ARL-TR-9074 ● SEP 2020

Fisher Information in Radio Frequency (RF) Field Data

by Vinod K Mishra

ARL-TR-9074 ● SEP 2020

Fisher Information in Radio Frequency (RF) Field Data

Vinod K Mishra
Computational and Information Sciences Directorate, CCDC Army Research Laboratory

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2020	Technical Report	May–July 2020

4. TITLE AND SUBTITLE
Fisher Information in Radio Frequency (RF) Field Data

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Vinod K Mishra

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CCDC Army Research Laboratory
ATTN: FCDD-RLC-NC
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-9074

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID(s): Vinod Mishra, 0000-0001-9432-9082

14. ABSTRACT
The extraction of useful information from large, disparate, and heterogeneous data sets requires a good set of metrics. In this work we use Fisher information as a tool to distinguish the modulation schemes of RF signals in the field.

15. SUBJECT TERMS
radio frequency, Fisher information, RF modulation, field data, information geometry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Vinod K Mishra
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	UU	15	(410) 278-0114

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

1. Introduction

The current digital revolution has resulted in a flood of data in many fields, which has necessitated the development of many tools for understanding their meanings and patterns. Some of them may be useful in analyzing underlying patterns in field-measured RF signals.

Information geometry (IG) is one of the very important tools for understanding big data. It arose from the considerations of applying concepts of differential geometry to statistical models. In IG, the familiar geometric concepts are redefined (e.g., "distances" are measured not in the coordinate space but in the space of functions such as density matrices describing different models). The Fisher information (FI) is one of the key IG metrics.[1–4] FI determines the direction in the parameter space for giving the largest change in the objective function for unit change in the parameter. Therefore, it is a kind of curvature.

In this report, we focus on FI and describe its usefulness in understanding RF data. Analytical expressions for FI for some modulation schemes have been derived in Delmas[3] and Bellili et al.[4] under general conditions. It was found that a modified version of those expressions could distinguish between some standard RF modulation schemes of the field-measured data.

2. Fisher Information (FI)

The FI is a very important concept of IG. Let there be

1) X = a random variable
2) θ = unknown parameter of a distribution that models the behavior of X
3) $f(x;\theta)$ = the probability density function of *X,* which gives the probability of a given outcome x (of random variable *X*), for a known value of θ

Then the FI for continuous variables is defined as

$$F(\theta) = \int \left(\frac{\partial}{\partial\theta} log f(x;\theta)\right)^2 f(x;\theta)dx \tag{1}$$

The inverse of FI is a lower bound on the variance of any unbiased estimator of θ.

$$Var(\hat{\theta}) \geq \frac{1}{F(\theta)} \tag{2}$$

This result is known as Cramér-Rao bound (CRB). In practical terms, it means that higher value of FI implies a more “true” estimate of the unknown parameter. This same property makes FI a faithful indicator of the change in stability and trend of a system in response to the changes in the underlying parameters.

Additionally, FI has many connections to some other well-known mathematical concepts.

1) FI can be considered as a Riemannian metric on the space of parametric distributions,
2) FI is equal to the negative expected Hessian of log likelihood of a given probability density function
3) FI is equal to the Jacobian of the gradient of a given probability density function.

As an example, let $f(x;\theta)$ be a Gaussian distribution with mean θ and variance σ^2.

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right) \tag{3}$$

Then the FI is calculated as

$$F(\theta) = \frac{1}{\sigma^2} \tag{4}$$

So for the Gaussian distribution, the FI is large if variance is small and it is small for large variance. In the situation with many parameters (e.g., $\theta_1, \theta_2, \ldots$), it is known as Fisher Information Matrix and is given by

$$F_{ij}(\theta) = \int \left[\frac{\partial}{\partial\theta_i} log f(x;\theta)\right] f(x;\theta) \left[\frac{\partial}{\partial\theta_j} log f(x;\theta)\right] dx \tag{5}$$

3. Application to RF Data

The RF data collected by Spectrum Analyzers in the field contain much information about the electromagnetic environment of the surrounding area. We need data analysis tools for uncovering the useful information hidden in this data. FI, or equivalently CRB, is one such tool.

3.1 Theoretical Background

We use the expression for CRB given in Delmas.[3] First, we define some relevant quantities:

1) Observation vector:

$$y_k = a s_k e^{i(2\pi k\nu+\phi)} + n_k,\ k = k_0, k_0+1, \ldots., k_0+K-1. \tag{6}$$

2) $\{s_k\}$ = a sequence of independent identically distributed (IID) data symbols taking values with equal probabilities.

   - ± 1 for binary phase shift keying (BPSK),
   - $(\pm 1 \pm i)/\sqrt{2}$ for quadrature phase shift keying (QPSK), and
   - $s_{k+1} = i s_k c_k$ for minimum shift keying (MSK) with c_k as sequence of BPSK symbols and $s_k = \{+1, i, -1, -i\}$ set

3) $\{n_k\}$ = a sequence of IID zero-mean complex circular Gaussian noise random variables of variance $\sigma^2 = E|n_k|^2$

4) $\{\boldsymbol{\theta}\} = \{\nu, \phi, a, \sigma\}$; the parameter set of the signal and noise

5) The probability distribution function is given as

$$p(y_k; \boldsymbol{\theta}) = \frac{1}{\pi\sigma^2} exp\left(-\frac{|y_k|^2 + a^2}{2\sigma^2}\right) c(y_k) \tag{7}$$

with

$$c(y_k) = \cosh\left(\frac{a}{g_1(y_k)\sigma^2}\right); \qquad \text{BPSK} \tag{8a}$$

$$c(y_k) = cosh\left(\frac{a}{g_1(y_k)\sqrt{2}\sigma^2}\right) cosh\left(\frac{a}{g_2(y_k)\sqrt{2}\sigma^2}\right), \qquad \text{QPSK} \tag{8b}$$

$$c(y_k) = cosh\left(\frac{a}{g_3(y_k)\sigma^2}\right) \qquad \text{MSK} \tag{8c}$$

and

$$g_1(y_k) = 2Re(e^{i(2\pi k\nu+\phi)} y_k^*); \tag{9a}$$

$$g_2(y_k) = 2Im(e^{i(2\pi k\nu+\phi)} y_k^*); \tag{9b}$$

$$g_3(y_k) = 2Re(i^{k-k_0} s_{k_0} e^{i(2\pi k\nu+\phi)} y_k^*) \tag{9c}$$

6) The FI is given as

$$F_{ij}(\boldsymbol{\theta}) = -\sum_{k=k_0}^{k_0+K-1} \left[\frac{\partial}{\partial\theta_i} \log p(y_k; \boldsymbol{\theta})\right] \log p(y_k; \boldsymbol{\theta}) \left[\frac{\partial}{\partial\theta_j} log \log p(y_k; \boldsymbol{\theta})\right], (i,j) = \{\nu, \phi\}, \{a, \sigma\} \tag{10}$$

The parameter set is partitioned in two separate sets using the properties of random Gaussian variables (Delmas[3], Appendix A). The final FI expressions mainly as a function of the ignal-to-noise ratio ρ for joint estimation of parameters are

1) Phase

$$F_{\phi\phi}(\boldsymbol{\theta}) = 2K\rho\{1 - f_1(\rho)\}, \text{ for BPSK and MSK, replace } f_1(\rho) \text{ by } (1+\rho)f_1(\rho/2) \text{ for QPSK} \tag{11a}$$

2) Frequency

$$F_{\nu\nu}(\boldsymbol{\theta}) = \frac{1}{3}\pi^2(K^2 - 1)F_{\phi\phi}(\boldsymbol{\theta}), \text{ for BPSK and MSK, replace } f_1(\rho) \text{ by } (1+\rho)f_1(\rho/2) \text{ for QPSK} \tag{11b}$$

3) Amplitude

$$F_{aa}(\boldsymbol{\theta}) = \frac{2K\rho\{1-(1+2\rho)f_2(\rho)\}}{a^2\{1-2\rho f_2(\rho)\}}, \text{ for BPSK and MSK, replace } f_2(\rho) \text{ by } f_2(\rho/2) \text{ for QPSK} \tag{12a}$$

4) Noise

$$F_{\sigma\sigma}(\boldsymbol{\theta}) = 2F_{aa}(\boldsymbol{\theta}), \text{ for BPSK and MSK, replace } f_2(\rho) \text{ by } f_2(\rho/2) \text{ for QPSK} \tag{12b}$$

Here

$$\begin{bmatrix} f_1(\rho) \\ f_2(\rho) \end{bmatrix} = e^{-\rho}\sqrt{\frac{2}{\pi}}\int_0^{\infty} \frac{e^{-u^2/2}}{\cosh(u\sqrt{2\rho})}\begin{bmatrix} 1 \\ u^2 \end{bmatrix} du \tag{13}$$

We normalize the values by division with $\rho = 1$ values of FI.

$$modified\ F_{\phi\phi}(\rho) = \frac{F_{\phi\phi}(\rho)}{F_{\phi\phi}(\rho=1)} = \frac{2\rho\{1-f_1(\rho)\}}{2\{1-f_1(\rho=1)\}}, \text{ for BPSK and MSK,} \tag{14a}$$

$$= \frac{2\rho\{1-f_1(\rho)\}}{2\{1-2f_1(\rho=1/2)\}}, \text{ for QPSK} \tag{14b}$$

Here

$$\begin{bmatrix} f_1(\rho = 1) \\ f_1(\rho = 1/2) \end{bmatrix} = \sqrt{\frac{2}{\pi}}\int_0^{\infty} e^{-u^2/2}\begin{bmatrix} e^{-1}/\cosh(u\sqrt{2}) \\ e^{-1/2}/\cosh(u) \end{bmatrix} du \tag{15}$$

3.2 Numerical Calculations

The particular data set we will consider is a stream of I and Q values for a random alphabet and is presented as a time series. The measured I_m and Q_m values have noise mixed with the corresponding noiseless values I and Q. The measured amplitude A_m is related to I_m and Q_m as

$$E(A_m^2) = I_m^2 + Q_m^2 \tag{16}$$

Here E denotes expectation value. The measured amplitude is a mixture of random noise and signal amplitude. A similar relation for noiseless amplitude is given by

$$E(A^2) = I^2 + Q^2 \tag{17}$$

It is assumed that the noise variable is a Gaussian with zero mean and σ^2 variance.

$$E(A_m^2) = E(A^2) + \sigma^2 \tag{18}$$

Define, $\rho_m = E(A_m^2)/\sigma^2$ and $\rho = E(A^2)/\sigma^2$, then we get

$$\rho = \rho_m - 1 \tag{19}$$

4. Numerical Results

We choose $\sigma^2 = 2 \times 10^{-9}$ and present the results for the modified classical Fisher Information (mCFI) for both BPSK and QPSK. The raw data correspond to the RF signals measured by a receiver and represent a sequence of random symbols of a speech. The mCFI is calculated as a function of ρ (Fig. 1). This metric has similar patterns for both BPSK/MSK and QPSK modulations but their values are very different. It is clear that this metric can distinguish between BPSK/MSK and QPSK modulations.

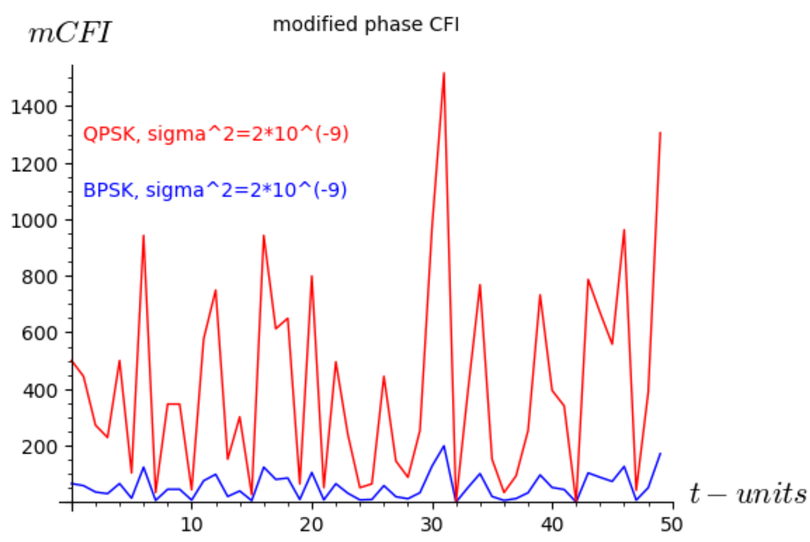

Fig. 1 Modified classical Fisher Information (mCFI) vs. time (arbitrary units) plots for BPSK/MSK and QPSK

It was found that for $\sigma^2 > 2 \times 10^{-9}$, ρ becomes negative and the integral in mCFI cannot be calculated. This observation indicates that the analytical approach has a built-in noise cutoff beyond which it is not applicable.

5. Conclusion and Next Steps

The simple model presented in this report is useful for finding out if the underlying signal modulation is BPSK/MSK or QPSK. The results indicate that it is applicable only for $\rho_m > 1$. In the future, this approach will expand to the following investigations in the mCFI behavior:

1) Modulation schemes other than BPSK/MSK and QPSK, including nonstandard waveforms
2) Presence of more than one waveform in the field (with same or different modulations), including both friendly and adversarial RF sources

In addition, one would like to understand the utility of this approach in the relevant machine learning algorithms. Ultimately, its usefulness in field scenarios will determine the utility of this approach.

6. References

1. Klein A. Matrix algebraic properties of the Fisher information matrix of stationary processes. Entropy. 2014;16:2023–2055. doi:10.3390/e16042023.

2. Ahmad N, Derrible S, Eason T, Cabezas H. Using Fisher information to track stability in multivariate systems. R Soc Open Sci. 2016;3:160582. http://dx.doi.org/10.1098/rsos.160582.

3. Delmas J-P. Closed-form expressions of the exact Cramer-Rao bound for parameter estimation of BPSK, MSK, or QPSK waveforms. IEEE Signal Processing Letters. 2008 Feb. doi: 10.1109/LSP.2008.921477.

4. Bellili F, Atitallah N, Affes S, Stephenne S. Cramér-Rao lower bounds for frequency and phase NDA estimation from arbitrary square QAM-Modulated signals. IEEE Transactions on Signal Processing. 2010 Oct. doI: 10.1109/TSP.2010.2051154.

List of Symbols, Abbreviations, and Acronyms

ARL	Army Research Laboratory
BPSK	binary phase shift keying
CCDC	US Army Combat Capabilities Development Command
CRB	Cramér-Rao Bound
FI	Fisher information
IG	information geometry
IID	independent identically distributed
mCFI	modified classical Fisher information
MSK	minimum shift keying
QPSK	quadrature phase shift keying
RF	radio frequency

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	CCDC ARL FCDD RLD DCI TECH LIB
1 (PDF)	CCDC ARL FCDD RLC NC V MISHRA