Observability, Covariance and Uncertainty of ICP Scan Matching

Martin Barczyk, Silvère Bonnabel and François Goulette

Abstract—In the first part, we study the observability and covariance of estimates of an ICP-based scan matching algorithm in both 2D and 3D cases. Using numerical examples and mathematical proofs we demonstrate that the Hessian of the point-to-plane ICPs linearized cost function correctly models the observability of the algorithm, whereas the point-to-point variant does not. In the second part, assuming noisy data, we obtain a closed-form expression for the covariance of the algorithm, which is equal to the (scaled) inverse of the Hessian for white noise, but which is completely different otherwise. This difference is very important, since between random noise and resolution errors in the scanned data, only the latter is shown to be relevant to the computation. While the analysis is based on mathematical proofs, the end goal of the paper is to provide the practitioner with simple to compute closed-form expressions.

I. INTRODUCTION

This paper considers the observability and covariance of estimates obtained by applying the Iterative Closest-Point (ICP) algorithm [1], [2] to pairs of successive point clouds captured by a scanning sensor (for instance a 2D laser or a Kinect camera) moving through a structured environment. This so-called scan matching [3] is used to estimate the pose of the robot carrying the scanner in either 2D (heading $R \in SO(2)$ and in-plane position $p \in \mathbb{R}^2$) or 3D (attitude $R \in SO(3)$ and position $p \in \mathbb{R}^3$). The observability and covariance of the scan matching estimates enables diagnosing their accuracy and fusing them with data from other sensors, using for instance an Extended Kalman Filter (EKF) e.g. [4], [5] or more recent filters such as the Unscented Kalman Filter (UKF) or the particle filter.

Observability of point-to-plane ICP (aka ICP stability or sensitivity analysis) was previously studied in various works, see e.g. [6], [7]. The first contribution of the present paper is to provide a formal proof that in the case of point-to-plane ICP — but not point-to-point — the lack of observability in under-constrained environments (such as corridors) is correctly captured by the Hessian matrix of the linearized cost function. This deterministic portion is presented throughout Section II with the proof given in Section IIA.

Section III is concerned with the different (probabilistic) problem of computing the covariance (i.e. variability) of the ICP output estimate over repeated experiments assuming noisy data. This has been studied in [8], [9] using repeated numerically costly offline computations, and in [10], [11] which deals with the slightly different problem of matching a scan with a known surface map, and focuses on the covariance projected on the observable directions. The second contribution of this paper is to derive a (easily computed) closed form for the covariance matrix, which is equal to the scaled inverse of the Hessian derived in Section II in the case of white noise, but which is completely different for other types of noise. We then demonstrate by numerical example that white noise actually has negligible impact on the covariance as compared to scanner resolution errors.

II. ICP OBSERVABILITY AND STABILITY

A. The ICP Algorithm

The ICP algorithm [1], [2] is an iterative procedure for finding the optimum rigid-body transformation $(R, T) \in SO(3) \times \mathbb{R}^3$ between two sets of points in $\mathbb{R}^3$ (aka clouds) $\{a_i\}$ and $\{b_i\}$, which do not necessarily have equal number of entries. The ICP algorithm consists of the following sequence of steps:

1) Select $N$ points from $\{a_i\}$
2) Match these with points in $\{b_i\}$
3) Minimize a distance cost function using $(\delta R, \delta T)$
4) Transform $\{a_i\}$ and update $(R, T)$ by $(\delta R, \delta T)$
5) If $(R, T)$ converged then quit, else goto 1.

The details of steps 1 and 2 can vary quite a bit between implementations as seen in the excellent survey paper [12]. Step 3 uses one of two possible approaches: the point-to-point cost function [1],

$$f_{pp}(\delta R, \delta T) = \sum_{i=1}^{N} ||\delta R a_i + \delta T - b_i||^2 \quad (1)$$

or the point-to-plane cost function [2]

$$f_{pp}(\delta R, \delta T) = \sum_{i=1}^{N} [(\delta R a_i + \delta T - b_i) \cdot n_i]^2 \quad (2)$$

where $n_i$ is the surface unit normal vector in the second cloud based at $b_i$ whose actual value needs to be numerically computed by fitting a plane through neighboring points [13]. In step 4 the computed $(\delta R, \delta T)$ is applied to $\{a_i\}$ in order to move it closer (in the least-squares sense) to $\{b_i\}$ for the next iteration. Step 5 can employ a threshold value for convergence and/or loop the sequence over $N_{ICP}$ iterations.

B. Cost Function Linearization

In practice, the two clouds input to the ICP are always either pre-aligned using an estimation of the rigid-body transformation obtained from e.g. numerical integration of the vehicle dynamics using on-board inertial sensors or a computation of their optical flow, or the sampling rate of the cloud scanning sensor is sufficiently high for cloud pairs
to start close to each other. This is necessary because the ICP algorithm guarantees convergence only locally.

Based on the previous paragraph, we are justified in taking \( \delta R \in SO(3) \) in either (1) or (2) as close to identity. We can thus employ the linearizing approximation

\[
\delta R \approx I + S(x_R), \quad x_R \in \mathbb{R}^3, \quad \delta R \text{ close to } I
\]

where \( S \) is the skew-symmetric matrix

\[
S(u) = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}
\]

with \( S(u)v = u \times v \), the \( \mathbb{R}^3 \) cross-product. We also write

\[
\delta T = x_T, \quad x_T \in \mathbb{R}^3
\]

Employing (3) in the point-to-point ICP cost function (1) and defining \( x \triangleq [x_R \ x_T]^T \) we obtain

\[
f(x) = \sum_i \|a_i + S(x_R)a_i + x_T - b_i\|^2
\]

which can be rewritten as the sum-of-squares cost function

\[
f(x) = \sum_i \|y_i - H_i x\|^2
\]

where in point-to-point ICP we have

\[
y_i = a_i - b_i, \quad H_i = [S(a_i) - I]
\]

Meanwhile for the point-to-plane ICP cost function (2), employing (3) gives

\[
f(x) = \sum_i \|x_R \times a_i + x_T + a_i - b_i\| \cdot n_i \|^2
\]

Using the scalar triple product circular property \((a \times b) \cdot c = (b \times c) \cdot a\), this can also be rewritten in form (5) with

\[
y_i = n_i^T(a_i - b_i), \quad H_i = [-(a_i \times n_i)^T \ -n_i^T]
\]

1) Specialization to 2D: Consider the case of a planar rotation \( \delta R \in SO(2) \) and translation \( \delta T \in \mathbb{R}^2 \). In this case

\[
\delta R \approx I + S_2(x_R), \quad x_R \in \mathbb{R}, \quad \delta R \text{ close to } I_2
\]

where for \( u \in \mathbb{R}^2 \)

\[
S_2(u) = \begin{bmatrix} 0 & -u_2 \\ u_2 & 0 \end{bmatrix}
\]

For \( v = [v_1 \ v_2]^T \in \mathbb{R}^2 \) we have

\[
S_2(u)v = \begin{bmatrix} 0 & -u_2 \\ u_2 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_2 \\ v_1 \end{bmatrix} = v^\perp \triangleq v \perp u
\]

where \( v^\perp = [-v_2 \ v_1]^T \) is perpendicular to \( v \). We also write

\[
\delta T = x_T, \quad x_T \in \mathbb{R}^2
\]

In the planar case the point-to-point cost (1) becomes

\[
f(x) = \sum_i \|a_i + S_2(x_R)a_i + x_T - b_i\|^2
\]

equivalent to form (5) with

\[
y_i = a_i - b_i, \quad H_i = [-a_i^\perp \ -I_2]
\]

while the point-to-plane cost (2) becomes

\[
f(x) = \sum_i \left[(a_i + S_2(x_R)a_i + x_T - b_i) \cdot n_i\right]^2
\]

also equivalent to (5) with

\[
y_i = n_i^T(a_i - b_i), \quad H_i = [n_i^T a_i^\perp \ n_i^T]
\]

C. Cost Function Minimization

All linearized ICP cost functions in Section II-B took the form of a sum-of-squares cost function (5). In order to find the minimizing solution \( x = \hat{x} \) we solve \( \partial f(\hat{x})/\partial x = 0 \). Taking the gradient of (5) in denominator layout we obtain

\[
\frac{\partial f}{\partial x} = 2 \sum_i H_i^T H_i x - 2 \sum_i H_i^T y_i \triangleq 2Ax - 2b
\]

Remark A is by construction symmetric; in fact provided \( H_i \) isn’t skinny, \( A \) is positive semi-definite [14, p. 74]. The second derivative of \( f \), known as the Hessian of \( f \), is

\[
\frac{\partial^2 f}{\partial x^2} = 2A
\]

From (9) we find the critical point

\[
\frac{\partial f}{\partial x}(\hat{x}) = 0 \implies \hat{x} = A^{-1}b
\]

and since \( A \geq 0 \) this \( \hat{x} \) is indeed the minimum of (5). We can analyze the stability of the solution by taking a Taylor series expansion of the cost function (5) about \( x = \hat{x} \):

\[
f(x) = f(\hat{x}) + \frac{\partial f}{\partial x}(\hat{x})(x - \hat{x}) + (x - \hat{x})^T \frac{\partial^2 f}{\partial x^2}(\hat{x})(x - \hat{x})
\]

Remark the expansion is exact because (5) is quadratic in \( f \). Defining \( f(x) - f(\hat{x}) \triangleq \Delta f, x - \hat{x} \triangleq \Delta x \) and substituting the computed gradient and Hessian terms we obtain

\[
\Delta f = \Delta x^T (2A) \Delta x
\]

Physically, \( \Delta f \) measures the change in the cost function \( f(x) \) when \( x \) moves away from its minimum point \( \hat{x} \) by \( \Delta x \). Any values(s) of \( \Delta x \) for which \( \Delta f = 0 \) correspond to unobservable parameters. Detecting these is equivalent to solving an eigenvalue problem. The Hessian \( 2A \geq 0 \) is diagonalizable. Thus if \( \hat{v} \) is a unit-length eigenvector of \( 2A \) with associated eigenvalue \( \lambda \) and \( \Delta x = \hat{v} \) then

\[
\Delta f = \hat{v}^T \lambda \hat{v} = \lambda \|\hat{v}\|^2 = \lambda
\]

and thus any eigenvectors \( \hat{v} \) associated to a zero eigenvalue \( \lambda = 0 \) of \( 2A \) (or equivalently \( A \)) necessarily indicate an un-observable degree of freedom. Conversely, large eigenvalues correspond to well-observable directions.

D. Application to ICP

Based on Section II-C it appears the observability of the ICP algorithm can be analyzed using its linearized cost function. However there’s an important caveat which we demonstrate by numerical example. For simplicity consider the 2D case. For point-to-point ICP

\[
A_{pp} = \sum_i (H_i)^T H_i = \sum_i \left[ ||a_i^\perp||^2 \ a_i^\perp \ a_i^\perp \right]
\]
and for point-to-plane ICP

\[
A_{pp} = \sum_i (H_i)^T H_i = \sum_i \left[ (n_i^T a_i^+)^2 (n_i^T a_i^-) n_i^T \right]
\]

Consider a scanning sensor moving parallel to a featureless wall as illustrated in Figure 1 known to be unobservable to scan matching along the direction of travel and observable for the other two degrees of freedom c.f. [10, Fig. 3].

![Figure 1. Scan matching along featureless wall](image)

Assuming the wall is located 1 unit away from the scanner, which captures three points from the centre and edges of its 90° field of view, the first scan points are \(a_1 = [-1 \ 1]^T\), \(a_2 = [0 \ 1]^T\), \(a_3 = [1 \ 1]^T\) and the second scan points and associated surface unit normals are \(b_1 = [-1 \ 1]^T\), \(n_1 = [0 \ -1]^T\), \(b_2 = [0 \ 0]^T\), \(n_2 = [0 \ 0]^T\) and \(b_3 = [1 \ 1]^T\), \(n_3 = [0 \ -1]^T\). We directly compute

\[
A_{pp} = \begin{bmatrix} 5 & -3 & 2 \\ -3 & 3 & 0 \\ 2 & 0 & 3 \end{bmatrix}, \quad A_{pn} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 3 \end{bmatrix}
\]

The eigenvalue problem for \(A_{pp}\) gives

\[
\lambda_1 = 0.2583, \quad \hat{v}_1 = [0.6053 \ 0.6023 \ -0.4415]^T
\]

\[
\lambda_2 = 3, \quad \hat{v}_2 = [0 \ 0.5547 \ 0.8321]^T
\]

\[
\lambda_3 = 7.7417, \quad \hat{v}_3 = [0.7960 \ -0.5036 \ 0.3358]^T
\]

while for \(A_{pn}\) it gives

\[
\lambda_1 = 0, \quad \hat{v}_1 = [0 \ 1 \ 0]^T
\]

\[
\lambda_2 = 0.4384, \quad \hat{v}_2 = [0.7882 \ 0 \ -0.6154]^T
\]

\[
\lambda_3 = 4.5616, \quad \hat{v}_3 = [0.6154 \ 0.7882]^T
\]

This example shows that the linearized point-to-point ICP cost function analysis leads to incorrect observability results c.f. Section II-C. Meanwhile the point-to-plane version does provide correct results in this case: the \(\lambda = 0\) eigenvalue corresponds to an eigenvector along the unobservable direction (recall \(x = [x_R \ x_T]\)). It can be verified (c.f. Section III-A) that in the 3D case of the scanner facing a planar featureless wall, which should yield three zero eigenvalues corresponding to translations parallel to the wall and rotations perpendicular to it, the linearized point-to-point expressions (6) give incorrect results while the point-to-plane expressions (6) yield correct ones.

We now explain the reason for the discrepancy of the results. As described in Section III-A, the ICP algorithm reforms pairs of corresponding points at every iteration (step 2). This is why when facing a featureless plane wall, the algorithm will ignore any movements parallel to it: as the scanner moves, step 2 of the algorithm will re-match point pairs such that the value of the distance cost function in step 3 will not change. However, for all the linearized ICP cost functions in Section II-B the pair correspondence is fixed as soon as \(f(x)\) is written down and does not get re-matched. Thus in the point-to-point case, \(f(x)\) will necessarily increase when moving in any direction, including the unobservable ones, such that the eigenvalue-based stability analysis in Section II-C will not reflect the behavior of the actual ICP. Meanwhile the point-to-plane cost function projects the error via the surface normal \(n_i\), such that \(f(x)\) is invariant to the unobservable motions parallel to the wall even though the pair correspondences are not updated. This invariance of the point-to-plane error metric was noted by [7], [12] who stated that it permits flat surfaces to “slide” past each other.

We now consider another example of scan matching observability [10, Fig. 3]: the scanner is placed at the middle of a circular room with featureless walls and rotates around the vertical axis as illustrated in Figure 2.

![Figure 2. Scan matching of featureless circular room](image)

Assuming the radius of the room is 1 unit and that the scanning field of view is 90° wide as shown, the points of the first scan are \(a_1 = [-\sqrt{2}/2 \ \sqrt{2}/2]^T\), \(a_2 = [0 \ 1]^T\) and \(a_3 = [\sqrt{2}/2 \ \sqrt{2}/2]^T\) while the second scan points and associated normals are \(b_1 = [-\sqrt{2}/2 \ \sqrt{2}/2]^T\), \(n_1 = [-\sqrt{2}/2 \ -\sqrt{2}/2]^T\), \(b_2 = [0 \ 0]^T\), \(n_2 = [0 \ 0]^T\) and \(b_3 = [\sqrt{2}/2 \ \sqrt{2}/2]^T\), \(n_3 = [-\sqrt{2}/2 \ -\sqrt{2}/2]^T\). We obtain

\[
A_{pp} = \begin{bmatrix} 3 & -3\sqrt{2}/2 & -\sqrt{2}/2 \\ -3\sqrt{2}/2 & 3 & -\sqrt{2}/2 \\ -\sqrt{2}/2 & 0 & 3 \end{bmatrix}, \quad A_{pn} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}
\]

By direct computation \(A_{pp}\) has

\[
\lambda_1 = 0.7639, \quad \hat{v}_1 = [0.7071 \ 0.6708 \ 0.2236]^T
\]

\[
\lambda_2 = 3, \quad \hat{v}_2 = [0 \ 0.3162 \ -0.9487]^T
\]

\[
\lambda_3 = 5.2361, \quad \hat{v}_3 = [0.7071 \ -0.6708 \ -0.2236]^T
\]

and \(A_{pn}\) has

\[
\lambda_1 = 0, \quad \hat{v}_1 = [1 \ 0 \ 0]^T
\]

\[
\lambda_2 = 1, \quad \hat{v}_2 = [0 \ 1 \ 0]^T
\]

\[
\lambda_3 = 2, \quad \hat{v}_3 = [0 \ 0 \ 1]^T
\]

Once again the linearized point-to-point cost function gives incorrect observability results while the point-to-plane version provides correct ones.
Section II-D represents these cases.

Section II-D correctly analyzed the observability of the point-to-plane ICP by minimizing its (linearized) cost function, but the real algorithm actually re-matches point pairs (step 2) before minimizing the resulting cost function (step 3) c.f. Section II-A. ICP can thus be viewed as a coordinate descent algorithm alternating a descent over matching (by choosing nearest neighbors) and a descent over rotation-translations (note the original proof of local convergence of the ICP [1] is based on this fact).

We now study the general problem in the 2D (plane) setting. A similar analysis could easily be carried out in the 3D case but it would complicate the exposition.

Theorem Consider a cloud of points \( \{a_i\}_{1 \leq i \leq N_a} \) obtained from scanning a 2D environment with maximum curvature \( \kappa \). To simplify the analysis of the stability of ICP at convergence, we create a second cloud \( \{b_i\}_{1 \leq i \leq N_a} \) by duplication, i.e. letting \( b_i = a_i \) and we let \( (\delta R, \delta T) \) be a small rotation-translation close to identity. If the clouds are sufficiently dense the point-to-plane ICP cost function (2) which omits re-matching (step 2 in Section II-A)

\[
 f(\delta R, \delta T) = \sum_{i=1}^{N} \left[ (\delta R_{a_i} + \delta T - b_i) \cdot n_i \right]^2
\]

and thus whose Hessian \( A_m \) is easy to compute, differs term by term from the “true” ICP cost function

\[
 f^\pi(\delta R, \delta T) = \sum_{i=1}^{N} \left[ (\delta R_{a_i} + \delta T - b_{\pi(\delta R, \delta T;i)}) \cdot n_i \right]^2
\]

where \( b_{\pi(\delta R, \delta T;i)} \) is the closest point to \( \delta R_{a_i} + \delta T \) (step 2 of ICP) as follows:

\[
 (\delta R_{a_i} + \delta T - b_i) \cdot n_i = (\delta R_{a_i} + \delta T - b_{\pi(\delta R, \delta T;i)}) \cdot n_i + \psi_i
\]

where \( \psi_i \) is already second order in the function arguments

\[
 |\psi_i| \leq \frac{1}{2} \kappa (2 \alpha(\|\delta R_{a_i} + \delta T - a_i\|) \|n_i\|)^2 \leq \frac{1}{2} \kappa (\alpha(\|x_R\| |a_i\| + \|x_T\|))^2
\]

with \( \alpha > 0 \) a known finite value. The Hessians of \( f(\delta R, \delta T) \) and \( f^\pi(\delta R, \delta T) \) around the minima are thus totally identical.

Note that the density assumption is in fact not a restriction, as in the regions where the cloud is not dense, the closest point is far enough so that \( f \) and \( f^\pi \) also remain locally equal. As a consequence of this Theorem we have:

Corollary In the 2D case, minimizing the point-to-plane ICP cost function \( f(\delta R, \delta T) \) is equivalent up to third-order terms in \( \|x_R\|, \|x_T\| \) to matching closest points and minimizing the resulting cost function \( f^\pi(\delta R, \delta T) \). Thus this leads to identical Hessians (second-order derivatives). Moreover, in the special cases of flat disjoint walls (\( \kappa = 0 \)) the two cost functions agree exactly since \( \psi_i = 0 \).

In other words minimization of the point-to-plane ICP cost function is an accurate model of the ICP’s behavior provided \( (\delta R, \delta T) \) are not excessively large. Remark the examples of Section II-D represent these cases.

Proof Assume the environment being scanned is modeled as a smooth curve \( \gamma(s) \) parameterized with respect to its length \( L \) such that \( s \in [0, L] \). Such a curve has tangent vector \( \gamma'(s) \) with \( \|\gamma'(s)\| = 1 \) and normal vector \( \gamma''(s) \) with \( \|\gamma''(s)\| \leq \kappa \), the curvature. The point cloud \( \{a_i\} = \{b_i\} \in \mathbb{R}^2 \) is obtained by scanning this environment at discrete points:

\[
 \forall t \exists s_i \ b_i = \gamma(s_i)
\]

A set of unit normals \( \{n_i\} \) to the curve (environment) is constructed such that each \( n_i \) originates at the point \( b_i \) on the curve. If the cloud is sufficiently dense, the closest point belong to the neighboring surface, i.e., both \( b_i \) and \( b_{\pi(\delta R, \delta T;i)} = \gamma(\pi(\delta R, \delta T;i)) \). The error made for each term \( i \) is

\[
 (\delta R_{a_i} + \delta T - b_{\pi(\delta R, \delta T;i)}) \cdot n_i = (\delta R_{a_i} + \delta T - b_i) \cdot n_i + (b_i - b_{\pi(\delta R, \delta T;i)}) \cdot n_i = (\delta R_{b_i} + \delta T - b_i) \cdot n_i - \psi_i
\]

To study \( \psi \) expand \( \gamma(s) \) about \( s = s_i \) using Taylor’s theorem with remainder:

\[
 \gamma(s) = \gamma(s_i) + \gamma'(s_i)(s - s_i) + \int_{s_i}^{s} \gamma''(u)(u - s_i) \, du
\]

Take \( s = s_{\pi(\delta R, \delta T;i)} \) and project along the normal \( n_i \):

\[
 (\gamma(\pi(\delta R, \delta T;i)) - \gamma(s_i)) \cdot n_i = (\gamma(\pi(\delta R, \delta T;i)) - s_i) \gamma'(s_i) \cdot n_i + \int_{s_i}^{s_{\pi(\delta R, \delta T;i)}} \gamma''(u) \cdot n_i(u - s_i) \, du
\]

Note \( \gamma'(s_i) \cdot n_i = 0 \) since \( \gamma'(s_i) \) is the (unit) tangent vector to the curve at \( s = s_i \). Taking absolute values of both sides, \( |\psi_i| = |(b_{\pi(\delta R, \delta T;i)} - b_i) \cdot n_i| \)

\[
 \leq \frac{1}{2} \max_{u} ||\gamma''(u)|| |s_{\pi(\delta R, \delta T;i)} - s_i|^2
\]

\[
 \leq \frac{1}{2} \kappa |s_{\pi(\delta R, \delta T;i)} - s_i|^2
\]

\[
 \leq \frac{1}{2} \kappa (\alpha(\|\delta R_{b_i} + \delta T - a_i\|))^2
\]

as claimed. Only the last inequality needs be justified. It stems from the fact that either \( \pi(\delta R, \delta T; i) = i \) if \( \delta R, \delta T \) is too small, and the inequality is obvious, or \( \pi(\delta R, \delta T; i) = j \) but this occurs only if \( \delta R_{b_j} + \delta T \) is closer to \( b_j \) than to \( b_i \), in which case the distance between \( \delta R_{b_j} + \delta T \) and \( b_i = a_i \) becomes greater than half the distance between \( b_j \) and \( b_i \):

\[
 \|b_j - b_i\| \leq 2 \|\delta R_{b_j} + \delta T - a_i\|
\]

Another Taylor expansion yields \( \gamma(s_j) - \gamma(s_i) = \gamma'(s_j)(s_j - s_i) + \int_{s_i}^{s_j} \gamma''(u)(u - s_i) \, du \). Using \( \|\gamma'(s)\| = 1 \) and \( \|\gamma''(s)\| \leq \kappa \) we get \( \|\gamma(s_j) - \gamma(s_i)\| \geq |s_j - s_i| - \frac{1}{2} \kappa (s_j - s_i)^2 \). The term \( \frac{1}{2} \kappa (s_j - s_i)^2 \) can be made as small as \( |s_j - s_i| \) as wanted if the cloud is sufficiently dense. Typically letting \( \frac{1}{2} \kappa (s_j - s_i)^2 \leq (1 - \frac{1}{2}) |s_j - s_i| \) (hence the definition of \( \alpha \) as a function of the density vs curvature) we get

\[
 \|b_j - b_i\| = \|\gamma(s_j) - \gamma(s_i)\| \geq \frac{\alpha}{\kappa} |s_j - s_i|
\]

This proves \( |s_{\pi(\delta R, \delta T;i)} - s_i| \leq \alpha \|\delta R_{b_i} + \delta T - a_i\| \).
In order to evaluate (14) we need to assume an error model and $y$. From Section II-C the estimate

Substituting $i.e.

A. Covariance of point-to-plane ICP estimates

As explained in Section II-B we assume clouds $\{a_i\}$ and $\{b_i\}$ start close to each other such that the ICP cost function takes the sum-of-squares form (5). Section II-B demonstrated that for the point-to-plane variant minimizing $f(x)$ accurately models the ICP algorithm.

Let $\hat{x}$ denote the estimate computed by the ICP from the sensed point clouds $\{a_i\}$ and $\{b_i\}$ and $x^*$ the true transformation, and assume the ICP is an unbiased estimator i.e. $E(\hat{x}) = x^*$. The covariance of the ICP estimate is thus

Based on cost function (5) we define the residuals

From Section II-C the estimate $\hat{x}$ minimizing (5) is

Substituting $y_i$ from (13) this becomes

and (12) becomes

In order to evaluate (14) we need to assume an error model on the residuals $r_i$ in (13). Employing the point-to-plane $H_i$ and $y_i$ terms from (8) yields

where $w_i \in \mathbb{R}^3$ represents the post-alignment error of the $i$th point pair due to sensor error. The residual $r_i = w_i - n_i = n_i^T w_i$ thus represents the projection of $w_i$ and (14) becomes

B. Random Noise Errors

Perhaps the most natural approach [15] is to assume the post-alignment errors $w_i$ are independent and identically normally isotropically distributed as

Under these assumptions $E(w_i w_i^T) = E(w_i)E(w_j^T) = 0$, $i \neq j$ and the double sum in (15) reduces to a single sum ($n_i^2 n_i = 1$ for unit normals):

We can validate (17) as follows: for $r_i = w_i \cdot n_i = n_x w_x + n_y w_y + n_z w_z$ with $w_k \sim N(0, \sigma^2)$ being independent Gaussian scalar variables, $r_i \sim \mathcal{N}(0, \sigma^2 (n_x^2 + n_y^2 + n_z^2)^2) = \mathcal{N}(0, \sigma^2)$ and $\{r_i\}$ are also independent random variables. Returning to (5) and stacking the $N$ terms $y_i$ and $H_i$ into columns $y$ and $H$, we now have $A = H^T H$, $b = H^T y$ in (9) such that $\hat{x} = (H^T H)^{-1} H^T y$ in (11). Stacking the $r_i$ terms into column vector $r$, (13) becomes $y = H \hat{x}^* + r$ where $r \sim \mathcal{N}(0, \sigma^2 I_{N \times N})$ by construction. For this last linear model, it is known [16, Thm. 4.1] that $\hat{x} = (H^T H)^{-1} H^T y$ is the minimum-variance unbiased estimator for which $\text{cov}(\hat{x}) = \sigma^2 (H^T H)^{-1}$, which precisely matches (17). Such an MVU estimator is efficient i.e. the covariance of its estimates attains the Cramér-Rao lower bound [16, Thm. 3.2]

where $\mathcal{I}(x)$ is the Fisher information matrix for this problem. In this case $A = (1/\sigma^2) \mathcal{I}(x)$ and the stability results of Section II-C can be restated in terms of the Fisher information matrix as done in [15], [10], [17].

Despite validation, it turns out (17) provides unrealistically optimistic results. Consider a 3D scan $\{a_i\}$ of a plane wall by a facing scanner placed $d$ units away as shown in Figure 3. A depth image of $N_H$ by $N_V$ pixels (function of the hardware) captures a surface measuring $H$ by $V$ units (function of the optical field of view and distance $d$) such that $a_i = [x_i \ y_i \ d]^T$ where $-H/2 \leq x_i \leq H/2$, $-V/2 \leq y_i \leq V/2$.

Assume a previous scan $\{b_i\}$ with associated surface normals $\{n_i\}$ was captured with the same camera orientation at distance $d'$ such that $b_i = [x'_i \ y'_i \ d'^T], n_i = [0 \ 0 \ -1]^T$.
where $-H'/2 \leq x' \leq H'/2, -V'/2 \leq y' \leq V'/2$. We have (8)

$$H_i = [- (a_i \times n_i)^T \quad -n_i^T ] = [y_i \quad -x_i \quad 0 \quad 0 \quad 0 \quad 1]$$

and compute

$$A = \sum_i H_i^T H_i = \sum_i \begin{bmatrix} y_i^2 & -x_i y_i & 0 & 0 & 0 & y_i \\
-x_i y_i & x_i^2 & 0 & 0 & 0 & -x_i \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
y_i & -x_i & 0 & 0 & 0 & 1 \end{bmatrix}$$

Due to $x_i$, $y_i$ being symmetrically distributed, $\sum x_i = \sum y_i = 0$. Next

$$\sum x_i^2 = 2 \sum_{k=1}^{N_H/2} \left( \frac{k H}{N_H} \right)^2 = \frac{H^2 (N_H + 1) (N_H + 2)}{12 N_H}$$

and $\sum y_i^2$ is obtained by replacing $H$ and $N_H$ with $V$ and $N_V$ respectively. Finally $\sum x_i^2 + \sum y_i^2 = N_H N_V \equiv N$, the number of scanned points. By inspection $A$ possesses three zero eigenvalues with associated eigenvectors $(e_3, e_4, e_5) \in \mathbb{R}^6$, indicating (as expected) that in this case rotations about the $z$ axis and translations along the $x$ and $y$ axes are unobservable to scan matching. $A$ is thus singular and so (17) cannot be evaluated. However we can proceed by removing $x_3$, $x_4$ and $x_5$ from the state vector $x = [x_R \quad x_T]$, thus deleting the 3rd, 4th and 5th column of $H_i$ or rows and columns of $A$ such that we work with the subspace of observable parameters. In this case

$$\text{cov}(\hat{x}) = \sigma^2 \begin{bmatrix} 12 N_V \quad 0 \quad 0 \\
0 & \frac{12 N_H}{H^2 (N_H + 1) (N_H + 2)} & 0 \\
0 & 0 & \frac{1}{N} \end{bmatrix}$$

Consider parameter values representative of a Kinect. Assume $d = 2$ m for which $\sigma \approx 1$ cm [18, Fig. 10]. The Kinect has a field of view of $57^\circ$ horizontally and $43^\circ$ vertically such that $H = 4 \tan 28.5^\circ \approx 2.17$ m, $V = 4 \tan 21.5^\circ \approx 1.58$ m, and its depth images have $N_H = 640, N_V = 480$. We obtain the standard deviations

$$\sqrt{\text{diag}(\text{cov}(\hat{x}))} = [0.06^\circ \quad 0.04^\circ \quad 0.02 \text{ mm}]$$

It is clear that the computed results are several orders of magnitude too optimistic — it is simply impossible to obtain displacement estimates with sub-millimeter precision by scan matching images from a low-cost Kinect scanner. The reason is that the independent Gaussian noise (16) which models random errors becomes a negligible effect as the number of scanned points increases. We thus need to focus on a different noise effect, namely resolution errors of the scanner.

C. Resolution Errors

Suppose that $N$ point pairs $\{(a_i, b_i)\}$ are scanned and that the environment consists of $K$ planes represented as disjoint subsets $M_1, M_2, \ldots, M_K$ of these pairs. We will assume the post-alignment errors $w_i$ created by resolution errors are correlated with each other on a given plane $M_k$ but independent of the other planes and among scanner axes, and denote the average axial resolution by $\delta$, such that

$$E(w_i^T w_i^T) = \delta^2 I_3 1_{M_k \times M_k} (i, j)$$

where $1$ is the indicator function defined as

$$1_M(x) = \begin{cases} 1 & \text{if } x \in M, \\ 0 & \text{if } x \notin M. \end{cases}$$

Using (18) and the definition $A = \sum H_i^T H_i$ (15) becomes

$$\text{cov}(\hat{x}) = \delta^2 A^{-1} \sum_i \sum_j \left( H_i^T H_i 1_{M_k \times M_k} (i, j) \right) A^{-1}$$

Using $H_i$ from (8) and denoting $M_k \supset \{n_k\} \equiv n_k$ as the normal of a given plane, the double sum in (19) becomes

$$\sum_k \sum_{i \in M_k} \left[ (a_i \times n_k)(\sum_{j \in M_k} a_j \times n_k)^T \quad n_k (\sum_{j \in M_k} a_j \times n_k)^T \right] n_k n_k^T$$

and $A$ can be expressed as

$$A = \sum_k \sum_{i \in M_k} \left[ (a_i \times n_k)(a_i \times n_k)^T \quad n_k (a_i \times n_k)^T \right]$$

In principle (19) computes the covariance of $\hat{x}$, although it is unwieldy to use. We can simplify by assuming that all planes $M_k$ are sufficiently compact such that the mean $1/N_k \sum_{j \in M_k} a_j \approx a_i \in M_k$. Under this assumption the double sum in (19) simplifies to

$$\sum_k N_k \sum_{i \in M_k} \left[ (a_i \times n_k)(a_i \times n_k)^T \quad n_k (a_i \times n_k)^T \right] n_k n_k^T$$

Finally by assuming $N_k = N/K$ i.e. each plane contains the same number of points, we obtain a much simpler version of (19):

$$\text{cov}(\hat{x}) = \delta^2 N/K A^{-1} = \delta^2 N/K \left( \sum_i H_i^T H_i \right)^{-1}$$

The assumption of planes $M_k$ having the same number of points can be met by implementing normal space directed sampling [12] at step 1 of the ICP algorithm (c.f. Section II-A). This consists of bucketing the computed normals of a scan based on their direction in angular space then
sampling evenly across these buckets. Typically, in a standard room, one can implement the formula as an approximation to the true covariance by letting \( K = 5 \), that is the image is essentially made of three visible walls, roof and floor.

We return to the earlier example of scan matching a plane wall. Since there is only one plane \( K = 1 \). Referring to earlier results, \( \delta \) yields

\[
\text{cov}(\hat{x}) = \delta^2 N \begin{bmatrix}
12N_V & 0 & 0 \\
0 & H^2(N_H+1)(N_H+2) & 0 \\
0 & 0 & \frac{1}{N}
\end{bmatrix}
\]

Taking \( \delta = 0.01 \) m as the resolution error of the Kinect at \( d = 2 \) m [18, Fig. 10] and \( H = 2.17 \) m, \( V = 1.58 \) m, \( N_H = 640 \), \( N_V = 480 \) as before we obtain

\[
\sqrt{\text{diag}(\text{cov}(\hat{x}))} = [31.8^\circ \quad 20.0^\circ \quad 1 \text{ cm}]
\]

which are reasonable numbers for Kinect scan matching performance at a 2 m distance. Recall from Section III-B that we are analyzing the observable subspace and that the covariance along unobservable directions is infinite.

IV. CONCLUSIONS

This paper has formally demonstrated that the linearized cost function of the point-to-plane ICP correctly models the observability of the algorithm, while the point-to-point version does not. We then derived a closed-form expression for computing estimate covariances [14], which correctly models the uncertainty due to sensing errors and the observability of the environment. Using parameter values representative of a Kinect, we demonstrated that random noise provides negligibly small covariance, whereas resolution error is the dominant source of uncertainty in scan matching estimates.

ACKNOWLEDGMENT

The work reported in this paper was partly supported by the Cap Digital Business Cluster TerraMobilita Project.

REFERENCES

[1] P. J. Besl and N. D. McKay, “A method for registration of 3-D shapes,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pp. 239–256, February 1992.
[2] Y. Chen and G. Medioni, “Object modelling by registration of multiple range images,” Image and Vision Computing, vol. 10, no. 3, pp. 145–155, April 1992.
[3] F. Lu and E. E. Milios, “Robot pose estimation in unknown environments by matching 2D range scans,” in Proceedings of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, June 1994, pp. 935–938.
[4] J. Nieto, T. Bailey, and E. Nebot, “Recursive scan-matching SLAM,” Robotics and Autonomous Systems, vol. 55, no. 1, pp. 39–49, January 2007.
[5] A. Mallios, P. Rida, D. Ribas, F. Maurelli, and Y. Pettitot, “EKF-SLAM for AUV navigation under probabilistic sonar scan-matching,” in Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, October 2010, pp. 4404–4411.
[6] S. Rusinkiewicz, ICP stability. [Online]. Available: http://www.cs.princeton.edu/~smr/papers/icpstability.pdf
[7] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, “Geometrically stable sampling for the ICP algorithm,” in Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling, Banff, Canada, October 2003, pp. 260–267.
[8] O. Bengtsson and A.-J. Baerveldt, “Localization in changing environments – estimation of a covariance matrix for the IDC algorithm,” in Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, Hawaii, USA, October 2001, pp. 1931–1937.
[9] ——, “Robot localization based on scan-matching — estimating the covariance matrix for the IDC algorithm,” Robotics and Autonomous Systems, vol. 44, no. 1, pp. 29–40, July 2003.
[10] A. Censi, “An accurate closed-form estimate of ICP’s covariance,” in Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, April 2007, pp. 3167–3172.
[11] ——, “On achievable accuracy for pose tracking,” in Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, May 2009, pp. 1–7.
[12] S. Rusinkiewicz and M. Levoy, “Efficient variants of the ICP algorithm,” in Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, Quebec City, Canada, May 2001, pp. 145–152.
[13] K. Klasing, D. Althoff, D. Wollherr, and M. Buss, “Comparison of surface normal estimation methods for range sensing applications,” in Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, May 2009, pp. 3206–3211.
[14] C.-T. Chen, Linear System Theory and Design, 3rd ed. Oxford University Press, 1999.
[15] T. Hervier, S. Bonnabel, and F. Goulette, “Accurate 3D maps from depth images and motion sensors via nonlinear kalman filtering,” in Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal, October 2012, pp. 5291–5297.
[16] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall, 1993.
[17] A. Censi, “On achievable accuracy for range-finder localization,” in Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, April 2007, pp. 4170–4175.
[18] K. Khoshelham and S. Oude Elberink, “Accuracy and resolution of Kinect depth data for indoor mapping applications,” Sensors, vol. 12, no. 2, pp. 1437–1454, February 2012.