

Ultrasonic Imaging

<http://uix.sagepub.com/>

Effect of Prior Probability Quality on Biased Time-Delay Estimation

Brett C. Byram, Gregg E. Trahey and Mark L. Palmeri

Ultrason Imaging 2012 34: 65

DOI: 10.1177/016173461203400201

The online version of this article can be found at:

<http://uix.sagepub.com/content/34/2/65>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Ultrasonic Imaging and Tissue Characterization Symposium](#)

Additional services and information for *Ultrasonic Imaging* can be found at:

Email Alerts: <http://uix.sagepub.com/cgi/alerts>

Subscriptions: <http://uix.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://uix.sagepub.com/content/34/2/65.refs.html>

>> [Version of Record](#) - Apr 1, 2012

[What is This?](#)

Effect of Prior Probability Quality on Biased Time-Delay Estimation

BRETT C. BYRAM,¹ GREGG E. TRAHEY^{1,2} AND MARK L. PALMERI^{1,3}

*Departments of ¹Biomedical Engineering, ²Radiology and ³Anesthesiology
Duke University
Durham, NC 27708
brett.byram@duke.edu*

When properly constructed, biased estimators are known to produce lower mean-square errors than unbiased estimators. A biased estimator for the problem of ultrasound time-delay estimation was recently proposed. The proposed estimator incorporates knowledge of adjacent displacement estimates into the final estimate of a displacement. This is accomplished by using adjacent estimates to create a prior probability on the current estimate. Theory and simulations are used to investigate how the prior probability impacts the final estimate. The results show that with estimation quality on the order of the Cramer-Rao lower bound at adjacent locations, the local estimate in question should generally exceed the Cramer-Rao lower-bound limitations on performance of an unbiased estimator. The results as a whole provide additional confidence for the proposed estimator.

KEY WORDS: Bayes' theorem; motion estimation; prior probability; speckle tracking; ultrasound.

I. INTRODUCTION

Traditionally ultrasound motion estimation has been limited by the Cramer-Rao lower bound, which is the minimum variance obtainable for an unbiased estimator.¹⁻³ Unbiased estimators may still result in biased estimates, which result from data characteristics rather than from the algorithm used for motion estimation. Evidence for this can be seen in papers where unbiased algorithms are used and bias is reported: two examples are Pinton et al and Byram.^{4, 5} Bias is appropriately combined with variance through the mean-square error (MSE). Bias and variance contribute equal parts to the total estimation error but in typical clinical ultrasound applications, the bias' contribution to error is significantly smaller than the variance contribution. Because the contribution of bias and variance to the total estimation error is usually weighted towards the variance, a large decrease in variance for a small increase in bias would almost always be considered a useful estimation property. This trade-off is often possible using biased estimators, which are usually implemented using Bayes' theorem.⁶ While a Bayesian approach to biased estimation has not previously existed for the ultrasound time-delay estimation problem, many *ad hoc* algorithms exist to restrict the search region around the expected displacement and these solutions may be considered nonBayesian approaches to biased estimation. A notable example is the approach by Zahiri-Azar and Salcudean.⁷ Their approach uses displacements estimated at adjacent locations in a medium to aggressively restrict the search window at the current estimate. This has advantages of decreasing computational load and decreasing peak hopping artifacts, which, in turn, results in decreased estimation variance without significantly changing estimation bias. Similar improvements in estimation can be realized using regularization methods first popularized in ultrasound by Pellet-Barakat et al.⁸ A more general approach to the problem of biased estimation was recently presented using Bayes' theorem.⁵ One of the challenges with Bayesian approaches is understanding how the prior information impacts the final esti-

mate. This is particularly true in the case when prior probabilities are an attempt at estimating the prior information rather than utilizing prior information known from other factors, such as experimental design.

This paper explores the effect that the quality of prior information has on displacement estimation; here, the notion of prior information quality will refer to the difference between the true displacement and the mean of the prior distribution as well as the width of the prior distribution (e.g. standard deviation, bandwidth, etc.). In order to provide intuition to the notion of prior quality, several examples are provided of various levels of prior probability quality. The highest quality prior would be a delta function located at the true displacement (although this is generally not an interesting case assuming the data has a region of support encompassing the prior delta function because the data become irrelevant). This high-quality prior can be contrasted against a prior with a small standard deviation but a grossly-incorrect mean value, which represents a worst case prior. An example midquality prior would have a mean that is close to or exactly the true displacement but with a sufficiently broad standard deviation for the final estimate to be dominated by the data. The quality of the prior distribution can be assessed quantitatively using the mean-square error of the displacement estimate relative to a baseline measure of quality. A possible baseline quality measure for time-delay estimation is the Cramer-Rao lower bound (CRLB). For this scenario, quality could be expressed quantitatively as

$$Quality_{p(\hat{d}_0)} = \frac{MSE_{\hat{d}_0}}{CRLB_{\hat{d}_0}} \quad (1)$$

where $p(\hat{d}_0)$ represents the prior, \hat{d}_0 is the estimated displacement and MSE is the mean-square error for the estimate of \hat{d}_0 . Based on this representation of quality, a prior is better as the quality becomes more negative. In reality, Eq. (1) is mostly useful to connect the quality of the prior to the unbiased performance restricted by the Cramer-Rao lower bound. The purpose here is to determine how well a true displacement has to be known in order to provide a better displacement estimate than that obtained using a noninformative prior. The noninformative prior for ultrasound time-delay estimation is a uniform probability distribution equivalent in size and position to the search region. The noninformative prior results in an algorithmically-unbiased estimate limited by the CRLB. In order to assist in determining the necessary quality to surpass performance dictated by the CRLB, analytic expressions that approximate estimator performance over a wide range of prior distributions will be derived. The analytic expressions themselves may prove useful in order to decide whether a displacement estimate is likely to have a quality that exceeds the CRLB.

II. METHODS

Theory

There are many ways to estimate parameters from probability distributions.⁹ The derivation to follow will focus on the minimum mean-square estimator (MMSE) of parameters and will derive the associated estimator bias, variance and mean-square error. The choice to focus on a particular method of estimating a parameter from a posterior distribution should not be considered particularly restrictive because several useful assumptions that will eventually be made in the derivation result in different parameter estimators having nearly identical behavior.

The minimum mean-square error estimator for the time-delay τ_0 is

$$\hat{\tau}_0 = \int \tau_0 p(\tau_0 | x) d\tau_0, \quad (2)$$

where $\hat{\tau}_0$ is the estimated time-delay and $p(\tau_0 | x)$ is the posterior distribution expressing knowledge of the displacement, which can be found using Bayes' theorem. Bayes' theorem is

$$p(\tau_0 | x) = \frac{p(x | \tau_0) p(\tau_0)}{\int p(x | \tau_0) p(\tau_0) d\tau_0}, \quad (3)$$

which shows that in order to obtain the necessary posterior distribution, a prior probability for τ_0 and a likelihood function are required. The likelihood function for the time-delay estimation problem is a canonical result that can be found in many texts.¹⁰⁻¹³ The likelihood function for this problem is

$$p(x | \tau_0) = \frac{1}{(4 \frac{\sigma_{noise}^2}{n})^{\frac{N}{2}}} \exp\left[-\frac{1}{(4 \frac{\sigma_{noise}^2}{n})} \sum_{n=0}^{N-1} s_1(n - \tau_0)^2 \right] \quad (4)$$

$$\exp\left[-\frac{1}{(4 \frac{\sigma_{noise}^2}{n})} \sum_{n=0}^{M-1} 2s_1(n - \tau_0) s_2(n - \tau_0) + s_2(n - \tau_0)^2 \right]$$

where s_1 and s_2 are the two signals with relative displacement, $\frac{\sigma_{noise}^2}{n}$ is the noise power (the noise power has been doubled based on the argument by Walker,¹⁴ τ_0 is the sampling period, and M is the number of samples in a kernel. Additionally, this is the likelihood function derived from the assumption of additive Gaussian-distributed band-limited noise.

For the purpose of deriving an analytic expression for the displacement estimates, the prior distribution will be assumed to be normally distributed,

$$p(\tau_0) = \frac{1}{\sqrt{2 \frac{\sigma_p^2}{p}}} \exp\left[-\frac{1}{2 \frac{\sigma_p^2}{p}} (\tau_0 - \mu_p)^2 \right] \quad (5)$$

where μ_p describes the expected location of τ_0 before we consider the data and $\frac{\sigma_p^2}{p}$ is a measure of the confidence in the knowledge of the displacement. The two parameters of the normal distribution, μ_p and $\frac{\sigma_p^2}{p}$, are at the heart of this paper and will be the parameters used to modulate the displacement estimation quality as described in the introduction. Normal distributions are often over-applied based on faulty assumptions but for this scenario, a normal distribution is appropriate because it represents the least informative distribution when only the mean and standard deviation are known.^{15, 16} The normal prior assumption nicely restricts the prior parameter space to two dimensions while assuming the least amount of additional information.

Inserting the likelihood function (Eq. (4)) and the prior distribution (Eq. (5)) into Eq. (3) canceling terms in the numerator and denominator lacking dependency on τ_0 , shifting the reference point of the correlation shift by $M/2$, and assuming τ_0 is small enough to replace the summation with an integral (mirroring the derivation by Kay¹⁰), the posterior distribution can be expressed as

$$p(x | \tau_0) = \frac{\exp\left[-\frac{1}{2} \frac{\tau_s}{\text{noise}} s_1(\tau_0) s_2(\tau_0)\right] \exp\left[-\frac{1}{2} \left(\frac{\tau_0}{p}\right)^2\right]}{\exp\left[-\frac{1}{2} \frac{\tau_s}{\text{noise}} s_1(\tau_0) s_2(\tau_0)\right] \exp\left[-\frac{1}{2} \left(\frac{\tau_0}{p}\right)^2\right] d_0} \tag{6}$$

where $T_s = M^{-10}$

A noise-free signal model is then asserted since noise is already modeled in Eq. (6),

$$s(t) = A e^{-\frac{t^2}{2 \sigma_{sig}^2}} \sin(\omega t). \tag{7}$$

In addition to the signal model, the assumption is made that $s_1(t)$ and $s_2(t)$ are identical except they are time-shifted by $\tilde{\tau}_0$. (It should be noted that this is a strong assumption that assumes perfect correlation between the two signals minus the thermal noise. While this scenario is rarely encountered *in vivo* or in phantoms, there are several similar classes of algorithms that aim to correct signal decorrelation and restore the ideal behavior obtained when correlating signals with only bulk motion shifts.^{17, 18}) Based on the signal model, the correlation between the two signals is

$$R(\tau_0) = s_1(\tau_0) s_2(\tau_0) d = A^2 \sqrt{\frac{2}{\sigma_{sig}^2}} e^{-\frac{(\tau_0 - \tilde{\tau}_0)^2}{4 \sigma_{sig}^2}} \cos(\omega(\tau_0 - \tilde{\tau}_0)). \tag{8}$$

$$p(\tau_0 | x) = \frac{\exp\left[A^2 \sqrt{\frac{2}{\sigma_{sig}^2}} e^{-\frac{(\tau_0 - \tilde{\tau}_0)^2}{4 \sigma_{sig}^2}} \cos(\omega(\tau_0 - \tilde{\tau}_0))\right] \exp\left[-\frac{1}{2} \left(\frac{\tau_0}{p}\right)^2\right]}{\exp\left[A^2 \sqrt{\frac{2}{\sigma_{sig}^2}} e^{-\frac{(\tau_0 - \tilde{\tau}_0)^2}{4 \sigma_{sig}^2}} \cos(\omega(\tau_0 - \tilde{\tau}_0))\right] \exp\left[-\frac{1}{2} \left(\frac{\tau_0}{p}\right)^2\right] d_0}. \tag{9}$$

So the posterior distribution of τ_0 with the introduced signal model is

Next, the unscaled posterior distribution is considered independent of the full posterior distribution expression shown above (Eq. (9)); the unscaled likelihood function is expressed more conveniently using a Taylor’s series representation of the outer exponential, resulting in

$$p(x | \tau_0) p(\tau_0) = \sum_{m=0} \frac{\left(\frac{A^4}{8} \frac{2}{\sigma_{sig}^2}\right)^{\frac{m}{2}} e^{-\frac{(\tau_0 - \tilde{\tau}_0)^2}{4 \sigma_{sig}^2}} \cos^m(\omega(\tau_0 - \tilde{\tau}_0))}{m!} \exp\left[-\frac{1}{2} \left(\frac{\tau_0}{p}\right)^2\right], \tag{10}$$

m is the only variable introduced and is the index for the terms in the Taylor series. This is a convenient method for expressing the unscaled posterior because it allows the two Gaussian distributions to be consolidated to give

$$p(x|_0)p(\tilde{0}) = \frac{1}{m!} \exp\left[-\frac{(\frac{p}{2}\tilde{0})^2}{2(\frac{p}{2}\frac{sig}{2})}\right] \left(\frac{A^4 \frac{2}{4} \frac{sig}{2}}{8 \frac{noise}{2}}\right)^{\frac{m}{2}} \quad (11)$$

$$\exp\left[\frac{\frac{m}{p} \frac{2\tilde{0}}{2} \frac{2}{2} \frac{sig}{p}}{\frac{m}{p} \frac{2}{2} \frac{sig}{2}}\right] \cos^m\left(\frac{2}{m} \frac{2}{p} \frac{2}{2} \frac{sig}{sig}\right)$$

This expression can be inserted into the following equations in order to calculate the moments that quantify the estimator. The scaling of the posterior is

$$K = \int_0^\infty p(x|_0)p(\tilde{0})d_0. \quad (12)$$

The bias of the estimate $\hat{0}$ is

$$bias_{\hat{0}} = [\hat{0} - \tilde{0}] = \frac{\int_0^\infty \tilde{0} p(x|_0)p(\tilde{0})d_0}{K} - \tilde{0}. \quad (13)$$

And the variance of the estimation error is

$$var_{\hat{0}} = [(\hat{0} - \tilde{0})^2] = \frac{\int_0^\infty \tilde{0}^2 p(x|_0)p(\tilde{0})d_0}{K} - \left(\frac{\int_0^\infty \tilde{0} p(x|_0)p(\tilde{0})d_0}{K}\right)^2. \quad (14)$$

If the order of the integral in Eqs. (12), (13) and (14) and the sum encountered in Eq. (11) are exchanged, the resulting integral form

$$\int_0^\infty \exp\left[\frac{\frac{m}{p} \frac{2\tilde{0}}{2} \frac{2}{2} \frac{sig}{p}}{\frac{m}{p} \frac{2}{2} \frac{sig}{2}}\right] \cos^m\left(\frac{2}{m} \frac{2}{p} \frac{2}{2} \frac{sig}{sig}\right) d_0 \quad (15)$$

has a closed form solution (only $r = 0, 1$, and 2 will be required). The necessary substitutions and identities used for solving the closed form solution to Eq. (15) can be found in the appendix of the work by Byram.¹⁹

The analytic forms for $bias_{\hat{0}}$ and $var_{\hat{0}}$ are long, but there are several common terms. The variable representing the variance of the posterior Gaussian weighting for each m of the summation is

$$\frac{2}{m} \frac{2}{p} \frac{2}{sig}. \quad (16)$$

The variance of the posterior is a combination of the variance of the prior and bandwidth of the original signal expressed as a variance. The interpretation is complicated by the m index from the series expansion but otherwise, the equation shows that if the prior's variance is large relative to the signal's bandwidth, then the variance of the posterior is dominated by the signal information and vice versa.

Similar to the combined variance, the variable representing the combined mean of the Gaussian from combining the likelihood function and the prior is

$$\mu_m = \frac{m \sigma_p^2 \tilde{\mu}_0 + 2 \sigma_{sig}^2 \mu_p}{m \sigma_p^2 + 2 \sigma_{sig}^2} \tag{17}$$

The form of this equation is similar to the form of the combined variance in Eq. (16). However, the new mean is exactly a weighted average of prior and signal correlation functions means, where the weights are determined by the variances of the prior and the signals' correlation function. Practically, the way the variances and means of the signal and prior interact is the primary determinant of the final quality of a displacement estimate. Finally, it was useful to lump several parameters together into a common scaling term that was useful for all the moments of the posterior distribution. The scaling factor is

$$C_m = \exp \left[-\frac{(\mu_p - \tilde{\mu}_0)^2}{2 \left(\frac{\sigma_p^2}{m} + \frac{\sigma_{sig}^2}{2} \right)} \right] \frac{A^4 \sigma_{sig}^2}{8 \left(\frac{\sigma_{noise}^2}{2} \right)^2} \sqrt{2 \sigma_m^2} \tag{18}$$

$$\exp \left[-\frac{(\mu_p - \tilde{\mu}_0)^2}{2 \left(\frac{\sigma_p^2}{m} + \frac{\sigma_{sig}^2}{2} \right)} \right] \frac{SNR^2}{2} \sqrt{2 \sigma_m^2}$$

The scaling factor includes normalization terms from the Gaussian distribution, SNR scaling and residuals from completing the square to form the new Gaussian distribution. Using the expressions just defined K , $E[\hat{\mu}_0 - \tilde{\mu}_0]$ and $E[(\hat{\mu}_0 - \tilde{\mu}_0)^2]$ can be expressed as

$$K = \frac{C_m}{m^0 m!} \frac{1}{2^{m-1}} \frac{m}{\frac{m}{2}} \tag{19}$$

$$\frac{C_m}{2^{m-1} m!} \sum_{k=0}^m \cos[(m-2k)(\mu_p - \tilde{\mu}_0)] \exp \left[-\frac{\sigma_m^2 (m-2k)^2}{2} \right],$$

$$E[\tilde{\mu}_0] = \frac{1}{K} \frac{C_m}{m^0 m!} \frac{1}{2^{m-1}} \sum_{k=0}^m \frac{C_m}{2^{m-1} m!} \sum_{k=0}^m \cos[(m-2k)(\mu_p - \tilde{\mu}_0)] \tag{20}$$

$$\sin[(m-2k)(\mu_p - \tilde{\mu}_0)] \sigma_m^2 (m-2k) \exp \left[-\frac{\sigma_m^2 ((m-2k)^2)}{2} \right],$$

Table I. Data simulation and displacement-tracking parameters.

Parameter	Value
Center frequency	5 MHz
Bandwidth	50%
A	1
c	1540 m/s
Sampling frequency	10 GHz
Kernel length	2.75
SNR	20 dB
m_{max}	170
Search offset	$\pm /4$

roduce an additional source of bias that would complicate the results. All the parameters used in the model and the simulations (unless stated otherwise) are summarized in table 1.

Technically, the model calculates the mean-squared error of the estimator while the simulation data are reported as the mean of the squared residuals,

$$MSR = \frac{1}{N} \sum_{n=1}^N (\hat{\alpha}_{0_n} - \tilde{\alpha}_{0_n})^2 \quad (23)$$

where $\tilde{\alpha}_{0_n}$ is the estimated value, $\hat{\alpha}_{0_n}$ is the true value known from the simulations and N is number of simulated displacements. Eq. (23) is related to the commonly-used residual sum of squares.

III. RESULTS

Some specifics of model implementation are shown first. The results show that the model cannot be implemented using double-precision computer arithmetic for SNR values larger than roughly 23 dB. This stems from the summation seen in Eqs. (19), (20) and (21). In these equations, as the SNR increases, the peak value of the summation occurs at higher values of m . This is shown in figure 1a, which shows the individual terms (indexed by m) before they are used to calculate the scaling value (K) as a function of m for several SNRs. This is important because factorials for integers greater than 170 cannot be expressed as traditional double-precision floating point numbers. The peak value of the summed terms occurs at nearly the identical location for Eqs. (19), (20) and (21). This occurs because the scaling is predominantly interplay between C_m shown in Eq. (18) and $m!$ This is shown graphically in figure 2. (In the future, this behavior may be exploitable to make approximations that are more easily implemented using double-precision computer arithmetic.)

The model was derived based on a simple Gaussian-weighted pulse and unnormalized cross-correlation. It is not obvious how this translates to kernel size. It is shown empirically in figure 3 that the model roughly corresponds to a kernel size of 2, which is what was used for the rest of the simulations.

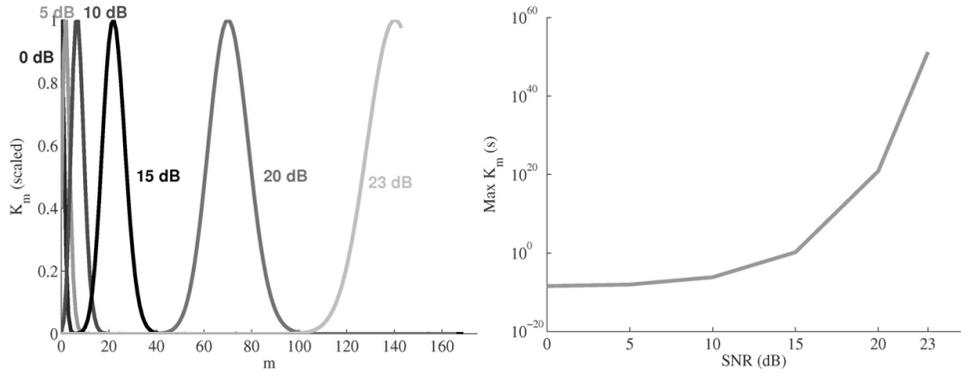


FIG 1 Plots demonstrating limitation of implementing the analytic model for predicting performance of biased time-delay estimation. Figure on the left shows the normalized components of the scaling term K as function of m before they are summed. K is the 0^{th} order moment shown in Eq. (19) and m indexes the terms in the Taylor series. The figure shows that the curve as a function of m broadens and the peak shifts to higher m as the SNR increases. The plot on the right shows the peak value used to scale the plots in the right figure. The peak values become extremely large as the center of the function moves towards higher values of m . The figures show why the model breaks down for SNRs larger than 22 dB when the model is implemented using double-precision arithmetic.

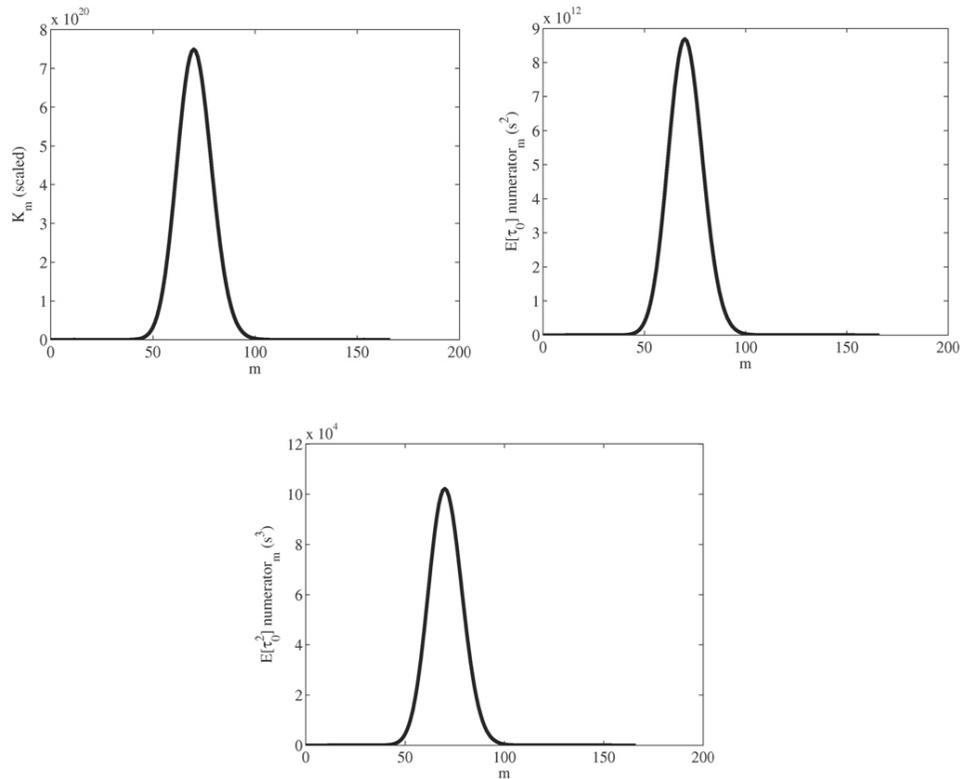


FIG. 2 Plots showing the presumed values of the scaling term, K and the numerators of the first two noncentral moments before they are scaled by K . This figure demonstrates relative orders of magnitude and also shows that the shape and position of the three plots are largely the same. The example shown is for an SNR of 20 dB, a bias of 1% of the period and a prior bandwidth that is 1% of the signal bandwidth.

Modeled and simulated results are shown in figure 4. The results in this figure are plotted as a function of the prior mean and variance. The mean and variance are both displayed rela-

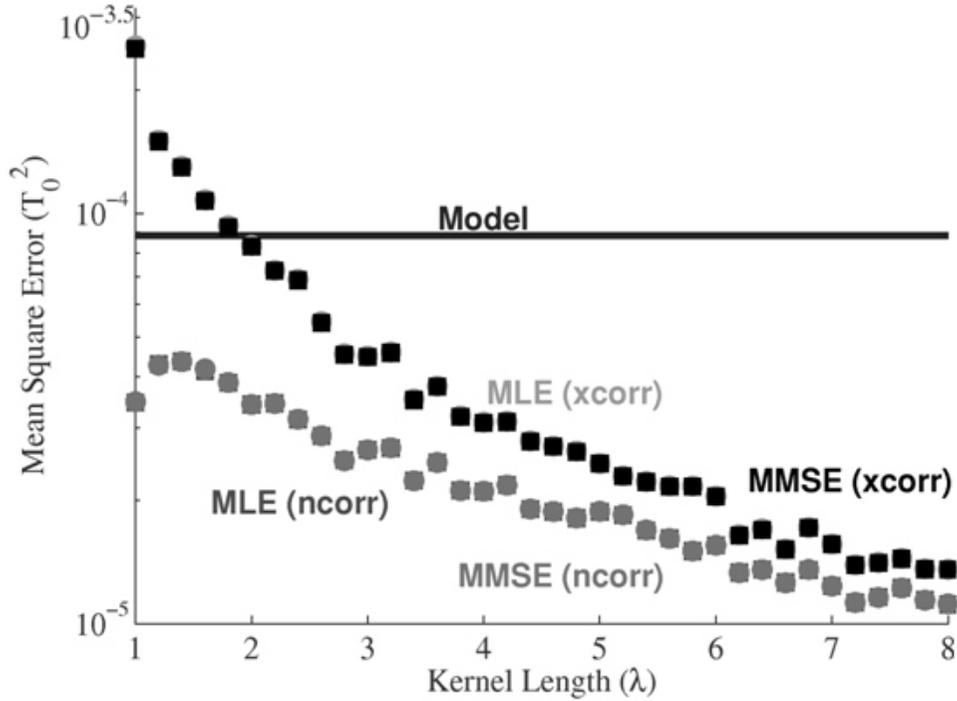


FIG. 3 In effect, the model is derived without any dependency on kernel length, which is a significant determinant in performance. In order to determine an appropriate kernel length for the simulations, a number of kernel lengths were compared to the performance of the model for a case with a broad prior that does not influence the final estimate. (The bandwidth of the prior was 10 times wider than the bandwidth of the pulse.) The figure shows where unnormalized cross-correlation-based methods are similar to the performance indicated by the model. As a point of reference, the performance of normalized cross-correlation is shown and the MMSE and the maximum likelihood estimate (MLE) are also shown. For the set of parameters chosen, the MMSE and the MLE yield indistinguishable results.

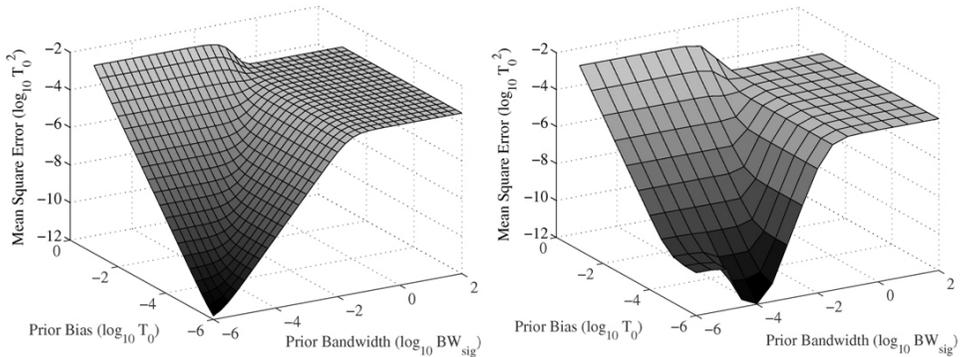


FIG. 4 Comparison of modeled and simulated MSE. The modeled MSE is shown on the left and the MSE resulting from the simulations is shown on the right. The MSE results are plotted as a function of the prior probability's bandwidth relative to the bandwidth of the signal and the bias is plotted relative to the center wavelength. The results are shown for the full range of the evaluated parameter space.

tive to the characteristics of the imaging pulse. To this end, the prior mean is expressed as a bias relative to the true displacement as a percentage of the period of the center frequency and the prior variance is plotted as a fraction of the bandwidth of the ultrasound pulse. The

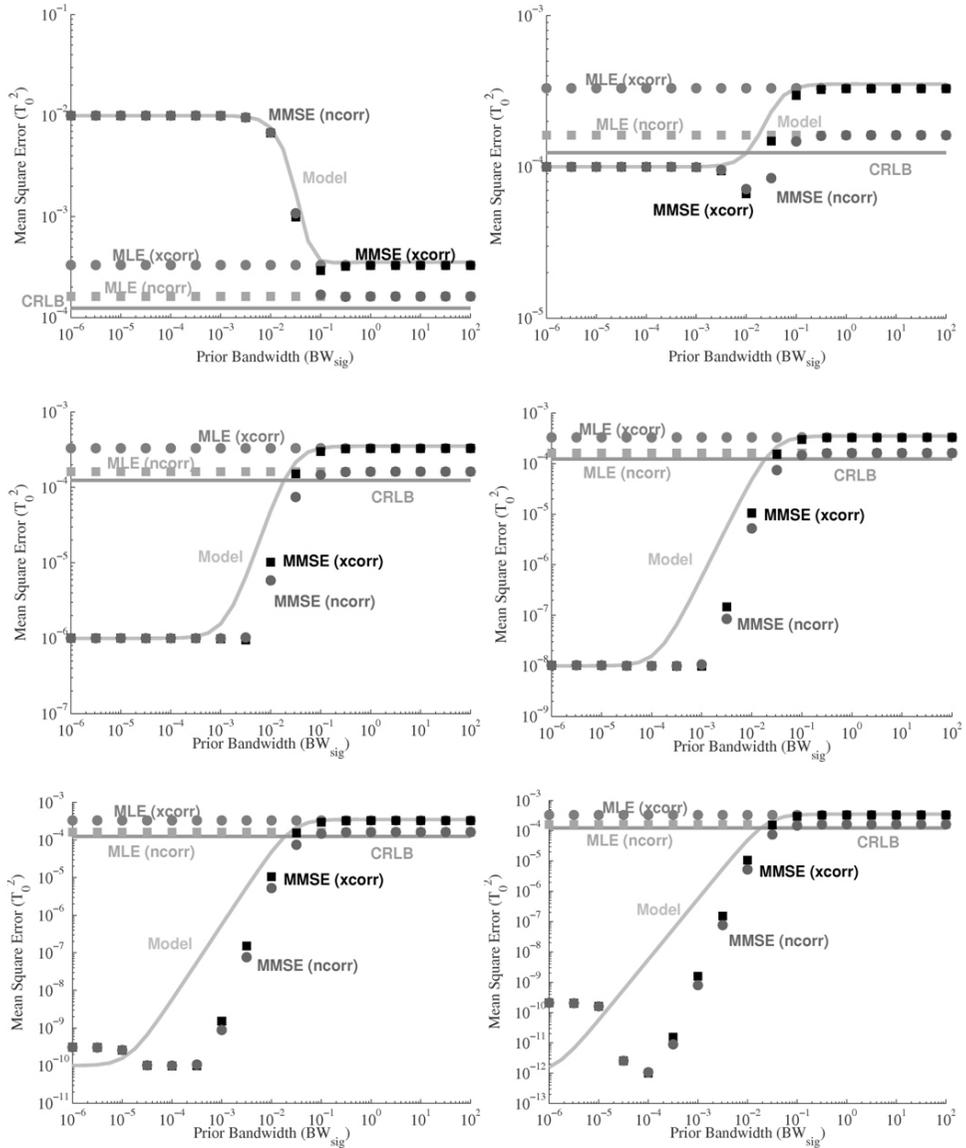


FIG. 5 Comparison of modeled and simulated MSE as a function of prior bandwidth relative to signal bandwidth. Plots are shown for six different biases between 0.000001 and 0.1 in units of T_0 . The plots more adequately highlight the differences between the model and simulated results. The differences between the two sets of results become more severe as the bias gets smaller.

results shown in the figure show that the model and the simulations are generally in agreement but are hard to compare directly when displayed as side-by-side surfaces. In order to facilitate more exact comparisons, the model and simulations are plotted as a single function of the prior's bandwidth and also the prior's bias. These plots also include the Cramer-Rao lower bound, which facilitates comparison in order to determine the domain where the new estimator is better than unbiased estimators.

The relative performance between the model and the simulations is more easily seen in figures 5 and 6. These two figures show that there is an approximate agreement between the model and simulations and this agreement is best when the bandwidth of the prior is only one or two orders of magnitude smaller than the bandwidth of the signal. While it is unfortunate

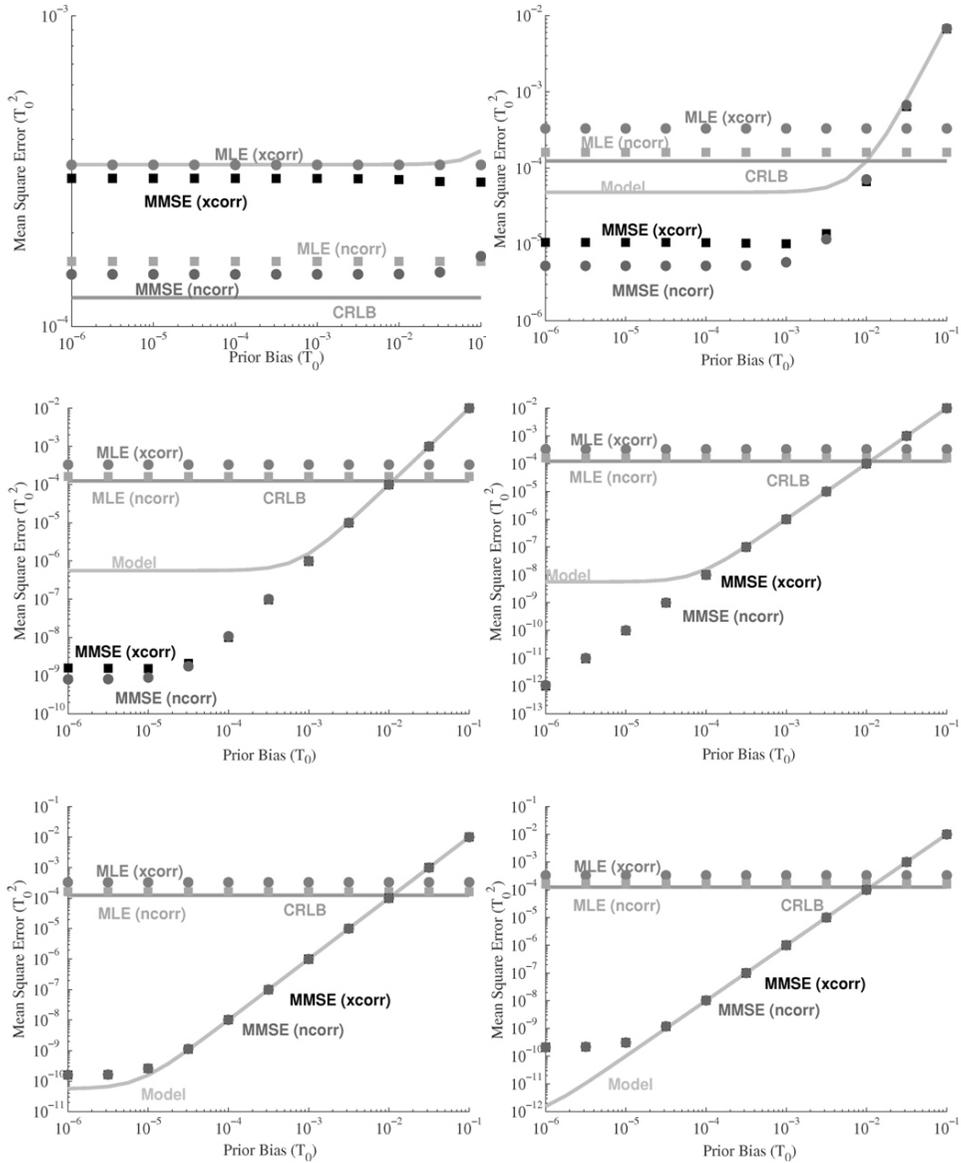


FIG. 6 Comparison of modeled and simulated MSE as a function of prior bias relative to signal wavelength. Plots are shown for six different prior bandwidths relative to signal bandwidth. The results show that for a relevant range of the parameter space investigated, the model closely predicts the simulation performance.

that the model does not match the simulations better in the ultrawide bandwidth regime, these regions represent a place where there is so much confidence in the prior information that the data has almost no influence on the final estimate. This is not expected to be a generally-useful domain of operation for Bayesian-style estimators in ultrasound TDE.

In order to show the effect of sampling frequency on the estimates from simulated data, several plots are shown for various sampling frequencies. These plots are shown in figure 7. The plots generally show that the sampling frequency does not affect the slope of the simulation in the transition region between bias-limited behavior and bandwidth- (variance) limited behavior. Additionally, the sampling frequency does not affect the final settling level for the bias limit as a function of prior bandwidth but it does affect the maximum prior band-

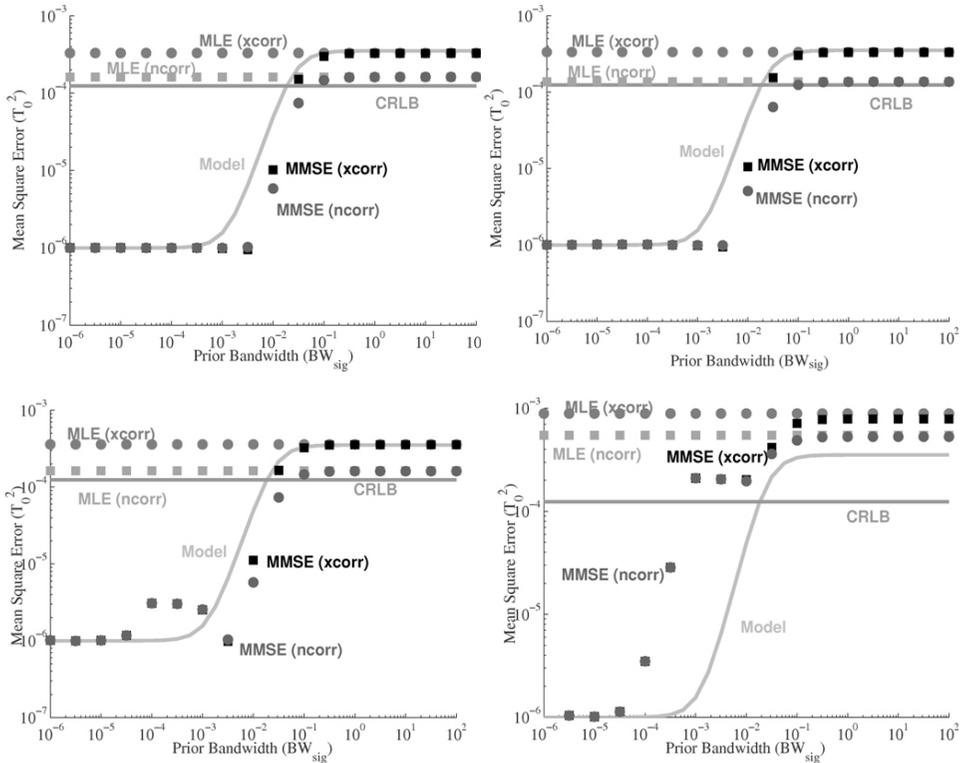


FIG. 7 Comparison of modeled and simulated MSE as a function of the prior bandwidth relative to the signal bandwidth for different sampling frequencies and a bias of $0.001T_0$. Plots are shown for four different sampling frequencies.

width where the estimates settle into the bias limit. This behavior suggests that the expected width of the prior information could be an important consideration when determining an appropriate sampling frequency for biased-motion estimation. These results will probably change if a subsample estimator were introduced.

IV. DISCUSSION

The utility of the results just presented is two-fold. First, a model was developed that could be used predictively to assess whether an estimate is bounded by the CRLB or surpasses the CRLB. Second, the necessary prior information to make estimates that surpass the CRLB and have improved quality was assessed.

The developed model does not agree exactly with simulation but it generally shows good agreement within the most relevant region of the tested parameter space. The model and the simulations diverge as the prior gets narrower by orders of magnitude relative to the signal's bandwidth. The primary difference between simulation and model is that the simulations show faster convergence towards the 'bias' noise floor as a function of prior bandwidth than the model results. There are two sources that seem to most credibly account for the difference between the simulation and modeled results. First, the model derivation makes several assumptions, specifically the introduction of a series expansion. The series expansion could be a source of error when the model is actually implemented because the bulk of the expansion terms that contribute to the final bias and variance are not near the beginning of the se-

ries, as seen in figure 1a. This may expose the model to numerical errors since the numerator and denominator at large m 's both end up being extremely large values, which combine to form very small numbers. Additionally, in general, the model and simulations match well but the likelihood function in Eq. (4) is for the case of white Gaussian noise. The simulations were implemented with the more realistic case of band-limited Gaussian noise. The likelihood function for white Gaussian noise is known to produce conservative results when used on signals with correlated noise,²² which is consistent with the presented results. The lack of better agreement between the simulated and modeled data is unfortunate but it is almost a moot point since the disagreement predominantly occurs in regions where the prior is a thousand to a million times narrower than the signal's pulse width. In this region, the data effectively become meaningless relative to the dominance of the prior information. Practically, this region of disagreement matters little for predicting whether estimates are CRLB limited or can be expected to be better than the CRLB. This should work since the results indicate that prior bandwidths that allow for improved estimator performance are independent of the bias. So, for a given set of pulse characteristics, it should be possible to predict the minimum prior bandwidth to produce estimates that exceed the CRLB.

The ability to predict whether the prior's bandwidth is sufficiently broad relative to the signal bandwidth to produce displacement estimates surpassing the CRLB is only useful if the bias is also sufficiently small. The qualification for 'sufficiently small' bias can be observed from the results. The results show that biases that are less than about $0.01T_0$ will result in displacement estimates that are better than the CRLB. In order to put this in perspective, the CRLB as a standard deviation is also about $0.01T_0$. This is significant because it indicates that there is enough information contained within a minimum variance unbiased estimate to appropriately influence a prior distribution to allow for an estimate that surpasses the Cramer-Rao lower bound about 68% of the time (assuming normal statistics). Nearly all of the estimates (99.7%) obtained using an unbiased minimum variance estimator are within $0.03T_0$. Based on the results, even the worst of the minimum variance estimates, if used as a mean for a prior distribution, should not be expected to have performance that is noticeably different than the CRLB.

Additionally, in order to be clear, a lower-bound on estimator performance for a Bayesian style mean posterior estimator has not been derived. These types of lower bounds do exist in theory but they can only be implemented for very specific practical applications,²³ which do not apply to the specific case presented here. However, an aggressive lower-bound for biased ultrasonic time-delay estimation using a different signal model or prior should not be ruled out.

Finally, it is appropriate to reiterate that restrictive assumptions were made in the course of this derivation. The derivations final results are mostly limited by the assumption of bulk motion but they are also limited by the assumption that the prior is a Gaussian distribution. The goal of the assumptions was to maintain analytic tractability so that the effect of prior information on the quality of displacement estimates could be assessed. One likely byproduct of the assumptions is that for motion estimated in the presence of significant scatterer decorrelation (e.g., blood flow or quasistatic elastography), the analytic derivations presented here are optimistic. Addressing more complicated displacement scenarios is an ongoing task.

V. CONCLUSIONS

A predictive model has been derived for Bayesian time-delay estimation. It was shown that biases with sufficient additional information to improve the quality over the CRLB are

obtainable from CRLB limited estimates. These results provide further support to previously-presented results for the usefulness and feasibility of biased time-delay estimators for clinical ultrasound.

ACKNOWLEDGMENTS

The authors thank Ned Danieley for computer support, Marko Jakovljevic for useful conversations and Albert Chang for formatting the final equations. This work was supported by NIH grants R37HL096023 and T32EB001040.

REFERENCES

1. Embree PM. *The Accurate Ultrasonic Measurement of Volume Flow of Blood by Time-Domain Correlation*, (Ph.D. thesis, University of Illinois, Urbana, IL, 1985).
2. Walker W. A fundamental limit on the performance of correlation based phase correction and flow estimation techniques, in *Adaptive Ultrasonic Imaging Performance for Near-Field Aberrating Layers*, pp. 18-46 (Ph.D. thesis, Duke University, Durham, 1995).
3. Walker W. A fundamental limit on delay estimation using partially correlated speckle signals, in *Adaptive Ultrasonic Imaging Performance for Near-Field Aberrating Layers*, pp. 47-68 (Ph.D. thesis, Duke University, Durham, 1995).
4. Pinton GF, Dahl JJ, Trahey GE. Rapid tracking of small displacements with ultrasound, *IEEE Trans Ultrason Ferroelec Freq Contr* 53, 1103-1117 (2006).
5. Byram BC. Biased ultrasonic time-delay estimation, in *Chronic Myocardial Infarct Visualization Using 3D Ultrasound*, pp. 47-135 (Ph.D. thesis, Duke University, Durham, 2011).
6. Kay SM. The Bayesian philosophy, in *Fundamentals of Statistical Signal Processing: Estimation Theory*, pp. 309-340 (Prentice-Hall, Inc., Upper Saddle River, 1993).
7. Zahiri-Azar R, Salcudean S. Motion estimation in ultrasound images using time domain cross correlation with prior estimates, *IEEE Trans. Biomed Engin* 53, 1990-2000 (2006).
8. Pellot-Barakat C, Frouin F, Insana M, Herment A. Ultrasound elastography based on multiscale estimations of regularized displacement fields, *IEEE Trans Med Imaging* 23, 153-163 (2004).
9. Tarantola A. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation* (Elsevier, Amsterdam, 1987).
10. Kay SM. Maximum likelihood estimation, in *Fundamentals of Statistical Signal Processing: Estimation Theory*, pp. 157-218 (Prentice-Hall, Inc., Upper Saddle River, 1993).
11. Bar-Shalom Y, Fortmann TE. Multitarget tracking using joint probabilistic data association, in *Statistical Signal Processing*, pp. 353-363, (Marcel Dekker, New York, 1984).
12. van Trees HL. Parameter estimation: slowly fluctuating point targets, in *Detection, Estimation and Modulation Theory, Part III* (Wiley, New York, 1971).
13. Urkowitz H. Unknown parameters and composite hypotheses, in *Signal Theory and Random Processes*, pp. 583-626 (Artech House, Dedham, 1983).
14. Walker W. Correction factor for a noisy reference signal, in *Adaptive Ultrasonic Imaging Performance for Near-Field Aberrating Layers*, pp. 123-124 (Ph.D. thesis, Duke University, Durham, 1995).
15. Jaynes ET. The central, Gaussian or normal distribution, in *Probability Theory: The Logic of Science (Vol 1)*, pp. 198-242 (Cambridge University Press, 2003).
16. Jaynes ET. Discrete prior probabilities: the entropy principle, in *Probability Theory: The Logic of Science (Vol 1)*, pp. 343-396 (Cambridge University Press, 2003).
17. Chaturvedi P, Insana MF, Hall TJ. 2-D companding for noise reduction in strain imaging, *IEEE Trans Ultrason Ferroelect Freq Contr* 45, 179-191 (1998).

18. Konofagou E, Ophir J. A new elastographic method for estimation and imaging of lateral displacements, lateral strains, corrected axial strains and Poisson's ratios in tissues, *Ultrasound Med Biol* 24, 1183-1199 (1998).
19. Byram BC. Supporting identities and substitutions, in *Chronic Myocardial Infarct Visualization Using 3D Ultrasound*, pp. 191-192 (Ph.D. thesis, Duke University, Durham, 2011).
20. Walker W, Trahey G. A fundamental limit on the performance of correlation-based phase correction and flow estimation techniques, *IEEE Trans Ultrason Ferroelect Freq Contr* 41, 644-654 (1994).
21. Palmeri ML, McAleavey SA, Trahey GE, Nightingale KR. Ultrasonic tracking of acoustic radiation force-induced displacements in homogeneous media, *IEEE Trans Ultrason Ferroelect Freq Contr* 53, 1300-1313 (2006).
22. Bretthorst GL. Single stationary sinusoid plus noise, in *Bayesian Spectrum Analysis and Parameter Estimation*, pp. 13-30 (Springer-Verlag, Berlin, 1988).
23. Weinstein E and Weiss AJ. A general class of lower bounds in parameter estimation, *IEEE Trans Info Theory* 34, 338-342 (1988).