Title: Robust statistics from multiple pings improves noise tolerance in sonar.

Authors: Nicola Neretti*, Nathan Intrator, Leon N Cooper

This work was supported in part by ARO (DAAD 19-02-1-0403), and ONR (N00012-02-C-02960).

N. Neretti is with the Institute for Brain and Neural Systems and Physics Department, Brown University, Providence RI 02912. N. Intrator is with the Institute for Brain and Neural Systems and Physics Department, Brown University, Providence RI 02912, and is on leave from Tel-Aviv University. L. N Cooper is with the Physics Department, Institute for Brain and Neural Systems and Neuroscience Department, Brown University, Providence RI 02912.

Edics: 1-DETE - Detection, Estimation, Classification, and Tracking; 1-NPAR - Parametric and Nonparametric Statistical Signal Processing.

Corresponding Author: Nicola Neretti, Box 1843, Brown University, Providence, RI 02912
Tel. (401) 863-2187, Fax (401) 863-3494, E-mail: Nicola_Neretti@brown.edu

Abstract

The ability of sonar to detect objects is strongly influenced by the operating signal-to-noise ratio (SNR). As sound amplitude decays very fast in water this sensitivity reduces the effective sonar range. It is well known that the range accuracy decays for increasing levels of noise until a breakpoint is reached after which accuracy deteriorates by several orders of magnitude. In this paper we present a robust fusion of time-delay estimates from multiple pings that reduces the SNR corresponding to the accuracy breakpoint. We show that a simple average of the time-delay estimates does not shift the breakpoint to a lower SNR. The method can improve the resilience to noise of a sonar system, hence increasing its potential range of operation.

I. INTRODUCTION

The theory of optimal receivers studies the design of pulses and receivers to obtain optimal detection in the presence of noise. Considerable work on the theoretical accuracy of range measurements has been done in the past starting with the Woodward equation, which has been derived using different methods. A comprehensive description of earlier work can be found in [1]. The validity of the Woodward equation depends on various assumptions in particular the assumptions of very low signal-to-noise ratios (SNR), therefore it must be reexamined for the case of low SNR's. The theory of optimal receivers shows that the *matched filter* receiver maximizes the output peak-signal-to-mean-noise (power) ratio [2, 3], and is the optimum method for the detection of signals in noise. Information about the distance of the target is extracted by computing the time at which the cross-correlation between the echo and a replica of the pulse is a maximum. This delay is converted into a distance by means of the sound velocity in the particular medium in consideration (e.g. water or air). This type of receiver is generally referred to as a *coherent* receiver.

The classical theory of optimal receivers describes the range accuracy of a sonar system via the well-known Woodward equation, which can be derived by using a variety of methods [1, 4-7]. For small SNR's, one of the parameters in the classical equation – i.e. the bandwidth – has to be modified, and the receiver is then called *semicoherent*. In [8] it was shown that the transition between the two types of behaviors occurs at different SNR's depending on characteristics of the pulses such as bandwidth and center frequency. With this observation, a novel system based on an adaptive choice of the pulse was proposed [8]that can improve accuracy in the case of relatively low SNR, when ambiguity in the choice of the correct peak of the cross-correlation function cannot be avoided.

Due to the nonlinear nature of the time-delay estimation problem, when the SNR drops below certain critical values *threshold effects* take place. Threshold effects can be characterized by a sharp deterioration of the time-delay estimator variance. Several statistical bounds have been used in the past to describe the accuracy of a matched filter receiver performance in between the above-mentioned SNR critical values. These include the Cramer-Rao lower bound [9-11], the Barankin bound [12], and the Ziv-Zakai bound [13-16]. Such bounds have been applied both to the problem of time-delay estimation [5, 17-23] and of frequency estimation [24-26]. In particular, the Barankin bound has been used to define the SNR breakpoints corresponding to the change in behavior of the optimal receiver as the SNR decreases for the case of a single ping and single echo [17, 19, 20], a single ping and multiple echoes [22, 23], and multiple pings and single echo [18, 21].

In this paper we analyze the signal-to-noise breakpoint, studying the probability of choosing the correct peak from the noisy cross-correlation function with a method similar to the one in [25], where the threshold effect was related to the existence of highly probable outliers far from the true time-delay value. This will enable us to extend the result to the case of multiple pings without *a priori* knowledge on the time-delay itself. This approach is different from that used in previous work on the multiple pings and single echo case [18, 21], where the multiple echoes for a single object are obtained artificially via multiple receivers and a unique ping. In fact, the bounds found in [18, 21] are valid only if the noise at the different receivers is totally uncorrelated or if the distance between transducer and receiver is constant, both conditions difficult to realize in practice.

Section II introduces two working models for the time delay estimation that are used to compute the probability of error form a single observation. In section III we show that while averaging of multiple observations improves accuracy, it does not increase noise tolerance. In section IV we derive a theoretical bound for the probability of correct time-delay estimation within a certain accuracy tolerance based on the mode of the observations. Section V describes a set of Monte Carlo simulations confirming the theoretical predictions.

II. SINGLE PING BREAKPOINT

A. A simplified model for the autocorrelation function: δ-function.

We first present a simple model where the noiseless cross-correlation function of the sonar ping is a δ-like function that is zero everywhere except from $t = 0$ where it is equal to the pulse energy (Fig. 1). This model requires a pulse with infinite bandwidth; therefore it is not realizable in practice. However, its study will give some insight on the mechanism leading to the sharp decay in range accuracy experienced by match-filtered receivers. When white Gaussian noise is added to the returning sonar ping, then the cross correlation vector has a Gaussian distribution with multidimensional centers at zero for all values that are outside of the bin t=0, and center at *a*=*E* > 0 for the value of the cross-correlation at *t*=0, where *E* is the energy of the noise-free echo (see Appendix B for a discussion on the validity of this approximation). Let x_i be the value at each bin of the cross-correlated signal and x_1 the value at t=0. We then have:

$$p\left(x_1,x_2,\ldots,x_N\right)=\frac{1}{\left(\sigma\sqrt{2\pi}\right)^N}e^{-\frac{\|\vec{x}-\vec{v}\|^2}{2\sigma^2}},\quad \vec{v}=\left[a,0,0,\ldots,0\right]^T \tag{1}$$

Using the Gaussian distribution, we can calculate the probability for the event $x_1 > x_i, i \neq 1$, which corresponds to the correct echo delay estimation, for a given noise level σ. Denoting this probability by α, we have:

$$\begin{aligned}P\left(X_1 > X_i, i \neq 1\right) = \alpha &= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{x_1} dx_2 \int_{-\infty}^{x_1} dx_3 \cdots \int_{-\infty}^{x_1} dx_N \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{\|\vec{x}-\vec{v}\|^2}{2\sigma^2}} \\ &= \int_{-\infty}^{+\infty} dx_1 \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{(x_1-a)^2}{2\sigma^2}} \left[\int_{-\infty}^{x_1} e^{-\frac{x^2}{2\sigma^2}} dx\right]^{N-1} \\ &= \frac{1}{2^{N-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\left(x-a/\sqrt{2}\sigma\right)^2} \left[1+erf(x)\right]^{N-1} dx\end{aligned} \tag{2}$$

Note that this probability depends only on $a/\sqrt{2}\sigma$, which is proportional to the signal to noise ration (SNR). $\beta = 1-\alpha$ denotes the probability of error, i.e., the probability that the amplitude of at least one of the points outside of the correct bin is larger than the amplitude correct peak.

B. A more realistic model for the autocorrelation function

For a ping with finite bandwidth it is more realistic to modify the model in equation (1) to include the finite extension in time of the autocorrelation function (Fig. 2). In this model the autocorrelation function is approximated by a piecewise constant function, with an amplitude equal to *a* within the central interval I_Δ of length Δ, and zero elsewhere. In this case, we need to consider the width of the *a priori* window of the cross-correlation. The width of the window corresponds to the sonar range. While the potential error in delay estimation is reduced when the width is reduced, so is the sonar range. If the *a priori* window has a length of $2L$ and the sampling frequency is f_s, then there will be $N = N_a + N_0 = 2L \cdot f_s$ points, $N_a = \Delta \cdot f_s$ of which will be within the central bin ("correct bin"), and N_0 outside the central bin but within the *a priori* window. Without loss of generality, we consider the case where N_a and N_0 are integers.

We define a random vector such that the first N_a random variables correspond to the amplitudes of the points within the correct bin, while the last N_0 correspond to the amplitudes of the points outside. As was derived above, for a white Gaussian noise, the joint probability density function for the vector of *n* random variables is given by:

$$p(x_1,x_2,\ldots,x_N)=\frac{1}{\left(\sigma\sqrt{2\pi}\right)^N}e^{-\frac{\|\vec{x}-\vec{a}\|^2}{2\sigma^2}},\quad \vec{v}=\Big[\underbrace{a,a,\ldots,a}_{N_a},\underbrace{0,0,\ldots,0}_{N_0}\Big]^T \tag{3}$$

The desired probability of time delay estimation within the correct bin of the center of the cross-correlation function is given by

$$\begin{aligned}
P\left(\underset{1\le j\le N}{\arg\max}\left\{X_j\right\}\in I_\Delta\right) &= \sum_{i=1}^{N_a}P\left(X_i>X_k,k\neq i\right)\\
&=\sum_{i=1}^{N_a}\int_{-\infty}^{+\infty}dx_i\int_{-\infty}^{x_i}dx_{k_1}\int_{-\infty}^{x_i}dx_{k_2}\cdots\int_{-\infty}^{x_i}dx_{k_{N-1}}\frac{1}{\left(\sigma\sqrt{2\pi}\right)^N}e^{-\frac{\|\vec{x}-\vec{v}\|^2}{2\sigma^2}}\\
&=N_a\int_{-\infty}^{+\infty}dx\frac{1}{\left(\sigma\sqrt{2\pi}\right)^N}e^{-\frac{(x-a)^2}{2\sigma^2}}\left[\int_{-\infty}^{x}e^{-\frac{(y-a)^2}{2\sigma^2}}dy\right]^{N_a-1}\left[\int_{-\infty}^{x}e^{-\frac{y^2}{2\sigma^2}}dy\right]^{N_0}\\
&=\frac{N_a}{2^{N-1}\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-\left(x-a/\sqrt{2}\sigma\right)^2}\left[1+erf\left(x-a/\sqrt{2}\sigma\right)\right]^{N_a-1}\left[1+erf(x)\right]^{N_0}dx
\end{aligned} \tag{4}$$

Thus, for a given noise level, the probability that the correct bin is selected is:

$$\alpha=\frac{N_a}{2^{N-1}\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-\left(x-a/\sqrt{2}\sigma\right)^2}\left[1+erf\left(x-a/\sqrt{2}\sigma\right)\right]^{N_a-1}\left[1+erf(x)\right]^{N_0}dx \tag{5}$$

C. Accuracy breakpoint

Let T be the R.V. whos probability distribution is given by Eq. 4. For SNR value that is above SNR_0, we have seen (Eq. 2) that the probability for falling in the correct bin α is above α_0. Let σ_Δ be the STD of the distribution of the echo location in the central (correct) bin, and let σ_0 be the STD of the distribution which is outside the central bin. Now, suppose we sample from the original distribution T whose cumulative function is given by equation (5). For *n* observations, a fraction of α of the *n* observations falls in the correct bin on average, while β*n* fall outside. The standard deviation of the distribution will then be give by:

$$\begin{aligned}
\text{std}^2(T)&=\alpha\ \text{std}^2(T^\Delta)+\beta\ \text{std}^2(T^0)\\
&=\alpha\sigma_\Delta^2+\beta\sigma_0^2
\end{aligned} \tag{6}$$

We define the breakpoint as the level of noise for which the contribution of T^0 to the total error becomes dominant. Thus, the RMSE will be significantly larger than the one given by the uniform distribution on I_Δ alone when

$$\alpha < \frac{\sigma_0^2}{\sigma_\Delta^2 + \sigma_0^2} \tag{7}$$

We then define the probability breakpoint to be:

$$\alpha_0 = \frac{(\sigma_0/\sigma_\Delta)^2}{1 + (\sigma_0/\sigma_\Delta)^2} \tag{8}$$

it is possible to find the SNR breakpoint as the SNR value for which equation (5) equals the value in (8).

III. Why does the mean fail?

The classical way to fuse information from multiple observations is to average. As the observations are independent and identically distributed, the central limit theorem (CLT) implies that the standard deviation (error) of the averaged R.V. should be $\sqrt{n}$ times smaller than the error made by each of the n observations separately. This is indeed the case before the breakpoint. However, this process does not improve the situation after the breakpoint and, in particular, does not shift the breakpoint to lower SNR's. Thus, while averaging improves accuracy, it does not increase noise tolerance. Below we provide a mathematical analysis which explains why the breakpoint does not change.

The measurement process described above is equivalent to sampling form a uniform distribution on the interval I_Δ with probability α and from a uniform distribution on the interval I_0, with a gap corresponding to I_Δ, with probability β. Suppose we sample n times to obtain $T_1, T_2, \ldots, T_n$ and use the sample mean as our estimate for the delay:

$$\overline{T} = \frac{1}{n}\sum_{i=1}^{n} T_i \tag{9}$$

On average, αn values will be in the correct bin, $T_1^\delta, T_2^\delta, \ldots, T_{\alpha n}^\delta$, while βn will be sampled from the uniform distribution, $T_1^u, T_2^u, \ldots, T_{\beta n}^u$. Then the sample mean can be decomposed into two parts:

$$\overline{T} = \frac{1}{n}\left(\sum_{i=1}^{\alpha n} T_i^\Delta + \sum_{i=1}^{\beta n} T_i^0 \right) \tag{10}$$

If σ_δ is the standard deviation of the δ-like distribution, and σ_u is the standard deviation of the uniform distribution, then, applying the central limit theorem to the two sums in equation (10) we obtain:

$$\begin{aligned} \left(\operatorname{std} \sum_{i=1}^{\alpha n} T_i^\Delta\right)^2 &= \alpha n \sigma_\Delta^2 \\ \left(\operatorname{std} \sum_{i=1}^{\beta n} T_i^0\right)^2 &= \beta n \sigma_0^2 \end{aligned} \tag{11}$$

The root-mean-square error will be significantly larger than the one given by the δ-like distribution alone when $\beta n \sigma_0^2 > \alpha n \sigma_\Delta^2$, i.e. when

$$\alpha < \frac{\sigma_0^2}{\sigma_\Delta^2 + \sigma_0^2} \tag{12}$$

It is observed that this bound does not improve with the number of pings and is equal to the bound found for a single ping in equation (7). This explains why averaging the echo delay estimates from multiple pings have not been found useful, and probably led to the wrong conclusion that no improvement can be achieved from multiple pings data.
It should be noted that it is possible to shift the breakpoint by estimating the time delay form the averaged cross-correlation functions of all the observations, a process that would require an extremely good alignment of the cross-correlation functions. This is basically equivalent to reducing the noise level by averaging the echoes [24] (see also discussion in Appendix A), which is not a realistic in a situation where the sonar is not completely still with respect to the target, or where it is not practical to store the entire echo waveforms for off line processing. However, a successful implementation of this averaging can be achieved by using multiple receivers [18, 21].

IV. USING THE MODE

Suppose we divide the *a priori* window into intervals of length Δ equal to the size of the correct bin, to obtain $m = \lfloor 2L/\Delta \rfloor$ intervals $B_1, B_1, \ldots, B_m$, with $B_1 = I_\Delta$ representing the correct bin (Fig. 3). If $p_1, p_2, \ldots, p_m$ are the probabilities for an estimate to fall in each of the intervals and $Y_1, Y_2, \ldots, Y_m$ are random variables representing the number of estimates falling in each interval, then

$$\sum_{j=1}^{m} Y_j = n$$
$$\sum_{j=1}^{m} p_j = 1 \tag{13}$$

The joint probability distribution for the number of estimates in each bin is given by the multinomial distribution

$$P\left(Y_1 = k_1, Y_2 = k_2, \ldots, Y_m = k_m\right) = \frac{n!}{k_1!k_2!\cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \tag{14}$$

The probability of choosing the correct bin using the mode is the probability that the number of estimates falling in the correct bin k_1, is greater than the number of estimates falling in any other bin $k_i, i \neq 1$:

$$P_{\text{correct}} = P\left(Y_1 > Y_j, \quad \forall j \neq 1\right) = \sum_{\substack{k_1, k_2, \ldots, k_m \\ k_1 > k_j, j \neq 1 \\ k_1 + k_2 + \cdots + k_m = n}} \frac{n!}{k_1!k_2!\cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \tag{15}$$

The sum in equation (15) can be decomposed into two parts: 1) the probability $P_{>50\%}$ that more that half of the n points falls into the correct bin; 2) the probability $P_{<50\%}$ that even if less than half of the n points fall in the correct bin, the number of points in it is greater than that of any other bin.

$$P\left(\text{correct bin}\right) = P_{>50\%} + P_{<50\%} \tag{16}$$

The first term in (16) can be written as:

$$\begin{aligned} P_{>50\%} &= \sum_{\substack{k_1, k_2, \ldots, k_m \\ k_1 > n/2 \\ k_1 + k_2 + \cdots + k_m = n}} \frac{n!}{k_1!k_2!\cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \\ &= \sum_{k_1 > n/2} \binom{n}{k_1} p_1^{k_1} \left(p_2 + p_3 + \cdots + p_m\right)^{n-k_1} \end{aligned} \tag{17}$$

The probability of an estimate to fall outside the correct bin is uniform over the a priori window with probability β, so that the probability for it to fall in any interval of size Δ is $\beta/(m-1)$:

$$p_1 = \alpha$$
$$p_j = (1-\alpha)/(m-1), \quad j \neq 1 \tag{18}$$

Substituting (18) into equation (17) we obtain

$$P_{>50\%} = \sum_{k>n/2} \binom{n}{k} \alpha^k (1-\alpha)^{n-k} , \tag{19}$$

where α is a function of the SNR through Eq. (5). The computation of $P_{<50\%}$ is more complicated. However, it is possible to derive an upper bound ($SNR_{>50\%}$) on the SNR breakpoint for the time-delay accuracy computed by using the mode of *n* estimates, as the SNR for which $P_{>50\%} = \alpha_0$, where α_0 is given in Eq. (8):

$$\sum_{k>n/2} \binom{n}{k} \alpha\left(SNR_{>50\%}\right)^k \left[1-\alpha\left(SNR_{>50\%}\right)\right]^{n-k} = \alpha_0 \tag{20}$$

The tighter bound corresponding to the total probability of choosing the correct bin given by Eq. (16) will always be lower than the one derived from Eq. (20):

$$SNR_{BP} \leq SNR_{>50\%} , \tag{21}$$

thus, the breakpoint which the mode can achieve will be for a lower SNR than the one calculated above, which is already significantly better than the breakpoint achieved by either a single ping or by averaging of the echo delay estimates of multiple pings (see Figure 4).

V. EXPERIMENTAL RESULTS

To test the mathematical results presented in the previous sections, we developed a set of Monte Carlo simulations using a cosine packet. We first analyzed the histograms of the errors in the delay estimate of the ideal receiver for different SNR's (figure 5). For high SNR (≥20dB) all the errors are small and follow the Woodward equation that corresponds to values within the central bin in figure 5a. As the level of the noise increases, large errors in the estimates appear. The errors are uniformly distributed over the entire a-priori window, and the relative ratio between the correct estimates (central bin) and the level of the uniform distribution decreases with SNR (figures 5b, 5c, and 5d). However, even for high levels of noise the central peak is significantly larger that the rest of the distribution.

Figure 6a shows the performance of ideal receiver for a single ping. For high SNR the accuracy follows the Woodward equation corresponding to a coherent ideal as expected from the theory of optimal receivers. The performance breaks for low SNR around 17 dB. Figures 6b, 6c and 6d show the analysis of the accuracy breakpoint for different number of pings – 10, 50 and 100 respectively. The blue line describes the optimal accuracy that can be achieved using cross-correlation from multiple pings. Its breaking point represents the optimal breaking point that could have been achieved using stationary sonar and target, and that could be predicted by using the Barankin bound as in [18, 21] (see also discussion in Appendix A). This breaking point however is not attainable, as it relies on careful registration of returns from different pings. Such careful registration can only be done if the distance between object to target is kept constant, or if it is known for each ping in advance. It can be seen that robust fusion of multiple pings based on the mode (light blue, and magenta lines) improves noise resiliency while retaining close to optimal achievable accuracy under multiple pings. In general there is no significant improvement in the resiliency to noise when a simple mean of the observations is used due the strong contamination of the distribution from outliers (red lines). This confirms the mathematical result presented in section III.

Figure 7, shows a summary of the results for the different methods. The breakpoint for the averaged cross-correlation function (blue squares) follows the ideal curve obtained by reducing the level of the noise ax explained in Appendix B (solid blue line). The breakpoint of the estimate obtained from the mean does not substantially change as the number of pings in increased. A more robust statistics such as the median improves the resiliency to noise as the number of pings increases (green triangles). The best results are obtained by using the mode of the estimates from the multiple pings (magenta diamonds).

APPENDIX A

In this section we show that estimating the time-delay from the averaged cross-correlation functions shifts the SNR breakpoint as if the noise level was reduced by averaging the echoes. We model each echo as a sum of two components:

$$\vec{y} = \vec{u} + \vec{\eta} \tag{22}$$

where $\vec{u}$ is an attenuated replica of the pulse, and $\vec{\eta}$ is white noise with standard deviation equal to σ_η. We define the signal to noise ration to be:

$$d^2 = \frac{2E}{N_0} \tag{23}$$

where E is the energy of the noise-free component $\vec{u}$ measured in Ws, and N_0 is the spectral density of the noise measured in W/Hz. If $f_s = 1/dt$ is the sampling frequency, then:

$$N_0 = \frac{\|\vec{\eta}\|^2 \, dt}{\frac{f_s}{2}} = \frac{\sigma_\eta^2}{\frac{f_s}{2}}$$
$$2E = 2\|\vec{u}\|^2 \, dt = \frac{\|\vec{u}\|^2}{\frac{f_s}{2}} \tag{24}$$

and the SNR can be expressed as:

$$d^2 = \frac{\|\vec{u}\|^2}{\sigma_\eta^2} \tag{25}$$

The SNR can be expressed in dB as:

$$SNR(dB) = 10\log_{10} d^2 = 10\log_{10} \frac{\|\vec{u}\|^2}{\sigma_\eta^2} \tag{26}$$

When averaging *n* echoes, the standard deviation of the gaussian process generating the noise for each echo is reduced by $\sqrt{n}$ due to the central limit theorem. Thus, the signal to noise ratio becomes:

$$SNR_n(dB) = 10\log_{10} d_n^2 = 10\log_{10} \frac{\|\vec{u}\|^2}{\left(\frac{\sigma_\eta}{\sqrt{n}}\right)^2} =$$
$$= 10\log_{10} \frac{\|\vec{u}\|^2}{\sigma_\eta^2} - 10\log_{10} n = SNR(dB) - 10\log_{10} n \tag{27}$$

If BP_1 is the breakpoint corresponding to a single ping, then the breakpoint for the average of *n* echoes will be:

$$BP_n = BP_1 - 10\log_{10} n \tag{28}$$

APPENDIX B

If $x(t)$ is the pulse and $y(t) = u(t) + \eta(t)$ is the echo, then their cross-correlation ϕ_{xy} can be expressed as:

$$
\begin{aligned}
\phi_{xy}(\tau) &= \int_{-\infty}^{+\infty} x(t)y(t+\tau)dt = \\
&= \int_{-\infty}^{+\infty} x(t)u(t+\tau)dt + \int_{-\infty}^{+\infty} x(t)\eta(t+\tau)dt = \\
&= \phi_{xu}(\tau) + \phi_{x\eta}(\tau)
\end{aligned} \tag{29}
$$

Moreover if the noise is white then the autocorrelation of the noise:

$$
R_{\eta\eta}(t_1,t_2) = \sigma_\eta^2 \delta(t_2 - t_1) \tag{30}
$$

If we treat $x(t)$ as a filter acting on the noise $\eta(t)$ to generate filtered version of the noise $\phi(t) = \phi_{x\eta}$, then the autocorrelation $R_{\phi\phi}$ is:

$$
\begin{aligned}
R_{\phi\phi}(t_1,t_2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} R_{\eta\eta}\left(t_1-\alpha, t_2-\beta\right) x\left(\alpha\right) x\left(\beta\right) d\alpha d\beta = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sigma_\eta^2 \delta\left(t_2-\beta-t_1+\alpha\right) x\left(\alpha\right) x\left(\beta\right) d\alpha d\beta = \\
&= \sigma_\eta^2 \int_{-\infty}^{+\infty} x\left(\beta+\Delta t\right) x\left(\beta\right) d\beta = \\
&= \sigma_\eta^2 \phi_{xx}\left(\Delta t\right)
\end{aligned} \tag{31}
$$

Thus, the noise on top of the cross-correlation function between the pulse and the noise-free component of the echo is not white, since its autocorrelation function is not a δ-function. However, since the support of ϕ_{xx} is small compared to the length of the a-priori window, we approximate $R_{\phi\phi}$ with $\sigma_\eta^2 \delta\left(\Delta t\right)$, and consider the noise to be white.

REFERENCES

[1] M. I. Skolnik, *Introduction to Radar Systems*, 1st ed: McGraw-Hill Book Company, 1962.

[2] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed: McGraw-Hill, 2000.

[3] D. O. North, "An Analysis of the Factors which Determine Signal/Noise Discrimination in Pulse-carrier Systems," RCA, Tech. Rept. PTR-6C, June 25 1943.

[4] A. J. Mallinckrodt and T. E. Sollenberger, "Optimum-pulse-time Determination," *IRE Trans.*, vol. PGIT-3, pp. 151-159, 1954.

[5] D. Slepian, "Estimation of Signal Parameters in the Presence of Noise," *IRE Trans.*, vol. PGIT-3, pp. 68-89, 1954.

[6] P. M. Woodward, *Probability and Information Theory, with Applications to Radar*. New York: McGraw-Hill Book Company, Inc., 1953.

[7] M. I. Skolnik, "Theoretical Accuracy of Radar Measurements," *IRE Trans.*, vol. ANE-7, pp. 123-129, 1960.

[8] N. Neretti, N. Intrator, and L. N. Cooper, "Adaptive pulse optimization for improved sonar range accuracy," *IEEE Signal Processing Letters*, 2003.

[9] H. Cramer, *Mathematical methods of statistics*. Princeton, NJ: Princeton Univ. Press, 1946.

[10] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81-91, 1945.

[11] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. Roy. Soc.*, vol. 222, pp. 309, 1922.

[12] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 20, pp. 477-501, 1946.

[13] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Trans. Commun.*, vol. COM-22, pp. 340-342, 1974.

[14] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bound on signal parameter estimation," *IEEE Trans. Information Theory*, vol. IT-21, pp. 90-93, 1975.

[15] L. P. Seidman, "Performance limitations and error calculations for parameter estimation," *Proc. IEEE*, vol. 58, pp. 644-652, 1970.

[16] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Information Theory*, vol. IT-15, pp. 386-391, 1969.

[17] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Information Theory*, vol. 43, pp. 1084-1093, 1997.

[18] S.-K. Chow and P. M. Schultheiss, "Delay estimation using narrow-band processes," *IEEE Trans. ASSP*, vol. ASSP-29, pp. 478-484, 1981.

[19] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Information Theory*, vol. IT-17, pp. 669-676, 1971.

[20] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Information Theory*, vol. IT-15, pp. 273-279, 1969.

[21] J. Tabrikian and J. L. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Processing*, vol. 47, pp. 2917 -2927, 1999.

[22] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation: Part 2 - Threshold phenomena," *IEEE Trans. Signal Processing*, vol. 42, pp. 1001-1007, 1994.

[23] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3102-3113, 1993.

[24] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. Signal Processing*, vol. 47, pp. 2398-2401, 1997.

[25] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Information Theory*, vol. IT-20, pp. 591-598, 1974.

[26] B. James, B. D. O. Anderson, and R. C. Williamson, "Characterization of threshold for single tone maximum likelihood frequency estimation," *IEEE Trans. Signal Processing*, vol. 43, pp. 817-821, 1995.

Figure Captions

Fig. 1 - Simple model where the noiseless cross-correlation function of the sonar ping is a δ-like function that is zero everywhere (within the *a priori* window of length *L*) except from $t = 0$ where it is equal to the pulse energy *a* (top). The bottom figure shows the probability of selecting a given time location in the cross-correlation function predicted by this model.

Fig. 2 - A more realistic model for the noiseless cross-correlation function includes the finite extension in time of the autocorrelation function. In this model the noiseless cross-correlation function is approximated by a piecewise constant function with an amplitude equal to *a* within the central interval I_Δ of length Δ and zero elsewhere (within the *a priori* window of length *L*). The bottom figure shows the probability of selecting a given time location in the cross-correlation function predicted by this model.

Fig. 3 – Model used in the computation of the probability of a correct bin choice in the case of multiple pings. We divide the a priori window into intervals of length Δ equal to the size of the correct bin, to obtain $m = \lfloor 2L/\Delta \rfloor$ intervals $B_1, B_1, \ldots, B_m$, with $B_1 = I_\Delta$ representing the correct bin.

Fig. 4 – Probability of making the correct choice as a function of SNR for different numbers of pings. The arrows indicate upper bounds on the SNR breakpoints.

Fig. 5 – Histograms of the errors in the delay estimate in a Monte Carlo simulation for different SNR's. For high SNR (≥20dB) all the errors are small and follow the Woodward equation that corresponds to values within the central bin in figure (a). As the level of the noise increases, large errors in the estimates appear. The errors are uniformly distributed over the entire a-priori window, and the relative ratio between the correct estimates (central bin) and the level of the uniform distribution decreases with SNR, see figures (b), (c), and (d). However, even for high levels of noise the central peak is significantly larger that the rest of the distribution.

Fig. 6 – RMSE as a function of SNR and number of pings – 1, 10, 100 and 200 respectively – for the Cosine Packet. Notice how the SNR breakpoint for the average of multiple pings (red line) does not decrease with the number of pings.

Fig. 7 – Breakpoint in dB as a function of number of pings for different methods. The solid line corresponds to the noise reduction due to averaging as discussed in Appendix C.

Figures

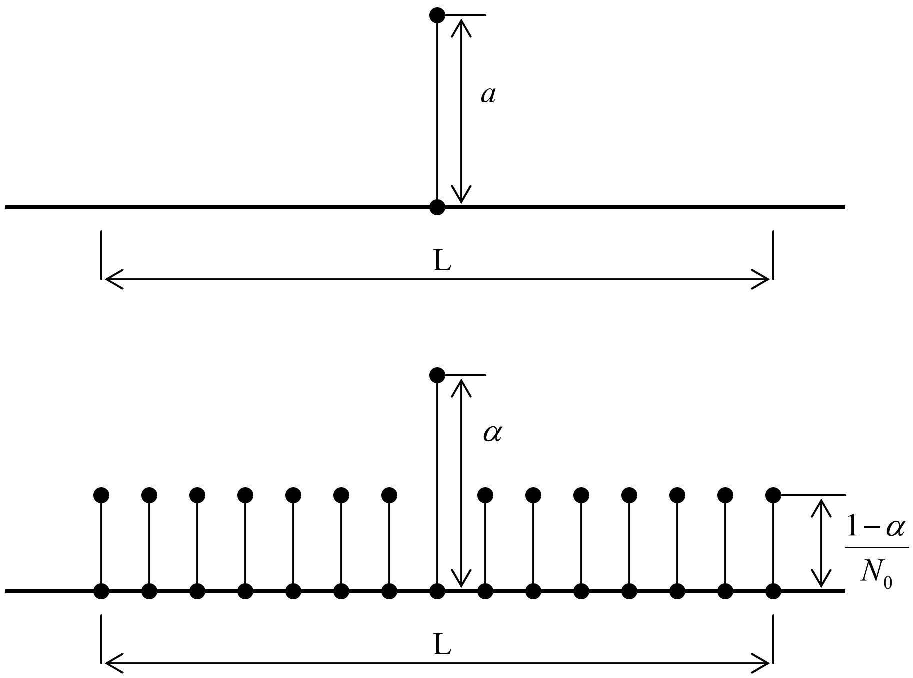


Fig. 1

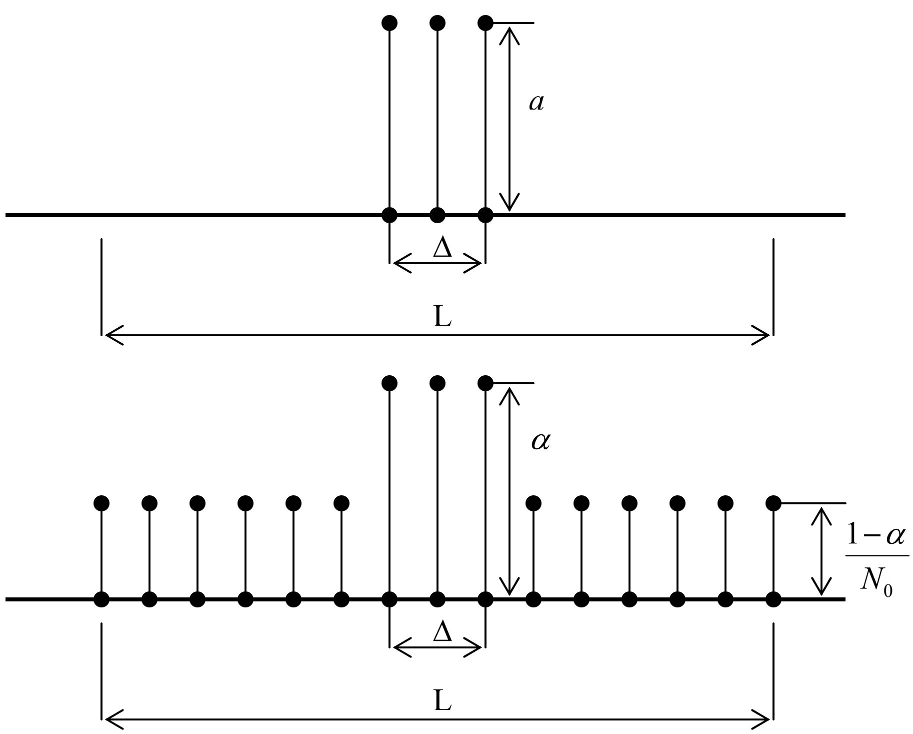


Fig. 2

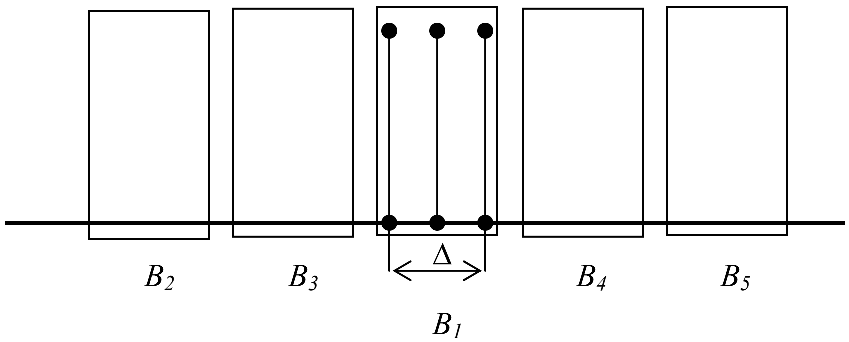

Δ
B2
B3
B4
B5
B1

Fig. 3

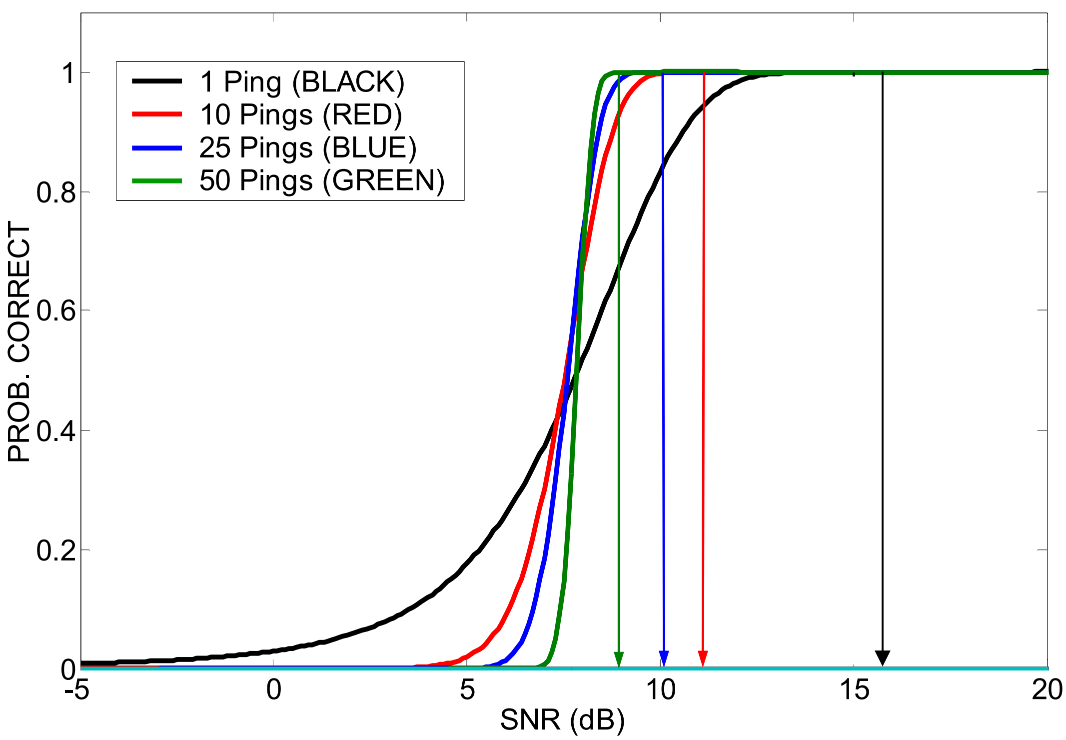

1 Ping (BLACK)
10 Pings (RED)
25 Pings (BLUE)
50 Pings (GREEN)
PROB. CORRECT
SNR (dB)
0
0.2
0.4
0.6
0.8
1
-5
0
5
10
15
20

Fig. 4

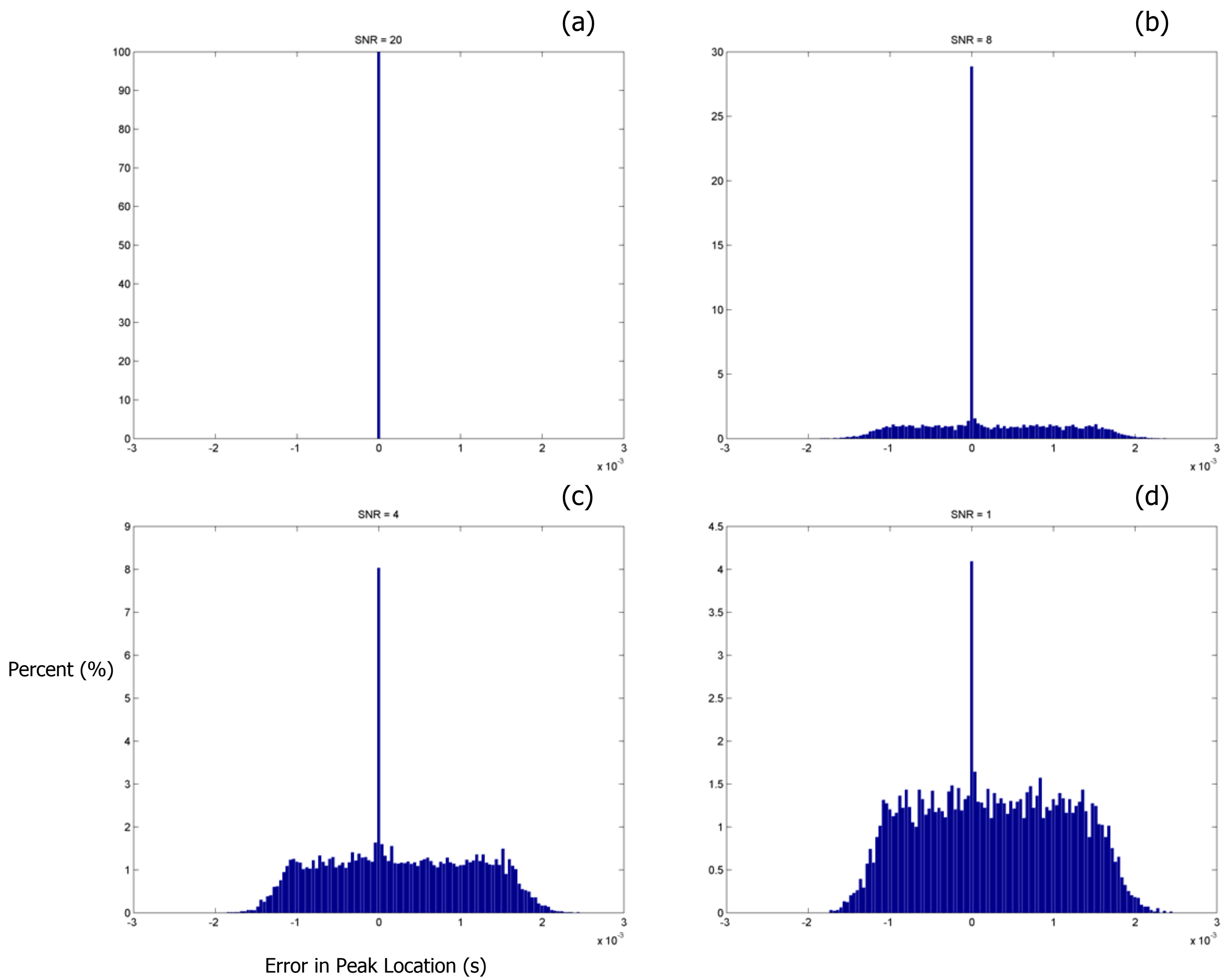
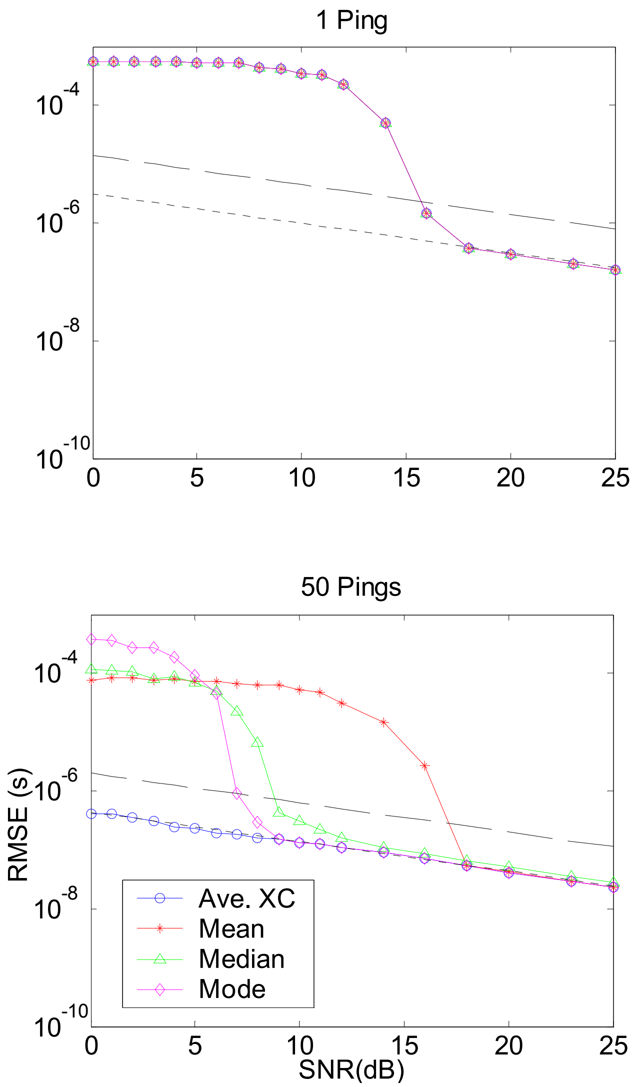
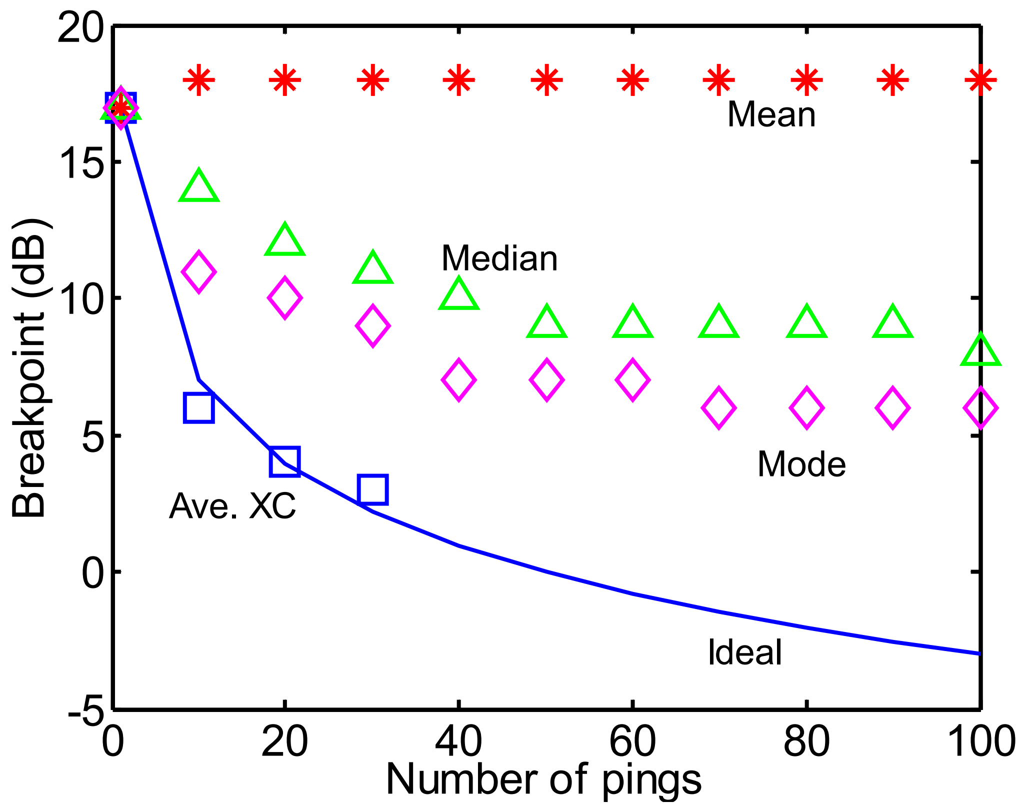

(a)
SNR = 20
(b)
SNR = 8
(c)
SNR = 4
(d)
SNR = 1
Percent (%)
Error in Peak Location (s)
x 10^-3

Fig. 5



1 Ping
10 Pings
50 Pings
100 Pings
RMSE (s)
SNR(dB)
Ave. XC
Mean
Median
Mode

Fig. 6



Mean
Median
Mode
Ave. XC
Ideal
Breakpoint (dB)
Number of pings
20
15
10
5
0
-5
0
20
40
60
80
100

Fig. 7