

Fisher-Information-Based Data Compression for Estimation Using Two Sensors

Mark L. Fowler* and Mo Chen
State University of New York at Binghamton
Department of Electrical and Computer Engineering
P. O. Box 6000
Binghamton, NY 13902-6000

Phone: 607-777-6973
Fax: 607-777-4464
E-mail: {mfowler, mchen0}@binghamton.edu

Abstract: A transform coding data compression method is developed for the case of making estimates from data collected using two sensors. We focus on the transfer of data from one sensor to another sensor, where the shared data is then used with the local data to estimate a parameter. Particular attention is paid to the case where neither sensor alone can estimate the parameter from its data. The method uses an operational rate-distortion viewpoint together with a distortion measure based on the Fisher information of the estimation problem. Explicit means of using the transformed data to compute operational measures of the Fisher information are given. An integer optimization version of the Lagrange multiplier method is used to efficiently determine the optimal operating point of the transform compression algorithm.

The advantages of the method lie in its ability to use transform coding to effectively capture the impact of compression on estimation accuracy in a way that lends itself to efficient optimization within the operational rate-distortion viewpoint. The applicability and effectiveness of the method are demonstrated for two illustrative examples: (i) estimation of time-difference-of-arrival (TDOA), and (ii) estimation of frequency-difference-of-arrival (FDOA). In these two cases it is shown that the Fisher-information-based method outperforms the standard MSE approach.

* Corresponding Author

I. Introduction

Often a single sensor can't estimate a desired signal parameter (e.g., [1], [2]); even when it can, using data from other sensors can improve the estimate. We focus on compression of data at one sensor (S_1), which is then transmitted to a second node (S_2) where it is used with the local data to estimate a single parameter; extension to multiple inferences is on-going [3]. A key to such compression problems is to use an inference-centric distortion measure such as Fisher information [14]. We assume a signal in AWGN model where the noises at the two sensors are independent – a simplifying but nonetheless useful scenario. In particular, work on compression for time/frequency-difference of arrival (TDOA/FDOA) systems goes back over two decades with limited success [5]-[7], and a better solution, even under simplifying assumptions, is useful.

Compression methods can be developed either under a classical rate-distortion viewpoint [8] or an operational rate-distortion viewpoint [9]. The former strives to develop methods that are optimal *on average*. The latter specifies a framework and optimizes its operating point for the particular signal at hand. Because a sensor system would likely operate in differing signal environments, we focus on the operational viewpoint, which uses numerically-computed allocations of bits (see [9],[23]) rather than classical closed forms such as reverse water-filling (see [8]).

We model the sensor data as a deterministic signal with deterministic parameter in AWGN. We compress data collected at S_1 using no more than a budgeted R bits while making the estimate at S_2 with the lowest possible mean-square estimation error. We assume that the estimation processing and compression processing are not jointly designed – this is motivated by our belief that a sensor may likely be called on to provide data to other sensors or systems that are independently designed.

Compression for distributed estimation has been considered in [11] – [20], which focus either on optimizing compression for (i) estimating the source *signal* [11],[12],[17],[18] or on (ii) estimating source *parameters* [13]-[16],[19],[20]. Many focus on only scalar quantization [11] -

[13],[15],[16],[20]. Others focus on more general compression structures [14],[17],[18],[19], although, most of these ([14],[18],[19]) focus on establishing information theoretic results. The closest to our approach are [15] and [16], which use FI but limit their designs to scalar quantizers. The result in [20] also limits its focus to scalar quantizers, but takes an quite different inference-centric metric specific to the TDOA case. In contrast, our approach considers transform-based compression with FI and our results show that proper choice of transform allows better exploitation of the FI structure.

II. Algorithm Development

Let the (real or complex) data vector \mathbf{x} have a probability density function (PDF) $p(\mathbf{x};\theta)$ that is parameterized by θ , which is to be estimated. The FI then is [10]

$$J(\theta;\mathbf{x}) = E \left\{ \left[\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} \right]^2 \right\}, \quad (1)$$

where the *notational* dependence on \mathbf{x} is included to show the data set for which the FI is computed. Lossy compression changes the PDF and thus changes the FI. We seek to retain the maximum FI in the compressed $\hat{\mathbf{x}}$ while satisfying a bit budget R . When the FI depends on θ , a question arises as to how to precisely characterize this desire to maximize the FI. Thus, in general, we propose a minimax approach. Let $J(\theta;\mathbf{x})$ and $J(\theta;\hat{\mathbf{x}})$ be the FI of the original data \mathbf{x} and the compressed data $\hat{\mathbf{x}}$, respectively, and compress to R bits such that we satisfy $\min_{\hat{\mathbf{x}}} \max_{\theta} [J(\theta;\mathbf{x}) - J(\theta;\hat{\mathbf{x}})]$, where the minimization is over all $\hat{\mathbf{x}}$ that meet bit budget R .

Clearly, when the FI does not depend on θ this becomes $\max_{\hat{\mathbf{x}}} J(\hat{\mathbf{x}})$, on which we will focus.

At sensor node S_k we model the received signal vectors \mathbf{x}_k as

$$\mathbf{x}_k = \mathbf{s}_k(\theta) + \mathbf{w}_k, \quad k = 1, 2, \quad (2)$$

where $\mathbf{s}_k(\theta)$ is an unknown deterministic vector dependent on the unknown deterministic scalar θ , and \mathbf{w}_k is a zero-mean, white Gaussian noise vector with \mathbf{w}_1 and \mathbf{w}_2 independent. We assume the variance σ_1^2 of \mathbf{w}_1 is known or estimated. For transform coding we use an orthonormal (ON) basis $\{\boldsymbol{\phi}_n\}_{n=1}^N$ with χ_n as the coefficients for \mathbf{x}_1 , and ON expansion conserves FI, so $J(\boldsymbol{\chi}_1) = J(\mathbf{x}_1)$. Only those χ_n with significant contribution to the FI should be selected and quantized. Let $\Omega \subset \{1, 2, \dots, N\}$ be a set of selection indices; let $B = \{b_n | n \in \Omega\}$ be a set of bit allocations; let $\{\hat{\chi}_n | n \in \Omega\}$ be the selected/quantized coefficients using allocation B . The compressed signal is $\hat{\mathbf{x}}_1 = \sum_{n \in \Omega} \hat{\chi}_n \boldsymbol{\phi}_n$; grouping these coefficients into vector form gives

$$\begin{aligned} \hat{\boldsymbol{\chi}}_1 &= \boldsymbol{\xi}_1(\theta) + \boldsymbol{\omega}_1 + \boldsymbol{\varepsilon} \\ &= \boldsymbol{\xi}_1(\theta) + \mathbf{v} \end{aligned} \quad (3)$$

where $\boldsymbol{\xi}_1(\theta)$ holds the selected noisy signal coefficients $\xi_n(\theta)$, $\boldsymbol{\omega}_1$ holds the corresponding noise coefficients ω_n , and $\boldsymbol{\varepsilon}$ is the quantization noise vector. By independence $J(\hat{\mathbf{x}}_1, \mathbf{x}_2) = J(\hat{\mathbf{x}}_1) + J(\mathbf{x}_2)$ so we need only maximize $J(\hat{\mathbf{x}}_1)$. We use a standard model for the quantization noise: white, uniformly distributed, zero mean, and independent of the sensor-noise vector $\boldsymbol{\omega}_1$; independence is valid due to dithering by the sensor noise [21]. The variance of the sum of these two noises is the sum of their two variances. But what PDF for $\hat{\boldsymbol{\chi}}_1$ should be used in (1)? For quantization to 1 bit we use the ideas of [22] to get an explicit result; for multi-bit quantization, the derivation of the FI seems intractable [22] – we will motivate the use of a Gaussian approximation.

Consider that $\hat{\chi}_n = \hat{\chi}_n^r + j\hat{\chi}_n^i$ is an element in $\hat{\boldsymbol{\chi}}_1$ that has been quantized to one bit. Using the ideas of [22] gives that the FI of this quantized coefficient can be computed according to

$$J^{(1)}(\hat{\chi}_n) = \sum_{q=\pm 1} \left(\frac{1}{f_{\hat{\chi}_n^r}(q; \theta)} \left(\frac{\partial f_{\hat{\chi}_n^r}(q; \theta)}{\partial \theta} \right)^2 + \frac{1}{f_{\hat{\chi}_n^i}(q; \theta)} \left(\frac{\partial f_{\hat{\chi}_n^i}(q; \theta)}{\partial \theta} \right)^2 \right) \quad (4)$$

where the superscript ⁽¹⁾ indicates that this is for the 1-bit case and where

$$\begin{aligned} f_{\hat{\chi}_k^r}(q; \theta) &= \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left(-\frac{(x - q \operatorname{Re}\{\xi_k\})^2}{\sigma^2}\right) dx \\ f_{\hat{\chi}_k^i}(q; \theta) &= \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left(-\frac{(x - q \operatorname{Im}\{\xi_k\})^2}{\sigma^2}\right) dx \end{aligned} \quad (5)$$

with ξ_k being the noise-free part of the unquantized coefficient χ_k .

For the multi-bit case, if $\operatorname{var}\{\varepsilon_k\} \ll \operatorname{var}\{\omega_k\}$, \mathbf{v} in (3) is approximately Gaussian, which follows from the relationship between their characteristic functions: $C_v(f) = C_\omega(f)C_\varepsilon(f)$. Namely, if $\operatorname{var}\{\varepsilon_k\} \ll \operatorname{var}\{\omega_k\}$, the sinc function $C_\varepsilon(f)$ is approximately flat where the Gaussian $C_\omega(f)$ is significant. Numerical results given in the Appendix indicate that the approximation is good at low SNR values (even for 2 bits) but gets worse at higher SNR values; note that our approximation is best at low SNR, where “every ‘ounce’ counts”. Then results for the FI for the complex Gaussian case (see (15.52) in [10]) give

$$J^{(m)}(\hat{\chi}_n) = \frac{2 \left| \frac{\partial \xi_n(\theta)}{\partial \theta} \right|^2}{\sigma_1^2 + q_n^2}. \quad (6)$$

where ^(m) indicates “multi-bit” and q_n^2 is the variance of the n^{th} element of the quantization noise.

Establishing the exact form for q_n^2 is generally not possible and it is common to use an approximate model [9]. A widely used model is:

$$q_n^2 = C \times \operatorname{Var}(X) \times 2^{-2b_n}, \quad (7)$$

where C is a constant (often determined heuristically) that depends on the PDF of the random variable X [24]. We use small blocks of coefficients to estimate $\operatorname{Var}(X)$.

In the operational setting we evaluate (4) and (6) using the noisy coefficients in $\hat{\chi}_1$. Thus the distortion function we use on the n^{th} coefficient quantized to b_n bits is

$$\hat{J}_n = \begin{cases} 0, & \text{if } b_n = 0 \\ J^{(1)}(\hat{\chi}_n), & \text{if } b_n = 1 \\ J^{(m)}(\hat{\chi}_n), & \text{if } b_n \geq 2 \end{cases}, \quad (8)$$

and we seek a bit allocation set $B = \{b_n \geq 0 \mid n \in 1, 2, \dots, N\}$ that solves

$$\max_B \left\{ \sum_{n=1}^N \hat{J}_n \right\} \quad \text{subject to} \quad \sum_{n=1}^N b_n \leq R. \quad (9)$$

Due to their efficiency, Lagrangian optimization methods are commonly used to determine the bit allocation in operational rate-distortion methods [9], [23] with an additive objective function like that in (9). We use the algorithm developed in [23].

An advantage of our approach is that once the functional form of the FI is found, it generally provides insight into the choice of a the transform. For example, (as seen in Section III) when the parameter to be estimated is the delay between two signals, the FI depends on a specific bandwidth measure, so a frequency domain transform makes sense in that case.

We are particularly interested in problems where data *must* be shared between sensors because neither sensor can estimate the parameter by itself. We call these types of problems “dual-sensor-critical” problems. Such problems often arise in passive systems due to lack of knowledge about a transmitted signal that has been perturbed by some parameter. It is important to keep in mind that the real essence of FI is that it captures the sensitivity of the data to a *change* in the parameter; this is due to the derivative in the definition of FI in (1). Thus, in a passive sensor setting the data at one sensor may be very sensitive to the parameter (i.e., the FI is non-zero) yet you may still be unable to estimate the *value* of the parameter. But with two sets of data (each having a different parameter value) you can estimate the difference in the two parameter values. For example, the data can be *sensitive* to a change in time delay but without a reference it is impossible

to estimate the delay. What makes an active sensor problem different is that you have a known reference signal: estimation of the parameter's change from the reference signal leads to estimation of the actual value. In a passive sensor scenario you often need the data at the other sensor to play the role of the reference. Thus, to make our approach work we only need that the data at a sensor is sensitive to the parameter, and that is what the Fisher information in (8) assesses.

III. Example Applications

We choose two examples (TDOA and FDOA estimation [1]) and illustrate (i) the effectiveness of our method relative to MSE-driven methods, and (ii) how the form of the FI drives the choice of the transform. We use cross correlation to estimate TDOA and FDOA [4]. The signal we use is a complex baseband FM signal with a pseudo-random modulating signal; a sample spectrum is shown in Figure 1. For each SNR and compression ratio (CR) of interest, we evaluated the estimation error over 400 Monte Carlo runs. We heuristically chose the value of C in (7) to be $C = \pi\sqrt{3}/2$, which gives good results in our simulations. To focus on the capabilities of the transform coding we performed no entropy coding, which would likely provide further improvement in the CR with no further accuracy degradation.

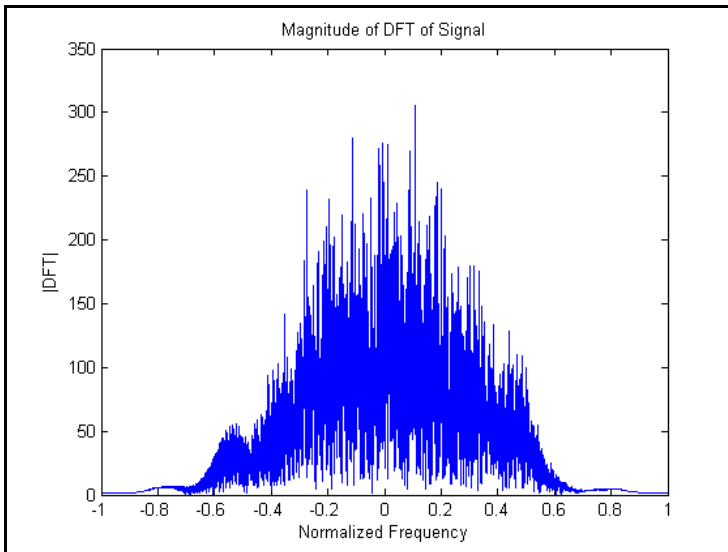


Figure 1: The spectrum of a typical FM signal used in the simulations.

A. Compression for TDOA Estimation

It is well known [4] that for TDOA estimation, the FI is proportional to $\int_{-\infty}^{\infty} f^2 |S(f)|^2 df$,

where $S(f)$ is the Fourier transform of the source signal. This view drives us to choose the DFT as our transform to allow discarding less useful frequency components. The continuous-time signal model for two passively-received complex baseband signals having an unknown TDOA of Δ is given by

$$\begin{aligned} x_1(t) &= s(t - (t_0 + \Delta/2)) + w_1(t) \\ x_2(t) &= s(t - (t_0 - \Delta/2)) + w_2(t) \end{aligned} \quad (10)$$

where t_0 is also an unknown parameter that can not be estimated, and $w_i(t)$ is complex bandlimited white Gaussian noise. In the frequency domain this model becomes

$$\begin{aligned} X_1(\omega) &= S(\omega)e^{-j\omega((t_0 + \Delta/2))} + W_1(\omega) \\ X_2(\omega) &= S(\omega)e^{-j\omega((t_0 - \Delta/2))} + W_2(\omega) \end{aligned} \quad (11)$$

Now consider samples of (10) such that the noise samples give discrete-time white noise with variances σ_i^2 and the aliasing of signal $s(t - (t_0 + \Delta/2))$ is negligible. Taking the DFT of these samples (taking care to reduce leakage errors) leads to the frequency domain model

$$X_1[k] \approx S[k] \exp[-jk \frac{2\pi}{N}(t_0 + \Delta/2)] + W_1[k] \quad k = -N/2, -N/2 + 1, \dots, N/2 - 1, \quad (12)$$

where the $S[k]$ are the DFT coefficients (for negative and positive frequencies) of the samples of signal $s(t)$ and $W_1[k]$ are the DFT coefficients of the noise.

This is clearly a dual-sensor-critical problem. Each data set is sensitive to changes in the time-of-arrival $t_0 + \Delta/2$ but neither sensor by itself can estimate $t_0 + \Delta/2$, t_0 alone, or Δ alone. In fact, even using both data sets it is impossible to estimate the nuisance parameter t_0 . Al-

though this appears to be a *two*-parameter problem that might require a 2-D FI *matrix*, there really is only *one* thing that matters: sensitivity of the first sensor's data to Δ for a fixed t_0 .

Note that because the DFT is an orthogonal, but not *orthonormal*, transform the DFT noise variance is $N\sigma_1^2$. Using (6), the FI after multi-bit quantization becomes

$$\hat{J}^{(m)}(X_1[n]) = \frac{2\pi^2 n^2 |X_1[n]|^2}{N\sigma_1^2 + q_n^2}, \quad n = -N/2, -N/2+1, \dots, N/2-1. \quad (13)$$

This shows that the DFT coefficients get quadratically weighted by frequency. For the 1-bit result application of (4) and (5) along the lines of those in [22] gives

$$\begin{aligned} \hat{J}^{(1)}(X_1[n]) = \frac{4}{\pi\sigma^2} n^2 & \left(\text{Im}^2\{X_1[n]\} \exp\left(-\frac{2}{N\sigma^2} \text{Re}^2\{X_1[n]\}\right) \frac{1}{1 - \text{erf}^2(\text{Re}\{X_1[n]\}/(\sqrt{N}\sigma))} + \right. \\ & \left. \text{Re}^2\{X_1[n]\} \exp\left(-\frac{2}{N\sigma^2} \text{Im}^2\{X_1[n]\}\right) \frac{1}{1 - \text{erf}^2(\text{Im}\{X_1[n]\}/(\sqrt{N}\sigma))} \right) \end{aligned} \quad (14)$$

Using (13) and (14) in (9) with the Lagrange optimization method produces our results. Figure 2 and Figure 3 show (for CRs of 4:1 and 8:1, respectively) the TDOA accuracy performance of our method (labeled “Fisher”) vs. the performance of standard MSE-optimum DFT-based transform compression (labeled “MSE”) as the SNR_1 at sensor S_1 is varied; the value of SNR_2 at sensor S_2 is fixed at 40 dB. We have included the case where just the signal at S_1 is compressed (labeled “S1”) as well as the case where the signals at S_1 and S_2 were both compressed (labeled “S1&S2”). The performance with no compression is labeled “w/o comp”. Figure 4 and Figure 5 show results for the case of both SNR's changing but set equal to each other.

In all cases shown, our method provides better TDOA accuracy than the MSE-optimized method; at moderately high SNR our method is nearly the same as when no compression is used – even when the CR is 8:1. Also note that when both sensor's signals have been compressed at 8:1, our method's performance is only degraded a small amount where as the performance of the MSE-based method is severely degraded.

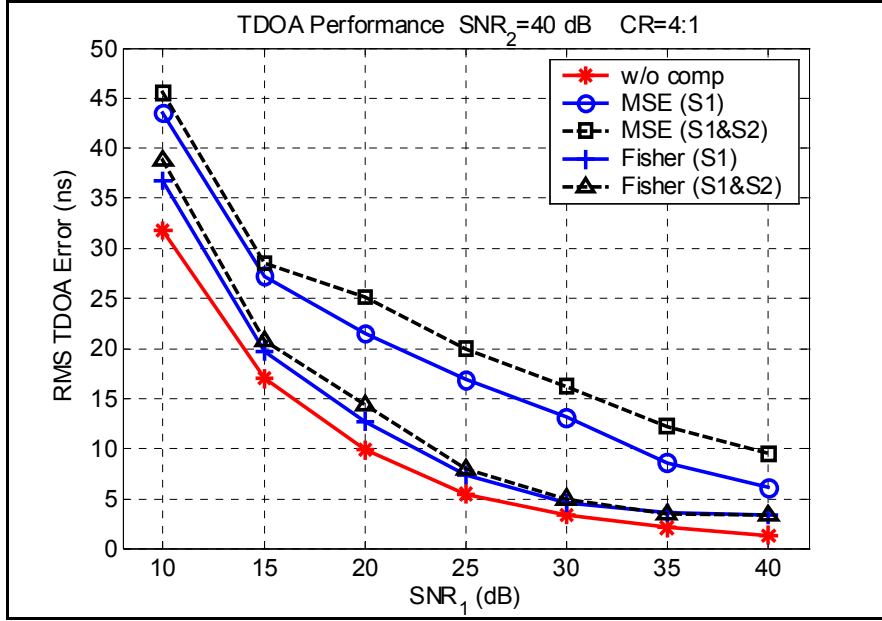


Figure 2: TDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 4:1; the SNR of the sensor S_2 signal was 40 dB.

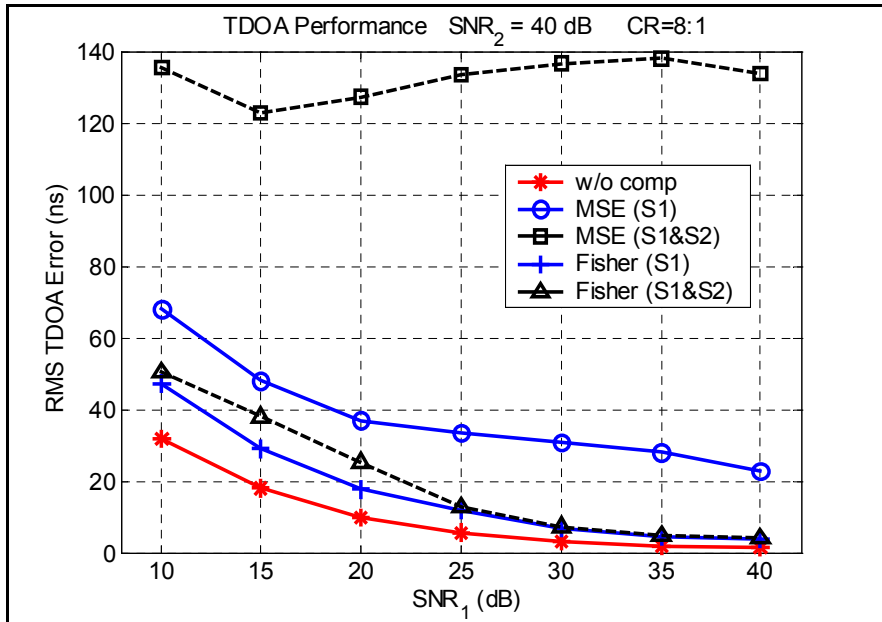


Figure 3: TDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 8:1; the SNR of the sensor S_2 signal was 40 dB.

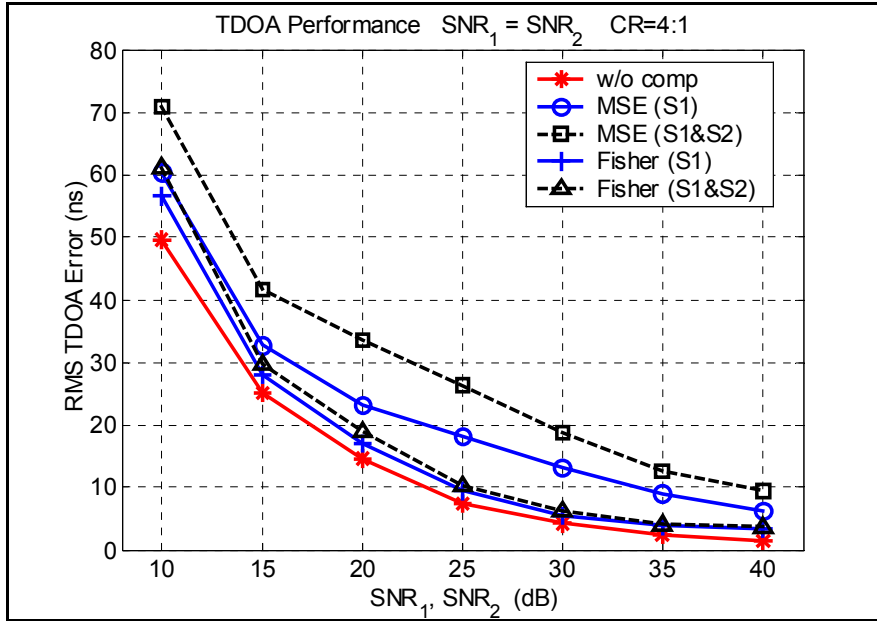


Figure 4: TDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 4:1; the SNR of the sensor S_2 signal was set equal to SNR_1 .

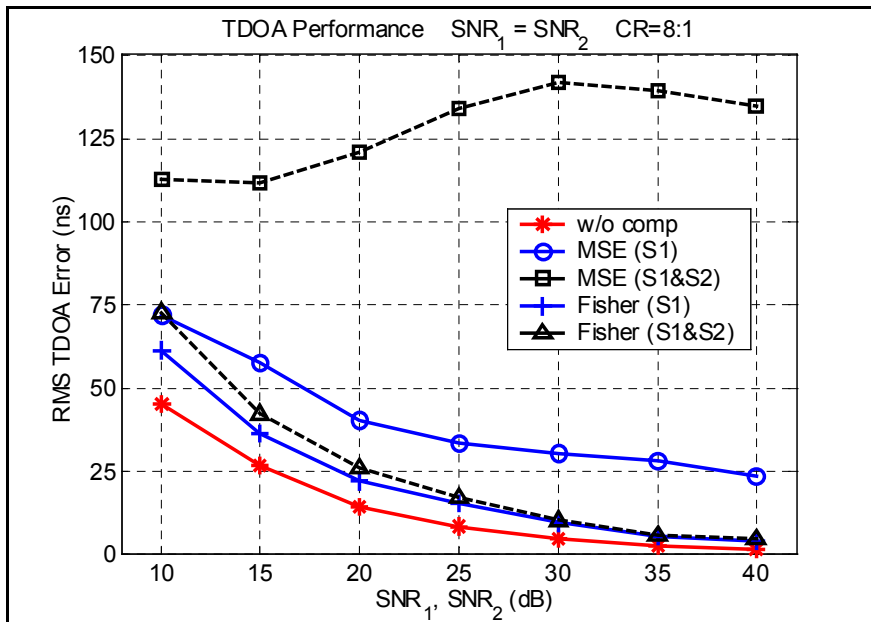


Figure 5: TDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 8:1; the SNR of the sensor S_2 signal was set equal to SNR_1 .

B. Compression for FDOA Estimation

It is well known [4] that for FDOA estimation, the FI is proportional to $\int_{-\infty}^{\infty} t^2 |s(t)|^2 dt$. This view drives us to choose the identity transform (which is an ON transform) to provide the ability to discard time components that contribute little to the FI. Thus we will directly quantize the complex-valued signal samples, using individual quantizers for the real and imaginary parts. The model for two passively-received signals having an unknown FDOA of Δ is given by

$$\begin{aligned} x_1[n] &= s[n]e^{j(\nu_0 + \Delta/2)n} + w_1[n] \quad n = -N/2, -N/2+1, \dots, N/2 \\ x_2[n] &= s[n]e^{j(\nu_0 - \Delta/2)n} + w_2[n] \quad n = -N/2, -N/2+1, \dots, N/2 \end{aligned} \quad (15)$$

where ν_0 is an unknown nuisance parameter that can not be estimated, and $w_i[n]$ is complex Gaussian noise with variance of σ_i^2 , with σ_1^2 assumed known. The model in (15) is mathematically identical to the TDOA model for the DFT transform in (12) and therefore we can use the previous results to immediately state that after multi-bit quantization the per-sample FI becomes

$$\hat{j}^{(m)}(x_1[n]) = \frac{2\pi^2 n^2 |x_1[n]|^2}{\sigma_1^2 + q_n^2}, \quad n = -N/2, -N/2+1, \dots, N/2. \quad (16)$$

For the 1-bit result application of (4) and (5) along the lines of those in [22] gives

$$\begin{aligned} \hat{j}^{(1)}(x_1[n]) &= \frac{4}{\pi\sigma^2} n^2 \left(\text{Im}^2 \{x_1[n]\} \exp\left(-\frac{2}{\sigma^2} \text{Re}^2 \{x_1[n]\}\right) \frac{1}{1 - \text{erf}^2(\text{Re}\{x_1[n]\}/\sigma)} + \right. \\ &\quad \left. \text{Re}^2 \{x_1[n]\} \exp\left(-\frac{2}{\sigma^2} \text{Im}^2 \{x_1[n]\}\right) \frac{1}{1 - \text{erf}^2(\text{Im}\{x_1[n]\}/\sigma)} \right) \end{aligned} \quad (17)$$

Using these results in (9) with the Lagrange optimization method produces the results shown in Figure 6 and Figure 7, which show results for the FDOA case. In all cases shown, our method provides better FDOA accuracy than the MSE-optimized method.

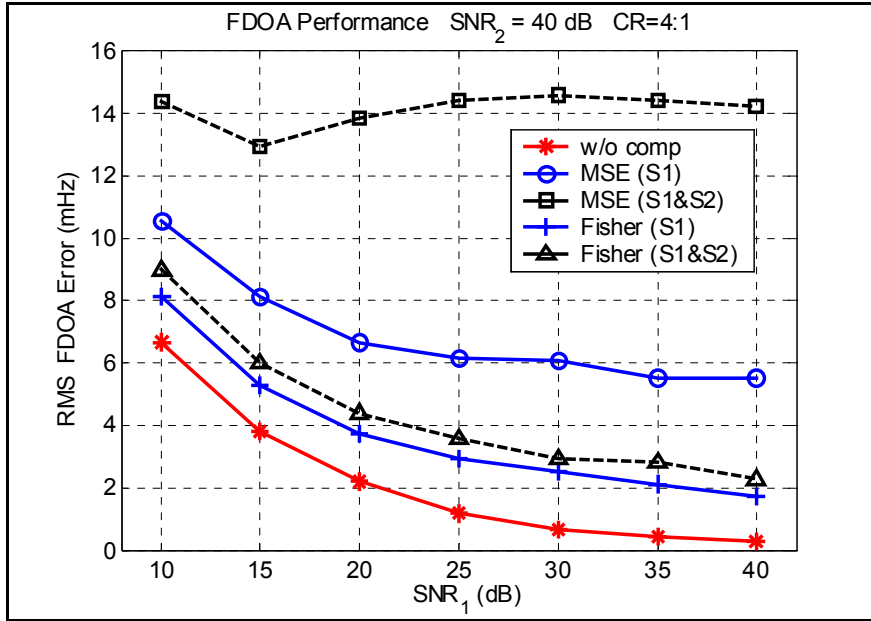


Figure 6: FDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 4:1; the SNR of the sensor S_2 signal was 40 dB.

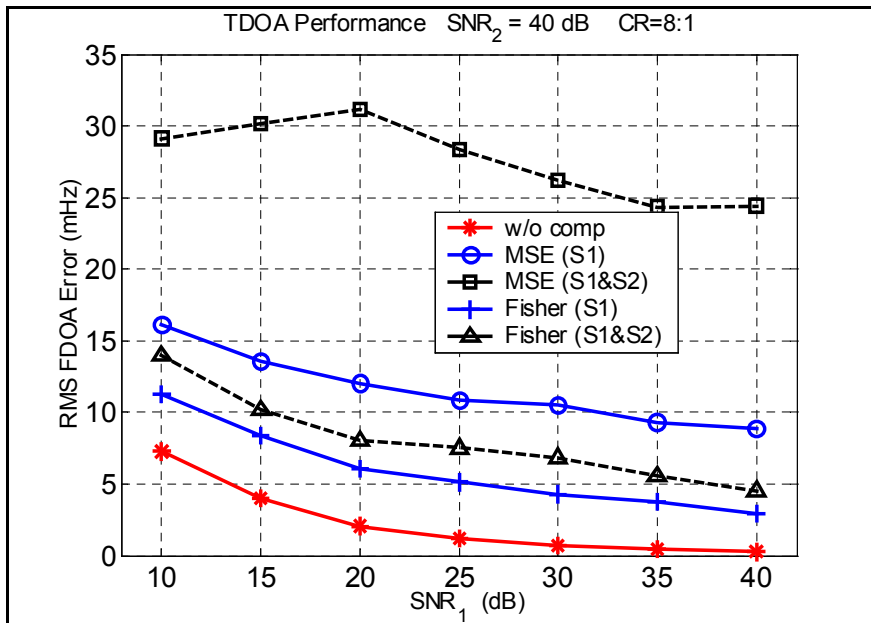


Figure 7: FDOA accuracy vs SNR of pre-compressed sensor S_1 signal for a CR of 8:1; the SNR of the sensor S_2 signal was 40 dB

IV. Concluding Remarks

As demonstrated in the example applications, the use of a distortion measure designed specifically for a specific estimation problem can lead to compression methods that outperform those using MSE-based distortion measures, especially when both sensor signals need to be compressed at high CRs. While MSE distortion accurately captures the effect of the compression on the compressed signal's SNR, it fails to capture the true impact of compression on the estimation accuracy. This is similar to the scenario in image and audio compression, where MSE distortion fails to capture the impact of compression perceptual quality of the compressed data. In those areas researchers have proposed effective distortion measures based on the psychology of perception. Of course, others have used such inference-centric distortion measures before for quantizer design; but here we see the power of combining this with transform coding, which leads to some new insights: (i) the structure of the FI provides insight into the proper choice of transform choice, (ii) the choice of transform and the optimal bit allocation can be in conflict for different parameter estimations – this is important as we extend to the case of multiple estimations, and (iii) it is possible to optimize FI-based measures within a specified operational compression framework. There are some directions for which further work is needed: (i) extension to the case of multiple estimates and decisions – although we have some preliminary results in this area [3], (ii) examination of the computational and implementation aspects; particularly, better models for the multi-bit post-compression FI, and (iii) generalization to the case when the FI depends on the parameter θ ; at the end of Section III-A we have proposed what we think is the correct approach but have not yet fully explored its application.

Appendix: Numerical Results for Gaussian Approximation

We assume that we have a set of noise-free coefficients that lie in the range $\pm A$. A noisy version of the coefficients having additive Gaussian noise of variance σ^2 is quantized to b bits using a mid-step uniform quantizer with quantization cell size given by

$$\Delta = \frac{\sqrt{A^2 + \sigma^2}}{2^{b-1} - 1/2}.$$

The true PDF can be numerically found via convolution of a Gaussian PDF with a uniform PDF and then plotted for various values of the peak SNR (PSNR) to be $PSNR = A^2 / \sigma^2$. Two such plots are given in Fig. A-1 for the case of 2 bit quantization at two values of $PSNR$. The results provide motivation that at least for low PSNR the approximation seems to be valid even down to the lowest number of bits for which it is applied. For higher PSNR the approximations at 2 bits will be poorer; nonetheless, we use the approximation. For sensor problems the interest generally lies at low SNR and it is good that we have a better approximation in that range.

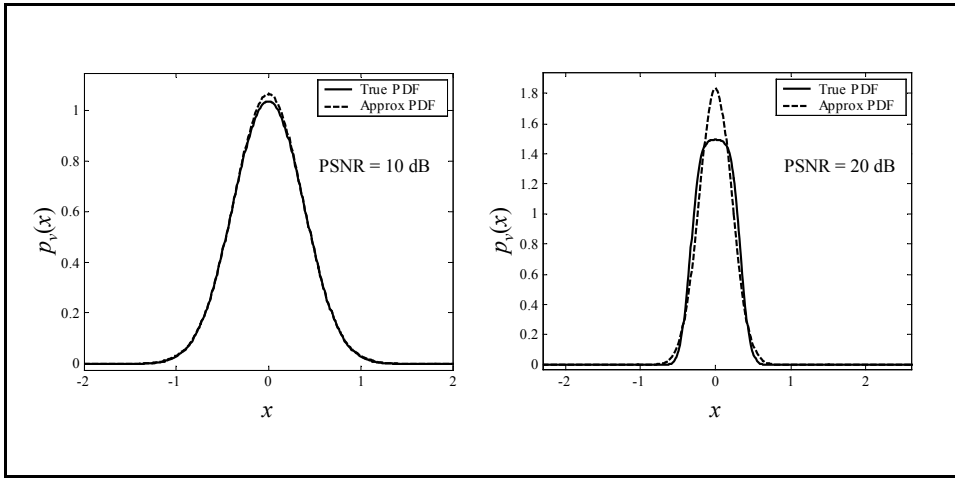


Figure A - 1 Numerical results for the true PDF compared to the approximation PDF for the case of 2 bit quantization for $PSNR$ of 10 dB and 20 dB.

References

- [1] D. Torrieri, "Statistical Theory of Passive Location Systems," *IEEE Trans. on Aerospace and Electronic Systems*, pp. 183 - 198, March 1984.
- [2] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, "A taxonomy of wireless micro-sensor network models," *ACM Mobile Computing and Communications Review (MC2R)*, vol. 6, pp. 1 - 8, April 2002.
- [3] M. Chen and M. L. Fowler, "Geometry-Adaptive Data Compression For TDOA/FDOA Location," *IEEE ICASSP 2005*, Philadelphia, PA, pp. IV1069 - IV1072, March 18 - 23, 2005.
- [4] S. Stein, "Differential Delay/Doppler ML Estimation with Unknown Signals," *IEEE Trans. on Signal Processing*, pp. 2717 - 2719, August 1993.
- [5] D. J. Matthiesen and G. D. Miller, "Data transfer minimization for coherent passive location systems," Report No. ESD-TR-81-129, Air Force Project No. 4110, June 1981.
- [6] G. Desjardins, "TDOA/FDOA technique for locating a transmitter," US Patent #5,570,099 issued Oct. 29, 1996, Lockheed Martin Federal Systems.
- [7] M. L. Fowler, "Coarse quantization for data compression in coherent location systems," *IEEE Trans. Aero. and Electr. Systems*, vol. 36, no. 4, pp. 1269 - 1278, Oct. 2000.
- [8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [9] A. Ortega and K. Ramchandran, "Rate distortion methods for image and video compression," *IEEE Signal Processing Magazine*, vol. 15, Nov. 1998, pp. 23 - 50.
- [10] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [11] M. Di Bisceglie and M. Longo, "Decentralized encoding of a remote source," *Signal Processing*, 55 (1996), pp. 15 - 29.
- [12] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Transactions on Information Theory*, vol. IT-33, no. 6, Nov. 1987, pp. 773 - 211.
- [13] J. A. Gubner, "Distributed estimation and quantization," *IEEE Transactions on Information Theory*, vol. 39, no. 4, July 1993, pp. 1456 - 1459.
- [14] T. S. Han and S. Amari, "Parameter estimation with multiterminal data compression," *IEEE Transactions on Information Theory*, vol. 41, no. 6, Nov. 1995, pp. 1802 - 1833.
- [15] H. V. Poor, "Fine Quantization in Signal Detection and Estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 5, Sept. 1988, pp. 960 - 972.

- [16] W. Lam and A. R. Reibman, "Design of quantizers for decentralized estimation systems," *IEEE Transactions on Communications*, vol. 41, no. 11, Nov. 1993, pp. 1602 – 1605.
- [17] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, pp. 51– 60, March 2002.
- [18] A. Scaglione and S. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," *MOBICOM'02*, Sept. 23 – 26, 2002, Atlanta, Ga.
- [19] Z. Zhang and T. Berger, "Estimation via compressed information," *IEEE Transactions on Information Theory*, vol. 34, no. 2, March 1988, pp. 198 – 211.
- [20] L. Vasudevan, A. Ortega, U. Mitra, "Application specific compression for time delay estimation in sensor networks," *Proceedings of the First ACM International Conference on Embedded Networked Sensor Systems*, Nov. 2003, Los Angeles, CA, pp. 243-253.
- [21] R. M. Gray and T. G. Stockham, Jr., "Dithered Quantizers," *IEEE Trans. on Inform. Theory*, vol. 39, pp. 805 - 812, May 1993.
- [22] A. Høst-Madsen and P. Händel, "Effects of Sampling and Quantization on Single-Tone Frequency Estimation," *IEEE Trans. Signal Processing*, Vol. 48, No. 3, pp. 650 – 662, March 2000.
- [23] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 36(9): 1445-1453, 1988.
- [24] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Information Theory*, Vol. 44, No. 6, pp. 1 – 63, October 1998.