Certifiability Analysis of the Global Optimality in Camera-Based Positioning with SEC-PnP Algorithm

Antonino Triolo, Chen Zhu, Michael Meurer

Institute of Communications and Navigation,
German Aerospace Center (DLR), Oberpfaffenhofen, Germany
Email: {Antonino.Triolo, Chen.Zhu, Michael.Meurer}@dlr.de

BIOGRAPHY

Antonino Triolo is a researcher at the German Aerospace Center (DLR, Institute of Communication and Navigation). He graduated in Computer Engineering in 2020 and in Automation and Control Engineering in 2022 from the University of Catania (Italy, Sicily), and did a short research period there. In the same year, after finishing his studies, he started his work at DLR. His research concerns the applicability of the concept of integrity in the case of camera-based navigation in the context of Unmanned Aerial Vehicles (UAVs).

Chen Zhu is a senior research fellow and the head of Visual and Terrestrial Augmentation research group at the Institute of Communications and Navigation, German Aerospace Center (DLR). He received his Ph.D. degree (Dr.-Ing.) and Master's degree (M.Sc) from Technical University of Munich, Germany and his Bachelor's degree from Tsinghua University, Beijing, China. He is interested in the research fields of visual navigation, multi-sensor fusion, and robotic swarm navigation, currently focusing on the system integrity.

Michael Meurer is the director of the Department of Navigation at the German Aerospace Center (DLR) and of the Center of Excellence for Satellite Navigation. In addition, he is a full professor of electrical engineering and director of the Chair of Navigation at the RWTH Aachen University. His current research interests include GNSS signals, GNSS receivers and navigation for safety-critical applications.

ABSTRACT

In GNSS degraded environments, cameras have great potential to be used as navigation sensors in various applications such as UAV landing, autonomous driving, indoor navigation, etc. Using a set of visible features with known locations on the map, the six degrees of freedom (DOF) pose (position and attitude) of the camera can be estimated. This is known as the Perspective-n-Point (PnP) problem. However, as the measurement equation is highly non-linear, the camera pose estimation result is affected by various error sources, such as an incorrect initial pose estimate or the presence of noise in the measurements (Zhu et al. (2022)).

Due to the aforementioned problems, there is still a lack of methods for camera-based navigation that provide results with certified global optimality and overbounded pose estimation error. In the previous work Triolo et al. (2024), we proposed the Slack-Eigen-Convexification-PnP (SEC-PnP) algorithm to estimate the camera pose reliably even when the initial pose estimation is erroneous and the measurements are noisy. In this work, we provide a method to analytically prove the global optimality of the pose estimation results obtained by the SEC-PnP algorithm, and discuss the conditions required to fulfill the global optimality. Also, a study on the accuracy of the certified optimal solution is performed. By comparing the highest achievable accuracy of the SEC-PnP algorithm with the Cramer-Rao Lower Bound (CRLB) for certain measurement noise levels and number of the 2D-3D associations, the global optimality of the results can be validated. The proposed method with its proven global optimality has great potential to be exploited in applications with high reliability requirements.

I. INTRODUCTION

In recent years there has been an increase in the use of cameras for navigation in a variety of systems ranging from automotive to robotics and aerospace. Each context has different requirements on the navigation system. For example, in augmented reality and many robotics applications, a fast and accurate camera-based localization algorithm is usually preferred. The real-time feasibility and the continuity of the system have a higher priority than the system integrity in the algorithm design. However, in safety-critical applications such as urban air mobility (UAM) or autonomous driving, the navigation system also needs to provide high integrity solutions.

Using a set of 2D visible features with known 3D locations, the 6-DOF pose (position and attitude) of the camera can be estimated in the reference frame of the 3D points with a single image. This is known as the Perspective-n-Point (PnP) problem. A few state-of-the-art algorithms can solve the PnP problem analytically, e.g., P3P in Gao et al. (2003), SQPNP in Terzakis and Lourakis (2020), EPNP in Lepetit et al. (2009). They have already been implemented in the prestigious open source library OpenCV (Bradski (2000)) and widely used in various visual navigation applications. However, most of the analytical PnP solvers are based on the noise-free assumption, but in practice the measurements are always noisy.

Nonlinear optimisation is required to solve the PnP problem with noisy measurements. Gradient or Jacobian-based methods such as Gauss-Newton or the Levenberg-Marquardt (Eade (2013)) algorithms are widely used to estimate camera poses iteratively. Nevertheless, the performance of the nonlinear optimization is significantly dependent on the initial pose estimation, since the measurement equations for solving the PnP problem are highly nonlinear. It has been shown that with exactly the same measurements and associations of 2D-3D points, the estimation results may differ for distinct initial guesses of the camera pose (Zhu et al. (2022)). It is in many cases challenging to validate whether the solver has converged to the global optimum or not. This introduces integrity risks to the navigation system and strongly limits their use in safety-critical applications.

In a recent work (Triolo et al. (2024)), we proposed the Slack-Eigen-Convexification-PnP (SEC-PnP) algorithm to estimate the camera pose reliably even when the initial pose estimation is erroneous and the measurements are noisy. SEC-PnP models the PnP problem as a Quadratic Constrained Quadratic Problem (QCQP) and iteratively convexifies the original nonlinear (and non-convex) problem using slack variables and linearization of the quadratic constraints based on the FPP-SCA algorithm (Mehanna et al. (2015)). It has been shown that the SEC-PnP algorithm significantly outperforms the state of the art in terms of accuracy and robustness to initial estimation errors using both simulated and real experimental data, at the cost of slightly increased computational time. In this work, we provide a method to analytically test and prove if the estimation result obtained by the aforementioned algorithm is the global optimal solution. The method is validated by comparing the highest achievable accuracy of the proposed method to the Cramer-Rao lower bound of the corresponding PnP problem.

The proposed SEC-PnP algorithm and the method to prove its global convergence, have great potential to be exploited in applications with high reliability requirements, such as urban air mobility, autonomous driving, etc. The most significant contribution of this work is the introduction of an integrity monitor for the convergence of the SEC-PnP algorithm. Quantifiable integrity and certifiability is available in many other navigation technologies, e.g., GNSS based positioning, but it has been a great challenge for camera-based navigation algorithms. This development will be a fundamental step in securing the visual navigation integrity in the future.

The paper is organized as follows: in Section II, the system model and reference frame definitions are introduced. In Section III, the global optimality of the SEC-PnP method is analytically demonstrated. Validations of the algorithm optimality and certifiability are shown in Section IV, using both simulated and real scenarios. Conclusions are drawn in Section V.

II. SYSTEM MODEL

In this section, the basics of camera-based positioning and the system models are briefly introduced. More detailed introduction of the theory behind visual navigation can be found in Hartley and Zisserman (2004), and the error sources and uncertainty in the visual positioning process are reviewed in Zhu et al. (2022).

The raw sensor measurement of digital cameras is a discrete image $\boldsymbol{I}$ which represents an amplitude measure of the illuminance during the exposure time on the image plane $\Omega \subset \mathbb{R}^2$. For visual navigation applications, it is essential to extract geometric information from the luminance of the images, and the processing should be in real time. For navigation purposes, gray-scale images are typically used because of computational reasons. It is important to underline that the intensity values in the image $\boldsymbol{I}$ can be noisy due to various error sources.

The location of pixels in the image plane and the corresponding light sources follows projective geometry. Most systems can be approximated using the classic pinhole model. For this model, the two-dimensional (2D) position of an image point $\boldsymbol{p}_i^{2D} = [x_i^{2D}, y_i^{2D}]^\top \in \Omega \subset \mathbb{R}^2$ can be related to the corresponding 3D point $\boldsymbol{p}_i^C = [x_i^C, y_i^C, z_i^C]^\top$ as :

$$\boldsymbol{p}_i^{2D} = \left[f_x \frac{x_i^C}{z_i^C} + c_x, f_y \frac{y_i^C}{z_i^C} + c_y \right]^\top \tag{1}$$

with f_x and f_y denoting the focal lengths in x and y direction, (c_x, c_y) denotes the coordinate of the projected principal point in the 2D image. The transformation between the camera frame C and the world reference frame O is dependent on the position and attitude of the camera. For a camera at position $\boldsymbol{t}_O \in \mathbb{R}^3$ and with attitude represented by $\boldsymbol{R}_O \in SO(3)$ in the world frame, the projective geometry of a 3D point in the world frame is described by:

$$\widetilde{\boldsymbol{p}}_i^{2D} = d_i \begin{bmatrix} \boldsymbol{p}_i^{2D} \\ 1 \end{bmatrix} = \boldsymbol{K}\,\boldsymbol{R}_O[\boldsymbol{I}| - \boldsymbol{t}_O] \begin{bmatrix} \boldsymbol{p}_i^{O} \\ 1 \end{bmatrix} = \boldsymbol{Q}_{cam}\widetilde{\boldsymbol{p}}_i^{O} \tag{2}$$

where $\widetilde{\boldsymbol{p}}_i^{O}$ is expressed in homogeneous coordinates in the world frame, and $\boldsymbol{K}$ is the intrinsic matrix. A simple example with four points in the space projected to the image plane is shown in Figure 1. By rewriting the last equation with variables in Euclidean space, we denote the 2D coordinates of a point in the image by a function of the camera pose and 3D position of the point as:

$$\boldsymbol{p}_i^{2D} = \pi(\widetilde{\boldsymbol{p}}_i^{O}, \boldsymbol{x}) \tag{3}$$

where $\boldsymbol{x} \in \mathbb{R}^6$ is the camera pose parameterized with a six degrees-of-freedom vector. The camera pose $\boldsymbol{x}$ can be estimated given a set of associated 3D points and their corresponding 2D projections in the image. The coordinates of the 3D points can be obtained, for example, by using georeferenced landmarks and maps.

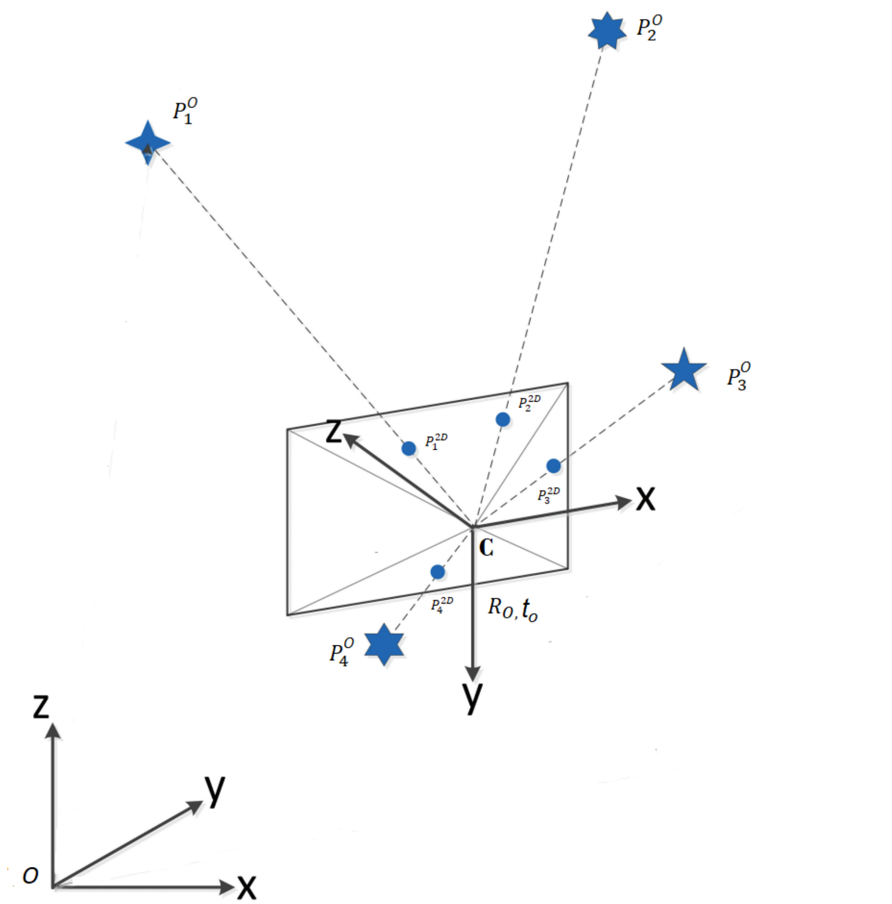


Figure 1: Reference frames in visual positioning.

By stacking all the N successfully associated 2D-3D points, the pose of the camera can be estimated by solving the following nonlinear optimization problem iteratively:

$$\boldsymbol{x}_{est} = arg\,min_x \, \|\boldsymbol{p}^{2D*} - \pi(\boldsymbol{x})\|^2 \tag{4}$$

with $\boldsymbol{p}^{2D*} = [\boldsymbol{p}_1^{2D*}, ..., \boldsymbol{p}_N^{2D*}]^\top$ and $\boldsymbol{p}_i^{2D*} = \boldsymbol{p}_i^{2D} + \boldsymbol{n}$, being $\boldsymbol{p}^{2D*}$ the vector composed by the stacked 2D measurements and $\boldsymbol{p}_i^{2D*}$ the single 2D measurement composed by the sum of the correct value given by the camera model and a noise vector $\boldsymbol{n}$ derived by multiple noise sources and disturbances, respectively. Equation 4 is the fundamental optimization problem for visual pose estimation. It is easy to verify that the measurement function $\pi(\boldsymbol{x})$ is highly nonlinear.

III. GLOBAL OPTIMALITY OF THE SEC-PNP ALGORITHM

In this section, the optimality of the pose estimation provided by the SEC-PnP algorithm is shown. For details of the construction of the cost function, the constraints and the iterative mechanism of the algorithm, please consult Triolo et al. (2024). In Subsection III.1, the structure of the PnP problem in Quadratic Constrained Quadratic Problem form will be briefly illustrated, followed by the convexified form used in SEC-PnP. Subsection III.2 will briefly describe the Karush-Kuhn-Tucker (KKT) conditions, which are necessary conditions for global optimality, and how they are applied to the QCQP problem representing the PnP problem. Finally, in Subsection III.3, the sufficient conditions that, once the KKT conditions are satisfied, analytically validate the pose estimation obtained from the SEC-PnP will be illustrated.

1. Equality between the Original Non-Convex Problem and the Sequence of Convex Problems

As a first step, we introduce the QCQP form of the PnP problem as

$$\begin{aligned} \min_{x} \quad & \boldsymbol{x}^\top \boldsymbol{W}_{tot} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} = c_i, \qquad i = 1, ..., 6. \end{aligned} \tag{5}$$

where $\boldsymbol{x} \in \mathbb{R}^{12}$ is the parameter vector with the rotation parameterised as matrix $\boldsymbol{R}_O \in SO(3)$ and $\boldsymbol{W}_{tot}$ represents the total geometric distance given by the difference between the actual measurements and the camera measurement model for each 2D-3D association. $\boldsymbol{R}_O$ is a 3×3 orthonormal matrix with determinant 1. This property can be briefly summarized in the expression $\boldsymbol{R}_O^\top \boldsymbol{R}_O = \boldsymbol{I}$. The latter can be expressed by a set of equations by using the vector of variables $\boldsymbol{x}$, so $\boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} = c_i$ with $\boldsymbol{c} = [1, 1, 1, 0, 0, 0]^\top$ and $i = 1, ..., 6$ represents the i-th constraint given to the quadratic optimization problem.

Non-convex QCQPs are NP-hard in general. Existing approaches relax the non-convexity using semi-definite positive (SDP) relaxation or linearize the non-convex part and solve the resulting convex problem. However, these techniques are seldom successful in even obtaining a feasible solution when the QCQP matrices are indefinite (Mehanna et al. (2015)). We propose to employ the FPP-SCA iterative algorithm for obtaining the optimal solution to the QCQP formulation in Equation (5), where we approximate the feasible region through a linear restriction of the non-convex parts of the constraints.

In order to guarantee feasibility of the modified problem, slack variables are added, and a penalty is used to ensure that slacks are sparingly used. The solution of the resulting optimization problem is then used to compute a new linearization, and the procedure is repeated until convergence. The previous procedure can be summarised in the following steps:

1. Set $k = 0$ and $\|\boldsymbol{s}_0\| = 0$ and generate a initial point $\boldsymbol{z}_0$ by knowing an initial estimate of the rotation $\boldsymbol{R}_0$ and the translation vector $\boldsymbol{t}_0$.
2. Solve the following problem:

$$\begin{aligned} \min_{x,\, s} \quad & \boldsymbol{x}^\top \boldsymbol{W}_{tot} \boldsymbol{x} + f(k, \|\boldsymbol{s}_k\|)\, \|\boldsymbol{s}\| \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{P}_i^{(+)} \boldsymbol{x} + 2\boldsymbol{z}_k^\top \boldsymbol{P}_i^{(-)} \boldsymbol{x} \leq c_i + \boldsymbol{z}_k^\top \boldsymbol{P}_i^{(-)} \boldsymbol{z}_k + s_i, \\ & \boldsymbol{x}^\top \boldsymbol{P}_i^{(-)} \boldsymbol{x} + 2\boldsymbol{z}_k^\top \boldsymbol{P}_i^{(+)} \boldsymbol{x} \geq c_i + \boldsymbol{z}_k^\top \boldsymbol{P}_i^{(+)} \boldsymbol{z}_k - s_i, \qquad i = 1, ..., 6, \end{aligned} \tag{6}$$

   where $\boldsymbol{P}_i^{(+)}$ and $\boldsymbol{P}_i^{(-)}$ are the Positive Semi-Definitive and Negative Semi-Definite components derived from the constraint matrices $\boldsymbol{P}_i$ respectively, and $\boldsymbol{s}$ is the slack variables vector.
3. Let $\boldsymbol{x}_k$ denote the optimal $\boldsymbol{x}$ obtained in Problem (6) at the k iteration, and set $\boldsymbol{z}_{k+1} = \boldsymbol{x}_k$ and $\|\boldsymbol{s}_{k+1}\| = \|\boldsymbol{s}_k\|$.
4. Set $k = k + 1$ and repeat the procedure from point 2. until convergence.
5. Once convergence is achieved, refine the estimate using Gauss-Newton or Levenberg-Marquardt iterations.

If the FPP-SCA algorithm converges and the converged slack variables turn out being all zero, then it is easy to show that the remaining variables satisfy the KKT conditions for the original Problem (5) (Mehanna et al. (2015),Beck et al. (2010)). In the infinitesimally small chance that the algorithm converges with slack variables not equal to 0, the result obtained is simply discarded.

After the convergence, to further improve the accuracy of the results and to eliminate the influence of slack variables in the optimization process, and to have the same consistency and asymptotic normality as a maximum likelihood (ML) estimator, we perform additional iterations with Gauss-Newton or or Levenberg-Marquardt algorithms, as proposed in Zeng et al. (2024).

2. Karush-Kuhn-Tucker Conditions

As previously stated, each solution provided by SEC-PnP is a KKT point, i.e. a point that satisfies the Karush-Kuhn-Tucker conditions, which are first derivative tests (sometimes called first-order necessary conditions) for a solution in nonlinear programming to be optimal, of the optimization problem. For more information on duality theory and the Lagrangian function, please consult Boyd and Vandenberghe (2004).

Let $\boldsymbol{x}^*$ and $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ be any primal and dual optimal points with zero duality gap for a general programming problem. Since $\boldsymbol{x}^*$ minimizes the Lagrangian function $L(\boldsymbol{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ over $\boldsymbol{x}$, it follows that its gradient must vanish at $\boldsymbol{x}^*$, i.e.,

$$\nabla f_0(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\boldsymbol{x}^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(\boldsymbol{x}^*) = 0 \tag{7}$$

Thus we have

$$\begin{aligned} f_i(\boldsymbol{x}^*) \leq 0, &\qquad i = 1, ..., m \\ h_i(\boldsymbol{x}^*) = 0, &\qquad i = 1, ..., p \\ \lambda_i^* \geq 0, &\qquad i = 1, ..., m \\ \lambda_i^* f_i(\boldsymbol{x}^*) = 0, &\qquad i = 1, ..., m \\ \nabla f_0(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\boldsymbol{x}^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(\boldsymbol{x}^*) = 0, & \end{aligned} \tag{8}$$

with f_0 representing the cost function, f_i representing the i-th inequality constraint function, h_i representing the i-th equality constraint function, and $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ representing the inequality and equality Lagrange multipliers.

For convenience, we transform the equations in the constraints from the original Problem (5) into two inequalities as

$$\begin{aligned} \min_{x} \quad & \boldsymbol{x}^\top \boldsymbol{W}_{tot} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} \leq c_i, \\ & \boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} \geq c_i, \qquad i = 1, ..., 6. \end{aligned} \tag{9}$$

Based on this formulation, we can fix $m = 6$ and define

$$\begin{aligned} f_0 &= \boldsymbol{x}^\top \boldsymbol{W}_{tot} \boldsymbol{x} \\ f_{i,1} &= \boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} - c_i \\ f_{i,2} &= -\boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} + c_i = -f_{i,1}, \end{aligned} \tag{10}$$

and finally rewrite the KKT conditions applied to the QCQP form of the PnP problem

$$\nabla(\boldsymbol{x}^{*\top} \boldsymbol{W}_{tot} \boldsymbol{x}^*) + \sum_{i=1}^{m} [\lambda_{i,1}^* \nabla(\boldsymbol{x}^{*\top} \boldsymbol{P}_i \boldsymbol{x}^* - c_i) + \lambda_{i,2}^* \nabla(-\boldsymbol{x}^{*\top} \boldsymbol{P}_i \boldsymbol{x}^* + c_i)] = 0 \tag{11}$$

that can be simplified into

$$\nabla(\boldsymbol{x}^{*\top} \boldsymbol{W}_{tot} \boldsymbol{x}^*) + \sum_{i=1}^{m} (\lambda_{i,1}^* - \lambda_{i,2}^*) \nabla(\boldsymbol{x}^{*\top} \boldsymbol{P}_i \boldsymbol{x}^* - c_i) = 0 \tag{12}$$

with

$$
\begin{aligned}
\boldsymbol{x}^{*\top}\boldsymbol{P}_i\boldsymbol{x}^* - c_i \leq 0, &\qquad i = 1, ..., m\\
-\boldsymbol{x}^{*\top}\boldsymbol{P}_i\boldsymbol{x}^* + c_i \leq 0, &\qquad i = 1, ..., m\\
\lambda^*_{i,1} \geq 0, &\qquad i = 1, ..., m\\
\lambda^*_{i,2} \geq 0, &\qquad i = 1, ..., m\\
\lambda^*_{i,1}(\boldsymbol{x}^{*\top}\boldsymbol{P}_i\boldsymbol{x}^* - c_i) = 0, &\qquad i = 1, ..., m\\
\lambda^*_{i,2}(-\boldsymbol{x}^{*\top}\boldsymbol{P}_i\boldsymbol{x}^* + c_i) = 0, &\qquad i = 1, ..., m.
\end{aligned} \tag{13}
$$

3. Sufficient Condition to Global Optimality

The Karush-Kuhn-Tucker conditions, as mentioned before, are only necessary (and not sufficient) to prove the global optimality of the results obtained, and consequently more stringent conditions are needed to certify the process. Fortunately, there are different mathematical tools to test the conditions for the global optimality of the solution of a QCQP type of problem. Such tools can be applied to analytically test if the output of our proposed SEC-PnP method is the global minimum of the nonlinear minimization. One of those tools is the so-called "Positive Semidefiniteness Condition" (Lu and Fang (2011), Gao (2008), Yeyakumar et al. (2007)). The global optimality condition is given as follows.

Theorem 1 (Positive Semidefiniteness Condition) *Let (14) be a general QCQP*

$$
\begin{aligned}
\min_{x} \quad & \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{q}^\top\boldsymbol{x}\\
\text{s.t.} \quad & \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{Q}_i\boldsymbol{x} + \boldsymbol{q}_i^\top\boldsymbol{x} \leq b_i, \qquad i = 1, ..., m,
\end{aligned} \tag{14}
$$

where $\boldsymbol{Q}$ and $\boldsymbol{Q}_i$ are $n \times n$ real symmetric matrices and $\boldsymbol{q}$ and $\boldsymbol{q}_i$ are real vectors in $\mathbb{R}^n$. For simplicity, we define $\boldsymbol{b} = [b_1, ..., b_m]^\top$ Let $\boldsymbol{x}^*$ be a Karush-Kuhn-Tucker solution of the given QCQP (14) with a corresponding Lagrangian vector $\boldsymbol{\lambda}^*$. If

$$
\boldsymbol{Q} + \sum_{i=1}^{m} \lambda_i^* \boldsymbol{Q}_i \succcurlyeq 0, \tag{15}
$$

then $\boldsymbol{x}^*$ is a global optimal solution of the QCQP (14).

Notice that the positive semidefiniteness condition is in fact a case of the classical Saddle Point optimality conditions as described in Bazaraa et al. (2005). It is possible to apply the aforementioned condition to the QCQP form of the PnP problem (9) with $m = 6$, and

$$
\boldsymbol{W}_{tot} + \sum_{i=1}^{m}(\lambda^*_{i,1}\boldsymbol{P}_i - \lambda^*_{i,2}\boldsymbol{P}_i) = \boldsymbol{W}_{tot} + \sum_{i=1}^{m}(\lambda^*_{i,1} - \lambda^*_{i,2})\boldsymbol{P}_i \succcurlyeq 0. \tag{16}
$$

To summarise the results found in the this section, we can introduce the following theorem. Please refer to (2) for a visual representation of the certification process.

Theorem 2 (Global optimality of SEC-PnP) *The results of the SEC-PnP algorithm can be proved analytically to be the global minimizers of the QCQP (5) representing the PnP problem if the critical point found $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ satisfies both:*

- *the KKT conditions (12) and (13),*
- *and the so-called "Positive Semidefiniteness Condition" (16).*

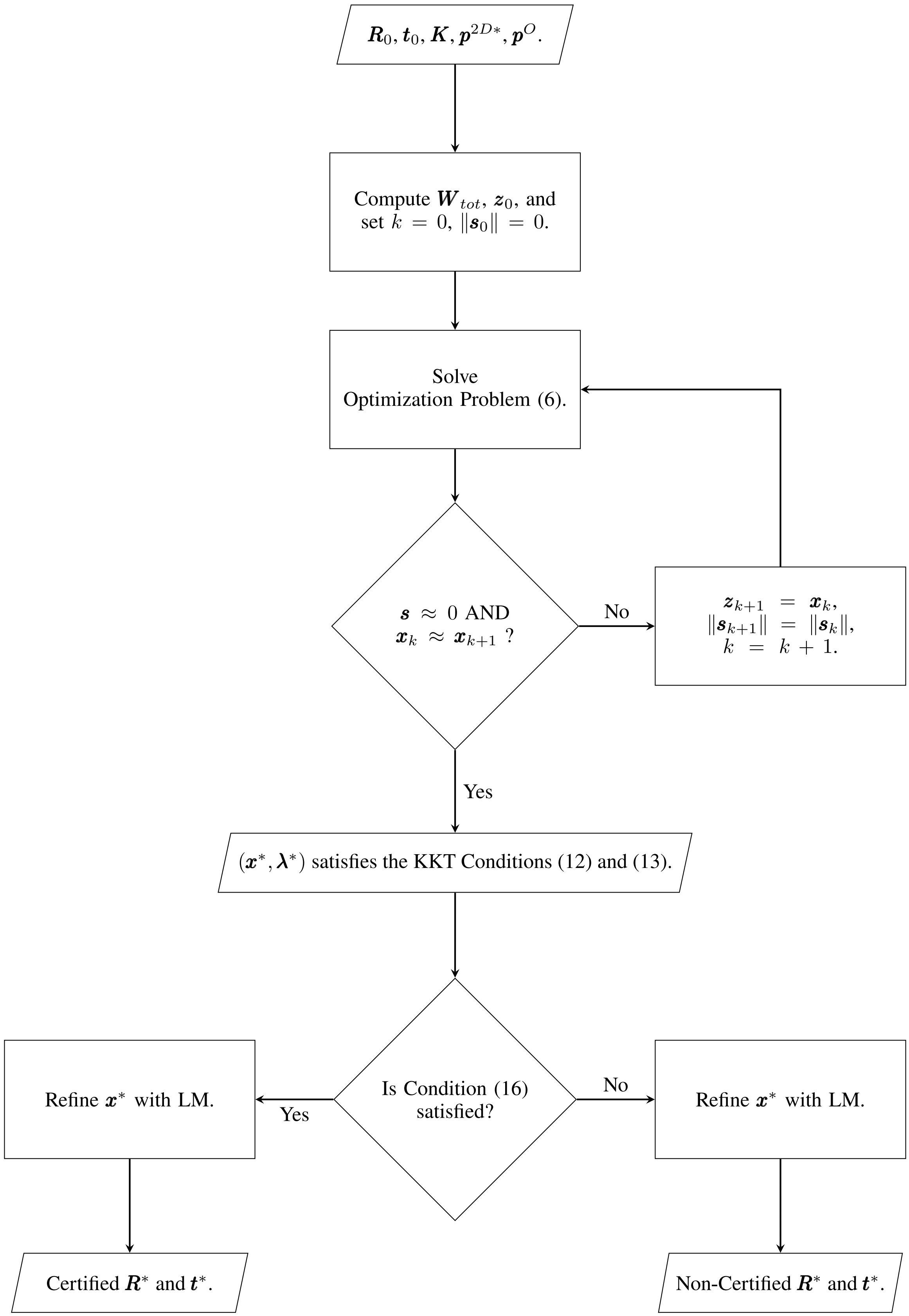


Figure 2: SEC-PnP flowchart.

IV. EXPERIMENTAL RESULTS

The certified optimal results from the SEC-PnP algorithm are compared to the state-of-the-art methods mentioned in Section I using real scenarios in IV.2. For clearness of the result, only selected methods ("classic" PnP solvers and non-linear iterative solvers such as Levenberg-Marquart) are shown in the comparison.

In addition, to validate the optimality of the certified SEC-PnP results, in IV.1 the Cramer-Rao Lower Bound applied to the PnP problem will be compared to the certified SEC-PnP results. The Cramer-Rao Lower Bound offers the theoretical lower bound for the RMSE of any unbiased estimator. This means that if the RMSE of the certified results provided by SEC-PnP matches the theoretical lower bound, then the results certified as global minima are also statistically validated. For the derivation of the general constrained CRLB, one can refer to Stoica (1998).

1. From Simulation

Let us first derive the constrained Cramer - Rao Lower Bound applied to the PnP problem. Starting from Equation (3) with the rotation matrix parametrization ($\boldsymbol{x} \in \mathbb{R}^{12}$) and considering the measurement noise $\boldsymbol{n}$ as a Gaussian variable with $\boldsymbol{\Sigma} = \sigma \boldsymbol{I}_2$, the likelihood function is

$$\mathcal{L}(\boldsymbol{R}_O, \boldsymbol{t}_O; \boldsymbol{p}^{2D*}) = \prod_{i=1}^{N} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\|\boldsymbol{p}_i^{2D*} - \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)\|_{\boldsymbol{\Sigma}}^2\right) \tag{17}$$

which further yields the log-likelihood function

$$\ell(\boldsymbol{R}_O, \boldsymbol{t}_O; \boldsymbol{p}^{2D*}) = N \ln \frac{1}{2\pi\sigma^2} - \sum_{i=1}^{N} \frac{1}{2}\|\boldsymbol{p}_i^{2D*} - \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)\|_{\boldsymbol{\Sigma}}^2. \tag{18}$$

The derivative of $\ell(\boldsymbol{R}_O, \boldsymbol{t}_O; \boldsymbol{p}^{2D*})$ is

$$\frac{\partial \ell}{\partial \boldsymbol{x}^\top} = \sum_{i=1}^{N} (\boldsymbol{p}_i^{2D*} - \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O))^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)}{\partial \boldsymbol{x}^\top}. \tag{19}$$

Hence, after setting $\boldsymbol{\epsilon}_i = \boldsymbol{p}_i^{2D*} - \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)$ the Fisher Information matrix can be calculated as

$$\boldsymbol{F} = \mathbb{E}\left[\frac{\partial \ell}{\partial \boldsymbol{x}} \frac{\partial \ell}{\partial \boldsymbol{x}^\top}\right] = \mathbb{E}\left[\left(\sum_{i=1}^{N} \frac{\partial \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)}{\partial \boldsymbol{x}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\epsilon}_i\right) \left(\sum_{i=1}^{N} \boldsymbol{\epsilon}_i^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)}{\partial \boldsymbol{x}^\top}\right)\right], \tag{20}$$

that finally yields to

$$\boldsymbol{F} = \sum_{i=1}^{N} \frac{\partial \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)}{\partial \boldsymbol{x}} \boldsymbol{\Sigma}^{-1} \frac{\partial \pi_i(\widetilde{\boldsymbol{p}}_i^O, \boldsymbol{R}_O, \boldsymbol{t}_O)}{\partial \boldsymbol{x}^\top}. \tag{21}$$

Note that what we have now derived is the unconstrained Fisher Information matrix. Since there are constraints on the rotation matrix $\boldsymbol{R}_O$, we need to calculate a constrained counterpart $\boldsymbol{F}_c$. We can use the constraint equations in (5) and arrange them in a matrix $\boldsymbol{h}(\boldsymbol{x}) \in \mathbb{R}^{m \times 1}$, setting $\boldsymbol{h}(\boldsymbol{x}) = 0$. Let

$$\boldsymbol{H}(\boldsymbol{x}) = \frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}^\top} \in \mathbb{R}^{6 \times 12}. \tag{22}$$

The gradient matrix $\boldsymbol{H}(\boldsymbol{x})$ has full row rank since the constraints are non-redundant, and hence there exists a matrix $\boldsymbol{U} \in \mathbb{R}^{12 \times 6}$ whose columns form an orthonormal basis for the nullspace of $\boldsymbol{H}(\boldsymbol{x})$, that is $\boldsymbol{H}(\boldsymbol{x})\boldsymbol{U} = \boldsymbol{0}$, where $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}$. Finally, the constrained Fisher Information is given as

$$\boldsymbol{F}_c = \boldsymbol{U}(\boldsymbol{U}^\top \boldsymbol{F}\boldsymbol{U})^{-1}\boldsymbol{U}^\top, \tag{23}$$

and the theoretical lower bound in variance terms is $CRLB = tr(\boldsymbol{F}_c)$.

Now let us compare the CRLB and the certified SEC-PnP results. In simulations, the rotation vector of the camera is set as $[\pi/3, \pi/3, \pi/3]$, and the translation vector is $[2, 6, 6]^\top$. For the intrinsic parameters of the pinhole camera model, the focal length is set as $f_x = f_y = 50$ mm (800 pixels), and the size of the image plane is 640×480 pixels, so we set $c_x = 320$ pixels and $c_y = 240$ pixels. For the 3D feature points under the camera frame, we randomly generate them from the region $[-2, 2] \times [-2, 2] \times [4, 16]$ m. After filtering out the features outside the range of the image plane, the remaining 3D points are projected onto the image plane by the projection equation (1). Specifically, the projection noise $\boldsymbol{n}_i$ is Gaussian noise whose standard deviation is σ pixels. For each fixed σ and number of points N, we execute $T = 2000$ Monte Carlo tests to evaluate the following root-mean-square errors (RMSE) of the SEC-PnP results:

$$RMSE_{\boldsymbol{R}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{R}_{est} - \boldsymbol{R}_{true}\|^2}, \tag{24}$$

$$RMSE_{\boldsymbol{t}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{t}_{est} - \boldsymbol{t}_{true}\|^2}. \tag{25}$$

We set the number of features as $N = 16, 64, 128$ respectively, and for each choice of N the noise intensity varies as $\sigma = 0.1, 1, 5, 10$ pixels. In the following plots (3), the logarithmic Y axis is used. In addition, for each iteration, we consider as t_{init} and r_{init} the ground truths influenced by Gaussian variables with Standard Deviation of $\sigma_t = 5$ m (for each translation coordinate) and Standard Deviation of $\sigma_r = \pi/3$ rad, respectively.

These results (3) suggest that, for any number of 2D - 3D associations N and level of noise in the measurements σ, the accuracy of the certified camera pose estimates provided by SEC-PnP matches the maximum achievable accuracy defined by the CRLB, and consequently we could define ”on-line” the variance of the parameters estimated by SEC-PnP if we know a priori, in some way, the magnitude of noise in the measurements.

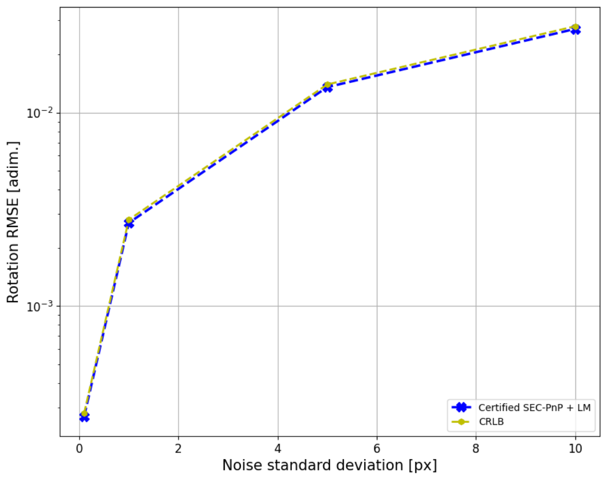
It is also interesting to compare the error distribution for each of the 12 estimated parameters (rotation matrix $\boldsymbol{R}$ coefficients and translation vector $\boldsymbol{t}$) from the SEC-PnP algorithm with the nominal theoretical distribution given by the Cramer Rao Lower Bound. For each subplot in Figures (4) and (5) $T = 2000$ error samples derived from SEC-PnP, the resulting Normal distribution derived from curve fitting and the nominal CRLB distribution are shown. For the sake of brevity, we will only show the results for $N = 64$ features and for $\sigma = 1, 10$ pixels. In the other cases, the error distributions look comparable, as in the case just shown.

It is important to note that, depending on the level of uncertainty in the initial estimate and noise in the measurements, the availability of certified results may vary. We emphasise that the SEC-PnP algorithm always provides a pose estimate, but not all estimates possess the property of global minimum certification. In Table 1 we show the availability of the certified estimates also for $\sigma_t = 3$ m and $\sigma_r = \pi/6$ rad to show that when the initial pose uncertainty decreases, the availability of the certified results increases. For the sake of brevity, we avoid showing the results for accuracy, since they mimic those in Figure (3) with $\sigma_t = 5$ m and $\sigma_r = \pi/3$ rad.

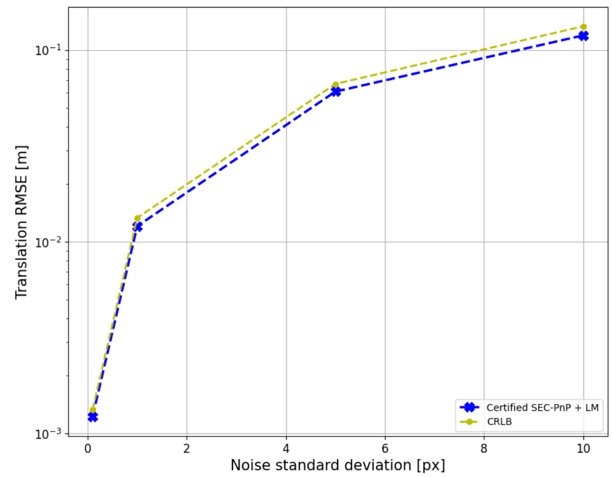
Table 1: Availability [%] of SEC-PnP certified pose estimates for different measurement noise levels and number of features with $T = 2000$, $\sigma_t = 5$ m and $\sigma_r = \pi/3$ rad (on the right) and $T = 2000$, $\sigma_t = 3$ m and $\sigma_r = \pi/6$ rad (on the left).

Points Number	0.1 px	1 px	5 px	10 px
16	92.5	93.1	70.5	18.2
64	99.5	99.4	95.3	25.9
128	100	100	99.2	24.7

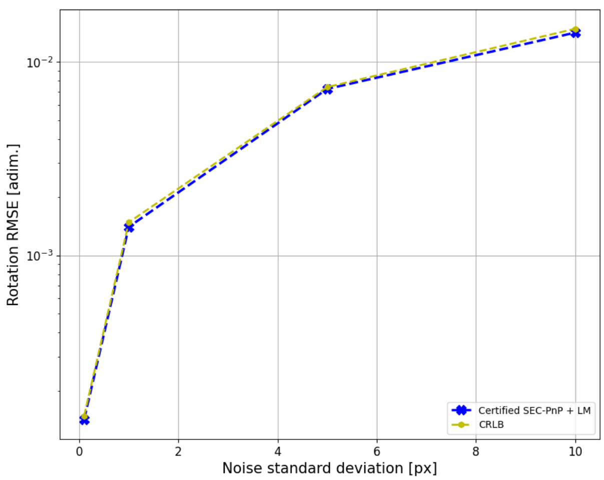
Points Number	0.1 px	1 px	5 px	10 px
16	58.5	57.9	34.5	10.3
64	79.6	82.3	76.3	23.6
128	91.6	90.0	83.8	23.6



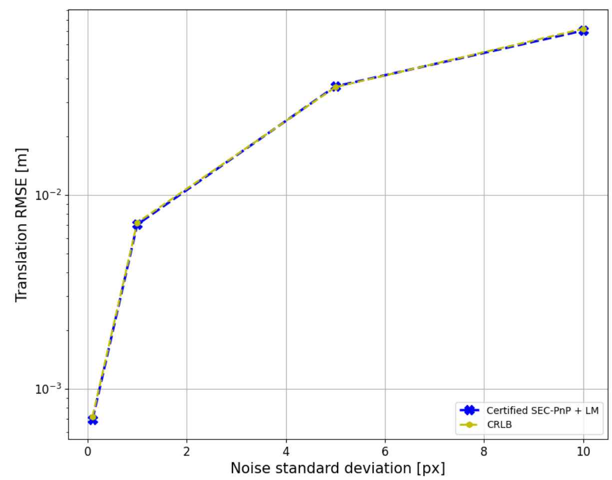
(a) *RMSE rotation for* $N = 16$ *points.*



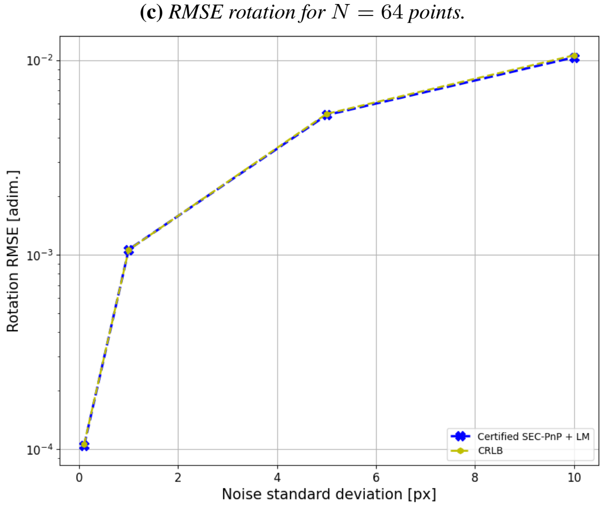
(b) *RMSE translation for* $N = 16$ *points.*



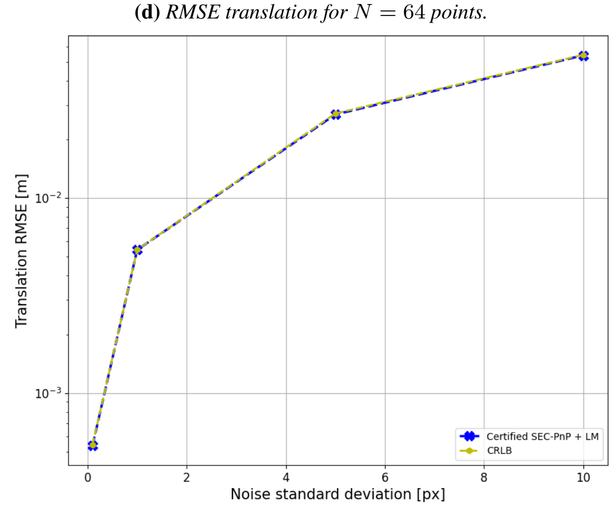
(c) *RMSE rotation for* $N = 64$ *points.*



(d) *RMSE translation for* $N = 64$ *points.*



(e) *RMSE rotation for* $N = 128$ *points.*



(f) *RMSE translation for* $N = 128$ *features.*

Figure 3: Comparison between CRLB and certified SEC-PnP pose estimates.

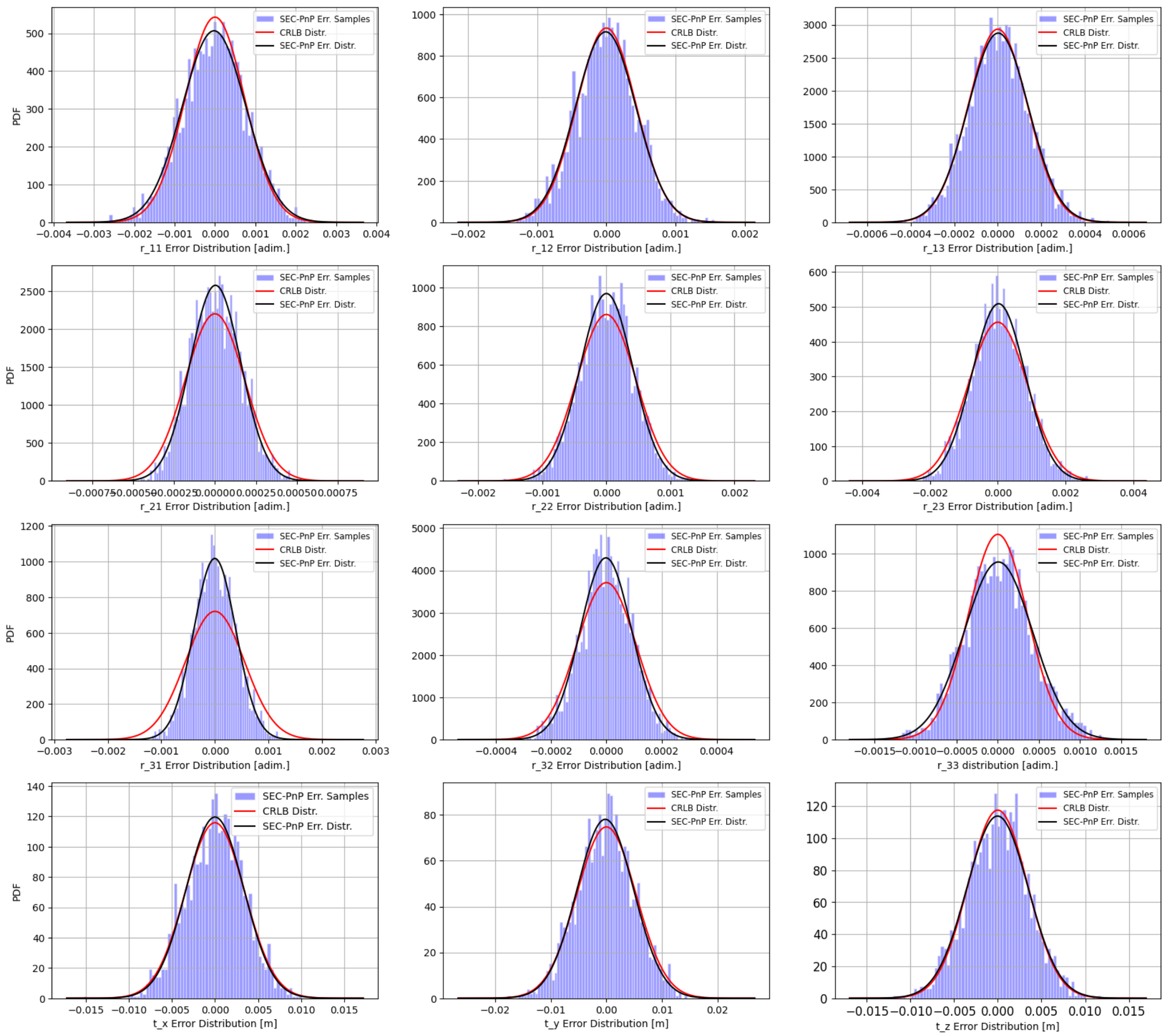


Figure 4: Error distribution for each parameter with $N = 64$ features, 640×480 px camera resolution and $\sigma = 1.0$ px STD measurement error.

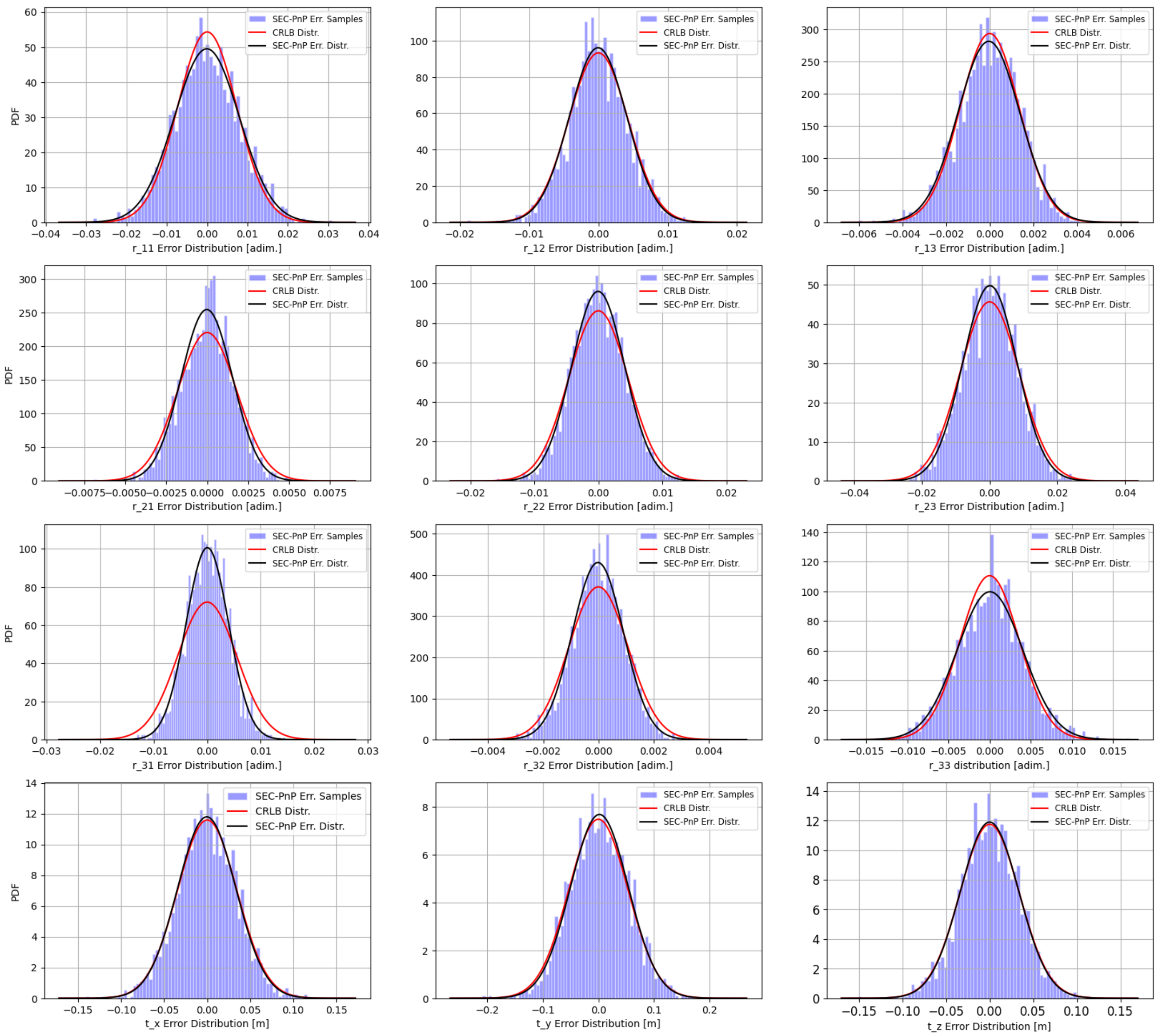


Figure 5: Error distribution for each parameter with $N = 64$ features, 640×480 px camera resolution and $\sigma = 10.0$ px STD measurement error.

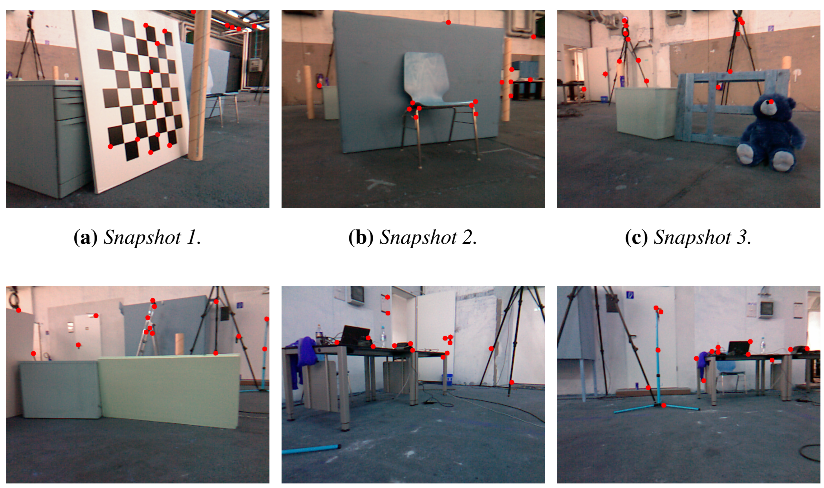
2. From Real-Data

For this test, we used data from the Technical University of Munich (TUM), where a wheeled robot must navigate an enclosed environment (Sturm et al. (2012)). The dataset consists of a sequence of snapshots (1.500 circa), and a point-cloud of 3D points. The latter are extracted from each snapshot using a Shi-Tomasi corner detector (Shi and Tomasi (1994)) provided by OpenCV (Bradski (2000)). The camera calibration matrix $\boldsymbol{K}$ is derived from an actual calibration process and is therefore subject to error. The focal length is approximately $f_x = f_y = 520$ pixels, and the size of the image plane is approximately 650×500 pixels. We noise the associations provided by the corner detector with a Gaussian variable of standard deviation of $\sigma = 8$ pixels combined with a random bias $\mu \in [-a, a]$, with $a = 4$ pixels. We now select 6 snapshots (Figure 6) and repeat the pose estimation process $T = 1000$ times. In addition, for each iteration, we consider as t_{init} and r_{init} the ground truths influenced by Gaussian variables with Standard Deviation of $\sigma_t = 1$ m (for each translation coordinate) and Standard Deviation of $\sigma_r = \pi/6$ rad, respectively. We set the number of features as $N = 50$ and we calculate $RMSE_{\boldsymbol{t}}$ as in (25). The average rotation error $avg(e)_{\boldsymbol{R}}$ will be calculated with

$$avg(e)_{\boldsymbol{R}} = \frac{1}{T} \sum_{t=1}^{T} \| rotvec(\boldsymbol{R}_{est} \boldsymbol{R}_{true}) \|, \tag{26}$$

where the operand $rotvec(\cdot)$ transforms a rotation from its matrix representation to its vector representation. For the results please refer to the Table (2). We denote the Levenberg-Marquardt algorithm as "LM" and the Gauss-Newton algorithm as "GN". In addition, to demonstrate the robustness of the SEC-PnP algorithm to the uncertainty of the initial estimate, we will include the results obtained from the ground truths "GT" (perturbed with error) coupled with Gauss-Newton.

From the results it can be seen that SEC-PnP demonstrates greater robustness and accuracy in each snapshot considered, despite the fact that there is large bias in the measurements provided. However, the EPnP and SQPnP analytical methods still provide relatively acceptable results, while the DLT (Direct Linear Transformation) method coupled with LM and the ground truths coupled with GN frequently provide remarkably inaccurate results.



(a) *Snapshot 1.* **(b)** *Snapshot 2.* **(c)** *Snapshot 3.*

(d) *Snapshot 4.* **(e)** *Snapshot 5.* **(f)** *Snapshot 6.*

Figure 6: Snapshots from the indoor dataset.

Table 2: $RMSE_{\boldsymbol{t}}$ [m] (left) and $avg(e)_{\boldsymbol{R}}$ [rad] (right) with $a = 4$ pixels and $\sigma = 8$ pixels.

Snap. #	SEC-PnP	EPNP	SQPNP	DLT + LM	GT + GN
1	**0.0411**	0.1077	0.0710	1.3697	15.764
2	**0.0641**	0.1161	0.1032	3.2174	1.1527
3	**0.0573**	0.0828	0.0808	0.0575	0.9765
4	**0.0632**	0.1269	0.0938	5.5761	0.6577
5	**0.0817**	0.1110	0.0901	3.0687	0.4189
6	**0.0668**	0.1178	0.0900	0.0861	1.1066

Snap. #	SEC-PnP	EPNP	SQPNP	DLT + LM	GT + GN
1	**0.0118**	0.0141	0.0136	0.1060	0.1060
2	**0.0110**	0.0131	0.0131	0.0644	0.0644
3	**0.0104**	0.0125	0.0127	0.0482	0.4820
4	**0.0127**	0.0185	0.0165	0.0352	0.0352
5	**0.0172**	0.0218	0.0188	0.0274	0.0274
6	**0.0104**	0.0148	0.0122	0.0563	0.0563

V. CONCLUSION AND FUTURE OUTLOOK

The SEC-PnP algorithm and the methods introduced in Section III to prove its global optimality, could be employed in applications with high reliability and certifiability requirements. The most significant contribution of the proposed work is the possible introduction of an integrity description for SEC-PnP. Quantifiable integrity is available in many other navigation technologies, e.g., GNSS based positioning, but it has been a great challenge for the state-of-the-art camera-based navigation algorithms because of the numerous uncertainty sources in the camera measurements and initial pose estimate described in Zhu et al. (2022).

Looking at Figure (7) we can see how the pose error is downstream from the other sources of error. This makes us think that if all other influences are limited or even nearly absent, the pose error could still make the estimation completely wrong, and it is in our interest to make sure that in the last stage there are no additional sources of uncertainty caused by convergence to a local minimum, and that the absolute minimum of the PnP problem is found.

Reflecting on this issue, we can think of the certification system of the SEC-PnP algorithm in Theorem 2 as an ”Integrity Monitor” of the convergence phase to the global minimum that, provided we know the side conditions such as the noise intensity in the measurements, can also provide a nominal distribution of the pose uncertainty based on equality with the Cramer Rao Lower Bound applied to the PnP problem, as demonstrated in Section IV. The equality of the error distribution provided by SEC-PnP with the CRLB not only allows us to have a nominal distribution, but also ensures that our pose estimator does not have any bias and provides the highest possible accuracy. Having a ”Monitor Test” that provides us with a nominal distribution of the pose error also allows us, as future work, to define protection levels, if we also can compute the corresponding integrity budget by knowing the probability of a fault in the analytical certification defined in Theorem 2. The developed algorithm paired with the conditions listed in this work will be a fundamental step in securing the visual navigation integrity in the future.

As already shown in Triolo et al. (2024), the possibility of certifiable results and better accuracy than state-of-the-art algorithms comes at the cost of longer computational times. As a future development, one of the priorities will be to optimize the algorithm respect its current state (the computational time is about 0.5 s regardless of the number of 2D-3D associations, with an 11-th generation 4-core Intel(R), and Python as the language used to write the algorithm code) in order to extend its use for applications with more stringent time requirements.

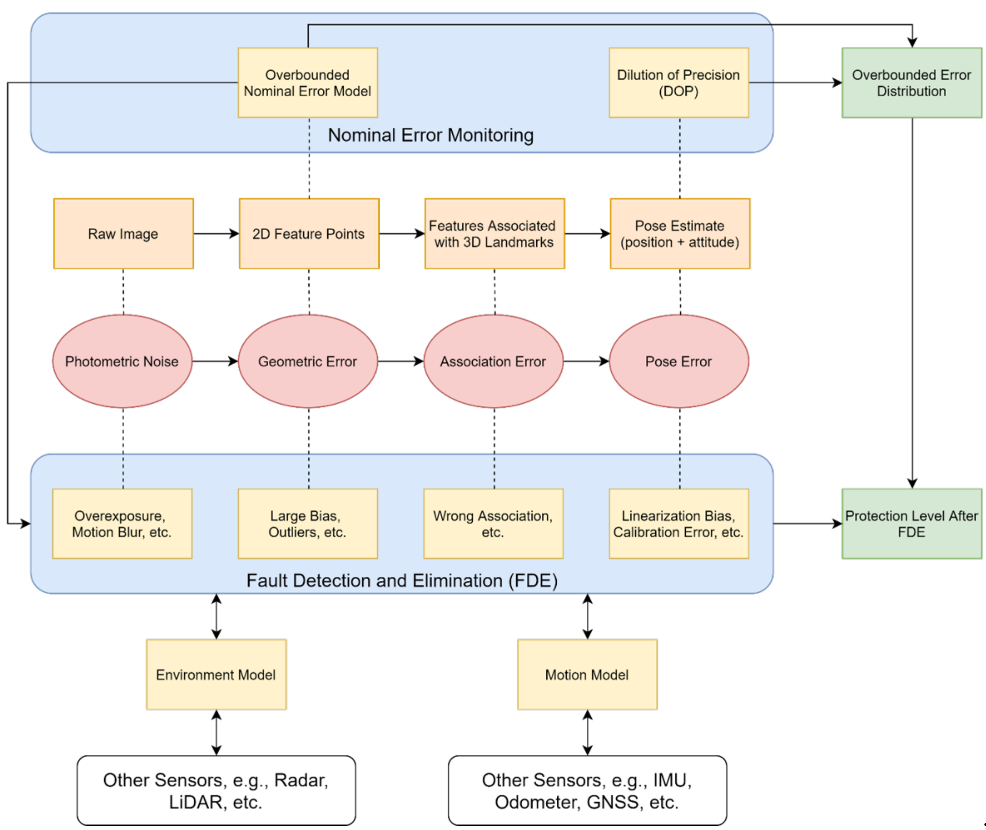


Figure 7: Integrity framework for visual positioning.

REFERENCES

Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. (2005). *Nonlinear Programming: Theory and Algorithms*. Wiley.

Beck, A., Ben-Tal, A., and Tetruashvili, L. (2010). A sequential parametric convex approximation method with applications to nonconvex truss topology design problems. *Journal of Global Optimization*, 47:29–51.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambdridge.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Eade, E. (2013). Gauss-newton /levenberg-marquardt optimization.

Gao, D. (2008). Advances in canonical duality theory with applications to global optimization. *Proceedings of the Fifth International Conference on Foundations of Computer-Aided Process Operation*.

Gao, X.-S., Hou, X.-R., Tang, J., and Cheng, H.-F. (2003). Complete solution classification for the perspective-three-point problem. *IEEE Transactions on pattern analysis and Machine Intelligence*, 25(8).

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambdridge.

Lepetit, V., Moreno-Noguer, F., and Fua, P. (2009). Epnp: An accurate o(n) solution to the pnp problem. *International Journal of Computer Vision*, 81:155–166.

Lu, C. and Fang, S.-C. (2011). Kkt solution and conic relaxation for solving quadratically constrained quadratic programming problems. *SIAM J.OPTIMATION*, 21(4):1475–1490.

Mehanna, O., Huang, K., Gopalakrishnan, B., Konar, A., and Sidiropoulos, N. D. (2015). Feasible point pursuit and successive approximation of non-convex qcqps. *IEEE Signal Processing Letters*, 22(7).

Shi, J. and Tomasi (1994). Good features to track. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.

Stoica, P. (1998). On the cramer-rao bound under parametric constraints. *IEEE Signal Processing Letters*.

Sturm, J., Engelhard, N., Endres, F., Burgard, W., and Cremers, D. (2012). A benchmark for the evaluation of rgb-d slam systems. In *Proc. of the International Conference on Intelligent Robot Systems (IROS)*.

Terzakis, G. and Lourakis, M. (2020). A consistently fast and globally optimal solution to the perspective-n-point problem. *European Conference on Computer Vision*.

Triolo, A., Zhu, C., and Meurer, M. (2024). Reliable camera-based positioning robust to initial pose error. *ION – ITM Conference Proceedings*.

Yeyakumar, V., Rubinov, A., and Wu, Z. (2007). Non-convex quadratic minimization problems with quadratic constraints: Global optimality conditions. *Math. Program.*

Zeng, G., Mu, B., Shi, L., and Chen, J. (2024). Consistent and asymptotically efficient localization from range difference measurements. *IEEE TRANSACTIONS ON INFORMATION THEORY*.

Zhu, C., Meurer, M., and Günther, C. (2022). Integrity of visual navigation —developments, challenges, and prospects. *Journal of the Institute of Navigation*, 69(2).