INTRINSIC QUADRATIC PERFORMANCE BOUNDS ON MANIFOLDS

Steven T. Smith[*]
MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420

Louis Scharf[†]
Elect. & Comp. Engr. and Statistics
Colorado State University
Fort Collins, CO 80523-1373

L. Todd McWhorter[‡]
Altius Research Associates
E. 7th Street, Suite 302
Loveland, CO 80537

ABSTRACT

Cramér-Rao bounds have been previously generalized to the class of nonlinear estimation problems on manifolds. This new approach can be used to derive a broad class of quadratic error performance bounds. A generalized intrinsic score function on the manifold-valued parameter space is introduced that distinguishes one bound from another. The derivation itself is invariant to transformations of the parameter space and score space. The resulting generalized Weiss-Weinstein bounds are shown to be invariant to certain transformations of the score. Applications of this work include cases where ambiguities, low signal-to-noise, or low sample support limit the utility of Cramér-Rao bounds, and more general quadratic bounds on manifold-valued parameters must be considered.

1. INTRODUCTION

Many signal processing applications involve estimation problems on manifolds, such as the sphere (e.g., unit-noise-gain filters), subspaces (e.g., interference suppression and signal detection), and covariance matrices (e.g., adaptive filtering and spectral estimation). Unlike estimation problems traditionally posed on vector spaces, the operations of vector addition and subtraction cannot be used to compare different points in the manifold-valued parameter space, and therefore the standard least-squares approaches cannot be used to derive the Cramér-Rao bounds or the class of more general Weiss-Weinstein quadratic bounds on the error covariance matrix. Cramér-Rao bounds have been previously generalized to the class of nonlinear estimation problems on manifolds [11], and this new approach can be used to derive a broad class of quadratic error performance bounds to establish new estimation bounds on manifolds, including the Weiss-Weinstein [13], Bhattacharyya [4], Barankin [2], and Bobrovsky-Zakai [5] bounds on manifolds, as well as the previously established Cramér-Rao bound [6,9,10,14]. A generalized, intrinsic score function on the manifold is introduced that distinguishes one bound from another. The derivation itself is invariant to transformations of the parameter space and score space. The generalized Weiss-Weinstein bounds are shown to be invariant to certain transformations of the score. Applications of this work include cases where ambiguities, low signal-to-noise, or low sample support limit the utility of Cramér-Rao bounds, and more general quadratic bounds on manifold-valued parameters must be considered.

[*] stsmith@ll.mit.edu. This work was sponsored by the United States Air Force under Air Force contract F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

[†] scharf@engr.colostate.edu. This work was sponsored by AFOSR under contract FA9550-04-1-0371 and by ONR under contract N0014-04-1-0084.

[‡] mcwhorter@altiusresearch.com. This work was sponsored by ONR contract N00014-C-04-0031.

2. THE ESTIMATION PROBLEM

2.1. The score function

Given the statistical model $f(\mathbf{z}|\boldsymbol{\theta})$, we are concerned with the problem of estimating the manifold-valued parameter $\boldsymbol{\theta} \in M$, where $\mathbf{z}$ is a random vector of measurements and M is a given n-dimensional real manifold. Common examples in signal processing include the unit-sphere, the space of orthogonal/unitary matrices, the Grassmann or Stiefel manifolds, or the space of covariance matrices [11]. For a given coordinate chart on M, we may imagine the manifold-valued parameter to be decomposed as the vector of real numbers $\boldsymbol{\theta} = (\theta^1, \theta^2, \ldots, \theta^n)^{\mathrm{T}} \in \mathbb{R}^n$. An estimate $\hat{\boldsymbol{\theta}}(\mathbf{z})$ of $\boldsymbol{\theta}$ is evaluated using the p-dimensional vector-valued *score function* $\mathbf{s}(\boldsymbol{\theta})$. The score function is a very general concept that expresses the multitude of quadratic performance bounds that will be considered. For the example of Euclidean space $M = \mathbb{R}^n$, the two most commonly encountered score functions are the *error score* $\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ and the Fisher score $\mathbf{s}_{\mathrm{F}}(\boldsymbol{\theta}) = \partial\ell/\partial\boldsymbol{\theta} = (\partial\ell/\partial\theta^1, \partial\ell/\partial\theta^2, \ldots, \partial\ell/\partial\theta^n)$, where $\ell(\mathbf{z}|\boldsymbol{\theta}) = \log f(\mathbf{z}|\boldsymbol{\theta})$ is the log-likelihood function.

2.2. Score functions on vector bundles

To define the score function in this very general intrinsic setting, we require the assignment of a p-dimensional vector space $V_{\boldsymbol{\theta}}$ to every point $\boldsymbol{\theta}$ in M. Such a construct is called a *vector bundle* over M [1], which we shall denote as F (Figure 1). In the simplest case $V_{\boldsymbol{\theta}} = \mathbb{R}^p$, however, there are im-

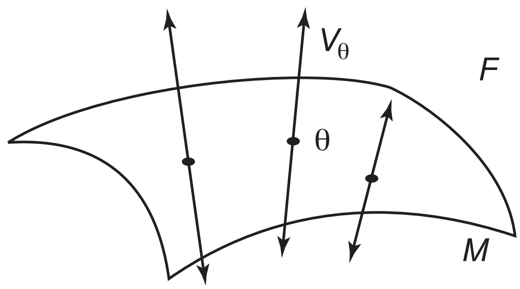

Fig. 1. The vector bundle F over the manifold M. Each point $\boldsymbol{\theta} \in M$ has an associated p-dimensional vector space $V_{\boldsymbol{\theta}}$ associated with it. The score $\mathbf{s}(\boldsymbol{\theta})$ is a vector lying in the vector space $V_{\boldsymbol{\theta}}$.

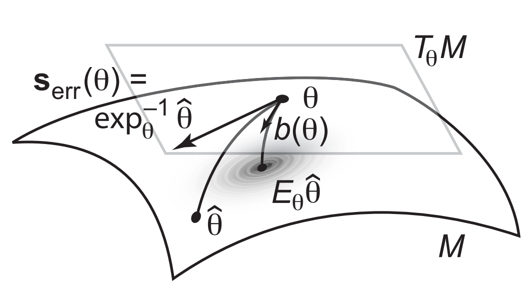

Fig. 2. The error score function. The error score function $\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \exp_{\boldsymbol{\theta}}^{-1} \hat{\boldsymbol{\theta}}$ is shown, which is a section of the tangent bundle TM. The mean of the error function is the estimator bias vector field $\boldsymbol{b}(\boldsymbol{\theta}) = E[\mathbf{s}_{\hat{\boldsymbol{\theta}}}]$.

portant cases where a generalization is necessary. For example, the most commonly encountered vector bundles over M are the tangent space TM, in which each point on M has a naturally defined tangent plane $T_{\boldsymbol{\theta}}M$, and the cotangent space T^*M with the naturally defined cotangent planes $T^*_{\boldsymbol{\theta}}M$. Using coordinates, we imagine the parameter $\boldsymbol{\theta}$ to be an n-by-1 column vector, tangent vectors in $T_{\boldsymbol{\theta}}M$ to also be n-by-1 column vectors, and cotangent vectors in $T^*_{\boldsymbol{\theta}}M$ to be 1-by-n row vectors. Of course, the vector bundle F need not be TM or T^*M, as the dimensionality of the score function is not necessarily that of the manifold. Furthermore, given a fixed vector space V, it may be easy to visualize the vector bundle F to be the so-called trivial bundle $M \times V$ (the vector space $V_{\boldsymbol{\theta}}$ at $\boldsymbol{\theta}$ is a copy of V); however, this is not necessarily the case.[1] A "point" $\boldsymbol{f}$ in the vector bundle F is a pair $(\boldsymbol{\theta}, \mathbf{v})$, where $\boldsymbol{\theta} \in M$ lies in the manifold and the vector $\mathbf{v} \in V_{\boldsymbol{\theta}}$ lies in the vector space above $\boldsymbol{\theta}$. For particular coordinates $\boldsymbol{\theta} = (\theta^1, \theta^2, \dots, \theta^n)$ on M and a particular basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\dots$, $\mathbf{e}_p$ of $V_{\boldsymbol{\theta}}$, the point $\boldsymbol{f} = (\boldsymbol{\theta}, \mathbf{v})$ in F is represented as the $(n+p)$-vector $\boldsymbol{f} = (\theta^1, \theta^2, \dots, \theta^n, v^1, v^2, \dots, v^p)$, where $\mathbf{v} = \sum_k v^k \mathbf{e}_k$. By abuse of notation, points in the vector bundle may also be written as the direct sum $\boldsymbol{f} = \boldsymbol{\theta} \oplus \mathbf{v}$, in which case it is understood that a particular set of coordinates on M and a basis for $V_{\boldsymbol{\theta}}$ have been specified. In the Euclidean setting $M = \mathbb{R}^n$, we have the trivial vector bundle $F = \mathbb{R}^n \times \mathbb{R}^p$.

Definition 1 *Let M denote a manifold and $F \supset M$ be a vector bundle over M. The* score function $\mathbf{s}$ *is a map*

$$\mathbf{s}\colon M \to F \tag{1}$$

that assigns to every point $\boldsymbol{\theta} \in M$ a unique vector $\mathbf{s}(\boldsymbol{\theta}) \in V_{\boldsymbol{\theta}}$. Such a function is also called a section *of F* [1].

For the case of an arbitrary Riemannian manifold [11], two examples of score functions are:

1. **The error score.** The error score

$$\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \exp_{\boldsymbol{\theta}}^{-1} \hat{\boldsymbol{\theta}} \in T_{\boldsymbol{\theta}}M, \tag{2}$$

is a section of the tangent bundle TM, where $\exp_{\boldsymbol{\theta}}^{-1}$ denotes the inverse exponential map (assuming that it is well-defined). In this case $V_{\boldsymbol{\theta}} = T_{\boldsymbol{\theta}}M$, the tangent plane at $\boldsymbol{\theta}$. The error score and the inverse exponential map are shown in Figure 2. For the Euclidean case $M = \mathbb{R}^n$, $\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$.

2. **The Fisher score.** The Fisher score is the section in the cotangent bundle T^*M defined as

$$\mathbf{s}_{\mathrm{F}}(\boldsymbol{\theta}) = d\ell|_{\boldsymbol{\theta}} \in T^*_{\boldsymbol{\theta}}M, \tag{3}$$

where $\ell(\mathbf{z}|\boldsymbol{\theta})$ is the log-likelihood function and $d\ell$ is the differential [1,11] of ℓ. In this case $V_{\boldsymbol{\theta}} = T^*_{\boldsymbol{\theta}}M$, the cotangent plane at $\boldsymbol{\theta}$. Note that for the Euclidean case $M = \mathbb{R}^n$,

$$d\ell = \frac{\partial \ell}{\partial \boldsymbol{\theta}} = \left(\frac{\partial \ell}{\partial \theta^1}, \frac{\partial \ell}{\partial \theta^2}, \dots, \frac{\partial \ell}{\partial \theta^n} \right) \in R^{1\times n}. \tag{4}$$

3. THE INTRINSIC WEISS-WEINSTEIN REPRESENTATION

Let $f(\mathbf{z}|\boldsymbol{\theta})$ be a statistical model for the manifold-valued parameter $\boldsymbol{\theta} \in M$, and let $\mathbf{s}\colon M \to F$ be a score defined on the vector bundle F. In this section we will derive the generalized Weiss-Weinstein representation [13] for quadratic covariance bounds on the error $\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \exp_{\boldsymbol{\theta}}^{-1} \hat{\boldsymbol{\theta}}$. Looking

[1] A simple counter-example is provided by comparing the infinite cylinder $S^1 \times \mathbb{R}$ (the set-theoretic product of the circle with the real line), and the Möbius strip $S^1 \ltimes \mathbb{R}$ (the so-called "twisted" product of the circle with the real line), in which the real line is twisted 180° around the circle. Topologically, the two-sided cylinder is not equivalent to the one-sided Möbius strip. Therefore, a vector bundle F is not necessarily equivalent to the trivial bundle $M \times V$, though the trivial bundle is in fact encountered most frequently.

ahead, it may be instructive for the reader to contrast the two-channel approach to performance bounds developed here and elsewhere [7, 8, 10, 13] to the one-channel approach used for Cramér-Rao bounds [11, 14]. The two-channel approach allows a direct comparison between the estimation error and the score function. Consider the *Whitney sum* $TM \oplus F$ of the tangent bundle TM with the vector bundle F [1]. This is a vector bundle over M whose vector space at $\boldsymbol{\theta}$ is the direct sum $T_{\boldsymbol{\theta}}M \oplus V_{\boldsymbol{\theta}}$. We will use these technicalities in the analysis of the covariance of the combined two-channel score

$$\mathbf{s}_{\hat{\boldsymbol{\theta}}} \oplus \mathbf{s}\colon M \to TM \oplus F. \tag{5}$$

The estimator bias and mean score vectors are given by the first moments

$$\boldsymbol{b}(\boldsymbol{\theta}) = E[\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})], \quad \text{and} \quad \boldsymbol{\mu}_{\mathbf{s}}(\boldsymbol{\theta}) = E[\mathbf{s}(\boldsymbol{\theta})]. \tag{6}$$

The covariance of the two-channel score $\mathbf{s}_{\hat{\boldsymbol{\theta}}} \oplus \mathbf{s}$ is given by the tensor product

$$\mathbf{C}_{\text{two}} \stackrel{\text{def}}{=} E\big[(\mathbf{s}_{\hat{\boldsymbol{\theta}}} \oplus \mathbf{s} - \boldsymbol{b} \oplus \boldsymbol{\mu}_{\mathbf{s}}) \otimes (\mathbf{s}_{\hat{\boldsymbol{\theta}}} \oplus \mathbf{s} - \boldsymbol{b} \oplus \boldsymbol{\mu}_{\mathbf{s}})\big]. \tag{7}$$

For particular coordinates $\boldsymbol{\theta} = (\theta^1, \theta^2, \ldots, \theta^n)$ on M and a particular basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p$ of $V_{\boldsymbol{\theta}}$, this tensor/Kronecker product becomes

$$\mathbf{C}_{\text{two}} = E\left[\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix}\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix}^{\mathrm{T}}\right] = \begin{pmatrix}\mathbf{C} & \mathbf{T} \\ \mathbf{T}^{\mathrm{T}} & \mathbf{J}\end{pmatrix}, \tag{8}$$

where $\mathbf{C} \stackrel{\text{def}}{=} E[(\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b}) \otimes (\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b})]$ is the n-by-n error covariance matrix, $\mathbf{T} \stackrel{\text{def}}{=} E[(\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b}) \otimes (\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}})]$ the n-by-p cross-covariance, and $\mathbf{J} \stackrel{\text{def}}{=} E[(\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}) \otimes (\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}})]$ the generalized p-by-p information matrix of the score $\mathbf{s}$.

Now we consider the problem of estimating the error $\mathbf{s}_{\hat{\boldsymbol{\theta}}} = \exp_{\boldsymbol{\theta}}^{-1}\hat{\boldsymbol{\theta}}$ of the estimator $\hat{\boldsymbol{\theta}}$ from the score $\mathbf{s}(\boldsymbol{\theta})$. In general, this error estimate is determined by an arbitrary mapping $\boldsymbol{\phi}\colon F \to TM$; however, we will restrict ourselves to a linearization of this mapping, that is, given coordinates and a basis for the vector bundle, an affine transformation of the form

$$\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} = \mathbf{A}(\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}) + \mathbf{a} \tag{9}$$

for some n-by-p matrix $\mathbf{A}$ and n-vector $\mathbf{a}$. A standard result from linear least-squares filtering theory asserts that the minimum mean-squared error estimate of $\mathbf{s}_{\hat{\boldsymbol{\theta}}}$ from $\mathbf{s}$ is

$$\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} = \mathbf{T}\mathbf{J}^{-1}(\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}) + \boldsymbol{b}. \tag{10}$$

It is important to observe that this estimate is intrinsically defined, meaning that it does not depend upon the choice of coordinates on the manifold M or basis in the vector bundle F. The change of coordinates $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}(\boldsymbol{\theta})$ and the change of basis $\mathbf{b}_k = \mathbf{B}\mathbf{e}_k$ for a nonsingular p-by-p matrix $\mathbf{B}$ induce the transformations

$$\mathbf{C} \mapsto \mathbf{A}\mathbf{C}\mathbf{A}^{\mathrm{T}}, \quad \mathbf{T} \mapsto \mathbf{A}\mathbf{T}\mathbf{B}^{\mathrm{T}}, \quad \mathbf{J} \mapsto \mathbf{B}\mathbf{J}\mathbf{B}^{\mathrm{T}}, \tag{11}$$

where $\mathbf{A} = \partial\boldsymbol{\vartheta}/\partial\boldsymbol{\theta}$. Therefore, this estimate transforms as

$$\mathbf{T}\mathbf{J}^{-1} \mapsto \mathbf{A}\mathbf{T}\mathbf{B}^{\mathrm{T}}\mathbf{B}^{-\mathrm{T}}\mathbf{J}^{-1}\mathbf{B}^{-1} = \mathbf{A}\mathbf{T}\mathbf{J}^{-1}\mathbf{B}^{-1} \tag{12}$$

which possesses the correct invariance for linear transformations from $V_{\boldsymbol{\theta}}$ to $T_{\boldsymbol{\theta}}M$.

Quantifying the covariance of the estimate $\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}}$ of the error $\mathbf{s}_{\hat{\boldsymbol{\theta}}}$, we consider the linear transformation of the Whitney sum $TM \oplus F$ defined by

$$\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix} = \begin{pmatrix}\mathbf{I} & -\mathbf{T}\mathbf{J}^{-1} \\ \mathbf{0} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix}. \tag{13}$$

The covariance of this error is given by

$$E\left[\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix}\begin{pmatrix}\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} \\ \mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}\end{pmatrix}^{\mathrm{T}}\right] = \begin{pmatrix}\mathbf{C} - \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}\end{pmatrix}. \tag{14}$$

From this result several fundamental conclusions may be drawn:

1. **Error covariance bound.** The covariance of the zero-mean vector $\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}}$ is $\mathbf{C} - \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}}$, and because this covariance is necessarily nonnegative definite, the error covariance bound is $\mathbf{C} \geq \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}}$.

2. **Estimator efficiency.** The error covariance is minimized when $\mathbf{C} = \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}}$, which occurs if and only if $\mathbf{s}_{\hat{\boldsymbol{\theta}}} = \exp_{\boldsymbol{\theta}}^{-1}\hat{\boldsymbol{\theta}} = \mathbf{T}\mathbf{J}^{-1}(\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}) + \boldsymbol{b}$, in which case $\hat{\boldsymbol{\theta}}$ is said to be an efficient estimator of $\boldsymbol{\theta}$.

3. **Statistically orthogonal errors.** The error $\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}}$ is orthogonal to the centered score $\mathbf{s} - \boldsymbol{\mu}_{\mathbf{s}}$.

4. **Statistically orthogonal decomposition.** The bias-compensated error $\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b}$ decomposes orthogonally as $(\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{b}) + (\mathbf{s}_{\hat{\boldsymbol{\theta}}} - \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}})$, with covariance $\mathbf{C} = \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}} + (\mathbf{C} - \mathbf{T}\mathbf{J}^{-1}\mathbf{T}^{\mathrm{T}})$.

5. **Intrinsic performance bounds.** The quadratic performance bound and all related results are invariant to the choice of coordinates on the parameter manifold M and the choice of basis in the vector bundle F.

4. THE FISHER SCORE

For the case of the Fisher score $\mathbf{s}_{\mathrm{F}} = d\ell$, the information matrix $\mathbf{J}$ is the Fisher information matrix $\mathbf{G} = E[d\ell \otimes d\ell] = -E[\nabla^2\ell]$, where ∇ is the Riemannian connection on M. It can be shown [11] that the cross-covariance in this case is

$$\mathbf{T} = \mathbf{I} + \nabla\boldsymbol{b}, \tag{15}$$

where $\nabla\boldsymbol{b}$ is the covariant differential of the bias vector field $\boldsymbol{b} = E[\mathbf{s}_{\hat{\boldsymbol{\theta}}}] = E[\exp_{\boldsymbol{\theta}}^{-1}\hat{\boldsymbol{\theta}}]$, and the higher-order terms involving the Riemannian curvature have been neglected. The

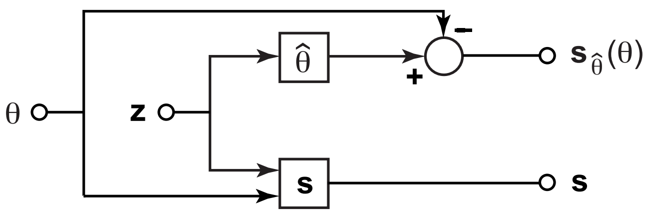

Fig. 3. The estimation error and the score for the Euclidean case $M = \mathbb{R}^n$. In general, the combined two-channel score $\mathbf{s}_{\hat{\boldsymbol{\theta}}} \oplus \mathbf{s}$ lies in the vector bundle $TM \oplus F$.

intrinsic quadratic performance bound is the intrinsic biased Cramér-Rao bound on the error covariance:

$$\mathbf{C} \geq (\mathbf{I} + \nabla \boldsymbol{b})\mathbf{G}^{-1}(\mathbf{I} + \nabla \boldsymbol{b})^{\mathrm{T}} \tag{16}$$

(again neglecting curvature). In the Euclidean, case $M = \mathbb{R}^n$, this simply becomes the biased Cramér-Rao bound [14] with $\nabla \boldsymbol{b} = \partial \boldsymbol{b} / \partial \boldsymbol{\theta}$.

If higher-order covariant derivatives of the log-likelihood ℓ are used, i.e., $\nabla^k \ell$, $k = 1, \ldots, K$, a generalization of the Bhattacharyya bound [4] is obtained.

5. INTRINSIC INVARIANCE AND THE BOBROVSKY-ZAKAI AND BARANKIN SCORES

Here we illustrate the importance of the fact that quadratic performance bounds in the Weiss-Weinstein class are invariant to affine transformations of the score. That is, the transformation $\mathbf{Bs}+\mathbf{c}$ transforms the cross-covariance matrix $\mathbf{T}$ to $\mathbf{TB}^{\mathrm{T}}$ and the matrix $\mathbf{J}$ to $\mathbf{BJB}^{\mathrm{T}}$, as in Equation (11). Therefore, for a change of basis of the vector bundle F represented by the nonsingular p-by-p matrix $\mathbf{B}$,

$$\mathbf{TB}^{\mathrm{T}}(\mathbf{BJB}^{\mathrm{T}})^{-1}\mathbf{BT}^{\mathrm{T}} = \mathbf{TJ}^{-1}\mathbf{T}^{\mathrm{T}}, \tag{17}$$

which is the original bound. In other words, this quadratic performance bound is intrinsically defined on the vector bundle, and does not depend on any arbitrary choice of basis used to represent the score function $\mathbf{s}$.

Given a set of "test points" $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p$, the Bobrovsky-Zakai score [3, 5] is defined as the p-vector

$$\mathbf{s}_{\mathrm{BZ}}(\boldsymbol{\theta}) = \left(\frac{f(\mathbf{z}|\boldsymbol{\theta}_i) - f(\mathbf{z}|\boldsymbol{\theta})}{f(\mathbf{z}|\boldsymbol{\theta})} \right)_{i=1,\ldots,p} = \mathbf{s}_{\mathrm{B}} - 1, \tag{18}$$

where

$$\mathbf{s}_{\mathrm{B}} = \left(\frac{f(\mathbf{z}|\boldsymbol{\theta}_i)}{f(\mathbf{z}|\boldsymbol{\theta})} \right)_{i=1,\ldots,p} \tag{19}$$

is the Barankin score [2]. Clearly, these two scores are simply affine translations of each other; therefore, their intrinsic quadratic performance bounds are the same.

6. CONCLUSIONS

The original quadratic covariance bounds of Weiss-Weinstein may be extended to quadratic performance bounds for manifold-valued parameters, producing intrinsic bounds that are invariant to the choice of coordinates on the manifold M and the choice of basis for the vector bundle over M. One of the key ideas in the extension is that the estimator error score $\mathbf{s}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$ may be viewed as a message channel and the (data) score $\mathbf{s}(\boldsymbol{\theta})$ may be viewed as a measurement channel in a virtual two-channel communication problem. Then the message is estimated as an affine function of the measurement, and standard least squares theory is used to compute the minimum error covariance. This error covariance is used to bound the covariance of the estimator error. This approach, originally developed for vector-value parameters is extended to manifold-valued parameters by generalizing the error score, data score, and the composite covariance structure to be faithful to the geometry of the underlying manifold.

REFERENCES

[1] R. ABRAHAM and J. E. MARSDEN. *Foundations of Mechanics*, 2d ed. Reading, MA: Addison-Wesley, 1985.

[2] E. W. BARANKIN. Locally best unbiased estimates, *Ann. Math. Stat.* **20** : 477–501, 1949.

[3] K. BELL, Y. STEINBERG, and H. L. VAN TREES. Extended Ziv-Zakai lower bound for vector parameter estimation, *IEEE Trans. Information Theory* **43** (2) : 624–637, March 1997.

[4] A. BHATTACHARYYA. On some analogues of the amount of information and their use in statistical estimation, *Sankhyā* **8** : 1–14, 1946.

[5] B. BOBROVSKY and M. ZAKAI. A lower bound on the estimation error for certain diffusion processes. *IEEE Trans. Information Theory* **22** : 45–52, January 1976.

[6] H. CRAMÉR. *Mathematical Methods of Statistics*. Princeton, NJ: Princeton U. Press, 1946.

[7] T. MCWHORTER and L. L. SCHARF. Properties of quadratic performance bounds. In *Proc. 27th Asilomar Conf. Signals, Systems, Computers*, 1993.

[8] ———. Quadratic covariance bounds: properties and representations. Unpublished manuscript.

[9] C. R. RAO. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37** : 81–89, 1945. See also *Selected Papers of C. R. Rao*, ed. S. DAS GUPTA, New York: John Wiley and Sons, Inc. 1994.

[10] L. L. SCHARF. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley Co., 1991.

[11] S. T. SMITH. Covariance, subspace, and intrinsic Cramér-Rao bounds. *IEEE Trans. Signal Processing* **53** (5) : 1610–1630, May 2005.

[12] E. WEINSTEIN and A. J.WEISS. A general class of lower bounds in parameter estimation, *IEEE Trans. Information Theory* **34** : 338–342, March 1988.

[13] A. J. WEISS and E. WEINSTEIN. A lower bound on the mean square error in random parameter estimation, *IEEE Trans. Information Theory* **31** : 680–682, September 1985.

[14] H. L. VAN TREES. *Detection, Estimation, and Modulation Theory*, Part 1. New York: John Wiley and Sons, Inc. 1968.