

THE RIEMANNIAN GEOMETRY OF CERTAIN PARAMETER ESTIMATION PROBLEMS WITH SINGULAR FISHER INFORMATION MATRICES

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

Many parametric statistical models suffer from “intrinsic ambiguities” in the sense that the distribution of the observation vector is invariant to smooth, structured changes in the model’s parameters. The fact that certain members of the parametric statistical family are locally undistinguishable makes the Fisher information matrix (FIM) associated to the given statistical model singular. We examine such degenerate deterministic parameter estimation problems from a Riemannian geometric perspective. We start by replacing the original (ambiguous) parameter set by a lower-dimensional Riemannian (non-ambiguous) parameter set. The new parameter set comes in the form of a quotient space and is obtained by identifying equivalent family members in the initial parameterization. We specialize recently developed extensions of the Cramér-Rao bound (CRB) for the Riemannian setup to this particular setting. This offers a re-interpretation of the CRB inequality involving the pseudo-inverse of the FIM. Also, we present a lower bound on the variance (computed with respect to the geodesic distance) of unbiased estimators taking values in the quotient space. Geometrically, this corresponds to a fundamental limit on the capability of these estimators in discriminating adjacent parameter equivalence classes in the original problem parameterization. A numerical example involving the blind identification of single-input multiple-output (SIMO) channels driven by a Gaussian source is worked out.

1. INTRODUCTION

Let $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ represent a parametric statistical model for the observation vector $x \in \mathbb{R}^n$. More precisely, \mathcal{F} denotes a family of positive probability density functions over \mathbb{R}^n , that is, $f_\theta : \mathbb{R}^n \rightarrow \mathbb{R}$, $f_\theta(x) > 0$ for all $x \in \mathbb{R}^n$ and $\int f_\theta d\mu = 1$, for all $\theta \in \Theta$. Here, μ represents Lebesgue measure. We also define $f_x : \Theta \rightarrow \mathbb{R}$ by $f_x(\theta) = f_\theta(x)$. The parameter θ takes values in an open subset Θ of some Euclidean space \mathbb{R}^p . Let the Fisher information matrix (FIM) associated to the statistical model \mathcal{F} at the point θ be denoted by

$$I_\theta = \mathbb{E}_\theta \left\{ \nabla l_x(\theta) \nabla l_x(\theta)^T \right\} \quad (1)$$

This work was supported by the FCT Programa Operacional Sociedade de Informao (POSI) in the frame of QCA III, under contract POSI/2001/CPS/38775

where $l_x : \Theta \rightarrow \mathbb{R}$ stands for the log-likelihood $l_x(\theta) = \ln f_\theta(x)$ and ∇ denotes the gradient operator. That is,

$$I_\theta = \int_{\mathbb{R}^n} \nabla l_x(\theta) \nabla l_x(\theta)^T f_\theta(x) d\mu(x).$$

Let $\hat{b} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $x \mapsto \hat{b}(x)$, denote some estimator with mean value $b : \Theta \rightarrow \mathbb{R}^m$,

$$b(\theta) = \mathbb{E}_\theta \left\{ \hat{b}(X) \right\} = \int_{\mathbb{R}^n} \hat{b}(x) f_\theta(x) d\mu(x).$$

If the FIM is non-singular then we have the classical Cramér-Rao bound (CRB) inequality

$$\text{Cov}_\theta(\hat{b}) \succeq Db(\theta) I_\theta^{-1} Db(\theta)^T, \quad (2)$$

where

$$\text{Cov}_\theta(\hat{b}) = \mathbb{E}_\theta \left\{ \left(\hat{b}(X) - b(\theta) \right) \left(\hat{b}(X) - b(\theta) \right)^T \right\}$$

is the covariance matrix of \hat{b} , Db represents the derivative of b and $A \succeq B$ means that $A - B$ is semidefinite positive. If the FIM is singular, we have the extended CRB inequality [1, 2]

$$\text{Cov}_\theta(\hat{b}) \succeq Db(\theta) I_\theta^+ Db(\theta)^T, \quad (3)$$

where I_θ^+ denotes the Moore-Penrose pseudoinverse of I_θ . We are interested in scenarios where I_θ is singular.

Local ambiguity. In this paper, we assume that the singularity of the FIM is reflecting local ambiguities in the statistical model \mathcal{F} . That is, certain smooth changes in the parameter θ leave the distribution of the observation vector x unaltered. This induces singular information matrices. For example, suppose that $\theta(t)$ denotes a smooth curve in Θ such that $\theta(0) = \theta_0$,

$$\dot{\theta}(0) = \frac{d}{dt} \theta(t) \Big|_{t=0} = v \neq 0,$$

and $f_{\theta(t)} = f_{\theta_0}$ (as real-valued functions over \mathbb{R}^n) for all t . That is, the distribution of the observation x is invariant to the motion $\theta(t)$ in the parameter set Θ . Taking the derivative with respect to t in the equality $f_x(\theta(t)) = f_x(\theta_0)$ and evaluating at $t = 0$ yields $\nabla l_x(\theta_0)^T v = 0$ (for all x). Thus,

$$v^T I_{\theta_0} v = \mathbb{E}_{\theta_0} \left\{ \left(\nabla l_x(\theta_0)^T v \right)^2 \right\} = 0,$$

and I_{θ_0} is singular.

Model reparameterization. The main idea in this paper consists in switching the analysis to a new parametric family $\mathcal{F}^* = \{f_{\theta^*} : \theta^* \in \Theta^*\}$ with a new parameter set Θ^* such that the associated FIM (more precisely, a tensor-like analogue of it) is no longer singular. The new parameter set Θ^* is obtained from Θ by “quotienting out” equivalent family members in \mathcal{F} . That is, introduce the equivalence relation \sim on Θ by declaring $\theta_1 \sim \theta_2$ if $f_{\theta_1} = f_{\theta_2}$ holds μ -almost everywhere. In words, the relation \sim identifies (groups together) members in \mathcal{F} which induce the same distribution on the data vector x . The equivalence relation \sim partitions the parameter space Θ into disjoint equivalence classes and we let $\Theta^* = \Theta / \sim$ be the corresponding quotient space. We denote by $\pi : \Theta \rightarrow \Theta^*$ the map which projects $\theta \in \Theta$ to its equivalence class $\pi(\theta) \in \Theta^*$, also written $[\theta]$. For $\theta^* \in \Theta^*$, the subset $F_{\theta^*} = \pi^{-1}(\theta^*) = \{\theta : \pi(\theta) = \theta^*\} \subset \Theta$ is termed the fiber over θ^* . See figure 1 for a sketch. The reparameterization of

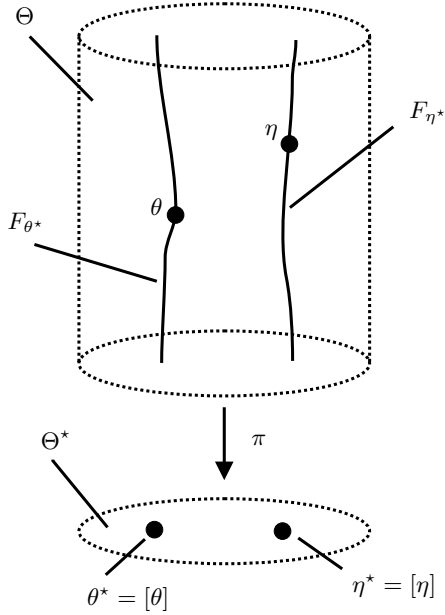


Fig. 1. Two points in Θ^* and corresponding fibers in Θ

\mathcal{F} by Θ^* corresponds to the new family $\mathcal{F}^* = \{f_{\theta^*} : \theta^* \in \Theta^*\}$, where $f_{\theta^*} = f_{\theta}$ for some θ in the fiber F_{θ^*} (note that the particular choice of θ is immaterial).

Assumptions. In this paper, we assume that Θ^* can be given the structure of a Riemannian manifold with the projection map π as a Riemannian submersion (the reader is assumed to be familiar with basic concepts of Riemannian geometry [3, 4]). An example of such a construction, in the context of blind identification of single-input multiple-output (SIMO) channels, is carried out in section 4. Loosely, π being a Riemannian submersion means that the geometries of Θ and Θ^* interface nicely (through π). In more precise terms, let $d\pi : T_{\theta}\Theta \rightarrow T_{\theta^*}\Theta^*$ denote the push-forward (derivative) of π . Then, the subspace $\text{Ker } d\pi = \{X_{\theta} \in T_{\theta}\Theta : d\pi(X_{\theta}) = 0\}$ is termed the vertical subspace of the tangent space $T_{\theta}\Theta$ and it is denoted in the sequel by V_{θ} . It is the tangent space to the fiber $F_{[\theta]}$ at θ . Its orthogonal complement (with respect to the metric on Θ) is called the horizontal

subspace, written H_{θ} . The projection map π is a Riemannian submersion if and only if $d\pi : H_{\theta} \rightarrow T_{[\theta]}\Theta^*$ is an isometry (with respect to the Riemannian metrics in Θ and Θ^*). Figure 2 illustrates these concepts. Moreover, we assume that $\text{Ker } I_{\theta} \simeq V_{\theta}$, that

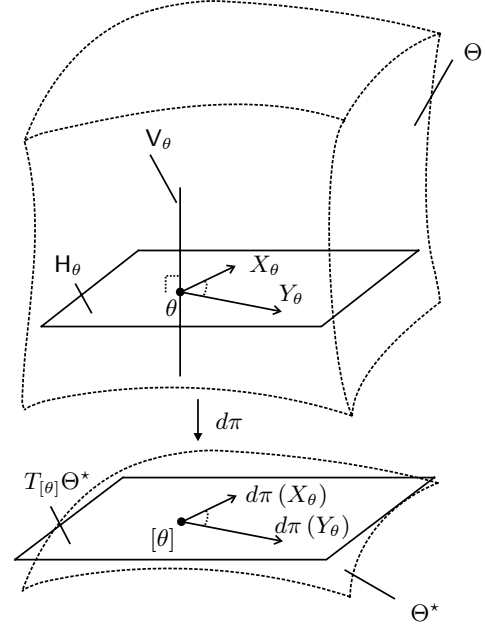


Fig. 2. The map $d\pi : H_{\theta} \rightarrow T_{[\theta]}\Theta^*$ is an isometry

is, the kernel of the FIM at the point θ can be identified with the vertical space V_{θ} . This technical condition captures the notion that motions along the fibers are ambiguous, whereas motions across the fibers are not.

Paper organization. In section 2, we investigate a first consequence of our analytical shift from the parameter set Θ to Θ^* . More precisely, we show that, when we apply to our context the smooth manifold generalization of the classical CRB inequality (2) in [5], the extended CRB inequality in (3) is recovered. This provides an interpretation of the extended inequality (3): it is nothing more than the “classical” CRB (2) read, not in the original parameter space Θ , but in the coset space Θ^* . In section 3, we concern ourselves with estimation over Θ^* . Remark that, although the original parameter θ is not identifiable, its equivalence class $[\theta]$ (or fiber) can be. We start by presenting a refinement of the intrinsic variance lower bound (IVLB) presented in [6] (see also [7]). The IVLB assumes a parametric statistical family indexed over a Riemannian manifold P and places a fundamental limit on the variance of estimators taking values in a (possibly distinct) Riemannian manifold M . The variance of the estimators is measured with respect to the intrinsic (geodesic) distance in M . We specialize the IVLB to our case ($P = M = \Theta^*$), thus obtaining a lower bound on the capability of estimators in resolving fibers in the original problem parameterization. In section 4, we analyse the problem of blind identification of SIMO channels excited by a complex, unit-power, white Gaussian source from the viewpoint proposed herein. Note that, as it is well-known, such channels can only be resolved up to a phase ambiguity. We compute the IVLB associated to this problem and present some computer simulations, using the subspace-based method in [8] as the estimator, to illus-

trate the tightness of the bound.

2. INTERPRETATION OF THE EXTENDED CRB INEQUALITY

In this section, we show that the extended CRB inequality (3) is nothing else than the ‘‘classical’’ CRB inequality (2) viewed in the coset space Θ^* . For this, we recall that [5] provides the generalization of the classical CRB (2) for scenarios where the parametric family is indexed by a smooth manifold P and the estimator takes values in another smooth manifold M . This CRB generalization is termed the information inequality in theorem 3.2 [5] and reduces to (2) when we let both P and M be Euclidean spaces.

It can be shown (using lemma 3.1 in [5] but details omitted due to lack of space) that the information inequality in [5] (with the squared Euclidean distance as the loss function) corresponds in our setting (that is, with $P = \Theta^*$ and $M = \mathbb{R}^m$) to

$$w^T \text{Cov}_\theta(\hat{b}) w \geq \max_{J_{[\theta]}(X_{[\theta]}, X_{[\theta]}) = 1} \langle w, db^*(X_{[\theta]}) \rangle^2 \quad (4)$$

for all $w \in \mathbb{R}^m$ and $\theta \in \Theta$. In (4), $J_{[\theta]} : T_{[\theta]}\Theta^* \times T_{[\theta]}\Theta^* \rightarrow \mathbb{R}$ denotes the Fisher information form associated with \mathcal{F}^* , see [5], $b^* : \Theta^* \rightarrow \mathbb{R}$ is defined by $b^*([\theta]) = b(\theta)$ (that is, we let b descend to the quotient), db^* denotes the derivative of b^* , and $\langle \cdot, \cdot \rangle$ represents the Riemannian metric in \mathbb{R}^m (that is, the usual Euclidean inner product under the canonical identification $T_y\mathbb{R}^m \simeq \mathbb{R}^m$). Now, exploiting the fact that i) $d\pi : \mathbf{H}_\theta \rightarrow T_{[\theta]}\Theta^*$ is a linear isomorphism, ii) $b^* \circ \pi = b$ and iii) $J_\theta(X_\theta, X_\theta) = J_{[\theta]}(d\pi(X_\theta), d\pi(X_\theta))$, the right-hand side of (4) can be written as

$$\max_{X_\theta \in \mathbf{H}_\theta} \langle w, db(X_\theta) \rangle^2, \quad (5)$$

where J_θ the Fisher information form associated with \mathcal{F} . Under the canonical identification $T_\theta\Theta = T_\theta\mathbb{R}^p \simeq \mathbb{R}^p$, any tangent vector in $T_\theta\Theta$, say X_θ , can be identified with a vector in \mathbb{R}^p , say u , and within this identification we have $db(x_\theta) = Db(\theta)u$ and $J_\theta(X_\theta, X_\theta) = u^T I_\theta u$. Moreover, if X_θ is horizontal, then u must be orthogonal to the kernel of the FIM I_θ (recall our assumption $\text{Ker } I_\theta \simeq V_\theta$), that is, if

$$I_\theta = \begin{bmatrix} H & V \end{bmatrix} \begin{bmatrix} \Lambda & \\ & 0 \end{bmatrix} \begin{bmatrix} H^T \\ V^T \end{bmatrix} \quad (6)$$

denotes an eigenvalue decomposition of I_θ , with H orthonormal and the diagonal matrix Λ positive definite, then $u = H\alpha$ for some vector α . Thus, (5) can be rewritten as

$$\max_{\alpha^T \Lambda \alpha = 1} \left(w^T Db(\theta) H \alpha \right)^2,$$

which evaluates to

$$w^T Db(\theta) H \Lambda^{-1} H^T Db(\theta)^T w = w^T Db(\theta) I_\theta^+ Db(\theta)^T w.$$

In sum, the information inequality in [5] applied to the statistical model \mathcal{F}^* in the quotient space Θ^* boils down to

$$w^T \text{Cov}_\theta(\hat{b}) w \geq w^T Db(\theta) I_\theta^+ Db(\theta)^T w$$

for all w . Thus, $\text{Cov}_\theta(\hat{b}) \succeq Db(\theta) I_\theta^+ Db(\theta)^T$, which is the extended CRB inequality in (3). The point here is that, as just seen,

the extended CRB inequality (3) possesses a (previously unsuspected) geometrical meaning: it is just a re-statement of the ‘‘classical’’ CRB inequality (more precisely, its generalization to the manifold setting) associated with the parameterization Θ^* but formulated in terms of the initial parameterization Θ .

3. INTRINSIC VARIANCE LOWER BOUND (IVLB)

The intrinsic variance lower bound (IVLB) presented in [6] (see also [7]) provides a fundamental limit on the variance of estimators taking values in a Riemannian manifold M , and associated to parametric families indexed by a Riemannian manifold P . The variance of the estimators is computed with respect to the geodesic distance in M . It can be shown (details omitted due to paper length restrictions) that the IVLB applied to our setting ($P = \Theta^*$, $M = \mathbb{R}^m$) states (within some omitted technical conditions, see [6]) that if $\hat{\theta}^* : \mathbb{R}^m \rightarrow \Theta^*$, $x \mapsto \hat{\theta}^*$, denotes an unbiased estimator for the equivalence class $[\theta]$ (see [7] for the precise meaning of the mean value of an estimator taking values in a Riemannian manifold), then

$$\text{var}_{[\theta]}(\hat{\theta}^*) = \mathbb{E}_\theta \left\{ d \left(\hat{\theta}^*(X), [\theta] \right)^2 \right\} \quad (7)$$

obeys the inequality

$$\text{var}_{[\theta]}(\hat{\theta}^*) \geq \frac{4C + 3\lambda_{[\theta]} - \sqrt{\lambda_{[\theta]} (9\lambda_{[\theta]} + 24C)}}{\frac{8}{3}C^2}, \quad (8)$$

where in (7) $d(\theta^*, \eta^*)$ denotes the geodesic distance between the points $\theta^*, \eta^* \in \Theta^*$ and in (8) C denotes an upper bound on the sectional curvature of Θ^* and $\lambda_{[\theta]}$ is the minimum eigenvalue of the Fisher information form $J_{[\theta]}$ associated with \mathcal{F}^* .

In our case, since $\pi : \Theta \rightarrow \Theta^*$ denotes a Riemannian submersion $\lambda_{[\theta]}$ can be computed as

$$\lambda_{[\theta]} = \min_{X_\theta \in \mathbf{H}_\theta} J_\theta(X_\theta, X_\theta),$$

or, in terms of the eigenvalue decomposition in (6), $\lambda_{[\theta]}$ is the minimum eigenvalue of Λ . Moreover, by O’Neill’s formula [4, theorem 3.61, pp. 127] and since the curvature of Θ is zero (flat Euclidean space) we have

$$C \leq \max_{X, Y} \frac{3}{4} \left| [X, Y]^V \right|^2, \quad (9)$$

where the maximum is over all pairs (X, Y) of smooth, horizontal, orthonormal vector fields in Θ ; here, $[X, Y]$ denotes the Lie bracket of X and Y , and $[X, Y]^V$ stands for the orthogonal projection onto the vertical subspace.

4. EXAMPLE

Consider a SIMO channel $x(t) = h(z)s(t) + w(t)$, where $x(t) \in \mathbb{C}^L$ denotes the output, $h(z) = h(0) + h(1)z^{-1} + \dots + h(D)z^{-D}$ is the z -transform of the finite-impulse response of the channel ($h(d) \in \mathbb{C}^L$ for all d), $s(t) \in \mathbb{C}$ denotes a zero-mean, unit-power, complex circular white Gaussian source and $w(t) \in \mathbb{C}^L$ denotes zero-mean, white spatio-temporal additive complex Gaussian noise with power σ^2 . If T consecutive data samples are observed and stacked in the LT -dimensional column vector x , then

x denotes a zero-mean Gaussian vector with a covariance matrix $C(h(z)) = \mathbb{E}\{xx^H\}$ (x^H stands for the conjugate transpose of x) which satisfies $C(h(z)) = C(h(z)e^{i\phi})$ for all ϕ . Thus, (as it is well-known) the channel is only identifiable up to a phase ambiguity.

The Riemannian manifold Θ . In this example, the parameter to be estimated is the SIMO channel $h(z)$ which can be viewed as a point in the parameter space $\Theta \simeq \mathbb{C}^{L(D+1)} - \{0\}$ (the origin has been excluded) by making the correspondence $h(z) \mapsto [h(0)h(1) \cdots h(D)]$. Note that, $T_{h(z)}\Theta$ can be naturally identified with $\mathbb{C}^{L(D+1)}$ and, within this identification, the canonical Riemannian metric is given by $\langle X_{h(z)}, Y_{h(z)} \rangle = \text{Re tr}(\Delta_y^H \Delta_x)$ if $X_{h(z)} \simeq \Delta_x \in \mathbb{C}^{L(D+1)}$ and $Y_{h(z)} \simeq \Delta_y \in \mathbb{C}^{L(D+1)}$ (Re z denotes the real part of the complex number z and $\text{tr}(A)$ is the trace of the matrix A). The parametric family is given by $\mathcal{F} = \{\mathcal{N}(0, C(h(z))) : h(z) \in \Theta\}$, where $\mathcal{N}(\mu, \Sigma)$ denotes the complex Gaussian distribution with mean μ and covariance Σ .

The smooth manifold Θ^* . Define a right action of the Lie group $\mathbb{S}^1 = \{u \in \mathbb{C} : |u| = 1\}$ on Θ as $\varphi : \Theta \times \mathbb{S}^1 \rightarrow \Theta$, $\varphi(h(z), u) = h(z)u$. In the sequel, we use the notation $\varphi(h(z), u) = h(z) \cdot u$. For given $h(z) \in \Theta$, the subset $h(z)\mathbb{S}^1 = \{h(z) \cdot u : u \in \mathbb{S}^1\} \subset \Theta$ is termed the orbit of $h(z)$ under the action of \mathbb{S}^1 . The space Θ is the disjoint union of all orbits generated by \mathbb{S}^1 . The set of orbits is called the orbit space and is denoted by Θ/\mathbb{S}^1 . Now, it can be seen that two members $h(z)$ and $g(z)$ in the statistical family \mathcal{F} are equivalent, that is, $h(z) \sim g(z)$, if and only if $h(z)$ and $g(z)$ are in the same orbit. Thus, the quotient space $\Theta^* = \Theta/\sim$ has a natural identification with the orbit space Θ/\mathbb{S}^1 . Moreover, it is readily seen that the action φ is smooth, free and proper. Thus, we are covered by theorem 1.95 in [4, page 32]. It states that the orbit space Θ/\mathbb{S}^1 , hence Θ^* by identification, has a unique smooth structure making the canonical projection $\pi : \Theta \rightarrow \Theta^*$ a smooth submersion.

The Riemannian manifold Θ^* . Up to this point, the quotient space is only a smooth manifold. We have not yet inserted in it a Riemannian structure. To induce such geometric structure, we note that the Lie group \mathbb{S}^1 acts on Θ by isometries, that is, for any $u \in \mathbb{S}^1$, the derivative $d\varphi_u : T_{h(z)}\Theta \rightarrow T_{h(z)u}\Theta$ of the smooth map $\varphi_u : \Theta \rightarrow \Theta$, $\varphi_u(h(z)) = \varphi(h(z), u)$ is a linear isometry, as it is easily checked. Thus, we can apply proposition 2.28 in [4, page 64] which asserts that Θ^* has a unique Riemannian metric making the canonical projection π a Riemannian submersion.

Geodesic distance and sectional curvature in Θ^* . The geodesic distance between the points $[h(z)]$ and $[g(z)]$ in Θ^* (with respect to the Riemannian structure in Θ^*) can be obtained in closed-form (details omitted) and it is given by

$$d([h(z)], [g(z)]) = \sqrt{\|H\|^2 + \|G\|^2 - 2|\text{tr}(G^H H)|}, \quad (10)$$

where $H = [h(0)h(1) \cdots h(D)]$, $G = [g(0)g(1) \cdots g(D)]$ and $\|A\|$ denotes the Frobenius norm of A . Moreover, by exploiting O'Neill's formula [4, theorem 3.61, pp. 127] it can be shown that the maximum of all sectional curvatures at the point $[h(z)] \in \Theta^*$ is given by $K_{[h(z)]} = 3/\|H\|^2$.

Computer simulations. We performed some computer simulations to assess the tightness of the IVLB equality in (8). A nominal SIMO channel $h_0(z)$ with $L = 4$ outputs and memory degree

$D = 2$ was randomly generated (and kept fixed throughout the simulations). We assume that the power of the channel $h_0(z)$ is known within a 50% error, that is, the value $C = 3/(0.5\|H_0\|)^2$ is used in the IVLB (8).

The signal-to-noise ratio is fixed at 20 dB and the data packet length is varied between $T = 25$ and $T = 100$. For each T , the variance of the subspace-based channel estimator in [8], computed in terms of the geodesic distance (10), is estimated by Monte-Carlo averaging. Figure 3 shows the results obtained. It can be seen that the performance of the subspace-based channel estimator in discriminating fibers in Θ is close to the fundamental limit placed by the IVLB.

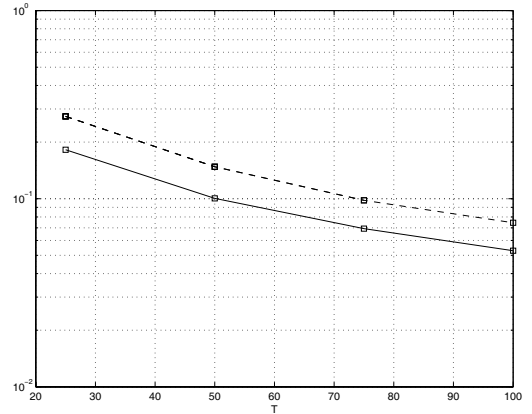


Fig. 3. Estimated $\text{var}_{[\Theta]}(\hat{\theta}^*)$ (dashed) and IVLB (solid)

5. REFERENCES

- [1] C. Rao. *Linear Statistical Inference and Its Application*. New York: Wiley, 1973.
- [2] P. Stoica and T. Marzetta, "Parameter estimation problems with singular information matrices" *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 87–90, January 2001.
- [3] J. Jost. *Riemannian Geometry and Geometric Analysis*. 2nd ed. Springer-Verlag.
- [4] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*, 2nd ed. Springer-Verlag.
- [5] H. Hicks, "A Cramér-Rao type lower bound for estimators with values in a manifold," *Journal of Multivariate Analysis*, no. 38, pp. 245–261, 1991.
- [6] J. Xavier and V. Barroso, "Intrinsic variance lower bound for inference problems on Riemannian manifolds," *Proceedings of the 10th Mediterranean Conference on Control and Automation (MED'2002)*, Lisbon, Portugal, July 2002.
- [7] 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)*, Orlando, Florida, EUA, May 2002.
- [8] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 516–525, February 1995.