Approximate Realization with Time Delay

Jan van Helvoirt, Okko Bosgra, Bram de Jager, and Maarten Steinbuch

Abstract—This paper describes the application of an approximate realization algorithm to dynamical systems with a time delay. First, a well-known algorithm is presented to obtain an approximate realization from an impulse response sequence. Then the limitation that a time delay imposes on the accuracy of this algorithm is discussed, and it is pointed out that time delays should be explicitly taken into account. Therefore, a time delay estimation method is proposed that utilizes the presented approximate realization algorithm. Simulation results show that the method is likely to provide an accurate estimate for the time delay in a dynamical system.

I. INTRODUCTION

In general, the realization problem refers to finding an equivalent internal description of a dynamical system from an external one. An external description can be any set of input-output data like an impulse response matrix or a frequency response function. An internal description on the other hand, comprises the mapping from inputs to internal state variables of the dynamical system and the mapping from these states to the outputs. So more specifically, the realization problem refers to finding a state-space representation for a given input-output mapping.

The formulation of the realization problem originates from work on describing linear time-invariant dynamical systems [1], [2]. The algorithm in [3] provides one of the first solutions to the minimal realization problem for a linear, finite-dimensional dynamical system. An overview of various results, related to the minimal state space realization problem, is given in [4].

Next to addressing a fundamental issue in system theory, the minimal realization problem also forms a basis for the more practical problem of system identification. Loosely speaking, system identification can be seen as a realization problem that is complicated by noise, nonlinear effects, modeling errors, and so on. In this context we mention the method given in [5] that, by applying a singular value decomposition (SVD), results in a low-order approximate realization for a measured impulse response.

The advantages of the mentioned approximate realization method is its straightforward numerical implementation, the manifest relation with the original data throughout the algorithm, and its computational simplicity. However, a problem not covered by the method explicitly is the presence of a time delay in the system input-output map.

Numerous strategies have been proposed to identify the time delay (dead time) in a dynamical system. In [6] a maximum likelihood estimator is proposed to estimate the transfer function of a linear continuous time system with time delay. A drawback of this method is the fact that a good initial estimate is required in order to guarantee convergence of the cost function to its global minimum. The method presented in [7] uses least squares estimation to determine model parameters from a time-domain residual. This approach requires an inverse Laplace transformation and knowledge of the model structure in advance. Adaptive identification schemes for systems with time delay are suggested in [8], [9]. These adaptive methods perform well but requirements like measurable states and/or oscillating inputs limit their practical value.

This paper presents the application of the approximate realization method to dynamical systems with a time delay. First we give a brief introduction on the approximate realization problem and we present an algorithm for obtaining an approximate realization from impulse response data. Then we will discuss the limitations that time delays impose on this realization algorithm. The main contribution of this paper is a method to determine the time delay in a linear time-invariant (LTI) dynamical system from an input-output map. In our pragmatic approach we calculate approximate realizations from input-output maps with different delay values. A time delay estimate is then obtained by determining the minimal error between the time-domain response of the resulting realizations and the original data. Simulation results for a test system are presented to show that the proposed method is suitable for estimating the time delay in a LTI system.

II. APPROXIMATE REALIZATIONS

In this section we recall the different steps of an algorithm known from the literature that calculates an approximate realization from impulse response data. We will demonstrate the use of this algorithm in a simulation example.

A. Approximate realization algorithms

We recall that the Markov parameters $G_k$ of a discrete-time LTI system are defined by

$$ G_0 = D \quad \text{and} \quad G_k = CA^{k-1}B \quad \text{for} \ k = 1, 2, \ldots $$

where $(A, B, C, D)$ represent the system matrices of a state space model. This model is said to be a realization of the sequence $\{G_k\}_{k=0}^{\infty}$ when (1) holds. It can be shown that the infinite sequence of Markov parameters corresponds to the impulse response of the discrete-time LTI system. An algorithm to obtain a minimal realization from the impulse response of a system was presented in [3].
In practice the exact Markov parameters are never available due to the presence of noise. Moreover, the number of available Markov parameters is finite in reality. A straightforward way to obtain an approximate realization for a discrete-time LTI system, is via the construction of a Hankel matrix from a finite sequence of measured Markov parameters. An approximate realization of appropriate order \( \rho \), is then calculated from the singular value decomposition (SVD) of this Hankel matrix [5]. The algorithm reads\(^1\)

1) Construct the Hankel matrix \( T \) according to
\[
T_{i,j} = \begin{cases} 
G_{i+j-1} & \text{for } i+j \leq n+1 \\
0 & \text{for } i+j > n+1 
\end{cases}
\]
from the impulse response sequence \( \{G_k\}_{k=0}^n \)

2) Compute the SVD \( H_e = U \Sigma V^T \) of the matrix \( H_e = T(1:r,1:r) \) with \( r = \lfloor n/2 \rfloor \)

3) Construct the matrices \( U_\rho = U(:,1:\rho), \ V_\rho = V(:,1:\rho) \), and \( \Sigma_\rho = \Sigma(1:\rho,1:\rho) \)

4) Construct the matrices
\[
H_a = T(2:1+r,1:1), \ H_b = T(1:1,r), \ H_c = T(1,1:\rho), \ \text{and} \ H_d = G_0
\]

5) Construct the system matrices of the realization
\[
A = \Sigma_\rho^{-\frac{1}{2}} U_\rho^T H_a V_\rho^T \Sigma_\rho^{-\frac{1}{2}} \\
B = \Sigma_\rho^{-\frac{1}{2}} U_\rho^T H_b \\
C = H_c V_\rho^T \Sigma_\rho^{-\frac{1}{2}} \\
D = H_d
\]

The presented algorithm only requires the selection of a value for the order \( \rho \) of the realization. The choice for this value \( \rho \) determines the singular values that are considered to be contributing to the order of the approximate realization. Note that when \( \rho \) is taken too large, the noise present in the Markov parameters is modelled as a part of the realization.

So far we have presented an approximate realization algorithm that can be used to obtain a state space model for a single input and single output (SISO), LTI system from impulse response data. The case of multi input and multi output (MIMO) systems will not be considered here. In order to evaluate the algorithm, simulations were carried out with a fourth order discrete-time LTI system of the form
\[
H(s) = \frac{9878s + 3.896 \cdot 10^4}{s^4 + 25.138s^3 + 11455s^2 + 6945s + 3.896 e^4}
\]
This continuous-time transfer function is transformed into a discrete time equivalent by using a zero-order hold (ZOH), yielding the transfer function given in (3). Note that the ZOH-discretization resulted in two additional zeros, see [10]. Furthermore, we point out that one of the zeros of (3) is located at \(-3.79\). The effect of this large non-minimum phase (NMP) zero is briefly discussed later on. For all simulations we have used the following settings unless stated otherwise: sample frequency \( f_s = 200 \) Hz, sequence length \( n = 1000 \), and model order \( \rho = 4 \).

The impulse response of this system was calculated and we supplied this sequence to the realization algorithm. Finally, the response of the resulting realization was compared with the original data.

The singular values \( \sigma \) of the Hankel matrix \( H_E \) are shown in Fig. 1 indicating that the appropriate order \( \rho \) is indeed equal to four. How to select the right value for \( \rho \) when the model order is unknown and the difference between the singular values is less profound, lies beyond the scope of this paper.

The time domain and frequency domain responses of the realization are depicted in Fig. 1 (right figure) and Fig. 2, showing that the algorithm results in a realization that accurately describes the original input-output data and the underlying dynamics.

Up till now we have considered delay free situations. We will now discuss the situation where the input-output data are shifted due to the presence of a time delay.

III. TIME DELAY ESTIMATION

This section deals with the problem of time delays in dynamical systems. We briefly discuss the notion of time delays and we will show the limitation that time delays impose

\(^1\)The used Matlab-like notation \( (:) \) is a shorthand for an entire row or column of a matrix. Similarly, \((a : b)\) indicates consecutive rows or columns.
A. Time delay

We consider the discrete time LTI systems with zero initial state and a time delay of $d$ samples at the input

\[
\begin{align*}
x_1(k+1) &= Ax_1(k) + Bu(k-d) \quad (4) \\
y(k) &= Cx_1(k) + Du(k-d) \quad (5)
\end{align*}
\]

or a time delay of $d$ samples at the output respectively

\[
\begin{align*}
x_2(k+1) &= Ax_2(k) + Bu(k) \quad (6) \\
y(k+d) &= Cx_2(k) + Du(k) \quad (7)
\end{align*}
\]

Note that the transfer function $H(z) = C(zI - A)^{-1}B + D$ is equal for both models, but the models have different state trajectories since $x_1(k)$ and $x_2(k)$ are related by

\[
x_2(k) = x_1(k-d)
\]

The time delay enters the transfer function $H(z)$ as $z^{-d}$ in both cases and hence it results in a pole at $z = 0$ with multiplicity $d$. In the time domain a delay of $d$ samples results in, for example, a step response $y(k)$ where the first $d$ samples are equal to zero, see also Fig. 3. Note that the case of a fractional delay (delay that is not equal to an integer number of samples) is not considered in this paper.

The effect of a time delay on the approximate realization for a system containing such a delay is illustrated in Fig. 4. From the depicted Bode diagrams we see that the approximate realization fails to describe the system dynamics properly when the system contains a time delay. This effect of a time delay on the accuracy of the approximate realization indicates that time delays should be taken into account explicitly.

B. Time delay estimation

With the observations from the previous section in mind, a natural way to obtain an approximate realization for an input-output data set with a delay of $d$ samples, is to increase the order of the realization from $\rho$ to $\rho + d$. In the remainder we will call this approach Method I. Note that this method results in a different model class than given in (4–5) and (6–7), since the parameter $d$ is replaced by additional states.

Another possibility is to remove the first $d$ samples from the input-output data and supply the remaining samples to one of the algorithms to obtain an approximate realization without any time delay. Afterwards, the time delay is included in the obtained realization by adding a term $z^{-d}$ to its transfer function. We will denote this approach by Method II. Note that for both approaches we assume that the length of the actual time delay $d$ is known.

This assumption imposes a limitation on both methods because in many cases only an estimate $\hat{d}$ of the time delay is available. Therefore, we defined the delay estimation error $e_d = \hat{d} - d$ and investigated the accuracy of the resulting realizations via simulations for cases where $e_d \neq 0$.

The accuracy of the approximate realizations was quantified by calculating the integrated absolute error (IAE) between the original input-output data and the corresponding response of the approximate realization. Next to this time domain error measure, we also calculated the $L_2$-norm of the difference $H_d(e^{j\omega}) = H_\rho(e^{j\omega}) - H(e^{j\omega})$. Here, $H_\rho(e^{j\omega})$ and $H(e^{j\omega})$ denote the frequency response functions (FRF) of the approximate realization and that of the actual system, respectively. Note that we introduce this second error measure as an additional check but that the FRF of the actual system is usually unknown in practice.

Simulations were performed for $e_d = [-5, 5]$, using the test system and simulation settings as mentioned in the previous section. In order to incorporate a 20 samples delay in the test system, we added the term $z^{-20}$ to the numerator of the transfer function in (3). The resulting errors for both methods are shown in Fig. 5 and 6.
From these results we see that Method I yields much smaller values for both the IAE and $||H_d(e^{jω})||_2$ in comparison with Method II. Close examination of the results showed that only for $e_d = 0$ the errors obtained with both methods were in the same order of magnitude. However, as can be seen from Fig. 7, the poles of the approximation obtained with Method I are not located exactly at $z = 0$. This explains the small errors resulting from Method I, since numerical inaccuracies and imperfections in the input-output data (e.g. a time delay) are compensated for by shifting the pole locations of the realization.

Furthermore, we point out that Fig. 8 only reveals 23 significant singular values instead of the expected $4 + 20 = 24$. This apparent order reduction is caused by the presence of the large NMP zero, introduced by the discretization of (2). Namely, a large NMP zero results in a large gap between $\bar{σ}(H_E)$ and $g(H_E)$ according to [11]. This large gap implies that there exists a subspace of signals that is poorly observable in the output data, resulting in a reduction of the number of significant singular values.

Nevertheless, the fact that the poles and zeros, associated with the time delay, are not located at $z = 0$ makes it difficult to distinguish them from the poles and zeros associated with the system dynamics. Furthermore, the shapes of the resulting error plots from Method II reveal a discernible difference between the correct and erroneous delay time estimates, which is not the case for Method I.

Based on these observations we propose the following time delay estimation method that is based on the previously introduced Method II. Suppose that an impulse response sequence $\{G_k\}_{k=0}^N$ of length $N$ is available and that an appropriate order $ρ$ for the corresponding dynamical system (without the time delay) is known. Supply the sequence $\{G_{k+d}\}_{k=0}^n$ with $n + \hat{d}_{\text{max}} < N$ to the approximate realization algorithm and determine for which value of $\hat{d}$ the IAE between the original data and the response of the resulting realization, including the added time delay $z^{-\hat{d}}$, is minimal. Note that a reasonable value for the upper bound $\hat{d}_{\text{max}}$ is helpful to minimize computational efforts. In the remainder we will denote this time delay estimation method by Method III.

We will now show the results of simulations that were performed to evaluate the proposed estimation method. For this purpose we have used the above-mentioned test system and simulation settings. In order to make the simulation more realistic, zero-mean white noise with variance of $4 \cdot 10^{-6}$ was added to the impulse response sequence and the proposed time delay estimation method was applied to the resulting data set. The results are shown in Fig. 9 and 10. For comparison we have included the results for the noisy data sequence, obtained with Method I.

From these results we conclude that Method III is capable to provide an estimate of the actual time delay. On the other hand, the results from Method I are not suitable to determine the time delay from noisy data.
Fig. 9. Integrated absolute error between impulse responses of the system with measurement noise (\(\sigma^2_v = 4 \cdot 10^{-6}\)) and the approximate realizations as function of \(e_d\): Method I (left figure) and Method III (right figure).

Fig. 10. \(L_2\)-norm of difference between frequency response functions of the system with measurement noise (\(\sigma^2_v = 4 \cdot 10^{-6}\)) and the approximate realizations as function of \(e_d\): Method I (left figure) and Method III (right figure).

Although the minimal values in the IAE and \(|H_d(e^{j\omega})|_2\) plots from Method III indicate the correct time delay value, the FRF of the resulting approximate realizations show large magnitude deviations at high frequencies from that of the actual system. See Fig. 11. Note that the phase is still accurate for the realization obtained with Method III.

This result indicates that the approximate realization algorithm is sensitive to noise in the input-output data set. Although this implies that the finally obtained approximate realization is not very accurate, especially at higher frequencies, additional simulations showed that the resulting time delay estimates are reliable for the test system under study, despite these amplitude mismatches at high frequencies. See also Fig. 12. Similar results were obtained for a modified test system that was obtained by manually removing the discrete zeros from (3). Successful application of the method to estimate the time delay in a process control valve is discussed in [12].

Fig. 12. Histogram of estimation errors \(e_d\) with Method III for 25 additional simulations.

IV. CONCLUSIONS

In this paper we have discussed the application of an approximate realization algorithm to dynamical systems with a time delay. After introducing the algorithm to obtain an approximate realization from input-output data, we have illustrated the limiting effect of time delays on the accuracy of this algorithm. Based on this discussion, we proposed a time delay estimation method that utilizes an approximate realization algorithm.

The time delay estimation method was then evaluated by performing simulations with a test system. The results of these simulations showed that the method is capable to provide an accurate time delay estimate from noisy impulse response sequences for this test system. This in contrast to the approximate realization algorithm itself, which appeared to be sensitive to measurement noise. Given these encouraging results, it seems worthwhile to investigate whether the proposed method is also suitable for other LTI dynamic systems with time delays.

Furthermore, additional research is needed to make the approximate realization algorithm more robust to measurement noise and to evaluate the statistical properties of the time delay estimator. This in order to increase the practical value of the method as an identification tool. Moreover, the calculation time of the time delay estimation method could be improved by implementing an optimization step to determine the delay value that minimizes the integrated absolute error between realization response and original data. Finally, the effect of NMP zeros and the extension of the method to unstable systems are interesting topics for future work.
REFERENCES


