Z_{1,T}$. To detect the knees, we propose to consider the common logarithm of the singular values of Ω and $Z_{1,T}$, compute both their first and second order derivatives and, finally, set r_Ω and r_Z based on the indexes associated with the maximum in the latter. This approach, summarized in Algorithm 2, allows us to tune all hyper-parameters, without endangering the plant with closed-loop calibration tests in an attempt to select them. Notice that data matrices are constructed with a single training set, thus no additional experiments are required. We point out that this calibration strategy tends to be slightly *conservative by design*, as it selects the truncation values not necessarily looking at the first knee in the singular values distribution. Nonetheless, our tests have shown that this choice allows us to generally retain

1. A complex matrix $M \in \mathbb{C}^{n\times n}$ is said to be unitary if $M^*M = MM^* = I_n$, with $M^\star$ being the conjugate transpose of M.

Algorithm 1 DDPC with DMD-based noise handling

Input: Data matrices $Z_{1,T}$,Ω; truncation thresholds r_Ω, r_Z.

Pre-processing phase

1. **Perform SVD** on $Z_{1,T}$ and Ω as in (9);
2. **Keep** the first r_Ω singular values of Σ_Ω and the first r_Z ones of Σ_Z and **set** the remaining diagonal terms to zero;
3. **Retrieve** the truncated matrices $\tilde{\Omega}$ and $\tilde{Z}_{1,T}$ as:

$$\tilde{\Omega} = U_\Omega \tilde{\Sigma}_\Omega V_\Omega^*, \quad \tilde{Z}_{1,T} = U_Z \tilde{\Sigma}_Z V_Z^*, \tag{10}$$

Control design

1. **for** $t = 0, 1 \dots,$
   - 1.1. **Solve** (7) with Ω and $Z_{1,T}$ in (7b) replaced with $\tilde{\Omega}$ and $\tilde{Z}_{1,T}$ in (10), respectively;
   - 1.2. **Set** $u(t)$ to the first element of the optimal input sequence, *i.e.,* $u(t) = \bar{u}^\star(0)$;
   - 1.3. **Discard** the rest of the optimal sequence;
   - 1.4. **Apply** $u(t)$ and **update** $\zeta(t)$;

Algorithm 2 An heuristic for singular values truncation

Input: Common logarithm of the singular values $\{\sigma_{\Omega,i}\}_{i=1}^{m+\zeta}$, $\{\sigma_{Z,i}\}_{i=1}^{\zeta}$.

1. **Compute** the first and second order discrete derivatives of the sequences $\{\sigma_{\Omega,i}\}_{i=1}^{m+\zeta}$, $\{\sigma_{Z,i}\}_{i=1}^{\zeta}$;
2. **Find** the maximum of the second order discrete derivatives, *i.e.,*

$$\bar{\sigma}_\Omega = \max_i D^2\sigma_{\Omega,i}, \quad \bar{\sigma}_Z = \max_i D^2\sigma_{Z,i};$$

3. **Retrieve** r_Ω as: $r_\Omega = \text{argmax}_i\, D^2\sigma_{\Omega,i} - 1,\ r_Z = \text{argmax}_i\, D^2\sigma_{Z,i} - 1;$

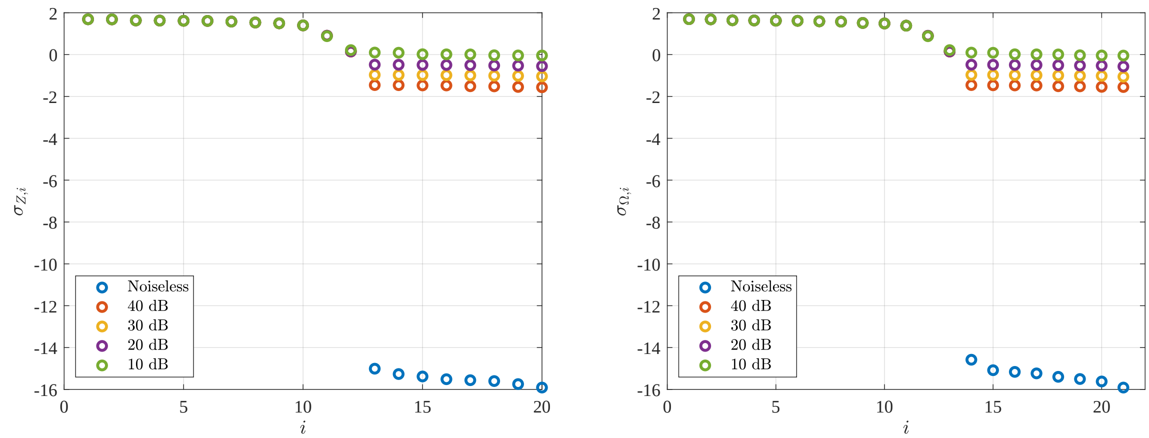
Output: Truncation values r_Ω and r_Z.

all the needed information on the unknown system, eventually slightly reducing the filtering effect of the DMD procedure.

5. A benchmark example

Consider the system introduced in Bemporad et al. (2002a), described by

$$\begin{cases} x(t+1) = \begin{bmatrix} 0.7326 & -0.0861 \\ 0.1722 & 0.9909 \end{bmatrix} x(t) + \begin{bmatrix} 0.0609 \\ 0.0064 \end{bmatrix} u(t), \\ y(t) = \begin{bmatrix} 0 & 1.4142 \end{bmatrix} x(t). \end{cases} \tag{11}$$



(a) $\{\sigma_{Z,i}\}_{i=1}^{\zeta}$ *vs* SNR

(b) $\{\sigma_{\Omega,i}\}_{i=1}^{m+\zeta}$ *vs* SNR

Figure 1: Logarithm of the singular values of $Z_{1,T}$ and Ω for increasing level of noise, dictated by the Signal-to-Noise Ratio (SNR). For $Z_{1,T}$, the first $\rho + n$ singular values are not affected by noise, while the number of noise-insensitive values grows to $\rho + n + m$ for Ω.

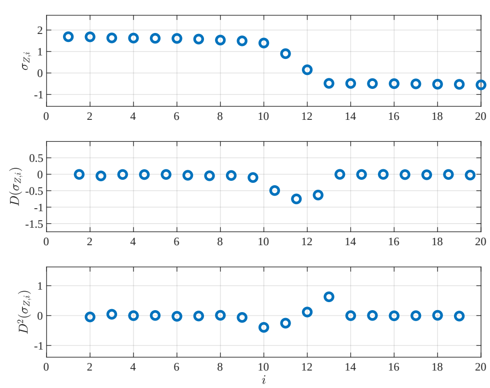
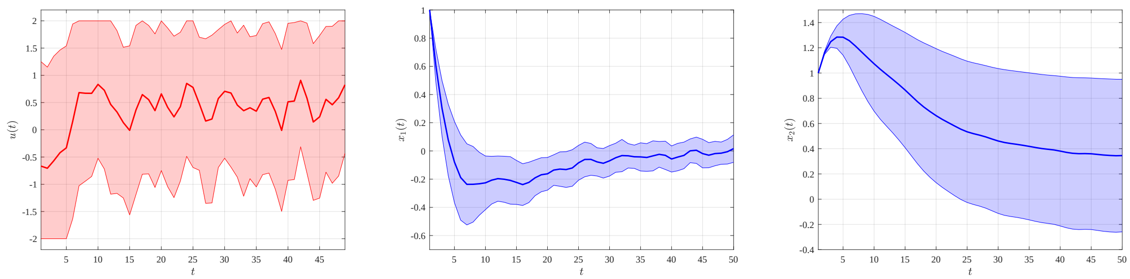


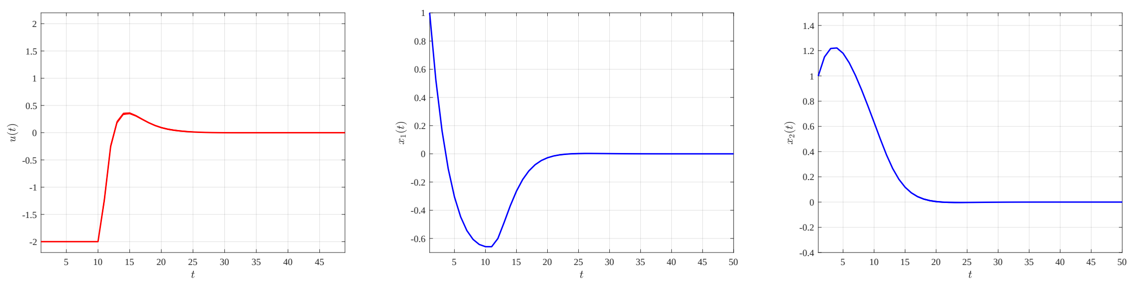
Figure 2: Logarithm of the singular values of $Z_{1,T}$ and their first and second order discrete derivatives, for a SNR around 20 dB. The knee in the singular values distribution can be effectively retrieved by looking at the maximum of the second order discrete derivative, according to the proposed heuristic.

Let us assume that we have no prior information on the system order and, as such, we conservatively select ρ in (2) equal to 10. Accordingly, we fed the system with an input sequence of length $T = 1000$, uniformly distributed within the interval $[-5, 5]$, so as to guarantee persistency of excitation. As introduced in (8), the output is corrupted by noise, here supposed to be normally distributed.

To understand whether the intuition founding the heuristic proposed in Section 4.1 is verified in this benchmark case, we initially evaluate how the singular values of $Z_{1,T}$ and Ω are impacted by noise. As shown in Figure 1, measurement noise causes changes after the twelfth singular value of $Z_{1,T}$, thus suggesting a truncation value $r_Z = 12$. Similar conclusions (see Figure 1) can be



(a) Early truncation: $r_Z = 6$, $r_\Omega = 7$



(b) Automatic truncation: $r_Z = 12$, $r_\Omega = 13$

Figure 3: Mean and standard deviation of the closed-loop input and states over Monte Carlo runs: early *vs* automatic truncation.

Table 1: Maximum absolute deviation from the average values of the closed-loop output (s_y), input (s_u) and first state (s_{x_1}): automatic *vs* late truncation. The second state is not considered, as it corresponds to the output.

	s_y	s_u	s_{x_1}
Automatic truncation: $r_Z = 12$, $r_\Omega = 13$	5.1 10^{-3}	4.1 10^{-2}	3.5 10^{-3}
Late truncation: $r_Z = 16$, $r_\Omega = 17$	5.3 10^{-3}	4.9 10^{-2}	3.9 10^{-3}

inferred for Ω, resulting in $r_\Omega = r_Z + m = 13$. Figure 2 reports the results obtaining by considering the successive derivatives of the singular value sequence of $Z_{1,T}$ for an intermediate noise level, *i.e.,* Signal-to-Noise Ratio (SNR) around 20 dB. This result supports our truncation logic, since explanatory modes and that mainly linked to noise can be actually distinguished by looking at the second derivative of the common logarithm of the singular values. Similar conclusions an be drawn for the singular values of Ω. Indeed, by applying Algorithm 2, we obtain $r_Z = 12$ and $r_\Omega = 13$.

Let us now consider the dataset obtained for a level of noise corresponding to a SNR of 20 dB. By following the steps of Algorithm 1 with these truncation values, we use the available data to design a data-driven predictive controller by imposing $N = N_c = N_u = 2$, $Q = I$, $R = 1$ and considering the input constraint

$$-2 \leq \bar{u}(k) \leq 2, \forall k = 0, \dots, N-1.$$

To select the terminal weight P, we instead find the data-driven Lyapunov matrix associated to the open-loop system, as suggested in Sassella et al. (2021).

Table 2: Experimental requirements and hyper-parameters: DMD-based approach *vs* averaging Sassella et al. (2021). The number of experiments L should be chosen as large as possible, according to the practical limitations characterizing the data collection phase.

	DMD-based	Averaging
Minimum experiment length T	$(m+1)\zeta+m$	$(m+1)\zeta+m$
Required persistence of excitation	$m+\zeta$	$m+\zeta$
Hyper-parameters	r_Z and r_Ω	L
Constraints on data collection	None	Repeated experiments with the same input

By solving the control problem with 30 dataset generated via Monte Carlo simulations[2], the average closed-loop behavior attained over noiseless tests by employing the proposed DMD-based DDPC and the heuristic summarized in Algorithm 2 is reported in Figure 3, along with the performance obtained when r_Z and r_Ω are smaller than the ones automatically tuned. Clearly the proposed approach allows us to mitigate the effect of noise, as shown by the negligible standard deviation computed over the input/output closed-loop trajectories. At the same time, it enables us to attain a result which tightly resembles the one obtained when the true model of the system is used to design the controller Bemporad et al. (2002a). On the other hand, truncating before the actual knee in the singular value distribution can be detrimental for the final performance, leading to consistent variations of the closed-loop output over the Monte Carlo runs and to worst performance in terms of both tracking and required control effort. This result was expected since *early truncation* removes modes that are relevant for the system dynamics. Instead, as shown in Table 1, a *late truncation* only leads to a slight increase in the variability of the closed-loop input and output, overall resulting in performance that is still close to the one obtained by using the proposed heuristic. This result is somehow encouraging, showing that the eventual conservatism of the proposed heuristic (see Section 4.1) would not have dramatic impacts on closed-loop performance. Both these outcomes are further aligned with well-known results on the choice of a model's order in subspace system identification, see *e.g.,* Bauer (2003).

5.1. A comparison with the strategy of Sassella et al. (2021)

We now compare the proposed DMD-based solution with the noise handling strategy introduced in Sassella et al. (2021). Instead of relying on a pre-processing phase, the latter requires one to collect data by replicating the same experiment L times, so as to construct a datatset on which the effect of noise is mitigated by averaging out the measured outputs. As shown in Table 2, the DMD-based approach proposed here and the averaging method of Sassella et al. (2021) share many features, but the latter heavily constraints the data collection phase, thus reducing its applicability in practice. To further compare the two approaches, we consider $L=10$ repeated experiments of length 1000, carried out according to the previously discussed setting, and we augment the dataset used when applying the DMD-based approach, so that $T=10000$ in both cases. By running 100 Monte Carlo simulations over data collection, we compare the performance attained with the DMD-based

2. Both the realization of the input and the measurement noise used to construct the data matrices in (6) are changed at each simulation.

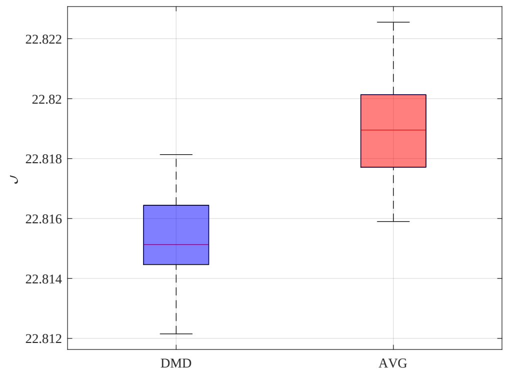


Figure 4: DMD-based *vs* averaging strategy: $\mathcal{J}$ in (12) over 100 dataset realizations.

approach and the averaging strategy through the following metric:

$$\mathcal{J} = \sum_{t=0}^{T_v} \left[\|x(t)\|_Q^2 + \|u(t)\|_R^2 \right], \tag{12}$$

where the state is exploited for performance assessment only and $T_v = 100$ is the horizon over which closed-loop performance is checked. As shown in Figure 5.1, both approaches result in similar values of the chosen performance index, with the DMD-based strategy leading to a slightly lower variability with respect to the average performance index. This result confirms that the DMD-based approach is a valid noise handling method, yet representing a more viable solution for noise mitigation in practical applications, where performing repeated experiments might be not allowed or too costly.

6. Conclusions

In this work, we have proposed a noise handling strategy for data-driven predictive control, that relies on the use of truncated DMD. To calibrate the hyper-parameters characterizing the approach, we have proposed a heuristic, whose effectiveness has been assessed over a benchmark case study. The presented heuristic allows us to reduce the burden on the user-side, while not requiring closed-loop calibrations. At the same time, it might result in conservative choices of the truncation value. When compared to an existing strategy based on averaging, the presented DMD-based approach results in a slightly improvement in performance, which is accompanied with a relevant benefit in terms of applicability.

Future research will be devoted to assess more sophisticated strategies for singular value selection and compare them with the proposed heuristic. At the same time, future works will juxtapose the DMD-based approach with other noise handling strategies presented in the literature, *e.g.,* the ones introduced in Dörfler et al. (2021).

Acknowledgments

This project was partially supported by the Italian Ministry of University and Research under the PRIN'17 project "Data-driven learning of constrained control systems", contract no. 2017J89ARP.

References

D. Bauer. Subspace algorithms. In *Proceedings of the IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 2003.

A. Bemporad, F. Borrelli, and M. Morari. Model predictive control based on linear programming - the explicit solution. *IEEE Transactions on Automatic Control*, 47(12):1974–1985, 2002a.

A. Bemporad, M. Morari, V. Dua, and E.N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, 2002b.

J. Berberich, J. Köhler, M.A. Müller, and F. Allgöwer. Data-driven model predictive control with stability and robustness guarantees. *IEEE Transactions on Automatic Control*, 66(4):1702–1717, 2020.

J. Berberich, J. Köhler, M.A. Müller, and F. Allgöwer. Linear tracking MPC for nonlinear systems part ii: The data-driven case. *arXiv preprint arXiv:2105.08567*, 2021.

V. Breschi, A. Sassella, and S. Formentin. On the design of regularized explicit predictive controllers from input-output data. *ArXiv*, abs/2110.11808, 2021.

J. Coulson, J. Lygeros, and F. Dörfler. Data-enabled predictive control: In the shallows of the deepc. In *2019 18th European Control Conference (ECC)*, pages 307–312. IEEE, 2019a.

J. Coulson, J. Lygeros, and F. Dörfler. Regularized and distributionally robust data-enabled predictive control. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 2696–2701, 2019b. doi: 10.1109/CDC40024.2019.9028943.

C. De Persis and P. Tesi. Low-complexity learning of linear quadratic regulators from noisy data. *Automatica*, 128:109548, 2021.

F. Dörfler, J. Coulson, and I. Markovsky. Bridging direct & indirect data-driven control formulations via regularizations and relaxations. *arXiv preprint arXiv:2101.01273*, 2021.

S. Formentin, K. Van Heusden, and A. Karimi. A comparison of model-based and data-driven controller tuning. *International Journal of Adaptive Control and Signal Processing*, 28(10):882–897, 2014.

M. Gavish and D.L. Donoho. The optimal hard threshold for singular values is 4/3. *IEEE Transactions on Information Theory*, 60(8):5040–5053, 2014.

J.N. Kutz, S.L. Brunton, B.W. Brunton, and J.L. Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

J.L. Proctor, S.L. Brunton, and J.N. Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.

A. Sassella, V. Breschi, and S. Formentin. Learning explicit predictive controllers: theory and applications, 2021.

P.J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656:5–28, 2010.

J.C. Willems, P. Rapisarda, I. Markovsky, and B.L.M. De Moor. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329, 2005.