INTRINSIC VARIANCE LOWER BOUND (IVLB): AN EXTENSION OF THE CRAMÉR-RAO BOUND TO RIEMANNIAN MANIFOLDS

João Xavier and Victor Barroso

Instituto Superior Técnico – Instituto de Sistemas e Robótica
Av. Rovisco Pais, 1049–001 Lisboa, Portugal
{jxavier,vab}@isr.ist.utl.pt

ABSTRACT

We consider parametric statistical models in which the parameter space Θ is a connected Riemannian manifold. This mathematical structure on the parameter space subsumes, as special cases, submanifolds of Euclidean spaces appearing in parametric estimation scenarios with a priori smooth deterministic constraints, and quotient spaces (such as Grassmann manifolds) which arise in certain parametric estimation scenarios with ambiguities. The Riemannian structure on the parameter space Θ turns it into a metric space and the associated Riemannian distance is used here to quantify estimation errors. We present the Intrinsic Variance Lower Bound (IVLB) which places a lower limit on the accuracy, measured in terms of the mean-square Riemannian distance, of unbiased estimators taking values in Θ. The IVLB depends both on the curvature of the parameter space and a coordinate-free extension of the well-known Fisher information matrix (FIM). We show that for flat Euclidean spaces, the IVLB collapses to the Cramér-Rao Bound (CRB). In this sense, we may interpret the IVLB as a generalization of the CRB for curved parameter spaces. Computer simulations illustrating the application of the IVLB are included.

1. INTRODUCTION

Let $\mathcal{F} = \{p(y;\theta) : \theta \in \Theta\}$ denote a parametric statistical model for the observed data vector $y \in \mathbb{R}^n$. That is, $\mathcal{F}$ stands for a collection of probability density functions with respect to the Lebesgue measure dy in $\mathbb{R}^n$. The statistical family $\mathcal{F}$ is indexed by the parameter $\theta \in \Theta$. The Cramér-Rao bound (CRB) [1, 2] is usually invoked in these setups to determine a fundamental limit on the accuracy of unbiased estimators. In the original theoretical development of the CRB it is assumed that the parameter space Θ represents an open subset of some Euclidean space $\mathbb{R}^p$. However, in many parametric estimation setups, one faces more sophisticated parameter spaces. Illustrative examples occur in (i) parametric estimation with constraints and (ii) parametric estimation over quotient spaces. In scenario (i), it is often the case that the parameter space Θ becomes an embedded submanifold of an Euclidean space, due to some *a priori* smooth deterministic constraints on the parameter θ, e.g. Θ is an ellipsoid, the special orthogonal group, etc [3, 4]. Case (ii) may arise in overparameterized statistical models or in parametric estimation affected by intrinsic ambiguities. For example, in single-input multiple-output (SIMO) blind channel identification based on 2nd order statistics, it is well

known that the channel is only identifiable modulo a phase factor. Consequently, the "right" parameter space should be the quotient space $\Theta/\sim$ obtained by quotienting out equivalent channels (i.e., differing only by a phase factor) in Θ, the original parameter space. Thus, conceptually, one switches from the identification of a channel (an ill-posed problem due the inherent phase ambiguity) to the identification of its equivalence class (a well-posed problem). This viewpoint has also advantages in practical identification algorithms as first noticed by Manton [11]. In the context of performance analysis, see [10] for other applications of such constructions.

Previous work. To tackle such sophistications in the structure of the parameter space Θ several generalizations of the CRB have been proposed. The works in [3, 4] address case (i) above, whereas [8] can cover both cases (i) and (ii) because it unifies their treatment by assuming that the parameter space Θ is a generic differentiable manifold (that is, not necessarily given as a submanifold of an ambient Euclidean space). However, no metric structure on Θ is assumed in [8]: the *covariance inequality* derived therein does not translate in a quantitative statement concerning the accuracy (measured in terms of some distance on Θ) of unbiased estimators. We introduced a preliminary version of the Intrinsic Variance Lower Bound (IVLB) in [9]. Similar to [8], the work in [9] assumes Θ to be a generic manifold and thus can be applied to cases (i) and (ii). However, it assumes Θ to be endowed with an extra piece of structure: a Riemannian layer. The existence of the Riemannian structure on Θ turns it into a metric space (the distance is induced from the Riemannian layer) and this canonical Riemannian distance can be used to measure the precision of estimators taking values in Θ: the magnitude of an estimaton error corresponds to the Riemannian distance between the "true" (unknown) family member $\theta \in \Theta$ and the estimated point $\widehat{\theta}(y) \in \Theta$ (here, $\widehat{\theta}(y)$ denotes a realization of an estimate of θ through the estimator $\widehat{\theta} : \mathbb{R}^n \to \Theta$). The preliminary version of the IVLB discussed in [9] establishes a lower-bound for the mean-square distance of unbiased estimators taking values in Θ. A clear drawback of the IVLB in [9] is the presence of an information gap relative to the CRB. More precisely, when applied to the special case of flat Riemannian manifolds, e.g., Euclidean spaces, the IVLB does not coincide with the Cramér-Rao bound: the CRB is more informative.

Contribution. In this paper, we improve our previous results in [9] by presenting a much tighter version of the IVLB. Contrary to the IVLB discussed in [9], this new version now coincides with the CRB for flat Riemannian manifolds. For generic curved spaces, it is substantially more informative than the previous version: a

particular example involving the unit-sphere illustrates this.

Paper organization. Section 2 reviews the generalizations of familiar statistical concepts such as mean, variance, Fisher information matrix, etc, to the Riemannian-geometric framework adopted herein. These concepts are used to formulate the IVLB in section 3 for parameter spaces Θ having a Riemannian structure. We also show that the IVLB coincides with the CRB for Euclidean spaces, thus closing the previous information gap. In section 4, we assess the tightness of the IVLB by examining an instance of case (i): parametric estimation with constraints. We consider an inference problem on the unit-sphere in order to make a direct comparison with the preliminary version of the IVLB in [9]. In section 5, we illustrate the IVLB in a parametric estimation scenario belonging to case (ii), namely, we consider subspace estimation from array data as an estimation problem posed on the Grassmann manifold.

2. PARAMETRIC ESTIMATION: RIEMANNIAN-GEOMETRIC FRAMEWORK

In this section, we briefly review the main concepts involved in parametric estimation from the perspective of Riemannian geometry (familiarity of the reader with basic concepts in Riemannian geometry is assumed [5, 6, 7]). More detailed expositions can be found in [8, 9]. Essentially, we review intrinsic (coordinate-free) extensions of familiar statistical objects. These are used in the formulation of the IVLB in the next section. For objects used in classical differential-geometry (e.g., tangent spaces, vector fields, etc) we (mostly) adhere to the notation in [5].

Tensors and traces. Let V be a finite-dimensional vector space over $\mathbb{R}$, with $\dim V = m$. A tensor Φ on V of type (k, l) is a multilinear map

$$\Phi : \underbrace{V \times \cdots \times V}_{k \text{ copies}} \times \underbrace{V^* \times \cdots \times V^*}_{l \text{ copies}} \to \mathbb{R},$$

where V^* stands for the dual space of V (the vector space of linear functionals on V, or, covectors). The trace of a $(1, 1)$-tensor Φ is defined as $\operatorname{tr} \Phi = \sum_{i=1}^{m} \Phi(X_i, \omega_i)$, where $\{X_1, X_2, \ldots, X_m\}$ and $\{\omega_1, \omega_2, \ldots, \omega_m\}$ denote any pair of dual basis in V and V^*, respectively, i.e., $\omega_j(X_i) = 1$ if $i = j$ and $\omega_j(X_i) = 0$ if $i \neq j$ (the definition of $\operatorname{tr} \Phi$ is independent of the particular choice of basis). An inner product on V is a $(2, 0)$-tensor Φ on V which is symmetric ($\Phi(X, Y) = \Phi(Y, X)$) and positive-definite ($\Phi(X, X) \geq 0$ with equality if and only if $X = 0$). An inner-product Φ provides a canonical identification between V and V^*: to each vector $X \in V$ corresponds the covector $\Phi^\flat(X) \in V^*$ defined by $(\Phi^\flat(X))(Y) = \Phi(X, Y)$ for all $Y \in V$. The correspondence $\Phi^\flat : V \to V^*$ is, in fact, a linear isomorphism. Its inverse map is denoted by $\Phi^\sharp : V^* \to V$. Let Φ be an inner-product; we define the $(0, 2)$-tensor $\Phi^{-1} : V^* \times V^* \to \mathbb{R}$ as $\Phi^{-1}(\omega, \sigma) = \Phi(\Phi^\sharp(\omega), \Phi^\sharp(\sigma))$. Let g be another inner-product on V (to be viewed as the "basic" inner-product). We define the trace of Φ^{-1} with respect to g, written $\operatorname{tr}_g \Phi^{-1}$, as the trace of the $(1, 1)$-tensor which operates as $(X, \omega) \in V \times V^* \mapsto \Phi^{-1}(g^\flat(X), \omega)$. One has the following result: if $\{X_1, X_2, \ldots, X_m\}$ is any basis for V, then

$$\operatorname{tr}_g \Phi^{-1} = \operatorname{tr}\left(M_\Phi^{-1} M_g\right), \tag{1}$$

where M_Φ and M_g are the $m \times m$ positive-definite matrices whose (i, j)th entry is given by $\Phi(X_i, X_j)$ and $g(X_i, X_j)$, respectively

(note that the symbol tr on the right-hand side of (1) denotes the usual trace of matrices).

Intrinsic mean and variance. Let $\widehat{\theta} : \mathbb{R}^n \to \Theta$ be an estimator. The intrinsic mean of $\widehat{\theta}$ with respect to $p(\cdot\,; \theta)$, written $\mathsf{E}_\theta\{\widehat{\theta}\}$, is the global minimizer (we assume uniqueness) in Θ of the mean-square Riemannian distance function $\phi : \Theta \to \mathbb{R}$,

$$\phi(\eta) = \mathsf{E}_\theta\{d(\eta, \widehat{\theta}(y))^2\} = \int_{\mathbb{R}^n} d(\eta, \widehat{\theta}(y))^2\, p(y; \theta)\, dy,$$

where $d : \Theta \times \Theta \to \mathbb{R}$ stands for the Riemannian distance on Θ. This intrinsic notion of mean-value is also called Fréchet mean or Riemannian center of mass [12]. Note that, by construction, $\mathsf{E}_\theta\{\widehat{\theta}\}$ is always a point in the manifold Θ. The intrinsic variance of $\widehat{\theta}$ with respect to $p(\cdot\,; \theta)$ is given by

$$\begin{aligned}
\mathsf{var}_\theta\{\widehat{\theta}\} &= \mathsf{E}_\theta\left\{d\left(\mathsf{E}_\theta\{\widehat{\theta}\}, \widehat{\theta}(y)\right)^2\right\} \\
&= \int_{\mathbb{R}^n} d\left(\mathsf{E}_\theta\{\widehat{\theta}\}, \widehat{\theta}(y)\right)^2 p(y; \theta) dy.
\end{aligned}$$

Note that these definitions coincide with the usual notions of mean value and variance when Θ is an Euclidean space with d as the usual Euclidean distance. An estimator $\widehat{\theta}$ is said to be unbiased if $\mathsf{E}_\theta\{\widehat{\theta}\} = \theta$ for all $\theta \in \Theta$.

Fisher information form. For $y \in \mathbb{R}^n$, we define the log-likelihood function $l_y : \Theta \to \mathbb{R}$, $l_y(\theta) = \log p(y; \theta)$, assumed smooth in θ. For each $\theta \in \Theta$, the Fisher information form, written I_θ, is the $(2, 0)$-tensor on $T_\theta\Theta$ (the tangent space to the manifold Θ at θ) given by

$$I_\theta(X_\theta, Y_\theta) = \mathsf{E}_\theta\{X_\theta l_y Y_\theta l_y\} = \int_{\mathbb{R}^n} (X_\theta l_y)\, (Y_\theta l_y)\, p(y; \theta)\, dy, \tag{2}$$

where X_θ and Y_θ denote tangent vectors in $T_\theta\Theta$. Recall that $X_\theta l_y$ represents the directional derivative of l_y at θ, taken in the direction defined by the tangent vector X_θ [5]. For each θ, we assume that I_θ denotes an inner-product on $T_\theta\Theta$; in particular, the $(1, 1)$-tensor I_θ^{-1} is defined (note that the Riemannian metric tensor g plays the role of the "basic" (fixed) inner-product on $T_\theta\Theta$).

3. IVLB

We now state our main result.

Theorem (IVLB). *Let the sectional curvatures of the Riemannian manifold Θ be bounded above by the constant $C \geq 0$. Let $\widehat{\theta} : \mathbb{R}^n \to \Theta$ denote an unbiased estimator and assume that, for each $\theta \in \Theta$, there exists $\epsilon > 0$, such that $\sqrt{C}\epsilon < \sqrt{2}$ and $\mathsf{Prob}_\theta\left\{\widehat{\theta}(y) \in V_\theta\right\} = 1$, where V_θ is a $\epsilon-$uniformly normal neighborhood of $\theta \in \Theta$. Then,*

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) \geq \begin{cases} \lambda_\theta & ,\ if\ C = 0 \\[2ex] \dfrac{\lambda_\theta C + 1 - \sqrt{2\lambda_\theta C + 1}}{C^2 \lambda_\theta / 2} & ,\ if\ C > 0 \end{cases}$$

where

$$\lambda_\theta = \operatorname{tr}_g I_\theta^{-1}. \tag{3}$$

The proof is omitted due to paper length restrictions (can be found in the companion paper [14]). We recall [9] that an open set V_θ is a $\epsilon-$uniformly normal neighborhood of θ if it contains θ and, for all $p \in V_\theta$, the geodesic ball centered at p with radius ϵ covers V_θ. This technical condition is certainly the main limitation of the IVLB both theoretically (it is usually difficult to check) and in practice: loosely, it means that the IVLB can only be applied to "accurate" enough estimators since it implies that $d\left(\theta, \widehat{\theta}(y)\right) < \epsilon$ with probability 1 with respect to $p(\,\cdot\,; \theta)$. However, Cramér-Rao bounds are usually applied to asymptotic regimes (high signal-to-noise ratios, large data samples, small errors, etc) in which this technical condition is "almost" fulfilled. Note that

$$\frac{\lambda_\theta C + 1 - \sqrt{2\lambda_\theta C + 1}}{C^2 \lambda_\theta / 2} \rightarrow \lambda_\theta \text{ as } C \rightarrow 0^+.$$

Thus, there is no discontinuity in the IVLB as we pass from curved spaces ($C > 0$) to flat spaces ($C = 0$).

Comparison with the CRB for flat spaces. Let Θ be the Euclidean space $\mathbb{R}^p$ (the following discussion also holds if Θ is some open, connected subset of it). Assuming the usual geometry, $\mathbb{R}^p$ is a flat space (sectional curvatures vanish everywhere) and we have $d(\theta, \eta) = \|\theta - \eta\|$ for $\theta, \eta \in \mathbb{R}^p$ ($\|\cdot\|$ is the usual Euclidean norm). For any unbiased estimator $\widehat{\theta}$, the IVLB yields the inequality $\mathsf{var}_\theta\left(\widehat{\theta}\right) = \mathsf{E}_\theta\{\|\widehat{\theta}(y) - \theta\|^2\} \geq \mathsf{tr}\, I_\theta^{-1}$, where $I_\theta : T_\theta \mathbb{R}^p \times T_\theta \mathbb{R}^p \rightarrow \mathbb{R}$ is the $(2,0)$-tensor defined in (2). It is well known that $T_\theta \mathbb{R}^p$ can be naturally identified with $\mathbb{R}^p$, by associating to $a = (a^1, a^2, \ldots, a^p) \in \mathbb{R}^p$ the directional derivative (i.e., an element of $T_\theta \mathbb{R}^p$) which acts on smooth functions f on $\mathbb{R}^p$ as $\sum_{i=1}^p a^i \frac{\partial f}{\partial x^i}(\theta)$, see [5]. Equivalently, this means that for each tangent vector $X_\theta \in T_\theta \mathbb{R}^p$ there is an (unique) $a = (a^1, a^2, \ldots, a^p) \in \mathbb{R}^p$ such that $X_\theta f = a^T \nabla f(\theta)$ for all smooth functions f defined on $\mathbb{R}^p$ (here, $\nabla f(\theta)$ denotes the gradient of f evaluated at θ). We let the notation $X_\theta \sim a$ represent this identification. Let $X_\theta \sim a$ and $Y_\theta \sim b$. Within this identification, we have

$$I_\theta(X_\theta, Y_\theta) = a^T E_\theta\{\nabla \log p(y; \theta) \nabla \log p(y; \theta)^T\}\, b = a^T I(\theta) b,$$

where $I(\theta)$ stands for the usual Fisher information matrix. Note also that $g(X_\theta, Y_\theta) = a^T b$ (recall that g stands for the Riemannian inner-product on $\mathbb{R}^p$). Thus, if we choose an orthonormal basis $\{q_1, q_2, \ldots, q_p\}$ for $\mathbb{R}^p$, the result in (1) asserts that

$$\mathsf{tr}\, I_\theta^{-1} = \mathsf{tr}\left(M_{I_\theta}^{-1}\right) = \mathsf{tr}\left((Q^T I(\theta) Q)^{-1}\right) = \mathsf{tr}\left(I(\theta)^{-1}\right),$$

where $Q = [\, q_1\, q_2\, \cdots\, q_p\,]$ (note that because the basis was chosen to be orthonormal, we have $M_g = I_p$, the $p \times p$ identity matrix, in (1)). In sum, the IVLB yields $\mathsf{var}_\theta\left(\widehat{\theta}\right) \geq \mathsf{tr}\left(I(\theta)^{-1}\right)$, which is precisely the CRB bound for the mean-square error of $\widehat{\theta}$.

4. PARAMETRIC ESTIMATION WITH CONSTRAINTS

We re-examine the inference problem on the unit-sphere $\mathbb{S}^{n-1}$ discussed in [9], in order to appreciate (through a simple example) the improvement introduced by the new version of IVLB as compared with the one in [9]. The parameter space Θ is the unit-sphere

$\mathbb{S}^{n-1} = \{\theta \in \mathbb{R}^n : \|\theta\| = 1\}$ and we observe $y = \theta + w \in \mathbb{R}^n$ where $w \sim \mathcal{N}(0, \sigma^2 I_n)$ is the observation noise. Here, $\mathcal{N}(\mu, \Sigma)$ is the Gaussian distribution with mean μ and covariance Σ. The Riemannian geometry of $\Theta = \mathbb{S}^{n-1}$ is assumed to be induced by the ambient Euclidean space; as a consequence, the Riemannian distance between $\theta, \eta \in \Theta$ is given by $d(\theta, \eta) = \mathsf{acos}(\theta^T \eta)$ and the sectional curvatures are everywhere equal to 1, see [5, 7]. For unbiased estimators $\widehat{\theta}$, the preliminary IVLB introduced in [9] yields the inequality

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) \geq \frac{4 + 3/\sigma^2 - \sqrt{\frac{1}{\sigma^2}\left(9/\sigma^2 + 24\right)}}{\frac{8}{3}}, \qquad (4)$$

see equation (16) in [9]. For the new version of IVLB, it can be seen (details omitted) that λ_θ in (3) is given by $\lambda_\theta = (n-1)\sigma^2$. Moreover, we can take $C = 1$ as a bound on the sectional curvatures of Θ. In sum, the new version of the IVLB yields

$$\mathsf{var}_\theta\left(\widehat{\theta}\right) \geq \frac{(n-1)\sigma^2 + 1 - \sqrt{2(n-1)\sigma^2 + 1}}{(n-1)\sigma^2/2}. \qquad (5)$$

Inspection of the old version of the IVLB in (4) reveals that the dimension of the manifold Θ (which is $n-1$) does not play a role. Indeed, the right-hand side of (4) does not change as n increases. This is in clear contrast with the new version in (5). In intuitive (loose) terms, the new version improves upon the old one because it takes into account the whole tangent space, whereas the previous version only considered a one-dimensional subspace in it.

We performed computer simulations to compare both bounds in (4) and (5) with the performance of the (unbiased) estimator $\widehat{\theta} : \mathbb{R}^n \rightarrow \Theta, \widehat{\theta}(y) = y / \|y\|$. We considered the case $n = 10$. A nominal point $\theta_0 \in \mathbb{S}^{n-1}$ was previously randomly generated and kept fixed during all Monte-Carlo experiments. We considered scenarios from $\mathsf{SNR}_{\min} = -10\,\mathsf{dB}$ to $\mathsf{SNR}_{\max} = 25\,\mathsf{dB}$, where $\mathsf{SNR} = \|\theta\|^2 / \mathsf{E}\left\{\|w\|^2\right\} = 1/(n\sigma^2)$. For each SNR, we considered $M = 1000$ statistically independent experiments and $\mathsf{var}_{\theta_0}(\widehat{\theta})$, the variance of $\widehat{\theta}$ with respect to $p(\,\cdot\,; \theta_0)$, was taken as as the mean value of $d(\theta_0, \widehat{\theta}(y_m))^2 = \mathsf{acos}\left(\theta_0^T \widehat{\theta}(y_m)\right)^2$, $m = 1, 2, \ldots, M$, where y_m denotes the mth realization of the observation vector y. Figure 1 plots the result of the experiments. The first (top) solid line corresponds to $\mathsf{var}_{\theta_0}(\widehat{\theta})$. The solid line with circles corresponds to the new IVLB in (5). The dashed line with squares corresponds to the old IVLB in (4). As seen, the new version of the IVLB provides a much tighter bound which coincides, asymptotically with SNR, with the variance of $\widehat{\theta}$.

5. PARAMETRIC ESTIMATION OVER QUOTIENT SPACES

To assess the tightness of the IVLB in a scenario belonging to case (ii), we consider a simplified subspace estimation problem. Let $y(k) = Us(k) + w(k)$ ($k = 1, 2, \ldots, K$) denote array snapshots where $U \in \mathbb{R}^{N \times P}$ is an unknown orthonormal frame ($U^T U = I_P$), $s(k) \in \mathbb{R}^P$ is a vector of i.i.d. zero-mean, unit-power, Gaussian sources and $w(k) \in \mathbb{R}^N$ represents zero-mean, white Gaussian observation noise with power σ^2. Both the sources and noise are assumed to be white in the temporal and spatial domains. The goal is to estimate the matrix U from the observed data vector $y = \left(y(1)^T, y(2)^T, \ldots, y(K)^T\right)^T \in \mathbb{R}^{NK}$. Thus, the parameter
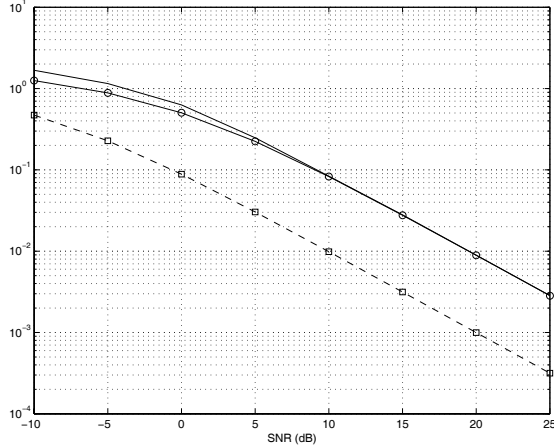
Fig. 1. $\mathrm{var}_{\theta_0}(\widehat{\theta})$ (solid), new IVLB (solid,circles) and old IVLB (dashed,squares)

Fig. 2. $\mathrm{var}_{\theta_0}(\widehat{\theta})$ (solid), new IVLB (solid,circles)

space is $\Theta = \{U \in \mathbb{R}^{N \times P} : U^T U = I_P\}$ (a Stiefel manifold). However, as the data vector is distributed as $\mathcal{N}(0, C(U))$, where $C(U) = I_K \otimes (UU^T + \sigma^2 I_N)$, ($\otimes$ denotes Kronecker product) we see that $C(U) = C(UQ)$ for any $P \times P$ orthogonal matrix Q. Thus, only the subspace spanned by U is identifiable from y. The "right" parameter space is thus $\mathsf{G}(N, P)$ the Grassmann manifold of P-dimensional subspaces in $\mathbb{R}^P$, see [5, 7]. See [13] for a more involved study on Grassmann spaces. We let $[U]$ denote the subspace spanned by U. We consider the case $N = 4$, $P = 2$, $K = 10$. It can be shown (details omitted) that the Riemannian distance is given by $d([U], [V]) = \sqrt{(\mathsf{acos}(\sigma_1))^2 + (\mathsf{acos}(\sigma_2))^2}$, where σ_1, σ_2 are the singular values of $U^T V$. Also, $C = 2$ can be used as an upper-bound on the sectional curvatures of $\Theta = \mathsf{G}(4, 2)$. To perform the simulations, we first randomly generate an $N \times P$ orthonormal matrix U_0, to which corresponds the point $\theta_0 = [U_0]$ in the parameter space $\Theta = \mathsf{G}(N, P)$. The matrix U_0 was kept fixed during the simulations, in which we varied the SNR from $\mathsf{SNR}_{\min} = 0\,\mathsf{dB}$ to $\mathsf{SNR}_{\max} = 35\,\mathsf{dB}$, where $\mathsf{SNR} = \mathsf{E}\{\|U_0 s(k)\|^2\}/\mathsf{E}\{\|w(k)\|^2\} = P/(N\sigma^2)$. We compared the IVLB (computation of λ_{θ_0} is straightforward but omitted here due to lack of space) with the performance of the SVD estimator $\widehat{\theta}$; more precisely, $\widehat{\theta}(y) = [\widehat{U}]$, where $\widehat{U}$ is the $N \times P$ orthonormal matrix containing the P singular vectors associated with the largest P singular values of $\widehat{R}_y = \frac{1}{K}\sum_{k=1}^{K} y(k)y(k)^T$. Figure 2 plots the result of the experiments. We see that the IVLB provides a reasonably tight bound for the observed mean-square Riemannian distance of the SVD estimator.

6. REFERENCES

[1] C. Rao. *Linear Statistical Inference and Its Application*. New York: Wiley, 1973.

[2] L. Scharf. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* Addison-Wesley Publishing Company, 1991.

[3] J. Gorman and A. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Transactions on Information Theory*, vol. 26, no. 6, November 1990.
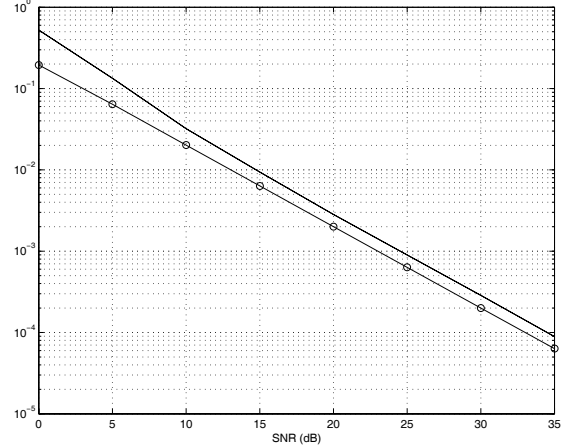
[4] P. Stoica and B. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177–179, July 1998.

[5] W. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* 2nd edition, New York: Academic Press.

[6] J. Jost. *Riemannian Geometry and Geometric Analysis*. 2nd ed. Springer-Verlag.

[7] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*, 2nd ed. Springer-Verlag.

[8] H. Hendricks, "A Cramér-Rao type lower bound for estimators with values in a manifold," *Journal of Multivariate Analysis*, no. 38, pp. 245–261, 1991.

[9] J. Xavier and V. Barroso, "Intrinsic distance lower bound for unbiased estimators on Riemannian manifolds," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'02)*, vol. 2, pp. 1141–1144, Florida, EUA, May 2002.

[10] J. Xavier and V. Barroso, "The Riemannian geometry of certain parameter estimation problems with singular Fisher information matrices," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'2004)*, vol. 2, pp. 1021–1024, Montreal, Canada, May 2004.

[11] J. Manton, "An Improved Least Squares Blind Channel Identification Algorithm for Linearly and Affinely Precoded Communication Systems," *IEEE Signal Processing Letters*, vol. 9, no. 9, pp. 282–285, September 2002.

[12] R. Bhattacharya and V. Patrangenaru, "Large sample theory of intrinsic and extrinsic sample means on manifolds I," *The Annals of Statistics*, vol. 31, no.1, pp.1–29, 2003.

[13] S. Smith, "Intrinsic Cramér-Rao bounds and subspace estimation accuracy," *1st IEEE Workshop on Sensor Arrays and Multichannel Signal Processing*, Boston, MA, 2000.

[14] J. Xavier and V. Barroso, "Intrinsic variance lower bound (IVLB) for unbiased estimators on Riemannian manifolds," *in preparation*.