Cross-correlation tracking for Maximum Length Sequence based acoustic localisation

Navinda Kottege
Research School of Information Sciences and Engineering
The Australian National University, ACT 0200, Australia
navinda@ieee.org

Uwe R. Zimmer
Research School of Information Sciences and Engineering
The Australian National University, ACT 0200, Australia
uwe.zimmer@ieee.org

Abstract

While Maximum Length Sequence based cross-correlation and localisation methods are specifically robust with respect to multiple forms of disturbances, the method is still subject of becoming 'side-tracked' by echoes, reverberations and other resonance effects. This article details a method to track the correct reading throughout a series of measurements which is specifically designed for the cross-correlation localisation technique as deployed in underwater environments by the Serafina project. The method is based on the selection of the most likely candidate rather than Kalman filters or other low-pass filters and extrapolation methods. Therefore every individual reading is an actual measurement, rather than an extrapolation.

As transducers with non-linear frequency responses are common in energy-efficient underwater acoustic setups, the resulting signal deformations are explicitly addressed and corrected by means of inverse frequency transformations.

The article discusses in detail a series of experiments performed in a highly reverberant underwater environment. The achieved performances include standard deviations of a less then 3 degrees in azimuth estimation, and 11 cm in range estimations throughout the presented sequence of experiments.

1. Introduction

In attempting to design and implement localisation systems for autonomous vehicles, the transducer characteristics as well as the operating environment contributes to the performance of the system. This is especially true in underwater environments where localisation is performed using acoustic signals.

The underwater acoustic localisation system proposed in [Kottege and Zimmer, 2007; Kottege and Zimmer, 2008a] and later developed in [Kottege, 2008; Kottege and Zimmer, 2008b] uses low cost transducers as projectors and hydrophones while using acoustically transmitted Maximum Length Sequence (MLS) signals [Peterson et al., 1972]. The accuracy and precision of the estimates produced by this system is affected by the non-linear frequency response of the low cost transducers as well as the cluttered underwater environments which produce multiple reflections of the acoustic signal.

The two-fold outlier rejection scheme described in the following sections manage to compensate the effects of the transducer frequency response while recovering estimates which otherwise would be lost due to reflected signals.

2. Cross-correlation of MLS signals

The main tool used in [Kottege, 2008; Kottege and Zimmer, 2008b] for estimating the azimuth, range and heading of a signal source is time-domain cross-correlation of MLS signals. The theory associated with the statistical properties of MLS signals [Borish and Angell, 1983; Borish, 1985a; Borish, 1985b; Bradley, 1996] recommends the use of a longer sequence to improve the definition of the cross-correlogram peak. However, using signal chirps with a longer duration attracts undesirable reverberation effects in cluttered and enclosed acoustical environments. As a compromise, a 127 point MLS signal with a duration of 1.3 ms when sampled at 96 kHz is used.

The peak position of the cross-correlogram remains unaffected by the introduction of uncorrelated noise on to the source signal channels while the peak height is reduced as a result. Figure 1 shows the cross-correlogram produced by cross-correlating two MLS signals with a relative shift of 8 samples which were mixed with white Gaussian noise with a signal to noise ratio of 0 dB.

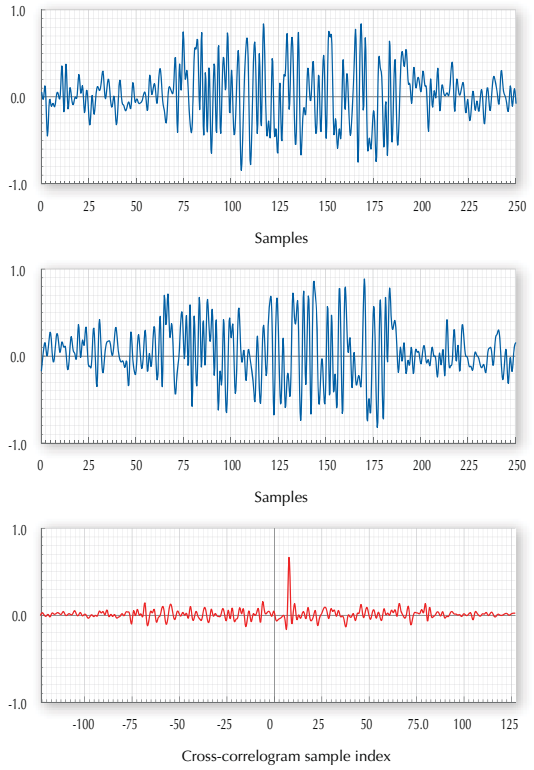

Figure 1: Plot shown in c) results from the cross-correlation of the shifted noisy MLS signals shown in a) and b). In each plot, the y-axis represents normalised amplitude.

3. Effect of transducer frequency response

As it was shown by Figure 1.c, contamination by additive white noise does not contribute to a noticeable deterioration of the cross-correlation performance of MLS signals, apart from a slightly lower height for the peak. In theory, addition of two *"flat"* frequency spectra should again result in a *"flat"* frequency spectrum, hence the spectral properties of the MLS signal which provides the narrowness of the peak are preserved. However, this is not necessarily the case when these signals are transmitted and received via transducers with a non-linear frequency response within the bandwidth of the signal. Due to the sampling rate used, the signals are band limited by 48 kHz. The frequency response of the Benthos AQ-2000 transducers which are used as transmitting projectors as well as receiving hydrophones by the localisation system is shown in Figure 2. As seen from this logarithmic plot, the transducers have a resonance near 20 kHz and an anti-resonance near 25 kHz which results in a highly non-linear response within the signal bandwidth.

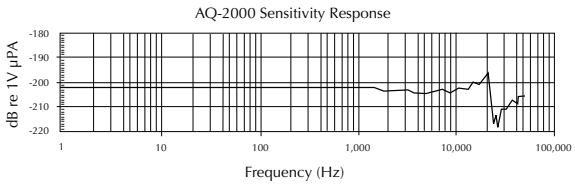

Figure 2: The frequency response curve of the hydrophone reproduced from the AQ-2000 data sheet []

3.1. Inverse frequency response filtering

To test the effect of this potential frequency filtering effect introduced by the transducers, the frequency response shown in Figure 2 was empirically modelled[1] and implemented as an FFT filter. The filter was then applied to white noise

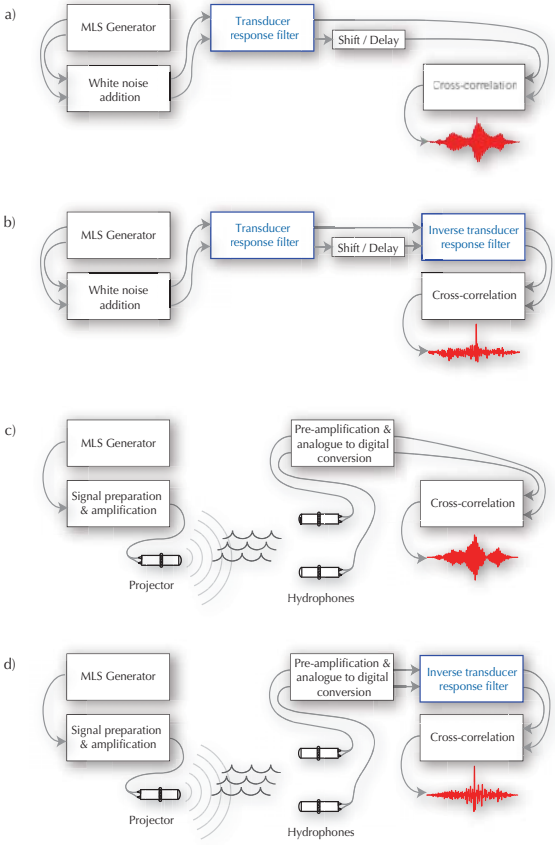

Figure 3: The four different setups producing the four different cross-correlograms shown in Figure 4.

1. The shape of the frequency response curve was replicated as the shape of a frequency response curve of an FFT filter.

contaminated MLS signals earlier shown in plots a) and b) of Figure 1 and the two filtered signals were cross-correlated. This setup is illustrated in Figure 3.a. The effect it has on the cross-correlogram shown in Figure 4.a can be compared to the cross-correlogram shown in Figure 1.c, which was produced by the same source signals albeit the frequency filtering. Resonance of the transducer near 20kHz appears as the dominant frequency in the resulting cross-correlogram.

For comparison, Figure 4.c shows the cross-correlogram of two signal channels received via two AQ-2000 transducers

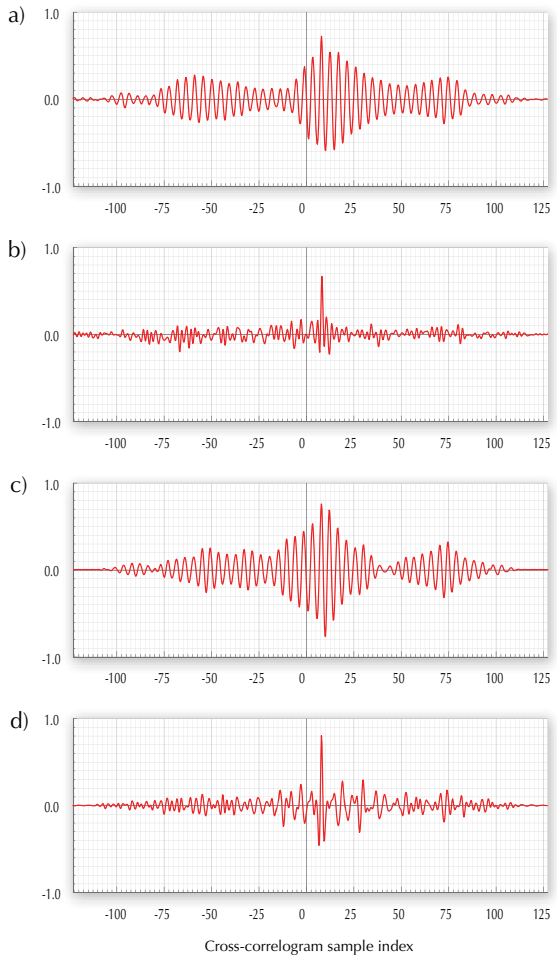

Figure 4: Plots in a) and b) represents the cross-correlograms of the shifted MLS signals contaminated with additive white noise, first filtered with the transducer frequency response, then filtered with the inverse of that filter. Plots c) shows the cross-correlogram resulting from two actual signal channels (with a shift of +8 samples) which were transmitted and received using the transducers. Plot d) shows the resulting cross-correlogram when the inverse transducer filter was applied to the signal channels prior to cross-correlation. In each plot, the *y-axis* represents the normalised amplitude of the signals.

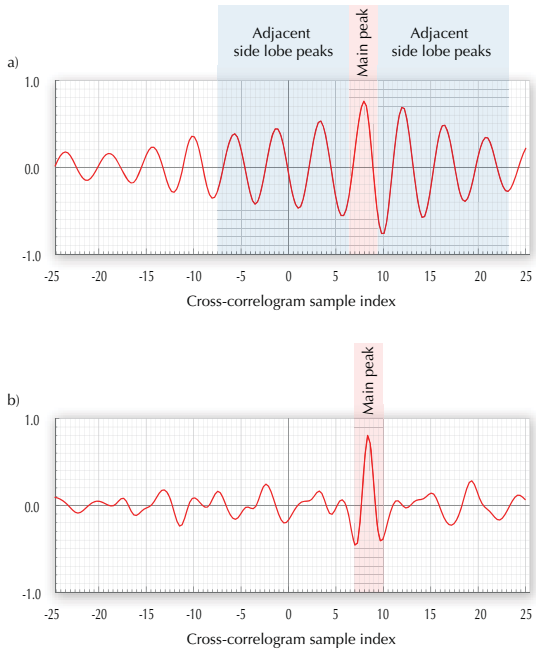

Figure 5: Cross-correlogram plots resulting from a) unfiltered signal channels and b) filtered signal channels, showing the main peak and adjacent side lobe peaks caused by the resonance frequency of the transducers. The *y-axes* on these plots represents the normalised amplitude.

(separated by 0.3m) when an MLS (length-127, duration 1.3ms) signal was transmitted via another AQ-2000 transducer.The signal travelled a distance of 2.0m underwater and the cross-correlation reveals a delay of +8 samples (83.33µs) between the channels. This setup is schematically illustrated in Figure 3.c.

Both the cross-correlograms, one from cross-correlating the 'real' and the other from cross-correlating the 'simulated' signals, shows the dominance of the 20kHz resonance of the transducer throughout the plots. Even though, the position and height of the peak is not affected, the uniqueness of the peak had been lost by being surrounded by an envelope of decaying side-lobes (Figure 5). This decreases the accuracy of the localisation system when used in enclosed and cluttered environments due to peaks caused by reflected signals.

In order to address this issue, another filter was empirically modelled which had the inverse frequency response of the one modelled earlier to represent the response of the transducer. The results of applying the new FFT filter[1] to the simulated and the real signals used earlier and cross-correlating the channels are shown in plots b) and d) of Figure 4 respec-

1. The filter was applied to the two signal channels prior to being cross-correlated.

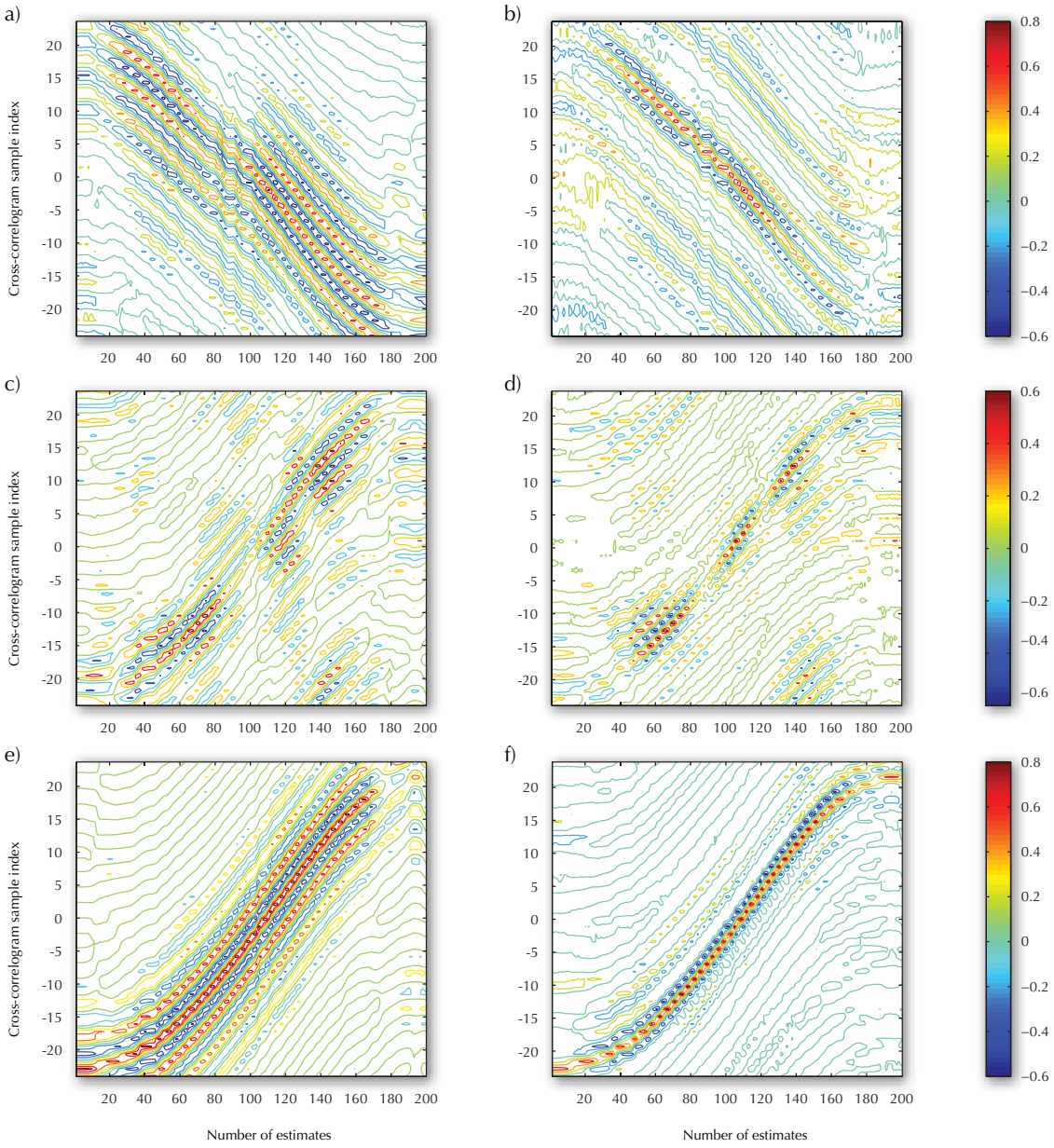

Figure 6: Contour plots shows series of cross-correlograms resulting from three different experiments shown on three rows.

tively. The setups used in these instances are schematically depicted in Figures 3.b and 3.d. As depicted by the plots the uniqueness and narrowness of the peak are restored making it possible to unambiguously locate it.

It must be emphasised that the filter process only accounts for the transducer characteristics and not those of the propagation medium. Since channel characteristics of the underwater medium greatly varies with depth, temperature and salinity as well as environmental features such as the composition and texture of the bottom (sediment/sand/vegetation), modelling the transducer characteristics are more practical. In the context of an autonomous underwater vehicle, it is far more convenient to account for the characteristics of onboard sensors than to have access to a model of the channel characteristics of the operating medium. However, the underwater channel does indeed have an effect on the trans-

mitted MLS signal as seen in Figure 4.d where the filtering process does not completely reconstruct the original cross-correlogram (Figure 1.c). Furthermore, the power of the received signals are greatly reduced by the inverse filtration process since most of the transmitted signal power is around the resonance frequency of the transducer. This attenuation affects the distance which the MLS signals can be effectively transmitted at a given transmission power.

3.2. Effectiveness of filtering

Figure 6 shows contour plots of normalised cross-correlograms resulting from cross-correlating the two experimentally recorded hydrophone channels corresponding to the MLS chirps emitted by a projector. The receiving hydrophone pair was rotated about the mid point of the line connecting them during the experiment[1]. Depending on the direction of rotation, the time difference of arrival (TDOA) between the two channels either increase or decrease during the experiment resulting in either an increasing or decreasing acoustic path-length difference. The three rows shows three different experimental runs while the first column is without any filtering and the second column is with the inverse frequency response filter applied.

In each of the contour plots, each vertical 'slice' is a normalised cross-correlogram such as the ones shown in Figure 4, with the peaks moving diagonally from top to bottom (first row) or vice versa (second and third row) as the experiment proceeds with increasing estimates. For clarity, closely cropped versions of the cross-correlogram plots c) and d) in Figure 4 is presented in Figure 5.

The three different experiments whose results are depicted in Figure 6 were selected to show varying levels of effectiveness and necessity of the filter being applied. The instance shown in plot a) has at least three adjacent peaks competing for prominence characterised by red parallel 'ridges' continuing diagonally across the plot. This is due to the dominant frequency component introduced by the resonance of the transducers. This behaviour of the cross-correlogram peaks causes the precision of the relative localisation system estimates to decrease (the standard deviation of errors increase). As a result of the filter, the side lobe peaks in the cross-correlograms subsides leaving only the main ridge of peaks as shown in plot b).

Plot c) depicts an experiment where the peaks of the cross-correlograms were affected by actual reflected signals (off the test tank walls in which the experiment was conducted in) apart from the multiple adjacent peaks due to resonance as in the previous case. In the first instance of reflection between estimates 80 to 100, the peaks are lost while in the second instance between estimates 140 to 160, a series of outlier peaks appear further away from the continuing ridge of peaks. As a result of the filter, the adjacent peaks subside leaving the main ridge as before but the outlier peaks still remain as seen from plot d).

In plot e), though there are multiple ridges flanking the main ridge caused by adjacent side lobe peaks in the cross-correlograms, the main ridge is continuously higher than its flanks. As expected, the filtered signals causes the side lobe peaks to subside leaving only the prominent main ridge shown in plot f).

The experiment represented in the first row clearly benefits from the filtering scheme as it helps to subdue unwanted side lobe peaks. In the experiment shown in the third row, due to the height of the main ridge compared to its flanks, filtering does not necessarily introduce any improvement to the estimates even though the side lobe peaks are suppressed in the process. In the experiment depicted in the second row, while filtering contributes to reducing the standard deviation of estimation errors, it does not improve performance in areas affected by reflections (i.e. peak drop-offs and outliers). Such situations necessitate an additional scheme for outlier rejection.

4. Outlier rejection via peak tracking

The localisation system suggested in [Kottege and Zimmer, 2007] and [Kottege and Zimmer, 2008a] uses a simple maximum search routine coupled with sub-sample spline interpolation for finding the peak of a cross-correlogram and the corresponding position of the peak. The following **peak tracking** scheme builds on this methodology and contributes to effectively reject outliers arising due to interference.

The main feature of this scheme is that it enforces continuity assumptions of the estimated quantities by considering the physically possible variation of each raw estimate (in sample space) within one estimation step.

Here $R_{s_1 s_2}$ refers to the full range cross-correlation of two length N signal channels $s_1(n)$ and $s_2(n)$ which includes both positive and negative discrete lags in sample space spanning $\{-N \ldots N\}$.

If $x(n)$ denotes an element of the resulting cross-correlation with a lag of n samples and can be expressed as:

$$x(n) \in \{x_{-N} \ldots x_N\} = R_{s_1 s_2} \tag{1}$$

$X(n)$ is defined as a set containing all ordered pairs of lags and corresponding value of the cross-correlogram as follows:

$$X(n) = \{(n, x(n)), \forall n \in \{-N \ldots N\}\} \tag{2}$$

Another set $X_{\text{Sorted}}(m)$ is formed by sorting $X(n)$ in descending order by the value of $x(n)$ as:

$$X_{\text{Sorted}}(m) = \{(m, y(m)), \forall m \in \{-N \ldots N\}\} \tag{3}$$

Therefore, the following conditions are satisfied by the elements of $X_{\text{Sorted}}(m)$ and $X(n)$:

[1] Experiments conducted in a cylindrical test tank (Diameter 4.2 m, depth 1.5 m) with corrugated metal walls and filled with freshwater. See [Kottege and Zimmer, 2008b] and [Kottege, 2008] for details of experimental setup.

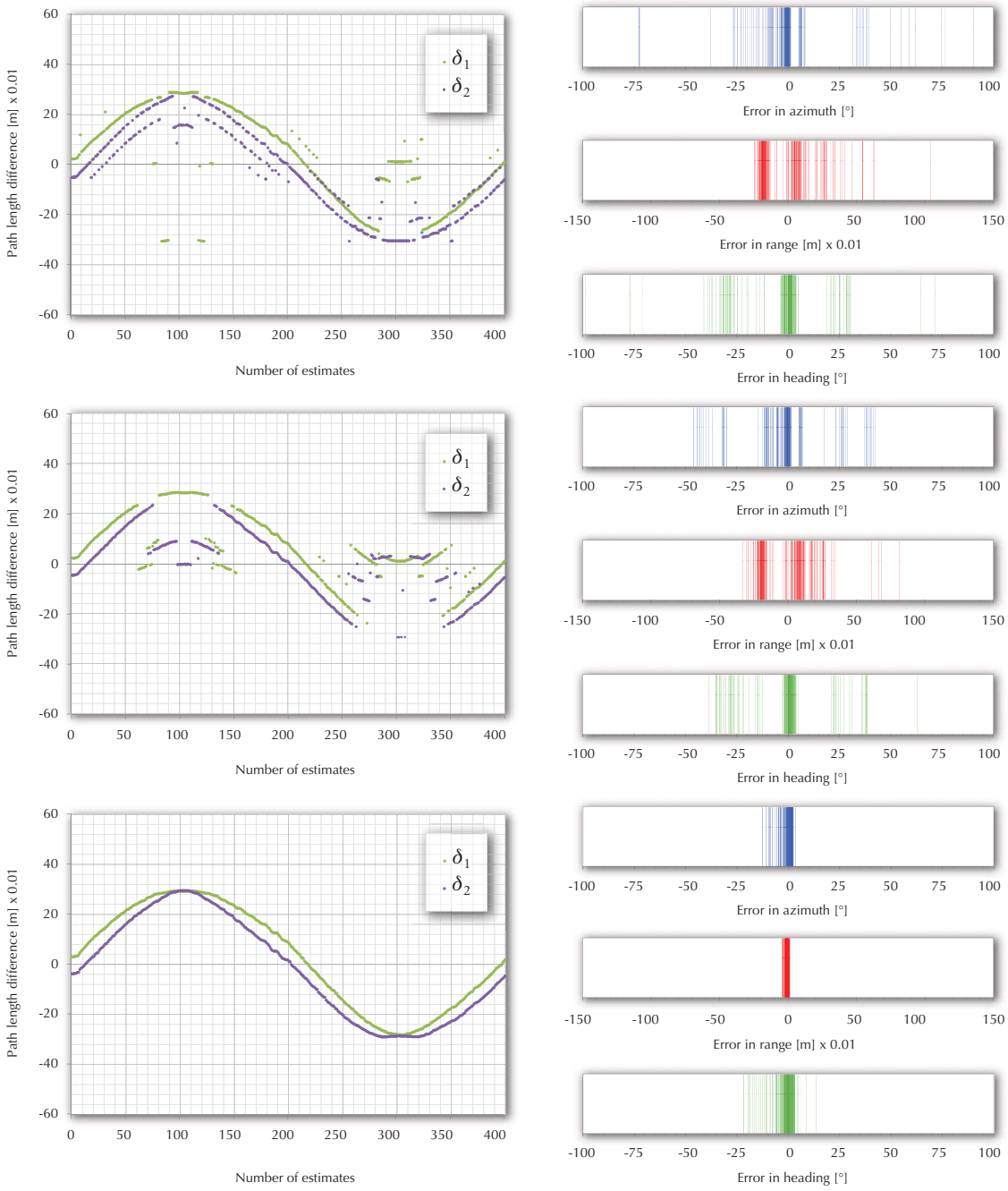

Figure 7: The intermediate values produced by the relative localisation system which consists of the acoustic path length differences are plotted along with the errors of the final estimates which are calculated using these values. The first row is with no filtering, the second row is with inverse frequency response filtering applied and the third row is with filtering and peak tracking on the cross-correlograms enabled.

$$\forall m \in \{-N \ldots N\}, \exists n \in \{-N \ldots N\} \text{ s.t. } x(n) = y(m) \tag{4}$$

$$\forall n \in \{-N \ldots N\}, \exists m \in \{-N \ldots N\} \text{ s.t. } y(m) = x(n) \tag{5}$$

$$\forall m \in \{-N \ldots (N-1)\}, y(m) \geq y(m+1) \tag{6}$$

	No filtering and no peak tracking	Filtered with no peak tracking	Filtered with peak tracking
Std. dev. of azimuth error ($\sigma_{\Delta\theta}$)	20.26°	17.05°	2.77°
Mean of azimuth error ($\mu_{\Delta\theta}$)	0.55°	–1.12°	0.00°
Avg. dev. of azimuth error ($\overline{\Delta\theta}$)	10.26°	10.38°	1.47°
Std. dev. of range error ($\sigma_{\Delta r}$)	214.7×10^{-3}m	197.0×10^{-3}m	11.0×10^{-3}m
Mean of range error ($\mu_{\Delta r}$)	356.7×10^{-3}m	366.1×10^{-3}m	-1.50×10^{-3}m
Avg. dev. of azimuth error ($\overline{\Delta r}$)	172.4×10^{-3}m	163.9×10^{-3}m	8.20×10^{-3}m
Std. dev. of heading error ($\sigma_{\Delta\alpha}$)	20.36°	16.30°	5.24°
Mean of heading error ($\mu_{\Delta\alpha}$)	–3.17°	–0.27°	–0.63°
Avg. dev. of heading error ($\overline{\Delta\alpha}$)	13.02°	9.69°	3.04°

Table 1 : Comparison of standard deviations, means and average deviations of errors associated with azimuth, range and heading estimates corresponding to the intermediate values plotted in Figure 7.

where $n, m \in \{-N \ldots N\} \subseteq \mathbb{Z},\ x(n), y(m) \in [-1, 1] \subseteq \mathbb{R}$ and (4), (5) maintains bijectivity between $X(n)$ and $X_{\text{Sorted}}(m)$.

A set $M(k)$ corresponding to estimation step[1] k is defined as follows:

$$M(k) = \{m \in \{-N \ldots N\} \text{ s.t. } |m - \tau_0^I(k-1)| < \Delta_{\text{Tolerance}}\} \quad (7)$$

where m is drawn from the ordered pairs in the set $X_{\text{Sorted}}(m)$, $\Delta_{\text{Tolerance}}$ being a tolerance value based on the continuity assumptions of the quantity being estimated and $\tau_0^I(k-1)$ being the sub-sample interpolated lag at estimation step $k-1$. The lag of the new 'tracked peak' of the cross-correlogram which maintains continuity with the previous estimates denoted by $\tau_0(k)$ is the minimum element of $M(k)$ given as:

$$\tau_0(k) = min\{M(k)\} \quad (8)$$

This procedure essentially performs a local maxima search of the cross-correlogram within the neighbourhood of lags around the previously estimated lag. The size of the neighbourhood is decided by the value selected for $\Delta_{\text{Tolerance}}$. The motivation behind the implementation described above which involves sorting the elements of the cross-correlogram, was to provide a facility to dynamically limit the search domain for lags by restricting maximum m to some $N_{\text{restricted}} \leq N$ in (7) (which in turn violates the bijectivity condition between $X(n)$ and $X_{\text{Sorted}}(m)$). This would specify a lower bound to the peak magnitude in the cross-correlogram which can correspond to the lag returned by (8). However, this approach was not implemented for the version of the localisation system experimentally evaluated in this text.

Once the discrete lag τ_0 is selected by (8), the sub-sample interpolated lag is returned by the cubic spline interpolation function I_{Spline} which takes in n_{Int} which is the number of sub-sample interpolation steps as a parameter to produce the 'real' lag at estimation step k:

$$\tau_0^I(k) = I_{\text{Spline}}(\tau_0(k)) \quad (9)$$

where $\tau_0(k) \in \{-N \ldots N\} \subseteq \mathbb{Z},\ \tau_0^I(k) \in [-N, N] \subseteq \mathbb{R}$.

There are multiple cross-correlations per estimation step which results in multiple sets for $M(k)$ used for estimating the azimuth, range and heading by the localisation system. For each of these quantities, there is a corresponding tolerance value $\Delta_{\text{Tolerance}}$ calculated using the maximum possible variation (constrained by the physical capabilities of the experimental setup or autonomous vehicle used) of the lags in the sample domain corresponding to variation of angles and distances within the duration between two estimation steps. The underlying assumption is that these variations maintain continuity between estimation steps. As a 'boot-strapping' technique, initially peak tracking is disabled and simple maximum searching is used to find the lags corresponding to the maximum magnitude peak in the cross-correlogram as no prior estimate is available at initialization. Once the position of the peak stabilises (*e.g.* detected by a result sequence which can be explained by the maximal relative speeds of the vehicles, *i.e.* no discontinuities over a certain number of estimation steps), peak tracking is enabled.

[1] This also corresponds to the k^{th} logical time-step as well as the k^{th} sending event.

5. Analysis

As mentioned earlier, unlike the simple search used to locate the peak which only considered the amplitude of the cross-correlogram, with peak tracking, the history of the previous peak positions is incorporated in to the search parameters and a higher prominence is given to the position rather than the amplitude of the peak. Peak positions estimated in this manner are refined as before using sub-sample interpolation.

The combination of inverse frequency response filtering and peak tracking significantly improves the accuracy and precision of the estimates produced by the relative localisation system. The effect of inverse frequency response filtering and peak tracking on the intermediate values produced (path length differences) by the localisation system and the corresponding errors associated with the final estimates are depicted in the plots shown in Figure 7. The first column of Figure 7 contains plots of path length differences corresponding to TDOAs between the two hydrophone channels during an experiment with two source signals transmitted in sequence. In the experiment depicted in Figure 7, the receiving hydrophone pair was rotated such that the azimuth of the source varied as $\theta : 0^\circ \rightarrow -90^\circ \rightarrow 0^\circ \rightarrow 90^\circ \rightarrow 0^\circ$.

Plot a.i) was obtained using unfiltered hydrophone channels as the inputs to the cross-correlation with no peak tracking on the cross-correlograms, plot b.i) was with the inverse frequency response filter applied but without any peak tracking while plot c.i) was with the filter applied and with peak tracking enabled. Plot a.i) shows the effects of the adjacent side lobe peaks which causes the outliers on either side of the main path length differences.

When the inverse frequency response filter is applied, a deterioration of the values around 100 and 300 estimates is clearly noticeable. These areas correspond to azimuths of -90° and 90° where the hydrophones on the observer are almost pointed in a direction perpendicular to projectors emitting the source signal. Due to the directivity pattern of the AQ-2000 hydrophones which are non-omni-directional, the direct-path signals that are received in these regions would carry most of their energy in frequencies within a narrow bandwidth centred at the resonance frequency of the transducers. The filtering process which aims to *"smoothen out"* the received frequency spectrum by attenuating spikes caused by resonance contributes to further reduce the energy of the direct-path signals. Under these circumstances the signals reflected off the curved metal walls of the test tank, which impinge on the hydrophones from the front, tend to carry more energy than the direct-path signals. Hence the cross-correlogram peaks due to reflections tends to have a higher amplitude than those due to direct-path signals in these regions causing the performance of the estimation to deteriorate when the filter is applied. Due to causality, peaks caused by direct-path signals occur *before* the peaks caused by reflected signals albeit with much lower amplitude. The peak tracking algorithm exploits this feature as explained in the previous section and manages to retrieve the peaks caused by direct path signals. The results can be seen by comparing plot b.i) and c.i) where applying filtering and peak tracking (with $\Delta_{\text{Tolerance}} = 1.25$) completely eliminates outliers in the path length difference measurements which are subsequently used for the angular estimations.

The second column in Figure 7 shows the errors in the final components of the pose vector, azimuth, range and heading calculated using the intermediate values plotted in the first column. As observed from these plots, the standard deviation of the errors are greatly reduced when peak tracking is enabled while filtering does not lead to significant improvements. The standard deviations and means of these errors are compared in Table 1.

6. Conclusions

As seen by the error plots given in Figure 7, the peak tracking scheme and inverse frequency response filtering scheme operating at an early stage in the estimation process is extremely effective in eliminating outliers and heavily contributes towards minimising estimation errors.

As a future work, the peak-tracking algorithm will be improved to accommodate dynamically changing parameters such as the search domain size which introduces a lower bound to the peak magnitude. Experiments need to be conducted to further analyse the effect of the parameter $\Delta_{\text{Tolerance}}$ on the estimates produced by the system.

References

[Borish and Angell, 1983] Borish, J. & Angell, J. B., An efficient algorithm for measuring the impulse response using pseudorandom noise; *Journal of the Audio Engineering Society*, **31**(7), August 1983, pp. 478-488.

[Borish, 1985a] Borish, J., An efficient algorithm for generating colored noise using a pseudorandom sequence; *Journal of the Audio Engineering Society*, **33**(3), March 1985, pp. 141-144.

[Borish, 1985b] Borish, J., Self-contained crosscorrelation program for maximum -length sequences; *Journal of the Audio Engineering Society*, **33**(11), November 1985, pp. 888-891.

[Bradley, 1996] Bradley, J. S., Optimizing the decay range in room acoustics measurements using maximum-length-sequence techniques; Journal of the Audio Engineering Society, **44**(4), 1996, pp. 266-273.

[Kottege and Zimmer, 2007] Kottege, N. & Zimmer, U. R.; Relative localisation for AUV swarms, i*n proceedings of the IEEE International Symposium on Underwater Technology 2007 (SUT '07);* Tokyo, Japan, April 2007.

[Kottege, 2008] Kottege, N.; *Underwater Acoustic Localisation in the context of Autonomous Submersibles;* submitted at The Australian National University, Canberra, ACT, Australia, April 2008.

[Kottege and Zimmer, 2008a] Kottege, N. & Zimmer, U. R.; MLS Based Distributed, Bearing, Range and Posture Estimation for Schools of Submersibles, *Experimental Robotics*; Springer: Berlin; Heidelberg, March 2008, pp. 377-385.

[Kottege and Zimmer, 2008b] Kottege, N & Zimmer, U. R.; Acoustical Methods for Azimuth, Range and Heading Estimation in Underwater Swarms, *In proceedings of Acoustics '08*, Paris, France, July 2008

[Peterson et al., 1972] Peterson, W. W. & Weldon, E. J. Jr., *Error-correcting codes*, MIT Press, Cambridge, Massachusetts, 1972.

[Serafina, 2008] http://serafina.anu.edu.au