Estimation of Subsample Time Delay Differences in Narrowband Ultrasonic Echoes Using the Hilbert Transform Correlation

Anders Grennberg, *Member, IEEE* and Magnus Sandell, *Student Member, IEEE*

Abstract— In many areas the time delay of arrival (TDOA) is desired. In the case of narrowband signals we propose a fast and simple method to estimate small time delays. This method is shown to have the same or better accuracy as the cross correlation methods for small delays in the order of fractions of the sample interval. It is based on using the Hilbert transform in correlation between two signals and consists of only one scalar product, which makes it fast. It may also be used in applications with narrowband signals where the measurements are repeatable, such as ultrasonic imaging and nondestructive testing. In ultrasonic applications, due to fluctuations in the insonified media, a small random time shift may be present causing the signals to be misaligned in time. Averaging signals under these conditions will result in a distortion of the signal shape. We propose an averaging method to avoid this and to accomplish a higher SNR without the distortion. Simulations and experiments from ultrasonic applications are presented.

I. INTRODUCTION

THE TIME delay of arrival (TDOA) is required in many areas, e.g., ultrasonic imaging [18]. Estimators of the TDOA are designed in different ways as can be seen in [3], [5], [19] and various articles in [6], [7], but they have a few things in common. Usually they have an SNR threshold under which they perform substantially less well [20]. For correlator based estimators this is due to peak ambiguity, i.e., the possibility of selecting the wrong peak of the correlation. These estimators also need some kind of interpolation to obtain subsample resolution, e.g., a parabolic fit [4], [11], [14]. In [12], a comparison of several methods is made. A special case of time delay estimation is when the signals are narrowband [8]. Under this assumption we propose a novel method to estimate the TDOA for small time delays. This method will be shown to be fast and simple. Apart from the increased resolution in TDOA, it can be used for other purposes. One such application is the problem of averaging. In areas where the measurements are repeatable, such as ultrasonic imaging and nondestructive evaluation (NDE), the process of averaging can be used to improve the SNR. If the useful part of the signal is the same for all measurements, it will remain the same after the averaging while the disturbances (e.g., noise and artifacts) will be attenuated since they occur randomly in each separate measurement. However, if the useful part

of the signal has a small random time shift due to physical reasons or inaccuracies in the measurement equipment, the averaging will distort this signal. This phenomenon can appear, e.g., in ultrasonic imaging. Airborne ultrasound suffers from inhomogeneities in the medium such as temperature gradients and air turbulence [10]. Using the above mentioned method we can estimate these small random time shifts and design an averaging algorithm that time aligns (i.e., causes the useful part of the signals to have the same position in time) so that very little signal shape distortion occurs when the averaging is carried out.

II. DERIVATION OF THE ESTIMATOR

We model all received echoes as having the same signal shape and only differing in the arrival time. Thus, using $r(t)$ as a reference echo we find that the received echo is

$$s(t) = r(t - \theta) \tag{1}$$

where the time delay θ is considered to be an unknown deterministic parameter. In many time delay estimators [3], [6], [7] the cross correlation

$$R_{sr}(\tau) = \int_{-\infty}^{\infty} s(t)r(t + \tau)dt \tag{2}$$

between the reference echo $r(t)$ and the received echo $s(t)$ is considered. The time lag maximizing this cross correlation is the estimated time delay. However, Cabot [5] suggested that the cross correlation between $s(t)$ and $\check{r}(t)$, the Hilbert transform of $r(t)$, should be used instead. This cross correlation is shown to be

$$R_{s\check{r}}(\tau) = \check{R}_{sr}(\tau) \tag{3}$$

In this case the cross correlation will not have a maximum at the time lag θ but a zero crossing. It can be argued that it is easier to find a zero crossing than a peak in a noisy signal. It introduces however an ambiguity since there can be several zero crossings, particularly if the cross correlation is narrowband. The Hilbert transform is defined as [2]

$$\check{r}(t) = \mathcal{H}\{r(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{r(x)}{x - t}dx = h(t) * r(t) \tag{4}$$

where the integral is a Cauchy Principal Value (CPV) and $*$ denotes convolution. The Hilbert kernel is denoted by $h(t) = -\frac{1}{\pi t}$. We observe that if $r(t) \in L^2(-\infty, \infty)$ then

it follows that $\check{r}(t) \in L^2(-\infty, \infty)$, see [9]. Consider the cross correlation between $s(t)$ and $\check{r}(t)$ at time lag zero and use Parseval's theorem

$$
\begin{aligned}
R_{s\check{r}}(0) &= \int_{-\infty}^{\infty} r(t-\theta)\check{r}(t)dt \\
&= \int_{-\infty}^{\infty} r(t-\theta)[h(t)*r(t)]dt \quad (5) \\
&= \int_{-\infty}^{\infty} \overline{e^{-j2\pi f\theta}R(f)}H(f)R(f)df \\
&= \int_{-\infty}^{\infty} e^{j2\pi f\theta}H(f)|R(f)|^2 df \quad (6)
\end{aligned}
$$

Here we let $H(f) = j \operatorname{sgn}(f)$ denote the Fourier transform of $h(t)$ [2]. Assuming that $r(t)$ is narrowband, i.e., its energy is concentrated in frequency intervals of size B around $\pm f_0$, and that $B\theta \ll 1$, $R_{s\check{r}}(0)$ can be approximated as

$$
\begin{aligned}
R_{s\check{r}}(0) &\approx \int_{-f_0-B/2}^{-f_0+B/2} e^{j2\pi f\theta}H(f)|R(f)|^2 df \\
&\quad + \int_{f_0-B/2}^{f_0+B/2} e^{j2\pi f\theta}H(f)|R(f)|^2 df \\
&\approx -je^{-j2\pi f_0\theta} \int_{-f_0-B/2}^{-f_0+B/2} |R(f)|^2 df \\
&\quad + je^{j2\pi f_0\theta} \int_{f_0-B/2}^{f_0+B/2} |R(f)|^2 df \\
&\approx \frac{E_r}{2}j\left(-e^{-j2\pi f_0\theta} + e^{j2\pi f_0\theta}\right) = -E_r \sin 2\pi f_0\theta \\
&= -E_r \sin \omega_0\theta \quad (7)
\end{aligned}
$$

where $E_r = \int_{-\infty}^{\infty} r^2(t)dt$ is the energy of $r(t)$ and $\omega_0 = 2\pi f_0$. We have assumed the exponential function to be constant over the integration interval. This gives us

$$
\sin \omega_0\theta \approx \frac{-R_{s\check{r}}(0)}{E_r} = \frac{-\int_{-\infty}^{\infty} s(t)\check{r}(t)dt}{\int_{-\infty}^{\infty} r^2(t)dt} \quad (8)
$$

This estimator can be expected to work well only for narrowband signals and small time delays θ. The assumption that $e^{j2\pi f\theta}$ is constant over the integration interval will not be valid otherwise. As we work with sampled signals, we give the discrete version of (8). The observation interval T, assumed to be larger than the duration of the echoes, is sampled at N points with interval $T_s = T/N$. Denoting these samples $s_n = s(nT_s)$, we may approximate the integrals in (8) with sums and yield

$$
\begin{aligned}
\hat{\theta} &= \frac{1}{\omega_0} \arcsin\left(\frac{-\int_0^T s(t)\check{r}(t)dt}{\int_0^T r^2(t)dt} \right) \\
&\approx \frac{1}{\omega_0} \arcsin\left(\frac{-T_s \sum_{k=0}^{N-1} s_k\check{r}_k}{T_s \sum_{k=0}^{N-1} r_k^2} \right) \quad (9)
\end{aligned}
$$

Using vectors and denoting them $\mathbf{s} = [s_0\, s_1 \cdots s_{N-1}]^T$, $\check{\mathbf{r}} = [\check{r}_0\, \check{r}_1 \cdots \check{r}_{N-1}]^T$ and $\mathbf{r} = [r_0\, r_1 \cdots r_{N-1}]^T$, we have

$$
\hat{\theta} = \frac{1}{\omega_0} \arcsin\left(-\frac{\mathbf{s}^T\check{\mathbf{r}}}{\mathbf{r}^T\mathbf{r}}\right) \quad (10)
$$

Since $r(t)$ is a reference echo, we may calculate $-\check{\mathbf{r}}/\mathbf{r}^T\mathbf{r} = \mathbf{b}$ one time for all and arrive at the fast and simple time delay estimator

$$
\hat{\theta} = \frac{1}{\omega_0} \arcsin(\mathbf{s}^T\mathbf{b}) \quad (11)
$$

If the angular centre frequency ω_0 is not known, it may be estimated by "inducing our own time delay". Introducing a constructed signal $x(t) = \mathbf{r}(t-T_s)$ we have $\theta = T_s$. In vector notation this is equal to $\mathbf{x}^T = \begin{bmatrix} 0 & \mathbf{r}^T \end{bmatrix}$. For this known θ we can rewrite (10) to estimate ω_0

$$
\hat{\omega}_0 = \frac{1}{T_s} \arcsin\left(\begin{bmatrix} 0 & \mathbf{r}^T \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} \right) \quad (12)
$$

III. PERFORMANCE OF THE ESTIMATOR

A. The Influence of Bandwidth

In the derivation of the estimator we assumed an infinitesimally small bandwidth. This is of course not the case for an arbitrary narrowband signal. To analyze the influence of bandwidth we assume that the signal $r(t)$ has the following Fourier transform:

$$
R(f) = \begin{cases} \sqrt{\dfrac{E_r}{2B}} & |f \pm f_0| < B/2 \\ 0 & \text{otherwise} \end{cases} \quad (13)
$$

This signal has energy E_r and single sided bandwidth B. In Appendix A we show that the bandwidth introduces a bias in the form of an multiplicative term:

$$
\sin \omega_0\hat{\theta} = \frac{-\int_{-\infty}^{\infty} s(t)\check{r}(t)dt}{\int_{-\infty}^{\infty} r^2(t)dt} = \operatorname{sinc}(B\theta)\sin \omega_0\theta \quad (14)
$$

We see that in order to keep this bias low, the product of bandwidth and time-delay must be small. We assume that the delay is small, $|\theta| < \frac{1}{4f_0}$, resulting in $|\omega_0\theta| < \pi/2$. Hence we have $|B\theta| \leq \frac{B}{4f_0} = \frac{1}{4\beta}$ where $\beta = f_0/B$. For narrowband signals β is in the order of 10–50, causing the bias factor $\operatorname{sinc}(B\theta)$ to be close to 1.

B. Noise Error

Assume that $r(t)$ and $s(t)$ are corrupted by independent additive white Gaussian noise. Since both $r(t)$ and $s(t)$ are bandlimited we can also assume that the noise has a limited bandwidth. If this is not the case, the received signals can always be lowpass filtered without any loss of information. Note that the Hilbert transform of a white Gaussian process $n(t)$ is also white and Gaussian:

$$
S_{\check{n}}(f) = S_n(f)|H(f)|^2 = S_n(f) \quad (15)
$$

Here we let $S_n(f)$ denote the power spectral density of the stochastic process $n(t)$, i.e., the Fourier transform of the autocorrelation function $R_n(\tau)$. If we include the noise in (8) we get

$$
\hat{\theta} = \frac{1}{\omega_0}\arcsin\left(\frac{-\int_0^T (s(t)+n(t))(\check{r}(t)+\check{n}_0(t))dt}{\int_0^T (r(t)+n_0(t))^2 dt} \right) \quad (16)
$$

where the reference echo is $r(t) + n_0(t)$ and the received echo $s(t) + n(t)$. The superscript $\check{}$ denotes Hilbert transformation and can numerically be calculated in the Fourier domain using FFT [15]. From the discrete version of the estimator (10), we have

$$\hat{\theta} = \frac{1}{\omega_0} \arcsin\left(\frac{-T_s \sum_{k=0}^{N-1} (s_k + n_k)(\check{r}_k + \check{n}_{0k})}{T_s \sum_{k=0}^{N-1} (r_k + n_{0k})^2}\right)$$
$$= \frac{1}{\omega_0} \arcsin\left(\frac{B}{A}\right) \tag{17}$$

The sampled processes n_k and n_{0k} are white and Gaussian [16] with variance σ^2. Since our estimator is nonlinear, it is difficult to analyze. However, the mean is much larger than the standard deviation for the denominator and we may linearize the estimator in (17), see Appendix B, to obtain:

$$E\{\hat{\theta}\} \approx \frac{1}{\omega_0} \arcsin\left(\frac{\mu_B}{\mu_A}\right)$$
$$E\left\{\left(\hat{\theta} - \theta\right)^2\right\} \approx \frac{\mu_A^2 \sigma_B^2 + \mu_B^2 \sigma_A^2}{\mu_A^2 (\mu_A^2 - \mu_B^2) \omega_0^2}$$
$$+ \left(\frac{1}{\omega_0} \arcsin\left(\frac{\mu_B}{\mu_A}\right) - \theta\right)^2 \tag{18}$$

where

$$\mu_A = E_r + T\sigma^2$$
$$\mu_B = E_r \sin \omega_0 \theta$$
$$\sigma_A^2 = 4E_r T_s \sigma^2 + 2TT_s \sigma^4$$
$$\sigma_B^2 = 2E_r T_s \sigma^2 + TT_s \sigma^4 \tag{19}$$

The mean of the estimate is

$$E\{\hat{\theta}\} = \frac{1}{\omega_0} \arcsin\left(\frac{E_r \sin \omega_0}{E_r + T\sigma^2}\right)$$
$$= \frac{1}{\omega_0} \arcsin\left(\frac{\text{SNR}}{\text{SNR} + 1} \sin \omega_0 \theta\right) \tag{20}$$

where we define

$$\text{SNR} = \frac{E_r}{T\sigma^2} \tag{21}$$

This definition of the SNR is the ratio between the energies of the signal and the noise in the observed time interval. It is equivalent to the definition in [11], which is SNR = $R_{rr}(0)/\sigma^2$, where $R_{rr}(\tau) = E\{r(t)r(t + \tau)\}$. For an ergodic stochastic process, $R_{rr}(\tau)$ can be approximated with $\frac{1}{T}\int_0^T r(t)r(t+\tau)dt$, if the observation interval is long enough. Hence, for large T the definitions of SNR's are equivalent. As we see from (20), the estimator can be expected to be asymptotically unbiased for large SNR but contain a bias for small SNR. As for the error variance, it can be shown from (18) that

$$\lim_{\text{SNR} \to -\infty \, \text{dB}} E\{(\hat{\theta} - \theta)^2\} = \frac{1}{N\omega_0^2} + \theta^2$$
$$\lim_{\text{SNR} \to \infty \, \text{dB}} E\{(\hat{\theta} - \theta)^2\} = 0 \tag{22}$$

From this we see that the standard deviation of the error for very low SNR will be approximately equal to the delay θ. Generally, for an arbitrary SNR, we will have an increase in error variance for larger θ.
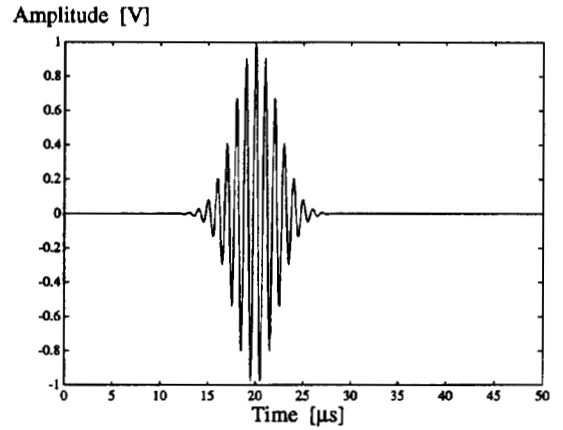
Amplitude [V]

Fig. 1.   Synthetic echo used in simulations.

C. Numerical Simulations

In order to verify the performance given by the previous section we simulated (10) numerically by using a synthetic echo in the form of

$$r(t) = e^{-a(t-t_0)^2} \cos \omega_0 t \qquad 0 < t < T \tag{23}$$

with the parameters $a = 10^{11} \, \text{s}^{-2}$, $\omega_0 = 2\pi 10^6 \, \text{rad/s}$, $t_0 = 20 \cdot 10^{-6} \, \text{s}$ and the sample frequency $F_s = 20$ MHz which corresponds to our equipment. The number of samples were $N = 1024$ corresponding to a measuring time of $T = 51.2 \, \mu s$. The synthetic echo is shown in Fig. 1. Since we used $f_s = 20$ MHz and the centre frequency $f_0 = 1$ MHz, we have 20 samples per period in the echo. For each value of the SNR, we simulated 1000 echoes and calculated the mean of the estimator (10). We added noise to both $r(t)$ and $s(t)$. The simplified expression of the mean (18) was compared to the mean of the simulation. In Fig. 2 , we show these curves as a function of SNR with $\theta = 0.1 T_s$. For comparison, we give the error variance for Cabot's estimator [11], i.e., the estimate $\hat{\theta}$ is the interpolated zero crossing of $R_{s\check{r}}(\tau)$.

$$E\left\{\left(\hat{\theta}_{\text{Cabot}} - \theta\right)^2\right\} = \frac{1}{B_n T \gamma_C b_s^2} \tag{24}$$

where

$$b_s \triangleq \frac{\int_{-\infty}^{\infty} |2\pi f| |S_r(f)|^2 df}{\int_{-\infty}^{\infty} |S_r(f)|^2 df} \tag{25}$$
$$\frac{1}{\gamma_C} \triangleq \frac{2\gamma + 1}{2\gamma^2} \tag{26}$$

and $\gamma = \frac{E_r}{T\sigma^2}$ is the SNR for $r(t)$ and $B_n = F_s/2$ is the single sided bandwidth of $n(t)$. Note that for a narrowband system we have $b_s \approx \omega_0$. The error variance in (24) can be rewritten as

$$E\left\{\left(\hat{\theta}_{\text{Cabot}} - \theta\right)^2\right\} = \frac{2\,\text{SNR} + 1}{N\,\text{SNR}^2\, b_s^2} \tag{27}$$

Due to approximations in [11], this error variance is however only valid for SNR $\gg -17$ dB in our simulation. In Fig. 3 the simulated and theoretical error variance of our estimator are plotted together with Cabot's. The curves are hard to tell
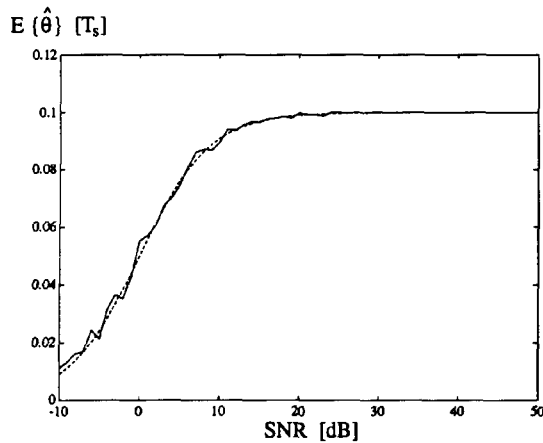
E {$\hat{\theta}$} [T_s]

Fig. 2. Simulated (solid) and theoretical (dashed) mean of the estimator for $\theta = 0.1\ T_s$.

E {$(\hat{\theta}-\theta)^2$} [T_s^2]

Fig. 3. Simulated (solid) and theoretical (dashed) error variance of the estimator for $\theta = 0.1\ T_s$. These are hard to distinguish. The dotted curve is the error variance of Cabot's estimator.

from each other for SNR > 10 dB. The reason why the error variance for the estimator (10) below this SNR is less than Cabot's is that it becomes biased for small SNR and hence not quite comparable.

IV. APPLICATION EXAMPLE: AVERAGING ALGORITHM

Using the estimator for small time delays (10), we can proceed to formulate an averaging algorithm. This method becomes useful when the echoes have small random time shifts. These could originate from fluctuations in the insonified medium. In our line of research, which contains surface profiling by using airbourne ultrasound [18], these fluctuations take the form of temperature variations and air turbulence. A change in the air temperature of 0.1°C results in a 0.06 m/s change of sound speed [13]. Using our transducer with a radius of curvature of 45 mm, i.e., a total path length of 90 mm for a surface in the focal plane, this means a time delay of 45 ns $= 0.95T_s$ in our case, where T_s is the sample interval. These small random time shifts have to be taken into account when an averaging process is applied. Otherwise we will introduce a distortion of the signal due to the misalignment of the echoes. Assume that we want to average over M echoes,

$$s_i(t) = r(t - \theta_i) + n_i(t) \qquad i = 0 \cdots M - 1 \qquad (28)$$

where we define $s_0(t) \stackrel{\triangle}{=} r(t) + n_0(t)$. In (28), θ_i are small independent random time shifts and $n_i(t)$ white Gaussian additive noise which are independent of θ_i. The first thing is to estimate the individual time delays between the echoes. This can be done as described in Section II. We use the first echo $s_0(t)$ as the reference echo. To be able to average over the M echoes, we need to align the echoes in time. This has been done by using an interpolation formula, which is approximately an all-pass filter with linear phase (corresponding to the time delay). Since we have samples of the echoes the problem is expressed as having $s_i(nT_s)$ and wanting $s_i(nT_s + \theta_i)$. We use the same notation as before, i.e., the samples are $s_{in} = s_i(nT_s)$. A simple interpolation method is Lagrange's
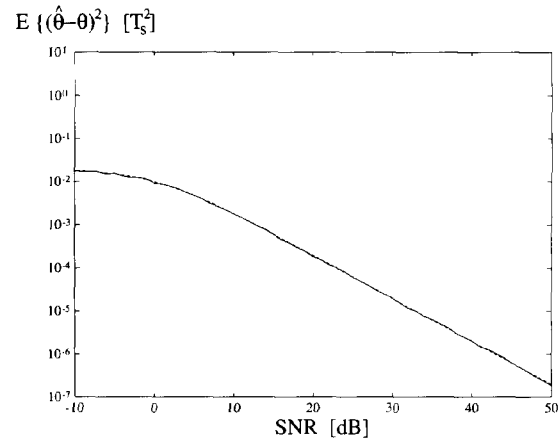
formula [1] of the order $2N + 1$

$$f(x + t\Delta) = \sum_{k=-N}^{N} h_k(t)f(x - k\Delta) \qquad (29)$$

which uses the fact that $f(x)$ is known in the equally spaced points $x_k = x - k\Delta$. The weights h_k depend on t which is the fraction of the distance Δ for which we want to find f. The interpolation (29) can be seen as a filter with taps h_k. We can now align s_{in} with s_{0n} by forming $s'_{in} = s_{in} * h_n(\theta_i)$ where $*$ denotes convolution. By doing this for all echoes s_{in} we can average over the M echoes:

$$\bar{s}_n = \frac{1}{M} \sum_{i=0}^{M-1} s'_{in} \qquad (30)$$

The averaged echo $\bar{s}_n$ will have the same position in time as s_{0n} since this is used as a reference echo. In order to find an absolute position in time we average over all estimated time delays

$$\bar{\theta} = \frac{1}{M} \sum_{i=0}^{M-1} \hat{\theta}_i \qquad (31)$$

This gives us an estimation of the correct relative position in time of $\bar{s}_n$. To obtain the final averaged signal we interpolate with the Lagrange interpolation filter with taps h_k which are determined by $\bar{\theta}$.

V. EXPERIMENTS

Apart from the difficulties of defining signal shape distortion, there are also practical problems. A noise-less reference echo is usually not available and thus it is difficult to measure the effect of the signal shape distortion. Instead we chose to get another quantitative measure of the performance of the algorithm. One area where it can be used is airborne ultrasound. We used a focused transducer with a radius of curvature of 45 mm, a diameter of 20 mm and a centre frequency of approximately 1.07 MHz . It was designed by H. W. Persson, Lund Institute of Technology, for the investigation presented in [18], and is acoustically adapted for air. We
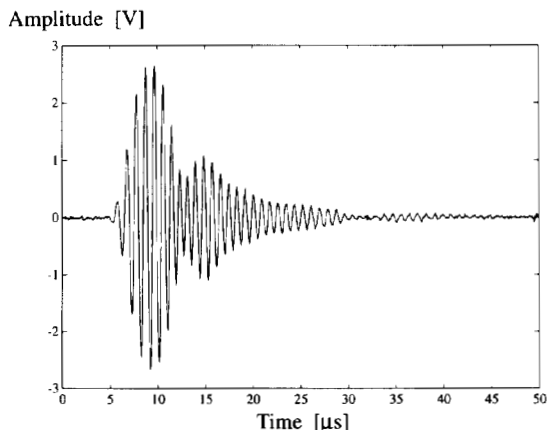
Amplitude [V]

Fig. 4.   Measured ultrasonic echo from the phantom.

Amplitude [V]

Fig. 5.   Cross section of the measured echoes for two values of t. The uppermost curve corresponds to $t = t_1$ which is a zero crossing. Total time for the 500 measurements was 30 seconds.

$\hat{\theta}$ [samples]

Fig. 6.   Estimated time shifts relative to the first echo for the measurements in Fig. 5.

Var$\{s(t)\}$ [V^2]

Fig. 7.   Variance of the cross section before (solid) and after (dashed) time alignment.

applied it to a phantom in the form of a massive brass block with smooth surfaces. The block was placed perpendicular to the transducer, the position of which was fixed throughout the whole experiment. We registered 500 echoes which took about 30 seconds with our equipment. These echoes, which should be identical since they were measured at the same spot, were then investigated. A measured echo is shown in Fig. 4 and the SNR was estimated to be approximately 30 dB.

VI. RESULTS

If we plot a cross section of the registered echoes, i.e., for a fixed value of the time t we plot how the amplitude varies for the different measurements, we can see how the misalignment effect can cause problems when we average the echoes. This cross-section plot is shown in Fig. 5. If the echoes did not have a random time shift, these curves would be straight horizontal lines apart from the measurement noise. The uppermost curve corresponds to a point in time where the reference echo has a peak, i.e., small derivative, and the lowermost to a point in time when the reference echo has a zero crossing, i.e., large derivative. Looking at the estimations of the time shifts relative to the reference (i.e., first) echo, Fig. 6, we can see that it has the same drift as the cross section. The magnitude of the derivative of the echo in Fig. 4 at the zero crossings is estimated to be almost 20 V/μs. From Fig. 6 and Fig. 5 we see that the delay of one sample, i.e., 50 ns, at a zero crossing results in an amplitude difference of approximately 1 V. If we linearize the echo at the zero crossing we get an amplitude of 20 V/μs $\times$50 ns $= 1$ V. From Fig. 6 we can see that it has a low-frequency shape which corresponds to the fluctuations in the medium [10] that appear during the total measurement time of 30 seconds. To see the effects of using our averaging algorithm, we look at the cross-section variance of $s(t)$ for each point of time, i.e., $v(t) = \text{Var}\{s(t)\}$. We compare the variance for the original echoes with the one where we estimated the individual time shifts and interpolated the echoes in order for them to be time aligned with the first echo. As we can see in Fig. 7, the alignment is very good. The peaks in variance for the original echoes correspond to points in time where the echoes' derivative is large, thus yielding large variations in the cross section. The variance after the
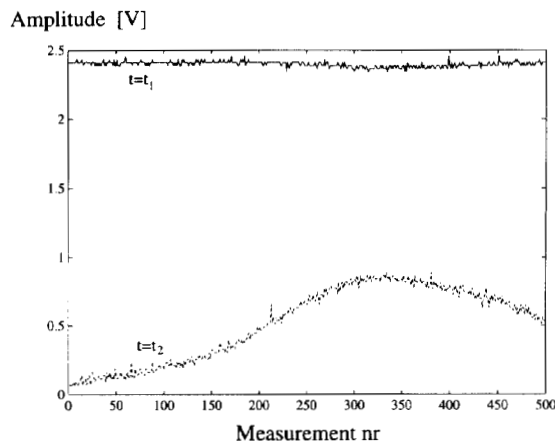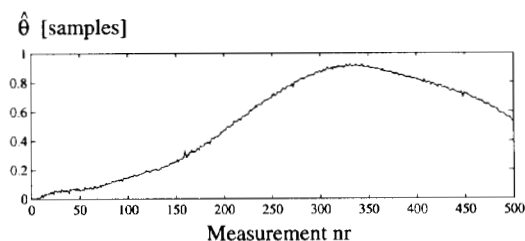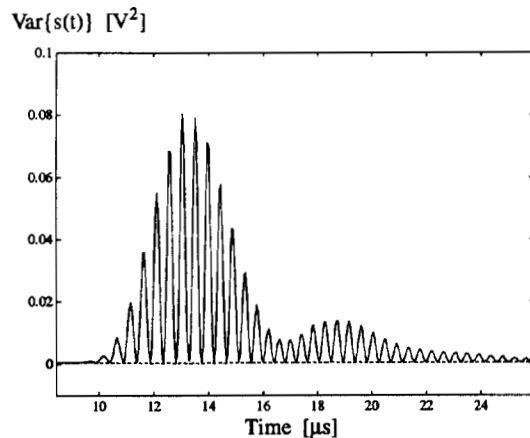
time alignment has a constant amplitude and does not vary with time. The small variations that are still present are mainly due to additive noise.

VII. DISCUSSION

We have proposed a new method for estimating small time delays in narrowband signals. Cabot [5] proposed to use the correlation with the Hilbert transform of the reference echo instead of the original echo. Our proposed estimator uses however only the Hilbert transform correlation for time lag zero. For narrowband signals we have shown that this cross

correlation has the shape of a sine wave in the neighbourhood of the true time delay θ. For the "standard" cross correlation this will be a cosine, but knowing only one point on that curve results in an ambiguity of the sign since cosine is an even function. Thus, the Hilbert transforming is necessary. For large SNR it can be shown from (18) that

$$E\{(\hat{\theta} - \theta)^2\} \sim \frac{2T_s}{T\omega_0^2 \cos^2 \omega_0 \theta} \cdot \frac{1}{\text{SNR}} \qquad (32)$$

while Cabot's estimator has the asymptotic shape [11]

$$E\{(\hat{\theta}_{\text{Cabot}} - \theta)^2\} \sim \frac{2T_s}{T\omega_0^2} \cdot \frac{1}{\text{SNR}} \qquad (33)$$

Hence, for large SNR and small time delays θ, these two methods have the same accuracy. For small time delays θ a bias is introduced in our method. Notice the lack of the usual "bumps" in the error variance curve. This is due to the effect that no peak ambiguity occurs since θ is assumed to be smaller than a quarter of a wavelength, i.e., $\frac{1}{4f_0}$. By using only one cross correlation value, the estimator consists of a scalar product with the normalized Hilbert transform of the reference echo, making it suitable for applications where multiple echoes are to be compared to a reference echo. Such applications include, e.g., sensor arrays and the proposed averaging algorithm. The latter can be used when the individual signals have a small random time shift between themselves. In this case we can have the advantage of using the above estimation method to align all signals in time before averaging them. The mean of the estimated time shifts can also be used to find a better absolute position in time for the averaged signal. A typical area of application for this method is ultrasonic measurements.

APPENDIX A
THE EFFECT OF BANDWIDTH

Assuming that the narrowbanded signal $r(t)$ has the Fourier transform

$$R(f) = \begin{cases} \sqrt{\frac{E_r}{2B}} & |f \pm f_0| < B/2 \\ 0 & \text{otherwise} \end{cases} \qquad (34)$$

where E_r denotes the energy and B the bandwidth, we have, using Parseval's theorem and the model $s(t) = r(t - \theta)$

$$\int_{-\infty}^{\infty} s(t)\breve{r}(t)dt = \int_{-\infty}^{\infty} \overline{e^{-j2\pi f\theta}R(f)} j\,\text{sgn}(f)R(f)df$$

$$= \int_{-f_0-B/2}^{-f_0+B/2} -je^{j2\pi f\theta}|R(f)|^2 df$$

$$+ \int_{f_0-B/2}^{f_0+B/2} je^{j2\pi f\theta}|R(f)|^2 df$$

$$= -j\frac{E_r}{2B}\left[\frac{e^{j2\pi f\theta}}{j2\pi\theta}\right]_{-f_0-B/2}^{-f_0+B/2}$$

$$+ j\frac{E_r}{2B}\left[\frac{e^{j2\pi f\theta}}{j2\pi\theta}\right]_{f_0-B/2}^{f_0+B/2}$$

$$= -j\frac{E_r}{2B}e^{-j2\pi f_0\theta}B\,\text{sinc}(B\theta)$$

$$+ j\frac{E_r}{2B}e^{j2\pi f_0\theta}B\,\text{sinc}(B\theta)$$

$$= j\frac{E_r}{2B}B\,\text{sinc}(B\theta)2j\sin(2\pi f_0\theta)$$

$$= -E_r\text{sinc}(B\theta)\sin(2\pi f_0\theta) \qquad (35)$$

Thus we have

$$\frac{-\int_{-\infty}^{\infty} s(t)\breve{r}(t)dt}{\int_{-\infty}^{\infty} r^2(t)dt} = \frac{E_r\text{sinc}(B\theta)\sin\omega_0\theta}{E_r}$$

$$= \text{sinc}(B\theta)\sin\omega_0\theta \qquad (36)$$

APPENDIX B
LINEARIZING THE ESTIMATOR

In order to linearize the estimator (17) we use the Taylor series of the function

$$f(x, y) = \frac{1}{\omega_0}\arcsin\left(\frac{x}{y}\right)$$

$$= f(x_0, y_0) + (x - x_0)\left.\frac{\partial f}{\partial x}\right|_{(x,y)=(x_0,y_0)}$$

$$+ (y - y_0)\left.\frac{\partial f}{\partial y}\right|_{(x,y)=(x_0,y_0)} + \cdots \qquad (37)$$

The partial derivatives are

$$\frac{\partial f}{\partial x} = \frac{1}{\omega_0}\frac{1}{\sqrt{1 - \left(\frac{x}{y}\right)^2}}\frac{1}{y} = \frac{1}{\omega_0}\frac{1}{\sqrt{y^2 - x^2}}$$

$$\frac{\partial f}{\partial y} = \frac{1}{\omega_0}\frac{1}{\sqrt{1 - \left(\frac{x}{y}\right)^2}}\frac{-x}{y^2} = -\frac{1}{\omega_0}\frac{x}{y}\frac{1}{\sqrt{y^2 - x^2}} \qquad (38)$$

Introducing ε_B, which is a zero-mean stochastic variable with variance $\sigma_B^2 = \text{Var}\{B\}$, and $\mu_B = E\{B\}$ we can approximate the estimator (17) by putting $B = \mu_B + \varepsilon_B$ and correspondingly $A = \mu_A + \varepsilon_A$ and use the above linearization at the point (μ_B, μ_A)

$$\hat{\theta} = \frac{1}{\omega_0}\arcsin\left(\frac{B}{A}\right) \approx$$

$$= \frac{1}{\omega_0}\arcsin\left(\frac{\mu_B}{\mu_A}\right) + \varepsilon_B\frac{1}{\omega_0\sqrt{\mu_A^2 - \mu_B^2}}$$

$$- \varepsilon_A\frac{\mu_B}{\omega_0\mu_A\sqrt{\mu_A^2 - \mu_B^2}} \qquad (39)$$

The mean of $\hat{\theta}$ is now found to be

$$E\{\hat{\theta}\} = \frac{1}{\omega_0}\arcsin\left(\frac{\mu_B}{\mu_A}\right) \qquad (40)$$

The error variance is

$$E\{(\hat{\theta} - \theta)^2\} = E\left\{\left(\frac{1}{\omega_0}\arcsin\left(\frac{\mu_B}{\mu_A}\right) - \theta\right.\right.$$

$$+ \frac{\varepsilon_B}{\omega_0 \sqrt{\mu_A^2 - \mu_B^2}}$$

$$\left. - \frac{\mu_B \varepsilon_A}{\omega_0 \mu_A \sqrt{\mu_A^2 - \mu_B^2}} \right)^2 \right\}$$

$$= \left(\frac{1}{\omega_0} \arcsin\left(\frac{\mu_B}{\mu_A}\right) - \theta \right)^2$$

$$+ \frac{\sigma_B^2}{\omega_0^2(\mu_A^2 - \mu_B^2)} + \frac{\mu_B^2 \sigma_A^2}{\omega_0^2 \mu_A^2 (\mu_A^2 - \mu_B^2)}$$

$$= \left(\frac{1}{\omega_0} \arcsin\left(\frac{\mu_B}{\mu_A}\right) - \theta \right)^2$$

$$+ \frac{\mu_A^2 \sigma_B^2 + \mu_B^2 \sigma_A^2}{\omega_0^2 \mu_A^2 (\mu_A^2 - \mu_B^2)}$$

$$(41)$$

if we assume that ε_A and ε_B are uncorrelated. The mean of B is found to be

$$E\{B\} = E\left\{ -T_s \sum_{k=0}^{N} s[k]\check{r}[k] + s[k]\check{n}_0[k] \right.$$

$$\left. + n[k]\check{r}[k] + n[k]\check{n}_0[k] \right\}$$

$$\approx -\int_{-\infty}^{\infty} s(t)\check{r}(t)dt \approx E_r \sin \omega_0 \theta \qquad (42)$$

since $n[k]$ and $\check{n}_0[k]$ are independent and have zero mean. The mean of A follows from

$$E\{A\} = E\left\{ T_s \sum_{k=0}^{N-1} r^2[k] + 2r[k]n_0[k] + n_0^2[k] \right\}$$

$$\approx \int_{-\infty}^{\infty} r^2(t)dt + T_s N\sigma^2 = E_r + T\sigma^2$$

$$(43)$$

For the variances of A and B we have

$$\sigma_B^2 = \text{Var}\{B\} = E\{(B - \mu_B)^2\}$$

$$= T_s^2 E\left\{ \left(\sum_{k=0}^{N-1} s[k]\check{n}_0[k] + n[k]\check{r}[k] + n[k]\check{n}_0[k] \right)^2 \right\}$$

$$= T_s^2 E\left\{ \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} s[k]\check{n}_0[k]s[l]\check{n}_0[l] + n[k]\check{r}[k]n[l]\check{r}[l] \right.$$

$$+ n[k]\check{n}_0[k]n[l]\check{n}_0[l] + 2s[k]\check{n}_0[k]n[l]\check{r}[l]$$

$$\left. + 2s[k]\check{n}_0[k]n[l]\check{n}_0[l] + 2n[k]\check{r}[k]n[l]\check{n}_0[l] \right\}$$

$$= T_s^2 \left(\frac{E_r}{T_s}\sigma^2 + \frac{E_r}{T_s}\sigma^2 + N\sigma^2\sigma^2 \right)$$

$$= 2E_r T_s \sigma^2 + NT_s^2\sigma^4 = 2E_r T_s \sigma^2 + T_s T\sigma^4 \qquad (44)$$

and

$$\sigma_A^2 = \text{Var}\{A\} = E\{(A - \mu_A)^2\}$$

$$= T_s^2 E\left\{ \left(\sum_{k=0}^{N-1} \left(2r[k]n[k] + n^2[k] - \sigma^2 \right) \right)^2 \right\}$$

$$= T_s^2 E\left\{ \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} 4r[k]n[k]r[l]n[l] + n^2[k]n^2[l] + \sigma^4 \right.$$

$$\left. + 4r[k]n[k]n^2[l] - 4r[k]n[k]\sigma^2 - 2n^2[k]\sigma^2 \right\}$$

$$= T_s^2 \left(4\frac{E_r}{T_s}\sigma^2 + (N^2 + 2N)\sigma^4 + N^2\sigma^4 - 2N^2\sigma^2\sigma^2 \right)$$

$$= 4E_r T_s \sigma^2 + 2NT_s^2\sigma^4 = 4E_r T_s \sigma^2 + 2T_s T\sigma^4 \qquad (45)$$

Here we have used the fourth moment of a zero-mean Gaussian stochastic variable $E\{x^4\} = 3\sigma_x^4$, [17], giving us

$$E\left\{ \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} n^2[k]n^2[l] \right\} = E\left\{ \sum_{k \neq l} n^2[k]n^2[l] + \sum_{k=l} n^4[k] \right\}$$

$$= (N^2 - N)\sigma^4 + N3\sigma^4$$

$$= (N^2 + 2N)\sigma^4 \qquad (46)$$

and with the third moment, $E\{x^3\} = 0$, [17],

$$E\left\{ \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} n[k]n^2[l] \right\} = E\left\{ \sum_{k \neq l} n[k]n^2[l] + \sum_{k=l} n^3[k] \right\}$$

$$= 0$$

$$(47)$$

REFERENCES

[1] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*, Nat. Bureau of Stand. Applied Math. Series 55, U.S. Govt. Printing Office, Wash., DC.
[2] R. N. Bracewell, *The Fourier Transform and its Applications*, 2nd ed. New York: McGraw-Hill, 1986.
[3] J. N. Bradley and R. L. Kirlin, "Delay estimation by expected value," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 1, pp. 19–27, 1984.
[4] R. E. Boucher and J. C. Hassab, "Analysis of discrete implementation of generalized cross correlator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 609–611, 1981.
[5] R. C. Cabot, "A note on the application of the Hilbert transform to time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, no. 3, pt. 2, pp. 607–609, Jun 1981.
[6] G. C. Carter, Ed., "Special issue on time delay estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pt 2, June 1981.
[7] ——, *Coherence and Time Delay Estimation* New York: IEEE Press, 1993.
[8] S.-K. Chow and P. M. Schultheiss, "Delay estimation using narrow-band processes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pt. 2, pp. 478–484, June 1981.
[9] H. Dym and H. P. McKean, *Fourier Series and Integrals*, New York: Academic, 1972.
[10] J. D. Fox, B. T. Khuri-Yakub and G. S. Kino, "High-frequency acoustic wave measurements in air," *Proc. IEEE 1983 Ultrason. Symp.*, pp. 581–584, Atlanta, GA, USA.
[11] D. Hertz, "Time delay estimation by combining efficient algorithms and generalized cross-correlation methods," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 1, pp. 1–7, Feb 1986.
[12] G. Jacovitti and G. Scarano, "Discrete time techniques for time delay estimation," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 525–533, 1993.
[13] L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, *Fundamentals of Acoustics*. New York: Wiley, 1980.
[14] R. Moddemeijer, "On the determination of the position of extrema of sampled correlators," *IEEE Trans. Signal Process.*, vol 39, no. 1, pp. 216–219, 1991.
[15] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[16] B. Picinbono, *Random Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[17] A. Papoulis, *Probability, random variables and stochastic processes*, 2nd ed. Singapore: McGraw-Hill, 1984.
[18] N. Sundström, P. O. Börjesson, N.-G. Holmer, L. Olsson, and H. W. Persson, "Registration of surface structures using airborne focused ultrasound," *Ultrasound in Med. and Biol.*, vol. 17, no. 5, 1991.
[19] H. L. van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley, 1968.
[20] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation," *IEEE Trans. Signal Process.*, vol. 41, no. 11, pp. 3102–3113, Nov. 1993.

Magnus Sandell (S'92) was born in Eksjö, Sweden in 1965. He received the M.Sc. degree in computer science in 1990 at Luleå University of Technology, Sweden.

Since 1992, he has been working toward the Ph.D. degree in signal processing at Luleå University of Technology. His main research interests are signal processing applied to ultrasonic imaging and image processing.

Anders Grennberg (M'79) was born in Härnösand, Sweden in 1939. He received the B.Sc. degree from Uppsala University in 1960 and the Ph.D. degree in mathematics from Umeå University in 1972, both in Sweden.

He has been a Senior Lecturer with Luleå University of Technology, Sweden since 1973, first in mathematics and then in signal processing. His research interest was Fourier analysis but has more recently been classification of transients and in inverse problems in ultrasonics.