Experimental Evaluation of some Thresholding Methods for Estimating Time-Delays in Open-Loop

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

14th July 2003

Report no.: LiTH-ISY-R-2525

Technical reports from the Control & Communication group in Linköping are available at http://www.control.isy.liu.se/publications.
Experimental Evaluation of some Thresholding Methods for Estimating Time-Delays in Open-Loop

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

The purpose with this work is to test and evaluate a certain class of methods for time-delay estimation, especially with automatic control applications in mind. Particularly interesting it is to determine the best method. Is one method best in all situations or should different methods be used for different situations? The tested class of methods consists essentially of thresholding the cross correlation between the output and input signals. This is a very common method for time-delay estimation. The methods are tested and evaluated experimentally with the aid of simulations and plots of RMS error, bias and confidence intervals.

The results are: The methods often miss to detect because the threshold is too high. The threshold has nevertheless been selected to give the best result. All methods over-estimate the time-delay. Nearly the whole RMS error is due to the bias. None of the tested methods is always best. Which method is best depends on the system and what is done when missing detections. Some form of averaging of the cross correlation, e.g. integration to step response or CUSUM, is advantageous. Fast systems are easiest. White noise input signal is easiest and steps is hardest. The RMS-errors are high in average (approximately greater than 6 sampling intervals). The error is lower for fast system or for high SNR.

time-delay, dead-time, estimation, system identification, linear dynamic systems, cross correlation, ANOVA, confidence intervals, simulations
# Contents

1  Introduction .......................................................... 1

2  Thresholding methods ............................................... 2

3  Simulation setup ....................................................... 3

4  Analysis .................................................................. 6

4.1  Simple analysis ....................................................... 6

4.2  Statistical analysis ................................................... 8

5  Discussion and conclusions .............................................. 16

5.1  Discussion ............................................................. 16

5.2  Conclusions ............................................................ 16

5.3  Recommendations .................................................... 16

5.4  Future work ........................................................... 17

References .................................................................. 1

A  Validation of confidence intervals .................................. 1
1 Introduction

The problem we address in this report is estimating time-delays in linear dynamical systems with additive noise. A synonym for time delay is dead-time. Estimating time-delays is a common engineering problem, e.g. in control performance monitoring of industrial processes [Hor00, Swa99], in design and tuning of controllers, in range estimation in radar [KQ92] and in direction estimation by time-delay of arrival in signal intelligence [HR97, Wik02]. Dead-time estimation is also a necessary part in all system identification [Lju99].

The purpose with this work is to test and evaluate a certain class of methods for time-delay estimation, especially with automatic control applications in mind. Particularly interesting is it to determine the best method. Is one method best in all situations or should different methods be used for different situations? The tested class of methods consists essentially of thresholding the cross correlation between the output and input signals. This is a very common method for time-delay estimation. The methods are tested and evaluated experimentally with the aid of simulations and plots of RMS error, bias and confidence intervals.

In the next chapter, the thresholding methods are briefly described. Then, Chapter 3 is about the simulation setup. After that, in Chapter 4 the analysis of the simulations is conducted. Following, Chapter 5 contains discussion, conclusions and suggestions for further work. After a literature reference list, Appendix A contains validation of required prerequisites for the analysis.
2 Thresholding methods

Thresholding methods is a subgroup of cross correlation methods. The steps of the methods are:

1. Estimate the impulse response and estimate the uncertainty of the impulse response estimate. The estimated impulse response is in principle the cross correlation between the output and input signals of the system [Bjö03].
2. Optionally, integrate to step response.
3. Thresholding.
   If the number zero is outside a certain confidence interval, then we consider the impulse (step) response to have started and this point of time is the time-delay estimate. The thresholding can be either
   - Direct thresholding [Bjö03] or.
   - Cumulative sum (CUSUM) thresholding [Bjö03]. For information about CUSUM see also [Gus00, GLM01].

There are some method parameters to choose. The most important are

- The relative threshold $h_{\text{std}}$ and relative drift $\nu_{\text{std}}$ [Bjö03].

The methods that we have used in this report are:

- idimp5. Direct thresholding of impulse response with prewhitening and $h_{\text{std}} = 5$. This would give a confidence interval with a confidence level of $0.999999713$ if the impulse response estimates are Gaussian distributed (which is a good assumption, see [Lju99]) and the estimate of the uncertainty in the confidence interval estimate is accurate.
- idimpCusum4. CUSUM thresholding of impulse response with prewhitening, $h_{\text{std}} = 3$, $\nu_{\text{std}} = 1$.
- idstepCusum4. CUSUM thresholding of step response with prewhitening, $h_{\text{std}} = 1$, $\nu_{\text{std}} = 6$.

The choices of $h_{\text{std}}$ and $\nu_{\text{std}}$ are the result of a statistical analysis in [Bjö03]. Note the opposite relation between $h_{\text{std}}$ and $\nu_{\text{std}}$ for idimpCusum4 and idstepCusum4. For more on the used methods in this report see [Bjö03].

In signal processing applications the system often consists of a pure time delay, maybe with an amplitude change. Then the right thing to do is to estimate the peak of the impulse response (cross correlation) instead.
3 Simulation setup

The setup for the simulations is the same as in [Bjö03] but with different time-delay estimation methods. Since the estimates are very non-Gaussian, so many as 4096 trials were simulated. From this 4 estimates of RMS error was calculated for the plots with confidence intervals. In this report we are studying properties in average.

There were many missed detections because the threshold was often not crossed. This complicates the analysis. It is not obvious what to do when missing. One possibility is to assign a value to the time-delay estimate. If this value is very incorrect, then we give a penalty because of the miss. Another possibility is to give an alarm. A third is to find the highest peak of the impulse response. This latter possibility will give a bias in the estimate. It, unfortunately, turns out that which method is the best depends on what we do when the detection is missed.

In this report, when a method missed to detect a uniform distributed random number in the range 20 to 30 was delivered as the time-delay estimate (as in Section 7.7-7.9 in [Bjö03]). The reason for this special choice of was to come closer to the required prerequisites for the confidence interval calculations. See appendix A and reference [Bjö03] for more on this subject.

Three environment factors were varied during the simulations: The system, the input signal type and the SNR [Bjö03]. The SNR was either 1 or 100. See [Bjö03] for the definition of the SNR. The impulse responses of the four used systems are depicted in Figure 3.1. Note that for all the systems the time delay will be 10 after the sampling. More information about the systems can be found in [Bjö03].

Figures 3.2-3.4 show the used input signals in the time and frequency domains. More information about the input signals can be found in [Bjö03].
Figure 3.1: Impulse response of system $G_1 - G_2$ and $G_5 - G_6$. True time-delay after sampling $T_d = 10$.

Figure 3.2: Time signal (left) and frequency spectrum (right) for a realization of the input signal type RBS 10-30%.
Figure 3.3: Time signal (left) and frequency spectrum (right) for a realization of the input signal type RBS 0-100%.

Figure 3.4: Time signal (left) and frequency spectrum (right) for a realization of the input signal type Steps.
4 Analysis

In this chapter we compare the tested methods and see in which cases they could be used. In Section 4.1 we start by a simple analysis that can give some feeling about the problems and how the methods work. Section 4.2 then has a statistical analysis, from which it is possible to draw conclusions.

4.1 Simple analysis

Figure 4.1 displays estimated impulse response for one of the systems but for different input signals and different SNRs. We see that in some cases the estimate is very inaccurate. This makes the time-delay estimation a hard job in these cases.

![Impulse response estimate](image1)

Figure 4.1: Impulse response estimate of the system $G_1$ by the function idimp4 for different input signal types and different SNRs [Bjo03]. The solid line is the true impulse response. The circles are the estimated impulse response and the triangles mark ±two estimated standard deviations. Note the different ranges of the vertical axes. (t130f1.m)

Direct thresholding of the impulse response estimate is illustrated in Figure 4.2 and CUSUM thresholding in Figure 4.3. Note that in these figures, the relative threshold $h_{std}$ and relative drift $\nu_{std}$ were not the same as in Chapter 2 and in the statistical analysis in Section 4.2.
Figure 4.2: Left: Estimated impulse response with uncertainty. Right: Estimated impulse response and threshold. Simulated input signal of type RBS 10-30%. SNR was 1. System $G_1$. The estimated time delay with idimp4 ($h_{\text{std}} = 3$) was $T_d = 11$. (t134d1.m)

Figure 4.3: Left: Impulse response with uncertainty. Right: Test statistics $g(t)$, threshold $h$, and drift $\nu$ for CUSUM on impulse response. Simulated input signal of type RBS 10-30%. SNR was 1. System $G_1$. The estimated time delay with idimpCusum3 ($h_{\text{std}} = 2$ and $\nu_{\text{std}} = 1$) was $T_d = 11$. (t146b1.m)
4.2 Statistical analysis

In this section we will perform a statistical analysis. Bear in mind that:

- All results are in average. Certain special cases can give a different result.
- Even if there is a statistically significant difference, it is not sure that the difference has any practical importance.

The statistical analysis is conducted in the same way as in Chapter 7 in [Bjö03]. The transformation was \((\text{RMS error})^{(0.73028)}\) (See [Bjö03]). This means “The lower the better”. Figure 4.4 shows confidence intervals for pair-wise comparisons of methods. We see that:

- Step response is significantly better (not overlapping confidence intervals) than impulse response in average.
- Step response: no significant difference (overlapping confidence intervals) between direct and CUSUM thresholding.
- Impulse response: significant difference between direct and CUSUM thresholding and CUSUM is better.

Figure 4.4: Confidence intervals (the lines in the circles) for pair-wise comparisons (95% simultaneous confidence level) for different thresholding methods. Positive transformation: \((\text{RMS error})^{(0.730028)} = \Rightarrow \text{“The lower the better”}\. (t149b1.m)

Figure 4.5 shows the average RMS error and bias for the different methods. We see that the RMS error \(\gtrapprox 6\) in average. It is high because:
- We are punishing missed detections.
- The case with SNR=1 is difficult difficult.
- The methods are overestimating the time delay. Nearly the whole RMS error is due to the bias.

![Figure 4.5: Average RMS error (left) and bias (right) of time-delay estimates for different thresholding methods.](image)

In Figure 4.6 it is obvious there is a large difference in performance between low and high SNR. Using step response estimates, the mean RMS error is 2.0 for high SNR but 10.4 for low SNR. Nearly the whole RMS error is due to the bias.

In Figure 4.7 we see how good the methods are for different input signal types. We notice:

- No method is always significantly best but step response is often better than impulse response.
- Wideband random signal easiest for all methods.
- Steps signal hardest for all methods.

Figure 4.8 shows the average RMS error and bias for different methods and input signal types. We see that:

- For the wideband random signal, RMS error ≈ 3.6 using step response.
- For steps as signal, RMS error ≈ 9.5 using step response.

In Figure 4.9 we see how good the methods are for different input signal types. We notice:
Figure 4.6: Average RMS error (left) and bias (right) of time-delay estimates for different thresholding methods and SNRs.

Figure 4.7: Confidence intervals (the lines in the circles) for pair-wise comparisons (95% simultaneous confidence level) for different thresholding methods and input signals. Positive transformation: (RMS error)\(^{(0.730028)}\) = "The lower the better". (t149b1.m)
Figure 4.8: Average RMS error (left) and bias (right) of time-delay estimates for different thresholding methods and input signals.

- Impulse response significantly best for the fast system.
- Step response significantly best for the other systems.
- The fast system is the easiest for all methods.

Figure 4.10 shows the average RMS error and bias for different methods and systems. We see that:

- Fast system: RMS error $\approx 3.5$ (step response), 2.9 (impulse response).
- High order system with complex poles: RMS error $\approx 7.8$ (step response), 10.5 (impulse response).

Figures 4.11 and 4.12 depict the RMS error and bias of the methods for all environment factors. We observe that:

- The bias is most often positive. It is often large. The few cases with negative bias has a very small bias. This indicate that we need a better estimation of the change time point in the impulse and step response.
Figure 4.9: Confidence intervals (the lines in the circles) for pair-wise comparisons (95% simultaneous confidence level) for different thresholding methods and systems. Positive transformation: (RMS error)^0.730028 => "The lower the better". (t149b1.m)
Figure 4.10: Average RMS error (left) and bias (right) of time-delay estimates for different thresholding methods and systems.
Figure 4.11: Average RMS error of time-delay estimates for different thresholding methods and different environment factors.
Figure 4.12: Average bias of time-delay estimates for different thresholding methods and different environment factors.
5 Discussion and conclusions

5.1 Discussion

There is a very high confidence level for the direct thresholding. This implies that we overestimate the time delay. All methods over-estimate the time-delay. The used thresholds have nevertheless been selected to give the best result [Bjo03]. A better estimation of the change time than simple thresholding is needed. In [KG81] a more sophisticated estimation method of the change time is used. This method has not been tested in this report.

5.2 Conclusions

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

- The methods often miss to detect. This is a problem in the analysis.
- No method is always best. Which is best depends on the system and what is done when missing detections.
- In the tested cases (the used penalty when missing), using step response was best in average. It was also best in most combinations of factor environments.
- In the tested cases, using impulse response was best in average for fast systems.
- No significant difference between direct and CUSUM thresholding for step response.
- CUSUM significantly better than direct thresholding for impulse response except for fast systems.
- Some form of averaging, e.g. integration to step response or CUSUM, of the impulse response enhances the time-delay estimate (except for fast systems). Integrating to step response and CUSUM are two ways to smooth the impulse response estimate. Using CUSUM on the step response estimate does not further improve the estimates in most cases.
- The wideband random input signal is easiest. Input signal with steps is hardest.
- Fast system is easiest.
- All methods over-estimate the time-delay. A better estimation of the change time is needed.
- Nearly the whole RMS error is due to the bias.
- The RMS-errors are high in average (≥ 6 sampling intervals). This is too high to useful. The error is lower for fast system or for high SNR.

5.3 Recommendations

- Use thresholding methods only for fast systems or when the SNR is high.
- For step input the SNR must be high.
- Integration to step response and/or CUSUM thresholding enhances the performance for the impulse response estimate, except for fast systems.
5.4 Future work

Some possible future work is:

- Estimate better the change point in the impulse and step responses to avoid over-estimating the time-delay.
- Are there really no better thresholds and drifts?
- Test logarithmic cross spectrum scale (cepstrum), see [Bjö03].
- Test the thresholding methods on random systems and make a statistical analysis.
- Test other methods in the same way as in this report.
- Compare groups of methods.
References


A Validation of confidence intervals

This appendix comments on the use and applicability of the confidence intervals in Section 4.2.

The 4 RMS estimates (Chapter 3) have first gone through a transformation to make the variance more constant [Bjö03]. The transformation became (RMS error)$^{(0.73028)}$ (Figure A.1). Then ANOVA [Mon97] was executed and confidence intervals for pairwise comparisons [Mat01] plotted. The traditional ANOVA table is given in Table A.1. Since all $p$-values (the column Prob>F) are very small, all factors and interactions have effect with a very high confidence level. The confidence intervals are plotted in Section 4.2. For the ANOVA and confidence intervals to be valid some prerequisites must be fulfilled [Mon97, Bjö03]. These are usually tested by studying some validation graphs [Mon97, Bjö03] (Figures A.2-A.3). We see in these graphs that the prerequisites are not completely fulfilled so we must be somewhat careful in the interpretation of the ANOVA and confidence interval. See [Mon97, Bjö03] for more information on this.

<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>Method</td>
<td>50.4689</td>
<td>3</td>
<td>16.823</td>
<td>1910.2693</td>
<td>0</td>
</tr>
<tr>
<td>InType</td>
<td>564.3458</td>
<td>2</td>
<td>282.1729</td>
<td>32041.1107</td>
<td>0</td>
</tr>
<tr>
<td>SNR</td>
<td>1421.6913</td>
<td>1</td>
<td>1421.6913</td>
<td>161434.944</td>
<td>0</td>
</tr>
<tr>
<td>Sys</td>
<td>527.0499</td>
<td>3</td>
<td>175.6833</td>
<td>19949.0399</td>
<td>0</td>
</tr>
<tr>
<td>Method*InType</td>
<td>21.1096</td>
<td>6</td>
<td>3.5183</td>
<td>399.5042</td>
<td>0</td>
</tr>
<tr>
<td>Method*SNR</td>
<td>5.9101</td>
<td>3</td>
<td>1.97</td>
<td>223.6997</td>
<td>0</td>
</tr>
<tr>
<td>Method*Sys</td>
<td>48.857</td>
<td>9</td>
<td>5.4286</td>
<td>616.42</td>
<td>0</td>
</tr>
<tr>
<td>InType*SNR</td>
<td>21.597</td>
<td>2</td>
<td>10.7985</td>
<td>1226.1859</td>
<td>0</td>
</tr>
<tr>
<td>InType*Sys</td>
<td>13.6246</td>
<td>6</td>
<td>2.2708</td>
<td>257.8478</td>
<td>0</td>
</tr>
<tr>
<td>SNR*Sys</td>
<td>45.817</td>
<td>3</td>
<td>15.2723</td>
<td>1734.1939</td>
<td>0</td>
</tr>
<tr>
<td>Method<em>InType</em>SNR</td>
<td>127.9938</td>
<td>6</td>
<td>21.3323</td>
<td>2422.3104</td>
<td>0</td>
</tr>
<tr>
<td>Method<em>InType</em>Sys</td>
<td>11.3337</td>
<td>18</td>
<td>0.62965</td>
<td>71.4977</td>
<td>0</td>
</tr>
<tr>
<td>Method<em>SNR</em>Sys</td>
<td>4.8604</td>
<td>9</td>
<td>0.54004</td>
<td>61.3222</td>
<td>0</td>
</tr>
<tr>
<td>InType<em>SNR</em>Sys</td>
<td>327.4812</td>
<td>6</td>
<td>54.5802</td>
<td>6197.6538</td>
<td>0</td>
</tr>
<tr>
<td>Method<em>InType</em>SNR*Sys</td>
<td>44.8829</td>
<td>18</td>
<td>2.4935</td>
<td>283.1397</td>
<td>0</td>
</tr>
<tr>
<td>Error</td>
<td>2.5363</td>
<td>288</td>
<td>0.0088066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3239.5586</td>
<td>383</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A.1: Analysis of Variance table for all methods. Constrained (Type III) sums of squares. Positive transformation: (RMS error)$^{(0.73028)}$. 
Before transformation: Cell std vs. mean

Figure A.1: Plot (without transform) for choosing a variance-stabilizing transform [Mon97] for ANOVA. The transformation is chosen by fitting a straight line to the data points by the least squares method. The outlier in the lower left corner is ignored when calculating the transformation. It is due to zero variance (always the same time-delay estimate) for one factor level combination.
Figure A.2: Residual analysis for ANOVA and confidence intervals. Ideally the residuals should be Gaussian and the residuals vs. time and fitted value should be within a horizontal band and be structureless [Mon97, Bjo03].
Figure A.3: Standard deviation of residuals versus factor levels for ANOVA and confidence intervals. Ideally the standard deviation for all factor levels should be equal [Mon97, Bjo03].