



# Correspondenceless scan-to-map-scan matching of 2D panoramic range scans<sup>☆</sup>

Alexandros Filotheou<sup>\*</sup>, Andreas L. Symeonidis, Georgios D. Sergiadis, Antonis G. Dimitriou

Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

## ARTICLE INFO

### Keywords:

Robot localisation  
Panoramic 2D LIDAR  
Scan-to-map-scan matching

## ABSTRACT

In this article a real-time method is proposed that reduces the pose estimate error for robots capable of motion on the 2D plane. The solution that the method provides addresses the recent introduction of low-cost panoramic range scanners (2D LIDAR range sensors whose field of view is 360°), whose use in robot localisation induces elevated pose uncertainty due to their significantly increased measurement noise compared to prior, costlier sensors. The solution employs scan-to-map-scan matching and, in contrast to prior art, its novelty lies in that matching is performed without establishing correspondences between the two input scans; rather, the matching problem is solved in closed form by virtue of exploiting the periodicity of the input signals. The correspondence-free nature of the solution allows for dispensing with the calculation of correspondences between the input range scans, which (a) becomes non-trivial and more error-prone with increasing input noise, and (b) involves the setting of parameters whose output effects are sensitive to the parameters' correct configuration, and which does not hold universal or predictive validity. The efficacy of the proposed method is illustrated through extensive experiments on public domain data and over various measurement noise levels exhibited by the aforementioned class of sensors. Through these experiments we show that the proposed method exhibits (a) lower pose errors compared to state of the art methods, and (b) more robust pose error reduction rates compared to those which are capable of real-time execution. The source code of its implementation is available for download.

## 1. Introduction

Mobile robot localisation on one plane is a well-studied field in robotics and several diverse approaches have been proposed in the past. Probabilistic methods, e.g. the Kalman filter [1] or Monte Carlo Localisation (MCL) methods [2–4] have been applied to the task of pose tracking and have proven their success with respect to tracking efficacy. At the same time, probabilistic methods are robust to sensor noise, discrepancies between the robot's environment and its corresponding map, motion model mismatch with regard to the true kinematics of the robot, and pose uncertainty [5–7]. These methods have also been employed for global localisation, where a system is tasked with estimating the robot's pose under global pose uncertainty [8–10].

In practice, the pose estimate of localisation methods is beset by an error which is often measured in centimeters or even decimeters [11, 12]. These errors are due to range scan measurements being distorted by noise, or the map of the environment not matching the latter adequately. Other reasons include the map being expressed as a finite

resolution grid, noisy or faulty and ever-drifting odometry (if at all available), and the nature of the observation model. In certain conditions such as industrial ones [13,14], the magnitude of the estimate's error is required to lie within constrained specifications. Therefore, standalone or prosthetic methods have been employed or used in tandem with well-established sturdy probabilistic (or otherwise) localisation methods, with many of them leveraging measurements from onboard pre-existing Light Detection And Ranging (LIDAR) sensors.

LIDAR sensors have become popular in robot localisation due to their high measurement precision, high update frequency, and almost no need for preprocessing. The use of panoramic LIDAR sensors was for a long time constrained to higher price ranges, low measurement noise, and in the context of industry. In recent years, however, cheaper but less accurate LIDAR sensors have become available. The former fact facilitates their adoption and usage in research, but the latter poses a challenge to both the robustness and accuracy of localisation methods.

A class of prosthetic localisation methods improves the robot's pose estimate by extracting the relative translation and orientation

<sup>☆</sup> This work was supported by the European Union and Greek National Funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call Research Create Innovate under Project T2EDK-02000.

<sup>\*</sup> Corresponding author.

E-mail address: [alefilot@auth.gr](mailto:alefilot@auth.gr) (A. Filotheou).

<https://doi.org/10.1016/j.array.2023.100288>

Received 23 February 2022; Received in revised form 15 October 2022; Accepted 19 April 2023

Available online 26 April 2023

2590-0056/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

between (a) the range scan captured from the robot's actual pose and (b) a virtual range scan derived by ray-casting the map of the robot's environment from the robot's pose estimate. Due to its operating principle, this technique may be termed *scan-to-map-scan matching*. Related methods rest on establishing correspondences between features, points, points and lines, distributions, or points and distributions. However, methods resting on locating features presuppose structured environments and the existence of features in a sufficiently undisturbed state. Furthermore different environments exhibit different features, and therefore the use of feature-finding methods must be tailored in an ad-hoc manner. The majority of the rest use scan-matching methods based on the Iterative Closest Point (ICP) method [15–17], research on which is ongoing for several decades [18–23]. However, ICP-based methods are subject to the perplexities delimited by the underlying process of establishing correspondences between the two input scans, which are exacerbated in high levels of measurement noise. Furthermore, their use and performance is hindered by the needed tuning of the plethora of parameters governing their response [24]. In any case, the methodology of approaches that solve the scan-to-map-scan matching task rests on establishing correspondences between the input scans.

In this article we propose a method that solves scan-to-map-scan matching in real time and in closed form that specifically targets panoramic 2D LIDAR sensors. The central contributions of this article are the following:

- To the best of the authors' knowledge, the first real-time method addressing the full 3D-matching of real-to-virtual 2D panoramic range scans that operates without establishing correspondences of any kind between input scans
- The introduction of a method that aims at reducing the orientation error to lower than the sensor's angle increment compared to relevant prior work
- The parameter set needed by the proposed method is smaller in size and more intuitive to tune than those of state-of-the-art methods, and trades execution time for accuracy
- The extensive and thorough evaluation of state-of-the-art scan-matching methods and of the proposed method on the task of scan-to-map-scan matching, over five public domain benchmark datasets and measurement noise levels from common-use commercially available panoramic sensors

The proposed method assumes that (i) a panoramic range scan, (ii) the map of the environment in which the robot operates, and (iii) a pose estimate residing in the vicinity of the robot's true pose are available. After computing a virtual range scan from the measurement sensor's pose estimate, the method updates it by reducing first the error of the orientation estimate and then that of the position estimate. The process is iterated until sufficient convergence conditions are met. The estimation of the 3D transformation between the robot's true and estimated pose is facilitated by the exclusive use of the first term of the Discrete Fourier Transform of the difference in ranges between the input real scan and computed virtual scans, where the range scan termed "real" is a measurement of a physical range finder and those termed "virtual" are generated by raycasting the map of the robot's environment.

In summary, (a) the orientation errors of the proposed method are independent of the initial angular displacement, and (b) it allows matching to preserve robustness in high levels of measurement noise and map distortions. Specifically, the approach proposed is shown to be more robust to measurement noise and map distortions than real-time state-of-the-art methods in the sense of proportion of cases where the pose estimate error is reduced after its application, and more accurate in terms of pose error magnitudes.

The remainder of this paper is structured as follows: Section 2 formulates the problem and the objective of its solution. Section 3 defines necessary notions. Section 4 provides a bibliographical ex-

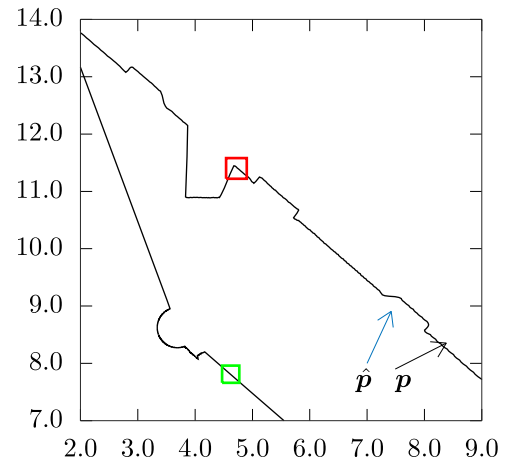


Fig. 1. In principle, a typical localisation condition: The robot's actual pose is  $p$  but its estimate  $\hat{p}$  is displaced in terms of position and orientation. The rate of changes in the portion of the environment circumscribed in red is greater than that of those in the portion enclosed in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

position of the current state-of-the-art solutions to the problem of performing scan-to-map-scan matching in order to improve the pose estimate of a robot capable of motion in the 2D plane equipped with a 2D range sensor. Section 5 illustrates the method of solving the stated problem that this paper proposes. Section 6 presents the experimental setup and the performance of the proposed method for robots using available panoramic sensors, in realistic conditions. Section 7 gives key characterisations and discusses the limitations of the proposed method, and Section 9 provides the aims of future work. Section 8 provides a number of applications where scan-to-map-scan matching is utilisable and useful. Finally, Section 10 summarises the work presented and concludes the paper.

## 2. Problem formulation

**Problem P.** Let a mobile robot capable of motion in the  $x$ - $y$  plane be equipped with a coplanarly mounted 2D panoramic range scan sensor emitting  $N_s$  rays. Let also the following be available or standing:

- The map  $M$  of the environment the robot operates in
- A 2D range scan  $S_R$ , captured from its range sensor's – unknown and sought-for – pose  $p(l, \theta)$ ,  $l = (x, y)$
- An estimate of the range scan sensor's pose  $\hat{p}(\hat{l}, \hat{\theta})$  in the map's frame of reference, where  $\hat{l} = (\hat{x}, \hat{y})$  is in a neighbourhood of  $l$

Then the objective is to reduce the 2-norm of the sensor's pose error  $e(p, \hat{p}) \triangleq p - \hat{p}$  from its initial value

$$\|e(p, \hat{p})\|_2 = ((x - \hat{x})^2 + (y - \hat{y})^2 + (\theta - \hat{\theta})^2)^{1/2} \quad (1)$$

by improving the sensor's pose estimate to  $\hat{p}'(\hat{x}', \hat{y}', \hat{\theta}')$  so that

$$\|e(p, \hat{p}')\|_2 < \|e(p, \hat{p})\|_2 \quad (*)$$

Assuming that the sensor's pose with respect to the robot's frame of reference is known, the correction of the sensor's pose estimate is equal to the correction of the robot's pose estimate with respect to the map's frame of reference. An example of a premise of Problem P is depicted in Fig. 1. Pose estimate  $\hat{p}$  is supplied externally from a localisation system in the case of pose tracking, or as a pose hypothesis in the case of global localisation.

## 3. Definitions

**Definition 1.** *Definition of a range scan captured from a conventional 2D LIDAR sensor.* A conventional 2D LIDAR sensor provides a finite

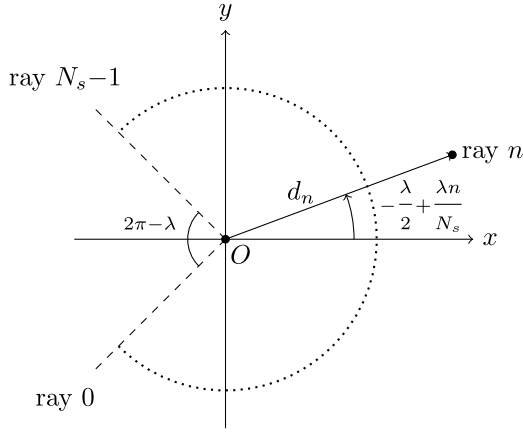


Fig. 2. The (local) frame of reference of a typical conventional range sensor. The sensor is located at  $O(0,0)$  and its heading is that of the  $x$  axis.

number of ranges, i.e. distances to objects within its range, on a horizontal cross-section of its environment, at regular angular and temporal intervals, over a defined angular range [25]. We define a range scan  $S$ , consisting of  $N_s$  rays over an angular range  $\lambda$ , to be an ordered map  $S : \Theta \rightarrow \mathbb{R}_{\geq 0}$ , where  $\Theta = \{\theta_n \in [-\frac{\lambda}{2}, \frac{\lambda}{2}) : \theta_n = -\frac{\lambda}{2} + \lambda \frac{n}{N_s}, n = 0, 1, \dots, N_s - 1\}$ . Angles  $\theta_n$  are expressed relative to the sensor's heading, in the sensor's frame of reference.

Fig. 2 depicts the geometry of a typical conventional 2D LIDAR sensor, where  $d_n = S[-\frac{\lambda}{2} + \frac{\lambda n}{N_s}]$  is the range returned by ray  $n$ .

**Definition II (Panoramic 2D Range Scan).** The angular range of a 2D LIDAR sensor is symmetrically distributed on either side of its  $x$ -axis. Each ray is equiangularly spaced from its neighbouring rays (with the exception of the first and last rays if  $\lambda < 2\pi$ ). When  $\lambda = 2\pi$ , the range scan returned by the sensor is termed panoramic.

**Definition III.** *Scan-matching using a 2D LIDAR sensor* (adapted for use in two dimensions from [17]). Let two range scans as defined by Definition I,  $S_R$  and  $S_V$ , be captured from a LIDAR sensor operating in the same environment at both capturing times. Let  $p_V(x_V, y_V, \theta_V)$  be the pose from which the sensor captured  $S_V$ , expressed in some coordinate system (usually a past pose estimate of the sensor). The objective of scan-matching in two dimensions is to find the roto-translation  $q = (t, \theta)$ ,  $t = (\Delta x, \Delta y)$  that minimises the distance of the endpoints of  $S_V$  roto-translated by  $q$  to their projection on  $S_R$ . Denoting the endpoints of  $S_V$  by  $\{p_V^i\}$ , in formula:

$$\min_q \sum_i \|p_V^i \oplus q - \Pi(\{S_R, p_V^i \oplus q\})\|^2 \quad (2)$$

The symbol " $\oplus$ " denotes the roto-translation operator  $p_V^i \oplus (t, \theta) \triangleq R(\theta)p_V^i + t$ , where  $R(\theta)$  is the 2D rotation matrix for argument angle  $\theta$ , and  $\Pi(\{S_R, p_V^i \oplus q\})$  denotes the Euclidean projector on  $S_R$ .

**Remark I.** Scan-matching is employed in robotics as a means to odometry, primarily in non-wheeled robots where no encoders can be utilised, or as a useful ameliorator of the ever-drifting encoder-ed odometry: scans captured at consecutive time instances, inputted to a scan-matching algorithm, convey an estimate as to the pose of the scan sensor at the second capture time relative to that captured first. Scan-matching is being successfully employed in the tasks of simultaneous localisation and mapping [26–28], local map construction [29–31], and in people-tracking systems [32].

**Definition IV.** *Definition of a map-scan.* A map-scan is a virtual scan that encapsulates the same pieces of information as a scan derived from

a physical sensor. Only their underlying operating principle is different due to the fact the map-scan refers to distances to obstacles within the map of the robot's environment rather than within the environment itself—hence its virtuality. A map-scan is captured from a virtual sensor and derived by means of locating intersections of rays emanating from the estimate of the sensor's pose and boundaries demarcating obstacles in the map.

**Definition V.** *Scan-to-map-scan matching in two dimensions.* Scan-to-map-scan matching is defined in the same way as scan-to-scan matching (Definition III) but with  $S_V$  now derived not from the physical environment of the robot but from its map.

**Remark II.** The benefit of matching (a) a map-scan derived from a virtual sensor from its estimated pose to (b) a scan derived from a physical sensor from its actual pose is that the correction of the sensor's pose estimate provides the correction of the robot's pose estimate: Assume that a pose hypothesis exists in the general vicinity of the true pose of a mobile robot equipped with a 2D range-scan sensor; assuming that the range sensor is fixed at the same pose relative to the robot in both real and virtual environments, the roto-translation of the virtual scan's endpoints that minimises their distance to their projection on the physical scan equals the roto-translation that, when applied to the robot's estimated pose, will minimise its displacement with respect to its real pose. Therefore extracting the relative roto-translation of the virtual scan with respect to the real scan can be used as a correction of the estimate of the robot's pose within the map. The significance of this correction lies in the fact that it may be used to reduce a robot's pose estimate error during pose-tracking, or to facilitate global localisation.

**Remark III.** In contrast to the problem of scan-matching, scan-to-map-scan matching is an inherently coupled problem: since the geometry of the endpoints of the real scan's rays can be matched, without loss of generality, only from the true pose of the robot, the robot's orientation can be extracted if and only if its location estimate coincides with its real location, and its location can only be extracted if and only if its orientation estimate equals its real orientation. However, both are, in principle, unequal. This coupling is why an iterative method is required, as we shall see in Section 5.3.

**Remark IV.** In contrast to scan-matching, where  $S_V$  is an immutable measurement laden with inevitable occlusions, in scan-to-map-scan matching  $S_V$  is generated from the map, which captures the robot's working environment in its entirety, thus transferring this property to the virtual scan. This subtle difference makes it in principle possible for the virtual scan  $S_V$  to match the immutable measurement  $S_R$

- exactly, i.e. to recover the robot's true pose with arbitrary accuracy (in the ideal case of perfect measurements and complete map-to-environment coincidence)
- without the need to establish correspondences between the two scans (as this function has been invented for, and primarily facilitates, the matching of sets that are in principle overlapping in some areas but not in others, i.e. for the task of scan-matching)

**Remark V.** The importance of dispensing with the establishing of correspondences – aside from the fact that, strictly, they are not necessary in the scan-to-map-scan matching task – is in the fact that by the same act the externally-configurable parameters which govern this function are also dispensed with. Matching methods that rely on the establishing of correspondences require the setting of these parameters, some of whose tuning has been proved to be non-intuitive, effort-consuming, and not globally-fitting to any environment or even to different poses in the same environment [24].

#### 4. State of the art approaches

This section serves as a recounting of approaches that aim to improve a robot's pose estimate during pose-tracking or perform global localisation that rest on the principle of scan-to-map-scan matching.

In general, scan-to-map-scan methods pertaining to 2D LIDAR range scan sensors compute the 3D transformation that aligns the input sensed scan (Definition I) to a map-scan (Definition IV) best, in the sense of minimising an error or alignment metric (e.g. Eq. (2)). In coarse classification this is achieved by (a) correlating features extracted from both input scans [11,33,34], (b) resting on scan-matching techniques [17,23,35–38] due to the indistinguishability of a virtual scan with respect to a real scan from the point of view of a scan-matching method (Definition V; Remark II) [14,39–42], and (c) by other means, e.g. spectral techniques [43], Gauss–Newton optimisation [44], Fourier analysis [24], or simply by randomly sampling the available pose space [45].

The entirety of real-time scan-to-map-scan matching methods mentioned above perform matching by establishing correspondences between input scans (whether they be between raw measurements, features, or other scan characteristics), and require the manual setting of parameters that govern it, while these do not hold universal or predictive validity (for an example pertaining to ICP-based methods see [24]). Establishing correspondences for facilitating matching, in particular, is a technique suitable for low-noise scans, which, in theory at least, becomes cumbersome and ineffective as input noise increases. The method proposed in this article addresses the above issues and exhibits the merits found in Section 1. The rest of this section delves deeper into each aforementioned method and describes its methodology.

In [33] a matching algorithm that deals in range scan features is introduced. The algorithm operates by detecting rotation- and translation-invariant features that are only computable in real-time (such as extreme values in the polar representation of a range scan) in both real and virtual scans. Subsequently, correspondences are established between them. The roto-translation between the two is then computed as the optimal transformation for mapping the latter's features to the former's.

In [45] an elementary stochastic search algorithm that corrects the robot's translational and rotational errors due to odometric drift is employed. This auxiliary localisation behaviour is activated whenever an error measure is found to be above a preset threshold. This measure is based on the relative deviation in detected ranges between rays from a real scan and a map-scan. To avoid having to correct for the motion of the robot while scan-matching, the robot is assumed to be standing still for the whole duration of its pose correction. Therefore whenever the error measure is found to be above its preset threshold the algorithm halts the robot's motion and picks a random pose in the neighbourhood of its estimated pose. It then takes a virtual range scan from that pose and computes the new error. If the error is lower than the one found for the previous estimated pose, a new iteration starts, this time centred around the newly found pose. If not, the algorithm keeps guessing poses until it finds one whose error is lower than the previous one. The final pose is then taken as the true pose of the robot, allowing for a correction of the odometry. Experiments performed with this method showed that it was able to correct a radial pose error of 0.3 m to 0.07 m, and an angular pose error of 0.393 rad to 0.01 rad.

The authors of [11] use scan-matching in order to improve the solution to the global localisation problem. Assuming that the robot's environment is structured and without any sort of symmetries, the method identifies the robot's global orientation by employing the HSM scan-matcher [46]. HSM is used to obtain the robot's heading by matching the lines in the map of the environment with the lines from the 2D range scan taken at the robot's initial pose. Having found the robot's orientation, they estimate the robot's location by calculating the likelihood that each location on the map's grid produced the input laser scan. This likelihood is extracted by using the beam endpoint model

[6]. The robot's position is the location from which the virtual scan that scored the maximum probability was captured.

Likewise, in the context of global localisation, the method in [43] first generates the generalised Voronoi diagram of the 2D grid map. Its nodes are taken to be initial hypotheses of where the robot is posed. From there virtual scans over an angular range of  $2\pi$  are computed using raycasting the grid map. Correspondences between each virtual scan and the scan captured from the physical sensor are then established by using a spectral technique [47]. The latter finds pairwise geometric relationships between its two input scans. These correspondences are then used to generate 2D geometric histograms that encode a sense of similarity between the true scan and all virtual scans. The nodes from which the latter were captured are then ranked according to this similarity measure and a threshold based on the correlation coefficient of all combinations of scans is used to extract a subset of candidate poses. This process is used to quickly sift between all candidate poses. The final pose is that which achieves the maximum number of correspondence pairs.

In [39] the occupancy grid map is first converted to a signed fitness map which encodes the distance the closest obstacle for a given sensor location. Through the fitness map the measurements of the 2D range sensor are related to the environment's map without extracting features from either. The global localisation problem is then formulated as an optimisation problem, where particle swarm optimisation is used to explore the pose space to search for the most likely solution. This is done by maximising the fitness function. In order to further improve the search precision, scan-to-map-scan matching is performed via ICP: from the poses of particles which hold the top fitness values virtual scans are captured and matched against the latest scan measurement. The output pose is that which resulted from scan-to-map-scan matching and whose updated fitness value is the maximum among all similarly treated particles.

By contrast, in order to solve Problem P in the context of localisation of autonomous forklifts, the method introduced in [14] solves scan-to-map-scan matching in two steps: Given a vehicle's pose estimate obtained through the use of MCL with KLD sampling [4], the orientation between the real and virtual scans is first estimated via scan-matching the two using an ICP variant. Specifically, scan-matching is performed through the all-encompassing, highly accurate, efficient, and outperformer of the state-of-the-art scan-matchers: PLICP [17].<sup>1</sup> The authors' findings indicate that the improvement of the location estimate through scan-matching with PLICP is unstable. Therefore they conclude that utilising PLICP in industrial warehouse settings, where milli-meter accuracy is required, in order to extract the relative translation between the two scans is precarious and unsuitable. Given that the forklift's orientation estimate error has decreased to as much as  $0.13^\circ$ , the position estimate error is corrected by iteratively performing scan-to-map-scan matching through a process that approximates the displacement error at each iteration by a function of the first element of the Discrete Fourier transform of the difference in ranges between the real scan and that iteration's map-scan.

A similar pipeline is presented in [44]. Instead of using PLICP off the shelf, the authors develop a scan-matching algorithm that aligns real scans with map-scans taken from MCL's pose estimate using the Gauss–Newton method. This alignment is performed layer-by-layer in increasing map resolution. Experiments conducted with a real robot in unstructured environments show that the scan-to-map-scan matcher achieves an average location accuracy of 0.017 m and an average orientation accuracy of  $0.5^\circ$ . In [40] and [41] PLICP is additionally used as odometry whenever an odometry error measure is found to be larger than a set threshold. However, in contrast to [44], scan-to-map-scan

<sup>1</sup> In principle, if the assumptions of Problem P are fulfilled, any of the available 2D scan matchers may be used to aim at attaining objective (\*). A comprehensive review of scan-matching methods may be found in [42].



matching is performed by chaining PLICP in tandem with GPM [48] in order to mitigate the effects of large angular errors on PLICP.

The method introduced in [49] jointly considers odometry, scan-matching and scan-to-map-scan matching of 2D LIDAR scans with cadastral maps for localisation of autonomous vehicles in outdoor scenarios. These are used as constraints in the solution of a graph optimisation problem that computes the most likely vehicle pose given measurements from the 2D range sensor. With regard to the cadastral plans, non-building objects are filtered-out from the real laser observation using a split and merge approach, which is combined with weighted line fitting. The input range scan and the one derived from the map are then aligned via Generalised ICP, and the resulting pose transform is then added to the graph if and only if ICP has converged. At the same time, a method for detecting the ambiguity regarding the longitudinal position of the vehicle arising in corridor-like environments is introduced. In [42] scan-to-map-scan matching is employed in tandem with a particle filter. From the pose estimate of the latter, a map-scan is computed and then matched against the range scan captured from the physical sensor using PLICP. Feeding back the resulting pose estimate to the population of the particle filter in the form of a multitude of particles is shown to exhibit lower pose errors compared to [44], where the resulting pose estimate is fed back in the form of only one particle. Furthermore it is shown that this method of feedback exhibits increased robustness compared to [14], where the particle filter is initialised anew around the resulting estimate.

In [34] the proposed global localisation method is divided into two phases: an offline and an online phase. During the offline phase, the input map is partitioned into a 2D grid according to a set resolution. A rotation-invariant location signature is then generated for the virtual panoramic 2D range scan that is captured from each traversable cell location within the map. All resulting signatures are then inserted into an ANN search tree. In the online phase, for each incoming laser scan, a signature of the input scan is generated in the same way as during the offline phase. Then the signature is used for retrieving the neighbouring candidate locations from the search tree: the output location is that whose virtual scan's signature is the closest neighbour of the signature of the input measurement scan. In order to obtain the orientation of the robot's pose, a virtual scan is generated from the determined location and registered to the measurement panoramic 2D range scan after pre-processing and pre-aligning steps. The angular registration is performed in 1-degree steps and the robot's orientation is the one that which records the minimum relative entropy between the virtual and real scan.

In [24] the solution to the global localisation problem is given entirely online. At first a dense cloud of hypotheses is generated within the unoccupied interior of the robot's map. Subsequently each hypothesis is inputted to a rotation subsystem, which at first captures a virtual scan from the hypothesis' pose, projects it to two dimensions, discretises it, and matches it to the similarly-treated real scan via the application of Fourier-Mellin Invariant matching [50]. The latter provides the orientation difference between the two scans and, most importantly, a measure of their similarity. After rotating the pose hypothesis the translation component displaces it in order to match the location of the sensor's real location. At the end all similarity measures are ranked and the pose hypothesis with the greatest similarity degree is outputted as the system's pose estimate.

In recent months a number of new scan-matching methods, offering improvements on established methods or introducing new innovations, have been introduced. In [36] NDT is used to model the sensor's environment in order to address its uncertainties and constraints. The pose transformation between successive poses—the solution to the optimisation problem of Eq. (2)—is given by a modified stochastic particle swarm optimisation approach that incorporates inertia weights in its formulation. These weights encode the momentum expressed by forces attracting the particle in keeping its current velocity, forces that

bias its motion towards its personal optimum, and forces that direct it towards the swarm's optimum pose thus far.

In contrast to NDT however, which establishes correspondences by considering the distance of point positions to voxel distributions, VGICP [23] aggregates the distribution of each point in the voxel and establishes correspondences between these distributions and target distributions, thus making VGICP a voxel-based distribution-to-multi-distribution approach. This approach yields valid voxel distributions even when there are few points in a voxel, resulting in an algorithm that is robust to changes in voxel resolution. VGICP extends GICP [37] in order to avoid costly nearest neighbour search, while reducing its execution time.

In [38] a certifiable scan-matching algorithm is introduced. The registration solution is first made insensitive to large number or spurious correspondences by reformulating the problem in manner that uses a truncated least squares cost. Rotation, translation, and scale between the two input scans is decoupled with the use of a general graph-theoretic framework, which allows for the pruning of outliers by finding the graph's maximum clique. Scale and translation are shown to be solvable in polynomial time via an adaptive voting scheme, while rotation is solved by being relaxed to a semidefinite program.

## 5. The proposed method

This section offers a dissection of the proposed method. In order to solve Problem P, a map-scan  $S_V$  is computed initially and each time the pose estimate  $\hat{p}$  is updated.

The problem is iteratively decomposed into two disjunctive sub-problems. The first is estimating the relative angle between  $S_R$  and  $S_V$  under the assumption that both are captured from the same location. The second is estimating the relative displacement of  $S_V$  with respect to  $S_R$  under the assumption that both are captured from poses of the same orientation. Solving the first sub-problem is followed by the solution to the second sub-problem. This process is iterated until termination conditions are met.

The orientation and location correction submethods are presented in Sections 5.1 and 5.2. Section 5.3 presents the method of how these two are woven together into the system that solves Problem P that is proposed in this study.

### 5.1. Orientation correction

Let the assumptions of Problem P be standing. Let additionally  $\hat{l} = l$ , that is, only the sensor's orientation needs to be estimated. Then let virtual scan  $S_V$  be computed via ray-casting  $\mathbf{M}$  from  $\hat{p}$ . An estimate of the rotation of  $S_V$  with respect to  $S_R$  may be found in the following way.

Let a panoramic range scan  $S$  be captured from pose  $(x, y, \theta)$  in some coordinate frame (Definition 1). The coordinates of the end-point of the scan's  $n$ th ray  $n = 0, 1, \dots, N_s - 1$  within that frame of reference are  $(x_n, y_n)$ :

$$x_n - x = d_n \cos(\theta + \frac{2\pi n}{N_s} - \pi) = -d_n \cos(\theta + \frac{2\pi n}{N_s}) \quad (3)$$

$$y_n - y = d_n \sin(\theta + \frac{2\pi n}{N_s} - \pi) = -d_n \sin(\theta + \frac{2\pi n}{N_s}) \quad (4)$$

Here we make the observation that  $-(x_n - x)$  and  $(y_n - y)$  are, respectively, the real and imaginary parts of the complex quantity

$$\begin{aligned} d_n e^{-i(\theta + \frac{2\pi n}{N_s})} &= d_n \cos(\theta + \frac{2\pi n}{N_s}) - i \cdot d_n \sin(\theta + \frac{2\pi n}{N_s}) \\ &\stackrel{(3),(4)}{=} -(x_n - x) + i \cdot (y_n - y) \end{aligned} \quad (5)$$

and, therefore, that

$$d_n e^{-i \frac{2\pi n}{N_s}} = e^{i\theta} (-(x_n - x) + i \cdot (y_n - y)) \quad (6)$$

The end product of summing (6) over  $N_s$  rays is equal to the first term of the Discrete Fourier Transform of the signal  $\{d_n\}$ ,  $n = 0, 1, \dots, N_s - 1$ ,  $F_1$ :

$$\begin{aligned} F_1 &= \sum_{n=0}^{N_s-1} d_n \cdot e^{-i \frac{2\pi n}{N_s}} \\ &\stackrel{(6)}{=} \sum_{n=0}^{N_s-1} e^{i\theta} (-(x_n - x) + i \cdot (y_n - y)) \\ &= e^{i\theta} \sum_{n=0}^{N_s-1} [(x - i \cdot y) + (-x_n + i \cdot y_n)] \\ &= e^{i\theta} N_s (x - i \cdot y) - e^{i\theta} \Delta \end{aligned} \quad (7)$$

where  $\Delta \triangleq \sum_{n=0}^{N_s-1} (x_n - i \cdot y_n)$ .

Denoting with the letter  $R$  quantities which correspond to the real scan  $S_R$ , which has been captured from the sensor pose  $p(x, y, \theta)$ , and with  $V$  those which correspond to the virtual scan  $S_V$ , which has been captured from pose  $\hat{p}(x, y, \hat{\theta})$ :

$$\begin{aligned} R_1 &= \sum_{n=0}^{N_s-1} d_n^R \cdot e^{-i \frac{2\pi n}{N_s}} \\ &\stackrel{(7)}{=} N_s e^{i\theta} (x - i \cdot y) - e^{i\theta} \Delta_R \end{aligned} \quad (8)$$

$$\begin{aligned} V_1 &= \sum_{n=0}^{N_s-1} d_n^V \cdot e^{-i \frac{2\pi n}{N_s}} \\ &\stackrel{(7)}{=} N_s e^{i\hat{\theta}} (x - i \cdot y) - e^{i\hat{\theta}} \Delta_V \end{aligned} \quad (9)$$

Let now  $\Delta_R - \Delta_V = \sum_{n=0}^{N_s-1} (x_n^R - x_n^V) - i \cdot \sum_{n=0}^{N_s-1} (y_n^R - y_n^V) = N_s (\delta_x - i \cdot \delta_y)$ , where

$$\delta_x \triangleq \frac{1}{N_s} \sum_{n=0}^{N_s-1} (x_n^R - x_n^V) \quad (10)$$

$$\delta_y \triangleq \frac{1}{N_s} \sum_{n=0}^{N_s-1} (y_n^R - y_n^V) \quad (11)$$

Then

$$\Delta_V = \Delta_R - N_s (\delta_x - i \cdot \delta_y) \quad (12)$$

The first term of the Discrete Fourier Transform of the signal that consists of the difference of the two signals (8) and (9) is  $X_1$ :

$$\begin{aligned} X_1 &= R_1 - V_1 \\ &= \sum_{n=0}^{N_s-1} (d_n^R - d_n^V) \cdot e^{-i \frac{2\pi n}{N_s}} \\ &\stackrel{(8),(9)}{=} N_s (x - i \cdot y) (e^{i\theta} - e^{i\hat{\theta}}) - e^{i\theta} \Delta_R + e^{i\hat{\theta}} \Delta_V \\ &\stackrel{(12)}{=} N_s (x - i \cdot y) (e^{i\theta} - e^{i\hat{\theta}}) - e^{i\theta} \Delta_R \\ &\quad + e^{i\hat{\theta}} (\Delta_R - N_s (\delta_x - i \cdot \delta_y)) \\ &= N_s (x - i \cdot y) (e^{i\theta} - e^{i\hat{\theta}}) - \Delta_R (e^{i\theta} - e^{i\hat{\theta}}) \\ &\quad - N_s e^{i\hat{\theta}} (\delta_x - i \cdot \delta_y) \\ &= (e^{i\theta} - e^{i\hat{\theta}}) [N_s (x - i \cdot y) - \Delta_R] - N_s e^{i\hat{\theta}} (\delta_x - i \cdot \delta_y) \\ &\stackrel{(8)}{=} (e^{i\theta} - e^{i\hat{\theta}}) \frac{R_1}{e^{i\theta}} - N_s e^{i\hat{\theta}} (\delta_x - i \cdot \delta_y) \\ &= (1 - e^{-i(\theta - \hat{\theta})}) R_1 - N_s e^{i\hat{\theta}} (\delta_x - i \cdot \delta_y) \end{aligned}$$

Therefore, since  $X_1 = R_1 - V_1$ :

$$\begin{aligned} -V_1 &= -e^{-i(\theta - \hat{\theta})} R_1 - N_s e^{i\hat{\theta}} (\delta_x - i \cdot \delta_y) \\ e^{-i(\theta - \hat{\theta})} &= \frac{V_1}{R_1} - \frac{N_s e^{i\hat{\theta}}}{R_1} (\delta_x - i \cdot \delta_y) \\ e^{-i(\theta - \hat{\theta})} &= \frac{|V_1|}{|R_1|} e^{i(\angle V_1 - \angle R_1)} - \frac{e^{i(\hat{\theta} - \angle R_1)}}{|R_1|} (N_s \delta_x - i \cdot N_s \delta_y) \end{aligned}$$

where the polar representation of complex  $A$  is  $A = |A|e^{i\angle A}$ .

Due to the fact that the sensor's orientation  $\theta$  is unknown, so are the endpoints  $\{(x_n^R, y_n^R)\}$ , and therefore quantities  $\delta_x, \delta_y$ . In order to gain an initial intuition as to the magnitudes of the latter we make the observation that, by definition,  $N_s \delta_x$  and  $N_s \delta_y$  quantify the difference of the approximation of line integrals over the closed paths provided by the two scans' endpoints over the two principal axes  $x$  and  $y$ . This approximation is due to the finiteness of  $N_s$ . Therefore, under the assumptions that (a) the map of the environment is its perfect representation and (b) the physical range scan is unaffected by disturbance, as  $N_s \rightarrow \infty$ ,  $N_s \delta_x, N_s \delta_y \rightarrow 0$ , which in turn means that  $|V_1| \rightarrow |R_1|$  and  $\theta - \hat{\theta} \rightarrow \angle R_1 - \angle V_1$ .

Updating the orientation estimate by

$$\hat{\theta}' = \hat{\theta} + \angle R_1 - \angle V_1 \quad (13)$$

results in a residual orientation error  $\phi$ :

$$\phi = \tan^{-1} \frac{N_s \delta_x \tan(\theta - \angle R_1) - N_s \delta_y}{|R_1| + N_s \delta_x + N_s \delta_y \tan(\theta - \angle R_1)} \quad (14)$$

whose magnitude is inversely proportional to the number of rays of the physical range sensor  $N_s$  in the case where both  $S_R$  and  $S_V$  are undisturbed by noise.

The finiteness of the physical sensor's emitted rays, coupled with the arbitrariness of the rate of changes in the environment (Fig. 1), may result in portions of the map being undersampled. What is more is that the number of emitted rays by the physical sensor is immutable. In order to mitigate the effects of these constraints on the orientation error, let  $2^v$  virtual scans of size  $N_s$  be generated at  $\gamma/2^v$  angular increments starting from  $\hat{\theta}$ ,  $v \in \mathbb{N}_{\geq 0}$ , where  $\gamma = 2\pi/N_s$  is the physical sensor's angle increment. Let then the orientation correction process (Eq. (13)) be carried out once between the real scan and the virtual scan  $S_V^k$  captured from orientation  $\hat{\theta}_k = \hat{\theta} + k \cdot \gamma/2^v$ ,  $k = 0, \dots, 2^v - 1$ , for a total of  $2^v$  times, resulting in  $2^v$  orientation estimates  $\hat{\theta}_k'$ . The angular alignment between the virtual scan captured from pose  $(x, y, \hat{\theta}_k')$  and the real scan is captured by the Cumulative Absolute Error per Ray (CAER) metric

$$\text{CAER}_k \triangleq \sum_{n=0}^{N_s-1} \left| S_R[n] - S_V[n] \right|_{(x, y, \hat{\theta}_k')} \quad (15)$$

which is proportional to the degree of misalignment between range scan  $S_R$  and map scan  $S_V$  captured from pose  $(x, y, \hat{\theta}_k')$ , and therefore between  $(x, y, \theta)$  and  $(x, y, \hat{\theta}_k')$ . A profile of the CAER metric is shown in Fig. 3, for the general case of location and orientation incoincidence between the sensor's pose and its estimate.

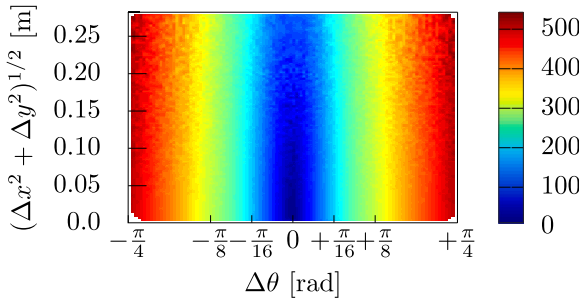
Let now  $k_{\min}$  denote the index of the  $k$ th virtual scan  $S_V^{k_{\min}}$  scoring the minimum  $\text{CAER}_k$ :

$$\text{CAER}_{k_{\min}} = \min\{\text{CAER}_k\}$$

$k = 0, \dots, 2^v - 1$ . Let also  $\hat{\theta}_{k_{\min}}$  denote the angle  $\hat{\theta}_{k_{\min}} = \angle R_1 - V_1^{k_{\min}}$ , where  $V_1^{k_{\min}}$  is the first term of the DFT of  $S_V^{k_{\min}}$  (Eq. (9)). Then, updating the sensor's orientation estimate by  $\hat{\theta}' = \hat{\theta} + \hat{\theta}_{k_{\min}} + k_{\min} \cdot \gamma/2^v$  results in an orientation error whose maximum is equal to that of updating it with (13) for  $v = 0$ .

## 5.2. Location correction

Let now the real and estimated poses be equal in terms of orientation but unequal in terms of position. If the map represents the environment perfectly and the physical range sensor reports faultless measurements then the estimate of the sensor's position can be driven arbitrarily close to its real position. In real conditions, when the rays of either or both real and virtual range sensors are corrupted by bounded additive noise, the position estimate can be made to be bounded in a neighbourhood of the sensor's real position. Theorems I and II formalise these statements [51].



**Fig. 3.** A profile of the CAER metric (Eq. (15)) from  $10^6$  pairs of unperturbed sample scans, depending on the distance  $(\Delta x^2 + \Delta y^2)^{1/2}$  and relative orientation  $\Delta\theta$  of the poses from where a real and a virtual scan were captured. Pose estimates closer to the true pose in terms of orientation (a) exhibit lower CAER values than those further away from it and (b) produce lower position errors once inputted to the Position Correction system.

**Theorem I.** Let the assumptions of Problem P hold. Additionally, let  $\hat{\theta} = \theta$ . Let a map-scan  $S_V$  be captured from  $\hat{p}$  within map  $M$  and be denoted by  $S_V|_{\hat{p}}$ . Assume that both  $S_R$  and  $S_V$  range scans are disturbance-free, that is, the distances to obstacles the rays of the real scan capture correspond to the true distance of the sensor to said obstacles, and that the map of the environment captures the latter perfectly. Then, treating the estimate of the location of the sensor as a state variable  $\hat{l}[k] = [\hat{x}[k], \hat{y}[k]]^T$  and updating it according to the difference equation

$$\hat{l}[k+1] = \hat{l}[k] + u[k] \quad (16)$$

where  $\hat{l}[0] = \hat{l} = [\hat{x}, \hat{y}]^T$ , i.e. the supplied initial location estimate,  $u$  being the two-dimensional vector hereafter referred to as the control vector:

$$u[k] = \frac{1}{N_s} \begin{bmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ \sin \hat{\theta} & -\cos \hat{\theta} \end{bmatrix} \begin{bmatrix} X_{1,r}(S_R, S_V|_{\hat{p}[k]}) \\ X_{1,i}(S_R, S_V|_{\hat{p}[k]}) \end{bmatrix} \quad (17)$$

where  $X_{1,r}(\cdot)$  and  $X_{1,i}(\cdot)$  are, respectively, the real and imaginary parts of the complex quantity  $X_1$ :

$$\begin{aligned} X_1(S_R, S_V|_{\hat{p}[k]}) &= X_{1,r}(S_R, S_V|_{\hat{p}[k]}) \\ &\quad + i \cdot X_{1,i}(S_R, S_V|_{\hat{p}[k]}) \\ &= \sum_{n=0}^{N_s-1} (S_R[n] - S_V[n]|_{\hat{p}[k]}) \cdot e^{-i \frac{2\pi n}{N_s}} \end{aligned} \quad (18)$$

where  $S_R[n]$  and  $S_V[n]|_{\hat{p}[k]}$  are, respectively, the ranges of the  $n$ th ray of the real  $S_R$  and virtual  $S_V|_{\hat{p}[k]}$  scans, and  $\hat{p}[k] = (\hat{l}[k], \hat{\theta})$ —then  $\hat{l}[k]$  converges to  $l$  uniformly asymptotically as  $k \rightarrow \infty$ .

In practice, the control system ((16),(17)) is let to iterate either until the norm of the control vector  $u[k]$  reaches a sufficiently small magnitude  $\|u[k]\|_2 < \epsilon_u$ , where  $\epsilon_u$  is sufficiently small—e.g.  $\epsilon_u < 10^{-3}$ —or for  $I > 0$  iterations (a sufficiently large, externally-supplied maximum iterations threshold—e.g.  $I \geq 20$ ). Therefore, if we denote by  $k_{stop} \in (0, I]$  the last index of iteration, and by  $\hat{l}' = \hat{l}[k_{stop}] \Rightarrow \|e(l, \hat{l}')\|_2 < \|e(l, \hat{l}[0])\|_2$ , and therefore objective (\*) is guaranteed.

**Remark VI.** Without loss of generality, subsequent to the application of Theorem I, the location error is proportional to the orientation error.

**Theorem II.** Let the assumptions of Theorem I hold. Assume additionally that the ranges of both real and virtual range scans  $S_R$  and  $S_V$  are affected by additive, bounded disturbances. Then  $\hat{l}[k]$  is uniformly bounded for  $k \geq k_0$  and uniformly ultimately bounded in a neighbourhood of  $l$ . Its size depends on the suprema of the disturbance corrupting the range measurements of the two scans.

Compared to the case where no disturbances are present, a solution satisfying objective (\*) is not strictly guaranteed for every starting

location  $\hat{l}[0]$ . Let us again denote by  $k_{stop} \in (0, I]$  the last index of iteration, by  $\hat{l}' = \hat{l}[k_{stop}]$  the final estimate of the sensor's location, and by  $B$  the ultimate bound of the pose error. If  $\|e(l, \hat{l}[0])\|_2 > B$ , Theorem II guarantees the satisfaction of objective (\*) if  $k_{stop} \geq k_0$ . If, on the other hand,  $\|e(l, \hat{l}[0])\|_2 \leq B$ , it is not certain that  $\|e(l, \hat{l}')\|_2 < \|e(l, \hat{l}[0])\|_2$ ; what is certain in this case, though, is that  $\|e(l, \hat{l}[k])\|_2 \neq B$  for all  $k \geq 0$ .

### 5.3. Joint correction of orientation and location

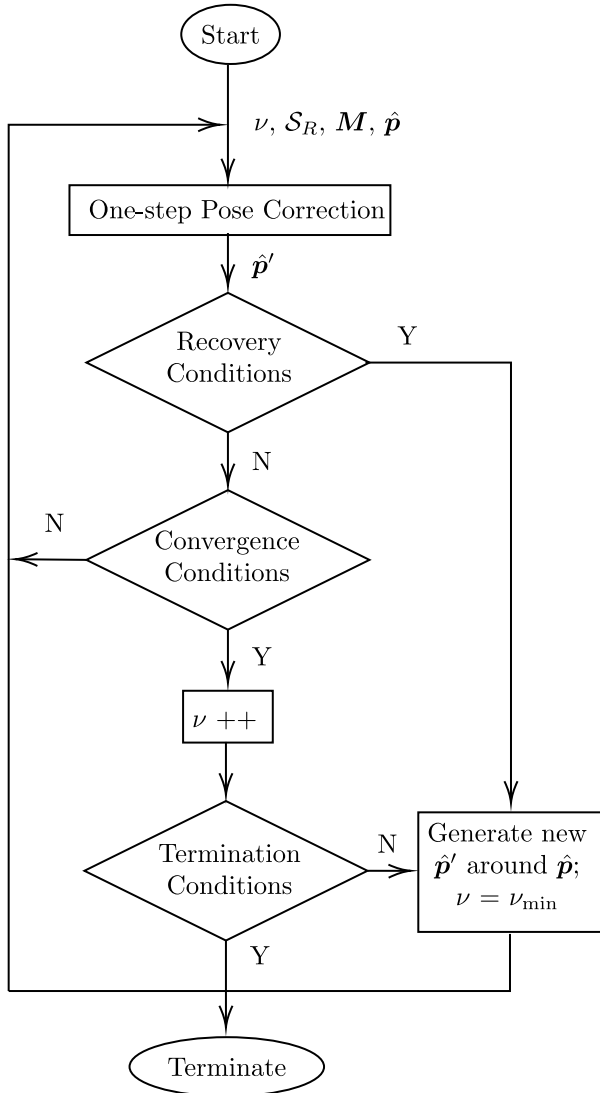
The previous two sections describe two methods of how it is possible to (a) reduce the error of the orientation estimate when the position estimate coincides with the sensor's position, and (b) reduce the error of the position estimate when the orientation estimate equals the sensor's orientation. In the general case, however, no equality stands. What is more is that the problem is coupled: the optimal orientation error cannot be attained in one step when the position error is not zero, and the optimal position error cannot be attained in one step when the orientation error is not zero. Therefore the first goal of a method reducing both would be to first reduce the orientation error and then reduce the location error. The second would be to iterate this process until some termination condition is met. The method proposed by this article is described in the following.

Given an input pose estimate  $\hat{p}(\hat{x}, \hat{y}, \hat{\theta})$ , the real scan  $S_R$ , and the map  $M$ , the pose correction method proposed (Fig. 4) reduces the error of the pose estimate by iteratively invoking the One-step Pose Correction process (Fig. 5) until a set of termination conditions is met. Denoting the former by X1SMSM, X1SMSM starts off with an initial degree of sampling the map  $v = v_{min}$ . The input pose estimate is processed by the One-step Pose Correction process, and its output  $\hat{p}'$  is examined with regard to Recovery and Convergence conditions. If the resulting pose estimate falls outside of the map  $M$  then a new pose estimate is generated from the initially supplied pose estimate, and the process is reset. If no significant pose estimate correction is observed  $\|\hat{p}' - \hat{p}\|_2 < \epsilon_{dp}$ , then the degree of map sampling  $v$  is increased. Its increase serves as a means of reducing the orientation and hence the position estimate error further. Otherwise, the One-step Pose Correction process is reiterated until no significant correction is observed. The process is iterated until a maximum degree of map sampling is reached  $v = v_{max}$ , at which point X1SMSM terminates if a terminal condition is met. This terminal condition facilitates the avoidance of local maxima. In the case where this condition is not met, a new pose is generated from the initially supplied pose estimate, and the process is reset.

Given an input pose estimate  $\hat{p}(\hat{x}, \hat{y}, \hat{\theta})$ , the real scan  $S_R$ , the map  $M$ , and a sampling degree  $v$ , the One-step Pose Correction system first calculates  $2^v$  pose estimates  $\hat{P}_{OC} = \{(\hat{x}, \hat{y}, \hat{\theta}_k)\}$ ,  $k = 0, \dots, 2^v - 1$  (Section 5.1). Its operation is denoted in Fig. 5 by the operator OC( $\cdot$ ). The location error of each pose estimate in set  $\hat{P}_{OC}$  is then attempted to be reduced by rehearsing the subsequent Position Correction operation for one iteration. This operation produces pose estimate set  $\hat{P}_{RPC}$ . Subsequently, the CAER metric of each pose in  $\hat{P}_{RPC}$  is calculated.

The pose estimate  $\hat{p} \in \hat{P}_{RPC}$  that records the minimum CAER among all pose estimates in  $\hat{P}_{RPC}$  is in principle the pose from which the computed virtual scan records the least deviation from the real scan among all poses in  $\hat{P}_{RPC}$ . By capturing the CAER for each displaced pose estimate in  $\hat{P}_{RPC}$  it is possible to establish a less erroneous pose error rank between orientation estimates in  $\hat{P}_{OC}$  than if position correction was not rehearsed, and simultaneously retain only one pose estimate for the subsequent step of Position Correction proper.<sup>2</sup> The pose estimate in  $\hat{P}_{RPC}$  that records the lowest CAER value is then inputted to the Position Correction subsystem, for a limited number

<sup>2</sup> Alternatively, correcting the position of  $2^v$  pose estimates and feeding them back to the One-step Pose Correction method would incur exponential costs in time of execution.



**Fig. 4.** The flow diagram of X1SMSM. Execution commences with an initial angular sampling degree  $\nu_{\min}$ , the scan captured by the physical range sensor  $S_R$ , and the map of the environment  $M$ . The initial pose estimate is provided by a tracking filter during pose tracking or in the form of a hypothesis during global localisation. The inner method One-step Pose Correction (Fig. 5) is called iteratively, updating the pose estimate until a maximum of angular sampling degree is reached.

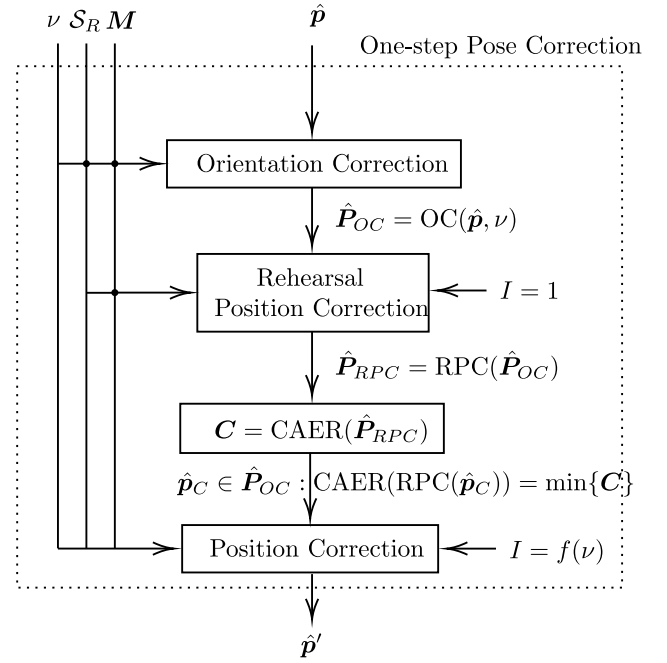
of iterations  $I$ . The output of the One-step Pose Correction system is set to its resulting output, denoted by  $\hat{p}'$ . In practice, the pose set  $\hat{P}_{OC}$  is supplemented with the pose that produces the minimum CAER over time. This addition introduces a form of memory to the system, which assists it in avoiding divergence and which, therefore, benefits speed of execution.

## 6. Experiments

This section serves to test the efficacy and performance of the proposed method, termed X1SMSM, against those of state-of-the-art methods utilisable in the scan-to-map-scan matching task.

### 6.1. Experimental procedure

The experimental procedure was conducted using five established and publicly available benchmark datasets provided courtesy of the



**Fig. 5.** The block diagram of the core pose alignment method of X1SMSM, termed One-step Pose Correction.

**Table 1**

The size of each dataset used during evaluation of the performance of X1SMSM, FastGICP, CSM, NDT, NDT-PSO, FastVGICP, and TEASER methods in the scan-to-map-scan matching task.

Dataset	# instances
aces	7373
fr079	4933
intel	13630
mit_csail	1987
mit_killian	17479

Department of Computer Science, University of Freiburg.<sup>3</sup> Each dataset comprises a collection of range scanner measurements and the pose  $r(x, y, \theta)$  from which these measurements were made. The signifier and size of each dataset used for the experimental procedure is shown in Table 1.

For purposes of comparison against scan-matching methods that may be utilised in scan-to-map-scan matching, the experimental procedure is extended to the Normal Distributions Transform (NDT) scan-matching method [35,52], FastGICP [37,53] and PLICP-GPM [17,54]. The latter shall be denoted hereafter by the acronym CSM. GPM was used initially in order to overcome the angular realignment problems [17] of PLICP. NDT, FastGICP, and CSM belong to the *established* state-of-the-art methods of scan-matching [23,55–59]. In addition, for comparison against *contemporary* state-of-the-art algorithms, the experimental procedure is extended to FastVGICP [23,60], NDT-PSO [36,61], and TEASER [38,62].

The experimental setup is the following. The rays of each dataset instance  $D_k^d$ ,  $k \in \{0, 1, \dots, 4\}$ ,  $d \in \{0, 1, \dots, |D_k|\}$  are first projected to the  $x$ - $y$  plane around  $r_k^d$ . The datasets' scans are not panoramic, therefore the remaining space is filled with a semicircular arc that joins the scan's two extreme ends. Its radius is set to the minimum range between the two extreme rays of  $D_k^d$ . Similar fashions for closing-off the environment have been found equivalent with respect to the

<sup>3</sup> The datasets are available at <http://ais.informatik.uni-freiburg.de/slamevaluation/datasets.php>; last accessed 25 Oct 2021



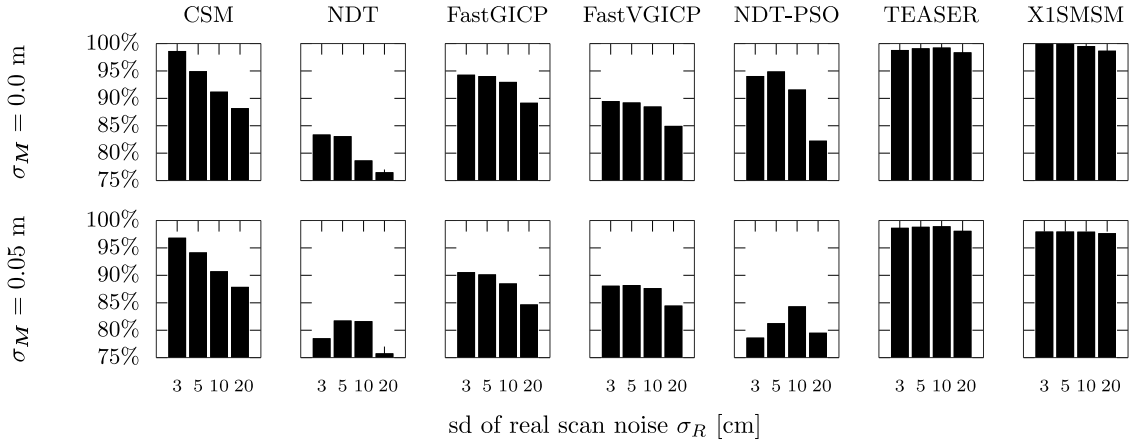


Fig. 6. The relative proportion of cases where objective (\*) was attained by each tested method as a function of sensor noise  $\sigma_R$  and map distortion  $\sigma_M$  levels over all conducted experiments.

performance of the tested methods. The resulting point-set is regarded as the environment  $\mathbf{W}_k^d$  in which the range sensor operates (e.g. the environment of Fig. 1). Then the map of the environment  $\mathbf{M}_k^d$  is set to be  $\mathbf{W}_k^d$ . In order to induce distortions in the map, each coordinate of all points in  $\mathbf{M}_k^d$  is perturbed by errors extracted from a normal distribution  $\mathcal{N}_{\mathbf{M}} \sim (0, \sigma_{\mathbf{M}}^2)$ . What is considered the sensor's actual pose  $\mathbf{p}_k^d$  is generated randomly within the polygon formed by  $\mathbf{W}_k^d$ . The range scan  $S_{R,k}^d$  that is considered to be reported by the physical sensor is then computed by locating the intersection points between  $N_s$  rays emanating from  $\mathbf{p}_k^d$  and the polygon formed by  $\mathbf{W}_k^d$  across an angular field of view  $\lambda = 2\pi$ . The initial pose estimate of the sensor  $\hat{\mathbf{p}}_k^d$  is then obtained by perturbing the components of  $\mathbf{p}_k^d$  with quantities extracted from uniformly distributed error distributions  $U_{xy}(-\bar{\delta}_{xy}, \bar{\delta}_{xy})$ ,  $U_{\theta}(-\bar{\delta}_{\theta}, \bar{\delta}_{\theta})$ ;  $\bar{\delta}_{xy}, \bar{\delta}_{\theta} \in \mathbb{R}_{\geq 0}$ .

In order to test the performance of the above methods four levels of noise acting on the range measurements of the real scan  $S_R^d$  are tested. The range measurements are perturbed by zero-mean normally-distributed noise with standard deviation  $\sigma_R \in \{0.03, 0.05, 0.10, 0.20\}$  m. The values of tested standard deviations were calculated from commercially available panoramic LIDAR scanners by identifying the magnitude of their reported maximum range errors and dividing it by a factor of three. The rationale is that 99.73% of errors are located within  $3\sigma$  around the actual range between a ray and an obstacle, assuming errors are distributed normally. These are reported for price-appealing but disturbance-laden panoramic sensors, e.g. the RPLIDAR A2M8, or the YDLIDAR G4, G6, TG30, and X4 scanners [63–67]. In addition, two levels of map distortion are tested:  $\sigma_M \in \{0.0, 0.05\}$  m. Maximal displacements  $\bar{\delta}_{xy}$  and  $\bar{\delta}_{\theta}$  are set to  $\bar{\delta}_{xy} = 0.20$  m and  $\bar{\delta}_{\theta} = \pi/4$  rad. The value of  $\bar{\delta}_{xy}$  was chosen as such from reports on positional errors in real conditions [44]. The value of  $\bar{\delta}_{\theta}$  was chosen as such in order to include orientation errors at the initialisation stage of pose tracking and errors induced due to diverging odometry readings. The size of the input real scan was set to  $N_s = 360$  rays. The minimum and maximum oversampling rates of X1SMSM were set to  $(\mu_{\min}, \mu_{\max}) = (2^{\nu_{\min}}, 2^{\nu_{\max}}) = (2^2, 2^4)$ . The number of iterations of the translational component were set to  $I = 2$  and  $\epsilon_{\delta p} = 10^{-5}$  (Section 5.3). X1SMSM's termination condition was set to  $\text{CAER}(\hat{\mathbf{p}}') \leq (\hat{\sigma}_R + \hat{\sigma}_V)^{1/2}$ , where  $\hat{\sigma}_R$  and  $\hat{\sigma}_V$  are estimates of the standard deviation of noise affecting the rays of  $S_R$  and  $S_V$  respectively.

For each experiment X1SMSM, CSM, NDT, FastGICP, and FastVGICP ran for  $E = 10$  times across all instances of  $D_k$ ,  $D = \{\text{aces}, \text{fr079}, \text{intel}, \text{mit\_csail}, \text{mit\_killian}\}$ ,  $k \in \{0, 1, \dots, 4\}$ . Therefore each method was tested a total of  $N_{\text{tot}} = 10 \times 2 \times 4 \times \sum |D_k| \approx 3.6 \cdot 10^6$  times. The execution times of NDT-PSO and TEASER in the scan-to-map-scan matching problem were measured in the order of seconds per pose input—approximately one order of magnitude larger than the

execution time of X1SMSM: the evaluation of NDT-PSO and TEASER was performed once for every instance of  $D$ .

Experiments with X1SMSM, FastGICP, FastVGICP, CSM, and NDT were carried out on a single thread, on a machine of a CPU frequency of 4.0 GHz. NDT-PSO and TEASER are parallel implementations; their experiments ran over four threads, with a machine of a CPU frequency of 2.2 GHz.

The criterion on which the evaluation of all tests rests is the 2-norm of the total pose error—Eq. (1) for  $\hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}}'$ , where  $\hat{\mathbf{p}}'$  is the output of each algorithm tested. For every pose estimate  $\hat{\mathbf{p}}_k^{d'}$  outputted by each algorithm,  $d = 1, 2, \dots, |D_k|$ ,  $k \in \{0, 1, \dots, 4\}$ , its offset from the actual pose  $\mathbf{p}_k^d$  is recorded in the form of the 2-norm total error. The unit of measurement of the total pose error is  $(\text{m}^2 + \text{rad}^2)^{1/2}$  and, where omitted in the figures of the following subsections, it has been so for purposes of economy of space and readability.

## 6.2. Results

Fig. 6 illustrates the percentage of cases where objective (\*) was attained by all tested algorithms per standard deviation of measurement noise and map corruption levels, over all conducted experiments. The performance of all algorithms was approximately invariant over any dataset: therefore their results are aggregated over all datasets in the figures of this section.

Fig. 7 illustrates the distribution each tested method's pose errors across all conducted experiments per sensor noise  $\sigma_R$  and map distortion  $\sigma_M$  levels. Fig. 8 illustrates the corresponding execution times.

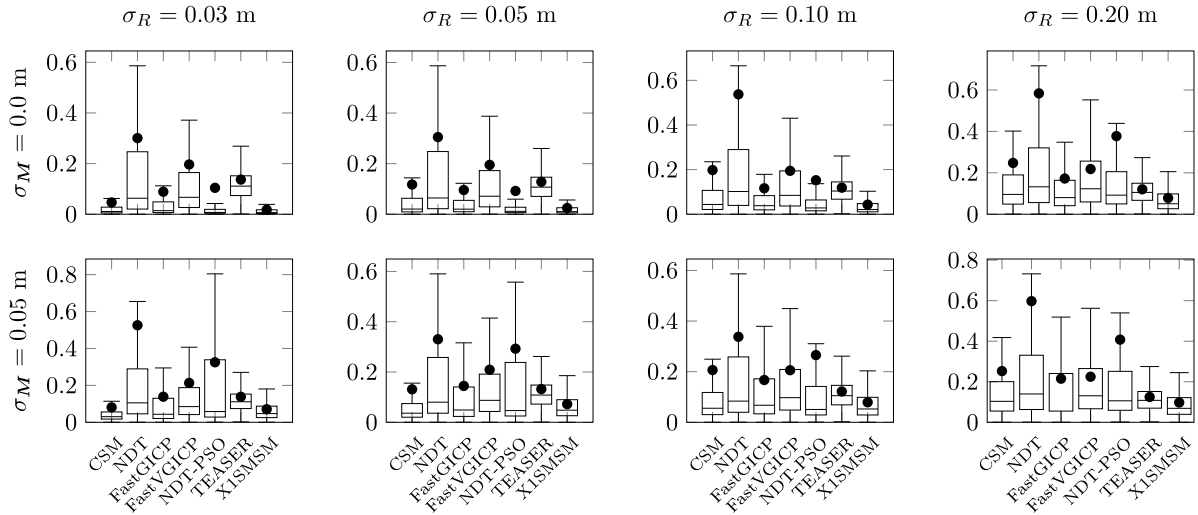
Fig. 9 shows a breakdown of X1SMSM's mean execution time for each tested configuration.

Fig. 10 provides a juxtaposition of the evolution of mean position and orientation errors of the three methods with the highest proportion of objective (\*) attainment cases for increasing real scan noise  $\sigma_R$  per map distortion level  $\sigma_M$  tested.

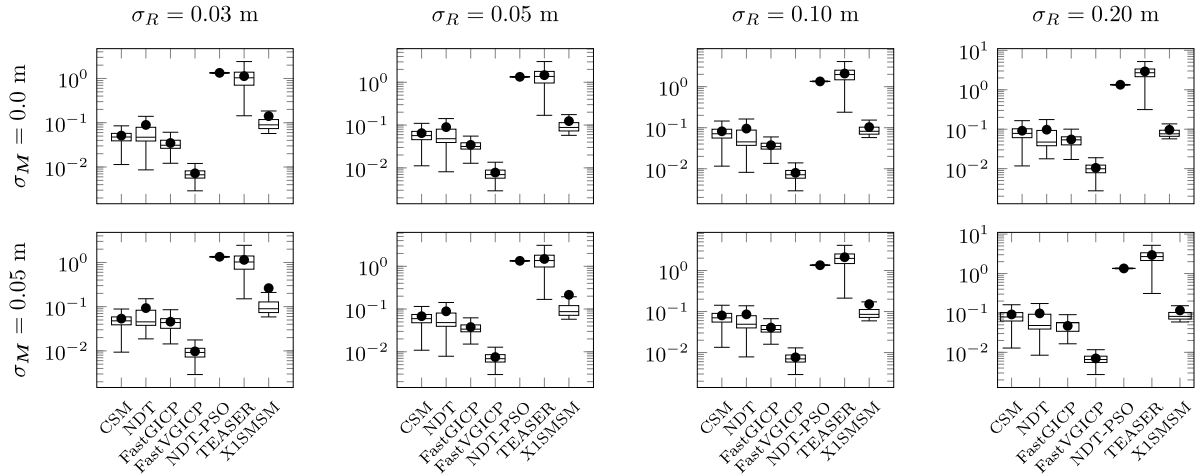
Fig. 11 summarises the mean core execution time (total time minus map-representation and implementation-specific intersection-finding time), and the mean total number of virtual scans captured by X1SMSM for  $\sigma_M \in \{0.0, 0.05\}$  m, over  $N_{\text{tot}}$  tests, as a function of the standard deviation of noise affecting the rays of the real scan  $\sigma_R$ .

## 6.3. Discussion

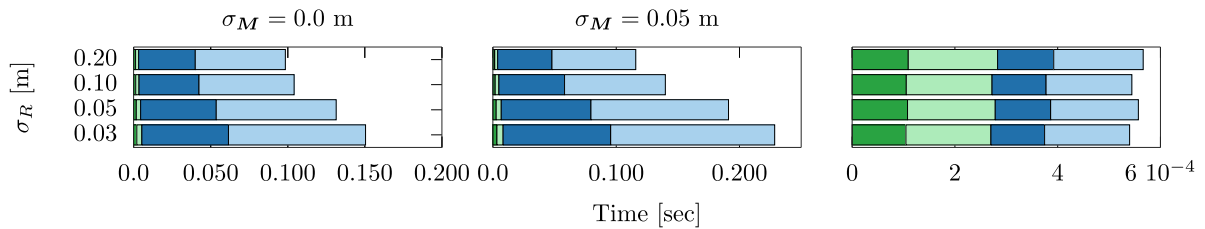
The proposed method's pose error reduction rate is largely invariant across real scan noise levels for a given level of map distortion. The method manages to improve the pose estimate of at least 97.5% of all input pose estimates with regard to the tested configurations (Fig. 6).



**Fig. 7.** Distribution of pose errors of state of the art scan-matching methods and of X1SMSM in the scan-to-map-scan matching task, for maximal uniform position displacements  $\delta_{xy} \in U_{xy}(-0.20, +0.20)$  m and maximal uniform orientation displacements  $\delta_\theta \in U_\theta(-\pi/4, +\pi/4)$  rad for  $\sigma_M = 0.0$  m (top) and  $\sigma_M = 0.05$  m (bottom) over all conducted experiments, per sensor noise level  $\sigma_R$  tested. Dots encode the mean pose error for each method and configuration. Unit of measurement is  $(\text{m}^2 + \text{rad}^2)^{1/2}$ .



**Fig. 8.** Distribution of execution times of state of the art scan-matching methods and of X1SMSM in the scan-to-map-scan matching task, for maximal uniform position displacements  $\delta_{xy} \in U_{xy}(-0.20, +0.20)$  m and maximal uniform orientation displacements  $\delta_\theta \in U_\theta(-\pi/4, +\pi/4)$  rad for  $\sigma_M = 0.0$  m (top) and  $\sigma_M = 0.05$  m (bottom) over all conducted experiments, per sensor noise level  $\sigma_R$  tested. Dots encode the mean execution time for each method and configuration. Unit of measurement is seconds.



**Fig. 9.** Breakdown of X1SMSM's execution time per each tested configuration. Rotation times are signified in green and translation times in blue. Light colours signify the time consumed in computing virtual scans and dark colours the core execution time of each component. The third column illustrates the timing breakdown for one iteration over each component. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

X1SMSM's pose error reduction rate is on par with that of TEASER but, in contrast, X1SMSM runs in real time, requires fewer computing resources, and exhibits lower pose errors. According to the results the performance of ICP-based methods in terms of the proportion of cases where the pose estimate error was reduced deteriorates as real scan noise increases. FastVGICP is the most consistent among ICP variants with regard to the proportion of cases where pose errors were reduced. NDT-PSO manages to increase the pose error reduction rate of NDT

but at a cost of significant increases in execution time and processing resources required.

In terms of pose estimate errors (Fig. 7) the behaviour of X1SMSM is more accurate than those of all tested methods across all configurations and datasets tested; only CSM's pose errors are comparable to those of X1SMSM, and only in low levels of measurement noise and when the map is disturbed. Additionally, the rate of increase of the interquartile range for X1SMSM's pose errors is consistently lower than that of all

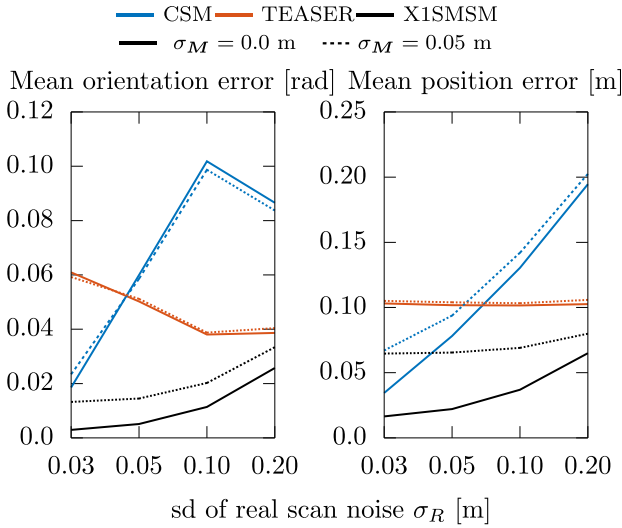


Fig. 10. Mean position and orientation error of the three methods with the highest proportion of objective (\*) attainment cases across all conducted tests for increasing levels of real scan noise  $\sigma_R$  per map distortion level  $\sigma_M$ .

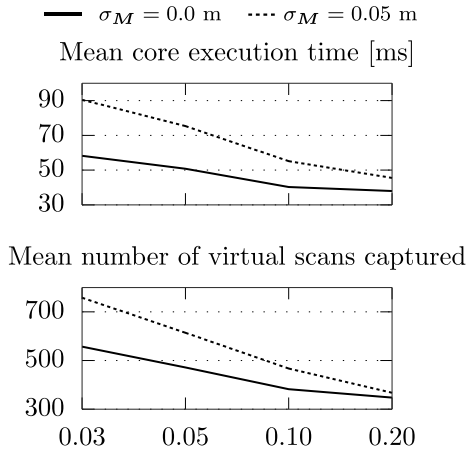


Fig. 11. Core execution time and total number of virtual scans captured by X1SMSM per standard deviation of noise affecting the rays of the physical sensor  $\sigma_R$  and standard deviation of noise affecting the maps' coordinates  $\sigma_M$  over all conducted experiments.

tested methods. By contrast, the behaviour of CSM, GICP, VGICP is less robust to measurement noise, and especially so at the upper range of the spectrum of measurement noise exhibited by available panoramic LIDAR sensors. It is presumed that the widening gap in performance between these methods and X1SMSM, in terms of increasing measurement noise, is a consequence of their modus operandi of establishing correspondences between a point and a (line)point in its two input scans [24]. This conjecture is supported by the fact that the more laden a scan is with noise, the more difficult it is for the algorithm to distinguish true correspondences from false. By contrast, X1SMSM does not deal in correspondences and, ipso facto, does not require the manual setting of parameters relating to establishing correspondences.

Focusing on the mean position and orientation errors of CSM, TEASER, and X1SMSM (Fig. 10), and according to the evidence, the orientation and position errors of CSM increase at a greater rate than those of X1SMSM for a given level or map distortion, while starting off at higher magnitudes. Interestingly, when the map is distorted, the position errors of TEASER are invariant to the noise affecting the ranges of the real scan.

X1SMSM's lowest processing speed was approximately 225 ms per pose input. In comparison, CSM's execution times ranged from 34 to

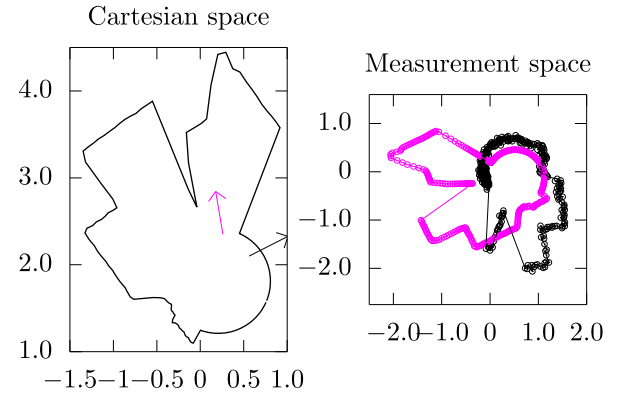


Fig. 12. In principle a typical localisation scenario. Left: the pose of the LIDAR sensor, denoted with black colour, and its estimate with magenta. Right: the real and virtual scans captured from their respective poses in the local coordinate system of each sensor. In this example  $\sigma_R = 0.05$  m and  $\sigma_M = 0.0$  m. Note how at this initial configuration there are portions of the environment's map visible only to the pose estimate, but the corresponding portions of the environment are not visible to the sensor itself. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

94 ms, NDT's 85–111 ms, FastGICP's 25–59 ms, and FastVGICP's were approximately constant at 7 ms. NDT-PSO's execution times were approximately constant as well at 1.34 s, while TEASER had the greatest variability and execution time at 1.3 – 3.3 seconds. This is due to the fact that TEASER requires the measurement and virtual scan vectors to be homologous, i.e. a reference and target point are required to be in the same position in both vectors.<sup>4</sup> The applicability of NDT-PSO and TEASER is hindered both by their execution times and the fact that their implementation requires multiple processor cores ( $> 4$ ) in order to run in real time—which may be unavailable in multi-subsystem systems with limited resources. In any case all execution times should be taken with a grain of salt, as they are specific to the map representation used, the number of rays of the input range scan, and the processing power available.

Finally, the proposed method's increasing iterations with respect to decreasing measurement noise (Fig. 11) suggests that the termination criterion used during the experimental procedure is unevenly strict across measurement noise levels.

## 7. Characterisation and limitations

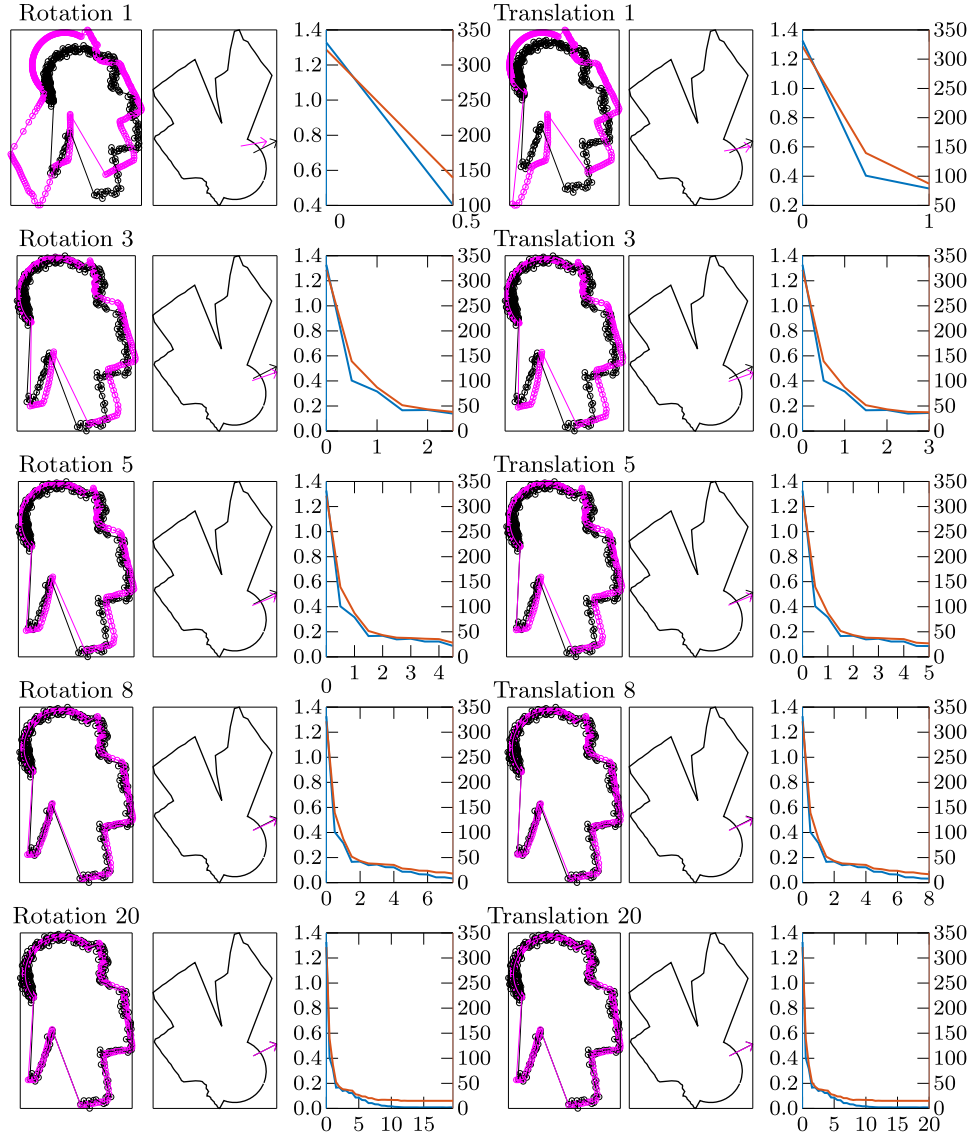
In this section we give key characterisations of X1SMSM and compare aspects of its performance against representative real-time scan-matching methods.

### 7.1. Characterisation

Fig. 12 demonstrates in principle a scenario where scan-to-map-scan matching may be applicable for pose estimate error reduction. In the left-hand side figure the robot's real pose is denoted with black, and its initial estimate with magenta. The right-hand side figure shows the real and virtual scans captured from sensor's respective poses in the local coordinate system of each sensor. Fig. 13 illustrates the interim and final outputs of the alignment process carried out by X1SMSM.

The first row of Fig. 14 illustrates the dependence of the output position errors on the initial position displacement (left) and the initial orientation displacement (right) across all experiments conducted (Section 6) for  $\sigma_R = 0.03$  m and  $\sigma_M = 0.0$  m. The second row

<sup>4</sup> <https://github.com/MIT-SPARK/TEASER-plusplus/issues/24#issuecomment-622652266>



**Fig. 13.** The inner outputs of the alignment process of X1SMSM. The first and fourth columns show the outputs of the orientation and location subsystems at each iteration respectively. The second and fourth columns show the respective subsequent configurations in the map's frame of reference. The third and sixth columns show the corresponding pose estimate error with blue colour and the value of the CAER metric with red. Notice how the virtual scan transforms at each iteration to increasingly resemble the real scan as the pose error is progressively reduced. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

illustrates the dependence of the output orientation errors with respect to the same initial configurations. According to the evidence the output position error is dependent on the initial location displacement between the real pose and its estimate, but independent of the initial orientation displacement between them, with regard to the tested displacement configurations. The output orientation error on the other hand is independent of both initial location and orientation displacements.

Figs. 15 and 16 depict the mean orientation errors of the orientation correction subsystem of X1SMSM, along with those of CSM and NDT, for varying levels of maximal initial orientation displacements  $\bar{\delta}_\theta$ , range sensor emitted rays  $N_s$ , and range sensor noise level  $\sigma_r$ , for two cases of map-corruption levels  $\sigma_M$ , for one iteration. Each method was tested 100 times over the 778 instances of the *laserazos* dataset.<sup>5</sup> These

results are captured in the case where the position of the sensor and its estimate are coinciding so that the effect of the methods' orientation correction subsystem is revealed.

The evidence shows that in complete coincidence of the map to its referent environment the orientation errors of X1SMSM decrease with increasing number of rays for a given level of sensor noise, regardless of the value of initial orientation displacement between the sensor's true orientation and its estimate. In contrast, the orientation errors of CSM and NDT (a) do not progressively decrease proportionally to the size of scan rays, and (b) are not invariant to increasing initial orientation displacement. Predictably, the orientation error of X1SMSM increases for increasing sensor noise for a given number of rays emitted. When the map is corrupted these correlations cease to exist for X1SMSM and, therefore, since (a) the accuracy of the position correction subsystem depends on the orientation error and (b) the processing time is proportional to the range sensor's size, from a computational-resources perspective it is more efficient for an input real scan whose size is greater than 360 rays to be downsampled to this size prior to matching.

<sup>5</sup> The dataset is available at <https://censi.science/pub/research/2007-plipclp/laserazosSM3.log.gz>



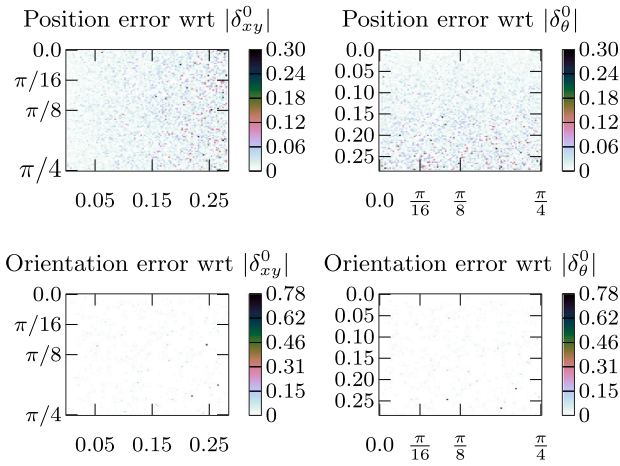


Fig. 14. The dependence of the output position (top) and orientation (bottom) errors on initial location (left) and orientation (right) displacement. Unit of measurement is meters and rad respectively.

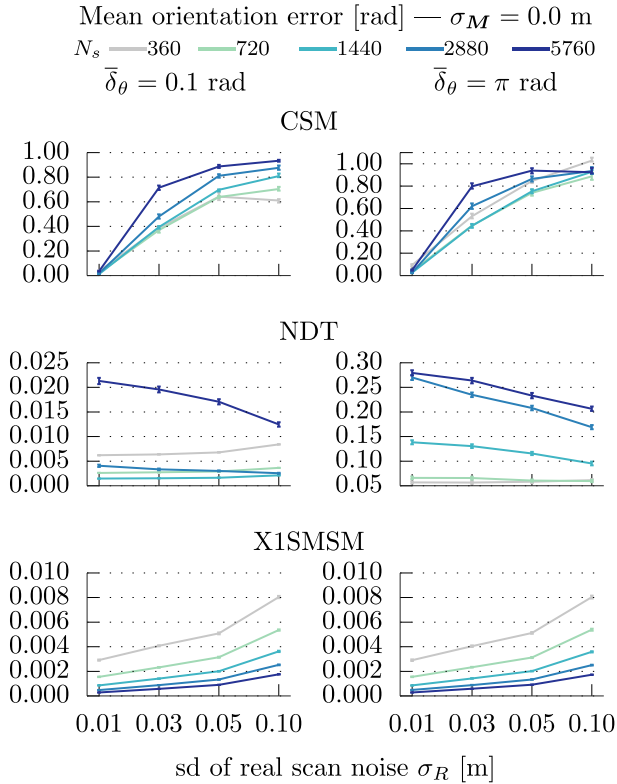


Fig. 15. Mean orientation errors of CSM, NDT, and the orientation correction subsystem of X1SMSM after one iteration for coinciding positions of the range sensor's pose and its estimate per two different maximal initial orientation displacements  $\bar{\delta}_\theta$ , varying number of range sensor emitted rays  $N_s$  and noise  $\sigma_R$  when  $\sigma_M = 0.0$  m.

## 7.2. Limitations

From the evidence of Section 6 X1SMSM is capable of addressing initial position errors ranging to 0.20 m per component, with its orientation correction subsystem being independent of initial location errors (Fig. 14). However, the orientation correction subsystem may fail to estimate the sensor's real orientation at large initial position errors. In Fig. 17 the pose estimate is displaced by 1.0 m in the x-wise direction; the first orientation correction iteration misplaces the pose's orientation, on whose accuracy the location correction depends, which

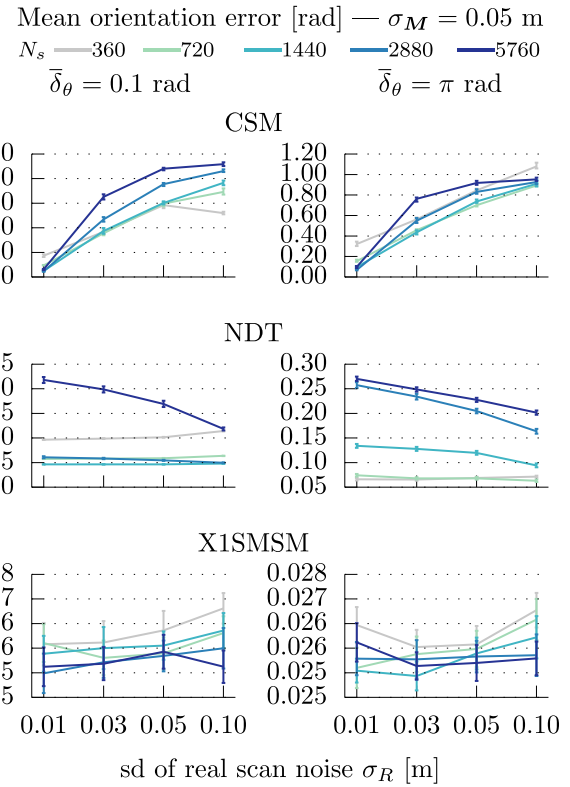


Fig. 16. Mean orientation errors of CSM, NDT, and the orientation correction subsystem of the proposed method after one iteration for coinciding positions of the range sensor's pose and its estimate per two different maximal initial orientation displacements  $\bar{\delta}_\theta$ , varying number of range sensor emitted rays  $N_s$  and noise  $\sigma_R$  when  $\sigma_M = 0.05$  m.

further misplaces the estimate in terms of location. Although in this example we have deactivated the recovery module, recovering from the initial pose may finally yield an accurate pose estimate but this would have happened at the expense of execution time. This situation is unrealistic in pose tracking, but highlights a key limitation of X1SMSM.

The greatest challenge for the performance of a scan(-to-map-scan) matching method, aside from corrupted input data, is missing data. Missing range information in range scan sensors manifests due to their innate constraint of inability to detect ranges to objects that lie beyond their maximum detectable range radius. The absence of usable data is a function of a range sensor's maximum range and the geometry of its surrounding environment at any given sensing instant. In principle, correspondence-finding methods ought to fare better in the face of missing ranges than correspondenceless methods due to their ability to establish true and discard false correspondences between separate segments in their input scans. In order to guard against missing range data and facilitate matching in their presence, X1SMSM duplicates a sensor's physical constraint of an upper range threshold in its generation of a virtual scan so that the latter is assisted in resembling its target real scan as much as achievable. We illustrate the limitations of the proposed method in the face of missing range measurements, and the difference in performance between it and established state-of-the-art correspondence-finding and correspondenceless methods in Fig. 18.

At the top row of the figure four distinct environments are depicted in white colour. The radius of concentric circles around the sensor's true position (depicted in blue) equals the set maximum range for the sensor. The colour of their perimeter signifies the proportion of real scan ranges which are within that maximum range radius according to the colourbar at the second row. The two latter rows of figures illustrate the mean pose error of CSM (red), NDT (blue), and X1SMSM (green), over

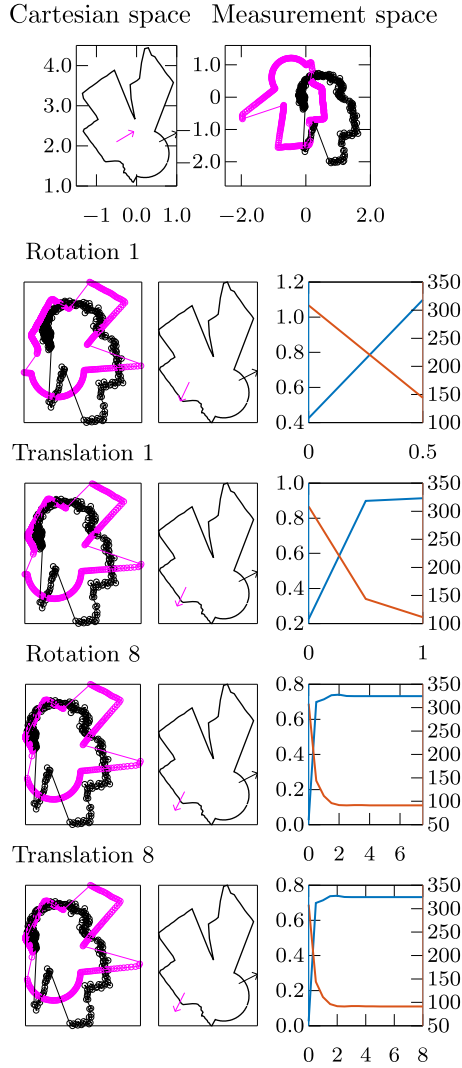


Fig. 17. An exemplary condition where X1SMSM fails to converge: the initial position estimate is displaced far from the sensor's real position. The orientation correction subsystem outputs an inaccurate orientation estimate and ipso facto the estimate's position diverges from its target. The third column shows the evolution of the pose error in blue colour and the value of the CAER metric in red for the first eight iterations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ten iterations at the same maximum range level for two levels of sensor noise, when the map of the environment is not corrupted with noise.

At low sensor noise levels, ICP-based methods seem to dominate their counterparts due to their aforementioned merits. However, this characteristic is reversed as measurement noise increases. CSM's, NDT's and the proposed method's performance deteriorates at irregular rates and according to the particular characteristics of the sensor's surroundings. CSM records the highest pose errors and the lowest robustness to maximum range reduction overall. Compared to NDT, X1SMSM exhibits greater accuracy at the lower and higher ends of the missing ranges scale. Qualitatively, X1SMSM records its lowest performance when it completely loses its footing over large areas at two opposite directions: the figures of the third column summarise this limitation of X1SMSM.

## 8. Applications

Scan-to-map-scan matching in two dimensions may be employed in various contexts. The most usual application is in pose-tracking, where

the output of the tracking filter is fed to a scan-to-map-scan method in order for it to provide a pose estimate of lower error [42] (Section 4). In RFID localisation, for example, one seeks to accurately localise the placement of tags in 3D space by relying on the accuracy of the pose of a robot's antennas, whose pose in space is provided by the LIDAR sensor equipped to the robot. In this case position accuracy has more gravity than orientation accuracy.

Scan-to-map-scan matching may also be used in shape-matching, i.e. estimating the transformation between two similar shapes or detecting those shapes from a collection that match a reference one. In the first, a source pose would need to be computed, preferably via the shapes' centroid, as the latter is located at the same point relative to all shape-constituting points. Then from the source pose of the reference shape a virtual scan would be captured, which in scan-to-map-scan matching terms would be considered the real scan. Capturing virtual scans initially from the source pose of the second shape and matching them with the real scan would then provide the rotation and translation transformation between the two. Detecting shapes that record similarity to a reference one would involve the same process, but with a limited number of recoveries so that the rejection of false samples be carried out efficiently. The experimental procedure of Section 6 constitutes a direct test on these applications.

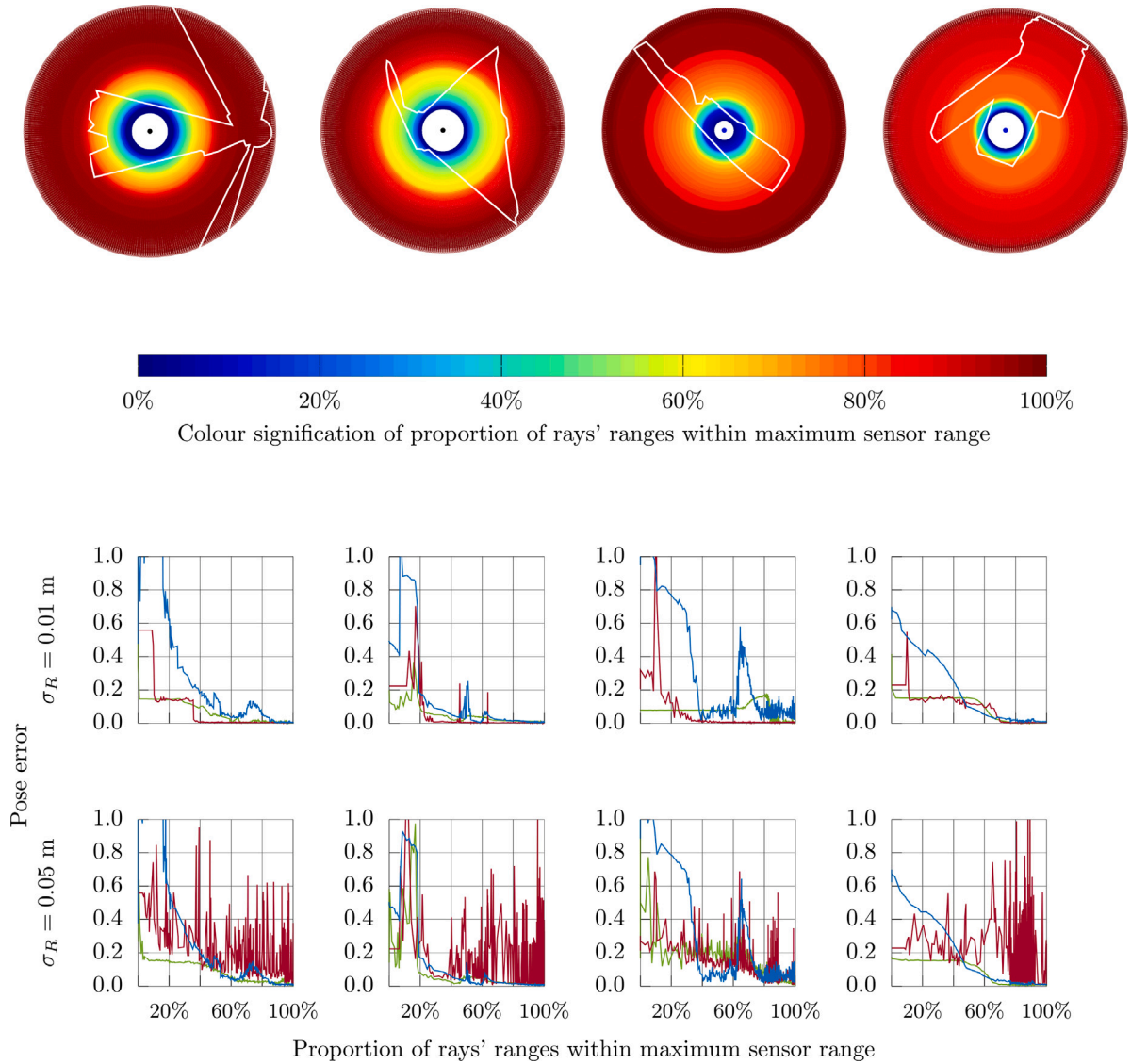
Recently we have also used scan-to-map-scan matching for the solution to the global localisation problem [24]. The introduction of the CAER metric (Fig. 3) may be used to accelerate the solution to the problem in the following manner: As standard a dense cloud of hypotheses is generated in the unoccupied space of the map. Then the CAER metric is computed for each pose hypothesis. The hypotheses with the lowest CAER values would then be handed over to X1SMSM for registration. The one whose final CAER value is the minimum would then be considered the robot's pose. Fig. 19 shows a sample global localisation scenario.

## 9. Future steps

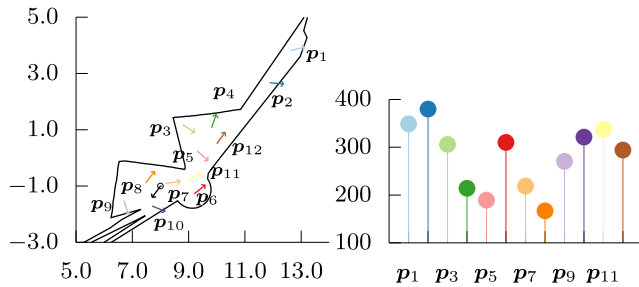
The limitations of X1SMSM dictate the course of future work. With regard to missing range measurements due to inadequacy of a sensor's maximum range, work focus is needed on devising mechanisms or methods whose pose error is more consistent with the case where there are none, and across a greater range of proportion of missing measurements. Furthermore, we conjecture that the estimation of discrepancies  $\delta_x, \delta_y$  (Eqs. (10), (11)) will result in lower orientation estimate errors (Eq. (14)) and therefore lower subsequent location estimate errors, and translate X1SMSM's ability in reducing its orientation error for increasing range scan size when the map is not corrupted to the case where it is. This will benefit speed of execution and will also aim at additional pose error reductions.

## 10. Conclusions

This article introduced a method for reducing the pose estimate error for robots equipped with LIDAR range sensors whose field of view is  $360^\circ$  in a scan-to-map-scan matching manner. Contrary to state-of-the-art approaches, the solution it provides to the problem of scan-to-map-scan matching does not require the establishing of correspondences between its two input scans, and is in closed form. These facts account for (a) the method's robustness against the typical measurement noise levels exhibited by low-cost panoramic 2D LIDAR sensors, whose measurement noise is significantly elevated compared to prior sensors, (b) its lower pose errors compared to prior art, and (c) its real-time execution. The proposed method assumes that a panoramic range scan, the map of the environment in which the robot operates, and a pose estimate residing in the vicinity of the robot's true pose are available. Therefore the proposed method's placement within a localisation system may be at the end of or at the same level as a pose tracking method, or at the heart of global localisation.



**Fig. 18.** Limitations in the performance of correspondence-finding and correspondenceless methods for scan-to-map-scan matching when the range scan sensor's range is progressively restricted, for different sensor noise levels and distinctive environments. In the figures of the last two rows CSM is denoted with red, NDT with blue, and X1SMSM in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 19.** A typical global localisation scenario solvable with the use of scan-to-map-scan matching. The robot's true pose is denoted with black colour. Pose hypotheses are dispersed in the unoccupied interior space of the map. Note how symmetries in the environment make the CAER of  $p_5$  lower than that of  $p_7$ , which is actually closer to the real pose than  $p_5$ . With enough pose hypotheses the global localisation problem can be solved in reduced time due to the use of the CAER metric before handing a hypothesis over to the matching algorithm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

After computing a virtual range scan from the measurement sensor's pose estimate, the method updates the estimate by reducing first the orientation estimate's error and then that of the position estimate. The process is iterated until sufficient convergence conditions are met. The correction of the orientation and location estimates is performed by utilising the first coefficient of the Fourier Transform of the difference between the two scans. The proposed method exhibits higher accuracy, consistent and greater pose error reduction rates compared to established and contemporary state-of-the-art methods that may be utilised in real time in the scan-to-map-scan task. The proposed method's implementation is available for download at <https://github.com/li9i/x1smsm-public>.

#### CRediT authorship contribution statement

**Alexandros Filotheou:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualisation. **Andreas L. Symeonidis:** Methodology, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition, Writing – review & editing. **Georgios D. Sergiadis:** Methodology, Writing – original draft, Supervision.

**Antonis G. Dimitriou:** Conceptualization, Resources, Validation, Supervision, Project administration, Funding acquisition, Writing – review & editing.

## Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work.

## Data availability

The link to the proposed method's implementation is mentioned within the manuscript.

## References

- [1] Maybeck P. *Stochastic models, estimation and control*, vol. 1. New York: Academic Press; 1979.
- [2] Dellaert F, Fox D, Burgard W, Thrun S. Monte Carlo localization for mobile robots. In: Proceedings 1999 IEEE international conference on robotics and automation (Cat. No. 99CH36288C), vol. 2. 1999, p. 1322–8. <http://dx.doi.org/10.1109/ROBOT.1999.772544>.
- [3] Thrun S. Particle filters in robotics. In: Proceedings of the 17th annual conference on uncertainty in AI. 2002.
- [4] Fox D. Adapting the sample size in particle filters through KLD-sampling. *Int J Robot Res* 2003;22(12):985–1003. <http://dx.doi.org/10.1177/0278364903022012001>.
- [5] Gustafsson F, et al. Particle filters for positioning, navigation, and tracking. *IEEE Trans Signal Process* 2002;50(2):425–37. <http://dx.doi.org/10.1109/78.978396>.
- [6] Thrun S, Burgard W, Fox D. *Probabilistic robotics, intelligent robotics and autonomous agents*. The MIT Press; 2005.
- [7] McKenna SJ, Nait-Charif H. Tracking human motion using auxiliary particle filters and iterated likelihood weighting. *Image Vis Comput* 2007;25(6):852–62. <http://dx.doi.org/10.1016/j.imavis.2006.06.003>.
- [8] Jensfelt P, Kristensen S. Active global localization for a mobile robot using multiple hypothesis tracking. *IEEE Trans Robot Autom* 2001;17(5):748–60. <http://dx.doi.org/10.1109/70.964673>.
- [9] Se S, Lowe D, Little J. Local and global localization for mobile robots using visual landmarks. In: Proceedings 2001 IEEE/RSJ international conference on intelligent robots and systems. expanding the societal role of robotics in the next millennium (Cat. No. 01CH37180), vol. 1. 2001, p. 414–20. <http://dx.doi.org/10.1109/IROS.2001.973392>.
- [10] Gasparri A, Panzneri S, Pascucci F, Ulivi G. A hybrid active global localisation algorithm for mobile robots. In: Proceedings 2007 IEEE international conference on robotics and automation. 2007, p. 3148–53. <http://dx.doi.org/10.1109/ROBOT.2007.363958>.
- [11] Zhu J, Zheng N, Yuan Z. An improved technique for robot global localization in indoor environments. *Int J Adv Robot Syst* 2011. <http://dx.doi.org/10.5772/10525>.
- [12] Kumagai I, Ueda R, Sugai F, Nozawa S, Kakiuchi Y, Okada K, et al. Achievement of localization system for humanoid robots with virtual horizontal scan relative to improved odometry fusing internal sensors and visual information. In: 2016 IEEE/RSJ international conference on intelligent robots and systems. IROS, <http://dx.doi.org/10.1109/iros.2016.7759124>.
- [13] Röwekämper J, Sprunk C, Tipaldi GD, Stachniss C, Pfaff P, Burgard W. On the position accuracy of mobile robot localization based on particle filters combined with scan matching. In: 2012 IEEE/RSJ international conference on intelligent robots and systems. 2012, p. 3158–64. <http://dx.doi.org/10.1109/IROS.2012.6385988>.
- [14] Vasiljević G, Miklič D, Draganjac I, Kovačić Z, Lista P. High-accuracy vehicle localization for autonomous warehousing. *Robot Comput-Integr Manuf* 2016;42:1–16. <http://dx.doi.org/10.1016/j.rcim.2016.05.001>.
- [15] Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* 1992;14(2):239–56. <http://dx.doi.org/10.1109/34.121791>, ISSN0162-8828.
- [16] Minguez J, Lamiraux F, Montesano L. Metric-based scan matching algorithms for mobile robot displacement estimation. In: Proceedings of the 2005 IEEE international conference on robotics and automation. 2005, p. 3557–63. <http://dx.doi.org/10.1109/ROBOT.2005.1570661>.
- [17] Censi A. An ICP variant using a point-to-line metric. In: 2008 IEEE international conference on robotics and automation. 2008, p. 19–25. <http://dx.doi.org/10.1109/ROBOT.2008.4543181>, Code available at <https://github.com/AndreaCensi/csm>.
- [18] Wang J, Zhao M, Chen W. MIM-SLAM: A multi-level ICP matching method for mobile robot in large-scale and sparse scenes. *Appl Sci* 2018;8:2432.
- [19] Yingzhong T, Xining L, Long L, Wenbin W. Intensity-assisted ICP for fast registration of 2D-LiDAR. *Sensors* 2019;19(9):2124.
- [20] Krzysztof N, Łukasz M. Use of a weighted ICP algorithm to precisely determine USV movement parameters. *Appl Sci* 2019;9(17):3530.
- [21] Tao L, Bui T, Ito T. Modified iterative closest point matching for 2D LiDAR laser data. In: 14th south east Asian technical university consortium 2020. (SEATUC 2020) 27th – 28th 2020, KX Building, KMUTT, Bangkok, Thailand.
- [22] Xu H, Zhou L, Zhao Y, Yuan Z. A two-dimensional point cloud matching method based on ICP improvement. In: China satellite navigation conference (CSNC) 2020 proceedings: vol. I. p. 390–8.
- [23] Koide K, Yokozuka M, Oishi S, Banno A. Voxelized GICP for fast and accurate 3D point cloud registration. In: 2021 IEEE international conference on robotics and automation. 2021, p. 11054–9. <http://dx.doi.org/10.1109/ICRA48506.2021.9560835>.
- [24] Filotheou A, Tzitzis A, Tsardoulas E, Dimitriou A, Symeonidis A, Georgiadis G, et al. Passive global localisation of mobile robot via 2D Fourier-Mellin invariant matching. *J Intell Robot Syst* 2022;104. <http://dx.doi.org/10.1007/s10846-021-01535-7>.
- [25] Cooper M, Raquet J, Patton R. Range information characterization of the Hokuyo UST-20LX LiDAR sensor. *Photonics* 2018;5:12.
- [26] Gutmann J, Konolige K. Incremental mapping of large cyclic environments. In: Proceedings 1999 IEEE international symposium on computational intelligence in robotics and automation. (Cat. No. 99EX375). 1999, p. 318–25. <http://dx.doi.org/10.1109/CIRA.1999.810068>.
- [27] Hahnel D, Burgard W, Fox D, Thrun S. An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In: Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 1 (Cat. No. 03CH37453). 2003, p. 206–11. <http://dx.doi.org/10.1109/IROS.2003.1250629>.
- [28] Wang C, Thorpe C, Thrun S. Online simultaneous localization and mapping with detection and tracking of moving objects: Theory and results from a ground vehicle in crowded urban areas. In: 2003 IEEE international conference on robotics and automation, vol. 1 (Cat. No. 03CH37422). 2003, p. 842–9. <http://dx.doi.org/10.1109/ROBOT.2003.1241698>.
- [29] Lacroix S, Mallet A, Bonnafoe D, Bazuil G, Fleury S, Herrb M, et al. Autonomous rover navigation on unknown terrains: Functions and integration. *Int J Robot Res* 2002;21(10–11):917–42. <http://dx.doi.org/10.1177/0278364902021010841>.
- [30] Minguez J, Montesano L, Montano L. An architecture for sensor-based navigation in realistic dynamic and troublesome scenarios. In: 2004 IEEE/RSJ international conference on intelligent robots and systems (IEEE Cat. No. 04CH37566), vol. 3. p. 2750–6, 3.
- [31] Montesano L, Minguez J, Montano L. Modeling dynamic scenarios for local sensor-based motion planning. *Auton Robots* 2008;25:231–51.
- [32] Schulz D, Burgard W, Fox D, Cremers AB. Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In: Proceedings 2001 ICRA. IEEE international conference on robotics and automation (Cat. No. 01CH37164), vol. 2. 2001, p. 1665–70. <http://dx.doi.org/10.1109/ROBOT.2001.932850>.
- [33] Kai L, Nüchter A, Hertzberg J, Surmann H. High-speed laser localization for mobile robots. *Robot Auton Syst* 2005;51:275–96.
- [34] Wang Z, Zhang L, Zhao S, Zhang S. Global localization with a single-line LiDAR by dense 2D signature and 1D registration. *IEEE Sens J* 2021;21(10):11497–506. <http://dx.doi.org/10.1109/JSEN.2020.3021049>.
- [35] Biber P, Strasser W. The normal distributions transform: A new approach to laser scan matching. In: Proceedings 2003 IEEE/RSJ international conference on intelligent robots and systems, vol. 3 (Cat. No. 03CH37453). 2003, p. 2743–8. <http://dx.doi.org/10.1109/IROS.2003.1249285>.
- [36] Bouraine S, Bougouffa A, Azouaoui O. NDT-PSO, a new NDT based SLAM approach using particle swarm optimization. In: 2020 16th international conference on control, automation, robotics and vision. 2020, p. 321–6. <http://dx.doi.org/10.1109/ICARCV50220.2020.9305519>.
- [37] Segal A, Hähnel D, Thrun S. Generalized-ICP. *Proc Robot: Sci Syst* 2009. <http://dx.doi.org/10.15607/RSS.2009.V.021>.
- [38] Yang H, Shi J, Carlone L. TEASER: Fast and certifiable point cloud registration. *IEEE Trans Robot* 2021;37(2):314–33. <http://dx.doi.org/10.1109/TRO.2020.3033695>.
- [39] Zhang Q, Wang P, Bao P, Chen Z. Mobile robot global localization using particle swarm optimization with a 2D range scan. In: 2017, in proceedings of the 2017 international conference on robotics and artificial intelligence. New York, NY, USA: Association for Computing Machinery; 2017, p. 105–9. <http://dx.doi.org/10.1145/3175603.3175618>.
- [40] Chen W, Huang T, Maalla A. Research on adaptive Monte Carlo location method based on fusion posture estimation. In: 2019 IEEE 3rd advanced information management, communication, electronic and automation control conference. 2019, p. 1209–13. <http://dx.doi.org/10.1109/IMCEC46724.2019.8983808>.
- [41] Liu X, Wei W, Gao Y, Yong, Wei X. Research on improved localization and navigation algorithm for automatic guided vehicle. *IOP Conf Ser: Mater Sci Eng* 2019;611:012076. <http://dx.doi.org/10.1088/1757-899X/611/1/012076>.



- [42] Filotheou A, Tsardoulas E, A. A Dimitriou, et al. Pose selection and feedback methods in tandem combinations of particle filters with scan-matching for 2D mobile robot localisation. *J Intell Robot Syst* 2020;100:925–44. <http://dx.doi.org/10.1007/s10846-020-01253-6>.
- [43] Park S, Park S. Global localization for mobile robots using reference scan matching. *Int J Control Autom Syst* 2014;12:156–68. <http://dx.doi.org/10.1007/s12555-012-9223-0>.
- [44] Peng G, Zheng W, Lu Z, Liao J, Hu L, Zhang G, et al. An improved AMCL algorithm based on laser scanning match in a complex and unstructured environment. *Complexity* 2018;2018:2327637. <http://dx.doi.org/10.1155/2018/2327637>, 11.
- [45] Sandberg D, Wolff K, Wahde M. A robot localization method based on laser scan matching. In: *Advances in robotics*. Berlin Heidelberg: Springer; 2009, p. 171–8.
- [46] Censi A, Iocchi L, Grisetti G. Scan matching in the hough domain. In: *Proceedings of the 2005 IEEE international conference on robotics and automation*. 2005, p. 2739–44.
- [47] Leordeanu M, Hebert M. A spectral technique for correspondence problems using pairwise constraints. In: *Tenth IEEE international conference on computer vision*, vol. 1, 2. 2005, p. 1482–9. <http://dx.doi.org/10.1109/ICCV.2005.20>.
- [48] Censi A. Scan matching in a probabilistic framework. In: *Proceedings 2006 IEEE international conference on robotics and automation*, vol. 2006. 2006, p. 2291–6. <http://dx.doi.org/10.1109/ROBOT.2006.1642044>.
- [49] Bresson G, Alsayed Z, Jonchery S. Graph-based map-aided localization using cadastral maps as virtual laser scans. In: *2019 IEEE intelligent transportation systems conference*. 2019, p. 4074–80. <http://dx.doi.org/10.1109/ITSC.2019.8917506>.
- [50] Chen Qin-Sheng, Defrise M, Deconinck F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Trans Pattern Anal Mach Intell* 1994;16(12):1156–68. <http://dx.doi.org/10.1109/34.387491>.
- [51] Filotheou A. Correspondenceless scan-to-map-scan matching of homoriented 2D scans for mobile robot localisation. *Robot Auton Syst* 2022;149:103957. <http://dx.doi.org/10.1016/j.robot.2021.103957>.
- [52] NDT implementation provided by the Point Cloud Library. 2021, <https://pointclouds.org/>. [Last Accessed 25 October 2021].
- [53] FastGICP implementation repository:. 2021, [https://github.com/SMRT-AIST/fast\\_gicp](https://github.com/SMRT-AIST/fast_gicp). [Last Accessed 25 October 2021].
- [54] CSM implementation repository:. 2021, <https://github.com/AndreaCensi/csm>. [Last Accessed 25 October 2021].
- [55] Xu P, Gu F, Song Z, Li J. Non-iterative multiple data registration method based on the motion screw theory and trackable features. In: *2018 24th international conference on pattern recognition*. 2018, p. 2428–32. <http://dx.doi.org/10.1109/ICPR.2018.8545473>.
- [56] Sobreira H, Costa C, Sousa I, Rocha L, Lima J, Farias P, et al. Map-matching algorithms for robot self-localization: A comparison between perfect match, iterative closest point and normal distributions transform. *J Intell Robot Syst* 2019;93. <http://dx.doi.org/10.1007/s10846-017-0765-5>.
- [57] Pishchvari A, Iurgel U, Lessmann S, Roesko-Koerner L, Tibken B. Radar scan matching using navigation maps. In: *2019 third IEEE international conference on robotic computing*. 2019, p. 204–11. <http://dx.doi.org/10.1109/IRC.2019.00038>.
- [58] Qingshan W, Jun Z. Point cloud registration algorithm based on combination of NDT and PLICP. In: *2019 15th international conference on computational intelligence and security*. 2019, p. 132–6. <http://dx.doi.org/10.1109/CIS.2019.00036>.
- [59] Pham Q-H, Tran N-H, Nguyen T-T, Tran T-P. Online robust sliding-windowed LiDAR SLAM in natural environments. In: *2021 international symposium on electrical and electronics engineering*. 2021, p. 172–7. <http://dx.doi.org/10.1109/ISEE51682.2021.9418728>.
- [60] VGICP implementation repository:. 2021, [https://github.com/SMRT-AIST/fast\\_gicp](https://github.com/SMRT-AIST/fast_gicp). [Last Accessed 25 October 2021].
- [61] NDT-PSO implementation repository:. 2021, [https://github.com/abougouffa/ndtpso\\_slam](https://github.com/abougouffa/ndtpso_slam). [Last Accessed 25 October 2021].
- [62] TEASER implementation repository:. 2021, <https://github.com/MIT-SPARK/TEASER-plusplus>. [Last Accessed 25 October 2021].
- [63] RPLIDAR A2M8 datasheet. 2021, [https://cdn.sparkfun.com/assets/e/a/f/9/8/LD208\\_SLAMTEC\\_rplidar\\_datasheet\\_A2M8\\_v1.0\\_en.pdf](https://cdn.sparkfun.com/assets/e/a/f/9/8/LD208_SLAMTEC_rplidar_datasheet_A2M8_v1.0_en.pdf). [Last Accessed 23 October 2021].
- [64] YDLIDAR G4 datasheet. 2021, <https://www.ydlidar.com/Public/upload/files/2021-07-24/YDLIDAR%20G4%20Data%20sheet%20V2.0.pdf>. [Last Accessed 23 October 2021].
- [65] YDLIDAR G6 datasheet. 2021, <https://www.ydlidar.com/Public/upload/files/2021-07-24/YDLIDAR%20G6%20Data%20sheet%20V1.6.pdf>. [Last Accessed 23 October 2021].
- [66] YDLIDAR TG30 datasheet. 2021, <https://www.robotshop.com/media/files/content/y/ydl/pdf/ydlidar-tg30-360-laser-scanner-30-m-datasheet.pdf>. [Last Accessed 23 October 2021].
- [67] YDLIDAR X4 datasheet. 2021, <https://www.ydlidar.com/Public/upload/files/2021-08-20/YDLIDAR%20X4%20Data%20sheet%20V2.0.pdf>. [Last Accessed 23 October 2021].