HAL open science

A note on the intrinsic Cramer-Rao bound

Axel Barrau, Silvère Bonnabel

To cite this version:

Axel Barrau, Silvère Bonnabel. A note on the intrinsic Cramer-Rao bound. 2013. hal-00828037v2

HAL Id: hal-00828037

https://hal.science/hal-00828037v2

Preprint submitted on 31 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A note on the intrinsic Cramer-Rao bound

Axel Barrau and Silvère Bonnabel

Centre de Robotique, Mathématiques et Systèmes, Mines ParisTech, 60 bd Saint-Michel, 75272 Paris cedex 06, [axel.barrau,silvere.bonnabel]@mines-paristech.fr

Abstract. We consider the intrinsic version of the Cramer-Rao lower bound (CRLB) as introduced by S.T. Smith in 2005. In the concerned paper, the derived lower bound on the intrinsic root-mean-square error (RMSE) of any sample covariance matrix (SCM) estimator is shown not to depend on the underlying parameter, and the author claims the result stems from the invariances of the Fisher Information Metric (FIM). But it also stems from the use of special coordinates used to derive the result. The goal of this paper is to address the following questions: 1- to what extent is the *intrinsic* CRLB bound independent of a specific choice of coordinates ? 2- when can the intrinsic CRLB be expected not to depend on the underlying parameter ? The paper is essentially tutorial and can be considered as an introduction to the intrinsic CRLB.

1 Introduction

The theory of statistical estimation aims at estimating parameters from measured or observed data, these parameters underlying the distribution of the data. In the present paper, we focus on data that are assumed to be random with a parametric probability distribution $p(x|\theta)$. In numerous applications, e.g. in signal processing, the space in which the parameter θ lives is rather a manifold than an Euclidian space. A typical example concerns the estimation of covariance matrices for normal multivariate distributions having known mean, where the set of parameters to be estimated is the cone of positive semi-definite matrices.

When the parameter belongs to a manifold, the classical Cramer-Rao analysis, that quantifies the expectable maximal accuracy of any estimator, depends upon a specific choice of coordinates. Yet, it seems reasonable to expect that the maximal accuracy an estimator can reach depends only upon the distribution $p(x|\theta)$, independently from the coordinates chosen by the user to parametrize θ. The intuitively appealing field of *intrinsic estimation* (see e.g. [3] and references therein), and more generally information geometry [1] allows to develop Cramer-Rao bounds that quantify the goodness of an estimator through intrinsic tools. For more information on Cramer-Rao analysis on manifolds see also [6] and the long list of references therein.

In the present paper, we consider the intrinsic version of the Cramer-Rao lower bound (CRLB) as introduced by S.T. Smith in [6]. In this latter paper, the author proposes a derivation of the intrinsic CRLB, and presents the results in a way that is relevant to the field of signal processing. In particular, the author proves the following striking result: intrinsic lower bound for the root-mean-square error (RMSE) of any sample covariance matrix (SCM) estimator is shown not to depend on the underlying parameter, and the author claims the result stems from the invariances of the Fisher Information Metric (FIM). But it can be noted it also stems from the use of special coordinates used to derive the result. The goal of the present paper is to clearly answer the following questions: 1- to what extent is the *intrinsic* CRLB bound independent of a specific choice of coordinates ? 2- when can the intrinsic CRLB be expected not to depend on the underlying parameter ?

The paper is essentially tutorial, and only deals with the unbiased case to simplify exposure. It consists of some remarks on intrinsic Cramer-Rao bounds, and can be viewed as an

introduction to the paper [6]. In particular, we show that the results contained in [6] allow to derive in a straightforward way the fact that intrinsic RMSE can *always* be expected not to depend on the underlying parameter. This is a worthy to note result, which is not explicitely stated in [6]. This is thus no surprise it is the case in the SCM estimator problem, independently from the invariances of the FIM. Note that, this result is in fact already known from [3], and thus here a link between the two latter papers is established. Moreover, we provide several other examples to illustrate this property.

The paper is organized as follows. Section 2 is mainly expository, and recalls a few facts of classical estimation theory, and introduces the notion of intrinsic CRLB. Section 3 deals with the dependance of the intrinsic Cramer-Rao bound with respect to a choice of local coordinates, and simply proves using normal coordinates that the lower bound of the intrinsic RMSE can always be expected to be constant, as long as curvature terms are neglected. Section 4 contains several examples, and in particular revisits the SCM estimation problem of [6].

2 A few preliminaries on the Intrinsic Cramer-Rao bounds

In this section, we recall a few facts of classical estimation theory, namely the notion of Fisher Information, and the well-known Cramer-Rao lower bound on the minimal dispersion an estimator can achieve. The notion of intrinsic estimation and intrinsic CRLB is also introduced. Throughout this article, we will focus on the unbiased case, in order to simplify the exposure and the calculations.

2.1 The Fisher Information Matrix and the Cramer-Rao lower bound in classical estimation theory

The measurement of a random variable $X \in \mathbb{R}^p$ produces a result x. The random variable X is assumed to admit a probability density function (pdf) denoted by $p(x|\theta)$ which depends on a paramater $\theta \in \mathbb{R}^n$. The goal of any estimator $\hat{\theta} : \mathbb{R}^p \to \mathbb{R}^n$ is to provide an estimate $\hat{\theta}(x)$ of the true paramter θ from the measurement x. As the measure is random, an estimator cannot be expected to give the true value of the parameter θ. If the estimator takes its values in a vetor space we define its bias $b_{\hat{\theta}}$ as $\mathbb{E}(\hat{\theta}) - \theta$. An estimator is said to be "unbiased" if the property $b_{\hat{\theta}} = 0$ is verified. It means that the retuned value is on average the true value of the parameter. A usual way to measure the accuracy of an unbiased estimator (at least in the vectorial case) is to consider its covariance matrix $C_\theta = \mathbb{E}((\hat{\theta}-\theta)(\hat{\theta}-\theta)^T)$. A natural question about estimators is to know if there is a limit to their precision. The answer is of course yes. For instance, it can not be expected to determine the variance of a Gaussian random variable using only one realization. Actually, a finite number of samples never allows to give an exact estimation. A quantification of this lack of information through the covariance matrix is given by the Cramer-Rao Lower Bound for unbiased estimators :

$$C_\theta \geqslant \mathcal{I}^{-1}(\theta) \tag{1}$$

Where the symbol $\geqslant$ means that $C_\theta - \mathcal{I}^{-1}(\theta)$ is semi-definite positive and $\mathcal{I}(\theta)$ designates the Fisher information matrix having for a given parameter θ, entries defined by:

$$\mathcal{I}_{i,j}(\theta) = \mathbb{E}(\frac{\partial}{\partial \theta_i} \log p(x|\theta) \frac{\partial}{\partial \theta_j} \log p(x|\theta))$$

A useful property of the Fisher Information is additivity. Indeed let $\mathcal{I}_K(\theta)$ denote the Fisher information matrix corresponding to K i.i.d. realizations of the random variable X. Then we have:

$$\mathcal{I}_K(\theta) = K\mathcal{I}_1(\theta)$$

As a trivial consequence, if we use K measurements, the Cramer-Rao bound decreases in $\frac{1}{K}$. In the sequel, we can thus omit this $\frac{1}{K}$ factor, and we will only consider situations with one observation. Finally, an alternative definition of the Fisher Matrix information has to be mentioned. If $p(x|\theta)$ has sufficient regularity we have :

$$\mathcal{I}_{i,j}(\theta) = -\mathbb{E}(\frac{\partial^2 \log p(x|\theta)}{\partial\theta_i \partial\theta_j}) \tag{2}$$

2.2 Derivation of the Cramer-Rao result in dimension 1

If the dimension of the parameter space is one, results (1) and (2) can be easily derived. We will mention the proof because of its simplicity. When the dimension is more than 1, the results can still be derived in the same way.

Cramer-Rao bound Consider an unbiased estimator θ for a parametric family $p(x|\theta)$. We have :

$$\int_x (\hat{\theta}(x) - \theta) p(x|\theta) d\theta = 0$$

Differentiating with respect to θ gives :

$$\int_x (\hat{\theta}(x) - \theta) \frac{\partial}{\partial\theta} p(x|\theta) d\theta - \int_x p(x|\theta) = 0$$

i.e.

$$\int_x (\hat{\theta}(x) - \theta) \frac{\partial}{\partial\theta} p(x|\theta) d\theta = \int_x (\hat{\theta}(x) - \theta) \frac{\partial}{\partial\theta} \log p(x|\theta) p(x|\theta) d\theta = 1$$

The Cauchy-Schwarz inequality gives then :

$$\int_x (\hat{\theta}(x) - \theta)^2 p(x|\theta) dx \int_x (\frac{\partial}{\partial\theta} \log p(x|\theta))^2 p(x|\theta) dx \geqslant 1$$

yielding the result $C_\theta \geqslant \mathcal{I}^{-1}(\theta)$.

Alternative definiton of the FIM The result (2) is obtained by a straightforward calculation:

$$\begin{aligned}
\int_x (\frac{\partial^2}{\partial\theta^2} \log p(x|\theta)) p(x|\theta) dx &= \int_x \frac{\partial^2}{\partial\theta^2} p(x|\theta) - \frac{1}{p(x|\theta)} (\frac{\partial}{\partial\theta} p(x|\theta))^2 dx \\
&= \frac{\partial^2}{\partial\theta^2} \int_x p(x|\theta) - \int_x (\frac{\partial}{\partial\theta} \log p(x|\theta))^2 p(x|\theta) dx \\
&= -\mathcal{I}(\theta)
\end{aligned}$$

2.3 Estimation on Riemannian manifolds

Suppose now, that the parameter θ that one seeks to estimate belongs to a n-dimensional smooth manifold $\mathcal{M}$. A typical example is the case of multivariate Gaussian distributions having null mean $X \sim \mathcal{N}(0, \Sigma)$ whose pdf is

$$\frac{1}{(2\pi|\Sigma|^{\frac{n}{2}})} \exp[-\frac{1}{2} x^T \Sigma^{-1} x]$$

where $x \in \mathbb{R}^p$ and where the parameter Σ is the covariance matrix of the variable X, that lives in the cone of positive definite matrices, which is a manifold, but not a vector space. We assume here we have a *particular* parameterization $(\theta_1, ..., \theta_n)$ of the manifold, i.e. a coordinate chart, and thus of the tangent space at any value of the parameter $\theta \in \mathcal{M}$. We also assume the exponential map from this space to $\mathcal{M}$ is globally defined and everywhere invertible. In order to derive an intrinsic CRLB on the manifold, the two following concepts from intrinsic estimation are needed. Indeed, we introduce a definition of the bias of an estimator on a manifold :

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}[\exp_\theta^{-1}(\hat{\theta})]$$

If θ is an estimator with zero bias, its covariance matrix C_θ is defined as :

$$C_\theta = \mathbb{E}[\exp_\theta^{-1}(\hat{\theta}) \exp_\theta^{-1}(\hat{\theta})^T]$$

In fact the exponential is generally not everywhere invertible, and the previously introduced notions of bias and covariances in the tangent space rely on several assumptions. The interested reader is referred to the tutorial paper [5] for more information. For intrinsic means on Riemannian manifold, the reader is referred to the seminal paper [4].

2.4 Intrinsic Cramer-Rao lower bound

The intrinsic Fisher Information Matrix *associated to a given coordinate chart* on the parameter space is a matrix whose entries are defined by:

$$G_{\theta,ij} = \mathbb{E}[(\frac{\partial}{\partial \theta_i} \log p(x|\theta))(\frac{\partial}{\partial \theta_j} \log p(x|\theta))] \tag{3}$$

Without bias, the result of [6] reads in those given coordinates

$$C_\theta \succeq G_\theta^{-1} + \text{curvature terms} \tag{4}$$

In the applications, the curvature terms are generally neglected assuming that the curvature is small in the range of the estimator, and the result simply becomes the classical Cramer-Rao result:

$$C_\theta \succeq G_\theta^{-1}$$

As this is a matrix inequality based on a specific choice of local coordinates, one can wonder to what extent the result is intrinsic. The next section deals with this issue.

3 To what extent does the Intrinsic CRLB depend on a choice of coordinates ?

3.1 The Fisher Information Matrix is a covariant tensor on $\mathcal{M}$

As a key result in information geometry, equation (3) actually defines a covariant tensor which could have been expressed in intrinsic terms as :

$$G_\theta(\Omega_1, \Omega_2) = \mathbb{E}(\frac{d}{dt}[\log p(x| \exp_\theta(t\Omega_1))]_{t=0} \frac{d}{dt}[\log p(x| \exp_\theta(t\Omega_2))]_{t=0})$$

Of course this expression does not depend on any choice of coordinates. Indeed, the entries of the Fisher matrix G_θ transform covariantly under changes of coordinates. Indeed, consider

a change of coordinates $\phi : U \to V$ where U, V are coordinate charts on the tangent space at θ. Then the Fisher matrix is transformed via the equation

$$G_\theta = J_\phi^T(\theta) G_{\phi(\theta)} J_\phi(\theta) \tag{5}$$

where J_ϕ denotes the Jacobian matrix of the transformation. This is the reason why the Fisher information matrix provides a Riemannian metric, the so-called Fisher metric. When the parameter space is endowed with this metric, its geometric properties are intrinsic, i.e., unchanged under different parametrizations of the space.

3.2 Choice of normal coordinates

As already said before, the Fisher matrix is (generally) positive definite, and transforms covariantly with respect to changes of coordinates, and thus defines a Riemannian metric on the parameter space. Locally, one can always define normal coordinates through the exponential map. Consider an orthonormal basis of the tangent space with respect to the metric. Then, the exponential map allows to map the tangent space onto the manifold which in turn inherits of this orthonormal coordinate system. Of course, when expressed in normal coordinates, the tensor metric is the identity matrix. As a result, if $(\theta_1, ..., \theta_n)$ is a set of normal coordinates for the FIM, and if $Id \in \mathbb{R}^{n \times n}$ denotes the identity matrix, the result of [6] simply becomes

$$C_\theta \succeq Id + \text{curvature terms} \tag{6}$$

and neglecting the curvature terms, the intrinsic CRLB finally writes:

$$C_\theta \succeq Id$$

Thus, we see that using a very special choice of coordinates on the manifold (and neglecting curvature additional terms), the intrinsic CRLB becomes *totally independent* of the underlying parameter. This result is particularly striking and worthy to note, though very intuitive and maybe obvious to the readers that are familiar with information geometry. Indeed, choosing normal coordinates for the FIM amounts to choosing coordinates based on the Cramer-Rao bound, i.e., parameters that truly account for the statistical discrepancy between nearby parameters in the following way: the easier to distinguish between two parameters, the larger their distance. Thus, in this geometry dictated by the statistical properties of the model, it is no surprise the minimal dispersion an estimator can achieve around a given value is a fixed quantity, which is the same in any direction.

Finally, note that, from a computational viewpoint, starting from a particular coordinate system, and finding (local) normal coordinates amounts to finding the transformation J_ϕ in equation (5) that allows to diagonalize and then scale the Fisher Matrix G_θ so that it is transformed into the identity matrix.

3.3 Consequences on the intrinsic Root Mean Square error

Define the intrinsic Root Mean Square error (IRMSE) associated to a given estimator $\hat{\theta}$ of a parameter $\theta \in \mathcal{M}$, as its average dispersion :

$$\epsilon^2 := \mathbb{E}(d(\theta, \hat{\theta})^2)$$

where d denotes the geodesic distance associated to the Fisher Information Metric (see the Appendix). This quantity is also known as the Riemannian risk [3], as it is the risk associated to the squared geodesic distance as a loss function. But, because of the definition of the exponential map, we have $d(\theta, \hat{\theta}) = \|\exp_\theta^{-1}(\hat{\theta})\|$. The following result is thus a straightforward consequence of the remarks above:

Proposition 1. *Neglecting curvature terms in* (4) *we have*

$$\epsilon^2 = \mathbb{E}[d^2(\theta, \hat{\theta})] = \mathbb{E}(\|\exp_\theta^{-1}(\hat{\theta})\|^2) \geq n \tag{7}$$

It states that the IRMSE is lower bounded by a fixed quantity, which is completely independent of a specific choice of coordinates used to do the calculations. We will from now refer to this quantity as the *lower bound of the IRMSE*. The proposition is easy to prove as follows. If one uses normal coordinates, we have :

$$\mathbb{E}(\|\exp_\theta^{-1}(\hat{\theta})\|^2) = \mathbb{E}(\exp_\theta^{-1}(\hat{\theta})^T \exp_\theta^{-1}(\hat{\theta})) = \mathrm{Tr}(\mathbb{E}(\exp_\theta^{-1}(\hat{\theta}) \exp_\theta^{-1}(\hat{\theta})^T)) = \mathrm{Tr}(C_\theta)$$

But in normal coordinates inequality (4) boils down to (6) and taking the trace of each member gives the result.

This result is already known from [3], where all the calculations are done including the curvature terms, and using other means. In particular, it is shown that when the sectional curvature is everywhere negative, the result of the proposition above is valid even without neglecting curvature terms. As a consquence, we see that in the examples, as long as one wishes to calculate the lower bound of the IRMSE, no calculations are needed, as this latter quantity can always be expected to be constant and equal to the dimension of the manifold. The following section presents a few example applications of this principle.

4 Examples

4.1 Location parameter estimation in dimension 1

As a first very simple example, we consider here the families of the form $p(x|\theta) = f(x-\theta)$. We have $\mathcal{I}(\theta+a) = \int_x (\frac{\partial}{\partial\theta} f(x-\theta-a))^2 = \int_y (\frac{\partial}{\partial\theta} f(y-\theta))^2 = \mathcal{I}(\theta)$ where we used the variable change $y = x - a$. As a consequence, normal coordinates are a linear function of θ. Consider the example of Gaussians of mean μ unknown and variance σ^2 known. We have

$$I(\mu) = \frac{1}{\sigma^2}$$

and normal coordinates are given by the dimensionless mean $\eta = \frac{\mu}{\sigma}$. Thus, for location parameter estimation problems, the ICRLB coincides, up to a scale factor, with the conventional CRLB, and the lower bound of the IRMSE (or Riemannian risk), as well as of the classical RMSE, can be expected to be constant over the whole parameter space.

4.2 Scale parameter estimation in dimension 1

Another standard class of parametric families consists of pdf's of the form $p(x|\theta) = \frac{1}{\theta} f(\frac{x}{\theta})$ with x and θ in $\mathbb{R}$. The variable change $y = \log(x), \chi = \log(\theta)$ returns us to the previous case:

$$p(y|\chi) = e^{y-\chi} f(y-\chi) = g(y-\chi)$$

Where $g(z) = \exp(z) f(z)$. The normal coordinates associated to these families are thus proportional to $\chi = \log(\theta)$. Using (7) we can say without any calculation that we can find a bound of the form:

$$\mathbb{E}[\log(\frac{\hat{\theta}}{\theta})^2] \succeq \alpha$$

Where α does not depend on θ.

4.3 SCM in dimension n

Consider now the family of centered multivariate normal laws in dimension $n > 1$:

$$p(x|\Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp(-\frac{1}{2}x^T \Sigma^{-1} x)$$

Following [6], the Fisher Information Metric can be derived through a second-order Taylor expansion. Let D be a small symmetric matrix. We have :

$$\begin{aligned}
\mathbb{E}[\log p(x|\Sigma + D)] &= \mathbb{E}[-\frac{1}{2}\log(|\Sigma + D|) - \frac{1}{2}x^T(\Sigma + D)^{-1}x] \\
&= -\frac{1}{2}\log(|\Sigma + D|) - \frac{1}{2}\mathrm{Tr}((\Sigma + D)^{-1}\Sigma) \\
&= -\frac{1}{2}\log(|\Sigma|) - \frac{1}{2}\mathrm{Tr}(I) - \frac{1}{2}\mathrm{Tr}(D\Sigma^{-1}) - \frac{1}{2}\mathrm{Tr}(D\Sigma^{-1})
\end{aligned}$$

We get finally :

$$G_\Sigma(D, D) = \mathrm{Tr}(D\Sigma^{-1})^2$$

Up to a factor 2, we obtain the natural metric of the cone of positive definite matrices (see e.g. [2]), whose geodesic distance is denoted by d_{cov} in [6]. Now, the inequalities found by S.T. Smith in Theorem 4 of [6] directly follow from equation (7), and require no further calculations than having derived the Fisher Metric:

$$\epsilon^2 = \mathbb{E}(d^2_{cov}(\hat{\Sigma}, \Sigma)) = 2\ \mathbb{E}(d^2_{FIM}(\hat{\Sigma}, \Sigma)) \geqslant n(n+1)$$

4.4 Some further remarks on the invariance properties of the FIM in the SCM estimation problem

In [6], the author claims the fact that the lower bound of the IRMSE does not depend on the underlying parameter stems from the invariance of the metric to invertible transformations by congruence (that is invariance under a finite dimensional Lie group action, not to be confused with invariance of the FIM to reparametrizations). But as shown by Proposition 1, this is to be always expected, and this independently from the invariance properties of the Fisher Metric. However, one must admit the invariance of the FIM to invertible transformations is another way to explain the result. Indeed, the ICRB inherits the invariances of the estimation problem. To specify the sense of this property, we introduce a transformation group G and two additionnal definitions. Let ρ_g be an action of G on $\mathcal{M}$ and ϕ_g an action of G on X.

Definition 1 *The parametric family $p(x|\theta)$, is told invariant under the action of G if we have :*

$$p(\phi_g^{-1}(x)|\theta)|J_{\phi_g^{-1}}| = p(x|\rho_g(\theta))$$

Where $J_{\phi_g^{-1}}$ designates the jacobian of the function $x \to \phi_g^{-1}(x)$.

Definition 2 *A quadratic form Q on the manifold $\mathcal{M}$ is told constant under an action ψ_g of G if for any $\theta \in \mathcal{M}$, $g \in G$ and $\Omega_1, \Omega_2 \in T_\theta$ (tangent plane to θ) we have:*

$$Q_{\psi_g(\theta)}(D\psi_g\Omega_1, D\psi_g\Omega_2) = Q_\theta(\Omega_1, \Omega_2)$$

Proposition 2. *If $p(x|\theta)$ is invariant under the actions ρ_g and ϕ_g, the Fisher matrix information C_θ is constant as a bilinear function under the action ρ_g.*

Proof. For $\theta \in \mathcal{M}$, $g \in G$ and $\Omega_1, \Omega_2 \in T_\theta$, $G_{\rho_g(\theta)}(D\rho_g\Omega_1, D\rho_g\Omega_2)$ is equal to :

$$\int_x \frac{d}{dt}_{t=0} \log p(x|\exp_{\rho_g(\theta)}(tD_{\rho_g}\Omega_1))\frac{d}{dt}_{t=0} \log p(x|\exp_{\rho_g(\theta)}(tD_{\rho_g}\Omega_2))p(x|\rho_g(\theta))dx$$
$$= \int_x \frac{d}{dt}_{t=0} \log p(x|\rho_g(\exp_\theta(t\Omega_1)))\frac{d}{dt}_{t=0} \log p(x|\rho_g(\exp_\theta(t\Omega_2)))p(x|\rho_g(\theta))dx$$
$$= \int_x \frac{d}{dt}_{t=0} \log p(\phi_g^{-1}(x)|\exp_\theta(t\Omega_1)\frac{d}{dt}_{t=0} \log p(\phi_g^{-1}(x)|\exp_\theta(t\Omega_2))p(x|\rho_g(\theta))dx$$

using the variable change $y = \phi_g^{-1}(x)$:

$$= \int_y \frac{d}{dt}_{t=0} \log p(y|\exp_\theta(t\Omega_1))\frac{d}{dt}_{t=0} \log p(y|\exp_\theta(t\Omega_2))p(\phi_g(y)|\rho_g(\theta))|J\phi_g(y)|dy$$

and using definition 1:

$$= \int_y \frac{d}{dt}_{t=0} \log p(y|\exp_\theta(t\Omega_1))\frac{d}{dt}_{t=0} \log p(y|\exp\theta(t\Omega_2))p(y|\theta)dy$$
$$= G_\theta(\Omega_1, \Omega_2)$$

5 Conclusion

In this paper we tried to give an gentle introduction to the main concepts of intrinsic estimation on manifolds. We explained why the intrinsic CRLB could be considered as independent from the parameterization chosen and focused on two notions : Intrinsic Root Mean Square Error and normal coordinates. We showed in particular that they are bounded by a result we derived very simply from [6]. We illustrated its consequences on several standard estimation problems where it allows to see immediately that the lower bound of the IRMSE is independent from the underlying parameter. It appeared in theses examples that this independence could be also related to invariances under certain transformations. An interesting issue to investigate in further work is the link between intrinsic variables and invariances under the action of a transformation group for a given estimation problem.

References

1. S.I. Amari, H. Nagaoka, and D. Harada. *Methods of information geometry.* American Mathematical Society, 2000.
2. S. Bonnabel and R. Sepulchre. Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM J. matrix anal. appl.*, 31, 2009.
3. G. Garca and J. M. Oller. What does intrinsic mean in statistical estimation?, 2006.
4. H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on pure and applied mathematics*, 30(5):509–541, 1977.
5. X. Pennec. Intrinsic statistics on riemaniann manifolds: basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25:127–164, 2006.
6. S.T. Smith. Covariance, subspace, and intrinsic cramer-rao bounds. *IEEE-Transactions on Signal Processing*, 53(5):1610–1629, 2005.