

Sensor Network Performance Evaluation In Statistical Manifolds

Yongqiang Cheng[†], Xuezhi Wang^{*}, 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 – *Information geometry, as a powerful though complex mathematical tool can provide additional insights in the analysis of sensor measurement. In this paper, the application of information geometry to the analysis of sensor networks is explored in an attempt to gain a better understanding of sensor system issues for target detection and tracking. In particular, the (integrated) Fisher information distance between two targets is used to measure target resolvability in the region covered by the sensor system and is approximately calculated. It is also compared with the Kullback Leibler divergence. The proposed analysis is elucidated via two simple sensor network examples in the context of target tracking. The preliminary analysis results presented in this paper provide evidence that information geometry is able to offer consistent but more comprehensive means to understand and solve sensor network problems which are difficult to deal with via conventional analysis methods.*

Keywords: Sensor Networks, Fisher Information Distance, Target Tracking, Information Geometry, Kullback Leibler Divergence.

1 Introduction

Information geometry studies intrinsic properties of manifolds of probability distributions [1]. 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 affine connection. The manifold is proved to have a unique Riemannian metric given by the Fisher information matrix (FIM), and a dual pairs of affine connections [1].

As a powerful tool, information geometry offers comprehensive results about statistical models simply by considering them as geometrical objects.

Information geometry has found many applications in the asymptotic theory of statistical inference [2], semiparametric statistical inference [3], the study of Boltzmann machine [4], the Expectation Maximisation (EM) algorithm [5], and learning of neural networks [6], all with certain degree of success. In the last two decades, its application has spanned several discipline areas such as information theory [7, 8], systems theory [9, 10], mathematical programming [11], and statistical physics [12, 13]. It also played a central role in the multiterminal estimation theory [14]. In neuroscience it was used for extracting higher-order interactions among neurons [15]. Many researchers around the world are extending their work on this theory to a new level of applications and interpretations.

In this paper, the application of information geometry to the analysis of sensor networks is explored in an attempt to gain a better understanding of sensor system issues for target detection and tracking. In particular, the Fisher information distance (FID) between two targets is used to measure target resolvability in the region covered by the sensor system and is approximately calculated. It is also compared with the Kullback Leibler divergence (KLD). The proposed analysis method is elucidated via two simple sensor network examples in the context of target tracking. The preliminary analysis presented in this paper provides evidence that the information geometry theory is able to offer consistent but more comprehensive means to understand and solve sensor network problems which are difficult to deal with via conventional analysis methods.

In the next section, the mathematical background of the problem is stated. Principles of information geometry are described in Section 3. In Section 4, sensor net-

work information measured in FID on statistical manifolds is analyzed via two basic types of sensor network problems. Analytical results are discussed in Section 5 followed by the conclusions in Section 6.

2 Problem Formulation

Without loss generality, we will start our analysis by considering the problem of multi-target tracking in a sensor network where the state of interest $\boldsymbol{\theta}_k$ at time k consists of target positions and velocity components and the evolution of $\boldsymbol{\theta}$ is modeled as a Markov process subject to a random noise, i.e.

$$\boldsymbol{\theta}_k = f(\boldsymbol{\theta}_{k-1}) + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{Q}_k) \quad (1)$$

where f is the system dynamical model and the noise \mathbf{v}_k is assumed to be Gaussian distribution with zero mean and covariance \mathbf{Q}_k . The measurement of the system \mathbf{x}_k at time k is modeled as

$$\mathbf{x}_k = \boldsymbol{\mu}(\boldsymbol{\theta}_k) + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(0, \mathbf{C}_k) \quad (2)$$

where the measurement function $\boldsymbol{\mu}$ connects the measurement \mathbf{x}_k with target state $\boldsymbol{\theta}_k$, and \mathbf{w}_k is the measurement noise which is also assumed to be a zero mean Gaussian with covariance \mathbf{C}_k .

The multi-target tracking problem is to find an estimate of the posterior probability density of target current state based on a sequence of measurements. However, in this paper the underlying problem will be viewed from the sensor network design perspective, where we are more interested in the information gathering capacity of a measurement and in optimization of this over time.

In particular, it would be desirable to calculate the distance, i.e., FID, between two targets on the statistical manifolds and thus understand how the Euclidean distance differs from the information distance between two targets. It is also interesting to obtain insight about how the KLD is related to the FID. These ideas form the crux of the research work described in this paper.

3 Principle Of Information Geometry

3.1 Definition of Statistical Manifold

We consider the parameterized family of probability distributions $S = \{p(\mathbf{x}|\boldsymbol{\theta})\}$, where \mathbf{x} is a random variable and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ is a real vector parameter to specify a distribution [15]. The family is regarded as a statistical manifold with $\boldsymbol{\theta}$ as its natural coordinate system.

For a given state of interest $\boldsymbol{\theta}$ in the parameter space $\Theta \in \mathbb{R}^n$, the measurement \mathbf{x} in the sample space $\mathbf{X} \in \mathbb{R}^m$ corresponds to a probability distribution $p(\mathbf{x}|\boldsymbol{\theta})$. Each probability distribution $p(\mathbf{x}|\boldsymbol{\theta})$ corresponds a point in the manifold S . The parameterized

family of probability distributions $S = \{p(\mathbf{x}|\boldsymbol{\theta})\}$ forms an n -dimensional statistical manifold where $\boldsymbol{\theta}$ plays the role of a coordinate system of S .

3.2 Fisher Information Distance

For a parameterized family of probability distributions on a statistical manifold, Fisher information matrix (FIM) plays the role of a Riemannian metric tensor¹ [16]. Denoted by $\mathbf{I}(\boldsymbol{\theta}) = \{[\mathbf{I}(\boldsymbol{\theta})]_{ij}\}$, the FIM measures the amount of information the random variable \mathbf{x} tells about the parameter $\boldsymbol{\theta}$, i.e.,

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = E \left\{ \frac{\partial \log p(\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_i} \cdot \frac{\partial \log p(\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_j} \right\} \quad (3)$$

To provide the nature of operations on the manifold, the important concepts such as distance, angle and tangent are defined in an analogous way to the case of a Cartesian coordinate system. Information geometry permits the definition of a distance between statistical distributions that is invariant to non-singular parameterisation transformations [17]. The squared distance between two closely spaced distributions $p(\mathbf{x}|\boldsymbol{\theta})$ and $p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})$ is given by the quadratic form of $d\boldsymbol{\theta}$ [15],

$$ds^2 = \sum_{ij} [\mathbf{I}(\boldsymbol{\theta})]_{ij} d\theta_i d\theta_j = d\boldsymbol{\theta}^T \mathbf{I}(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (4)$$

The FID between two distributions $p(\mathbf{x}|\boldsymbol{\theta}_1)$ and $p(\mathbf{x}|\boldsymbol{\theta}_2)$ is defined by the following integral [18]

$$\mathcal{D}_F(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \triangleq \arg \min_{\boldsymbol{\theta}(t)} \int_{t_1}^{t_2} \left(\sqrt{\left(\frac{d\boldsymbol{\theta}}{dt} \right)^T \mathbf{I}(\boldsymbol{\theta}) \left(\frac{d\boldsymbol{\theta}}{dt} \right)} \right) dt \quad (5)$$

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

where $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ is the parameter path along the parameter space \mathbb{R}^n . Essentially, (5) amounts to finding the length of the shortest path, i.e. the geodesic, on the statistical manifold S connecting coordinates $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$.

In general, it is very difficult to find the geodesic on a statistical manifold. While the FID (5) is difficult to be exactly computed, the distance between two distributions $p(\mathbf{x}|\boldsymbol{\theta})$ and $p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})$ may be approximated with a variety of alternatives such as the Kullback-Leibler divergence [18],

$$\begin{aligned} KL[p(\mathbf{x}|\boldsymbol{\theta})||p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})] \\ &= \int p(\mathbf{x}|\boldsymbol{\theta}) \log \frac{p(\mathbf{x}|\boldsymbol{\theta})}{p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})} d\mathbf{x} \\ &= E \left\{ \log p(\mathbf{x}|\boldsymbol{\theta}) - \log p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta}) \right\} \quad (6) \end{aligned}$$

¹When it is non-singular at every parameter point.

It is well known that the following relationship between the KLD and differential Fisher information distance hold [15].

$$ds^2 = 2KL [p(\mathbf{x}|\boldsymbol{\theta})||p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})] \quad (7)$$

From the Taylor expansion around the distribution $p(\mathbf{x}|\boldsymbol{\theta})$, we have

$$\begin{aligned} \log p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta}) &= \log p(\mathbf{x}|\boldsymbol{\theta}) + d\boldsymbol{\theta}^T \nabla_{\boldsymbol{\theta}} [\log p(\mathbf{x}|\boldsymbol{\theta})] \\ &+ \frac{1}{2} d\boldsymbol{\theta}^T \nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T [\log p(\mathbf{x}|\boldsymbol{\theta})] d\boldsymbol{\theta} + O(|d\boldsymbol{\theta}|^3) \end{aligned} \quad (8)$$

where $\nabla_{\boldsymbol{\theta}} [\log p(\mathbf{x}|\boldsymbol{\theta})]$ denotes the gradient of $\log p(\mathbf{x}|\boldsymbol{\theta})$ and $\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T [\log p(\mathbf{x}|\boldsymbol{\theta})]$ the Hessian matrix, both evaluated at $\boldsymbol{\theta}$.

Note that the regular condition of the PDF [19] is

$$E \{ \nabla_{\boldsymbol{\theta}} [\log p(\mathbf{x}|\boldsymbol{\theta})] \} = 0, \quad \forall \boldsymbol{\theta} \in \Theta \quad (9)$$

and

$$E \{ \nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T [\log p(\mathbf{x}|\boldsymbol{\theta})] \} = -\mathbf{I}(\boldsymbol{\theta}) \quad (10)$$

Hence, in view of (8)–(10) the KLD (6) is given by

$$KL [p(\mathbf{x}|\boldsymbol{\theta})||p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})] = \frac{1}{2} d\boldsymbol{\theta}^T \mathbf{I}(\boldsymbol{\theta}) d\boldsymbol{\theta} + O(|d\boldsymbol{\theta}|^3) \quad (11)$$

As $p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta}) \rightarrow p(\mathbf{x}|\boldsymbol{\theta})$, (7) follows.

The KLD allows for the approximation of the information distance in the absence of the geometry of the statistical manifold, though it is not a genuine metric. A closed form KLD for Gaussian distributions is available in [20] and was used to compute KLD in this paper.

Concepts of tangent and orthogonality can be defined by analogy with the same concepts in a Euclidean space since they are local phenomena [15]. These topics are out of the scopes of this paper. Interested readers may refer to the references [21, 22, 23].

4 Application Examples In Sensor Networks

Two simple sensor network examples are given in this section to demonstrate the analysis of sensor networks by using statistical manifold techniques. As we mentioned earlier, it is not trivial to evaluate FID, we approximate this distance by calculating the integral of the differential FID (5) along a straight line (rather than the geodesic) between the locations of the two targets in the parameter space. This idea is illustrated in Fig. 1. We use this distance to measure the resolvability of closely spaced targets for a given sensor network.

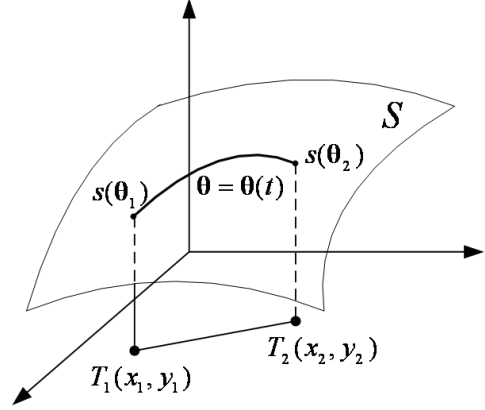


Figure 1: Geometry used for the approximation of the geodesic path.

4.1 Example 1: Network of Single Conventional Radar with Static Targets

In this example, the state of a target is simply represented by its location, i.e., $\boldsymbol{\theta} = [x, y]^T$. The sensor can observe both range and bearing of the target and the measurement model (2) is written as

$$\mathbf{x} = \begin{bmatrix} r \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \arctan(\frac{y}{x}) \end{bmatrix} + \begin{bmatrix} w_r \\ w_\phi \end{bmatrix} \quad (12)$$

where r and ϕ denote the range and bearing components of the measurement subject to an additive zero-mean Gaussian noise $\mathbf{w} = [w_r, w_\phi]^T$ with covariance $\mathbf{C}(\boldsymbol{\theta})$. Therefore, the measurement \mathbf{x} obeys a normal distribution,

$$\mathbf{x}|\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta})) \quad (13)$$

where

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \arctan(\frac{y}{x}) \end{bmatrix}, \quad \mathbf{C}(\boldsymbol{\theta}) = \begin{bmatrix} r^4 \sigma_r^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix} \quad (14)$$

In $\mathbf{C}(\boldsymbol{\theta})$, the term r^4 is appeared in the diagonal of range component to take into account the fact that the amplitude of the radar echo signal attenuates according to the fourth power of the target range, σ_r and σ_ϕ are the standard deviations of range and bearing measurement noise, respectively.

As shown in [24], the FIM $[\mathbf{I}(\boldsymbol{\theta})]_{ij}$ for the underlying measurement model (13) is of the form

$$\begin{aligned} [\mathbf{I}(\boldsymbol{\theta})]_{ij} &= \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \right]^T \mathbf{C}^{-1}(\boldsymbol{\theta}) \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} \right] \\ &+ \frac{1}{2} \text{tr} \left[\mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}^{-1}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} \right] \end{aligned} \quad (15)$$

Therefore, the FIM can be calculated and the squared differential FID in (4) is calculated as

$$\begin{aligned}
ds^2 &= \sum_{i,j=1}^2 [\mathbf{I}(\boldsymbol{\theta})]_{ij} d\theta_i d\theta_j \\
&= \frac{1}{(x^2 + y^2)^2} \left[\frac{x^2}{(x^2 + y^2)\sigma_r^2} + \frac{y^2}{\sigma_\phi^2} + 8x^2 \right] dx^2 \\
&+ \frac{1}{(x^2 + y^2)^2} \left[\frac{y^2}{(x^2 + y^2)\sigma_r^2} + \frac{x^2}{\sigma_\phi^2} + 8y^2 \right] dy^2 \\
&+ \frac{2}{(x^2 + y^2)^2} \left[\frac{xy}{(x^2 + y^2)\sigma_r^2} - \frac{xy}{\sigma_\phi^2} + 8xy \right] dx dy
\end{aligned} \tag{16}$$

The integral in (5) is done along the geodesic that connects the locations of the two targets. Since computing the geodesic is not trivial, the geodesic path as shown in Fig. 2 is approximated via a parametric equation $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ of a straight line between the two targets, i.e.,

$$\begin{aligned}
x &= x_1 + t \cos \alpha \\
y &= y_1 + t \sin \alpha
\end{aligned} \tag{17}$$

where the parameter t , $t \in [0, d]$, $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ represents the chord length from T_1 to T_2 , and as shown in Fig. 2 α is the inclination angle of this line. We have

$$\tan \alpha = \frac{y_2 - y_1}{x_2 - x_1} \tag{18}$$

Substituting (17) into (16) as a function of the param-

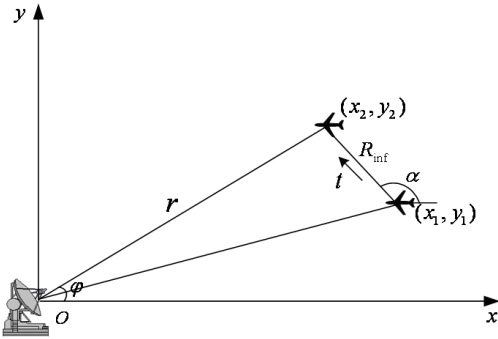


Figure 2: The parametric model for approximately computing the FID between two closely spaced targets.

eter t , we see that the squared differential FID becomes

$$\begin{aligned}
ds^2 &= \frac{dt^2}{A^2(t)} \left\{ \left[\frac{(x_1 + t \cos \alpha)^2}{A(t)\sigma_r^2} + \frac{(y_1 + t \sin \alpha)^2}{\sigma_\phi^2} \right. \right. \\
&+ 8(x_1 + t \cos \alpha)^2 \left. \right] \cos^2 \alpha + \left[\frac{(y_1 + t \sin \alpha)^2}{A(t)\sigma_r^2} \right. \\
&+ \frac{(x_1 + t \cos \alpha)^2}{\sigma_\phi^2} + 8(y_1 + t \sin \alpha)^2 \left. \right] \sin^2 \alpha \\
&+ \left[\frac{1}{A(t)\sigma_r^2} - \frac{1}{\sigma_\phi^2} + 8 \right] 2(x_1 + t \cos \alpha) \\
&\times (y_1 + t \sin \alpha) \sin \alpha \cos \alpha \left. \right\} \equiv \frac{B(t)}{A^2(t)} dt^2
\end{aligned} \tag{19}$$

where

$$A(t) = x_1^2 + y_1^2 + 2(x_1 \cos \alpha + y_1 \sin \alpha)t + t^2 \tag{20}$$

Therefore,

$$ds = \frac{\sqrt{B(t)}}{A(t)} dt \tag{21}$$

and the FID between the two targets T_1 and T_2 is approximated by the following integral

$$\mathcal{D} = \int_L ds = \int_0^d \frac{\sqrt{B(t)}}{A(t)} dt \tag{22}$$

Assume that the area of interest is 40×40 and $\sigma_r^2 = 1$, $\sigma_\phi^2 = 0.04$. The sensor is located at the origin of coordinates. The target T_1 is located at (20, 10).

Fig. 3 shows the information distance between two closely spaced targets T_1 and T_2 . In the calculation we fix the location of target T_1 and move target T_2 around the area centered at the location of T_1 with a circular radius of 15. A same plot is generated in Fig. 4 in KLD.

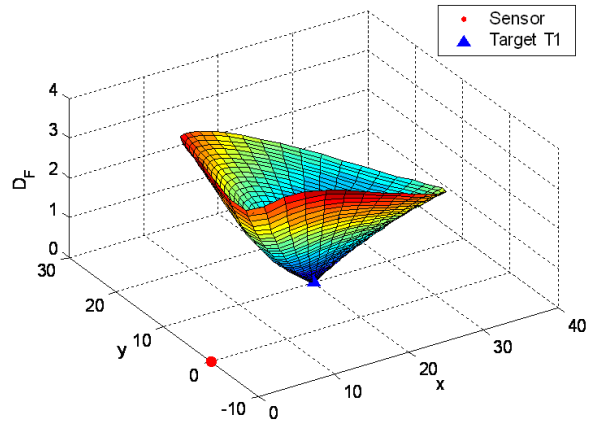


Figure 3: FID between two closely spaced targets.

Fig. 5 is the contour map of Fig. 3. It illustrates the sensing ability of the sensor network to distinguish two closely spaced targets. A minimum detectable information distance may be identified in the map. If

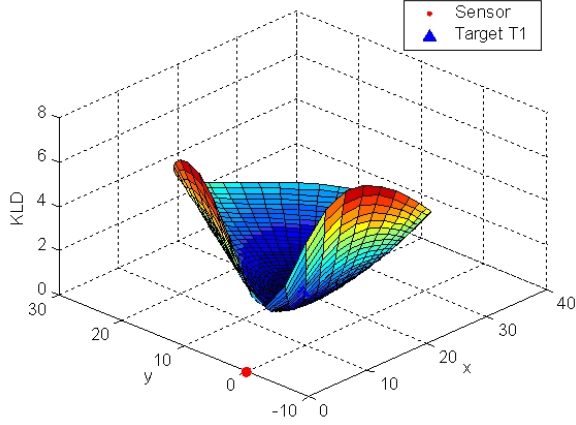


Figure 4: KLD between two closely spaced targets.

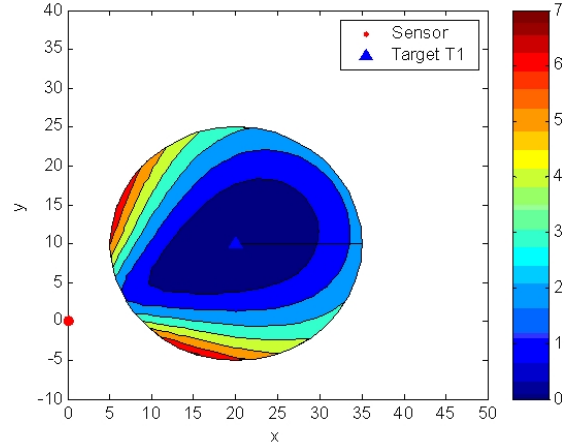


Figure 6: KLD between two closely spaced targets.

information distance between two targets is below the threshold, the two closely spaced targets may not be distinguished by the sensor system. The contour map

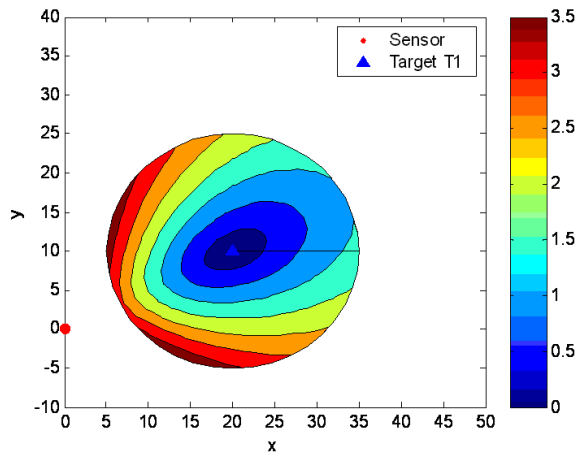


Figure 5: The contour map of the information distance.

generated via KLD in the same scenario as in Fig. 5 is given in Fig. 6.

4.2 Sensor network with range-only measurements from three sensors

Secondly, we consider an extended example of the target localization problem in a sensor network involving three range-only sensors. As shown in Fig. 7, these three distributed sensors are located at (η_i, ξ_i) , $i = 1, 2, 3$ and may only observe the range of a target subject to a random range noise. In this network configuration, the measurement model (2) is written as follows,

$$\mathbf{x} = \begin{bmatrix} \sqrt{(x - \eta_1)^2 + (y - \xi_1)^2} \\ \sqrt{(x - \eta_2)^2 + (y - \xi_2)^2} \\ \sqrt{(x - \eta_3)^2 + (y - \xi_3)^2} \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad (23)$$

where $\mathbf{w} = [w_1, w_2, w_3]^T$ is an additive zero-mean Gaussian noise with covariance $\mathbf{C}(\boldsymbol{\theta})$. As in the first exam-

ple, the likelihood function is also described by Equation (13), where

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \begin{bmatrix} \sqrt{(x - \eta_1)^2 + (y - \xi_1)^2} \\ \sqrt{(x - \eta_2)^2 + (y - \xi_2)^2} \\ \sqrt{(x - \eta_3)^2 + (y - \xi_3)^2} \end{bmatrix}, \quad (24)$$

$$\mathbf{C}(\boldsymbol{\theta}) = \text{diag} [r_1^4 \sigma_1^2, r_2^4 \sigma_2^2, r_3^4 \sigma_3^2]$$

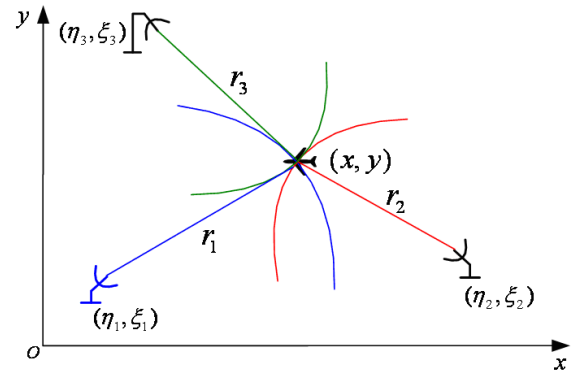


Figure 7: Example of sensor network of three range only sensors for target localisation

As before, the FIM for this measurement model can be derived using (15) and the squared differential information distance ds^2 between two target measurements $p(\mathbf{x}|\boldsymbol{\theta})$ and $p(\mathbf{x}|\boldsymbol{\theta} + d\boldsymbol{\theta})$ can be written in the following quadratic form of $d\boldsymbol{\theta}$,

$$ds^2 = \sum_{ij} [\mathbf{I}(\boldsymbol{\theta})]_{ij} d\theta_i d\theta_j$$

$$= \sum_{i=1}^3 \left[\frac{1}{r_i^2 \sigma_i^2} + 8 \right] \left[\frac{x - \eta_i}{r_i^2} dx + \frac{y - \xi_i}{r_i^2} dy \right]^2 \quad (25)$$

and it is parameterised as a function of t using (17):

$$ds^2 = \mathcal{F}(t)dt^2 = dt^2 \sum_{i=1}^3 \left\{ \left[\frac{1}{A_i(t)\sigma_i^2} + 8 \right] \times \left[\frac{x_1 - \eta_i + t \cos \alpha}{A_i(t)} \cos \alpha + \frac{y_1 - \xi_i + t \sin \alpha}{A_i(t)} \sin \alpha \right]^2 \right\} \quad (26)$$

where

$$A_i(t) = (x_1 - \eta_i + t \cos \alpha)^2 + (y_1 - \xi_i + t \sin \alpha)^2, \quad i = 1, 2, 3 \quad (27)$$

Then the approximate FID between two closely spaced targets T_1 and T_2 is the following integral

$$\mathcal{D} = \int_0^d \sqrt{\mathcal{F}(t)} dt \quad (28)$$

Figures 8 and 9 show both a 3D plot and its contour map for the FID between two targets where three range only sensors were used for taking target measurement. Clearly, in this case the uncertainty area where the value of FID below a threshold becomes isolated and target T_2 outside this area can be distinguished from target T_1 .

In the calculations, we assume that the three sensors were located at $(0, 0)$, $(15, 30)$ and $(50, 10)$ and $\sigma_1 = 1, \sigma_2 = 2$, and $\sigma_3 = 3$, respectively. Target T_1 was located at $(20, 10)$ on the location plane.

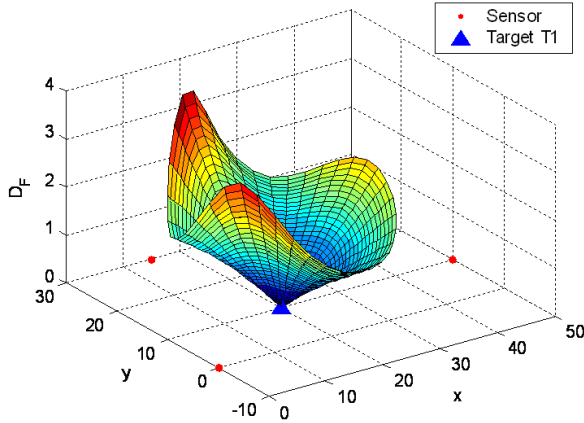


Figure 8: The information distance between targets T_1 and T_2 for the three range-only sensor network.

5 Discussions

In Section 4, the analysis of sensor networks on the statistical manifolds is demonstrated via two basic forms of sensor network scenarios. In contrast to the deterministic case where measurements are real-valued feature vectors in the Euclidean space, we consider the stochastic case where measurements can be represented as the points on a statistical manifold, i.e., a manifold of

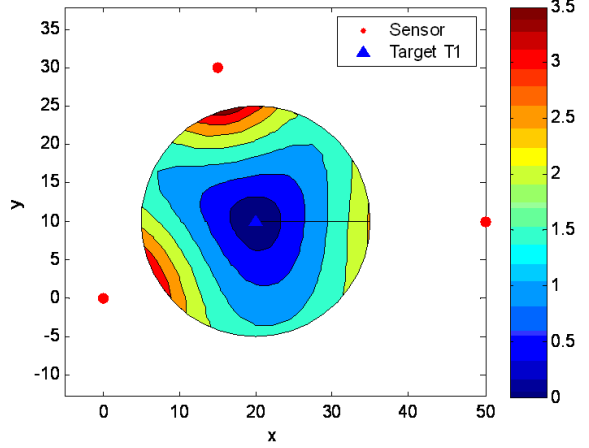


Figure 9: The contour map of the information distance plotted in Figure 8.

PDFs. In other words, it should be realised that both targets and their measurements are the realizations of some PDFs which lie on a statistical manifold. Using the properties of information geometry and statistical manifolds, the resolvability of two closely spaced targets is graphically illustrated. Further discussions and observations for the analysis results are presented below.

1. When the parameters to be estimated are considered as the points on a statistical manifold, information geometry elucidates the hierarchical structures of random variable reference to geometric concepts such as distance, tangent as well as orthogonality between probability distributions. In this paper, the FID on the statistical manifolds is calculated and used to measure the information difference between two targets for a given sensor system. Demonstrated in Figures 3, and 8, the FID between two targets describes the resolvability in the presence of multi-target using network sensors, which provides more informative than the Euclidean distance in the context of target detection.
2. On the other hands, the FID may be regarded as a description of the information associated with measurement uncertainty. The information acquired by sensor networks determines the estimation accuracy of target states. The contour maps of the FID presented in our examples such as Figures 5, 9, reflect the estimation accuracy of target states (positions) which can be achieved by an unbiased estimator. In practice, a threshold corresponding to the minimum detectable information may be set. Two closely spaced targets will not be resolved with the measurement by the sensor system alone if the FID is below the threshold. Fig.

10 illustrates several location regions for the three range-only sensor network. All regions represent identical FID between two targets with one target located at the center of the region and the other at the edge of the region. It indicated that the best achievable estimation accuracy is quite different when targets are at different locations across the surveillance area of interest.

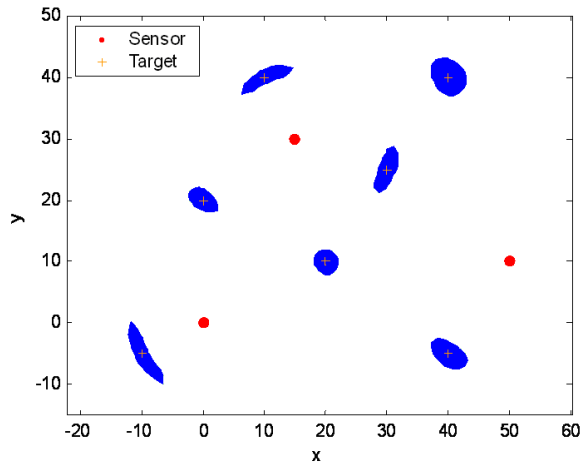


Figure 10: Target localisation accuracy at different locations in terms of FID.

3. In both sensor network examples, the computed FID was found consistent with KLD measure when two targets are reasonably close. Therefore, the KLD plots for the second sensor network scenario were omitted.

6 Conclusions

In this paper, the potential application of information geometry in the analysis and design of networked sensor systems is highlighted via two basic types of sensor systems for target tracking, where the Fisher information distance is used to measure the information difference between two targets and is approximately computed. It is found that the calculated information distance is consistent with the KLD measure when the two targets are reasonably close. The initial analysis results demonstrated that the projection of sensor measurements onto the statistical manifolds provides additional judgement to target resolvability issue for a given sensor system. Indeed, information geometry offers a powerful tool for sensor network design, evaluation and re-configuration.

It should be pointed out that the proposed approximation for the calculation of FID is only valid for closely spaced targets and the exact FID must be integrated along the geodesic connecting the two target states. The work is currently undertaken and will be reported elsewhere.

References

- [1] S. Amari. "Information geometry of statistical inference - an overview", *IEEE Information Theory Workshop*, Bangalore, India 2002.
- [2] R. E. Kass and P. W. Vos. "Geometrical Foundations of Asymptotic Inference", New York: Wiley, 1997.
- [3] S. Amari and M. Kawanabe. "Information geometry of estimating functions in semi parametric statistical models", *Bernoulli*, vol. 3, pp. 29-54, 1997.
- [4] S. Amari, K. Kurata, and H. Nagaoka. "Information geometry of Boltzmann machines", *IEEE Transactions on Neural Networks*, vol. 3, pp. 260-271, 1992.
- [5] S. Amari. "Information geometry of the EM and em Algorithms for neural networks", *Neural Networks*, vol. 8, pp. 1379-1408, 1995.
- [6] S. Amari. "Natural gradient works efficiently in learning", *Neural Computation*, vol. 10, pp. 251-276, 1998.
- [7] S. Amari. "Fisher information under restriction of Shannon information", *Annals of the Institute of Statistical Mathematics*, vol. 41, pp. 623-648, 1989.
- [8] L. L. Campbell. "The relation between information theory and the differential geometry approach to statistics", *Information Sciences*, vol. 35, pp. 199-210, 1985.
- [9] 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.
- [10] A. Ohara, N. Suda, and S. Amari. "Dualistic differential geometry of positive definite matrices and its applications to related problems", *Linear Algebra And Its Applications*, vol. 247, 1996.
- [11] A. Ohara. "Information geometric analysis of an interior point method for semidefinite programming", *Geometry in Present Day Science*, pp. 49-74, 1999.
- [12] S. Amari, S. Ikeda, and H. Shimokawa. "Information geometry of α -projection in mean-field approximation", in *Recent Developments of Mean Field Approximation*, M. Opper and D. Saad, Eds. Cambridge, MA: MIT Press 2000.
- [13] T. Tanaka. "Information geometry of mean field approximation", *Neural Computation*, vol. 12, pp. 1951-1968, 2000.

- [14] S. Amari and T. S. Han. “Statistical inference under multiterminal rate restrictions: a differential geometric approach”, *IEEE Transactions on Information Theory*, vol. 35, pp. 217-227 1989.
- [15] S. Amari. “Information geometry on hierarchy of probability distributions”, *IEEE Transactions on Information Theory*, vol. 47, pp. 1701-1711, 2001.
- [16] C. R. Rao. “Information and accuracy attainable in the estimation of statistical parameters”, *Bulletin of the Calcutta Mathematical Society*, vol. 37, pp. 81–91, 1945.
- [17] F. Barbaresco. “Innovative tools for radar signal processing Based on Cartan’s geometry of SPD matrices & information geometry”, in *2008 IEEE Radar Conference*, Rome, Italy, 2008.
- [18] 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.
- [19] S. M. Kay. “*Fundamentals of Statistical Signal Processing: Volume I: Estimation Theory*”, Prentice Hall, PTR, 1993.
- [20] S. Kullback. *Information Theory and Statistics*, Dover Publications, New York, 1968.
- [21] S. Amari. “*Differential-Geometrical Methods of Statistics (Lecture Notes in Statistics)*”, Berlin, Germany: Springer, 1985.
- [22] S. Amari and H. Nagaoka. “*Methods of Information Geometry*”, New York: AMS and Oxford Univ. Press, 2000.
- [23] H. Nagaoka and S. Amari. “Differential geometry of smooth families of probability distributions”, Univ. Tokyo, Tokyo, Japan, METR 1982.
- [24] X. Wang, Y. Cheng and B. Moran. “Bearings-only tracking analysis via information geometry”, *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July, 2010.