Bearings-Only Tracking Analysis via Information Geometry

Xuezhi Wang*, Yongqiang Cheng†, and Bill Moran*

†School of Electronic Science and Engineering,
National University of Defense Technology, Changsha, Hunan, 410073, P.R.China
Email: nudtyqcheng@gmail.com

* Melbourne Systems Laboratory,
Faculty of Engineering, University of Melbourne, Australia
Email: xwang@unimelb.edu.au, b.moran@unimelb.edu.au

Abstract – *In this paper, the problem of bearings-only tracking with a single sensor is studied via the theory of information geometry, where Fisher information matrix plays the role of Riemannian metric. Under a given tracking scenario, the Fisher information distance between two targets is approximately calculated over the window of surveillance region and is compared to the corresponding Kullback Leibler divergence. It is demonstrated that both "distances" provide a contour map that describes the information difference between the location of a target and a specified point. Furthermore, an analytical result for the optimal heading of a given constant speed sensor is derived based on the the properties of statistical manifolds.*

Keywords: Bearings-Only Tracking, Fisher Information Distance, Information Geometry, Optimal Sensor Heading.

1 Introduction

The complexity of bearings-only tracking (BOT) with a single sensor combines the potential unobservability of the underlying target state with strong nonlinearity between measurement and and target state [1]. It is well understood in the literature that a necessary condition for the position and velocity of a constant velocity target to be fully observable is the availability of measurements acquired before and after an ownship manoeuvre [2]. Interestingly, necessary and sufficient conditions derived in [3] show that there is a class of sensor maneuvers for which full observability of the target state is not achieved. The potential lack of observability, and the nonlinearity in the measurement equation have constantly challenged the development of appropriate filtering techniques [4] to solve the BOT problem optimally [5].

Information geometry pioneered by Cramer and Rao from 40s [6] and brought to maturity in the works of Amari [7], offers comprehensive results about statistical models simply by considering them as geometrical objects and the statistical structures as geometrical structures. As a powerful mathematical tool, it can also provide additional perspectives in the analysis of measurement of systems. The main tenet of information geometry is that many important structures in probability theory, information theory and statistics can be treated as structures in differential geometry by regarding a space of probabilities as a differentiable manifold endowed with a Riemannian metric and a family of affine connections distinct from the canonical Levi-Cevita connection [8]. The manifold is proved to have a unique Riemannian metric given by the Fisher information matrix (FIM), and a dual pairs of affine connections [7].

For a given target tracking system for which the Fisher information matrix can be calculated, the Fisher information distance (FID) on the statistical manifold between two target states is well defined and can be used as a measure for target resolvability over the region of interest. Such an advantage was demonstrated in [9] and motivated this work. Alternatively, the Kullback Leibler divergence provides a simple way to approximate the information distance without the knowledge of the statistical manifold.

In this paper, the problem of bearings-only tracking with a single sensor is studied via information geometry, where the Fisher information matrix plays the role of a Riemannian metric. Under a given tracking scenario, the FID between two targets is approximately calculated over the window of surveillance region and is compared to the corresponding Kullback Leibler divergence. It is demonstrated that both "distances" provide a contour map that describes the information difference between the location of a target and a specified point. Furthermore, an analytical result for the optimal heading of a given constant speed sensor is derived based on the properties of statistical manifolds.

Following the introduction, the BOT problem is briefly described in Section 2 where our analysis method is also introduced. A detailed analysis based on a particular BOT scenario is presented in Section 3 and the significance of the analysis in the statistical manifolds is also highlighted. Finally, conclusions are drawn in Section 4.

2 Problem and Analysis Method

Tracking a constant velocity (CV) target with a sequence of bearings-only measurements is equivalent to the problem of localising a CV target at given time points via multiple bearings-only sensors of known locations. At every scan t_k, knowledge of previous target location $[x_{k-1}, y_{k-1}]^T$, velocity $[\dot{x}, \dot{y}]^T$ and sensor measurement φ_{k-1} taken at location $[\eta_{k-1}, \xi_{k-1}]^T$ at the previous scan t_{k-1} are assumed known and the sensor is taking a new measurement φ_k at location $[\eta_k, \xi_k]^T$ from the target.

The BOT task is thus to estimate target location $[x_k, y_k]^T$ at scan t_k from the prior and sensor measurements. The crux of the underlying problem corresponds to that of localizing a static target with two sensors and such a scenario is illustrated in Fig. 1.

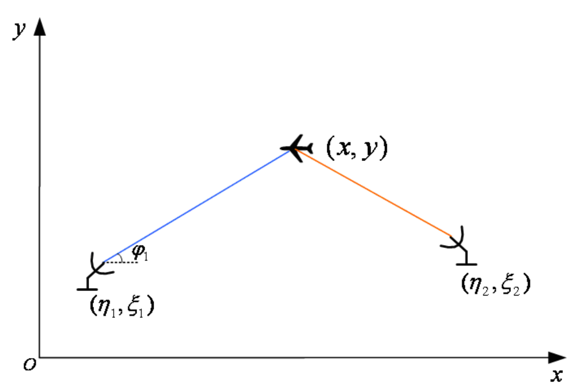

Figure 1: The measurement model of bearing-only measurements of two sensors.

For simplicity we ignore the system process noise and the two passive sensors are assumed to be able to observe the bearing of a target subject to a Gaussian zero-mean random noise $\boldsymbol{w}_k \sim \mathcal{N}(0,\ \boldsymbol{C}_k)$. Therefore, the system and measurement models are given by

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} \tag{1}$$

$$\boldsymbol{x}_k = \boldsymbol{\mu}(\boldsymbol{\theta}_k) + \boldsymbol{w}_k, \tag{2}$$

where the target state is represented by $\boldsymbol{\theta}_k = [x_k, y_k]^T$, the measurement function $\boldsymbol{\mu}_k$ connects the measurement $\boldsymbol{x}_k$ with target state $\boldsymbol{\theta}_k$. The measurement likelihood function for this system is of the form

$$\boldsymbol{x}_k|\boldsymbol{\theta}_k \sim \mathcal{N}\Big(\ \boldsymbol{\mu}_k,\ \boldsymbol{C}_k\Big) \tag{3}$$

where

$$\boldsymbol{\mu}_k = \begin{bmatrix} \arctan\left(\frac{y-\xi_1}{x-\eta_1}\right) \\ \arctan\left(\frac{y-\xi_2}{x-\eta_2}\right) \end{bmatrix},\quad \boldsymbol{C}_k = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \tag{4}$$

σ_1 and σ_2 signify the standard deviations of measurement noise for sensors 1 and 2, respectively.

The FIM in this case is derived (see Appendix A) as

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{11} = \sum_{i=1}^{2} \frac{1}{\sigma_i^2} \frac{(y-\xi_i)^2}{[(x-\eta_i)^2 + (y-\xi_i)^2]^2} \tag{5}$$

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{12} = [\boldsymbol{I}(\boldsymbol{\theta})]_{21} = \sum_{i=1}^{2} -\frac{1}{\sigma_i^2} \frac{(x-\eta_i)(y-\xi_i)}{[(x-\eta_i)^2 + (y-\xi_i)^2]^2} \tag{6}$$

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{22} = \sum_{i=1}^{2} \frac{1}{\sigma_i^2} \frac{(x-\eta_i)^2}{[(x-\eta_i)^2 + (y-\xi_i)^2]^2} \tag{7}$$

The squared differential FID ds^2 between the distributions $p(\boldsymbol{x}|\boldsymbol{\theta})$ and $p(\boldsymbol{x}|\boldsymbol{\theta}+d\boldsymbol{\theta})$ is given by the quadratic form of $d\boldsymbol{\theta}$,

$$\begin{aligned} ds^2 &= \sum_{ij} [\boldsymbol{I}(\boldsymbol{\theta})]_{ij}\, d\theta_i d\theta_j \\ &= \sum_{i=1}^{2} \frac{1}{\sigma_i^2} \Big[\frac{y-\xi_i}{(x-\eta_i)^2 + (y-\xi_i)^2} dx \\ &\quad - \frac{x-\eta_i}{(x-\eta_i)^2 + (y-\xi_i)^2} dy \Big]^2 \end{aligned} \tag{8}$$

The FID between two distributions $p(\boldsymbol{x}|\boldsymbol{\theta}_1)$ and $p(\boldsymbol{x}|\boldsymbol{\theta}_2)$ is defined by the following integral [10]

$$\mathcal{D}_F(\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2) \triangleq \arg\min_{\boldsymbol{\theta}(t)} \int_{t_1}^{t_2} \left(\sqrt{\left(\frac{d\boldsymbol{\theta}}{dt}\right)^T \boldsymbol{I}(\boldsymbol{\theta}) \left(\frac{d\boldsymbol{\theta}}{dt}\right)} \right) dt$$

$$\boldsymbol{\theta}(t_1) = \boldsymbol{\theta}_1,\ \boldsymbol{\theta}(t_2) = \boldsymbol{\theta}_2 \tag{9}$$

where $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ is the parameter path along the parameter space $\mathbb{R}^n$. Essentially, (9) amounts to finding the length of the shortest path, i.e. the geodesic, on the statistical manifold S connecting coordinates $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. However, finding the geodesic on a statistical manifold is not trivial.

Following the approach in [9], we approximate the FID by calculating the integral of the differential FID along the shortest path between two targets on target state space (rather than a geodesic) via a parametric equation. Equation (8) is then given by

$$\begin{aligned} ds^2 = \mathcal{F}(t)dt^2 = dt^2 \sum_{i=1}^{2} \Big[&\frac{y_1 - \xi_i + t\sin\alpha}{A_i(t)} \cos\alpha \\ &- \frac{x_1 - \eta_i + t\cos\alpha}{A_i(t)} \sin\alpha \Big]^2 \end{aligned} \tag{10}$$

where

$$A_i(t) = (x_1 - \eta_i + t\cos\alpha)^2 + (y_1 - \xi_i + t\sin\alpha)^2 \quad i = 1, 2 \tag{11}$$

Then the approximate FID between two targets T_1 and T_2 is given by the following integral

$$\mathcal{D} = \int_0^d \sqrt{\mathcal{F}(t)}\, dt \tag{12}$$

3 Localisation Analysis

The analysis includes three parts. Firstly, a second target T_2 is assumed and the FID between the target T_1 and T_2 is computed and a contour map that illustrates the change rate of FID around target T_1 is obtained. Secondly, the determinant of FIM is calculated over the surveillance region. The related plots display the amount of information which can be acquired from the sensor network of current configuration for a given target location. Finally, a simple one step sensor scheduling problem is analysed and a closed form expression for the optimal sensor course is presented.

In the calculations, we assume that the two passive sensors are located at $(0, 0)$, $(50, 10)$ with the standard deviations of noise $\sigma_1 = \sigma_2 = 0.2$. The position of target T_1 is $(20, 30)$.

3.1 The information distance between two closely spaced targets

Using (12) the FID is approximately calculated in a given window to highlight the information difference between two closely spaced targets T_1 and T_2. A contour map which shows the information difference given by the sensor system is presented in Fig. 2.

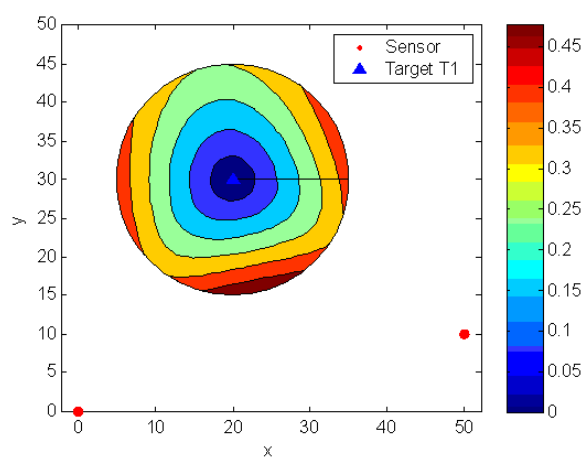

Figure 2: The contour map of the FID between targets T_1 and T_2 with bearing-only measurements of two sensors.

For comparison, a similar contour map calculated using Kullback Leibler divergence is presented in Fig. 3.

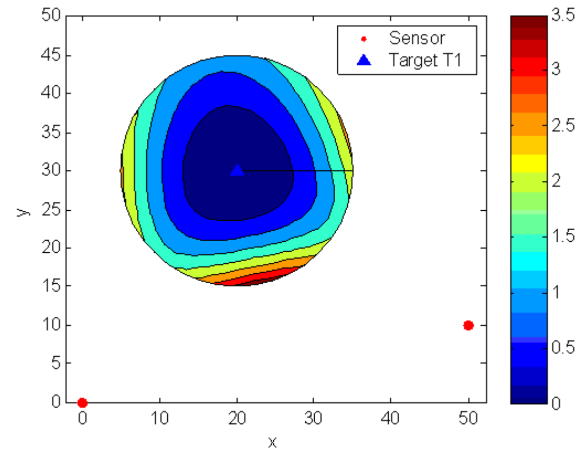

Figure 3: The contour map of the KLD between targets T_1 and T_2 with bearing-only measurements of two sensors.

It is observed that

- in both figures, the shapes of the equal information contours are very similar. In this case, the measures of FID and KLD are virtually consistent. It has been shown in [9] that the squared differential FID is twice the KLD. It is expected that the FID and KLD are equivalent for measuring closely spaced targets.

- the information to identify the difference between two targets increases more rapidly in the direction that is parallel to rather than vertical to the line joining the two sensors as targets move apart from each other. Intuitively, the resolvability of this sensor system has the same general trend as the variation of information difference.

3.2 The amount of information acquired from measurements

It is well known that the value of the determinant of the Fisher information matrix represents the volume of the amount of information [11] that can be acquired by the underlying sensor system.

Figures 4 and 5 are the surface and contour plots respectively of the determinant of FIM on a logarithm scale. Both figures indicate, as expected, that it difficult to localize targets when they fall in the narrow region along the straight line passing through the two sensors. The surface plot in Fig. 4 reveals a similar shape as the statistical manifold defined by the Fisher information metric and its height is proportional to the amount of information available from the sensor system.

The graph shown in Fig. 5 illustrates clearly where where the best estimation accuracy can be achieved and where target location is unobservable in applications to sensor networks.

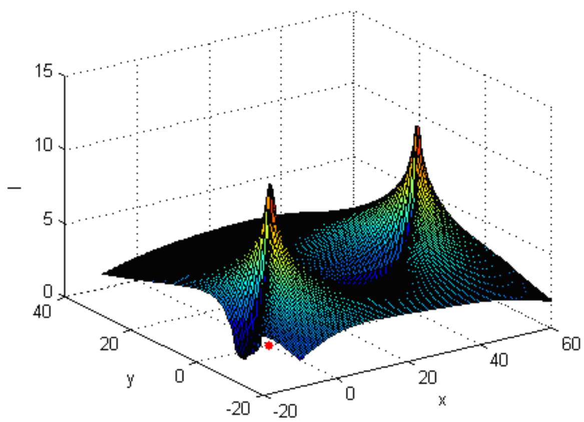

Figure 4: Target information map for the sensor network with two bearing-only passive sensors.

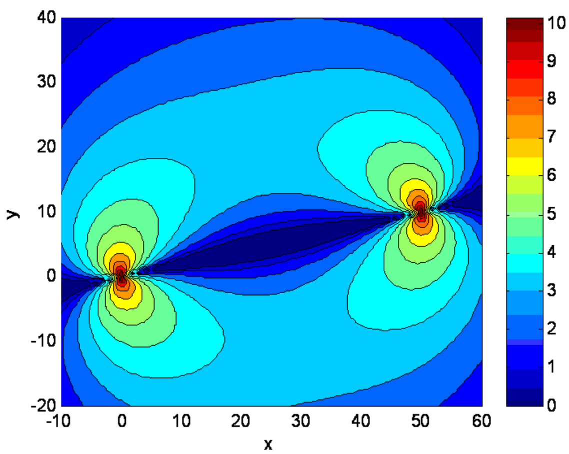

Figure 5: The contour map of Fig. 4.

3.3 Optimal sensor course

In view of the BOT scenario shown in Fig. 1, the problem of a single passive bearings-only sensor scheduling for target tracking [5] can also be formulated. In this problem, the sensor course remains unchanged until a measurement is taken and one need to optimise the sensor course before each data sampling such that the sensor will maximize the information it achieves from a measurement. In other words, we need to find the optimal sensor course such that maximum target information can be acquired when the sensor moves along this course.

Fig. 6 is the polar plots describing the the amount of information gained when the underlying sensor moves from the center toward to different directions at a “constant” speed. The target is located at 0^o with a distance to the sensor first location $d = 10$. It is assumed that at the second data sampling time, the sensor reaches a new location with a distance to its previous location $Radius = 5, 15$ and 30 respectively. From Fig. 6, we can clearly see the optimal sensor heading direction at which the maximum gain is achieved in these three cases. In accordance with the information gains

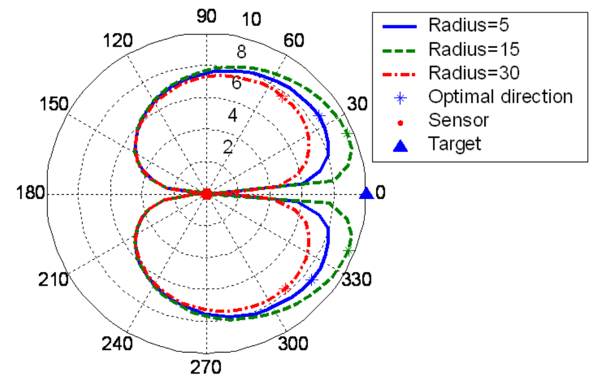

Figure 6: Information gain achievable when the passive sensor is moving in different directions at a fixed speed.

represented by the lines in Fig. 6, the sensor travel distances (which we describe as the “Radius”) between data sampling are illustrated in Fig. 7.

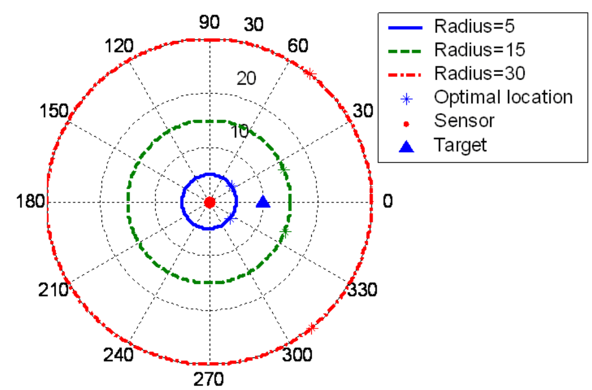

Figure 7: Geometry of sensor motion corresponding to Fig. 6. Lines represent the locations to which the sensor moves from the center at different speeds (or time).

Clearly, the optimal sensor course varies with the length of the radius (i.e., the distance the sensor travels during one scan). In an attempt to achieve some insight, we derived a closed form relationship between the optimal sensor heading φ_{opt}, the sensor travel distance r and the distance between the sensor and target d based on the maximum determinant of FIM criterion, i.e.,

$$\varphi_{opt} = \pm \arctan\left(\frac{|r^2 - d^2|}{2rd}\right) \tag{13}$$

The derivation detail is given in Appendix B. The variation of optimal sensor course with the length of radius is plotted in Fig. 8.

4 Conclusions

In this paper, the application of information geom-

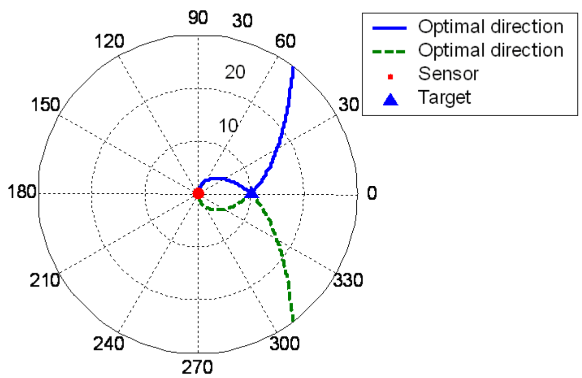

Figure 8: Optimal sensor movement direction for given radius constraints to maximize information gain in tracking a stationary target. Two symmetric solutions are indicated in different colors.

etry in bearings-only sensor network analysis is studied via a static target BOT scenario. It is evidenced that Fisher information distance between two target distributions on the statistical manifolds provides additional measures for the evaluation of a sensor (network) system. Although the example discussed is based on a static noiseless target tracking problem, it can be straightforwardly extended to the more general cases where both target motion dynamics and process noise are taken into account. However, it is found that the approximated calculation for FID is only valid when two targets are reasonably close. The exact FID can only be integrated along the geodesic that connect the two targets. We are currently working on this issue.

Appendix

A Derivation of the Fisher information matrix of general Gaussian distribution

Assume that the measurement likelihood function is of the following generic form:

$$\boldsymbol{x} \sim p(\boldsymbol{x}|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{C}(\boldsymbol{\theta})) \tag{14}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta})$ is a $n \times 1$ vector and $\boldsymbol{C}(\boldsymbol{\theta})$ is a $n \times n$ matrix, both depend on the $n \times 1$ vector valued parameter $\boldsymbol{\theta}$.

The first order derivative of the log-density function is given by

$$\begin{aligned}\frac{\partial \log p(\boldsymbol{x}|\boldsymbol{\theta})}{\partial \theta_i} &= -\frac{1}{2}\frac{\partial \log|\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \theta_i} \\ &\quad - \frac{1}{2}\frac{\partial}{\partial \theta_i}\left[(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{C}^{-1}(\boldsymbol{\theta})(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right]\end{aligned} \tag{15}$$

where $|\boldsymbol{C}|$ signifies the determinant of $\boldsymbol{C}$. The first term of (15) is derived as

$$\begin{aligned}\frac{\partial \log|\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \theta_i} &= \frac{1}{|\boldsymbol{C}(\boldsymbol{\theta})|}\frac{\partial |\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \theta_i} \\ &= \mathrm{tr}\left\{\boldsymbol{C}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}(\boldsymbol{\theta})}{\partial \theta_i}\right\}\end{aligned} \tag{16}$$

in (16) the following results were used:

$$\begin{aligned}\frac{\partial|\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \theta_i} &= \mathrm{tr}\left\{\frac{\partial|\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \boldsymbol{C}(\boldsymbol{\theta})}\frac{\partial \boldsymbol{C}^T(\boldsymbol{\theta})}{\partial \theta_i}\right\}, \\ \frac{\partial|\boldsymbol{C}(\boldsymbol{\theta})|}{\partial \boldsymbol{C}(\boldsymbol{\theta})} &= \boldsymbol{C}^{-1}(\boldsymbol{\theta})|\boldsymbol{C}(\boldsymbol{\theta})|\end{aligned}$$

The second term of (15) is derived as

$$\begin{aligned}&\frac{\partial}{\partial \theta_i}\left[(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta}))^T \boldsymbol{C}^{-1}(\boldsymbol{\theta})(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta}))\right] \\ &\quad = -2\frac{\partial \boldsymbol{\mu}^T(\boldsymbol{\theta})}{\partial \theta_i}\boldsymbol{C}^{-1}(\boldsymbol{\theta})(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})) \\ &\quad + \Big(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\Big)^T \frac{\boldsymbol{C}^{-1}(\boldsymbol{\theta})}{\partial \theta_i}\Big(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\Big)\end{aligned} \tag{17}$$

By combining (16) and (17) with (15) we obtain

$$\begin{aligned}\frac{\partial \log p(\boldsymbol{x}|\boldsymbol{\theta})}{\partial \theta_i} &= -\frac{1}{2}\mathrm{tr}\left\{\boldsymbol{C}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}(\boldsymbol{\theta})}{\partial \theta_i}\right\} \\ &+ \frac{\partial \boldsymbol{\mu}^T(\boldsymbol{\theta})}{\partial \theta_i}\boldsymbol{C}^{-1}(\boldsymbol{\theta})\left(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right) \\ &- \frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right)^T \frac{\boldsymbol{C}^{-1}(\boldsymbol{\theta})}{\partial \theta_i}\left(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right)\end{aligned} \tag{18}$$

Finally, by noting that $E\left(\boldsymbol{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right) = 0$, the Fisher information matrix can be written as

$$\begin{aligned}\left[\boldsymbol{I}(\boldsymbol{\theta})\right]_{ij} &= \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\right]^T \boldsymbol{C}^{-1}(\boldsymbol{\theta})\left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right] \\ &+ \frac{1}{2}\mathrm{tr}\left[\boldsymbol{C}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\boldsymbol{C}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j}\right]\end{aligned} \tag{19}$$

B Derivation of the optimal sensor heading for bearings-only tracking with a single sensor

If the sensor is traveling at a constant speed, a closed form expression for the optimal sensor heading during the next measurement period under the maximum determinant of FIM criterion can be derived as follows. Again, the basic BOT tracking scenario as in Fig. 1 is considered, where we wish to find the optimal angle φ_{opt} between the line connecting the two sensors and the x-axis in a 2D Cartesian coordinate system such that the target information which can be observed by the "two" sensors is maximised. The result may be extended to the case of a moving target with known dynamics. We assume that the distance r between the two sensors is fixed and given by

$$r = \sqrt{(\eta_2 - \eta_1)^2 + (\xi_2 - \xi_1)^2} \tag{20}$$

The distance between the first sensor and the target is given by

$$d = \sqrt{(x-\eta_1)^2 + (y-\xi_1)^2} \tag{21}$$

The determinant of the FIM, as a function of the second sensor location, is derived as

$$f_c(\eta_2, \xi_2) = \frac{1}{\sigma_1^2 \sigma_2^2} \times \frac{\left[(x-\eta_1)(y-\xi_2) - (x-\eta_2)(y-\xi_1)\right]^2}{\left[(x-\eta_1)^2 + (y-\xi_1)^2\right]^2 \left[(x-\eta_2)^2 + (y-\xi_2)^2\right]^2} \tag{22}$$

Assume the following parameter replacement using the polar coordinate (r, φ)

$$\begin{cases} \eta_2 = \eta_1 + r\cos\varphi \\ \xi_2 = \xi_1 + r\sin\varphi \end{cases} \tag{23}$$

where φ denotes the direction of the sensor movement (i.e. sensor heading). Substituting (23 into (22), we see that the determinant of the FIM becomes a function of r and φ, i.e., $f_P(r, \varphi)$.

The optimal sensor heading course φ_{opt}, which maximise the target information observed by the sensor, can be obtained by taking the partial derivative of $f_P(r, \varphi)$ with respect to φ and setting it to zero, which gives the following expression:

$$\tan\varphi_{opt} = \frac{\left(\frac{|r^2-d^2|}{2rd} + \frac{y-\xi_1}{x-\eta_1}\right)}{\left(1 - \frac{|r^2-d^2|}{2rd}\frac{y-\xi_1}{x-\eta_1}\right)} \tag{24}$$

Without loss generality, let the Line-of-Sight (LOS) from the sensor to the target be the baseline direction of the coordinate systems. The optimal sensor moving direction based on (24) can then be simplified as

$$\begin{aligned} \varphi_{opt} &= \pm\arctan\left(\frac{|r^2-d^2|}{2rd}\right) \\ &= \pm\arctan\left(\frac{|\kappa^2-1|}{2\kappa}\right), \end{aligned} \tag{25}$$

which depends only on the parameter κ, where $\kappa = \frac{r}{d}$ denotes the ratio of r and d.

References

[1] Y. Bar-Shalom, and T. E. Fortmann. *Tracking and Data Association*, Academic Press, 1988.

[2] S. C. Nardone and V. J. Aidala. Observability criteria for bearings-only target motion analysis", *IEEE trans. Automatic Control*, vol. 17, no. 2, pp. 162–166, 1981.

[3] C. Jauffret and D. Pillon. "Observability in passive target motion analysis", *IEEE trans. Aerospace and Electronic Systems*, 32(4):1290–1300, 1996.

[4] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

[5] X. Wang, M. Morelande and B. Moran. "Sensor Scheduling for Bearings-Only Tracking with A Single Sensor", *Proceedings of The Fifth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 2009)*, Melbourne, Australia, pp. 67–72, 7-10, December 2009.

[6] C. R. Rao. "Information and accuracy attainable in the estimation of statistical parameters", *Bulletin of the Calcutta Mathematical Societ*, vol. 37, pp. 81–91, 1945.

[7] S. Amari. "Information geometry of statistical inference - an overview", *IEEE Information Theory Workshop*, Bangalore, India 2002.

[8] S. Amari. "Differential geometry of a parametric family of invertible linear systems-Riemannian metric, dual affine connections and divergence", *Mathematical Systems Theory*, vol. 20, pp. 53-82, 1987.

[9] Y. Cheng, X. Wang, and B. Moran. "Sensor network performance evaluation in statistical manifolds", *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July, 2010.

[10] K. M. Carter, R. Raich, W. G. Finn, and A. O. Hero. "FINE: fisher information nonparametric embedding", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 2093-2098, 2009.

[11] L. Pronzato and E. Walter. Robust experiment design via stochastic approximation. *Mathematical Biosciences*, 75(1):103–120, 1985.