On achievable accuracy for pose tracking

Andrea Censi

***Abstract*—This paper presents Cramér-Rao bound-like inequalities for pose tracking, which is defined as the problem of recovering the robot displacement given two successive readings of a relative sensor. Computing the exact Fisher Information Matrix (FIM) for pose tracking is hard, because the state comprises the map, which is infinite-dimensional and unknown. This paper shows that the FIM for pose tracking can be bounded by a function of the FIM for localization on a known map, thereby reducing the analysis to a finite-dimensional problem. The resulting bounds are independent of the map prior and representation. The results are valid for any relative sensor; the experimental verification is done for the particular case of pose tracking using range-finders (scan matching).**

I. Introduction

Localization, pose tracking, and simultaneous localization and mapping (SLAM) using "dense" sensors, such as laser range-finders, are attractive solutions for robot navigation, because they offer good precision and do not rely on assumptions about the environment. However, it is hard to analyze such algorithms and obtain proofs of convergence, optimality, consistency, etc. For example, consider the problem of pose tracking with range-finders, usually called scan matching. While there are dozens of different algorithms that attempt to solve the same problem (too many to be all cited here!), it is not known what is the limit on the achievable accuracy; whether the available methods reach that limit; and whether their accuracy can be improved.

One way to study the achievable accuracy is by deriving statistical bounds such as the Cramér-Rao Bound (CRB). This paper focuses on finding such bounds for pose tracking, defined as the problem of estimating the sensor displacement given two readings of a relative sensor. While the case of range-finders was the initial motivation, the theoretical results hold for other relative sensors, such as cameras — a precise definition of "relative sensor" is given later.

The CRB is computed from the Fisher Information Matrix (FIM). In the case of localization on a known map, there is a prescribed procedure that allows to compute the FIM from the sensor model. For example, in the case of range-finders [1], the FIM so obtained can be used to derive both quantitative results, such as the CRB, and qualitative results, such as the characterization of under-constrained situations. In the case of pose tracking, it is not that easy to derive the FIM, because the environment in unknown; in particular, the achievable accuracy for recovering the displacement depends on the prior for the map. Such prior can be easily defined only for "finite-dimensional" environments (polygons or landmarks) but not for the infinite-dimensional, not necessarily structured environments which are of more interest. In this respect, pose tracking presents the same difficulties as the full SLAM problem.

This paper shows that an approximation to the FIM for pose tracking can be found independently of the prior of the map. In particular, two results are presented. The first result (Theorem 1) concerns arbitrary displacements: an optimistic estimate to the FIM for pose tracking can be expressed through the FIM for localization; this is significant because the analysis of an infinite-dimensional problem (pose tracking) is reduced to the analysis of a finite-dimensional problem (localization). The second result (Theorem 2) strengthens the first in the limit for infinitesimal displacements. These results allow for the first time to compare the accuracy of a pose tracking algorithm to a well-defined baseline. Experiments are shown for the case of scan matching using the Iterative Closest/Corresponding Point algorithm (ICP).

The author is with the Control & Dynamical Systems department, Division of Engineering and Applied Science, California Institute of Technology. Address: MC 107-81, 1200 E. California Blvd., 91125, Pasadena, CA (USA). E-mail: andrea@cds.caltech.edu

II. Statistical background

The Cramér-Rao bound (CRB) is the simplest bound for the performance of estimators [2].

Proposition 1: (Cramér-Rao bound) Let $\boldsymbol{y}$ be the available observations, and $\boldsymbol{x} \in \mathbb{R}^n$ be an unknown fixed parameter to be estimated. Define the $n \times n$ symmetric positive semidefinite Fisher Information Matrix as

$$\boldsymbol{\mathcal{I}}[\boldsymbol{x}] = \mathrm{E}\left\{ \frac{\partial \log p(\boldsymbol{y}, \boldsymbol{x})}{\partial \boldsymbol{x}}^{\mathrm{T}} \frac{\partial \log p(\boldsymbol{y}, \boldsymbol{x})}{\partial \boldsymbol{x}} \right\} \tag{1}$$

If the density $p(\boldsymbol{y}, \boldsymbol{x})$ satisfies certain regularity conditions, and $\boldsymbol{\mathcal{I}}[\boldsymbol{x}]$ exists and is non-singular, then for any estimator $\hat{\boldsymbol{x}}$,

$$\mathrm{cov}[\hat{\boldsymbol{x}}] \geq \left[\boldsymbol{I} + \frac{\partial \boldsymbol{b}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right] (\boldsymbol{\mathcal{I}}[\boldsymbol{x}])^{-1} \left[\boldsymbol{I} + \frac{\partial \boldsymbol{b}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right]^{\mathrm{T}} \tag{2}$$

where $\boldsymbol{b}(\boldsymbol{x})$ is the bias of the estimator. In particular, for unbiased estimators, $\mathrm{cov}[\hat{\boldsymbol{x}}] \geq (\boldsymbol{\mathcal{I}}[\boldsymbol{x}])^{-1}$. An estimator which reaches the CRB is said to be *efficient*.

See, e.g., [2] for details about the regularity conditions, for the corner-case of a singular FIM, and other extensions such as the Bayesian CRB. Note that the FIM depends on the particular "true" value of $\boldsymbol{x}$: a more correct expression for the FIM is $\boldsymbol{\mathcal{I}}[\boldsymbol{x}](\boldsymbol{x}_*)$, in the sense that it is the information for the variable $\boldsymbol{x}$, but it also depends on the particular point $\boldsymbol{x}_*$ being considered — this notation will be used later when it is needed. Note that, in general, it is not possible to precisely evaluate (2) because of the bias function $\boldsymbol{b}(\boldsymbol{x})$ which is specific to the particular estimator. Nevertheless, the FIM $\boldsymbol{\mathcal{I}}[\boldsymbol{x}]$ depends only on the problem, and can be studied independently.

In general, the CRB is tight only in special cases, for example if the model is linear with Gaussian noise ($\boldsymbol{y} = \mathbf{L}\boldsymbol{x} + \boldsymbol{\epsilon}$).

In the case of a nonlinear Gaussian measurement model: $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{\Sigma})$, the FIM depends on the first derivative of $\boldsymbol{f}$:

$$\mathcal{I}[\boldsymbol{x}] = (\partial \boldsymbol{f}/\partial \boldsymbol{x})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\partial \boldsymbol{f}/\partial \boldsymbol{x}) \quad (3)$$

There exist tighter bounds that consider higher-order derivatives of $\boldsymbol{f}$, but the computations quickly become convoluted [2]. The CRB can be considered approximately tight at high signal-to-noise-ratios, but this must be usually verified experimentally.

A. Properties of the FIM

The FIM can be partitioned according to logical groups of variables. For example, in SLAM, $\boldsymbol{x}$ comprises the pose(s) $\boldsymbol{q}$ and the world $\boldsymbol{w}$. The FIM can be partitioned accordingly:

$$\mathcal{I}[\boldsymbol{q}, \boldsymbol{w}] = \begin{bmatrix} \mathcal{I}[\boldsymbol{q}] & \mathcal{I}[\boldsymbol{q}|\boldsymbol{w}] \\ * & \mathcal{I}[\boldsymbol{w}] \end{bmatrix} \quad (4)$$

The blocks on the diagonal represent the information for a variable, *assuming that all the other variables are known.* For example, the block $\mathcal{I}[\boldsymbol{q}]$ represents the information in the data about the pose in the case of localization; that is, assuming $\boldsymbol{w}$ known. Analogously, the block $\mathcal{I}[\boldsymbol{w}]$ represents the available information for mapping (defined as the problem of estimating the map given known poses). The off-diagonal terms represent the mutual information between variables; in this case the upper-right block $\mathcal{I}[\boldsymbol{q}|\boldsymbol{w}]$ represents the "simultaneous" in Simultaneous Localization and Mapping.

To compute the CRB for a subset of the state, say for $\boldsymbol{q}$, one must invert the full matrix $\mathcal{I}[\boldsymbol{q}, \boldsymbol{w}]$ and consider the upper left block of the inverse $\left[\mathcal{I}[\boldsymbol{q}, \boldsymbol{w}]^{-1}\right]_{11}$. This is not the same as just inverting the block $\mathcal{I}[\boldsymbol{q}]$, because, in general, $\mathcal{I}[\boldsymbol{q}]^{-1} \neq \left[\mathcal{I}[\boldsymbol{q}, \boldsymbol{w}]^{-1}\right]_{11}$. The former is the CRB on $\boldsymbol{q}$ assuming $\boldsymbol{w}$ known, while the latter is the CRB for $\boldsymbol{q}$ assuming $\boldsymbol{w}$ *not* known. The matrix $\left[\mathcal{I}[\boldsymbol{q}, \boldsymbol{w}]^{-1}\right]_{11}$ coincides with the inverse of the Schur complement of $\mathcal{I}[\boldsymbol{q}]$ — this follows from one of the matrix inversion lemmas. This motivates the following is a non-standard but useful notation.

Definition 1: For any two groups of variables $\boldsymbol{q}$ and $\boldsymbol{w}$, define $\mathcal{I}[\boldsymbol{q}/\boldsymbol{w}]$ as that Schur complement[1] of $\mathcal{I}[\boldsymbol{q}]$ with respect to $\mathcal{I}[\boldsymbol{w}]$:

$$\mathcal{I}[\boldsymbol{q}/\boldsymbol{w}] \triangleq \mathcal{I}[\boldsymbol{q}] - \mathcal{I}[\boldsymbol{q}|\boldsymbol{w}] \left(\mathcal{I}[\boldsymbol{w}]\right)^{-1} \mathcal{I}[\boldsymbol{q}|\boldsymbol{w}]^{\mathrm{T}}$$

The intuitive interpretation of $\mathcal{I}[\boldsymbol{q}/\boldsymbol{w}]$ is as the information in the data about $\boldsymbol{q}$, after considering that the nuisance parameter $\boldsymbol{w}$ is unknown. One has that $\mathcal{I}[\boldsymbol{q}/\boldsymbol{w}] = \mathcal{I}[\boldsymbol{q}]$ only if 1) $\boldsymbol{q}$ and $\boldsymbol{w}$ were statistically "orthogonal" ($\mathcal{I}[\boldsymbol{q}|\boldsymbol{w}] = 0$); or 2) $\boldsymbol{w}$ is precisely known ($\mathcal{I}[\boldsymbol{w}] \to \infty$).

The following proposition shows how FIMs behave under linear transformations, and is used later in one of the proofs.

Proposition 2: Let $\boldsymbol{z} = \mathbf{A}\boldsymbol{x}$. Then, if $\mathcal{I}[\boldsymbol{x}]$ is non-singular,

$$\mathcal{I}[\boldsymbol{z}] \leq \left(\mathbf{A} \left(\mathcal{I}[\boldsymbol{x}]\right)^{-1} \mathbf{A}^{\mathrm{T}}\right)^{-1},$$

with equality if $\boldsymbol{x}$ is Gaussian or $\mathbf{A}$ is square and non-singular [4].

[1]Note that a similar formula is used with *covariance* matrices for *conditioning*, which expresses the opposite idea ("what would be the covariance of $\boldsymbol{q}$ if $\boldsymbol{w}$ is known?") to what $\mathcal{I}[\boldsymbol{q}/\boldsymbol{w}]$ is ("what is the information on $\boldsymbol{q}$ if $\boldsymbol{w}$ is unknown?"). See for example [3, section 10].

III. Problem setting

Let $\boldsymbol{q}_k \in \mathrm{SE}(2)$ be the pose of the robot with respect to a fixed world frame. Between times k and $k+1$, the robot displacement is $\boldsymbol{\delta}_k \in \mathrm{SE}(2)$. Thus $\boldsymbol{q}_k = \boldsymbol{q}_0 \oplus \boldsymbol{\delta}_0 \oplus \cdots \oplus \boldsymbol{\delta}_{k-1}$, where $\oplus$ is the pose composition operator[2]. The robot observes $\boldsymbol{\delta}_k$ through proprioceptive measurements $\boldsymbol{u}_k$ (such as odometry/IMU). Moreover, the robot obtains exteroceptive measurements $\boldsymbol{y}$. These are usually described by the distribution $p(\boldsymbol{y}|\boldsymbol{w}, \boldsymbol{q})$, where $\boldsymbol{w} \in \mathcal{W}$ is the "world" or "map", but in this paper it is convenient to use an explicit functional expression for the measurements. Assuming additive Gaussian noise, the sensor model is

$$\boldsymbol{y}_k = \boldsymbol{r}(\boldsymbol{q}_k, \boldsymbol{w}) + \boldsymbol{\epsilon}_k \quad (5)$$

The function $\boldsymbol{r}$ is assumed to model a "relative" sensor; the exact meaning will be formalized later (Definition 2).

This paper uses the following nomenclature: *localization* is the problem of estimating $\boldsymbol{q}$ (either $\boldsymbol{q}_k$ or the full trajectory $\boldsymbol{q}_{0:k}$) given $\boldsymbol{y}_{0:k}$, $\boldsymbol{u}_{0:k}$ and a *known* $\boldsymbol{w}$; *mapping* is the problem of estimating $\boldsymbol{w}$ given $\boldsymbol{q}_{0:k}$ and $\boldsymbol{y}_{0:k}$; *pose tracking* is the problem of estimating $\boldsymbol{\delta}_{0:k}$ given $\boldsymbol{y}_{0:k}$ and $\boldsymbol{u}_{0:k}$, with *unknown* $\boldsymbol{w}$; *SLAM* is the problem of jointly estimating $\boldsymbol{w}$ and $\boldsymbol{q}$ given $\boldsymbol{y}_{0:k}$ and $\boldsymbol{u}_{0:k}$. This, of course, is not a clear-cut distinction, especially because there are algorithms doing both pose-tracking an localization, and SLAM algorithms may contain submodules solving one of the other simpler problems. Still, it is a useful conceptual distinction.

Localization is the only problem which is finite-dimensional as the robot poses are the only unknowns. In this case, the FIM is a well-defined finite-dimensional matrix, which can be computed from the model (5). For example, the paper [1] studied the FIM for the case of 2D localization with range-finders. The observations $\boldsymbol{y}$ are a set of ranges $\{\tilde{\rho}_i\}$, each corresponding to a direction φ_i. Letting $\boldsymbol{q} = (x, y, \theta)$, the observation model (5) is particularized as

$$\tilde{\rho}_i = r(x, y, \theta + \varphi_i) + \epsilon_i \qquad i = 1 \ldots n \quad (6)$$

where ϵ_i is a Gaussian noise with variance σ_i^2, and r is the "ray-tracing" function: $r(x, y, \psi)$ is the distance from point (x, y) to the first obstacle in direction ψ. The FIM depends on the geometry of the environment: in the following, α_i is the surface orientation at the point intercepted by the i-th ray, $\boldsymbol{v}(\alpha_i) = [\cos(\alpha_i) \ \sin(\alpha_i)]$ is the corresponding versor, and β_i is the angle of incidence, defined as $\beta_i \triangleq \alpha_i - (\theta + \varphi_i)$. The FIM is

$$\mathcal{I}[\boldsymbol{q}] = \sum_{i=1}^{n} \frac{1}{\sigma_i^2 \cos^2 \beta_i} \begin{bmatrix} \boldsymbol{v}(\alpha_i)\boldsymbol{v}(\alpha_i)^{\mathrm{T}} & r_i \sin\beta_i \boldsymbol{v}(\alpha_i) \\ * & r_i^2 \sin^2 \beta_i \end{bmatrix} \quad (7)$$

[2]This paper uses the standard notation in the SLAM literature. The symbol "$\oplus$" denotes pose composition, which is the group operation on SE(2). Please note the order of the arguments to $\oplus$: $(x_1, y_1, \theta_1) \oplus (x_2, y_2, \theta_2) = (x_1 + x_2\cos\theta_1 - y_2 \sin\theta_1, y_1 + x_2 \sin\theta_1 + y_2\cos\theta_1, \theta_1 + \theta_2)$. The symbol "$\ominus$" denotes the group inverse: $\ominus(x, y, \theta) = (-x\cos\theta, -y\sin\theta, x\sin\theta - y\cos\theta, -\theta)$. The following basic properties of $\oplus$ and $\ominus$ are used in the proofs: 1) $(\boldsymbol{a} \oplus \boldsymbol{b}) \oplus \boldsymbol{c} = \boldsymbol{a} \oplus (\boldsymbol{b} \oplus \boldsymbol{c})$; 2) $\ominus(\boldsymbol{a} \oplus \boldsymbol{b}) = (\ominus \boldsymbol{b}) \oplus (\ominus \boldsymbol{a})$; 3) $\ominus \ominus \boldsymbol{a} = \boldsymbol{a}$; 4) $\boldsymbol{a} \oplus (\ominus \boldsymbol{a}) = \boldsymbol{0}$.

The greatest difficulty for extending such analysis to pose tracking, mapping and SLAM is defining a proper prior for the world. A simple extreme example can show the importance of the prior in defining accuracy bounds. Suppose the robot lives in a world which has two admissible shapes: triangle ($\triangle$) and circle ($\bigcirc$); i.e., let $\mathcal{W} = \{\triangle, \bigcirc\}$. A sensor precise enough allows to perfectly distinguish $\triangle$ from $\bigcirc$. In this example, even if the map is unknown in advance, the achievable accuracy for *mapping* is infinite, because $\boldsymbol{w}$ can be recovered exactly also in the presence of (moderate) sensor noise. To appreciate the importance of the prior, consider a case in which the set $\mathcal{W}$ is a superset of $\{\triangle, \bigcirc\}$ (for example, let $\mathcal{W}$ be the space of all closed curves). The achievable accuracy for mapping is still infinite if the prior is non-zero only for those two elements. The prior for the map is important also for pose tracking, because the achievable accuracy for recovering the displacement from two sensor readings ultimately depends on the accuracy with which one can estimate the map from those readings.

The immediate technical difficulty for computing the CRB for problems involving an unknown map is that one should compute the FIM $\boldsymbol{\mathcal{I}}[\boldsymbol{q}, \boldsymbol{w}]$ as in (4). The block $\boldsymbol{\mathcal{I}}[\boldsymbol{q}]$ depends on the sensor model only, and is equal to (7) in the case of range-finders (one block like (7) for each pose), plus additional terms accounting for the odometry information. The block $\boldsymbol{\mathcal{I}}[\boldsymbol{w}]$ is the problematic one: except in the trivial cases, such as polygonal environments, it is infinite-dimensional, therefore it cannot be computed even in principle. It is also unclear how expressive can the FIM be for defining the prior for the world; given that it considers essentially the second moments, it should be as expressive as Gaussian processes.

In the case of pose tracking, the world is a *nuisance parameter*: it is unknown but not required as part of the estimation. This paper shows how this can be exploited. In particular, the next section shows two results that allow to obtain bounds on $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$ (the information for the robot displacement $\boldsymbol{\delta}$, given that the world $\boldsymbol{w}$ is unknown) as a function of $\boldsymbol{\mathcal{I}}[\boldsymbol{q}]$, which is a well-defined object, without knowing $\boldsymbol{\mathcal{I}}[\boldsymbol{w}]$ or $\boldsymbol{\mathcal{I}}[\boldsymbol{q}|\boldsymbol{w}]$. The rest of this section discusses some general properties on which the proofs are based.

A. Geometric properties of $\mathcal{W}$

In the robotics literature, the intrinsic algebraic and geometric properties of the world $\mathcal{W}$ are seldom discussed. Consider the set of worlds $\mathcal{W}$: for any world $\boldsymbol{w}$, there will be many others that have the same "shape" but differ for position and orientation. Therefore the space $\mathcal{W}$ must have a factorization into two components: pose and shape.

Axiom 1: The set $\mathcal{W}$ is isomorphic to $\mathrm{SE}(2) \times \mathbb{S}$, where $\mathbb{S}$ is the set of "shapes".

The idea of investigating the intrinsic properties of shape spaces, regardless of a particular representation, has been explored in other fields, both for distribution of points [5] (similar to a landmark-based formalization) and for curves [6], [7] (similar to "dense" mapping). The factorization $\mathrm{SE}(2) \times \mathbb{S}$ means that one considers *shape* all the properties that are invariant to roto-translations. In robotic mapping, scaling is important — a door's frame is about 1m, etc. — but in fields such as vision, scale can be part of the factorization, by setting $\mathcal{W} \simeq \mathrm{SE}(2) \times \mathbb{R} \times \mathbb{S}$.

This factorization idea can be used to formally define a relative sensor. An element $\boldsymbol{w} \in \mathcal{W}$ can be written as $\langle \boldsymbol{p}, \boldsymbol{s} \rangle$, with $\boldsymbol{p} \in \mathrm{SE}(2)$ and $\boldsymbol{s} \in \mathbb{S}$. Let $\overline{\boldsymbol{r}}$ be the function $\boldsymbol{r}$ after the change of coordinates $\boldsymbol{w} \to \langle \boldsymbol{p}, \boldsymbol{s} \rangle$; i.e., let $\boldsymbol{r}(\boldsymbol{q}, \boldsymbol{w}) = \overline{\boldsymbol{r}}(\boldsymbol{q}, \langle \boldsymbol{p}, \boldsymbol{s} \rangle)$; this can be thought of a "canonical" form for $\boldsymbol{r}$. It is possible to think of $\boldsymbol{q}$ and $\boldsymbol{p}$ as the poses of the robot and the shape with respect to some fixed frame. If the sensor is relative, then the output does not change if $\boldsymbol{q}$ and $\boldsymbol{p}$ are jointly roto-translated. This is captured by the following definition.

Definition 2: The function $\overline{\boldsymbol{r}}$ corresponds to a relative sensor if $\overline{\boldsymbol{r}}(\boldsymbol{q}, \langle \boldsymbol{p}, \boldsymbol{s} \rangle) = \overline{\boldsymbol{r}}(\boldsymbol{a} \oplus \boldsymbol{q}, \langle \boldsymbol{a} \oplus \boldsymbol{p}, \boldsymbol{s} \rangle)$ for all displacements $\boldsymbol{a}$.

By letting $\boldsymbol{a} = \ominus \boldsymbol{p}$, it is clear that this is equivalent to saying that $\overline{\boldsymbol{r}}$ depends only on the relative pose $\ominus \boldsymbol{p} \oplus \boldsymbol{q}$.

IV. On achievable accuracy for pose tracking

In this paper, pose tracking is defined as the problem of recovering the displacement $\boldsymbol{\delta}$ given two relative sensor readings taken at poses $\boldsymbol{q}_*$ and $\boldsymbol{q}_* \oplus \boldsymbol{\delta}$. The mathematical model is the following:

$$\begin{aligned} \boldsymbol{y}^a &= \boldsymbol{r}(\boldsymbol{q}_*, \boldsymbol{w}) + \text{noise} \\ \boldsymbol{y}^b &= \boldsymbol{r}(\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \boldsymbol{w}) + \text{noise} \end{aligned} \tag{8}$$

Given the two readings $\boldsymbol{y}^a$ and $\boldsymbol{y}^b$, how precise can the estimate of $\boldsymbol{\delta}$ be? The CRB can be computed from $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$, but as discussed, an exact expression involves $\boldsymbol{\mathcal{I}}[\boldsymbol{w}]$ and $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}|\boldsymbol{w}]$, which depend on the prior and the particular representation of $\boldsymbol{w}$. The results in this section provide bounds on $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$ which only depend on the FIM for localization $\boldsymbol{\mathcal{I}}[\boldsymbol{q}]$, hence reducing the analysis of pose tracking to the analysis of localization on a known map. In the following, the information given by the odometry is neglected for simplicity; its only effect is to add a constant term to $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$.

A. A result for arbitrary $\boldsymbol{\delta}$

Theorem 1: Given the system (10), without additional information on the prior for the world, the particular sensor model or the size of $\boldsymbol{\delta}$, the best bounds on the FIM for $\boldsymbol{\delta}$ are:

$$\mathbf{0} \leq \boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}] \leq \left(\boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_*, \boldsymbol{w})^{-1} + \boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \boldsymbol{w})^{-1} \right)^{-1} \tag{9}$$

Here the notation must be slightly verbose to be precise: the result concerns the information on $\boldsymbol{\delta}$ when $\boldsymbol{w}$ is unknown ($\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$, according to Definition 1) and it is expressed through the FIM for localization $\boldsymbol{\mathcal{I}}[\boldsymbol{q}]$. In turn, $\boldsymbol{\mathcal{I}}[\boldsymbol{q}]$ is evaluated at the particular points $\boldsymbol{q}_*$ and $\boldsymbol{q}_* \oplus \boldsymbol{\delta}$ and for the particular world $\boldsymbol{w}$. As one would expect, the result is symmetric in the two poses.

Intuitive proof: It is clear that if $\boldsymbol{\delta}$ is large enough, the two sensor readings will not sample the same parts of the environment, and it would be impossible to estimate $\boldsymbol{\delta}$. Therefore the best lower bound on $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$ must be $\mathbf{0}$.

To prove the upper bound, consider that the accuracy of pose tracking depends on the prior for the world. To obtain a lower bound on the error (i.e., an upper bound on the information), it is possible to assume to be in the most favorable situation: that the world $\boldsymbol{w}$ is known. If the world is known, then the model (10) looks like two independent localization problems, with unknowns $\boldsymbol{q}_*$ and $\boldsymbol{q}_* \oplus \boldsymbol{\delta}$. To obtain $\boldsymbol{\delta}$, one can solve the two localization problems for the two poses and then compute the difference. The accuracy of each pose estimation problem is bounded by the CRB for localization, which is $(\boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_*, \boldsymbol{w}))^{-1}$ and $(\boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \boldsymbol{w})|))^{-1}$ respectively. When computing the difference, one must add the minimum covariances, thereby obtaining (9).

Moreover, there are situations in which this upper bound is achieved. Consider the case of uni-dimensional scan-matching, where the robot pose is $\boldsymbol{q} \in \mathbb{R}$, and the range-finder returns the distance to a wall at pose $\boldsymbol{w} \in \mathbb{R}$ with Gaussian noise $\boldsymbol{\epsilon}$ with covariance σ^2. In this case, the sensor model is $\boldsymbol{y} = \boldsymbol{r}(\boldsymbol{q}, \boldsymbol{w}) + \boldsymbol{\epsilon} = \boldsymbol{w} - \boldsymbol{q} + \boldsymbol{\epsilon}$. Given the two measurements $\boldsymbol{y}^a = \boldsymbol{w} - \boldsymbol{q}_* + \boldsymbol{\epsilon}^a$ and $\boldsymbol{y}^b = \boldsymbol{w} - (\boldsymbol{q}_* + \boldsymbol{\delta}) + \boldsymbol{\epsilon}^a$, it is easy to derive that $\boldsymbol{\delta}$ can be estimated with accuracy $2\sigma^2$, which corresponds to the upper bound for the information given by the theorem.

Proof: The mathematical argument is more refined than the intuitive proof, as it uses the world factorization (Axiom 1) into a pose-shape pair, and assumes only the world shape to be known. Following Axiom 1, the world $\boldsymbol{w}$ can be factored as the pose-shape pair $\langle \boldsymbol{p}, \boldsymbol{s} \rangle$, with $\boldsymbol{p} \in \mathrm{SE}(2)$ and $\boldsymbol{s} \in \mathbb{S}$:

$$\begin{aligned} \boldsymbol{y}^a &= \overline{\boldsymbol{r}}(\boldsymbol{q}_*, \langle \boldsymbol{p}, \boldsymbol{s} \rangle) + \text{noise} \\ \boldsymbol{y}^b &= \overline{\boldsymbol{r}}(\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \langle \boldsymbol{p}, \boldsymbol{s} \rangle) + \text{noise} \end{aligned} \tag{10}$$

Because $\langle \boldsymbol{p}, \boldsymbol{s} \rangle$ is just a reparametrization for $\boldsymbol{w}$, $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}] = \boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\langle \boldsymbol{p}, \boldsymbol{s} \rangle]$. To obtain an upper bound on $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\langle \boldsymbol{p}, \boldsymbol{s} \rangle]$, one can assume $\boldsymbol{s}$ to be known. Formally, the upper bound used is

$$\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\langle \boldsymbol{p}, \boldsymbol{s} \rangle] \leq \boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{p}], \tag{11}$$

which follows from Definition 1. The matrix $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{p}]$ is computed by assuming $\boldsymbol{s} = \overline{\boldsymbol{s}}$ known. By using Definition 2, the system (10) becomes:

$$\begin{aligned} \boldsymbol{y}^a &= \overline{\boldsymbol{r}}(\ominus \boldsymbol{p} \oplus \boldsymbol{q}_*, \langle \boldsymbol{0}, \overline{\boldsymbol{s}} \rangle) + \text{noise} \\ \boldsymbol{y}^b &= \overline{\boldsymbol{r}}(\ominus \boldsymbol{p} \oplus \boldsymbol{q}_* \oplus \boldsymbol{\delta}, \langle \boldsymbol{0}, \overline{\boldsymbol{s}} \rangle) + \text{noise} \end{aligned} \tag{12}$$

This looks like a localization problem: the only unknowns are the *two* poses $\boldsymbol{q}_1 \triangleq \ominus \boldsymbol{p} \oplus \boldsymbol{q}_*$ and $\boldsymbol{q}_2 \triangleq \ominus \boldsymbol{p} \oplus \boldsymbol{q}_* \oplus \boldsymbol{\delta}$, while the world $\langle 0, \overline{\boldsymbol{s}} \rangle$ is known. The variable of interest, $\boldsymbol{\delta}$, can be recovered from the two poses by the formula $\boldsymbol{\delta} = (\ominus \boldsymbol{q}_1) \oplus \boldsymbol{q}_2$. The model (10) and (12) describe exactly the same measurements: $\boldsymbol{y}^a = \overline{\boldsymbol{r}}(\ominus \boldsymbol{p} \oplus \boldsymbol{q}_* \oplus \boldsymbol{\delta}, \langle 0, \overline{\boldsymbol{s}} \rangle) = \boldsymbol{r}(\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \boldsymbol{w})$. Therefore, the FIM for $\boldsymbol{q}_1$ can be assumed to correspond to be the FIM for localization, at pose $\boldsymbol{q}_*$, on the known map $\boldsymbol{w}$: $\boldsymbol{\mathcal{I}}[\boldsymbol{q}_1] = \boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_*, \boldsymbol{w})$, and equivalently for $\boldsymbol{q}_2$, $\boldsymbol{\mathcal{I}}[\boldsymbol{q}_2] = \boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_* \oplus \boldsymbol{\delta}, \boldsymbol{w})$.

Up to now, the proof has treated the poses and $\boldsymbol{\delta}$ as intrinsic objects; the following discussion uses coordinates instead. For the purpose of writing the FIM for $\boldsymbol{\delta}$, the expression $\boldsymbol{\delta} = (\ominus \boldsymbol{q}_1) \oplus \boldsymbol{q}_2$ is inconvenient to handle because it is nonlinear in the components of $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ (nevertheless, a bound on the covariance of $\boldsymbol{\delta}$ could be derived by using this nonlinear relation [8, Lemma 1.3]). However, note that, in the discussion so far, the actual numerical value of $\boldsymbol{p}$ could have been changed arbitrarily, by properly roto-translating the shape $\overline{\boldsymbol{s}}$ (in fact, $\ominus \boldsymbol{p}$ is contained in both $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$). The particular choice $\boldsymbol{p} = \boldsymbol{q}_*$ — but it suffices for $\boldsymbol{p}$ and $\boldsymbol{q}_*$ to have the same rotational part — allows to write, in coordinates, $\boldsymbol{\delta} = \boldsymbol{q}_2 - \boldsymbol{q}_1$. This corresponds to choosing the coordinate frame aligned with the first pose. The FIM properties can then be used on this linear relationship. From Proposition 2, by letting $\boldsymbol{x} = [\boldsymbol{q}_1; \boldsymbol{q}_2]$ and $\mathbf{A} = [\mathbf{I} \mid -\mathbf{I}]$, it follows that the FIM for the difference of two variables $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ has the following upper bound:

$$\boldsymbol{\mathcal{I}}[\boldsymbol{q}_2 - \boldsymbol{q}_1] \leq \left(\boldsymbol{\mathcal{I}}[\boldsymbol{q}_2]^{-1} + \boldsymbol{\mathcal{I}}[\boldsymbol{q}_1]^{-1} \right)^{-1} \tag{13}$$

The thesis follows by substituting the expressions found above for $\boldsymbol{\mathcal{I}}[\boldsymbol{q}_1]^{-1}$ and $\boldsymbol{\mathcal{I}}[\boldsymbol{q}_2]^{-1}$, and then considering (11).

The bound (9) is tight if (11) and (13) hold with equality, and this happens in some special cases. The bound (11) holds with equality if 1) the world shape is known; or 2) it can be disambiguated perfectly from the sensor data. This depends ultimately on the prior for the world, and is valid, for example, in the case of the toy world $\mathcal{W} = \mathrm{SE}(2) \times \{\triangle, \bigcirc\}$ discussed in the previous section. The bound (13) holds with equality if the estimates of $\boldsymbol{q}_2$ and $\boldsymbol{q}_1$ are Gaussian. ■

B. *A stronger result for* $\boldsymbol{\delta} \to \mathbf{0}$

As a corollary to Theorem 1, note that for $\boldsymbol{\delta} \to \mathbf{0}$,

$$\lim_{\boldsymbol{\delta} \to 0} \boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}](\boldsymbol{q}_*, \boldsymbol{w}) \leq \frac{1}{2} \boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_*, \boldsymbol{w}) \tag{14}$$

A priori, this bound might be arbitrarily smaller than the true value of $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$, as it was derived with the very strong assumption of knowing the map shape; however, the following theorem can be proved:

Theorem 2: In the limit $\boldsymbol{\delta} \to \mathbf{0}$, the accuracy of pose tracking is exactly half that of localization, regardless of prior or representation of $\boldsymbol{w}$:

$$\lim_{\boldsymbol{\delta} \to 0} \boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}](\boldsymbol{q}_*, \boldsymbol{w}) = \frac{1}{2} \boldsymbol{\mathcal{I}}[\boldsymbol{q}](\boldsymbol{q}_*, \boldsymbol{w}) \tag{15}$$

To change a "$\leq$" in (14) to a "$=$" in (15), it takes a long and tedious proof, which is relegated in the Appendix. The proof relies on the fact that for $\boldsymbol{\delta} \to \mathbf{0}$ the two scans tend to coincide, and the unobservable part of the world is unimportant.

V. Simulations

This section presents some experiments corroborating the theoretical results, for the particular case of pose tracking using range-finders, using the Iterated Closest/Corresponding Point (ICP) algorithm. The reason for using simulated data is that these experiments require a very precise ground-truth. A single scan matching operation has sub-millimeter accuracy using the common range-finders, such as the latest Hokuyo and Sick models, and to observe the distribution of errors reliably, an even more accurate ground-truth would be necessary.

The following results have been found to be consistent for different environments shapes — of course, the actual numerical values and the shape of the distributions change; but the results are qualitatively consistent — therefore, in the interest of experimental repeatability, a simple square map has been used (Fig. 1). Some of the results depend on the density of the scans, so a "dense" and a "sparse" sensor are simulated. The first simulated sensor has 360 rays over a field of view of 180deg. The second sensor simulated has only 60 rays over a field of view of 180deg. The sensor noise is $\sigma = 1\mathrm{cm}$. The odometry error is assumed to be Gaussian, with uncorrelated errors on x, y, θ, and with standard deviations $\sigma[x] = \sigma[y] = 5\mathrm{cm}$, $\sigma[\theta] = 5\mathrm{deg}$. In these simple cases, where ICP surely converges in the vicinity of the true solution, the influence of the odometry error on the result is negligible.

Three different $\boldsymbol{\delta}$ have been considered. The first is $\boldsymbol{\delta} = \mathbf{0}$, to verify Theorem 2. The second is a pure rotation $\boldsymbol{\delta} = (0, 0, \Delta/2)$, where Δ is the angular resolution of the sensor. This is a special $\boldsymbol{\delta}$ because the sensed surface points in the second scan are maximally far away from the points in the first scan. The third displacement is $\boldsymbol{\delta} = (1\mathrm{m}, 0, 45\mathrm{deg})$, which is an arbitrary "far" position: there is nothing special about this pose and the results are consistent with a different one.

For $\boldsymbol{\delta} = \mathbf{0}$, Fig. 2 shows the results for the sparse sensor. The error covariance of ICP is compared to $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]^{-1}$ computed according to Theorem 1. This corresponds to the CRB, when the bias of the algorithm is neglected in (2). The accuracy of ICP appears to be very close to the theoretical bound. Fig. 3 shows the result for the case of the dense sensor. The multi-modality is probably due to wrong correspondences, which happen more easily for a dense scan. Note the very interesting fact that the CRB predicts the covariance of each mode. This makes it clear that the analysis in this paper concerns only the error intrinsic in the problem, due to the stochasticity of the data, but it cannot predict the error due to the quirks of the particular algorithm.

For $\boldsymbol{\delta} \neq \mathbf{0}$, the upper bound to $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$ given by Theorem 1 is used for computing the CRB. For $\boldsymbol{\delta} = (0, 0, \Delta/2)$, the multi-modality disappears (Fig. 4-5). However, for the sparse sensor, a clear bias appears, which is in the same order of magnitude of the variance. The results in the case of a far displacement of $\boldsymbol{\delta} = (1\mathrm{m}, 0, 45\mathrm{deg})$ (Fig. 6 and 7) are similar, with strong bias for the sparse sensor, and a slightly larger error covariance than the CRB.

In the case $\boldsymbol{\delta} = \mathbf{0}$, there is an exact expression for $\boldsymbol{\mathcal{I}}[\boldsymbol{\delta}/\boldsymbol{w}]$ given by Theorem 2 and hence an exact expression for the CRB. Because an actual algorithm (ICP) has very close performance to the CRB, then we can conclude that 1) the CRB is approximately strict; and 2) the ICP is approximately efficient. The word *approximately*, in the sense of "for engineering purposes", is needed, because the CRB is exactly tight only in the linear/Gaussian case.

In the case $\boldsymbol{\delta} \neq \mathbf{0}$, there is no exact expression for the CRB, but only an estimate given by Theorem 1. In this case, in particular in the "far" case of Fig. 6-7, the performance of the actual algorithm is slightly larger than the CRB. This could be due to several reasons: 1) The estimate of the CRB given by Theorem 1 is too optimistic; 2) The estimate is correct, but the CRB itself is optimistic for the scan-matching problem; 3) The estimate is correct, the CRB is tight, but ICP is suboptimal; 4) The derivative of the bias cannot be neglected in the computation of (2). Assessing the relative importance of these reasons is part of future work.

VI. Conclusions and future work

Statistical bounds such as the CRB can be used to derive hard performance limits for estimation algorithms, to compare the performance of algorithms to a well-defined reference, and to check how close is the algorithm to optimality and whether it can be improved. Studying these bounds for pose tracking and SLAM might be hard, because the achievable accuracy depends on the prior for the map. Nevertheless, this paper showed non-trivial bounds for pose tracking through the FIM for localization, therefore reducing the analysis of an infinite-dimensional problem to a finite-dimensional one. The theoretical results hold for pose tracking with any relative sensor and are the best possible, without additional information on the prior for the world or a particular sensor.

The experimental verification considered the particular case of 2D scan matching with range-finders, and a particular algorithm (ICP). The bounds derived in a completely general setting are almost tight in this particular situation; this confirms the validity of the approach. For what concerns the optimality of ICP, a preliminary analysis comparing the accuracy to the CRB shows that the margin for improving the covariance is very small with respect to the other errors, such as the bias and the error due to false correspondences.

There are many improvements on the theory to be considered future work. For example, one might consider multiple steps for pose tracking. This is not as easy as just propagating and adding the CRB for a single matching [9], because the same scan $\boldsymbol{y}_k$ is used for estimating both $\boldsymbol{\delta}_{k-1}$ and $\boldsymbol{\delta}_k$; this interaction must be taken into account when propagating the covariance [10], and therefore also for deriving the achievable accuracy. The other line of work is trying to derive bounds for a particular relative sensor, such as a range-finder. For the more general problem of deriving Bayesian bounds for (dense) SLAM, the next step might be considering just *mapping* (estimating the map given known poses), and the best way to formally specify a prior for the world; a possibility is using Gaussian processes [11]. In this context, many ideas, originally used essentially for synthesis, can be useful for analysis; examples are the insights on the sparse structure of view-based SLAM [12] and the natural factorization induced by using relative *local* sensors [13].

References

[1] A. Censi, "On achievable accuracy for range-finder localization," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2007.

[2] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley-IEEE Press, 2007.

[3] U. Frese, "A discussion of simultaneous localization and mapping," *Autonomous Robots*, vol. 20, no. 1, pp. 25–42, 2006.

[4] C. Vignat and J.-F. Bercher, "On Fisher Information inequalities and score functions in non-invertible linear systems," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 4, no. 4, 2003.

[5] H. Le and D. G. Kendall, "The Riemannian structure of Euclidean shape spaces: A novel environment for statistics," *Annals of Statistics*, vol. 21, no. 3, pp. 1225–1271, 1993.

[6] P. W. Michor and D. Mumford, "Riemannian geometries on spaces of plane curves," *J. of the European Math. Soc.*, vol. 8, pp. 1–48, 2004.

[7] S. Joshi, D. Kaziska, A. Srivastava, and W. Mio, "Riemannian structures on shape spaces: A framework for statistical inferences," *Statistics and Analysis of Shapes*, pp. 313–333, 2006.
[8] V. Papathanasiou, "Some characteristic properties of the Fisher Information Matrix via Cacoullos-type inequalities," *Journal of Multivariate Analysis*, vol. 44, pp. 256–265, 1993.
[9] A. Mourikis. Personal communication, 2008.
[10] A. Mourikis and S. Roumeliotis, "On the treatment of relative-pose measurements for mobile robot localization," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, (Orlando, FL), 2006.
[11] C. Plagemann, K. Kersting, P. Pfaff, and W. Burgard, "Gaussian beam processes: A nonparametric bayesian measurement model for range finders," in *Robotics: Science and Systems (RSS)*, 2007.
[12] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based SLAM," *IEEE Transactions on Robotics*, vol. 22, pp. 1100–1114, Dec. 2006.
[13] J.-L. Blanco, J. González, and J.-A. Fernández-Madrigal, "Subjective local maps for hybrid metric-topological SLAM," *Robotics and Autonomous Systems*, 2008.

APPENDIX

For simplicity, in the following proofs we assume that sensor readings are uncorrelated, identically distributed and with normalized variance $\sigma = 1$: in other words, we set $\mathbf{\Sigma} = \mathbf{I}$ in (3).

Lemma 1: Assume a model of the kind $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w})$, and suppose there is a reparametrization $\boldsymbol{w} \mapsto \boldsymbol{m} = \varphi(\boldsymbol{w})$ with φ invertible on its range. Then $\mathcal{I}[\boldsymbol{x}/\boldsymbol{m}] = \mathcal{I}[\boldsymbol{x}/\boldsymbol{w}]$.

Proof: The new measurement equation is $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{m}) = \boldsymbol{f}(\boldsymbol{x}, \varphi^{-1}(\boldsymbol{m}))$. To evaluate $\mathcal{I}[\boldsymbol{x}/\boldsymbol{m}]$ one needs the terms $\mathcal{I}[\boldsymbol{x}|\boldsymbol{m}] = \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}^{\mathrm{T}} \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}}$ and $\mathcal{I}[\boldsymbol{m}] = \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}}^{\mathrm{T}} \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}}$.

Note that $\partial \boldsymbol{g}/\partial \boldsymbol{x} = \partial \boldsymbol{f}/\partial \boldsymbol{x}$, and because of the chain rule:

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}} \frac{\partial}{\partial \boldsymbol{m}} p^{-1}(\boldsymbol{m}) = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}} \left[\frac{\partial p(\boldsymbol{w})}{\partial \boldsymbol{w}} \right]^{-1}$$

Let $J = \dfrac{\partial \varphi(\boldsymbol{w})}{\partial \boldsymbol{w}}$, and compute

$$\mathcal{I}[\boldsymbol{m}] = \mathbf{J}^{-\mathrm{T}} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}}^{\mathrm{T}} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}} \mathbf{J}^{-1} = \mathbf{J}^{-\mathrm{T}} \mathcal{I}[\boldsymbol{w}] \mathbf{J}^{-1}$$

$$\mathcal{I}[\boldsymbol{x}|\boldsymbol{m}] = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}^{\mathrm{T}} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}} \mathbf{J}^{-1} = \mathcal{I}[\boldsymbol{x}|\boldsymbol{w}] \mathbf{J}^{-1}$$

Substitute these in $\mathcal{I}[\boldsymbol{x}/\boldsymbol{m}]$ to obtain

$$\begin{aligned} \mathcal{I}[\boldsymbol{x}/\boldsymbol{m}] &= \mathcal{I}[\boldsymbol{x}] - \\ & \left[\mathcal{I}[\boldsymbol{x}|\boldsymbol{w}]\mathbf{J}^{-1}\right] \left[\mathbf{J}^{-\mathrm{T}} \mathcal{I}[\boldsymbol{w}] \mathbf{J}^{-1}\right]^{-1} \left[\mathcal{I}[\boldsymbol{x}|\boldsymbol{w}]\mathbf{J}^{-1}\right]^{\mathrm{T}} \\ &= \mathcal{I}[\boldsymbol{x}] - \mathcal{I}[\boldsymbol{x}|\boldsymbol{w}] \left[\mathcal{I}[\boldsymbol{w}]\right]^{-1} \mathcal{I}[\boldsymbol{x}|\boldsymbol{w}]^{\mathrm{T}} = \mathcal{I}[\boldsymbol{x}/\boldsymbol{w}] \end{aligned}$$

■

We are now ready to prove Theorem 2.

Proof: Write the model in the observable form:

$$\begin{aligned} \boldsymbol{y}^a &= \boldsymbol{r}(\mathbf{0}, \boldsymbol{w}) + \text{noise} \\ \boldsymbol{y}^b &= \boldsymbol{r}(\boldsymbol{\delta}, \boldsymbol{w}) + \text{noise} \end{aligned}$$

We do a change of coordinates for the map: $\boldsymbol{w} \mapsto \langle \boldsymbol{w}^1, \boldsymbol{w}^2 \rangle$ where $\boldsymbol{w}^1 \triangleq \boldsymbol{r}(\mathbf{0}, \boldsymbol{w})$, and $\boldsymbol{w}^2$ arbitrary containing the missing information to make the transformation invertible. In a degenerate case, if the first scan observes all the world's parameters, $\boldsymbol{w}^2$ might be empty. Define also the following function $\boldsymbol{h}$:

$$\boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) \triangleq \boldsymbol{r}(\boldsymbol{\delta}, \boldsymbol{w}) - \boldsymbol{r}(\mathbf{0}, \boldsymbol{w})$$

After the transformation, the model looks like this:

$$\begin{aligned} \boldsymbol{y}^a &= \boldsymbol{w}^1 + \text{noise} \\ \boldsymbol{y}^b &= \boldsymbol{w}^1 + \boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) + \text{noise} \end{aligned}$$

Note that $\boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta})$ represents the change in the sensor output as $\boldsymbol{\delta} \neq 0$: therefore $\boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \mathbf{0}) = \mathbf{0}$.

The FIM for $\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}$ is as follows:

$$\mathcal{I}[\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}] = \begin{bmatrix} \mathcal{I}[\boldsymbol{w}^1] & \mathcal{I}[\boldsymbol{w}^1|\boldsymbol{w}^2] & \mathcal{I}[\boldsymbol{w}^1|\boldsymbol{\delta}] \\ * & \mathcal{I}[\boldsymbol{w}^2] & \mathcal{I}[\boldsymbol{w}^2|\boldsymbol{\delta}] \\ * & * & \mathcal{I}[\boldsymbol{\delta}] \end{bmatrix}$$

We are now going to compute all these terms.

- **Computation of $\mathcal{I}[\boldsymbol{w}^1]$.** The element i, j of the matrix $\mathcal{I}[\boldsymbol{w}^1]$ is given by:

$$[\mathcal{I}[\boldsymbol{w}^1]]_{ij} = \frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_j} + \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_j}$$

Because $\boldsymbol{y}^a = \boldsymbol{w}^1 +$noise, we have that $\frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_i} = \boldsymbol{v}_i$, where $\boldsymbol{v}_i$ is defined as the zero vector with a 1 at position i:

$$\boldsymbol{v}_i \triangleq [0 \quad \cdots 0 \quad \underset{\stackrel{\uparrow}{i}}{1} \quad 0 \cdots \quad 0]^T$$

Equivalently $\frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_i} = \boldsymbol{v}_j$. Their product is $\boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_j = \delta_{ij}$, where δ_{jk} is Kronecker's delta function. For the second term, remember that $\boldsymbol{y}^b = \boldsymbol{w}^1 + \boldsymbol{h}$ and so

$$\frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_i} = \boldsymbol{v}_i + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}$$

The product of the two terms is:

$$\begin{aligned} \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_j} &= \left[\boldsymbol{v}_i + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}\right]^{\mathrm{T}} \left[\boldsymbol{v}_j + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_j}\right] \\ &= \boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_j + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_j} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \boldsymbol{v}_j + \boldsymbol{v}_i^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_j} \\ &= \delta_{ij} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_j} + \frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{w}^1_i} + \frac{\partial \boldsymbol{h}_i}{\partial \boldsymbol{w}^1_j} \end{aligned}$$

Hence the generic element is

$$[\mathcal{I}[\boldsymbol{w}^1]]_{ij} = 2\delta_{ij} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_j} + \frac{\partial \boldsymbol{h}_j}{\partial \boldsymbol{w}^1_i} + \frac{\partial \boldsymbol{h}_i}{\partial \boldsymbol{w}^1_j}$$

and going back to the matrix expression one obtains:

$$\mathcal{I}[\boldsymbol{w}^1] = 2\mathbf{I} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}}$$

Note that this is a symmetric matrix.

- **Computation of $\mathcal{I}[\boldsymbol{w}^2]$, $\mathcal{I}[\boldsymbol{\delta}]$, $\mathcal{I}[\boldsymbol{w}^2|\boldsymbol{\delta}]$.** These terms are readily computed as:

$$\mathcal{I}[\boldsymbol{w}^2] = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2} \qquad \mathcal{I}[\boldsymbol{\delta}] = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}$$

$$\mathcal{I}[\boldsymbol{w}^2|\boldsymbol{\delta}] = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}$$

- **Computation of $\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{w}^2]$:** As before, write the expression for the generic element and do some vector algebra:

$$\begin{aligned}
[\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{w}^2]]_{ij} &= \frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^2_j} + \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{y}^b}{\partial \boldsymbol{w}^2_j} \\
&= \frac{\partial \boldsymbol{y}^a}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \mathbf{0} + \left[\boldsymbol{v}_i + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}\right]^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2_j} \\
&= \boldsymbol{v}_i^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2_j} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2_j} \\
&= \frac{\partial \boldsymbol{h}_i}{\partial \boldsymbol{w}^2_j} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1_i}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2_j}
\end{aligned}$$

The matrix form in this case is

$$\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{w}^2] = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2}$$

Note that off-diagonal blocks such as $\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{w}^2]$ are not necessarily symmetric (nor necessarily square blocks).

- **Computation of $\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{\delta}]$:** The same procedure can be repeated to obtain:

$$\mathcal{I}[\boldsymbol{w}^1|\boldsymbol{\delta}] = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}$$

At this point we have an expression for the FIM: $\mathcal{I}[\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}] =$

$$\begin{bmatrix}
2\mathbf{I} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2} & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}} + \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^1}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}} \\
* & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2} & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{w}^2}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}} \\
* & * & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}^{\mathrm{T}} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\delta}}
\end{bmatrix}$$

Where all of these are functions of $\boldsymbol{w}^1$, $\boldsymbol{w}^2$ and $\boldsymbol{\delta}$.

Recall that the function $\boldsymbol{h}$ has the following property:

$$\lim_{\boldsymbol{\delta} \to \mathbf{0}} \boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) = \mathbf{0}$$

From this it follows that

$$\begin{aligned}
\lim_{\boldsymbol{\delta} \to \mathbf{0}} \frac{\partial}{\partial \boldsymbol{w}^1} \boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) &= \mathbf{0} \\
\lim_{\boldsymbol{\delta} \to \mathbf{0}} \frac{\partial}{\partial \boldsymbol{w}^2} \boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) &= \mathbf{0}
\end{aligned}$$

Moreover, from the definition of $\boldsymbol{h}$ it follows that

$$\lim_{\boldsymbol{\delta} \to \mathbf{0}} \frac{\partial}{\partial \boldsymbol{\delta}} \boldsymbol{h}(\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}) = \frac{\partial}{\partial \boldsymbol{q}} \boldsymbol{r}(\boldsymbol{q}, \boldsymbol{w})$$

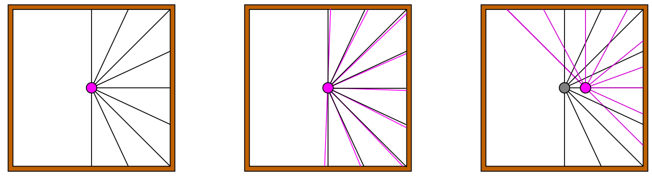
The limit of the FIM as $\boldsymbol{\delta} \to \mathbf{0}$ is then

$$\lim_{\boldsymbol{\delta} \to \mathbf{0}} \mathcal{I}[\boldsymbol{w}^1, \boldsymbol{w}^2, \boldsymbol{\delta}] = \begin{bmatrix} 2\mathbf{I} & \mathbf{0} & \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}} \\ * & \mathbf{0} & \mathbf{0} \\ * & * & \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}}^{\mathrm{T}} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}} \end{bmatrix}$$

Note the zeros for the entries relative to $\boldsymbol{w}^2$: as $\boldsymbol{\delta} \to 0$, there is no information about $\boldsymbol{w}^2$ in the data as $\boldsymbol{w}^1$ alone can be used to predict both scans. These elements can be dropped from the matrix, and the information $\mathcal{I}[\boldsymbol{\delta}/\boldsymbol{w}^1, \boldsymbol{w}^2]$ can be computed as:

$$\lim_{\boldsymbol{\delta} \to \mathbf{0}} \mathcal{I}[\boldsymbol{\delta}/\boldsymbol{w}^1, \boldsymbol{w}^2] = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}}^{\mathrm{T}} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}} - \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}}^{\mathrm{T}} [2\mathbf{I}]^{-1} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{q}} = \frac{1}{2} \mathcal{I}[\boldsymbol{q}]$$

Finally, by the previous lemma, $\lim_{\boldsymbol{\delta} \to \mathbf{0}} \mathcal{I}[\boldsymbol{\delta}/\boldsymbol{w}] = \lim_{\boldsymbol{\delta} \to \mathbf{0}} \mathcal{I}[\boldsymbol{\delta}/\boldsymbol{w}^1, \boldsymbol{w}^2]$ ■



(a) $\boldsymbol{\delta} = (0, 0, 0)$ (b) $\boldsymbol{\delta} = (0, 0, \Delta/2)$ (c) $\boldsymbol{\delta} = (1\mathrm{m}, 0, 45\mathrm{deg})$

Figure 1. The environment with the three displacements $\boldsymbol{\delta}$ considered. In the following figures, the red dots are the sample errors, the red ellipse is the sample covariance, the dashed blue ellipse is the CRB.

Figure 2. SPARSE SENSOR — $\boldsymbol{\delta} = (0, 0, 0)$

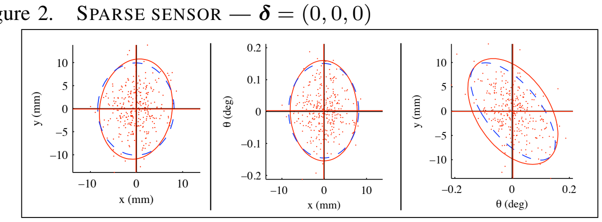


Figure 3. DENSE SENSOR — $\boldsymbol{\delta} = (0, 0, 0)$

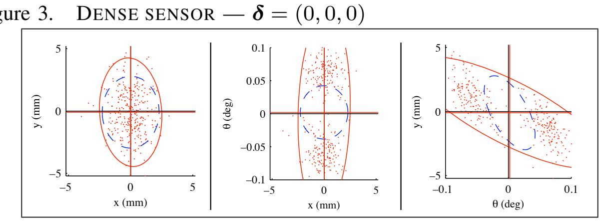


Figure 4. SPARSE SENSOR — $\boldsymbol{\delta} = (0, 0, \Delta/2)$

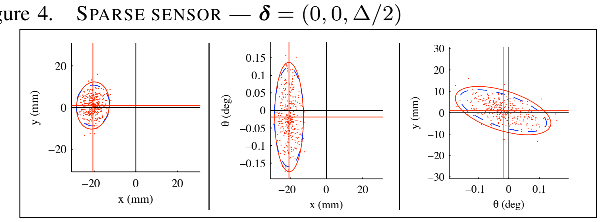


Figure 5. DENSE SENSOR — $\boldsymbol{\delta} = (0, 0, \Delta/2)$

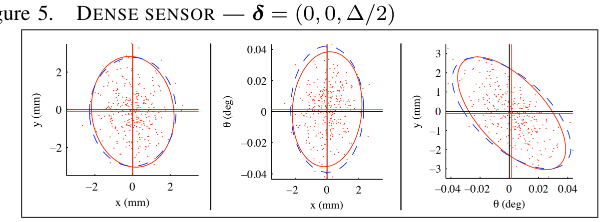


Figure 6. SPARSE SENSOR — $\boldsymbol{\delta} = (1\mathrm{m}, 0, 45\mathrm{deg})$

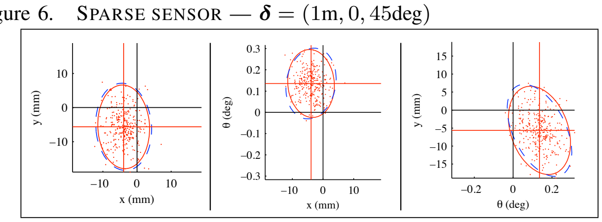


Figure 7. DENSE SENSOR — $\boldsymbol{\delta} = (1\mathrm{m}, 0, 45\mathrm{deg})$