Optimality and limit behavior of the ML estimator for Multi-Robot Localization via GPS and Relative Measurements

Marco Todescato*, Andrea Carron*, Ruggero Carli*, Antonio Franchi†, Luca Schenato*

***Abstract*— This work addresses the problem of distributed multi-agent localization in presence of heterogeneous measurements and wireless communication. The proposed algorithm integrates low precision global sensors, like GPS and compasses, with more precise relative position (i.e., range plus bearing) sensors. Global sensors are used to reconstruct the absolute position and orientation, while relative sensors are used to retrieve the shape of the formation. A fast distributed and asynchronous linear least-squares algorithm is proposed to solve an approximated version of the non-linear Maximum Likelihood problem. The algorithm is provably shown to be robust to communication losses and random delays. The use of ACK-less broadcast-based communication protocols ensures an efficient and easy implementation in real world scenarios. If the relative measurement errors are sufficiently small, we show that the algorithm attains a solution which is very close to the maximum likelihood solution. The theoretical findings and the algorithm performances are extensively tested by means of Monte-Carlo simulations.**

I. MATHEMATICAL PRELIMINARIES

Resorting to standard graph theory, the estimation problem can be naturally associated with an *undirected measurement graph* $\mathbf{G} = (\mathbf{V};\mathbf{E})$ where $\mathbf{V} \in \{1,\dots,N\}$ represents the nodes and $\mathbf{E} \subset \mathbf{V}\times\mathbf{V}$ contains the unordered pairs of nodes $\{i,j\}$ which are connected to and measure each other. We denote with $\mathcal{N}_i \subseteq \mathbf{V}$ the set $\{j \mid \{i,j\} \in \mathbf{E}\}$, i.e. the neighboring set of node i. An undirected graph $\mathbf{G}$ is said to be connected if for any pair of vertices $\{i,j\}$ a path exists, connecting i to j. In the problem at hand, we consider a communication graph among the nodes which coincides with the measurements graph $\mathbf{G}$. Moreover, broadcast and asynchronous communications are assumed among the nodes. We denote with $|\cdot|$ the modulus of a scalar. Assuming M to be the cardinality of $\mathbf{E}$, the incidence matrix $A \in R^{M\times N}$ of $\mathbf{G}$ is defined as $A = [a_{ei}]$, where $a_{ei} = \{1,-1,0\}$, if edge e is incident on node i and directed away from it, is incident on node i and directed toward it, or is not incident on node i, respectively. We denote with the symbol $\|\cdot\|$ the vector 2-norm and with $[\cdot]^T$ the transpose operator. The symbol $\odot$ represents the *Hadamard* product. Given a vector $\mathbf{v} \in \mathbb{R}^2$, the function $\text{atan2}(\cdot) : \mathbb{R}^2 \to [0,2\pi]$ returns its angle, i.e., $\mathbf{v} = \|\mathbf{v}\| e^{j\,\text{atan2}(\mathbf{v})}$. Given a matrix $\mathbf{v} \in \mathbb{R}^{2\times n}$, with $v_{\text{ctr.}}$ we denote the vector centroid, i.e., $v_{\text{crt.}} = \frac{1}{n}\sum_{i=1}^{n} v_i$, where v_i is the i- th row of the matrix. The symbol σ_x denotes the standard deviation of the generic measurement x. The operator $\mathbb{E}[\cdot]$ denotes the expected value, while $\text{proj}(\cdot) : \mathbb{R} \to \mathbb{R}^2$ denotes the function $\text{proj}(\theta) = \begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}^T$. Finally, $\mathbb{I}$ denotes the identity matrix of suitable dimensions.

II. PROBLEM FORMULATION

Consider the problem of estimating the 2D positions, expressed in a common reference frame, of N nodes of a sensor network. Each node of the network is endowed with a set of sensors that provide both relative and absolute measurements.

In the following, firstly, we introduce the statistical models exploited for each type of measurements. Secondly, we formulate the non linear *Maximum-Likelihood* estimation problem. Thirdly, we introduce an suitable linear and convex reformulation.

A. Measurement Model

We assume that the N nodes are provided with a GPS module, a compass, a relative range sensor, and a relative bearing sensor. We denote with $p_i = (x_i, y_i),\ i \in \mathbf{V}$, the 2D position of node i in a common inertial frame, and with θ_i its orientation with respect to the inertial North axis, which in the following we assume to coincide with the x-axis. Each sensor is described by the following statistical model:

1) The GPS measurement $p_i^{\text{GPS}} = (x_i^{\text{GPS}}, y_i^{\text{GPS}})$ represents a noisy measurement of $p_i = (x_i, y_i)$. We assume a normal distribution of the GPS measurements, that is $p_i^{\text{GPS}} \sim \mathcal{N}(p_i, \sigma_p^2 \mathbb{I})$.
2) The compass provides a noisy measurement θ_i^C of θ_i. This is modelled according to an angular Gaussian distribution (see, e.g., [1]) which approximates the *Langevin* distribution [2]. This reads as $\text{proj}(\theta_i^C) \sim \mathcal{N}\left(\text{proj}(\theta_i), \sigma_\theta^2 \mathbb{I}\right)$.
3) The range sensor returns a noisy measurement r_{ij} of the distance between nodes i and j, which is modelled according to a normal distribution, that is $r_{ij} \sim \mathcal{N}(\|p_i - p_j\|, \sigma_r^2)$.
4) The bearing sensor returns a noisy measurement δ_{ij} of the bearing angle of the node j in the local frame of node i. For δ_{ij} we adopt an angular Gaussian distribution model which reads as $\text{proj}(\delta_{ij}) \sim \mathcal{N}\left(\text{proj}(\text{atan2}(p_j - p_i) - \theta_i), \sigma_\delta^2 \mathbb{I}\right)$.

Remark II.1. Observe that, in order to reduce the set-up cost, each node has access to highly noisy absolute measurements together with relative measurements that are less prone to noise than the absolute ones. In particular, the GPS sensors are usually characterized by a standard deviation $\sigma_p = 2$ [m]

†A. Franchi is with LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse CEDEX 4, France. antonio.franchi@laas.fr

*R. Carli, A. Carron, L. Schenato and M. Todescato are with the Department of Information Engineering (DEI) of the University of Padova, Via Gradenigo 6/B, 35100 Padova, Italy. [carlirug | carronan | schenato | todescat]@dei.unipd.it

[3], [4], while the compass by a standard deviation $\sigma_\theta = 0.05$ [rad] [5]. To retrieve information about range and bearing different methods can be used, e.g., depth-camera, laser, ultrasound. Acceptable values for the standard deviation of these measurements might be $\sigma_r = 0.1$ [m] and $\sigma_\delta = 0.03$ [rad]. Due to the variability in the accuracy of the available sensors, we will test our algorithm in a sufficiently wide range of standard deviation values.

For the sake of simplicity, we consider that all the nodes are endowed with a GPS module. However, a simple reformulation of the problem would still guarantee that all the results hold even if a reduced number of nodes are provided with a GPS.

B. Maximum-Likelihood Estimator

We assume that all the measurements are independent and their probability distributions are given in the previous section. It is possible to formulate the localization problem as a *Maximum-Likelihood* (ML) estimation problem [6]. Let us define the state and measurements sets, respectively, as

$$\begin{aligned}\mathbf{x} &= \{\mathbf{p}, \theta\} = \{p_i, \theta_i \text{ with } i \in \mathbf{V}\}, \\ \mathbf{y} &= \left\{p_i^{\text{GPS}}, \theta_i^C, r_{hk}, \delta_{hk} \text{ with } i \in \mathbf{V},\ (h,k) \in \mathbf{E}\right\},\end{aligned}$$

where $\mathbf{p} := [p_1, \ldots, p_N]^T$ and $\theta := [\theta_1, \ldots, \theta_N]^T$. Then, the negative log-likelihood cost function can be written as

$$J(\mathbf{x}) := -\log f\big(\mathbf{y} \,|\, \mathbf{x}\big) = J_p + J_\theta + J_r + J_\delta + c, \qquad (1)$$

where

$$\begin{aligned}J_p &= \sum_{i=1}^{N} \frac{\|p_i - p_i^{\text{GPS}}\|^2}{2\sigma_p^2}, \\ J_\theta &= \sum_{i=1}^{N} \frac{\left\|\text{proj}(\theta_i^C) - \text{proj}(\theta_i)\right\|^2}{2\sigma_\theta^2}, \\ J_r &= \sum_{(i,j)=1}^{M} \frac{(r_{ij} - \|p_i - p_j\|)^2}{2\sigma_r^2}, \\ J_\delta &= \sum_{(i,j)=1}^{M} \frac{\left\|\text{proj}(\delta_{ij}) - \text{proj}\left(\text{atan2}(p_j - p_i) - \theta_i\right)\right\|^2}{2\sigma_\delta^2},\end{aligned}$$

and c is a constant term that does not depend on $\mathbf{x}$ and $\mathbf{y}$. The minimization of the function in (1) would provide the maximum-likelihood estimator for the nodes absolute positions and orientations, i.e.:

$$\widehat{\mathbf{x}}^{\text{ML}} = \text{argmin}_{\mathbf{x}}\, J(\mathbf{x}). \qquad (2)$$

The ML estimator benefits of some properties regarding its mean and its asymptotic behavior. In particular, consider the following equivalent parametrization of agents' positions using their centroid $p_{\text{ctr.}}$ and corresponding deviation Δp_i. This reads as

$$p_i = p_{\text{ctr.}} + \Delta p_i, \qquad \sum_i \Delta p_i = 0, \qquad (3)$$

Let us also define $\Delta\mathbf{p} = (\Delta p_1, \ldots, \Delta p_N)$. Thanks to the new parametrization, equation (2) is equivalent to:

$$\begin{aligned}\left\{\widehat{p}_{\text{ctr.}}^{\text{ML}}, \Delta\widehat{\mathbf{p}}^{\text{ML}}, \widehat{\theta}^{\text{ML}}\right\} = \underset{\{p_{\text{ctr.}}, \Delta\mathbf{p}, \theta\}}{\text{argmin}} \quad & J(p_{\text{ctr.}}, \Delta\mathbf{p}, \theta), \qquad (4) \\ \text{s.t.} \quad & \sum_i \Delta p_i = 0.\end{aligned}$$

The previous reformulation allows us to prove the following lemma, which suggests how the ML estimator exploits the GPS information to solve for the absolute positioning of the formation centroid:

Lemma II.1. *Consider the negative* log-*likelihood cost function* (1)*. Then, the maximum likelihood solution* $\widehat{\mathbf{x}}^{\text{ML}}$ *which solves* (4) *is such that*

$$\widehat{p}_{\text{ctr.}}^{\text{ML}} = p_{\text{ctr.}}^{\text{GPS}}, \qquad (5)$$

where $\widehat{p}_{\text{ctr.}}^{\text{ML}} := \frac{1}{N}\sum_{i=1}^N \widehat{p}_i$ *and* $p_{\text{ctr.}}^{\text{GPS}} := \frac{1}{N}\sum_{i=1}^N p_i^{\text{GPS}}$.

Proof. Observe that only the term J_p of the log-likelihood cost function depends on $p_{\text{ctr.}}$. Indeed, J_θ is not a function of p_i; while, both J_r and J_δ depend only on the difference between p_i and p_j which, thanks to the equation (3) reads as

$$p_i - p_j = p_{\text{ctr.}} + \Delta p_i - p_{\text{ctr.}} - \Delta p_j = \Delta p_i - \Delta p_j.$$

It is then possible to consider only the log-likelihood relative to the GPS measurements. Specifically, if we define $p_i^{\text{GPS}} = p_{\text{ctr.}}^{\text{GPS}} + \Delta p_i^{\text{GPS}}$, it is possible to write

$$\begin{aligned}2\sigma_p^2 J_p &= \sum_{i=1}^{N} \|p_{\text{ctr.}} + \Delta p_i - (p_{\text{ctr.}}^{\text{GPS}} + \Delta p_i^{\text{GPS}})\|^2 \\ &= \sum_{i=1}^{N} \big(\|p_{\text{ctr.}} - p_{\text{ctr.}}^{\text{GPS}}\|^2 + \|\Delta p_i - \Delta p_i^{\text{GPS}}\|^2 + \\ &\quad +2(\Delta p_i - \Delta p_i^{\text{GPS}})^T(p_{\text{ctr.}} - p_{\text{ctr.}}^{\text{GPS}})\big) \\ &= N\|p_{\text{ctr.}} - p_{\text{ctr.}}^{\text{GPS}}\|^2 + \sum_{i=1}^{N} \|\Delta p_i - \Delta p_i^{\text{GPS}}\|^2,\end{aligned}$$

where we used the facts $\sum_i \Delta p_i = 0$ and $\sum_i \Delta p_i^{\text{GPS}} = 0$. To minimize the first term on the right hand side we must have

$$p_{\text{ctr.}} = p_{\text{ctr.}}^{\text{GPS}},$$

which proves the lemma. □

We can also state some limit behavior in a scenario where range, bearing and compass noises are very large or very small:

Lemma II.2. *For fixed GPS variance* σ_p *we have*

1) $\lim_{\max\{\sigma_\theta, \sigma_r, \sigma_\delta\} \to 0} \widehat{p}_i^{\text{ML}} = p_{\text{ctr.}}^{\text{GPS}} + \Delta p_i$,
2) $\lim_{\min\{\sigma_r, \sigma_\delta\} \to +\infty} \widehat{p}_i^{\text{ML}} = p_i^{\text{GPS}}$.

Proof. In the first scenario $\max\{\sigma_\theta, \sigma_r, \sigma_\delta\} \to 0$. This implies that the distributions for compass, range and bearing measurements converge to delta distributions, implying that

$$r_{ij} \to \|p_i - p_j\|, \quad \theta_i^C \to \theta_i, \quad \delta_{ij} \to \theta_i + atan2(p_j - p_i).$$

From these expressions it easily follows that

$$\widehat{p}_j - \widehat{p}_i \to p_j - p_i = r_{ij} e^{j(\delta_{ij} - \theta_i^C)}, \quad \{j,i\} \in \mathbf{E} \ ,$$

i.e., the relative vectorial distances among the communicating nodes are perfectly known. Since the graph is connected, it is possible to compute the exact vectorial difference among any two agents in the network, and therefore also the exact distance of any agent from the true centroid since:

$$\Delta\widehat{p}_i = \widehat{p}_i - \frac{1}{N}\sum_j \widehat{p}_j = \frac{1}{N}\sum_j (\widehat{p}_i - \widehat{p}_j) \to \frac{1}{N}\sum_j (p_i - p_j) = \Delta p_i.$$

Since $\widehat{p}_i = \widehat{p}_{\text{ctr.}} + \Delta\widehat{p}_i$ and from Lemma II.1 we have $\widehat{p}_{\text{ctr.}} = p^{\text{GPS}}_{\text{ctr.}}$, then it follows the first part of the lemma.
In the second scenario when $\min\{\sigma_r, \sigma_\delta\} \to +\infty$ becomes arbitrary large, the probability distribution of range and bearing degenerate into an uniform distribution with infinite support. As so, the terms J_r and J_δ become negligible as compared to J_p and J_θ. Since the positions p_i do not appear in J_θ, it follows that $\widehat{p}_i$ results from the minimization of J_p, which gives $\widehat{p}_i = p_i^{\text{GPS}}$ and, therefore, the claim of the lemma. □

Scenario 1) of Lemma II.2 states that in the case where $\max\{\sigma_\theta, \sigma_r, \sigma_\delta\} \to 0$, the shape of the formation is perfectly retrieved. In this case the only source of error between the estimated formation and the ground-truth is given by the error between GPS centroid and the true centroid. Scenario 2) states that if the relative measurements accuracies deteriorate, the ML estimator will "trust" the GPS measurements only.
Unfortunately problem (2) is highly non linear and hard to solve. In particular, it is known that, if the angles are noise-free, the problem is linear [7]. Conversely, if the angles are not known, the problem presents many local minima [8], [9]. One possible way to tackle it, is using a standard gradient descent approach since the gradient vector of the log-likelihood function can be computed in closed form using (1). However, such approach heavily suffers of bad initialization. In fact, the presence of multiple local minima in the cost function (1) causes the algorithm to stop in the wrong minimizer.
In the following, we resort to a suitable approximation which let us reformulate the problem in a classical linear-least square framework.

C. An Approximated Linear Least-Squares Formulation

An approximated solution for the problem stated in (2), which exploits a suitable model linearization, is now presented. The idea is to move from the polar coordinate system to the equivalent Cartesian representation.
Indeed, assuming a perfect knowledge of range, bearing and compass, it is possible to express the *displacement* d_{ij} between agent i and j as

$$d_{ij} := p_i - p_j = r_{ij} \begin{bmatrix} \cos(\delta_{ij} + \theta_i) \\ \sin(\delta_{ij} + \theta_i) \end{bmatrix}. \tag{6}$$

Since the measurements are affected by noise, it is necessary to map the noise of range, bearing and compass into the equivalent noise in Cartesian coordinates. Namely, given the noisy version of (6), that is

$$d_{ij} = p_i - p_j + n_{ij}, \tag{7}$$

where n_{ij} is the noise in Cartesian coordinate, we want to find the expression for its covariance, $\mathbb{E}[n_{ij} n_{ij}^T] = \Sigma_{ij}$, in terms of the statistical description of range, bearing and compass measurements noises. After a first order expansion we obtain

$$\Sigma_{ij} = \begin{bmatrix} \sigma_x^2(i,j) & \sigma_{xy}(i,j) \\ \sigma_{yx}(i,j) & \sigma_y^2(i,j) \end{bmatrix}, \tag{8}$$

where

$$\begin{aligned}
\sigma_x^2(i,j) &= \sigma_r^2 \cos^2(\delta_{ij} + \theta_i) + r_{ij}^2(\sigma_\delta^2 + \sigma_\theta^2)\sin^2(\delta_{ij} + \theta_i),\\
\sigma_y^2(i,j) &= \sigma_r^2 \sin^2(\delta_{ij} + \theta_i) + r_{ij}^2(\sigma_\delta^2 + \sigma_\theta^2)\cos^2(\delta_{ij} + \theta_i),\\
\sigma_{xy}(i,j) &= \left(\sigma_r^2 - r_{ij}^2(\sigma_\delta^2 + \sigma_\theta^2)\right)\sin(\delta_{ij} + \theta_i)\cos(\delta_{ij} + \theta_i).
\end{aligned}$$

Remark II.2. Since the linear approximation introduced is based on a first order expansion, its validity holds under the assumption of sufficiently small measurement errors.

Remark II.3. Note that Σ_{ij} is a function of the true values of range, bearing and compass. Since it is not possible to have access to these data, in a real setup these quantities must be replaced by their corresponding measured values.

Once computed the displacements, it is possible to define the weighted residuals as

$$J_d = \frac{1}{2}\sum_{\{i,j\}\in\mathbf{E}} \|p_i - p_j - d_{ij}\|^2_{\Sigma_{ij}^{-1}}.$$

Thanks to this, it is possible to define an approximation of the negative log-likelihood in (1), which accounts for the GPS measurements and the displacements, as

$$J_{\text{LS}}(\mathbf{p}) = J_p + J_d. \tag{9}$$

The minimization problem becomes

$$\widehat{\mathbf{p}}^{\text{LS}} = \text{argmin}_{\mathbf{p}} \ J_{\text{LS}}(\mathbf{p}) \ , \tag{10}$$

which is a linear least-squares problem, thus convex, which can be solved in closed form. Specifically, assuming $\mathbf{G}$ connected, the optimal estimate is given by

$$\widehat{\mathbf{p}}^{\text{LS}} = (\Sigma_{\text{GPS}}^{-1} + A^T\Sigma^{-1}A)^{-1}(\Sigma_{\text{GPS}}^{-1}\mathbf{p}^{\mathbf{GPS}} + A^T\Sigma^{-1}\mathbf{d}), \tag{11}$$

where $\Sigma_{\text{GPS}} = \sigma_p^2\mathbb{I}$, Σ is the matrix which accounts for all the Σ_{ij}, and $\mathbf{d}$ and $\mathbf{p}^{\text{GPS}}$ are the vectors obtained stacking together all the relative distances defined in (7) and the GPS absolute positions, respectively.

Remark II.4. Note that the LS estimates only the absolute positions $\mathbf{p}$ without providing any estimate of the absolute orientations. These are retrieved using the compass and exploited to project the noise in rectangular coordinates.

Remark II.5. Observe that, even if the linear least-squares problem returns an approximate solution for the problem of equation (2), since the problem of equation (10) is convex, its solution is unique.

For the LS estimator it is possible to show an optimal result similar to the one stated in Lemmas II.1 and II.2 for the ML estimator. We state the following:

Lemma II.3. *Consider the cost function* (9)*. Then, the optimal solution* $\widehat{\mathbf{p}}^{\mathrm{LS}}$ *which solves* (10) *is such that*

$$\widehat{p}^{\mathrm{LS}}_{\mathrm{ctr.}} = p^{\mathrm{GPS}}_{\mathrm{ctr.}}. \tag{12}$$

Moreover, for fixed GPS variance σ_p *we have*

$$\lim_{\max\{\sigma_\theta,\sigma_r,\sigma_\delta\}\to 0} \widehat{p}^{\mathrm{LS}}_i = p^{\mathrm{GPS}}_{\mathrm{ctr.}} + \Delta p^{\mathrm{LS}}_i \ ,$$

$$\lim_{\min\{\sigma_r,\sigma_\delta\}\to +\infty} \widehat{p}^{\mathrm{LS}}_i = p^{\mathrm{GPS}}_i \ .$$

Proof. The result follows with arguments similar to those used in Lemma II.1 and II.2. □

Observe that, to compute $\widehat{\mathbf{p}}^{\mathrm{LS}}$ as in equation (10), one needs all the measurements, their covariances and the topology of $\mathbf{G}$ to be available to a central computation unit. In the following section we present a solution which is amenable for a distributed and asynchronous implementation. We assume that a nodes i and j can communicate with each other only if $\{i, j\} \in \mathbf{E}$. Remarkably, the solution is robust to packet losses and delays in the communication channel.

References

[1] N. I. Fisher, T. Lewis, and B. J. J. Embleton, *Statistical analysis of spherical data*. Cambridge Univ Pr, 1993.

[2] P. Langevin, "Sur la thorie du magntisme," *Journal de physique thorique et applique*, 1905.

[3] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global Positioning Systems, Inertial Navigation, and Integration*. Wiley-Interscience, 2007.

[4] H. Abbott and D. Powell, "Land-vehicle navigation using gps," *Proceedings of the IEEE*, 1999.

[5] R. Racz, C. Schott, and S. Huber, "Electronic compass sensor," in *Sensors, 2004. Proceedings of IEEE*, 2004.

[6] J. Pfanzagl, *Parametric statistical theory*. Walter de Gruyter, 1994.

[7] P. Barooah, "Estimation and control with relative measurements: Algorithms and scaling laws," Ph.D. dissertation, University of California, Santa Barbara, 2007.

[8] H. Wang, G. Hu, S. Huang, and G. Dissanayake, "On the structure of nonlinearities in pose graph slam." in *Robotics: Science and Systems*, 2012.

[9] L. Carlone and A. Censi, "From angular manifolds to the integer lattice: Guaranteed orientation estimation with application to pose graph optimization," *IEEE Trans. on Robotics*, 2014.