Analysis of a Phase Method for Time-Delay Estimation

Svante Björklund

Control & Communication
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: http://www.control.isy.liu.se
E-mail: svabj@isy.liu.se

15th October 2002

Report no.: LiTH-ISY-R-2467

Technical reports from the Control & Communication group in Linköping are available at http://www.control.isy.liu.se/publications.
Abstract

In this report we study estimation of time delays in linear dynamical systems with additive noise. Estimating the time delay is a common engineering problem, e.g. in automatic control, system identification and signal processing.

This report reviews and analyzes a certain time-delay estimation method: A discrete-time model of a continuous-time system is identified. The non-minimum phase zeros form the allpass part of the model. The dead-time is estimated by looking at the slope at low frequencies of the phase of the allpass part.

We have experimentally studied the estimation method with the help of simulations in open loop and closed loop. We have looked at the estimates, poles, zeros and the phase. We have also performed statistical analyses with RMS (root mean square) error plots and with confidence interval calculations.

Our results show that the estimation method can be used with several different model structures of the discrete-time model. The estimation method is, however, non-robust and can totally fail in some cases. The probability of failure depends on the used model structure, the input signal type and the signal-to-noise ratio. The estimation method can be made more robust by removing false non-minimum phase zeros, caused by the noise, in a certain way. It is usually better not to prewhite the data applied to the estimation method. For input (reference) signals of the form of steps, the Laguerre model structure seems to be an appropriate choice in open (closed) loop.

Keywords: time-delay, dead-time, estimation, system identification, Laguerre, linear dynamic system, non-minimum phase zero, ANOVA, confidence interval
In this report we study estimation of time delays in linear dynamical systems with additive noise. Estimating the time delay is a common engineering problem, e.g. in automatic control, system identification and signal processing.

This report reviews and analyzes a certain time-delay estimation method: A discrete-time model of a continuous-time system is identified. The non-minimum phase zeros form the allpass part of the model. The dead-time is estimated by looking at the slope at low frequencies of the phase of the allpass part.

We have experimentally studied the estimation method with the help of simulations in open loop and closed loop. We have looked at the estimates, poles, zeros and the phase. We have also performed statistical analyses with RMS (root mean square) error plots and with confidence interval calculations. Our results show that the estimation method can be used with several different model structures of the discrete-time model. The estimation method is, however, non-robust and can totally fail in some cases. The probability of failure depends on the used model structure, the input signal type and the signal-to-noise ratio. The estimation method can be made more robust by removing false non-minimum phase zeros, caused by the noise, in a certain way. It is usually better not to prewhite the data applied to the estimation method. For input (reference) signals of the form of steps, the Laguerre model structure seems to be an appropriate choice in open (closed) loop.
Contents

1 Introduction

2 Dead-time estimation by the phase of the allpass part
   2.1 Using the phase of the discrete-time allpass part
   2.2 Laguerre and other model structures
   2.3 Examples of using the DAP method
      2.3.1 A successful estimation
      2.3.2 A failing estimation
   2.4 A solution to the problem with the DAP method
      2.4.1 The uncertainty in the zeros and the dead-time estimate
      2.4.2 A solution

3 Statistical analysis
   3.1 Open loop simulations
      3.1.1 Open loop simulation setup
      3.1.2 Evaluation methods
      3.1.3 Open loop simulation results
   3.2 Closed loop simulations
      3.2.1 Closed loop simulation setup
      3.2.2 Closed loop simulation results

4 Discussion and conclusions
   4.1 Discussion
      4.1.1 Model structures
      4.1.2 The DAP method
      4.1.3 Zero guarding
      4.1.4 The simulation results
      4.1.5 Comparison with other’s results
      4.1.6 Materials and methods
   4.2 Conclusions
   4.3 Future work
References

A ANOVA and confidence intervals

A.1 ANOVA and confidence intervals to find the best zero guarding
1 Introduction

The problem we address in this report is the problem of estimating *time delays* in linear dynamical systems with additive noise. A synonym for time delay is *dead-time*. Estimating the time delay is a common engineering problem, e.g. in control performance monitoring of industrial processes, in design and tuning of controllers, in range estimation in radar and in direction estimation by time delay of arrival in signal intelligence. Dead-time estimation is also a necessary part in all system identification.

In [Isa97] a method for dead-time estimation is described. A discrete-time model of a continuous-time system is identified. The zeros of the model are translated to continuous-time. By comparing the dead-time with a Padé approximation, the dead-time may be estimated from the continuous-time non-minimum phase zeros. An improved method is described in [Hor00, IHD01] in which the discrete-time non-minimum phase zeros form the allpass part of the model, which directly represents the dead-time. The dead-time is estimated by looking at the slope at low frequencies of the phase of the allpass part. In [Isa97] the discrete-time model is identified using general orthonormal bases. In [Hor00, IHD01] a particular such basis is used, the Laguerre basis. The method in [Hor00, IHD01] shows very good results in both open loop and closed loop for input and reference signals in the form of steps. This method will in this report be called *Laguerre DAP*. The portion of the method from the discrete-time model to the dead-time estimate we call the *DAP* (Discrete-time Allpass part Phase) method.

The purpose with this work is to

- Review the Laguerre DAP method.
- Describe why the Laguerre DAP method appears to work so well.
- Investigate if the Laguerre DAP method has some drawbacks.

In Section 2 we describe the Laguerre DAP method and give some examples of its use. We will discover some problems and suggest a solution. In Section 3 we will perform statistical analyses of some DAP methods. After that, in Section 4 we have a discussion and give some conclusions.
2 Dead-time estimation by the phase of the allpass part

We will start by describing the DAP method in the first section. In Section 2.2 we describe the Laguerre model structure, which uses the Laguerre basis. In the following section, Section 2.3, we will give examples of dead-time estimation with the DAP method and we will see that it has an essential problem. Finally, in Section 2.4 we suggest a method to circumvent the problem.

2.1 Using the phase of the discrete-time allpass part

Assume the true continuous-time system is $G(s) = G_1(s) \cdot e^{-sT_d} = G_1(s) \cdot G_{ap}(s)$. The system $G_1(s)$ is linear and time invariant. $G_{ap}(s)$ is the dead-time. We estimate a discrete-time rational linear model $G(z)$ of $G(s)$ and factorize $G(z)$ into a minimum-phase system $G_1(z)$ and an allpass system $G_{ap}(z)$: $G(z) = G_1(z)G_{ap}(z)$. (Allpass means $|G_{ap}(e^{j\omega})| \equiv 1$)

We then consider $G_{ap}(e^{j\omega T_d})$ as an approximation of the dead-time $G_{ap}(e^{j\omega}) = e^{-j\omega T_d}$ [Hor00, IHD01]. A pure dead-time is of course an allpass system. We get the allpass part $G_{ap}(z)$ of $G(z)$ if we put all non-minimum phase (outside the unit circle) zeros of $G(z)$ into $G_{ap}(z)$ and add poles to $G_{ap}(z)$ which are the non-minimum phase zeros mirrored in the unit circle.

If we approximate a continuous-time linear system $G_{ap}(s)$ with a discrete-time linear system $G_{ap}(z)$ by sampling, then the frequency functions $G_{ap}(j\omega)$ and $G_{ap}(e^{j\omega T_s})$ will agree well for low frequencies [GL97, p. 115-116]. [GL97] has the rule-of-thumb that frequencies up to $1/10$ of the sampling frequency gives a good agreement.

From the allpass model $G_{ap}(e^{j\omega T_s})$ the dead-time can be estimated by taking the derivative of the phase $\varphi(\omega) = \arg G_{ap}(e^{j\omega T_s})$. The motivation for this is that the dead-time of the true system is given in this way:

$$d\varphi = d\arg G_{ap}(j\omega) = \frac{d}{d\omega}e^{-j\omega T_d} = \frac{d}{d\omega}\{-j\omega T_d\} = -T_d$$ (2.1)

Our dead-time estimate $\hat{T}$ is given by an approximation of the derivative of the phase $\varphi(\omega) = \arg G_{ap}(e^{j\omega T_s})$:

$$T_d \approx -\frac{\varphi(\omega T_s)}{\omega T_s} = -\frac{\arg G_{ap}(e^{j\omega T_s})}{\omega T_s}$$ (2.2)

for a small $\omega$ and adding a 1 because of the extra time delay that is created by the sampling:

$$\hat{T}_d = -\frac{\arg G_{ap}(e^{j\omega T_s})}{\omega T_s} + 1; \quad \omega \text{ small}$$ (2.3)

This approach is described in [Hor00, IHD01]. (In [Hor00, IHD01] a different and more complicated derivation is done.) In [Hor00, IHD01] $\omega = 10^{-4}$ is suggested. The method in equation 2.3 is in this report called DAP (Discrete-time Allpass Phase). Other forms of approximations of the derivative should also be possible.
2.2 Laguerre and other model structures

To use the estimation method in the previous section we first must estimate a discrete-time model of the true system with dead-time. One interesting model structure to use is the Laguerre model structure. In this section we briefly describe the Laguerre model structure.

If a discrete-time system $G(z)$ is strictly proper, asymptotic stable and continuous in $|z| \geq 1$, then it can be written [Wah91, p. 552]:

$$G(z) = \sum_{k=1}^{\infty} d_k \frac{\sqrt{(1-\alpha^2)T_s}}{z-\alpha} \left( \frac{1-\alpha z}{z-\alpha} \right)^{k-1} = \sum_{k=1}^{\infty} d_k L_k(z)$$  \hspace{1cm} (2.4)

with $|z| \geq 1$ and $-1 < \alpha < 1$. $T_s$ is the sampling interval and $d_k$ are some coefficients.

We change from the Z-transform variable $z$ to the delay operator $q^{-1}$ and write the functions $L_k(z)$ as

$$L_k(q) = \frac{\sqrt{(1-\alpha^2)T_s q^{-1}}}{1-\alpha q^{-1}} \left( \frac{-\alpha + q^{-1}}{1-\alpha q^{-1}} \right)^{k-1}.$$  \hspace{1cm} (2.5)

These functions are called the discrete-time Laguerre functions in the frequency domain.

A model with a finite number of Laguerre functions looks like

$$y(t) = \hat{G}(q)u(t) + v(t)$$  \hspace{1cm} (2.6)

$$\hat{G}(q) = \sum_{k=1}^{n_{\text{lag}}} d_k L_k(q)$$  \hspace{1cm} (2.7)

with $y(t)$, $u(t)$ and $v(t)$ being the output, input and noise signals respectively. The model $\hat{G}(q)$ is an approximation of the true system $G(z)$ in equation (2.4).

In [Hor00, IHD01] a Laguerre model was used. There is, however, nothing that prevents us from using a model of a different model structure and estimate the dead time from its allpass part. For example, FIR, ARX or output error (OE) model structures [Lju99] could be used. We will see examples on this later.

2.3 Examples of using the DAP method

In order to better understand how the DAP method works, we will in this section in detail calculate the DAP estimate using a Laguerre model for two simulated trials. In one of the trials the method will work well but in the other it will fail completely.

By looking at how the phase of a model $G(e^{i\omega T_s})$ can be calculated from the poles and zeros of $G(z)$ we can get some insight into what happens in the calculation of the dead-time estimate in equation (2.3).

If we have a rational discrete-time transfer function

$$G(z) = \beta \frac{(z-z_1)(z-z_m)}{(z-p_1)\ldots(z-p_n)}.$$  \hspace{1cm} (2.8)
its frequency function will be
\[
G(e^{j\omega T_s}) = \beta \frac{(e^{j\omega T_s} - z_1) \cdots (e^{j\omega T_s} - z_m)}{(e^{j\omega T_s} - p_1) \cdots (e^{j\omega T_s} - p_n)}. \tag{2.9}
\]

Let \( \bar{Z}_k(\omega) \) be the vector \((e^{j\omega T_s} - z_k)\) from \(z_k\) to the point \(e^{j\omega T_s}\) on the unit circle. Write \( \bar{Z}_k(\omega) = Z_k(\omega)e^{i\phi_k(\omega)} \). We do the same for the poles. \( \bar{P}_k(\omega) \) is the vector \((e^{j\omega T_s} - p_k)\) from \(p_k\) to the point \(e^{j\omega T_s}\) on the unit circle and \( \bar{P}_k(\omega) = P_k(\omega)e^{i\theta_k(\omega)} \). Then, we can write the phase of \( G(e^{j\omega T_s}) \) as
\[
\arg G(e^{j\omega T_s}) = \arg \beta + \sum_{k=1}^{m} \psi_k(\omega) - \sum_{k=1}^{n} \vartheta_k(\omega). \tag{2.10}
\]

Now, if we write the phase of the allpass part \( G_{ap}(z) \) of an identified model as in equation (2.10) and use equation (2.3) we receive at the expression
\[
\hat{T}_d = 1 - \frac{\text{unwrap}(\arg G_{ap}(e^{j\omega_1 T_s}))}{\omega_1 T_s} = 1 - \frac{1}{\omega_1 T_s} \text{unwrap}\left(\arg \beta + \sum_{k=1}^{m} \psi_k(\omega_1 T_s) - \sum_{k=1}^{n} \vartheta_k(\omega_1 T_s)\right) \tag{2.11}
\]
for the dead-time estimate. \( \text{unwrap}(\cdot) \) means replacing a phase outside \([-\pi, \pi]\) with the corresponding phase inside \([-\pi, \pi]\).

2.3.1 A successful estimation

We will give an example of when the DAP method works well. The linear continuous-time system
\[
G_2(s) = e^{-\theta_s} \frac{1}{(s + 1)(0.1s + 1)} \tag{2.12}
\]
(same as system \(G_2\), equation (3.3) in Section 3.1.1) was simulated by the function \texttt{lsim} in \([\text{CST}]\) in continuous-time with a random binary input signal with frequency contents between 10%-30% of the Nyquist frequency. The signal was generated by the function \texttt{idinput} in \([\text{SIT}]\) and was the same as the signal \( \text{“RBS 10-30%” in Section 3.1.1.} \) To the output signal, white Gaussian noise was added and the signal-to-noise ratio (SNR) was 10. The sampling interval \(T_s\) was \(T_s = 1\) and \(\omega_1 = 10^{-4}\). A Laguerre model (Section 2.2) was identified from the input-output data and the DAP method was used to estimate the dead-time. We will now step for step go through the calculation of \( \hat{T}_d \).

The simulation resulted in the poles and zeros of the identified Laguerre model depicted in Figure 2.1. Table 2.1 lists \(\psi_k(\omega_1)\) and \(\vartheta_k(\omega_1)\).

The dead-time estimate is
Figure 2.1: Pole-zero plot of the identified Laguerre model for the subsequent successful dead-time estimation ($\hat{T}_d = 9.7114$). The model steady-state gain is $-2.366$. The setup is as in Section 3.1.1: The system $G_2$ was simulated with the input signal RBS 10-30%. The signal-to-noise ratio (SNR) was 10. The Laguerre model was identified from the input-output data. No prewhitening or zero guarding was used. The poles of the whole model should all be located at $\alpha = 0.8$ but due to well-known numerical problems with multiple poles they are somewhat spread in the figure.

<table>
<thead>
<tr>
<th>$\psi_k(\omega_1)$</th>
<th>$\vartheta_k(\omega_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.70495</td>
<td>0.86837</td>
</tr>
<tr>
<td>1.70491</td>
<td>-0.86813</td>
</tr>
<tr>
<td>3.14135</td>
<td>0.000347</td>
</tr>
</tbody>
</table>

Table 2.1: $\psi_k(\omega_1)$ and $\vartheta_k(\omega_1)$ in equation 2.10 for $\omega_1 = 10^{-4}$ for the successful estimation. $\arg \beta = 3.14159$. 
\begin{align*}
\hat{T}_d &= 1 - \frac{\text{unwrap}(\text{arg } G_{\text{gap}}(e^{i\omega_1 T_s}))}{\omega_1 T_s} \\
&= 1 - \frac{1}{\omega_1 T_s} \text{unwrap} \left( \text{arg} \beta + \sum_{k=1}^{m} \psi_k(\omega_1 T) - \sum_{k=1}^{n} \vartheta_k(\omega_1 T_s) \right) \\
&= 1 - 10^4 \cdot \text{unwrap} \left( 3.14159 - 1.70495 + 1.70491 + 3.14135 \right. \\
&\left. - 0.86837 + 0.86813 - 0.00035 \right) = 1 - 10^4 \cdot \text{unwrap} \left( 6.28231 \right) \\
&= 1 - 10^4 \cdot (-0.00087114) = 9.7114,
\end{align*}

which is a good estimate since the true dead-time is 10.

### 2.3.2 A failing estimation

We will now study a case when the dead-time estimation failed. The only difference to the successful trial in the simulation setup is a different noise realization. The zeros and poles of the identified Laguerre model are shown in Figure 2.2. We note that the zero on the real axis just inside the unit circle in the successful simulation has moved to just outside the unit circle. We will later see what impact this will have on the dead-time estimate. Table 2.2 lists \( \psi_k(\omega_1) \) and \( \vartheta_k(\omega_1) \).

![Pole-zero plot of the identified Laguerre model for a subsequent failing dead-time estimation (\( \hat{T}_d = 1321.86 \)). The simulation setup is the same as in Figure 2.1. This simulation was one of only 3 out of 1024 simulations with failing dead-time estimation for SNR = 10. With a lower SNR the percentage of failing estimations is much higher. See Section 2.4.1. See Figure 2.1 for a comment on the location of the poles.](image)

The dead-time estimate will now be

\begin{align*}
\hat{T}_d &= 1 - \frac{\text{unwrap}(\text{arg } G_{\text{gap}}(e^{i\omega_1 T_s}))}{\omega_1 T_s} \\
&= 1 - 10^4 \cdot \text{unwrap} \left( 3.14159 - 1.70495 + 1.70491 + 3.14135 \right. \\
&\left. - 0.86837 + 0.86813 - 0.00035 \right) = 1 - 10^4 \cdot \text{unwrap} \left( 6.28231 \right) \\
&= 1 - 10^4 \cdot (-0.00087114) = 9.7114,
\end{align*}
Table 2.2: $\psi_k(\omega_1)$ and $\theta_k(\omega_1)$ in equation 2.10 for $\omega_1 = 10^{-4}$ for the failing estimation. arg $\beta = 0$. The values in italics are new compared to the successful estimation in Table 2.1. The new values are depend on the extra zero outside the unit circle. The other values are similar to those in Table 2.1. The value of arg $\beta$ is changed.

$$
\begin{array}{|c|c|}
\hline
\psi_k(\omega_1) & \theta_k(\omega_1) \\
\hline
-1.71186 & 0.84682 \\
1.71182 & -0.84658 \\
3.14125 & 0.000445 \\
3.07613 & 0.06556 \\
\hline
\end{array}
$$

As can be seen, the estimate is completely wrong.

2.4 A solution to the problem with the DAP method

As already mentioned in the previous section, a difference between the identified models in the successful and the failing trials was that the zero on the real axis just inside the unit circle in the successful trial had moved to outside the unit circle. Let us see what happens if we simply remove the moved zero from the model with the failing estimate. This results in the $\psi_k(\omega_1)$ and $\theta_k(\omega_1)$ in Table 2.3. The remaining $\psi_k(\omega_1)$ and $\theta_k(\omega_1)$ values in Table 2.3 are similar to the ones for the successful estimation in Table 2.1.

Table 2.3: $\psi_k(\omega_1)$ and $\theta_k(\omega_1)$ in equation 2.10 for $\omega_1 = 10^{-4}$ for the model in Section 2.3.2 when the real zero just outside the unit circle has been removed. arg $\beta = 3.14159$.

$$
\begin{array}{|c|c|}
\hline
\psi_k(\omega_1) & \theta_k(\omega_1) \\
\hline
-1.71186 & 0.84682 \\
1.71182 & -0.84658 \\
3.14125 & 0.000445 \\
\hline
\end{array}
$$

The dead-time estimate will now be

$$
\hat{T_d} = 1 - \frac{\text{unwrap} \left( \text{arg} G_{\text{ap}}(e^{i\omega_1 T_s}) \right)}{\omega_1 T_s} =
\begin{align*}
&= 1 - \frac{1}{\omega_1 T_s} \text{unwrap} \left( \text{arg} \beta + \sum_{k=1}^{m} \psi_k(\omega_1 T) - \sum_{k=1}^{n} \theta_k(\omega_1 T_s) \right) = \\
&= 1 - 10^4 \cdot \text{unwrap} \left( 3.14159 - 1.71186 + 1.71182 + 3.14125 \right) \\
&\quad -0.84682 + 0.84658 - 0.000445 = 1 - 10^4 \cdot \text{unwrap} (6.28212) = 1 - 10^4 \cdot (-0.00106586) = 11.6586,
\end{align*}
$$
which is an acceptable estimate. It is much better than in Section 2.3.2 with the failing estimation. It seems like the only significant difference between the models with successful and failing estimation is if a certain zero is inside or outside the unit circle.

### 2.4.1 The uncertainty in the zeros and the dead-time estimate

We note in equation (2.10) that if we move a zero from inside to outside of the unit circle when \( \omega = 0 \), then \( \arg G(e^{i\omega T_s}) \) will increase by \( \pi \). If \( \omega \) is not zero but very small, the phase increase will be between 0 and \( \pi \). When this extra phase is divided by a very small \( \omega_1 \) in equation (2.3), the dead-time estimate will get a large bias that in most cases will dominate the estimate.

The reason for zeros falling on the incorrect side of the unit circle is the noise. Figures 2.3, 2.4 and 2.5 show examples of the spread of zeros and poles due to the noise.

![Figure 2.3: Pole-zero plot with estimated uncertainty regions (3 standard deviations) and with zeros and poles from 1024 simulated trials for SNR = 10. The arrows indicate the extent of the uncertainty region for real zeros and poles. The estimated uncertainty regions seem to be too large. The zeros and poles from the 1024 trials give a better view of the uncertainty. There appears to be a risk of the zeros falling on the erroneous side of the unit circle. The simulation setup is as in Section 3.1.1 with system \( G_2 \) in open loop, signal RBS 10-30%, no prewhitening and no zero guarding.](image)

Figure 2.6 displays the dead-time estimate for different locations of the zero closest to +1. If the zero is inside the unit circle the estimates are reasonable. Outside the unit circle the estimates are poor. The worst estimates are close to the unit circle but outside it. Farther away from the unit circle (outside it) the estimate become better and better. The maximum possible estimation error occurs for the maximum phase error in \( \arg G_{ap}(e^{i\omega T_s}) \), which is \( \pi \). The estimation error will then be \( \hat{T}_d = T_d - T_d = \pi/\omega_1 = 3.1416/10^{-4} = 31416 \).
Figure 2.4: Pole-zero plot with estimated uncertainty regions (3 standard deviations) and with zeros and poles from 1024 simulated trials for SNR = 1. As can be seen, the risk of a zero falling on the wrong side of the unit circle is significant. The simulation setup is as in Section 3.1.1 with system $G_2$ in open loop, signal RBS 10-30%, no prewhitening and no zero guarding.

Figure 2.5: Location of the zero closest to $+1$ of a Laguerre model for 1024 trials. The complex conjugated zeros of the complex zeros are not shown. The simulation setup is as in Section 3.1.1 with system $G_2$ in open loop, signal RBS 10-30%, SNR=1, no prewhitening and no zero guarding.
Figure 2.6: Dead-time estimate for different locations of the zero closest to +1 for a Laguerre model. Zoom in on Figure 2.5.

In this calculation we have assumed that only one zero falls on the wrong side of the unit circle.

2.4.2 A solution

It appears that moving zeros located close to but outside the unit circle (back) to the inside of the unit circle is a solution to the problem with the DAP method. The motivation is that we assume that these zeros actually should be located inside the unit circle. Since we only need the allpass part in the DAP methods, we just remove some zeros outside the unit circle without putting them somewhere inside the unit circle. We call this technique for zero guarding.

Now we have the following questions:

- Shall only the zeros outside the unit circle close to +1 or all zeros outside outside and close to the unit circle be removed?
- How many zeros shall be removed? Only one or perhaps several?
- What is meant by “close”.

In the next section we will use a statistical technique to answer these questions.
3 Statistical analysis

In this chapter we will perform a statistical analysis of the DAP method and investigate the choice of zero guarding. We start by looking at open loop dead-time estimation in Section 3.1.3 and then continue with closed loop estimation in Section 3.2.2.

In order to acquire data for the statistical analysis some simulations were conducted in MATLAB. Both open loop and closed loop simulations were executed. The circumstances or factors were slightly different for the open loop and closed loop simulations because the simulation programs for open loop and closed loop were initially two different programs and they are not yet fully merged to a single program. We start by describing the open loop simulations and continue then with the closed loop simulations.

3.1 Open loop simulations

3.1.1 Open loop simulation setup

In the open loop simulations the output signal $y(t)$ was simulated as

$$y(t) = G(s)u(t) + v(t)$$

where $y(t)$, $u(t)$ and $v(t)$ are the output, input and noise signals respectively. $G(s)$ is a linear system with dead-time.

The factors of the simulations can be divided in fixed factors and varied factors. The different possible choices for the same factor are in this report called levels of the factor. The choice of level for all factors is called a factor level combination.

**Fixed factors.** Fixed factors for the open loop simulations were

- The noise $v(t)$ was white and Gaussian.
- The system $G(s)$ was simulated by the function lsim in [CST] in continuous-time.
- The sampling interval for the dead-time estimation was $T_s = 1$.

**Varied factors.** The varied factors for the open loop simulations are given the names Method, Prewhite, Sys, ZType, ZNo, ZSize, InType and SNR.

**Method.** Several dead-time estimation methods were tested:

- *Lag.dap.* The DAP method applied to a Laguerre model with 10 terms in the sum ($m_{lag} = 10$) and $\alpha = 0.8$. See Section 2.1 for a description of the DAP method and Section 2.2 for the Laguerre model structure.
- *FIRdap.* The DAP method applied to a FIR model with 15 taps.
- **ARXdap.** The DAP method applied to an ARX model \( y(t) = (B(q)/A(q))u(t) + (1/A(q))w(t) \) with \( n_a = 4 \) (the order of the \( A(q) \) polynomial), \( n_b = 15 \) (the order of the \( B(q) \) polynomial) and \( n_k = 0 \) (the number of time delays).

- **OEdap.** The DAP method applied to an OE model \( y(t) = (B(q)/F(q))u(t) + w(t) \) with \( n_b = 15 \) (the order of the \( B(q) \) polynomial), \( n_f = 2 \) (the order of \( F(q) \) polynomial) and \( n_k = 0 \) (the number of time delays).

**Prewhite.** For each of the methods above the input and output signals \( u(t) \) and \( y(t) \) were either

- **pw.** Filtered through a filter that made the input white:
  1. Remove the mean value of \( u(t) \) and \( y(t) \).
  2. Estimate a 10th order AR model for \( u(t) \).
  3. Filter \( u(t) \) and \( y(t) \) through the AR model.

- **nopw.** Not filtered through a prewhitening filter.

It was possible to adjust the DAP methods by removing one or several non-minimum-phase (outside the unit circle for discrete-time systems) zeros close to the point +1 or close to the unit circle. We call this zero guarding. The use of it will be discussed in Section 3.1.3. There are three factors, called ZType, ZSize and ZNo related to zero guarding:

- **ZType.** Remove non-minimum-phase zeros close to the point +1 or close to the unit circle (ucirc) or remove no zeros.
- **ZSize.** How far away from +1 or the unit circle to remove non-minimum-phase zeros.
- **ZNo.** The maximum number of zeros to remove within ZSize. The zeros closest to +1 or the unit circle are removed first.

**Sys.** The following systems were simulated:

- A slow second order system:
  \[
  G_1(s) = e^{-9s} \frac{1}{(10s + 1)(s + 1)}
  \]  
  (3.2)

- A fast second order system:
  \[
  G_2(s) = e^{-9s} \frac{1}{(s + 1)(0.1s + 1)}
  \]  
  (3.3)

- A fourth order system with real poles. The poles and zeros lie between the poles of system \( G_1 \):
  \[
  G_5(s) = e^{-9s} \frac{0.05(s + 0.9)(s + 0.4)}{(s + 1)(s + 0.6)(s + 0.3)(s + 0.1)}
  \]  
  (3.4)
A fourth order system with complex poles. The poles have the same distance to the origin as in system $G_1$:

$$G_6(s) = e^{-0.9s} \frac{1/36(s + 0.9)(s + 0.4)}{(s - 0.1\sqrt{2}(-1 + i))(s - \sqrt{2}(-1 + i))}$$

(3.5)

Note that for all the systems above, the time delay will be 10 after the sampling.

**InType.** The input signal was 500 samples long and could be of three different types:

- RBS 10-30%. (Random Binary Signal) with frequency contents between 10%-30% of the Nyquist frequency. It was generated by the function `idinput` in [SIT].
- RBS 0-100%. RBS with frequency contents between 0%-100% of the Nyquist frequency, i.e white noise. It was generated by the function `idinput` in [SIT].
- Steps. Three steps of the form (in MATLAB code): `[zeros(50,1); ones(150,1); -ones(150,1); zeros(150,1)]`.

**SNR.** The signal to noise ratio could be chosen. The used SNR has been in the interval 1 to 100.

Many of the factors and levels of the open loop simulations were chosen to be the same as in [Hor00], which makes comparisons easy. The names of the factors and factor levels in this section are also used in the following plots.

### 3.1.2 Evaluation methods

The following methods for evaluating performance and properties of dead-time estimation methods were utilized:

- Study of dead-time estimates, poles, zeros and phases of identified models.
- Pole-zero plots of identified models.
- Plot of dead-time estimates versus time. This makes it easy to see if some estimates are very large or very small and for which factor level combinations they are so.
- Root Mean Square (RMS) error of dead-time estimates for different factor level combinations. This is our main measure of how good an estimation method is.
- Mean of dead-time estimates for different factor level combinations.
- Standard deviation of dead-time estimates for different factor level combinations. Mean and standard deviation give a more detailed picture of a method’s performance than the RMS error.
- Analysis of variance (ANOVA) and confidence intervals. These analysis methods make it possible to give statements and conclusions with a level of confidence or a specified risk of being false. See appendix A and for example [Mon97] for more information.
When calculating the mean value, standard deviation and RMS error, normally the average of the levels are used for the factors which are not displayed in the graphs. This gives an average behavior over the factors which are not studied in the graphs. If this is to give a fair picture of the mean value, standard deviation and RMS error, it is important that the averaged levels are representative of the intended use of the dead-time estimation. For example, assume a graph depicts the RMS error for all combinations of estimation method and system. The levels of all other factors will be averaged out. If we know that the SNR typically lies between 10 and 100, then taking the average for SNR equal to 0.1 and 100 might give an RMS error which is too high in the graph (if the RMS error increases when the SNR decreases). We have tried to choose factor levels in the simulations which are representative of typical control applications.

When calculating the standard deviation and RMS error it is also possible to take the maximum over the levels of the factors not displayed in the graphs. In this way the graph will show the “worst case” behavior. The worst-case over the different systems $G(s)$ has been used in some graphs.

Not all evaluation methods above will be shown in the graphs in this report but they have been used during the work preceding this report.

### 3.1.3 Open loop simulation results

The result of the first open loop simulation we will look at is depicted in Figure 3.1. There, the RMS error of the DAP dead-time estimate is displayed for several model structures ($\text{Method}$). Estimation both with and without prewhitening ($\text{Prewhite}$, Section 3.1.1) was used. The SNR was high (SNR $\approx$ 100) or low (SNR $\approx$ 1). The three different input signal types ($\text{InType}$) of Section 3.1.1 (RBS 10-30%, RBS 0-100% and Steps) were used. The four systems ($\text{Sys}$) in Section 3.1.1 were employed. For each level combination of the factors $\text{Method}$, $\text{Prewhite}$, $\text{InType}$ and $\text{Sys}$ the RMS error was estimated from 1024 trials. Then for each level combination of the other factors, the system with the worst, i.e. highest, RMS error was plotted. Which systems that were chosen is not shown in the figure.

As can be seen in Figure 3.1, some of the DAP methods exhibit very large RMS errors for some combinations of $\text{InType}$, $\text{Prewhite}$ and $\text{SNR}$ . The methods totally fail for these factor level combinations. If we look at the estimates in these cases, we notice that some estimates are very large, up to about 30000 (not shown here). Compare with Section 2.4.1.

We notice in the figure that Laguerre DAP without prewhitening is the best method for the input signal type Steps at high SNR. It is also among the best methods for low SNR. The methods $\text{FIRdap}$ and $\text{ARXdap}$ with prewhitening succeed with all input types for high SNR but not for low SNR. $\text{ARXdap}$ at high SNR also succeeds with all input types without prewhitening. $\text{FIRdap}$ without prewhitening succeeds with RBS 0-100% and Steps but not with RBS 10-30% at high SNR. The only input type $\text{FIRdap}$ and $\text{ARXdap}$ succeed with at low SNR is Steps and then using no prewhitening. The method $\text{OEdap}$ fails in all cases.

In the above simulation, the only system that could be estimated without failure was the fast second order system $G_2$ and then only with $\text{FIRdap}$ and $\text{ARXdap}$ . This is not shown in the graphs in this report.
Figure 3.1: RMS error for some dead-time estimation methods. The DAP methods have no zero guarding. 1024 trials. For each combination of the other factors, the worst system (with the largest RMS error) of the four systems in Section 3.1.1 is used in the plot. The axis label “SNR*InType” means that on this axis there are different (factor level) combinations of SNR and input signal type. In the same way “Method*Prewhite” means different combinations of estimation method and prewhitening. The tick mark labels tell us what factor level combinations there are in the different “rows” and “columns”. The levels of the factors are separated by an asterisk “*”. The level “10-30%” is an abbreviation of “RBS 10-30%” and “0-100%” of “RBS 0-100%”. Axis labels and tick mark labels are further explained in Section 3.1.1.
As a summary of Figure 3.1 we can say that the DAP methods seem to be non-robust and they can totally fail in some cases.

In order to mitigate the failures of the DAP methods we have tried zero guarding according to Section 2.4.2 on the Laguerre model structure. We have tried two different types of zero guarding (ZType): Remove zeros outside the unit circle close to the point +1 or remove zeros outside but close to the unit circle (ucirc). We have tried different distances (ZSize) from +1 or the unit circle to be considered as “close”: 0.05, 0.10, 0.15, 0.20, 0.25, 0.25 and 0.30. If there are several zeros within the zero guarding distance, we have tested different maximum number (ZNo) of zeros to remove: 1 to 4. The closest zeros are removed first.

We have utilized the statistical method ANOVA (Analysis of Variance) and confidence intervals (see Section 3.1.2, appendix A and [Mon97]) to discover significant differences between the different combinations of the levels of ZType, ZSize and ZNo. A simulation with 1024 trials was conducted. All three input signal types and all system types in Section 3.1.1 were used. No prewhitening was employed. The SNR was 1. The SNR was low because we want the Laguerre method with zero guarding to manage low SNRs. The trials were split into four groups of 256 trials each. Each group was used to compute an estimate \( \hat{x} \) of the RMS error. Each estimate was transformed by \( \tilde{x} = x^{-0.87} \) and applied to ANOVA as four replicates. The reason for the transformation is to (approximately) fulfill the requirement of ANOVA (see [Mon97]). Finally, confidence intervals of confidence level 95% for different combinations of zero guarding type (ZType), number (ZNo) and size (ZSize) for Laguerre DAP estimation in open loop were computed. See appendix A for details.

The result is shown in Figure 3.2 and Table 3.1. We see that there are several “good” combinations (intervals to the right in the figure) with no significant differences (with overlapping confidence intervals). It is, however, clear that removing zeros close to +1 is better than removing zeros close to the unit circle (ucirc). It is also clear that allowing for removing more than one zero is necessary. We choose one of the good combinations: ZType=+1, ZSize=0.15 and ZNo=3 for the next simulation we will present. This is some form of statistical optimization of an estimation method.

Figure 3.3 displays the result of the last open loop simulation we will show. In same manner as in Figure 3.1 it shows the RMS error for some estimation methods. The difference compared to Figure 3.1 is that here the DAP methods use the zero guarding above. We immediately observe that now no DAP methods fail but give reasonable and low RMS errors. Even OEdap works and give good estimates. Still Laguerre DAP gives the best result for Steps at high SNR. At low SNR it is among the best. The “optimum” choice of ZType, ZNo and ZSize for Laguerre DAP appears to work also for the other DAP methods.

In the above simulation with zero guarding, the fast second order system \( G_2 \) gave the best result in average (over input signal type and SNR) for all methods (model structure and with/without prewhitening). The fourth order system \( G_4 \) with complex poles gave the worst result in average for all methods. This is not shown in the graphs in this report.
43 groups have population marginal means significantly different from Group 39

Figure 3.2: Confidence intervals for different combinations of zero guarding type (ZType), number (ZNo) and size (ZSize) for Laguerre DAP estimation in open loop. The factor level combination numbers 1-48 on the vertical axis are explained in Table 3.1. The combinations of the rightmost intervals are the best since a reciprocal transformation ($\tilde{x} = x^{-0.87}$) of the RMS error estimates was used. See appendix A for details. The text just below the figure tells us that there are 43 combinations of ZType, ZNo and ZSize that are significantly worse than the one we are choosing (the 39th combination: ZType=+1, ZNo=3 and ZSize=0.15, marked with the vertical dotted lines).
Figure 3.3: RMS error of different dead-time estimation methods for open loop. The DAP methods have employed the zero guarding \(Z_{Type}=+1\), \(Z_{No}=3\) and \(Z_{Size}=0.15\). 1024 trials. For each combination of the other factors, the worst system (with the largest RMS error) of the four systems in section 3.1.1 is used in the plot. Axis labels and tick mark labels are explained in Figure 3.1 and in Section 3.1.1. Compare with Figure 3.1.
<table>
<thead>
<tr>
<th>No.</th>
<th>ZSize</th>
<th>ZNo</th>
<th>ZType</th>
<th>No.</th>
<th>ZSize</th>
<th>ZNo</th>
<th>ZType</th>
<th>No.</th>
<th>ZSize</th>
<th>ZNo</th>
<th>ZType</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.05</td>
<td>1</td>
<td>ucirc</td>
<td>17</td>
<td>0.25</td>
<td>3</td>
<td>ucirc</td>
<td>33</td>
<td>0.15</td>
<td>2</td>
<td>+1</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>1</td>
<td>ucirc</td>
<td>18</td>
<td>0.3</td>
<td>3</td>
<td>ucirc</td>
<td>34</td>
<td>0.2</td>
<td>2</td>
<td>+1</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>1</td>
<td>ucirc</td>
<td>19</td>
<td>0.05</td>
<td>4</td>
<td>ucirc</td>
<td>35</td>
<td>0.25</td>
<td>2</td>
<td>+1</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>1</td>
<td>ucirc</td>
<td>20</td>
<td>0.1</td>
<td>4</td>
<td>ucirc</td>
<td>36</td>
<td>0.3</td>
<td>2</td>
<td>+1</td>
</tr>
<tr>
<td>5</td>
<td>0.25</td>
<td>1</td>
<td>ucirc</td>
<td>21</td>
<td>0.15</td>
<td>4</td>
<td>ucirc</td>
<td>37</td>
<td>0.05</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>6</td>
<td>0.3</td>
<td>1</td>
<td>ucirc</td>
<td>22</td>
<td>0.2</td>
<td>4</td>
<td>ucirc</td>
<td>38</td>
<td>0.1</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>2</td>
<td>ucirc</td>
<td>23</td>
<td>0.25</td>
<td>4</td>
<td>ucirc</td>
<td>39</td>
<td>0.15</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>2</td>
<td>ucirc</td>
<td>24</td>
<td>0.3</td>
<td>4</td>
<td>ucirc</td>
<td>40</td>
<td>0.2</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>9</td>
<td>0.15</td>
<td>2</td>
<td>ucirc</td>
<td>25</td>
<td>0.05</td>
<td>1</td>
<td>+1</td>
<td>41</td>
<td>0.25</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>2</td>
<td>ucirc</td>
<td>26</td>
<td>0.1</td>
<td>1</td>
<td>+1</td>
<td>42</td>
<td>0.3</td>
<td>3</td>
<td>+1</td>
</tr>
<tr>
<td>11</td>
<td>0.25</td>
<td>2</td>
<td>ucirc</td>
<td>27</td>
<td>0.15</td>
<td>1</td>
<td>+1</td>
<td>43</td>
<td>0.05</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>12</td>
<td>0.3</td>
<td>2</td>
<td>ucirc</td>
<td>28</td>
<td>0.2</td>
<td>1</td>
<td>+1</td>
<td>44</td>
<td>0.1</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>13</td>
<td>0.05</td>
<td>3</td>
<td>ucirc</td>
<td>29</td>
<td>0.25</td>
<td>1</td>
<td>+1</td>
<td>45</td>
<td>0.15</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>14</td>
<td>0.1</td>
<td>3</td>
<td>ucirc</td>
<td>30</td>
<td>0.3</td>
<td>1</td>
<td>+1</td>
<td>46</td>
<td>0.2</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>15</td>
<td>0.15</td>
<td>3</td>
<td>ucirc</td>
<td>31</td>
<td>0.05</td>
<td>2</td>
<td>+1</td>
<td>47</td>
<td>0.25</td>
<td>4</td>
<td>+1</td>
</tr>
<tr>
<td>16</td>
<td>0.2</td>
<td>3</td>
<td>ucirc</td>
<td>32</td>
<td>0.1</td>
<td>2</td>
<td>+1</td>
<td>48</td>
<td>0.3</td>
<td>4</td>
<td>+1</td>
</tr>
</tbody>
</table>

Table 3.1: This table should be read together with Figure 3.2. The table is an explanation of the factor level combination numbers 1-48 on the vertical axis in Figure 3.2. “ucirc” means zeros closest to unit circle and “+1” means zeros closest to +1. ZSize, ZNo and ZType are explained in Section 3.1.1
3.2 Closed loop simulations

In automatic control of systems and processes, usually feedback is used, resulting in closed loop systems. In this section we study dead-time estimation by DAP methods in closed loop by the aid of Monte Carlo simulations and statistical analysis.

3.2.1 Closed loop simulation setup

The control system of the closed loop simulations is depicted in Figure 3.4.

![Control System Diagram](image)

Figure 3.4: The control system in closed loop simulations. \(F(s)\) is the controller, \(G(s)\) the system with the dead-time and \(H(s)\) is the noise model.

**Fixed factors.** Fixed factors for the closed loop simulations were

- The noise source \(w(t)\) was white and Gaussian.
- The continuous-time system \(G(s)\) was converted to discrete-time with zero-order hold conversion before the simulations.
- The simulation were performed in discrete time.
- The sampling interval was 1.

**Varied factors.** Most of the factors in Section 3.1.1 could also be varied in the closed loop case but sometimes the levels were different. Only one input signal type was used so the factor \(\text{InType}\) was not varied. Extra factors were \(\text{Ctrl}, \text{NoiseMod}\) and \(\text{InputPlace}\).

The same dead-time estimation methods as in Section 3.1.1, both with and without prewhitening could also be used for the closed loop simulations. For the DAP methods one or several non-minimum-phase zeros could be removed as in Section 3.1.1.

**Sys.** The following systems were simulated:
slow. A slow second order system:

\[ G_{1b}(s) = 0.5e^{-9s} \frac{1}{(10s + 1)(s + 1)} \]  

This is the same as system \( G_1(s) \) in Section 3.1.1 except for the static gain being half of \( G_1(s) \)'s gain.

fast. A fast second order system:

\[ G_{2b}(s) = 0.5e^{-9s} \frac{1}{(s + 1)(0.1s + 1)} \]  

This is the same as system \( G_2(s) \) in Section 3.1.1 except for the static gain being half of \( G_2(s) \)'s gain.

Again, the time delay will be 10 after the sampling.

**Ctrl.** The possible controllers were all PI-controller but different fast:

- A slow PI controller:
  \[ F_1(q) = 0.6 \frac{1 + Ts/50 - q^{-1}}{1 - q^{-1}} \]  

- A medium fast PI controller:
  \[ F_2(q) = 0.25 \frac{1 + Ts/5 - q^{-1}}{1 - q^{-1}} \]  

- A fast PI controller:
  \[ F_3(q) = 0.9 \frac{1 + Ts/12 - q^{-1}}{1 - q^{-1}} \]  

\( T_s \) is the sampling interval.

**NoiseMod.** The noise model \( H(q) \) could be chosen in three different ways:

- White noise filtered by
  \[ H_1(q) = \frac{0.71}{1 - 0.7q^{-1}} \]  

- Unfiltered white noise giving an output error model structure \( y(t) = G(q)u(t) + w(t) \):
  \[ H_2(q) = 1 \]  

- A noise model \( H_3(q) \) with the same denominator as system \( G_{1b} \). This would give an ARX model structure \( y(t) = (B(q)/A(q))u(t) + (1/A(q))w(t) \) with \( G_{1b} \) except for a different normalization in the numerator of \( H_3(q) \).

\[ H_3(q) = \frac{0.92}{1 - 0.368q^{-1} + 1.671 \cdot 10^{-5}q^{-2}} \]
**InType.** In the closed loop simulations, only one reference signal \( r(t) \) was used. It is here called \( \text{Steps2} \) and consisted of two steps and 2000 samples and was generated in MATLAB by: \([4 \cdot \text{ones}(1,1000), 2 \cdot \text{ones}(1,1000)]'\). This signal is not exactly the same as in [IHD01] but of a similar character.

**InputPlace.** The input to the dead-time estimation could be either

- \( u \). The input signal \( u(t) \) to the system and the output signal \( y(t) \) or
- \( r \). The reference signal \( r(t) \) and the output signal \( y(t) \).

**SNR.** Only a high SNR (mean SNR=961 with standard deviation 67) was used.

Many of the factors and levels of the closed loop simulations were chosen to be the same as in [IHD01], which makes comparisons easy.

### 3.2.2 Closed loop simulation results

We now turn to the closed loop simulations. Figure 3.5 displays the RMS error for the DAP method for several model structures. Compared to Figure 3.1 and 3.3 we now only have one type of input signal and only one SNR, see Section 3.2.2. On the other side we have a new factor, **InputPlace.** We can either let the reference signal \( r \) or the control signal \( u \) be input to the estimation. We notice in the figure that one of the DAP methods, \( OEdap \), fails for all level combinations of **InputPlace** and **Prewhite**. We have not investigated the cause of the failure but we assume it is the same as in the open loop simulations: A zero has by the noise been moved from the inside to the outside of the unit circle.

The results of a closed loop simulation with the same zero guarding as derived in Section 3.1.3 is shown in Figure 3.6. We find out that the zero guarding has reduced the error for the DAP methods. The same zero guarding as for open loop appears to work also here.
Figure 3.5: Average RMS error over 128 trials and different systems, controllers and noise models (Section 3.2.1) in closed loop for some dead-time estimation methods. No zero guarding was used. SNR according to Section 3.2.1. The reference signal was Steps2, also defined in Section 3.2.1. The axis label “inputPlace*Prewhite” means that on this axis there are different (factor level) combinations of input signal place (system input $u$ or reference signal $r$) and prewhitening. In the same way “Method” means different methods. The tick mark labels tell us what factor level combinations there are in the different “rows” or “columns”. The levels of the factors are separated by an asterisk “*”. Axis labels and tick mark labels are further explained in Section 3.1.1 and 3.2.1.
Figure 3.6: Average RMS error over 128 trials and different systems, controllers and noise models (Section 3.2.1) in closed loop for some dead-time estimation methods. The zero guarding $ZType=+1$, $ZNo=3$, $ZSize=0.15$ was used. SNR according to Section 3.2.1. The reference signal was Steps2, which is defined in Section 3.2.1. The same noise realization as in Figure 3.5. Axis labels and tick mark labels are explained in Figure 3.5 and in Section 3.1.1 and 3.2.1.
4 Discussion and conclusions

4.1 Discussion

4.1.1 Model structures

Of the tested model structures, FIR, ARX and Laguerre models can be quickly estimated by linear regression. The OE model structure, on the other hand, requires a numerical search which makes the estimation slower. If it is enough to estimate the time-delay only once, this is probably no problem but in applications where the the time delay must be estimated in real-time at a high sampling speed, the computational burden may be too high. Also in Monte Carlo simulations, like in this report, it is troublesome with the long computation time of the OE model structure. For example the computation for Figure 3.5 was about 30 hours and the most of the time was spent by estimating OE models.

The location $\alpha$ of the pole of the Laguerre model probably affects the locations of the zeros and therefore also how sensitive they are for falling on the wrong side of the unit circle due to noise.

4.1.2 The DAP method

As seen in this report, the DAP method fails if a zero erroneously falls outside the unit circle. The error in the dead-time estimation is not proportional to the error in the position of the zero. A small error in the position which does not cause the zero to fall on the wrong side of the unit circle will result in a small estimation error. But when the zero just falls on the incorrect side of the unit circle the estimation error will be large. The estimation error is larger close to the point $+1$ and smaller further away. This makes the DAP method sensitive to where the true zero is located.

When a zeros is moved from the inside to the outside of the unit circle by noise, the dead-time estimate seems to be unpredictable. The only we can say is that the estimate will be much to large. It appears like it is not possible to perform Monte Carlo simulations to see the performance in average.

An advantage of the DAP method compared to many other methods is that it can estimate time-delays that are a fraction of a sampling interval. In, for example, TDOA (time-delay of arrival) estimation of the direction of impinging electro-magnetic waves in signal intelligence or radar this is necessary.

4.1.3 Zero guarding

The zero guarding chosen in Section 3.1.3 for Laguerre in open loop seems to work also for closed loop. It also appears to function for the other model structures in both open loop and closed loop. Thus, the choice of zero guarding appears to be robust.

In Section 3.1.3 we saw that sometimes it is not enough to remove only one zero. The reason is likely that because complex zeros come in complex conjugated pairs, both zeros in the pair must be removed.
The location of the true non-minimum phase zeros probably depends on the dead-time and the sampling interval. The longer dead-time or the shorter sampling interval, the closer to the point +1 will the true zeros be. Therefore $Z_{\text{Size}}$ depends on the maximum possible dead-time and on the sampling interval.

### 4.1.4 The simulation results

A reason to why Laguerre DAP works well for step input signals could be as follows. It is well known that an estimated model will become better in frequency ranges where the input signal has much energy [Lju99]. Since a step-like input signal has its most energy at the frequencies which are used by the dead-time estimation, i.e. at low frequencies, it will result in a good dead-time estimate.

In the open-loop simulations the fast second order system gave the lowest RMS error in average. This is natural because this system exhibits a clearer start of the rise in the step response. The fourth order system with complex poles gave the highest RMS error in average. It seems that a more complex system makes the dead-time estimation more difficult.

### 4.1.5 Comparison with other’s results

For step signals in open and closed loop, we in this report agree with [Hor00, IHD01] that Laguerre DAP is a suitable estimation method. [Hor00, IHD01] have, however, only tested with step signals and not discovered that the DAP method sometimes fails.

### 4.1.6 Materials and methods

The use of ANOVA and confidence intervals for choosing a suitable signal processing method is a statistical optimization of the choice of a signal processing method. An advantage with this is that we can with a certain degree of confidence say which method is the best. ANOVA and confidence intervals can be used also on other problems than the one in this report. A drawback is that it requires many simulations that can take a long time to execute. For example, the execution time was about 90 hours to generate Figure 3.2. It is also important to choose factor levels which are representative of the intended application. Otherwise the optimization can be biased. We would optimize for the wrong application. See also the discussion in Section 3.1.2.

In this report it has not been the purpose to find the best dead-time estimation method. If we would like find the best method, we could proceed as follows:

- Choose representative factor levels for the intended application of the dead-time estimation.
- Take the worst case over environment factors. By environment factors we mean factors whose levels the estimation must manage but we cannot control, e.g. all system types.
Display the RMS error for all level combinations of method factors. By method factors we mean all factors which are choices in the estimation method, e.g. with or without prewhitening.

Use ANOVA and confidence intervals to choose an appropriate method.

In this report, it would have been better to take the worst case RMS error of the environment factors (systems, controllers and noise models) instead of the average RMS error in the statistical analysis of closed loop. Some of the analyses in open loop took the worst case RMS error of the systems. We should have had the same factors and levels for closed loop as for open loop: SNRs, systems and input (reference) signals.

We could in the plots of the RMS errors also have computed confidence intervals to see if the differences between methods are significant. However, also without confidence intervals, the estimated RMS errors become more accurate the more trials are used to estimate them. Therefore, the RMS error plots for open loop (1024 trials) should be more trustworthy than the plots for closed loop (128 trials). The closed loop simulations were more compute-intensive than the open loop simulations. That is the reason for only 128 trials in these simulations.

4.2 Conclusions

We draw the following conclusions from the work presented in this report:

- The DAP method is not restricted to the Laguerre model structure but can be used with any linear model structure, e.g. OE, ARX or FIR.
- DAP methods are inherently non-robust and can totally fail in some cases.
- The probability of failure of a DAP method depends on the model structure of the estimated model, the input signal type and the SNR.
- In failing cases, DAP methods can be made more robust, e.g. by zero guarding. A means to choose the zero guarding for a certain application is using ANOVA and confidence intervals on estimates from simulated signals.
- An appropriate choice of zero guarding appears to be robust and work for several model structures in both open loop and closed loop.
- For zero guarding of Laguerre DAP, removing zeros outside the unit circle close to +1 works better than removing zeros at other places outside but close to the unit circle. We must also allow more than one zero (a complex conjugated pair) to be removed.
- Most often the DAP methods work better without prewhitening the data.
- Laguerre DAP without prewhitening seems to be a suitable dead-time estimation method for open loop with steps as input signal.
- Laguerre DAP also seems to be a suitable dead-time estimation method for closed loop with steps as reference signal. The estimation should use the system input (not the reference signal) and no prewhitening.
- OE DAP requires much more computation time than FIR, ARX or Laguerre DAP.
4.3 Future work

Some possible future work is:

- Use the same circumstances (factor levels) for both open loop and closed loop.
- Choose zero guarding for closed loop also.
- Derive and calculate a lower bound on the RMS error.
- Test the DAP methods on real data.
- Compare the DAP methods with other time-delay estimation methods.
References

[CST] CSTB. Matlab control system toolbox, v. 5.0 (r12) or ...(r11.1). The Mathworks Inc. Requires MATLAB.


[SIT] SITB. Matlab system identification toolbox v. 4.0.5 (r11). The Mathworks Inc. Requires MATLAB.

A ANOVA and confidence intervals

A statistical method for analysis of experiments used in some scientific disciplines is ANOVA (Analysis of variance) and subsequent computation of confidence intervals, see for example [Mon97]. These analysis methods make it possible to give statements and conclusions with a certain level of confidence or a specified risk of being false. We have used ANOVA and confidence intervals to select a suitable zero guarding in Section 3.1.3.

In ANOVA the following linear statistical model is used (here for just two factors):

$$x_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk},$$

(A.1)

where $x_{ijk}$ is the response variable (the observation) $\mu$ is the mean effect, $\tau_i$ is the effect of the $i$th level of the first factor, $\beta_j$ is the effect of the $j$th level of the second factor, $(\tau\beta)_{ij}$ is the effect of the interaction of the first and second factors and $\epsilon_{ijk}$ is a random error. The factor and interaction effects are assumed to be constant. Moreover, the following is assumed:

$$\sum_i \tau_i = 0, \sum_j \beta_j = 0, \sum_i (\tau\beta)_{ij} = 0 \text{ and } \sum_j (\tau\beta)_{ij} = 0.$$

A.1 ANOVA and confidence intervals to find the best zero guarding

We have utilized ANOVA and confidence intervals to discover significant differences between the different combinations of the levels of the factors ZType, ZSize and ZNo. A simulation with 1024 trials was performed. All three input signals types and all system types in Section 3.1.1 were employed. No prewhitening was used. The SNR was low (SNR=1) because we wanted the estimation method, Laguerre DAP, with zero guarding to manage low SNRs. The trials were split into four groups of 256 trials each. Each group was used to compute an estimate $\hat{x}$ of the RMS error (“the lower error the better”). This estimate of the RMS error was used as the response variable $x$ in the ANOVA analysis. Each estimate (response variable) was transformed by $\tilde{x} = x^{0.87}$ and applied to ANOVA as four replicates. The reason for the transformation was to (approximately) fulfill the requirement of ANOVA: The errors $\epsilon_{ijk}$ in equation (A.1) should be independent and Gaussian distributed with zero mean and constant variance [Mon97]. We trust the random number generator of MATLAB to give us independent experiments.

To examine if the requirements are fulfilled it is customary to study some plots of the residuals: $\varepsilon_{ijk} = x_{ijk} - \hat{x}_{ijk}$, where $\hat{x}_{ijk}$ is an estimate of the observation $x_{ijk}$. $\hat{x}_{ijk}$ is also called the fitted value. The residuals should be Gaussian distributed (and therefore also the response variable for a fixed factor level combination), be “structure-less” (show no obvious patterns) and have constant variance [Mon97].

The residuals Figure A.1-A.2 show such plots for our problem. In Figure A.1 we observe that the residuals increase with increasing fitted value. This is a sign of that the variance is not constant ([Mon97]). Also, the residuals plotted versus time is not “structure-less”. It is harder to say whether the Gaussianity condition is met. Furthermore, in Figure A.2 we see that the size of the residuals are dependent on the factor levels for the factors InType, ZType and ZNo. A solution to this could be a variance-stabilizing transform [Mon97, p. 84-90]. Therefore the plot in Figure A.3 has been used to choose the transform $\tilde{x} = x^{-0.89}$ (see [Mon97]). Note that a reciprocal transformation was used and accordingly the response variable is inverted. Now we have “the larger the better”. In Figure A.4, A.5
and A.6 we see that the transformation was successful. See [Mon97] for the interpretation of Figure A.1-A.6 and Table A.1.

The ANOVA method for the balanced (equal number of replicates for all factor level combinations) fixed (constant) effects model, which is used in this report, is rather robust against violations of the Gaussianity condition [Mon97, p. 82] and the constant variance condition [Mon97, p. 85]. This together with the plots in this appendix assures us that ANOVA and confidence intervals can be applied at the problem in this report.

Finally, confidence intervals of confidence level 95% for different combinations of zero guarding type (ZType), number (ZNo) and size (ZSize) for Laguerre DAP estimation in open loop were computed, see Figure 3.2 and the text in Section 3.1.3.

Figure A.1: Residual analysis (without transformation).
Figure A.2: Residuals versus factor levels (without transformation).
<table>
<thead>
<tr>
<th>Source</th>
<th>Sum Sq.</th>
<th>d.f.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>InType</td>
<td>3.2947</td>
<td>2</td>
<td>1.64735</td>
<td>2628.31</td>
<td>0</td>
</tr>
<tr>
<td>ZSize</td>
<td>0.7147</td>
<td>5</td>
<td>0.14293</td>
<td>228.04</td>
<td>0</td>
</tr>
<tr>
<td>ZNo</td>
<td>1.3878</td>
<td>3</td>
<td>0.46259</td>
<td>738.05</td>
<td>0</td>
</tr>
<tr>
<td>ZType</td>
<td>3.0921</td>
<td>1</td>
<td>3.09207</td>
<td>4933.32</td>
<td>0</td>
</tr>
<tr>
<td>Sys</td>
<td>6.1532</td>
<td>3</td>
<td>2.05105</td>
<td>3272.4</td>
<td>0</td>
</tr>
<tr>
<td>InType*ZSize</td>
<td>0.6516</td>
<td>10</td>
<td>0.06516</td>
<td>103.96</td>
<td>0</td>
</tr>
<tr>
<td>InType*ZNo</td>
<td>1.487</td>
<td>6</td>
<td>0.24783</td>
<td>395.4</td>
<td>0</td>
</tr>
<tr>
<td>InType*ZType</td>
<td>0.8199</td>
<td>2</td>
<td>0.40995</td>
<td>654.06</td>
<td>0</td>
</tr>
<tr>
<td>InType*Sys</td>
<td>1.1538</td>
<td>6</td>
<td>0.19231</td>
<td>306.82</td>
<td>0</td>
</tr>
<tr>
<td>ZSize*ZNo</td>
<td>0.1412</td>
<td>15</td>
<td>0.00941</td>
<td>15.02</td>
<td>0</td>
</tr>
<tr>
<td>ZSize*ZType</td>
<td>0.2516</td>
<td>5</td>
<td>0.05031</td>
<td>80.27</td>
<td>0</td>
</tr>
<tr>
<td>ZSize*Sys</td>
<td>0.7899</td>
<td>15</td>
<td>0.05266</td>
<td>84.02</td>
<td>0</td>
</tr>
<tr>
<td>ZNo*ZType</td>
<td>0.6827</td>
<td>3</td>
<td>0.22758</td>
<td>363.09</td>
<td>0</td>
</tr>
<tr>
<td>ZNo*Sys</td>
<td>0.3814</td>
<td>9</td>
<td>0.04238</td>
<td>67.61</td>
<td>0</td>
</tr>
<tr>
<td>ZType*Sys</td>
<td>0.4674</td>
<td>3</td>
<td>0.15581</td>
<td>248.59</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZNo</td>
<td>0.1002</td>
<td>30</td>
<td>0.00334</td>
<td>5.33</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZType</td>
<td>0.0392</td>
<td>10</td>
<td>0.00392</td>
<td>6.26</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZSize</em>Sys</td>
<td>0.7176</td>
<td>30</td>
<td>0.02392</td>
<td>38.16</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZNo</em>ZType</td>
<td>0.5405</td>
<td>6</td>
<td>0.09008</td>
<td>143.72</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZNo</em>Sys</td>
<td>0.2114</td>
<td>18</td>
<td>0.01175</td>
<td>18.74</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZType</em>Sys</td>
<td>0.0467</td>
<td>6</td>
<td>0.00778</td>
<td>12.42</td>
<td>0</td>
</tr>
<tr>
<td>ZSize<em>ZNo</em>ZType</td>
<td>0.0433</td>
<td>15</td>
<td>0.00289</td>
<td>4.61</td>
<td>0</td>
</tr>
<tr>
<td>ZSize<em>ZNo</em>Sys</td>
<td>0.0612</td>
<td>45</td>
<td>0.00136</td>
<td>2.17</td>
<td>0</td>
</tr>
<tr>
<td>ZSize<em>ZType</em>Sys</td>
<td>0.022</td>
<td>15</td>
<td>0.00147</td>
<td>2.34</td>
<td>0.0025</td>
</tr>
<tr>
<td>ZNo<em>ZType</em>Sys</td>
<td>0.2164</td>
<td>9</td>
<td>0.02405</td>
<td>38.37</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZNo*ZType</td>
<td>0.0268</td>
<td>30</td>
<td>0.00089</td>
<td>1.43</td>
<td>0.0629</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZNo*Sys</td>
<td>0.0852</td>
<td>90</td>
<td>0.00095</td>
<td>1.51</td>
<td>0.0018</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZType*Sys</td>
<td>0.0519</td>
<td>30</td>
<td>0.00173</td>
<td>2.76</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>ZNo</em>ZType*Sys</td>
<td>0.217</td>
<td>18</td>
<td>0.01206</td>
<td>19.24</td>
<td>0</td>
</tr>
<tr>
<td>ZSize<em>ZNo</em>ZType*Sys</td>
<td>0.0556</td>
<td>45</td>
<td>0.00124</td>
<td>1.97</td>
<td>0.0001</td>
</tr>
<tr>
<td>InType<em>ZSize</em>ZNo<em>ZType</em>Sys</td>
<td>0.0629</td>
<td>90</td>
<td>0.00007</td>
<td>1.11</td>
<td>0.2224</td>
</tr>
<tr>
<td>Error</td>
<td>1.0831</td>
<td>1728</td>
<td>0.00063</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25.0501</td>
<td>2303</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A.1: ANOVA table for comparing different zero guardings. Transformation $\tilde{x} = x^{-0.87}$ was used.
Figure A.3: Plot for choosing a variance-stabilizing transform (without transformation).
Figure A.4: Residual analysis with transformation $\tilde{x} = x^{-0.87}$. The residuals do not look like normal distributed in the normal plot up to the left but there are so many residuals so that also the tails look thick. The tails are, however much thinner than the middle of the curve.
Figure A.5: Residuals versus factor levels with transformation $\tilde{x} = x^{-0.87}$. 
Figure A.6: Plot for choosing a variance-stabilizing transform. The transformation is chosen by fitting a straight line to the data points. See [Mon97] for more information.