Maximum Likelihood Attitude and Position Estimation from Pseudo-Range Measurements using Geometric Descent Optimization

A. Alcocer, P. Oliveira, A. Pascoal, and J. Xavier

Abstract— This paper addresses the problem of estimating the attitude and the position of a rigid body when the available measurements consist only of pseudo-ranges between a set of body fixed beacons and a set of earth fixed landmarks. To this effect, a Maximum Likelihood (ML) estimator is formulated as an optimization problem on the parameter space $\Theta = SE(3) \times \mathbb{R}^p$ corresponding to the attitude and position of the rigid body as well as a set of biases present in the pseudo-range equations. Borrowing tools from optimization on Riemannian manifolds, intrinsic gradient and Newton-like algorithms are derived to solve the problem. The rigorous mathematical set-up adopted makes the algorithms conceptually simple and elegant; furthermore, the algorithms do not require the artificial normalization procedures that are recurrent in other estimation schemes formulated in Euclidean space. Supported by recent results on performance bounds for estimators on Riemannian manifolds, the Intrinsic Variance Lower Bound (IVLB) is derived for the problem at hand. Simulation results are presented to illustrate the estimator performance and to validate the tightness of the IVLB in a wide range of signal to noise ratio scenarios.

Index Terms— Navigation systems, attitude/positioning systems, maximum likelihood estimation, optimization on Riemannian manifolds

I. INTRODUCTION

Joint attitude and position estimation systems based on pseudo-range measurements are becoming popular and have received the attention of the research community as an alternative to more complex, expensive, and sophisticated Inertial Navigation Systems. An advantage of such systems is that they are drift-less and insensible to magnetic disturbances. Examples of applications include multiple GPS receiver systems, indoor wireless network positioning systems, and acoustic systems to determine the attitude/position of a body underwater. Range measurements are usually obtained by measuring the time it takes an electromagnetic or acoustic signal to travel between an emitter and a receiver given that the speed of propagation of the signals is assumed to be known. Pseudo-range measurements arise when the emitter and receiver are not synchronized, resulting in an unknown bias term in the range measurements. The range only positioning problem has been extensively treated in the literature, see for example [1] and the references therein.
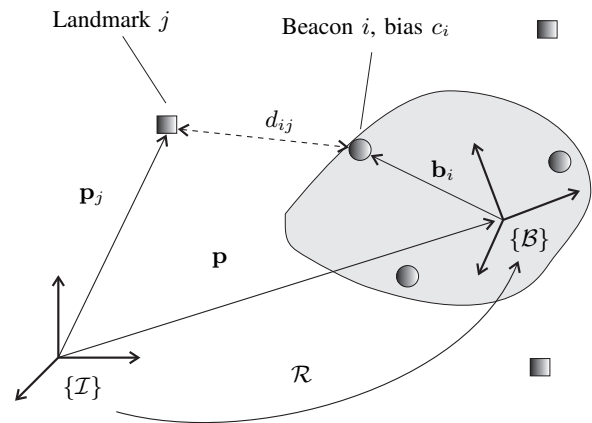
Fig. 1. Geometry of estimation problem.

It is also common to find in the literature work addressing the problem of estimating the attitude (that is the relative orientation of a reference frame with respect to another reference frame) with vector observations [2],[3]. In some applications, vector observations can be obtained from range observations by making the planar waveform appoximation [4],[5]. Despite the fact that the attitude and positioning problems are strongly coupled, a simultaneous treatment of the problem is seldom encountered (see [6] for an example based on line of sight measurements).

This paper addresses the problem of simultaneous attitude and position estimation with pseudo-range measurements. To this effect, a Maximum Likelihood (ML) estimator is formulated by solving an optimization problem on a parameter space containing the attitude and the position of the rigid body as well as a set of bias terms present in the pseudo-range equations. Intrinsic gradient and Newton-like algorithms are derived borrowing tools from optimization on Riemannian manifolds. The rigorous mathematical set-up adopted makes the algorithms conceptually simple and elegant. Furthermore, the algorithms do not require artificial normalization procedures that are recurrent in other estimation schemes formulated in Euclidean space.

II. PROBLEM FORMULATION

Suppose that one is interested in estimating the configuration (that is, position and attitude) of a rigid body in space. Define a reference frame $\{\mathcal{B}\}$ attached to the rigid body and an Inertial reference frame $\{\mathcal{I}\}$. The position of the origin of $\{\mathcal{B}\}$ with respect to $\{\mathcal{I}\}$ can be represented by vector $\mathbf{p} \in \mathbb{R}^3$, and the relative orientation between $\{\mathcal{B}\}$ and $\{\mathcal{I}\}$

can be represented by a rotation matrix $\mathcal{R} \in SO(3)$, where $SO(3)$ is the Special Orthogonal group defined as

$$SO(3) = \left\{ \mathcal{R} \in \mathbb{R}^{3\times 3} : \mathcal{R}^T \mathcal{R} = I_3,\ \det(\mathcal{R}) = 1 \right\}. \quad (1)$$

In the above expression I_3 stands for the 3×3 identity matrix and $\det(\cdot)$ is the matrix determinant operator. The configuration space of the rigid body can then be identified with the Special Euclidean group $SE(3) = SO(3) \times \mathbb{R}^3$ (as a set). Elements of $SE(3)$ will be identified with the pair $(\mathcal{R}, \mathbf{p})$ or with the vector $\left[\text{vec}(\mathcal{R})^T\ \mathbf{p}^T \right]^T$ (where $\text{vec}(\cdot)$ is the operator that stacks the columns of a matrix from left to right) depending on what representation is more suitable for the computations.

Suppose the rigid body has p beacons and assume that the location of the beacons with respect to $\{\mathcal{B}\}$ is known. The beacons could be GPS antennas, or underwater acoustic transponders, arranged with a certain known geometry in the rigid body. Let us further consider that there are m fixed landmarks distributed in the ambient space with known positions: for example, GPS satellites or surface buoys equipped with hydrophones. Let $\mathbf{b}_i \in \mathbb{R}^3, i \in \{1, \ldots, p\}$ denote the positions of the p beacons in the rigid body expressed in $\{\mathcal{B}\}$ and let $\mathbf{p}_j \in \mathbb{R}^3, j \in \{1, \ldots, m\}$ denote the positions of the m landmarks expressed in $\{\mathcal{I}\}$ (see Fig. 1).

Let d_{ij} denote the distance between the i'th beacon and the j'th landmark. Let $c_i \in \mathbb{R}; i \in \{1, \cdots, p\}$ be an unknown beacon dependent bias term. Then the pseudo-ranges are defined as

$$\rho_{ij} = d_{ij} + c_i = \| \mathcal{R}\mathbf{b}_i + \mathbf{p} - \mathbf{p}_j \| + c_i \quad (2)$$

with $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots, m\}$, and $\| \cdot \|$ stands for the Euclidean norm. The observations y_{ij}, defined by $y_{ij} = \rho_{ij} + w_{ij}$ consist of the pseudo-ranges ρ_{ij} corrupted by an additive zero mean Gaussian disturbances w_{ij}. Note that the terms *beacon* and *landmark* are used to illustrate the problem but they do not impose any particular role (in terms of emitter/receiver). They should be simply viewed as *pseudo-range measuring devices*.

It is convenient to stack all the ij elements in a more compact form by defining the vectors

$$\mathbf{d} \triangleq \left[[d_{11}\ \cdots\ d_{p1}]\ \cdots\ [d_{1m}\ \cdots\ d_{pm}] \right]^T \in \mathbb{R}^{mp}, \quad (3)$$

and $\rho, \mathbf{y}, \mathbf{w} \in \mathbb{R}^{mp}$ in a similar way. Define also a vector $\mathbf{c} = [c_1\ \cdots\ c_p]^T \in \mathbb{R}^p$ containing all the bias terms. With this arrangement, the observations can then be written in a more compact form as

$$\rho = \mathbf{d} + (\mathbf{1}_m \otimes I_p)\mathbf{c} \quad (4)$$

$$\mathbf{y} = \rho + \mathbf{w}, \quad \mathrm{E}\left\{ \mathbf{w}\mathbf{w}^T \right\} \triangleq \mathbf{R} \in \mathbb{R}^{mp \times mp} \quad (5)$$

where $\otimes$ denotes the Kronecker product of matrices, $\mathbf{1}_m$ is a $m \times 1$ vector of ones, and $\mathbf{R}$ is the covariance matrix of $\mathbf{w}$. Note that no assumption is made on the structure of $\mathbf{R}$, thus allowing for a number of interesting mission scenarios and configurations. For instance, the measurements of the distances between an Earth fixed landmark and all the body beacons can suffer from highly correlated disturbances. This might be due to the fact that the observations originate from signals that have travelled almost through the same propagation channel. In these cases the covariance matrix $\mathbf{R}$ is close to being block diagonal.

In order to determine the attitude and the position of the rigid body from the pseudo-range measurements, the vector of unknown biases $\mathbf{c} \in \mathbb{R}^p$ needs also to be estimated. It is common in the literature to avoid estimating the biases by subtracting pseudo-range equations among them to obtain range-differences. This approach will not be pursued in this paper. Instead, an augmented parameter space Θ will be considered containing the unknown bias terms. The final parameter space will be defined as the Cartesian product $\Theta = SE(3) \times \mathbb{R}^p = SO(3) \times \mathbb{R}^3 \times \mathbb{R}^p$.

A. ML Estimator Formulation

The Maximum Likelihood (ML) Estimator determines the triad $(\widehat{\mathcal{R}}, \widehat{\mathbf{p}}, \widehat{\mathbf{c}})_{ML} \in \Theta$ that maximizes the *likelihood* function, that is, the probability $p(\mathbf{y}|\mathcal{R}, \mathbf{p}, \mathbf{c})$ of obtaining the observations $\mathbf{y}$ given the parameters $(\mathcal{R}, \mathbf{p}, \mathbf{c})$ [7] [8]. According to the Gaussianity assumption on the disturbances, the ML estimator can be found by solving the optimization problem

$$\left(\widehat{\mathcal{R}}, \widehat{\mathbf{p}}, \widehat{\mathbf{c}} \right)_{\mathrm{ML}} = \arg \min_{(\mathcal{R}, \mathbf{p}, \mathbf{c}) \in \Theta} f(\mathcal{R}, \mathbf{p}, \mathbf{c}) \quad (6)$$

where $f : \Theta \to \mathbb{R}$ is given by

$$f(\mathcal{R}, \mathbf{p}, \mathbf{c}) = \frac{1}{2}(\mathbf{y} - \rho)^T \mathbf{R}^{-1}(\mathbf{y} - \rho) \quad (7)$$

and, from (2), $\rho = \rho(\mathcal{R}, \mathbf{p}, \mathbf{c})$. Note that the cost function f is not differentiable when some of the ranges d_{ij} vanish, that is, when the position of a beacon coincides with the position of a landmark. It is realistic to assume that this situation never occurs in practice.

III. OPTIMIZATION ON RIEMANNIAN SUBMANIFOLDS

In order to determine the ML Estimator, the function f needs to be minimized over the parameter space Θ. At this point, it is not obvious how to continue and effectively solve this constrained optimization problem. The problem does not admit a closed form solution so it is necessary to resort to an iterative scheme. There has been some work on generalizing the classic gradient descent and Newton methods to Riemannian manifolds [9] [10]. The key idea involved in [9] is to generalize the classic gradient and Newton-like directions and to perform line searches along geodesics, as depicted in Fig. 2.

In many cases, such as the one considered in this paper, the parameter space Θ can be shown to be an embedded Riemannian submanifold of some ambient Euclidean space $\mathbb{R}^n$. For instance, when Θ is characterized as the regular level set of some smooth function i.e., $\Theta = \{x \in \mathbb{R}^n : h(x) = 0\}$. It turns out that many intrinsic objects in Θ, such as intrinsic gradients and Hessians, can be easily obtained from the more familiar corresponding extrinsic objects in $\mathbb{R}^n$. In what
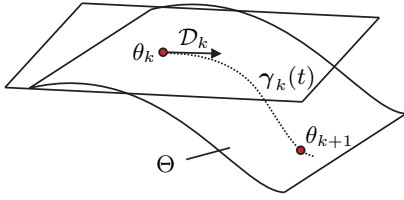
Fig. 2. Geometric descent optimization. Intrinsic gradient or Newton directions followed by geodesic line searches.

follows it is assumed that the reader is familiar with the concepts of Riemannian geometry [11], [12].

Let (Θ, g) be an embedded Riemannian submanifold of $\mathbb{R}^n$. The tangent space at $\theta \in \Theta \subset \mathbb{R}^n$, denoted $T_\theta\Theta$, can be identified with some linear subspace of $\mathbb{R}^n$. This means that the tangent vectors of Θ, which are usually quite abstract objects, can be identified with very familiar objects: vectors in $\mathbb{R}^n$. For each $\theta \in \Theta \subset \mathbb{R}^n$, the tangent space $T_\theta\mathbb{R}^n \simeq \mathbb{R}^n$ can be split into the direct sum $T_\theta\Theta \oplus N_\theta\Theta$ where $N_\theta\Theta$ is the orthogonal complement of the linear subspace $T_\theta\Theta$ in $\mathbb{R}^n$ [12, p.125].

Let $\widetilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ be a smooth function and let $f : \Theta \rightarrow \mathbb{R}$, be the restriction of $\widetilde{f}$ to Θ, that is, $f = \widetilde{f}|_\Theta$. The extrinsic gradient of f at θ is defined as the usual vector of partial derivatives $\nabla \widetilde{f}|_\theta = [\frac{\partial}{\partial x_1}\widetilde{f}|_\theta \cdots \frac{\partial}{\partial x_n}\widetilde{f}|_\theta]^T$. Moreover, according to the above discussion, the extrinsic gradient can be decomposed as the sum of two vectors in $\mathbb{R}^n$ belonging to $T_\theta\Theta$ and $N_\theta\Theta$. The intrinsic gradient of f at θ, denoted $grad f|_\theta$ is a tangent vector in $T_\theta\Theta$. It provides a way of locally approximating functions and obtaining generalized (intrinsic) gradient descent search directions. The extrinsic Hessian of f at θ is a map $\text{Hess}\widetilde{f} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ that can be obtained from the usual matrix of second order partial derivatives of $\widetilde{f}$. The intrinsic Hessian of f at θ is also a map $\text{Hess}f : T_\theta\Theta \times T_\theta\Theta \rightarrow \mathbb{R}$.

The next propositions summarize fairly well known results, the proofs of which are straightforward exercises in Riemannian geometry (see [13] or [14]):

Proposition 1 (Intrinsic Gradient): The intrinsic gradient of f at θ, denoted $grad f|_\theta \in T_\theta\Theta$, is exactly the projection of the extrinsic gradient $\nabla \widetilde{f}|_\theta$ onto the tangent space $T_\theta\Theta$.

Proposition 2 (Intrinsic Hessian): The intrinsic Hessian of the smooth function f at θ satisfies

$$Hess f(\Delta_1, \Delta_2) = Hess\widetilde{f}(\Delta_1, \Delta_2) + \langle II(\Delta_1, \Delta_2), \nabla\widetilde{f}|_\theta\rangle \tag{8}$$

for any tangent vectors $\Delta_1, \Delta_2 \in T_\theta\Theta$. In the above expression, $II : T_\theta\Theta \times T_\theta\Theta \rightarrow N_\theta\Theta$ is the second fundamental form, which relates the geometries of Θ and $\mathbb{R}^n$ [12]. These propositions will provide a simple way of determining intrinsic gradient descent and Newton-like search directions for the problem at hand.

IV. THE GEOMETRY OF THE PARAMETER SPACE

The parameter space Θ was defined as the Cartesian product $SO(3) \times \mathbb{R}^3 \times \mathbb{R}^p$. It will be regarded as an embedded

Riemannian submanifold of the ambient Euclidean space $\mathbb{R}^{3\times3} \times \mathbb{R}^3 \times \mathbb{R}^p$. There are many references in the literature about $SE(3)$ and $SO(3)$, the reader is referred to [15], [16], [17], and [18].

Let $\theta = (\mathcal{R}, \mathbf{p}, \mathbf{c})$ be an element of Θ. The tangent space $T_\theta\Theta$ can be identified with the linear subspace

$$\left\{(\mathcal{R}S, \mathbf{v}, \mathbf{u}) \in \mathbb{R}^{3\times3} \times \mathbb{R}^3 \times \mathbb{R}^p : S \in K(3, \mathbb{R})\right\}, \tag{9}$$

where $K(3, \mathbb{R})$ stands for the set of 3×3 skew symmetric matrices with real entries (recall that a matrix S is skew symmetric if $S + S^T = 0$). Recall also that sometimes it will be convenient to identify $T_\theta\Theta$ with vectors of the form $\left[\text{vec}(\mathcal{R}S)^T \ \mathbf{v}^T \ \mathbf{u}^T\right]^T$ in $\mathbb{R}^{12+p}$.

A Riemannian metric g is a smooth assignment of an inner product to each tangent space. The parameter space Θ can be made an embedded Riemannian submanifold by providing it with the canonical Riemannian metric inherited from its ambient Euclidean space. Let two tangent vectors $\Delta_1, \Delta_2 \in T_\theta\Theta$ have the form

$$\Delta_1 = \begin{bmatrix} \text{vec}(\Omega_1) \\ \mathbf{v}_1 \\ \mathbf{u}_1 \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} \text{vec}(\Omega_2) \\ \mathbf{v}_2 \\ \mathbf{u}_2 \end{bmatrix}. \tag{10}$$

Then $g(\Delta_1, \Delta_2)$, also denoted $\langle\Delta_1, \Delta_2\rangle$, becomes

$$\langle\Delta_1, \Delta_2\rangle = \Delta_1^T\Delta_2. \tag{11}$$

Note that, when restricted to $SE(3)$, the induced Riemannian metric of (11) is in fact equivalent to the *scale-dependent left-invariant metric* of [17] (with $c = 2$, and $d = 1$).

With this choice of metric, the geodesics of Θ have a simple closed form expression (see [17], [18] for a related discussion). Let $\theta = (\mathcal{R}, \mathbf{p}, \mathbf{c}) \in \Theta$ and $\Delta = (\Omega, \mathbf{v}, \mathbf{u}) \in T_\theta\Theta$. The geodesic emanating from θ in the direction of Δ is given by $\gamma : \mathbb{R} \rightarrow \Theta$,

$$\gamma(t) = (\mathcal{R}\exp(\mathcal{R}^T\Omega t), \mathbf{p} + \mathbf{v}t, \mathbf{c} + \mathbf{u}t) \tag{12}$$

where $\exp$ is the matrix exponential. Note that $\gamma(0) = \theta$ and $\dot{\gamma}(0) = \Delta$.

The second fundamental form, which will be used to determine the intrinsic Hessian, can also be simply computed (see [13] for details). Let $\Delta_1, \Delta_2 \in T_\theta\Theta$ be tangent vectors at some point $\theta \in \Theta$ as in (10). The second fundamental form $II : T_\theta\Theta \times T_\theta\Theta \rightarrow N_\theta\Theta$ can be found to be

$$II(\Delta_1, \Delta_2) = (-\mathcal{R}\text{symm}(\Omega_1^T\Omega_2), 0, 0) \tag{13}$$

where symm is the operator that extracts the symmetric part of a matrix, that is symm$(A) = \frac{1}{2}(A + A^T)$.

V. INTRINSIC GRADIENT AND NEWTON ALGORITHMS

In order to determine the maximum likelihood (ML) estimate of the attitude and position of a rigid body, a constrained optimization problem on the parameter space Θ needs to be solved. The geometric descent optimization algorithms proposed in this paper are of iterative nature. At each iteration a gradient or Newton-like search direction $\mathcal{D}_k$ is determined. Then, an intrinsic geodesic line search is performed along $\mathcal{D}_k$ to update the estimated parameter.

A. Intrinsic gradient and Newton directions

According to Proposition 1, the intrinsic gradient of a smooth function defined on an embedded Riemannian submanifold of some Euclidean space can be found by determining the extrinsic gradient (which is the usual vector of partial derivatives as if the function were defined on the ambient Euclidean space $\widetilde{f} : \mathbb{R}^{3\times3}\times\mathbb{R}^3\times\mathbb{R}^p \to \mathbb{R}$, instead of $f : \Theta \to \mathbb{R}$) and projecting it orthogonally onto the tangent space of the parameter space. The derivation of the extrinsic gradient for the problem at hand is done in detail in [13]. Let $\theta = (\mathcal{R}, \mathbf{p}, \mathbf{c})$ be the point in Θ at which to evaluate the extrinsic gradient. Define the matrices

$$\mathbf{F} = \begin{bmatrix} \mathbf{b}_1^T \otimes (\mathbf{p} - \mathbf{p}_1)^T \\ \vdots \\ \mathbf{b}_i^T \otimes (\mathbf{p} - \mathbf{p}_j)^T \\ \vdots \\ \mathbf{b}_p^T \otimes (\mathbf{p} - \mathbf{p}_m)^T \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \mathbf{b}_1^T \mathcal{R}^T + (\mathbf{p} - \mathbf{p}_1)^T \\ \vdots \\ \mathbf{b}_i^T \mathcal{R}^T + (\mathbf{p} - \mathbf{p}_j)^T \\ \vdots \\ \mathbf{b}_p^T \mathcal{R}^T + (\mathbf{p} - \mathbf{p}_m)^T \end{bmatrix},$$
(14)

$$\mathbf{D} = \mathrm{diag}(\mathbf{d}) = \begin{bmatrix} d_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{mp} \end{bmatrix}$$
(15)

where $\mathbf{F} \in \mathbb{R}^{mp\times9}$, $\mathbf{C} \in \mathbb{R}^{mp\times3}$, and $\mathbf{D} \in \mathbb{R}^{mp\times mp}$. Then the extrinsic gradient has the form

$$\nabla\widetilde{f}|_\theta \triangleq \begin{bmatrix} \mathrm{vec}(\widetilde{G}_\mathcal{R}) \\ \widetilde{G}_\mathbf{p} \\ \widetilde{G}_\mathbf{c} \end{bmatrix} = - \begin{bmatrix} \mathbf{F}^T\mathbf{D}^{-1} \\ \mathbf{C}^T\mathbf{D}^{-1} \\ (\mathbf{1}_m^T \otimes I_p) \end{bmatrix} \mathbf{R}^{-1}(\mathbf{y} - \rho).$$
(16)

According to Proposition 1 the intrinsic gradient can be found by solving the projection problem

$$\mathrm{grad}f|_\theta = \arg\min_{\Delta \in T_\theta\Theta} \langle \Delta - \nabla\widetilde{f}|_\theta, \Delta - \nabla\widetilde{f}|_\theta \rangle.$$
(17)

Using some algebraic manipulations and the fact that $\Delta = [\mathrm{vec}(\mathcal{R}S)^T \ \mathbf{v}^T \ \mathbf{u}^T]^T$ for some skewsymmetric matrix S and vectors $\mathbf{v}, \mathbf{u}$, it can be shown that (see [13] and also [9] for a related result):

$$\mathrm{grad}f|_\theta \triangleq \begin{bmatrix} \mathrm{vec}(G_\mathcal{R}) \\ G_\mathbf{p} \\ G_\mathbf{c} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\mathrm{vec}(\widetilde{G}_\mathcal{R} - \mathcal{R}\widetilde{G}_\mathcal{R}^T\mathcal{R}) \\ \widetilde{G}_\mathbf{p} \\ \widetilde{G}_\mathbf{c} \end{bmatrix}.$$
(18)

The Newton-like search direction is the unique tangent vector $N \in T_\theta M$ satisfying $\mathrm{Hess}f(X, N) = -\langle X, \mathrm{grad}f\rangle$ for all tangent vectors $X \in T_\theta M$ (assuming that $\mathrm{Hess}f$ is nonsingular at θ). It can be found by determining the usual matrix of second order partial derivatives and solving a certain linear system of equations. Its derivation is omitted due to space constraints. A similar derivation is done in detail in [13] and [14].

B. Intrinsic geodesic line search

Let $\theta_k \in \Theta$ be the parameter estimate at iteration k. Let $\mathcal{D}_k = (\mathcal{D}_\mathcal{R}, \mathcal{D}_\mathbf{p}, \mathcal{D}_\mathbf{c}) \in T_{\theta_k}\Theta$ be the search direction at iteration k, i.e, $\mathcal{D}_k = -\mathrm{grad}f|_{\theta_k}$ (gradient descent) or

$\mathcal{D}_k = N$ (Newton descent). A line search can be performed along the geodesic

$$\gamma_k(t) = (\mathcal{R}\exp(\mathcal{R}^T\mathcal{D}_\mathcal{R}t), \mathbf{p} + \mathcal{D}_\mathbf{p}t, \mathbf{c} + \mathcal{D}_\mathbf{c}t).$$
(19)

Ideally, the line search procedure aims at finding the optimal stepsize t_k^* satisfying

$$t_k^* = \arg\min_{t\in\mathbb{R}} f(\gamma_k(t)).$$
(20)

This optimization subproblem can be hard to solve. Alternatively, it is common to solve this problem only approximately using for instance the Armijo rule [19, p.29]. The Armijo rule selects $t_k = \beta^{m_i}s$, where $m_i \in \{0, 1, 2, \dots\}$ is the first integer satisfying

$$f(\theta_k) - f(\gamma_k(\beta^{m_i}s)) \geq -\sigma\beta^{m_i}s \langle \mathrm{grad}f|_{\theta_k}, \mathcal{D}_k\rangle$$
(21)

for some constants $s > 0$, and $\beta, \sigma \in (0, 1)$.

C. Algorithm implementations: an outline

The theoretical results of the previous sections lead to the following implementations of the proposed intrinsic gradient and Newton algorithms:

---

1) *Start at initial estimate $\theta_0 = (\mathcal{R}_0, \mathbf{p}_0, \mathbf{c}_0) \in \Theta$. Set $k = 0$.*
2) *Determine a search direction $\mathcal{D}_k$:*
   a) *Intrinsic gradient: take $\mathcal{D}_k = -\mathrm{grad}f|_{\theta_k}$ according to (18).*
   b) *Intrinsic Newton: If $\langle N, \nabla f|_{\theta_k}\rangle < 0$ (descent condition) take $\mathcal{D}_k = N$, otherwise take $\mathcal{D}_k = -\mathrm{grad}f|_{\theta_k}$.*
3) *Line search along the geodesic $\gamma_k(t)$ (19), using Armijo rule to determine step size t_k.*
4) *update estimate $\theta_{k+1} = \gamma_k(t_k)$. Set $k = k + 1$.*
5) *If $\|\mathrm{grad}f|_{\theta_k}\| \leq \epsilon$, stop. Otherwise return to 2.*

VI. SIMULATION RESULTS

Simulations were performed to validate the proposed algorithms. A simulation setup was chosen that is similar to a multi-antenna GPS attitude/positioning system. Four ($m = 4$) earth fixed landmarks were located at positions corresponding to a favorable GPS satellite constellation. That is,

$$[\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3 \ \mathbf{p}_4] = R_{sat}\begin{bmatrix} 0.966 & 0 & 0.259 & 0 \\ -0.483 & 0.837 & 0.259 & 0 \\ 0.259 & 0.259 & 0.259 & 1 \end{bmatrix}$$
(22)

where $R_{sat} = 2.656 \cdot 10^7$ m. Three ($p = 3$) body fixed beacons were located at positions $[\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3] = 5\,I_3$ m. The observation error covariance was set to $\mathbf{R} = \sigma^2 I_{mp}$, with $\sigma = 0.1$ m. In order to illustrate the attitude estimation errors (in Fig. 3 and Fig. 4) exponential coordinates for $SO(3)$ were used. That is, the vector $\theta = [\theta_1 \ \theta_2 \ \theta_3]^T$ is used to
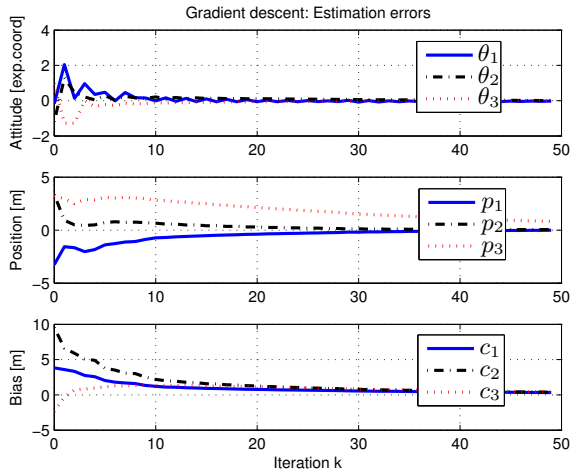
Fig. 3. Intrinsic gradient algorithm estimation errors: Attitude, Position and Bias. The attitude error is represented as the entries of vector $\tilde{\theta}$, the exponential coordinates of $\tilde{\mathcal{R}}_k = \hat{\mathcal{R}}_k^T \mathcal{R}$. The position and bias errors are shown as the entries of vectors $\tilde{\mathbf{p}} = \mathbf{p} - \tilde{\mathbf{p}}_k$ and $\tilde{\mathbf{c}} = \mathbf{c} - \tilde{\mathbf{c}}_k$, respectively. The algorithm has a very slow convergence.

represent the orthogonal matrix $\mathcal{R} = \exp\left\{ \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix} \right\}$.
The actual and initial parameters were:

$$\mathcal{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \hat{\mathcal{R}}_0 = \begin{bmatrix} 0.032 & -0.542 & 0.704 \\ 0.710 & 0.604 & -0.584 \\ -0.363 & -0.363 & 0.404 \end{bmatrix} \quad (23)$$

$$\mathbf{p} = \begin{bmatrix} 0.738 \\ 0.358 \\ -0.075 \end{bmatrix}, \quad \hat{\mathbf{p}}_0 = \begin{bmatrix} 3.991 \\ -2.993 \\ -3.299 \end{bmatrix} \quad (24)$$

$$\mathbf{c} = \begin{bmatrix} -160.331 \\ 33.937 \\ -13.113 \end{bmatrix}, \quad \hat{\mathbf{c}}_0 = \begin{bmatrix} -164.155 \\ 24.403 \\ -10.778 \end{bmatrix} \quad (25)$$

Fig. 5 shows the evolution of the cost function together with the norm of the gradient for the intrinsic gradient and Newton methods. Fig. 3 and Fig. 4 show the evolution of the attitude, position, and bias estimation errors for both methods. The gradient method shows a very poor (linear) convergence rate whereas the Newton method achieves quadratic convergence when close to the minimum.

Note that the cost function has local minima that depend highly on the number and position of the beacons and landmarks, as well as the intensity of the measurement noise. This is an important issue, as well as on many optimization problems, that prevents the methods from being global. Preliminary experiences suggest that local minimum solutions can be easily detected as they yield final likelihoods some orders of magnitude different from the global minimum.

VII. PERFORMANCE BOUNDS

Given a set of noisy pseudo-range measurements, there is a fundamental limitation on the size of the estimation errors that can be achieved. When the state space is Euclidean, the Cramér-Rao Bound (CRB) is the classic tool to determine
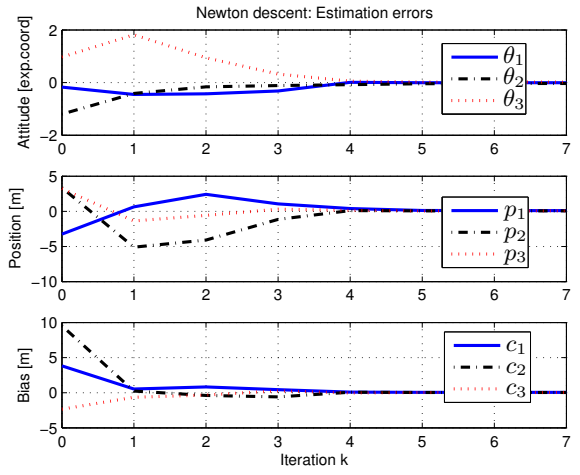

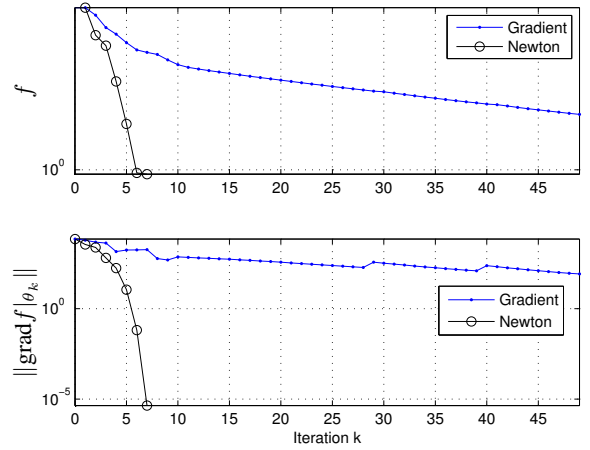Fig. 4. Intrinsic Newton algorithm estimation erros: Attitude, Position and Bias.

Fig. 5. Intrinsic gradient and Newton algorithms: performance evaluation. Value of the cost function f (top) and norm of the intrinsic gradient $\|\mathrm{grad}f|_{\theta_k}\|$ (bottom).

such a limitation [7]. When the state space is not Euclidean, as in our case, one should resort to generalizations of the CRB as the Intrinsic Variance Lower Bound (IVLB) [20]. The IVLB uses the intrinsic (geodesic) Riemannian distance to quantify the estimation errors. More specifically, it sets a lower bound on the intrinsic variance of unbiased estimators. Let θ be the true parameter and $\widehat{\theta}$ an estimate of it. The intrinsic variance is defined as

$$\mathrm{var}\left\{ \widehat{\theta} \right\} = \mathsf{E}\left\{ d_\ominus(\widehat{\theta}, \theta)^2 \right\} \quad (26)$$

where $d_\ominus(\cdot)$ is the intrinsic (geodesic) distance function in Θ. Considering the canonical metric in (11) it becomes

$$d_\ominus(\theta_1, \theta_2)^2 = d_{SO(3)}(\mathcal{R}_1, \mathcal{R}_2)^2 + \|\mathbf{p}_1 - \mathbf{p}_2\|^2 + \|\mathbf{c}_1 - \mathbf{c}_2\|^2, \quad (27)$$

where $\theta_1 = (\mathcal{R}_1, \mathbf{p}_1, \mathbf{c}_1)$, $\theta_1 = (\mathcal{R}_2, \mathbf{p}_2, \mathbf{c}_2)$, and

$$d_{SO(3)}(\mathcal{R}_1, \mathcal{R}_2) = \sqrt{2} \arccos\left(\frac{\mathrm{tr}\left(\mathcal{R}_1^T \mathcal{R}_2\right) - 1}{2} \right) \quad (28)$$
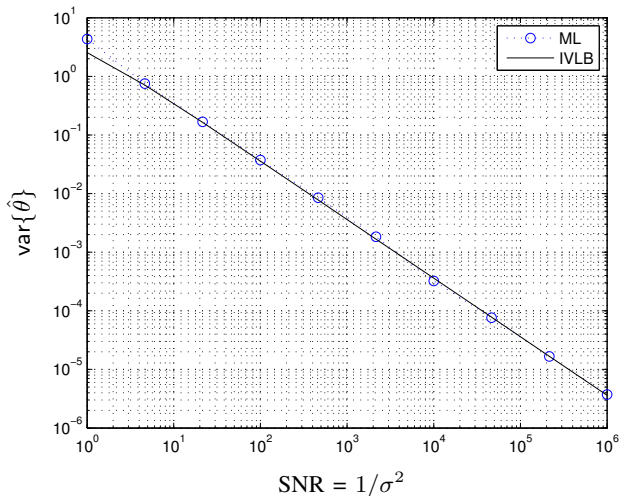
Fig. 6. IVLB compared to Monte carlo Newton ML at different SNR scenarios.

is the canonical metric on $SO(3)$ (as a submanifold of $\mathbb{R}^{n \times n}$) [17][15].

The (extrinsic) Fisher Information Matrix for the problem at hand, that is used in the computation of the IVLB (see [20] and [14] for a similar derivation), can be computed as

$$
\begin{aligned}
I(\theta) &= E_\theta \left\{ (\nabla_\theta \log p(\theta|\mathbf{y})(\nabla_\theta \log p(\theta|\mathbf{y})^T \right\} \\
&= \begin{bmatrix} \mathbf{F}^T \mathbf{D}^{-1} \\ \mathbf{C}^T \mathbf{D}^{-1} \\ (\mathbf{1}_m^T \otimes I_p) \end{bmatrix} \mathbf{R}^{-1} \begin{bmatrix} \mathbf{F}^T \mathbf{D}^{-1} \\ \mathbf{C}^T \mathbf{D}^{-1} \\ (\mathbf{1}_m^T \otimes I_p) \end{bmatrix}^T .
\end{aligned}
\tag{29}
$$

It can also be used as a measure of the observability of the problem. Let $U_\theta \in \mathbb{R}^{12+p \times 6+p}$ be a matrix whose columns form an orthonormal basis for $T_\theta \Theta$. If the matrix $U_\theta^T I(\theta) U_\theta \in \mathbb{R}^{6+p \times 6+p}$ is singular or badly conditioned, this indicates that with the present estimation set up, in terms of the number of landmarks/beacons and their relative geometry, it is not possible to determine all of the unknown parameters.

The IVLB for the same simulation setup of Section VI is shown in Fig. 6 and compared with experimental Monte Carlo ML Newton runs at different signal to noise ratio (SNR) conditions, where SNR$\simeq 1/\sigma^2$, and $\mathbf{R} = \sigma I_{mp}$. At each point, 100 Monte Carlo ML algorithm runs were performed initialized at the true value of the parameter. The SNR range plotted corresponds to standard deviations of $\sigma \in [10^{-3}, 1]$ m. It is hard to distinguish between the experimental ML performance and the IVLB. This shows the good performance of the estimator and the tightness of the derived bound.

VIII. CONCLUSIONS AND FUTURE WORK

In this paper intrinsic gradient and Newton like algorithms were derived to solve the problem of simultaneous position and attitude estimation with pseudo-range only measurements. The algorithms are relatively simple and avoid the need for any normalization procedure since the iterates evolve naturally on the parameter space. Simulation results showed that the intrinsic gradient algorithm has a slow convergence. However, the intrinsic Newton algorithm converged in a few iterations and attained performances very close to the Intrinsic Variance Lower bound (IVLB). Hence, the intrinsic Newton algorithm is suitable for real time implementation. Future work will include the study of nonlinear filters which take into account the dynamics of the rigid body. Another topic of interesting research is the inclusion of other sensorial data such as bearing, doppler, and vision information. Testing the algorithms with real experimental data is also a subject that warrants further efforts.

REFERENCES

[1] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-35 Issue: 12*, pp. 1661–1669, December 1987.

[2] G. Wahba, "A Least-Squares Estimate of Spacecraft Attitude," *SIAM Review, Vol. 7, No.3*, p. 409, July 1965.

[3] A. Nadler, Y. Bar-Itzhack, and H. Weiss, "On Algorithms for Attitude Estimation using GPS," in *Proceedings of the 2000 IEEE Conference on Decision and Control, Sidney, Australia*, December 2000, pp. 3321–3326.

[4] X. Yu, K. Ling, and N. Nagaragan, "LEO Satellite Attitude Determination Based on GPS," in *Proceedings of the 1999 IEEE Conference on Decision and Control, Phoenix, Arizona USA*, December 1999, pp. 1998–1992.

[5] J. L. Crassidis and F. L. Markley, "New Algorithm for Attitude Determination Using Global Positioning System Signals," *AIAA Journal of Guidance, Control, and Dynamics, Vol. 20, No. 5*, pp. 891–896, Sept.-Oct. 1997.

[6] J. L. Crassidis, R. Alonso, and J. L. Junkins, "Optimal Attitude and Position Determination from Line-of-Sight Measurements," *The Journal of the Astronautical Sciences, Vol. 48, Nos. 2-3*, pp. 391–408, April-Sept. 2000.

[7] H. L. V. Trees, *Detection, Estimation, and Modulation Theory (Part I)*. John Wiley and Sons, 1968.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey, USA: Prentice-Hall, 1993.

[9] A. Edelman, T. A. Arias, and S. T. Smith, "The Geometry of Algorithms with Orthogonality Constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303–353, 1998.

[10] J. Manton, "Optimization Algorithms Exploiting Unitary Constraints," *IEEE Trans. Signal Processing, 50(3):635–650*, March 2002.

[11] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. New York: Academic Press, 1975.

[12] M. do Carmo, *Riemannian Geometry*. Boston: Birkhäuser, 1992.

[13] A. Alcocer, P. Oliveira, A. Pascoal, and J. Xavier, "Estimation of Attitude and Position from Range only Measurements using Geometric Descent Optimization on the Special Euclidean Group," January 2006, Tech. Report, ISR, Lisbon.

[14] ——, "Estimation of Attitude and Position from Range only Measurements using Geometric Descent Optimization on the Special Euclidean Group," in *International Conference on Information Fusion, FUSION'06, Florence, Italy.*, July 2006.

[15] M. Moakher, "Means and Averaging in the Group of Rotations," *SIAM Journal on Matrix Analysis and Applications 24/1*, pp. 1–16, 2002.

[16] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL, USA: CRC Press, 1994.

[17] F. C. Park, "Distance Metrics on the Rigid Body Motions with Applications to Mechanism Design," *Transactions of the ASME Journal of Mechanical Design, Vol.117(1), pp:48–54.*, March 1995.

[18] M. Žefran, V. Kumar, and C. B. Croke, "On the Generation of Smooth Three-Dimensional Rigid Body Motions," *IEEE Trans. Robotics and Automation, 14(4):576–589*, August 1998.

[19] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995, 2nd edition 1999.

[20] J. Xavier and V. Barroso, "Intrinsic Variance Lower Bound (IVLB): an extension of the Cramér-Rao bound to Riemannian manifolds," in *IEEE Int. Conf. on Acoust., Sp. and Sig. Proc. (ICASSP)*, March 2005.