Conference on Information Sciences and Systems, Princeton University, March 22-24, 2004

Evaluating Fisher Information From Data for Task-Driven Data Compression

Mark L. Fowler and Mo Chen

***Abstract*—Sensor networks and other multi-sensor systems collect data upon which estimations are based. How a particular sensor's data is to be used within the sensor system depends on the quality of that sensor's data relative to the other sensor's data. Because Fisher information (FI) is a natural way to assess the quality of a sensor's data, it is desirable to be able to numerically compute the FI for a collected set of data for a specific estimation task. We explore this issue for the specific scenario of FI-driven data compression for a sensor system tasked with estimating an RF emitter's location. The data compression scheme uses sub-band coding and therefore for FI-driven data compression it is important to numerically assess the FI of each filter bank output sample. However, as we demonstrate, an orthogonal filter bank well-suited to the compression task seems ill-suited to the FI-assessment task. Alternatively we demonstrate that the STFT is well-suited to FI-assessment but – as is well known – is ill-suited to the compression task. This leads to a hybrid structure for the sub-band coding scheme that uses the STFT for time-frequency domain assessment of the FI but retains the orthogonal filter bank for the compression processing.**

Index Terms— Fisher Information, Data Compression, Filter Banks, TDOA/FDOA Emitter Location

I. Introduction

Sensor networks and other sensor systems collect data upon which estimations are based. How a particular sensor's data is to be used within the sensor system depends on the quality of that sensor's data relative to the other sensor's data. Because Fisher information (FI) is a natural way to assess the quality of a sensor's data, the FI can be used to determine: (i) how to compress to optimally share the data, (ii) how to choose a subset of sensors, (iii) how to choose the sensor trajectory, etc. Thus, in all of these cases it is desirable to be able to numerically compute the FI for a specific estimation task from the collected data.

We explore this issue for the specific scenario of FI-driven data compression for a sensor system tasked with estimating an RF emitter's location via TDOA/FDOA methods [1],[2]; "TDOA" = Time-Difference-of-Arrival and "FDOA" = Frequency-Difference-of-Arrival The data compression scheme uses sub-band coding and therefore for FI-driven data compression it is important to numerically assess the contribution to the FI from each filter bank output sample. However, as we demonstrate, an orthogonal filter bank well-suited to the compression task seems ill-suited to the FI-assessment task. Alternatively we demonstrate that the STFT is well-suited to FI-assessment but – as is well known – is ill-suited to the compression task. This leads to a hybrid structure for the sub-band coding scheme that uses the STFT for time-frequency domain assessment of the FI but retains the orthogonal filter bank for the compression processing.

This work was supported in part by the Air Force Summer Faculty Fellowship Program.

M. Chen and M. L. Fowler are with the Department of Electrical and Computer Engineering, SUNY Binghamton, Binghamton, NY 13902 USA (e-mail: mfowler@binghamton.edu).

II. Problem Set-Up

Let noisy measurement vector **x** be the set of data collected at a sensor to support the estimation of the parameters θ_i, $i = 1, 2, \ldots, p$. The data vector **x** can be represented as

$$\mathbf{x} = \mathbf{s}(\boldsymbol{\theta}) + \mathbf{n}\ , \tag{1}$$

where vector **s(θ)** holds samples complex-valued $s[n;\theta]$ for the unknown signal that depends on unknown parameter vector θ to be estimated, and **n** is a vector of complex white Gaussian measurement noise values with variance σ^2. The TDOA/FDOA parameters can be estimated from signal data collected by pairs of sensors [2] and the TDOA/FDOA estimates are then used to estimate the location of the emitter [1].

For simplicity consider estimating the 2-D location of an emitter; let the location be $\mathbf{x} = [x_e, y_e]^T$ and consider that there are K pairs of sensors available. Within each pair one sensor shares its intercepted data with the other to enable estimation of the TDOA/FDOA for the pair [2]. Thus we have the K TDOA/FDOA measurements

$$\boldsymbol{\theta} = \begin{bmatrix} \tau_i \\ \\ v_i \end{bmatrix} = \begin{bmatrix} f_{TDOA,i}(\mathbf{x}) \\ \\ f_{FDOA,i}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} w_{\tau_i} \\ \\ w_{v_i} \end{bmatrix}, \quad i = 1,2,\ldots,K\ .$$

where the function $f_{TDOA,i}(\mathbf{x})$ maps the location into the TDOA value for sensor pair i, the function $f_{FDOA,i}(\mathbf{x})$ maps the location into the FDOA value for sensor pair i, and w_{τ_i} and w_{v_i} are the random TDOA and FDOA measurement errors at the i^{th} pair of sensors, respectively. Because the TDOA/FDOA estimates are obtained using the maximum likelihood (ML) estimator of cross correlation [2], the asymptotic properties of ML estimators says that $[w_{\tau_i}, w_{\nu_i}]$ follows the Gaussian distribution $\mathcal{N}(0, \mathbf{J}_i^{-1})$, where $\mathbf{J}_i$ is the 2×2 Fisher information matrix for the i^{th} TDOA/FDOA pair. The J_{11} element of the TDOA/FDOA FIM is the Fisher information relative to the TDOA estimate, the J_{22} element of the TDOA/FDOA FIM is the Fisher information relative to the FDOA estimate, and the $J_{12} = J_{21}$ elements are the cross-FI between the two estimates. This matrix captures the quality of the data in (1) relative to the task of estimating TDOA/FDOA. If the noise at the two sensors in the i^{th} sensor pair is independent then $\mathbf{J}_i = \mathbf{J}_{i1} + \mathbf{J}_{i2}$, where $\mathbf{J}_{i1}$ and $\mathbf{J}_{i2}$ are Fisher information matrices independently computed from the two sensors in the pair. If $\mathbf{s}_{ij}(\boldsymbol{\theta})$ is the signal vector from the j^{th} sensor in the i^{th} pair, then

$$\mathbf{J}_{ij} = \frac{2}{\sigma^2}\begin{bmatrix} \dfrac{\partial \mathbf{s}_{ij}^H}{\partial \theta_1}\dfrac{\partial \mathbf{s}_{ij}}{\partial \theta_1} & \text{Re}\left\{\dfrac{\partial \mathbf{s}_{ij}^H}{\partial \theta_1}\dfrac{\partial \mathbf{s}_{ij}}{\partial \theta_2}\right\} \\ \text{Re}\left\{\dfrac{\partial \mathbf{s}_{ij}^H}{\partial \theta_2}\dfrac{\partial \mathbf{s}_{ij}}{\partial \theta_1}\right\} & \dfrac{\partial \mathbf{s}_{ij}^H}{\partial \theta_1}\dfrac{\partial \mathbf{s}_{ij}}{\partial \theta_1} \end{bmatrix}. \quad (2)$$

Letting

$$\mathbf{G}_i = \begin{bmatrix} \dfrac{\partial f_{TDOA,i}}{\partial x_e} & \dfrac{\partial f_{TDOA,i}}{\partial y_e} \\ \dfrac{\partial f_{FDOA,i}}{\partial x_e} & \dfrac{\partial f_{FDOA,i}}{\partial y_e} \end{bmatrix},$$

which is the i^{th} pair's Jacobian matrix [1], the impact of data quality on the location estimate quality can be characterized through

$$\mathbf{J}_{geo} = \sum_{i=1}^{K} \mathbf{G}_i^{\mathbf{T}} \mathbf{J}_i \mathbf{G}_i \ . \quad (3)$$

As we have shown in [5], a viable approach to optimizing data compression with respect to geo-location is to compress to maximize the trace of $\mathbf{J}_{\text{geo}}$. However, the trace of FIM $\mathbf{J}_{\text{geo}}$ depends on the off-diagonal elements of the TDOA/FDOA FIM as the following shows. For notational purposes let $G_{i,kl}$ be the *kl* element in matrix $\mathbf{G}_i$ and similarly let $J_{i,kl}$ be the *kl* element in matrix $\mathbf{J}_i$. Then

$$\begin{aligned}\text{Tr}\{\mathbf{J}_{geo}\} = \sum_{i=1}^{K} & (G_{i,11}^2 + G_{i,12}^2)J_{i,11} + (G_{i,21}^2 + G_{i,22}^2)J_{i,22} \\ & + 2(G_{i,11}G_{i,21} + G_{i,12}G_{i,22})J_{i,12}\end{aligned}, \quad (4)$$

where the first two terms in the sum depend on the diagonal elements of the TDOA/FDOA FIMs but the last term depends on the off-diagonal elements. Thus, the measure defined in (4) captures how (i) data compression, (ii) sensor selection, (iii) sensor trajectory, etc. all combine to impact the performance of the location estimate – thus, optimization should be done relative to (4).

III. Numerically Evaluating the FIM From Data

The Fisher information measures how much information is available from the observation $\mathbf{x}$ relevant to estimation of parameters θ_i, $i = 1, 2, \ldots, p$. The result given in (2) requires knowledge of the signal and analytical results for the derivatives. However, in [3] we explored using data-evaluated versions of the Fisher information for single parameter cases. For example, in TDOA-only location, the contribution to the data-evaluated FI for TDOA estimation from the n^{th} DFT coefficient $X[n]$ of the observed data a single sensor is

$$J_{TDOA} = \sum_{k=-N/2}^{N/2-1} \frac{2\pi^2 k^2 |X[k]|^2}{\sigma^2}, \quad (5)$$

where σ^2 is the known variance of the sensor noise. Each quantity in (5) can then be used as a measure of the quality of that DFT coefficient for estimating TDOA. The measure in (5) can be used to implement a data-adaptive compression scheme that minimizes the impact on TDOA accuracy for a given data rate constraint [3]. Similarly we have explored using

$$J_{FDOA} = \sum_{n=-N/2}^{N/2-1} \frac{2\pi^2 n^2 |x[n]|^2}{\sigma^2} \quad (6)$$

as the data-evaluated FI for optimizing compression with respect to FDOA estimation accuracy.

Extending this idea, in [4] we showed how compression trade-offs for *joint* TDOA/FDOA estimation can be accomplished by optimizing the compression with respect to $\alpha J_{TDOA} + (1 - \alpha) J_{FDOA}$, where α controls the relative importance between TDOA and FDOA. As argued in [4], this motivates using a time-frequency decomposition such as a filter bank to gain joint access to the time and frequency characteristics of the signal. However, in this case it becomes more challenging to numerically compute the FIMs diagonal elements because such time-frequency representations lack the mathematical characteristics with respect to derivatives that are used to derive results like (5) and (6). Thus, in principle it is possible to write a mathematical equation for the filter bank samples and then consider the derivatives with respect to parameters, as needed for the FIM; but this approach is intractable.

Nonetheless it was demonstrated that it is possible to get good results using standard ON wavelet packet filter banks, for which each channel in the filter bank has both positive and negative frequency content (so-called "two-sided" filter banks; "one-sided" filter banks have separate channels for positive and negative frequencies). Such filter banks are satisfactory for computing (5) and (6) in terms of time-frequency components because of the inherent negative-positive frequency symmetry due to the k^2 term. For example, let $S_1[n,k]$ be the time-frequency (T-F) representation of the signal $s_1[n]$ from such a filter bank. Then instead of (5) and (6) we can use

$$J_{TDOA} = \frac{1}{\sigma_1^2} \sum_{n=-N/2}^{N/2} \sum_{k=0}^{K/2} f_k^2 |S_1[n,k]|^2 \tag{7}$$

and

$$J_{22} = \frac{1}{\sigma_1^2} \sum_{n=-N/2}^{N/2} \sum_{k=0}^{K/2} t_n^2 |S_1[n,k]|^2 \tag{8}$$

where k is the index of two-sided channels.

However, as shown in Section II, for emitter location processing we also need the cross-term elements of the TDOA/FDOA FI matrix (FIM) and we wish to assess the FIM-quality of the output samples of the filter bank. As above, finding an analytical result for the FIM elements is intractable. Also, as we will see below, two-sided channels in the filter bank are not consistent with computing the cross-terms in the FIM.

The derivation of (5) makes use of the property that the Fourier transform of a time derivative gives the frequency variable times the Fourier transform; since there are two such terms multiplied (one conjugated) we get the squared-frequency dependence shown in (5). That squared dependence is what allows use of two-sided filter banks when computing the diagonal elements of the FIM. However, following the same mathematical development would lead to the cross-term in the FIM looking like

$$J_{12} \sim \text{Re}\left\{ \sum_{n=-N/2}^{N/2-1} \sum_{k=-N/2}^{N/2-1} knx[n]X^*[k], \right\}. \tag{9}$$

From this we see that there is no frequency symmetry in (9) and therefore two-sided filter banks are unsuitable for evaluating this measure. Furthermore, recognizing $x[n]X^*[k]$ as one particular time-frequency representation we resort to the following brazen conjecture, verified through the testing described below.

Brazen Conjecture: Let $S_1[n,k]$ be a time-frequency (T-F) representation of the signal $s_1[n]$; e.g., the short-time Fourier transform or a one-sided filter bank representation. It is important that the frequency range of the T-F representation is $[-\pi,\pi]$ rad/sample, meaning that it explicitly shows separate channels for positive and negative frequencies. Then the FIM can be computed (up to a multiplicative factor) using:

$$J_{11} = \frac{1}{\sigma_1^2} \sum_{n=-N/2}^{N/2} \sum_{k=-K/2}^{K/2} f_k^2 |S_1[n,k]|^2 , \tag{10}$$

$$J_{22} = \frac{1}{\sigma_1^2} \sum_{n=-N/2}^{N/2} \sum_{k=-K/2}^{K/2} t_n^2 |S_1[n,k]|^2 , \tag{11}$$

$$J_{12} = J_{21} = \frac{1}{\sigma_1^2} \sum_{n=-N/2}^{N/2} \sum_{k=-K/2}^{K/2} t_n f_k \, |S_1[n,k]|^2 , \tag{12}$$

where f_k and t_n are the frequency and time centers of the cells of the time-frequency representation.

Our test results (see below) show that the choice of the T-F representation has a significant impact on the accuracy of this conjecture. To test this issue we needed a signal that would impinge on each channel of the filter but for which it would still be possible to *analytically* compute the FIM elements. Then we could apply the filter bank and equations (10) – (12) to numerically compute the FIM elements and then compare the results to the analytically derived theoretical results for this signal.

A good choice for this signal is a linear chirp signal. We now derive the FIM for such a signal, and for simplicity assume unit variance noise. Let the received C-T signal be $r(t) = s(t+\tau)e^{j\nu t} + w(t)$ so that after sampling at intervals of T we have

$$r[n] = \underbrace{s(nT+\tau)e^{j\nu nT}}_{\to \mathbf{s}} + w(nT) \tag{13}$$

where the arrow notation means that samples of the indicated signal go into a signal vector called **s**. Then the partial derivatives of this signal vector w.r.t. the TDOA/FDOA parameters are given by

$$\frac{\partial \mathbf{s}}{\partial \tau} = \begin{bmatrix} \vdots \\ e^{j\nu nT} \dfrac{\partial s(nT+\tau)}{\partial \tau} \\ \vdots \end{bmatrix} \tag{14}$$

and

$$\frac{\partial \mathbf{s}}{\partial \nu} = \begin{bmatrix} \vdots \\ jnTe^{j\nu nT} s(nT+\tau) \\ \vdots \end{bmatrix} \tag{15}$$

The elements of the FIM are given by

$$J_{11} = 2\left(\frac{\partial \mathbf{s}}{\partial \tau}\right)^H \left(\frac{\partial \mathbf{s}}{\partial \tau}\right) = 2\sum_n \left|\left(\dot{s}(nT)\right)\right|^2 \tag{16}$$

$$J_{22} = 2\left(\frac{\partial \mathbf{s}}{\partial \nu}\right)^H \left(\frac{\partial \mathbf{s}}{\partial \nu}\right) = 2\sum_n (nT)^2 |s(nT)|^2 \tag{17}$$

$$J_{12} = 2\,\mathrm{Re}\left\{\left(\frac{\partial \mathbf{s}}{\partial \nu}\right)^H \left(\frac{\partial \mathbf{s}}{\partial \tau}\right)\right\}$$

$$= 2\,\mathrm{Re}\left\{\sum_n - jnTs^*(nT)\dot{s}(nT)\right\} \tag{18}$$

Let the signal be a complex linear chirp signal given by

$$s(t) = e^{j\phi} e^{j(\alpha/2)t^2}$$

so that we get

$$\dot{s}(nT) = j\alpha nTe^{j\phi} e^{j(\alpha/2)(nT)^2} = j\alpha nTs(nT)$$

Using this with (16) – (18) gives

$$J_{\tau\tau} = 2\sum_n \left|\left(\dot{s}(nT)\right)\right|^2 = 2\sum_n \left| j\alpha nTe^{j\phi} e^{j(\alpha/2)(nT)^2}\right|^2$$

$$= 2(\alpha T)^2 \sum_n n^2$$

$$J_{\nu\nu} = 2\sum_n (nT)^2 \left|s(nT)\right|^2$$

$$= 2T^2 \sum_n n^2$$

$$J_{\nu\tau} = 2\,\mathrm{Re}\left\{\sum_n - jnTs^*(nT)\left(j\alpha nTs(nT)\right)\right\}$$

$$= 2\alpha T^2 \sum_n n^2$$

or in matrix form

$$\mathbf{J} = \left(2T^2 \sum_n n^2\right)\begin{bmatrix} \alpha^2 & \alpha \\ \alpha & 1 \end{bmatrix} \tag{19}$$

where T is the sampling interval used to sample the chirp signal. From this result we see that J_{22} is a constant with respect to α, so we can look at the other terms relative to this term: $J_{11}/J_{22} = \alpha^2$ and $J_{12}/J_{22} = \alpha$.

The two-sided filter bank based on wavelet packets that was used in [4] was extended to a one-sided ON filter bank as follows. It can be shown that starting with a standard half-band lowpass filter for a two-sided ON filter bank (e.g., standard ON wavelet lowpass filter) and then modulating it up by π/2 rad/sample gives a $g_0[n]$ that will generate a one-sided two-channel ON filter bank by using the relationships shown in Figure 1, where the $g_i[n]$ are the two-channel prototypes for a cascaded analysis filter bank, and $h_i[n]$ are the two-channel prototypes for the cascaded synthesis filter bank.

Plots of the numerical and theoretical FIM results for the case of using such a complex orthogonal filter bank are shown in Figure 2, where it is seen that the numerical results are not very accurate: in the top plot the numerical results should be a constant (as stated by the theory), in the middle plot the numerical results should follow the theoretical curve of α^2 from (19) and in the bottom plot the numerical results should follow the theoretical curve of α from (19).

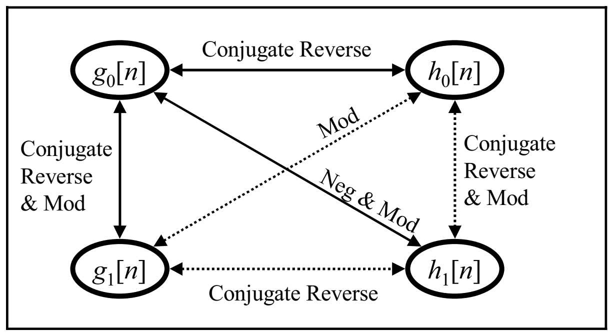

Figure 1: Relationships between the filters for complex PR filter banks.

To demonstrate that it ***is*** at least possible to evaluate the FIM from signal data, the short-time Fourier transform (STFT) was investigated as a replacement for the complex-valued orthogonal PR filter bank. The results obtained are given in Figure 3; notice that the numerical results match the theoretical results. Although the STFT enables quite accurate evaluation of the FIM, the STFT is not well-suited for compression due to the fact that it is not an orthogonal representation. Thus, we use it simply as an auxiliary parallel mechanism to allow evaluation of the FIM for bit allocation to the complex orthogonal PR filter bank, as shown in Figure 4. Each sample coming out of the filter bank represents a known time-frequency region and the FIM elements can be evaluated over this time-frequency region using the STFT to provide the FIM evaluation for the filter bank sample.

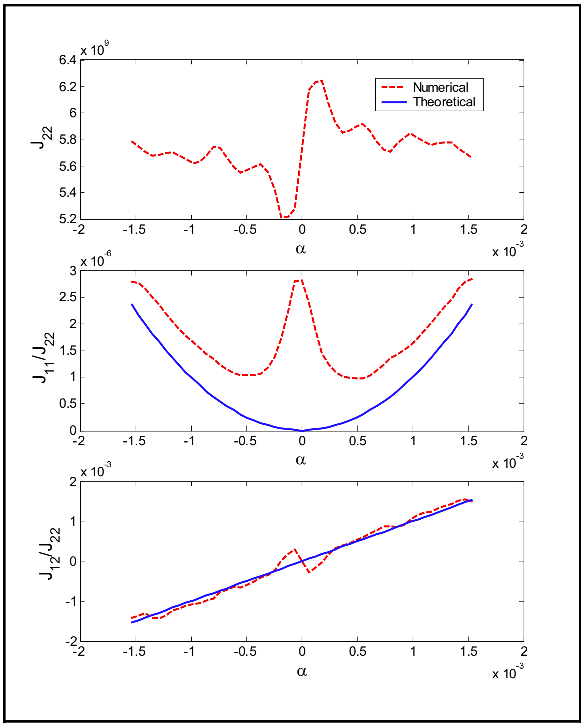

Figure 2: Quality of Numerical Evaluation of FIM via Classical Filter Bank

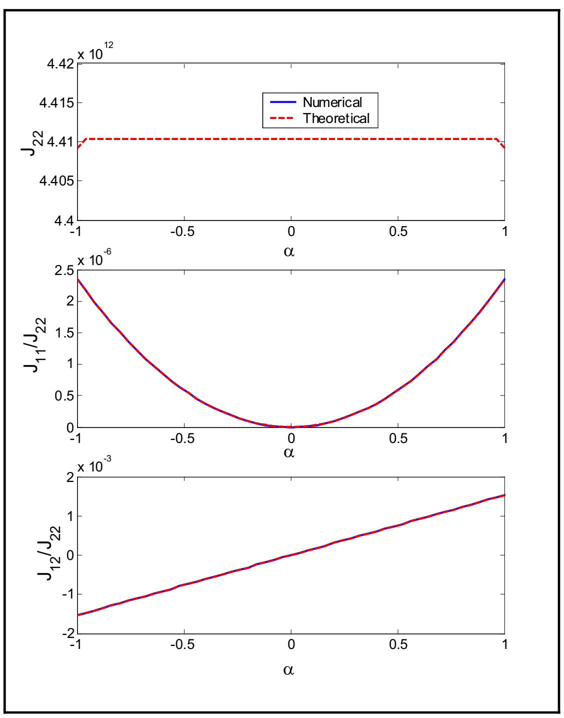

Figure 3: Quality of Numerical Evaluation of FIM via STFT

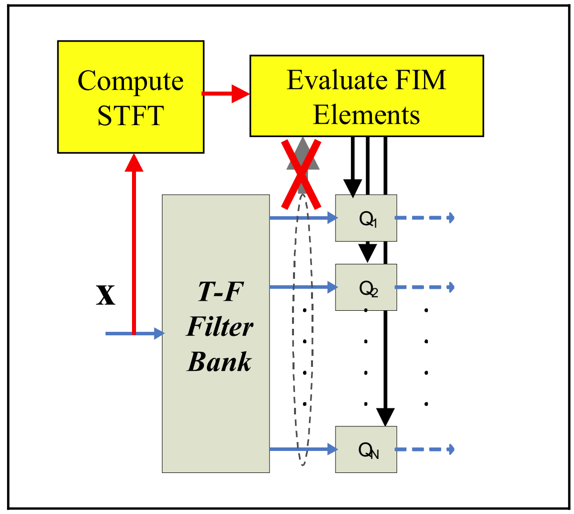

Figure 4: Our compression framework using parallel auxiliary STFT processing to evaluate the FIM elements.

IV. Conclusions

Estimation processing within sensor systems can be optimized by considering the FIM-based measure. However, evaluating this measure from collected data can be challenging. We have demonstrated how it is possible to use the STFT to numerically evaluate the FIM for the task of emitter location and have demonstrated the usefulness of this to developing a task-optimal data compression scheme. This has further ramifications in more grand schemes of optimizing with respect to sensor selection, sensor trajectory, etc.

We have demonstrated that one particular one-sided complex ON perfect reconstruction filter bank performs poorly when used to evaluate the FIM elements needed to assess the filter bank's various output samples. Some interesting questions arise. Is there a fundamental prohibition against complex ON filter banks performing well for the purpose of computing FIM elements? What are the general conditions needed on a filter bank to ensure it performs well for the purpose of computing FIM elements? These are open questions we hope to address in the future.

References

[1] D. J. Torrieri, "Statistical theory of passive location system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, March 1984, pp. 183 – 198.

[2] S. Stein, "Differential delay/Doppler ML estimation with unknown signals," *IEEE Trans. Signal Processing*, vol. 41, August 1993, pp. 2717-2719.

[3] M. L. Fowler and M. Chen, "Fisher-Information-Based Data Compression for Estimation Using Two Sensors,"

IEEE Transactions on Aerospace and Electronic Systems, July 2005, pp. 1131 – 1137.
[4] M. Chen and M. L. Fowler, "Geometry-Adaptive Data Compression For TDOA/FDOA Location," Proceedings of IEEE ICASSP 2005, Philadelphia, PA, March 2005.
[5] M. Chen and M. Fowler, "Data Compression for Simultaneous/Sequential Inference Tasks in Sensor Networks," to appear in Proceedings of IEEE ICASSP 2006, Invited Paper for Special Session on Task-Driven Processing, Toulouse, France, May 2006.