Beyond Persistency of Excitation:
Online Experiment Design for Data-Driven Modeling and Control

Henk J. van Waarde

Abstract—This paper presents a new experiment design method for data-driven modeling and control. The idea is to select inputs online (using past input/output data), leading to desirable rank properties of data Hankel matrices. In comparison to the classical persistency of excitation condition, this online approach requires less data samples and is even shown to be completely sample efficient.

I. INTRODUCTION

Recently, there has been an increasing interest in the direct design of controllers using data [1]–[12]. Several contributions study how controllers can be obtained from a given batch of (informative) data, even in the presence of noise [4], [6], [7] and for classes of nonlinear systems [11], [12]. The question of how to obtain such informative data sets, however, is largely open. For data-driven control to become an end-to-end solution, there is a need for new experiment design methods to empower the data-based design. This is true especially for settings including noise and nonlinear dynamics. However, even for linear systems with exact data, current experiment design methods are not sample efficient.

In this paper we will explore a new idea for designing experiments for data-driven modeling and control. Experiment design is a classical problem that has been mostly studied in the parametric identification literature. An established idea is to optimize a measure of the expected accuracy of the parameter estimates subject to input power constraints [13], [14]. This problem is usually tackled in the frequency domain and convex formulations have been provided in [15]. The dual problem of finding the “least costly” input achieving a fixed level of parameter accuracy has also been studied [16], in a closed-loop setting.

Less results are known in the context of non-parametric methods. In this area, a state-of-the-art result for linear time-invariant systems is Willems et al.’s fundamental lemma [17]. The idea is to select an input that is persistently exciting, which implies that a Hankel matrix of measured inputs and outputs satisfies a rank condition. This rank property is important, since it guarantees that all trajectories of the system can be parameterized in terms of the measured trajectory. Essentially, the Hankel matrix of measured inputs and outputs serves as a non-parametric model of the system. This idea is simple yet powerful, and has been successfully employed in a number of recent publications on data-driven simulation and control, cf. [1]–[3].

Despite its elegance and clear impact on data-driven methods, the fundamental lemma has some limitations.

1) It is not sample efficient in the sense that it uses more data than strictly necessary. This is because the fundamental lemma works with inputs that are persistently exciting of a certain order, which imposes a conservative lower bound on the number of samples. Large data sets are challenging from a computational point of view, especially for the real-time implementation of control laws, see e.g. [18].

2) It is not applicable to noisy data. A possible approach to deal with this is to perform approximate data-driven simulation and control in a maximum likelihood framework [19], for which an experiment design method was developed in [20]. Another line of work establishes data-driven control design methods with guaranteed stability and performance in the presence of bounded process noise [4], [6], [7]. However, for this, experiment design methods are missing and it is unclear whether persistency of excitation is helpful in this context.

The purpose of this paper is to introduce another angle of attack to experiment design. In contrast to Willems’ fundamental lemma, we do not use persistently exciting inputs, but instead design the inputs online. This means that at each time step, the input is selected on the basis of inputs and outputs that have been collected at previous time steps. Such an online approach is natural because the first samples of an experiment already contain valuable partial information about the system, which can be exploited in the design of the remaining samples. Although online approaches have been studied in various contexts [21], [22], such a method has not been explored in the setting of Willems’ lemma.

As our main contributions, we provide online experiment design methods for both input/state and input/output systems. In both settings, we formally prove that the methods are sample efficient. This completely resolves the problem in 1), and thus shows an advantage of online experiment design over the classical persistency of excitation condition. The main idea underlying our approach is to reduce the dimension of the left kernel of a data Hankel matrix at each time step. Analogous to the fundamental lemma, our technical results are presented for noise-free data. Nonetheless, we believe that online experiment design will be crucial also in tackling problems involving noisy data as mentioned in 2). We will discuss this issue in more detail in Section IV.

Outline: In Section II we provide background material. In Section III we present our online experiment design method. Finally, Section IV contains our conclusions and discussion.
A. Notation and terminology

The left kernel of a real matrix \( M \) is denoted by \( \text{leftker} \ M \). Consider a signal \( f : \mathbb{Z} \to \mathbb{R}^n \) and let \( i, j \in \mathbb{Z} \) be integers such that \( i \leq j \). We denote by \( f_{[i,j]} \) the restriction of \( f \) to the interval \([i,j]\), that is,

\[
 f_{[i,j]} := \begin{bmatrix} f(i)^T & f(i + 1)^T & \cdots & f(j)^T \end{bmatrix}^T.
\]

With slight abuse of notation, we will also use the notation \( f_{[i,j]} \) to refer to the sequence \( f(i), f(i + 1), \ldots, f(j) \). Let \( k \) be a positive integer such that \( k \leq j - i + 1 \) and define the Hankel matrix of depth \( k \), associated with \( f_{[i,j]} \), as

\[
 \mathcal{H}_k(f_{[i,j]}) := \begin{bmatrix}
 f(i) & f(i + 1) & \cdots & f(j - k + 1) \\
 f(i + 1) & f(i + 2) & \cdots & f(j - k + 2) \\
 \vdots & \vdots & \ddots & \vdots \\
 f(i + k - 1) & f(i + k) & \cdots & f(j)
 \end{bmatrix}.
\]

Note that the subscript \( k \) refers to the number of block rows of \( \mathcal{H}_k \). The sequence \( f_{[i,j]} \) is called persistently exciting of order \( k \) if \( \mathcal{H}_k(f_{[i,j]}) \) has full row rank.

II. Rank conditions on data matrices

Consider the linear time-invariant (LTI) system

\[
 \begin{align}
 x(t + 1) &= Ax(t) + Bu(t) \quad (1a) \\
 y(t) &= Cx(t) + Du(t), \quad (1b)
\end{align}
\]

where \( x \in \mathbb{R}^n \) denotes the state, \( u \in \mathbb{R}^m \) is the input and \( y \in \mathbb{R}^p \) is the output. Throughout the paper, we will assume that (1) is minimal, i.e., the pair \((A, B)\) is controllable and \((C, A)\) is observable. The lag \( \ell \) of (1) can be defined as the smallest integer \( i \) for which the observability matrix

\[
 O_i := \begin{bmatrix} C^T & (CA)^T & (CA^2)^T & \cdots & (CA^{i-1})^T \end{bmatrix}^T
\]

has rank \( n \). Now, let \((u[0:T-1], y[0:T-1])\) be an input/output trajectory of (1) of length \( T \geq \ell \). In addition, let \( L \geq \ell \) be an integer. The main goal of this article is to provide a new method to design the input sequence \( u[0:T-1] \) so that the resulting input/output Hankel matrix

\[
 \begin{bmatrix} \mathcal{H}_L(y[0:T-1]) \\ \mathcal{H}_L(u[0:T-1]) \end{bmatrix}
\]

has rank \( n + mL \). This rank condition plays a fundamental role in modeling and control using data. Indeed, it implies that any length-\( L \) input/output trajectory of (1) can be obtained from (3). More specifically, if (3) has rank \( n + mL \) then \((\bar{u}[0:L-1], \bar{y}[0:L-1])\) is an input/output trajectory of (1) and only if

\[
 \begin{bmatrix} \bar{y}[0:L-1] \\ \bar{u}[0:L-1] \end{bmatrix} \in \text{im} \begin{bmatrix} \mathcal{H}_L(y[0:T-1]) \\ \mathcal{H}_L(u[0:T-1]) \end{bmatrix}.
\]

This parameterization of input/output trajectories has been used extensively to simulate and control dynamical systems, see e.g., [1]-[3]. If \( L > \ell \), the rank condition on (3) even implies that all input/output trajectories of (1) (not just those of length \( L \)) can be obtained from \((u[0:T-1], y[0:T-1])\), see [23]. In this case, the input/output behavior of (1) is identifiable [24] and \((A, B, C, D)\) can be computed up to similarity transformation, e.g., using subspace methods [25].

A notable special case is that of full state measurement, i.e., \( y(t) = x(t) \) and \( L = \ell = 1 \). In this case, (3) reduces to

\[
 \begin{bmatrix} \mathcal{H}_1(x[0:T-1]) \\ \mathcal{H}_1(u[0:T-1]) \end{bmatrix} = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \\ u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}.
\]

If (3) has rank \( n + m \), the input/state trajectory \((u[0:T-1], x[0:T])\) fully captures the behavior of (1a) which allows the unique identification of \( A \) and \( B \).

A. Recap of Willems et al.’s fundamental lemma

Willems et al.’s fundamental lemma [17] is an important experiment design result. It reveals that the rank condition on (3) is satisfied if the input sequence is chosen to be persistently exciting. In a state-space setting, this result can be stated as follows [17], [26].

\[
 \text{Proposition 1:} \text{ Consider the minimal system (1). Let } (u[0:T-1], y[0:T-1]) \text{ be an input/output trajectory of (1) and } L \geq \ell. \text{ If the input } u[0:T-1] \text{ is persistently exciting of order } n + L \text{ then (3) has rank } n + mL. \]

We note that persistency of excitation of order \( n + L \) requires at least \( T \geq (m + 1)(n + L) - 1 \) data samples.

III. Online input design

The main contribution of the paper will be to provide a new input design technique to guarantee the rank property of (3). In contrast to Willems’ fundamental lemma, this method will not rely on applying inputs that are persistently exciting of order \( n + L \). Rather, each input \( u(t) \) is selected online, by making use of the previous samples \( y(0), y(1), \ldots, y(t-1) \) and \( u(0), u(1), \ldots, u(t-1) \), see Figure 1.

![Fig. 1. Illustration of the online input design approach.](image-url)
A. Input design for input/state systems

To explain the idea, we will start with the simplest setting of input/state data collected from system (1a). In this setting, the purpose is to design a sequence of inputs \( u(0), u(1), \ldots, u(T - 1) \) so that the matrix \( \mathcal{H}_1 \) has full row rank \( n + m \). As mentioned, we opt for the online design of the inputs. Specifically, this means that for the design of \( u(t) \) we make use of the collected states \( x(0), x(1), \ldots, x(t) \) and inputs \( u(0), u(1), \ldots, u(t - 1) \). The following theorem shows how such inputs, as well as a time horizon \( T \), can be designed.

**Theorem 1:** Consider the controllable system (1a). Define \( T := n + m \). Select a nonzero \( u(0) \in \mathbb{R}^m \), and design the input \( u(t) \) (for \( t = 1, 2, \ldots, T - 1 \)) as follows:

- If \( x(t) \notin \text{im} \mathcal{H}_1(\{x[0, t-1]\}) \), select \( u(t) \in \mathbb{R}^m \) arbitrarily.
- If \( x(t) \in \text{im} \mathcal{H}_1(\{x[0, t-1]\}) \) there exists \( \xi \in \mathbb{R}^n \) and a nonzero \( \eta \in \mathbb{R}^m \) such that

\[
\begin{bmatrix}
\xi^T \\
\eta^T
\end{bmatrix}
\begin{bmatrix}
\mathcal{H}_1(\{x[0, t-1]\}) \\
\mathcal{H}_1(\{u[0, t-1]\})
\end{bmatrix} = 0. \quad (5)
\]

In this case, select \( u(t) \) such that \( \xi^T x(t) + \eta^T u(t) \neq 0 \).

Then, we have that

\[\text{rank} \begin{bmatrix}
\mathcal{H}_1(\{x[0, t-1]\}) \\
\mathcal{H}_1(\{u[0, t-1]\})
\end{bmatrix} = n + m. \quad (6)\]

We will not provide a proof of Theorem 1 at this stage, since the statement will follow from the more general Theorem 2 that will be proven in Section III-B. Instead, we will explain the main ideas behind the result. The essence of Theorem 1 is that the proposed input sequence increases the rank of the input/state Hankel matrix at every time step. To be specific, the input sequence guarantees that

\[\text{rank} \begin{bmatrix}
\mathcal{H}_1(\{x[t]\}) \\
\mathcal{H}_1(\{u[t]\})
\end{bmatrix} > \text{rank} \begin{bmatrix}
\mathcal{H}_1(\{x[0, t-1]\}) \\
\mathcal{H}_1(\{u[0, t-1]\})
\end{bmatrix} \quad (7)\]

for all \( t = 1, 2, \ldots, T - 1 \). This increase in rank obviously occurs for any input \( u(t) \in \mathbb{R}^m \) if \( x(t) \notin \text{im} \mathcal{H}_1(\{x[0, t-1]\}) \). Therefore, \( u(t) \) can be chosen arbitrarily in this case. However, it is not evident that (7) can be guaranteed if \( x(t) \in \text{im} \mathcal{H}_1(\{x[0, t-1]\}) \). Here, the important intermediate step of Theorem 1 is to show that there exists a \( \xi \) and a nonzero \( \eta \) satisfying (5). The existence of these vectors heavily relies on controllability of \((A, B)\), and can be proven using a geometric argument that draws some inspiration from Hautus’ proof of Heymann’s lemma [27]. Once the existence of \( \xi \) and \( \eta \neq 0 \) has been established, the construction of the input \( u(t) \) should come without surprise. Indeed, choosing \( u(t) \) such that \( \xi^T x(t) + \eta^T u(t) \neq 0 \) ensures that \( \begin{bmatrix}
\xi^T \\
\eta^T
\end{bmatrix} \begin{bmatrix}
\mathcal{H}_1(\{x[0, t]\}) \\
\mathcal{H}_1(\{u[0, t]\})
\end{bmatrix} \]

is not in the left kernel of the input/state Hankel matrix up to time \( t \). By (5), this implies that

\[
\dim \text{leftker} \begin{bmatrix}
\mathcal{H}_1(\{x[t]\}) \\
\mathcal{H}_1(\{u[t]\})
\end{bmatrix} < \dim \text{leftker} \begin{bmatrix}
\mathcal{H}_1(\{x[0, t-1]\}) \\
\mathcal{H}_1(\{u[0, t-1]\})
\end{bmatrix},
\]

which is equivalent to (7).

Note that to guarantee the rank condition (6), we require at least \( T \geq n + m \) samples. A surprising fact is that Theorem 1 always guarantees (6) with exactly \( n + m \) samples, despite the a priori lack of knowledge of the system matrices \( A \) and \( B \). This makes our design method completely sample efficient in the sense that any experiment design method requires at least as many samples as the one in Theorem 1. In particular, our method outperforms the usual condition of persistency of excitation of order \( n + 1 \), which requires at least \( T \geq nm + n + m \) samples to guarantee (6). Specifically, Theorem 1 saves at least \( nm \) samples compared to persistency of excitation. The number of samples in Theorem 1 is linear (instead of quadratic) in the dimensions of the system variables.

We emphasize that the success of Theorem 1 is due to its online nature: the input \( u(t) \) is computed by making use of past inputs and states, which contain valuable information about the unknown system. The method is thereby fundamentally different from the classical persistency of excitation condition, which is a purely offline condition: one can design a persistently exciting input before collecting any data. In the language of Sontag [28], persistently exciting inputs (of order \( n + 1 \)) are universal in the sense that they guarantee (6) for any controllable system (1a). In contrast, the input sequence of Theorem 1 is tailored to the specific data-generating system (1a) that has produced the past inputs and states. This allows for a reduction of the required number of data samples, at the small cost of some simple online computations.

B. Input design for input/output systems

Next, we turn our attention to input/output data generated by system (1). In this setting, we want to find a sequence of inputs \( u(0), u(1), \ldots, u(T - 1) \) so that (3) has rank \( n + mL \). As in the previous section, we opt for the online design of the inputs, meaning that for the design of \( u(t) \) we use the collected outputs \( y(0), y(1), \ldots, y(t - 1) \) and inputs \( u(0), u(1), \ldots, u(t - 1) \).

The following theorem is a building block in our approach, and shows how full row rank of the input/state Hankel matrix

\[
\begin{bmatrix}
\mathcal{H}_I(x[0, T-L]) \\
\mathcal{H}_L(u[0, T-1])
\end{bmatrix} = \begin{bmatrix}
x(0) & x(1) & \cdots & x(T-L) \\
u(0) & u(1) & \cdots & u(T-L) \\
\vdots & \vdots & \ddots & \vdots \\
u(L-1) & u(L) & \cdots & u(T-1)
\end{bmatrix}
\]

can be guaranteed if the state of (1a) is measured.

**Theorem 2:** Consider the controllable system (1a). Define \( T := n + (m + 1)L - 1 \) for \( L \geq 1 \). Choose \( u(0), u(1), \ldots, u(L-1) \in \mathbb{R}^m \) arbitrarily, but not all zero. Furthermore, design \( u(t) \) (for \( t = L, L+1, \ldots, T - 1 \)) as follows:

- If \( x(t - L + 1) \notin \text{im} \mathcal{H}_I(x[0, t-L]) \), then choose \( u(t) \in \mathbb{R}^m \) arbitrarily.
- If \( x(t - L + 1) \in \text{im} \mathcal{H}_I(x[0, t-L]) \), then

\[
\begin{bmatrix}
x(t - L + 1) \\
u(t-L+1, t-1)
\end{bmatrix} \notin \text{im} \begin{bmatrix}
\mathcal{H}_I(x[0, t-L]) \\
\mathcal{H}_{L-1}(u[0, t-2])
\end{bmatrix}, \quad (9)
\]

which is equivalent to (7).

\[
\dim \text{leftker} \begin{bmatrix}
\mathcal{H}_I(x[t]) \\
\mathcal{H}_I(u[t])
\end{bmatrix} < \dim \text{leftker} \begin{bmatrix}
\mathcal{H}_I(x[0, t-1]) \\
\mathcal{H}_I(u[0, t-1])
\end{bmatrix},
\]

which is equivalent to (7).

In this setting, we use \( x(t) \) in the design since it is independent of \( u(t) \).
then there exist $\xi \in \mathbb{R}^n$ and $\eta_1, \eta_2, \ldots, \eta_L \in \mathbb{R}^m$ with $\eta_1 \neq 0$ such that

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_2^T & \eta_1^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-1]}) \end{bmatrix} = 0. \quad (11)$$

In this case, choose $u(t)$ such that

$$\xi^T x(t-L+1) + \eta_L^T u(t-L+1) + \cdots + \eta_1^T u(t) \neq 0.$$

Then, the input sequence $u(0), u(1), \ldots, u(T-1)$ is such that

$$\text{rank} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-1]}) \end{bmatrix} = n + mL. \quad (12)$$

**Proof:** Since $u(0), u(1), \ldots, u(L-1)$ are not all zero, we clearly have

$$\text{rank} \begin{bmatrix} x(0) \\ u_{[0, L-1]} \end{bmatrix} = 1.$$

The idea of the proof is to show that in each time step, we can increase the rank of the Hankel matrix. That is, for each $t = L, L + 1, \ldots, T - 1$, we want to prove that

$$\text{rank} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-1]}) \end{bmatrix} > \text{rank} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-2]}) \end{bmatrix}. \quad (13)$$

Let $t \in \{L, L + 1, \ldots, T - 1\}$. First consider the case that

$$\eta_1 \neq 0$$

holds. Clearly, (13) is satisfied for any $u(t) \in \mathbb{R}^m$.

Next, consider the case that

$$\eta_1 = \cdots = \eta_k = 0$$

for some $1 \leq k < L$. Our goal is to show that (11) holds for $i = k + 1$ as well. Let

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_1^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-1]}) \end{bmatrix} = 0. \quad (16)$$

where $\xi \in \mathbb{R}^n$ and $\eta_1, \ldots, \eta_L \in \mathbb{R}^m$. By the induction hypothesis, we have $\eta_1 = \cdots = \eta_k = 0$ so that

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_1^T \\ \xi^T & \eta_L^T & \cdots & \eta_{k+1}^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k-1]}) \end{bmatrix} = 0. \quad (17)$$

By (10), we obtain

$$\begin{bmatrix} x(t-L+1) \\ u_{[t-L+1, t-1]} \end{bmatrix} \in \text{im} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k-1]}) \end{bmatrix}.$$ 

Combining the latter inclusion with (17) yields

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_{k+1}^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k]}) \end{bmatrix} = 0. \quad (18)$$

Note that the laws of system (1a) imply that

$$\begin{bmatrix} H_{1}(x_{[1, t-L+1])} \\ H_{L}(u_{[1, t-k])} \end{bmatrix} = \begin{bmatrix} A & B & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k])} \end{bmatrix}.$$ 

Therefore, (18) implies

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_{k+1}^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k]}) \end{bmatrix} = 0.$$ 

This yields

$$\begin{bmatrix} \xi^T & \eta_L^T & \cdots & \eta_{k+1}^T \end{bmatrix} \begin{bmatrix} H_{1}(x_{[0, t-L])} \\ H_{L}(u_{[0, t-k]}) \end{bmatrix} = 0.$$ 

Finally, by the induction hypothesis we see that $\eta_{k+1} = 0$, as desired. Thus, (15) holds for all $i = 1, L$. In particular, for $i = L$ we obtain

$$\text{leftker} H_{1}(x_{[0, t-L])} = \text{leftker} H_{1}(x_{[0, t-L]} \times \{0, m \}).$$

equivalently (using the fact that $(\text{im } X)^\perp = \text{leftker } X$)

$$\text{im} H_{1}(x_{[0, t-L])} = \text{im} H_{1}(x_{[0, t-L]} \times \mathbb{R}^m). \quad (19)$$

Multiplication on both sides by the matrix $[A \ B \ 0]$ yields

$$A \text{im} H_{1}(x_{[0, t-L])} + B = \text{im} H_{1}(x_{[1, t-L+1]}).$$

Using (10), we obtain

$$A \text{im} H_{1}(x_{[0, t-L])} + B \subseteq \text{im} H_{1}(x_{[0, t-L])}.$$ 

In particular, we see that

$$A \text{im} H_{1}(x_{[0, t-L])} \subseteq \text{im} H_{1}(x_{[0, t-L])}$$

and by (19)

$$\text{im} H_{1}(x_{[0, t-L])} = \mathbb{R}^n + mL.$$ 

In other words, $\text{im} H_{1}(x_{[0, t-L])}$ is an $A$-invariant subspace containing $B$. Since the reachable subspace $(A | \text{im } B)$ of the pair $(A, B)$ is the smallest $A$-invariant subspace containing $B$ (cf. [29, Ch. 3]), we see that

$$\mathbb{R}^n = \langle A | \text{im } B \rangle \subseteq \text{im} H_{1}(x_{[0, t-L])},$$

where we made use of the fact that $(A, B)$ is controllable. As such, $\text{im} H_{1}(x_{[0, t-L])} = \mathbb{R}^n$ and by (19) we conclude that

$$\text{im} H_{1}(x_{[0, t-L])} = \mathbb{R}^n + mL.$$ 

This implies that $t \geq n + (m + 1)L - 1$ which leads to a contradiction as $t \in \{L, L + 1, \ldots, T - 1\}$. As such, we conclude that (11) does not hold. Therefore, there exist $\xi \in \mathbb{R}^n$ and $\eta_1, \eta_2, \ldots, \eta_L \in \mathbb{R}^m$ with $\eta_1 \neq 0$ such that (11) holds. Clearly, this means that there exists a $u(t)$ such that

$$\xi^T x(t-L+1) + \eta_L^T u(t-L+1) + \cdots + \eta_1^T u(t) \neq 0.$$ 

For such an input, we have that

$$\dim \text{leftker} H_{1}(x_{[0, t-L])} > \dim \text{leftker} H_{1}(x_{[0, t-L+1]}).$$

and consequently, (13) holds. This proves the theorem.  

Note that in the special case $L = 1$, Theorem 2 reduces to Theorem 1. Next, we turn our attention to the situation
in which we measure inputs and outputs and want to ensure that (3) has rank \( n + mL \). This is the topic of the following theorem.

**Theorem 3:** Consider the controllable and observable system (1). Let \( L > \ell \) and define \( T := n + (m+1)L - 1 \). Choose \( u(0), u(1), \ldots, u(L-1) \in \mathbb{R}_m^m \) arbitrarily, but not all zero. Furthermore, design \( u(t) \) (for \( t = L, L + 1, \ldots, T - 1 \)) as follows:

- If \[
\begin{bmatrix}
y_{[t-L+1, t-1]}^T \\
u_{[t-L+1, t-1]}^T
\end{bmatrix} \notin \text{im } \frac{\mathcal{H}_{L-1}(y_{[0, t-2]})}{\mathcal{H}_{L-1}(u_{[0, t-2]})},
\]
  then choose \( u(t) \in \mathbb{R}_m^m \) arbitrarily.
- If \[
\begin{bmatrix}
y_{[t-L+1, t-1]}^T \\
u_{[t-L+1, t-1]}^T
\end{bmatrix} \in \text{im } \frac{\mathcal{H}_{L-1}(y_{[0, t-2]})}{\mathcal{H}_{L-1}(u_{[0, t-2]})},
\]
  there exist \( \xi_1, \ldots, \xi_{L-1} \in \mathbb{R}^p \) and \( \eta_1, \ldots, \eta_{L} \in \mathbb{R}_m^m \) with \( \eta_1 \neq 0 \) such that
  \[
  \begin{bmatrix}
  \xi_{L-1}^T \\
  \vdots \\
  \xi_1^T \\
  \eta_L^T \\
  \vdots \\
  \eta_1^T
\end{bmatrix} \begin{bmatrix}
  \mathcal{H}_{L-1}(y_{[0, t-2]}) \\
  \mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = 0.
\]

In this case, choose \( u(t) \) such that
\[
\begin{align*}
\xi_{L-1}^T y(t-L+1) + \cdots + \xi_1^T y(t-1) \\
+ \eta_L^T u(t-L+1) + \cdots + \eta_1^T u(t) \neq 0.
\end{align*}
\]

Then we have that
\[
\text{rank } \begin{bmatrix}
\mathcal{H}_{L}(y_{[0, T-1]}) \\
\mathcal{H}_{L}(u_{[0, T-1]})
\end{bmatrix} = n + mL. \quad (24)
\]

Proof: Since \( u(0), u(1), \ldots, u(L-1) \) are not all zero, we clearly have
\[
\text{rank } \begin{bmatrix}
y_{[0, L-2]}^T \\
u_{[0, L-1]}^T
\end{bmatrix} = 1.
\]

The idea of the proof is to show that for each time step \( t = L, L + 1, \ldots, T - 1 \), we have
\[
\text{rank } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-1]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = n + mL. \quad (25)
\]

This would prove that
\[
\text{rank } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = n + mL,
\]
and consequently, (24) holds.

Let \( t \in \{L, L + 1, \ldots, T - 1\} \). First consider the case that (20) holds. Clearly, (25) is satisfied for any \( u(t) \in \mathbb{R}^m \).

Next, consider the case that (21) holds. We claim that
\[
\text{leftker } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} \neq \text{leftker } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L-1}(u_{[0, t-2]})
\end{bmatrix} \times \{0_m\}.
\]

We will prove this claim by contradiction. Thus, suppose that (26) holds with equality, equivalently,
\[
\text{im } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = \text{im } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L-1}(u_{[0, t-2]})
\end{bmatrix} \times \mathbb{R}^m. \quad (27)
\]

This implies that
\[
M \text{ im } \begin{bmatrix}
\mathcal{H}_{L}(x_{[0, t-L]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = M \left( \text{ im } \begin{bmatrix}
\mathcal{H}_{L-1}(x_{[0, t-L]}) \\
\mathcal{H}_{L-1}(u_{[0, t-2]})
\end{bmatrix} \times \mathbb{R}^m \right),
\]
where \( M \) is defined as
\[
M := \begin{bmatrix}
N_0 & \ldots & 0 \\
0 & I_m
\end{bmatrix}, \text{ with } N := \begin{bmatrix}
O_{L-1} & T_{L-1}
\end{bmatrix},
\]
\( O_i \) defined in (2) and \( T_i \) given by
\[
\begin{bmatrix}
D & 0 & 0 & \cdots & 0 \\
CB & D & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \cdots \\
CAB \cdot 4 B & \cdots & \cdots & \cdots & D
\end{bmatrix}.
\]

Since \((C, A)\) is observable and \( L > \ell \), the matrix \( M \) has full column rank. As such, we conclude that
\[
\text{im } \begin{bmatrix}
\mathcal{H}_{L}(x_{[0, t-L]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = \text{im } \begin{bmatrix}
\mathcal{H}_{L}(x_{[0, t-L]}) \\
\mathcal{H}_{L-1}(u_{[0, t-2]})
\end{bmatrix} \times \mathbb{R}^m,
\]
i.e., (14) holds. Similarly, we note that (21) implies
\[
\text{im } \begin{bmatrix}
\mathcal{H}_{L}(y_{[0, t-L]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} = \text{im } \begin{bmatrix}
\mathcal{H}_{L}(y_{[0, t-L]}) \\
\mathcal{H}_{L-1}(u_{[0, t-2]})
\end{bmatrix} \times \mathbb{R}^m.
\]

Since \( N \) has full column rank, this implies that (10) holds. It now follows from the proof of Theorem 2 that \( t \) satisfies \( t \geq n + (m+1)L - 1 \). This yields a contradiction since \( t \in \{L, L + 1, \ldots, T - 1\} \). As such, (20) holds. Therefore, there exist \( \xi_1, \xi_2, \ldots, \xi_{L-1} \in \mathbb{R}^p \) and \( \eta_1, \eta_2, \ldots, \eta_L \in \mathbb{R}^m \) with \( \eta_1 \neq 0 \) such that (22) is satisfied. This means that we can choose an input \( u(t) \) such that (23) holds. For this choice of \( u(t) \), we obtain
\[
\text{dim leftker } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-2]}) \\
\mathcal{H}_{L}(u_{[0, t-1]})
\end{bmatrix} > \text{dim leftker } \begin{bmatrix}
\mathcal{H}_{L-1}(y_{[0, t-1]}) \\
\mathcal{H}_{L}(u_{[0, t]})
\end{bmatrix}.
\]

We conclude that (25) holds, which proves the theorem.  

We note that (24) can only hold if \( T \geq n + (m + 1)L - 1 \). By Theorem 3, we can always design an informative experiment of length exactly \( n + (m + 1)L - 1 \). In other words, the design procedure of Theorem 3 is again sample efficient. In comparison, the usual persistency of excitation condition of order \( n + L \) (Proposition 1) requires \( T \geq n + (m + 1)L + mn - 1 \) samples. Theorem 3 thus saves at least \( mn \) samples in comparison to persistency of excitation.

Finally, it is noteworthy that exact knowledge of the state-space dimension is not required for Theorem 3 even though \( T \) is defined in terms of \( n \). The only requirement is an upper bound \( L > \ell \) on the lag of the system. The reason is that all the steps of Theorem 3 can be executed without knowing \( n \). One should only take some care with the stopping criterion since \( T \) is a priori unknown. One simple approach to deal with this is the following: apply the steps of Theorem 3 for \( t = 0, 1, \ldots \) until \( t \) is such that (27) is satisfied. By the proof of Theorem 3, the smallest \( t \) for which (27) holds is \( t = n + (m+1)L - 1 \). From this relation, \( n \) can be recovered, and for \( T := t \), the rank condition (24) holds.
IV. CONCLUSIONS AND DISCUSSION

The purpose of this paper has been to provide a new angle of attack to the problem of input design for data-driven modeling and control of linear systems. Instead of the classical persistency of excitation condition, we have proposed an online method to select the inputs. We have studied both input/state and input/output systems. For both types of systems, we have provided an online design method that guarantees an important rank condition on a data Hankel matrix. A notable feature of our approach is that it is completely sample efficient in the sense that it guarantees these rank conditions with the minimum number of samples. Compared to the usual persistency of excitation conditions, the online approach saves at least \( nm \) samples, where \( n \) and \( m \) are the dimensions of the state and input, respectively.

The results presented in this paper are applicable to linear time-invariant systems and exact data. In the case of noisy measurements, rank conditions on Hankel matrices are in general not sufficient to perform successful data-driven modeling and control. Nonetheless, under suitable conditions, controllers with guaranteed stability and performance can be obtained from data. In fact, data-based linear matrix inequality (LMI) conditions were studied in [4] and [6], and a necessary and sufficient condition was provided in [7] using a generalization of Yakubovich’s S-lemma [30]. However, the problem of designing inputs that guarantee feasibility of the LMI’s in these papers is still open.

It is important to note that such an input design problem is impossible to solve using offline conditions on the input. The reason is that feasibility of the data-based LMI’s is dependent on a signal-to-noise ratio, which is determined by the true -a priori unknown- system. Therefore, making use of past data, and designing inputs online will be necessary. In fact, the experiment design method proposed in this paper is promising also for control using noisy data. Indeed, recall that the main idea behind our approach is to reduce the dimension of the left kernel of the Hankel matrix at each time step. In a similar spirit, we envision the possibility to reduce the dimension of the convex solution set to the alternative LMI [32] of the data-based LMI in [7]. This will be studied in detail in our future work.

REFERENCES

[1] I. Markovsky and P. Rapisarda, “Data-driven simulation and control,” International Journal of Control, vol. 81, no. 12, pp. 1946–1959, 2008.
[2] T. M. Maapong and P. Rapisarda, “Data-driven control: A behavioral approach,” Systems & Control Letters, vol. 101, pp. 37–43, 2017.
[3] J. Coulson, J. Lygeros, and F. Dörfler, “Data-enabled predictive control: In the shadows of the DeepPC,” in Proceedings of the European Control Conference, June 2019, pp. 307–312.
[4] C. De Persis and P. Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness,” IEEE Transactions on Automatic Control, vol. 65, no. 3, pp. 909–924, March 2020.
[5] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, “Data informativity: a new perspective on data-driven analysis and control,” IEEE Transactions on Automatic Control, vol. 65, no. 11, pp. 4753–4768, 2020.

\( ^2 \)It is actually common that optimal excitation inputs depend on the data-generating system. In the identification literature, this has been referred to as the chicken-and-egg problem, see e.g. [31].