Towards Global Optimality in Stereo Localization

by

Ben Agro

Supervisor: Timothy D Barfoot May 5, 2023

B.A.Sc. Thesis



Towards Global Optimality in Stereo Localization

by

Ben Agro

Supervisor: Timothy D Barfoot May 5, 2023 **Abstract:** There is a growing demand for robotics systems in safety-critical applications, including transportation, policing, defense, medicine, and home care. These systems must reliably and efficiently estimate their state and that of the environment by solving an optimization problem. Many contemporary state estimation systems employ local algorithms to solve these optimization problems, which may return sub-optimal estimates called *local minima*. This can have dangerous consequences, such as a self-driving car thinking it is in the wrong lane.

Recent work in *certifable algorithms* has attempted to solve this problem by developing methods that can find a globally optimal solution or prove that a solution is globally optimal. Certifiable algorithms are important for detecting autonomy stack failures and increasing the robustness of safety-critical systems. Many problems in robotics have been certified, including rotation averaging, pose-graph optimization, multiple point-cloud registration, landmark-based simultaneous localization and mapping (SLAM), calibration, and image segmentation.

In this work, we attempt to find a globally optimal algorithm for the problem of stereo localization with re-projection error. This problem involves estimating the pose of a stereo camera given known landmarks in the world by minimizing matrix-weighted pixel-space errors. Prior work has yet to develop a certifiable algorithm for this problem because of its non-polynomial cost function. This certificate, or globally optimal solver, would have applications in visual odometry, localization, SLAM, and all the many other autonomy systems that solve this optimization problem. Further, the techniques developed to find a globally optimal algorithm for this problem may help develop globally optimal algorithms for other non-polynomial optimization problems.

The code for this project is available here: https://github.com/BenAgro314/bagro_engsci_thesis

Acknowledgements: In completing my undergraduate thesis, I express my deepest gratitude to several individuals who have supported, inspired, and challenged me throughout this journey. First and foremost, I want to thank Prof. Timothy D. Barfoot for being my first academic mentor, believing in me as a first-year student, and providing unwavering inspiration and guidance during my thesis work. I am also grateful to Connor Holmes and Frederike Dumbgen for their collaboration and contributions, which enriched the project and provided me with a snapshot of graduate school work. Prof. Florian Shkurti deserves special mention for his support and for opening up numerous opportunities that have shaped my academic path. I sincerely appreciate my friends Arthur Allshire, Shrey Jain, Rassam Yazdi, Eric Liu, Jason Liu, Ritvik Singh, and Camille Bruckman for their camaraderie and inspiration during our time in Engineering Science. Lastly, I would like to thank my family—Karen, Paul, Rachel, and Jacquie—for their unwavering love and support throughout this endeavor.

Contents

1	Intr	oduction	1
2	Bacl	kground	3
	2.1	Theory	3
		2.1.1 Localization	3
		2.1.2 Stereo Camera Model	3
		2.1.3 Optimization	5
		2.1.4 Least-Squares Pose Optimization	6
		2.1.5 Stereo Localization Optimization Problem	7
		2.1.6 Lagrangian Duality Theory	8
		2.1.7 QCQP	11
	2.2	Related Work	13
		2.2.1 Research Gap	16
3 Methods and Results			
	3.1	Datasets	18
	3.2	Stereo Localization as a QCQP	19
	3.3	Tightness of the Stereo Localization QCQP	22
	3.4	Redundant Constraints	24
		3.4.1 Tightness with Redundant Constraints	27
	3.5	Iterative SDP	27
	3.6	Solver Comparison	30
		3.6.1 Global Optimality	30
		3.6.2 Solver Time Efficiency	31
	3.7	Visualizing Solution Trajectories	32
4	Cha	llenges and Future Work	34
	4.1	1D Problem	34

	4.2	2D Problem	36	
5	Con	clusion	39	
Aŗ	opend	ices	46	
A	A Back Projection Error Tightness			
B	1D p	problem	48	
	B .1	QCQP Formulation	48	
	B.2	Local Solver	49	
C	2D J	oroblem	50	
	C .1	QCQP Formulation	50	
	C .2	Local Solver	51	

List of Figures

2.1	An illustration of the localization problem. We want to find the rigid transfor-	
	mation T between $\vec{\mathcal{F}}_w$ and $\vec{\mathcal{F}}_s$ given sensor measurements of the landmarks \mathbf{p}_i .	4
2.2	A depiction of the stereo camera rig	4
2.3	An example problem demonstrating a local minimum in the stereo localization	
	problem. The green dots are the landmarks, the blue frame is the global mini-	
	mum, and the red frame is a local minima	9
2.4	An illustration of Lagrangian duality in the cases of weak and strong duality,	
	assuming Slater's constraint qualifications. Note the curves are for illustration	
	purposes only. Under weak duality, we cannot certify the problem nor solve the	
	dual of the dual for the globally optimal solution. If strong duality holds, the	
	dual problem provides us with a means to certify solutions, and the dual of the	
	dual is a tight convex relaxation of the primal problem.	10
2.5	A summary of duality theory applied to a quadtratically constrained quadratic	
	program (QCQP)	13

2.6	An illustration of stereo camera triangulation (the area in blue) uncertainty and	
	its Gaussian approximation (the ellipse in red), adapted from [36]. Observe that	
	the Gaussian approximation does support the long tail of the true distribution	17
3.1	STEREOSIM: Left is a depiction of the landmark points, the sensor frame \mathcal{F}_s ,	
	and the world frame \mathcal{F}_w .	19
3.2	A visualization of the camera trajectory and the landmark positions in the STAR-	
	RYNIGHT dataset.	19
3.3	Each dot represents one solution from GAUSSNEWTON. The blue dots are glob-	
	ally optimal solutions, and the red dots are local minima. We observe that while	
	there is a gap between the minimum eigenvalues of $\mathbf{H}(\boldsymbol{\lambda})$ for the local and global	
	minima, this gap decreases as the noise level increases. This result indicates that	
	strong duality does not hold in practice.	23
3.4	A plot of the duality gap of the primal relaxation. The plot on the left has no	
	redundant constraints, while the plot on the right includes all of our redundant	
	constraints. Notice that in both cases, the gap is non-zero for all problems indi-	
	cating that in both cases, the primal relaxation is not tight.	27
3.5	An example problem where the iterative SDP method does not find the globally	
	optimal solution. In this case, one landmark is at a significantly larger depth	
	than the others.	29
3.6	For many problem instances, this figure plots the gap between the global mini-	
	mum cost p^* and the cost obtained from the solution found by the iterative SDP	
	algorithm denoted with q^* .	30
3.7	A comparison of the localization methods measuring the percentage of globally	
	optimal solutions on StereoSim at various pixel-space noise levels.	31
3.8	A comparison of the localization methods measuring the percentage of globally	
	optimal solution on StereoSim at various pixel-space noise levels	32
3.9	Solver runtime as a function of the number of landmarks.	33
3.10	Visualizing the solution trajectories of the GAUSSNEWTON, ITERSDP, and PRI-	
	MALRELAX. The solver pose at the i^{th} iteration is labeled \mathcal{F}_i . PRIMALRELAX	
	has more than one pose in its trajectory because the output from the primal	
	relaxation is refined with GAUSSNEWTON.	33
4.1	1D problem	34

- 4.2 Investigating tightness of the 1D problem. The blue dashed line denotes the global minimum cost for a single problem instance, and the bars indicate the cost of the primal relaxation both with and without redundant constraints. . . . 35
- 4.4 Investigating tightness of the 2D problem. The blue dashed line denotes the global minimum cost for a single problem instance, and the bars indicate the cost found by the primal relaxation both with and without redundant constraints. 37

List of Tables

3.1	Average solver runtime on StarryNight	32		
4.1	1 Comparing the computational complexity of the primal relaxation with and			
	without the Lasserre hierarchy. All problems in both 2D and 3D used three			
	landmarks	38		

Acronyms and Abbreviations

MPCR multiple point cloud registration. 1, 14

- PGO pose graph optimization. 1, 14
- PnP Perspective-n-Point. 15, 27
- **QCQP** quadtratically constrained quadratic program. v, 8, 11, 13–16, 19, 21–24, 26, 31, 34–36, 39, 47, 48, 50, 51
- **RA** rotation averaging. 1, 14
- **SDP** semi-definite program. 12, 14, 28, 30, 37, 39
- SLAM simultaneous localization and mapping. 1, 14, 16, 39
- VO visual odometry. 16

Notation

- *a* This typeface is used for real scalars
- a This typeface is used for real column vectors
- A This typeface is used for real matrices
- I The identity matrix
- 0 The zero matrix
- $A \ge 0$ denotes that the matrix A is positive-semi definite

 $\mathbf{a} \ge \mathbf{0}$ denotes that the elements of \mathbf{a} are all non-negative

 $\mathbb{R}^{M \times N}$ The vector space of real $M \times N$ matrices

- $\vec{\mathcal{F}}_a$ A vectrix representing a reference frame in three dimensions
- O(3) The orthogonal group
- SO(3) The special orthogonal group, used to represent orientations and rotations
- SE(3) The special Euclidean group, used to represent poses and rigid transformations
- \mathbf{C}_{ba} A 3 × 3 rotation matrix that takes points expressed in $\vec{\mathcal{F}}_a$ and reexpresses them in $\vec{\mathcal{F}}_b$, which is rotated with respect to $\vec{\mathcal{F}}_a$
- \mathbf{T}_{ba} A 4 × 4 rotation matrix that takes points expressed in $\vec{\mathcal{F}}_a$ and reexpresses them in $\vec{\mathcal{F}}_b$, which is rotated and/or translated with respect to $\vec{\mathcal{F}}_a$
- **T** A shorthand notation expressing \mathbf{T}_{sw} which is the transformation matrix that takes points expressed in the world frame $\vec{\mathcal{F}}_w$ and reexpresses them in the sensor frame $\vec{\mathcal{F}}_s$
- \mathbf{e}_i Denotes the i^{th} column of \mathbf{I}

1 Introduction

Robotic state estimation involves understanding the physical environment and the robot's state through sensor input [1]. The algorithms underlying perception and state estimation will participate in increasingly important — *and safety critical* — roles in society, e.g., autonomous driving technology and space robotics. Typically these algorithms have to solve an *optimization problem*, where an objective function is minimized with respect to some state variables, subject to a set of constraints [2]. These optimization problems are typically highly complex with many *local minima*; solutions that are optimal within a local region of the state-variable space but may not be the globally optimal solution (the global minimum) [3]. Most optimization algorithms are solved using efficient local search methods (e.g., gradient descent) and may return poor local minima [4]. In real-world robotics, getting stuck in a local minimum can have disastrous consequences, e.g., an autonomous car thinking it is in the wrong lane [3].

While local search methods are not guaranteed to find the globally optimal solution, contemporary *certifiably optimal algorithms* can determine if the local solution is also the global solution [5]. This *certificate* of optimality allows the robot to make informed decisions based on its confidence in the optimality of the state estimate. Prior works have developed certifiably optimal algorithms for many problems in robotics, including rotation averaging (RA) [6]–[9], pose graph optimization (PGO) [10]–[12], multiple point cloud registration (MPCR) [13], [14], simultaneous localization and mapping (SLAM) [4], robust estimation [15]–[21], extrinsic calibration [22]–[24], and segmentation [25], to name a few. As of the time of writing, there is no general method for developing a certifiable algorithm, nor is there a method for determining if a given optimization problem admits a certificate. Each problem is tackled individually, usually drawing on previous techniques from the literature [5].

Localization is the problem of, given known landmarks in the world and sensor observations of those landmarks, determining where the sensor is in the world [1]. In the *stereo localization problem*, a stereo camera is used; a sensor with two or more lenses and a separate image sensor for each lens [1]. No one has yet developed a true certifiable algorithm for stereo localization

minimizing pixel-space errors, an important problem given the prevalence of this algorithm in robotic systems. Further, developing more certifiable algorithms will add to the 'cookbook' of techniques applicable to similar problems in the future [5].

This project aims to develop methods to achieve or certify global optimality in the stereo localization problem. Our goals are to (i) develop algorithms that solve the stereo localization problem with guaranteed global optimality and (ii) develop and test an efficient certificate of the stereo localization problem.

This document details our progress towards certifiably optimal stereo localization. We provide the necessary theoretical background to understand the problem and our approach, discuss related work, and identify the research gap. We describe and derive our proposed methods, assess them experimentally, and draw insights from them. Finally, we highlight future research opportunities building on our work.

2 Background

This section provides the theory and related work necessary to understand our method and the research gap.

2.1 Theory

2.1.1 Localization

The localization problem involves determining the pose (position and orientation) of a sensor frame $\vec{\mathcal{F}}_s$ with respect to the world frame $\vec{\mathcal{F}}_w$ from sensor measurements of features in the world with a known position (e.g., a known map) [1]. Refer to fig. 2.1 for an illustration of the localization problem. Formally, given the positions of N points in world coordinates and their corresponding sensor measurements, the localization problem seeks the transformation matrix that re-expresses points in $\vec{\mathcal{F}}_w$ in $\vec{\mathcal{F}}_s$:

$$\mathbf{T} = \mathbf{T}_{sw} = \begin{bmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0}^T & 1 \end{bmatrix} \in SE(3), \tag{2.1}$$

where $\mathbf{C} \in SO(3)$ represents the rotation and $\mathbf{r} \in \mathbb{R}^3$ the translation.

2.1.2 Stereo Camera Model

A stereo camera consists of two cameras rigidly attached to one another with a known transformation between them [1]. The difference in pixel coordinates of a feature observed by both cameras can be used to estimate the depth of that feature relative to the camera. See fig. 2.2 for a schematic of the stereo camera rig.

In the stereo camera model, the observation of a point takes the form of a pair of pixel



Figure 2.1: An illustration of the localization problem. We want to find the rigid transformation T between $\vec{\mathcal{F}}_w$ and $\vec{\mathcal{F}}_s$ given sensor measurements of the landmarks \mathbf{p}_i .



Figure 2.2: A depiction of the stereo camera rig.

measurements in each camera, which we stack and express as

$$\mathbf{y} = \begin{bmatrix} u_{\ell} \\ v_{\ell} \\ u_{r} \\ v_{r} \end{bmatrix}, \qquad (2.2)$$

which are illustrated in fig. 2.2. We locate the sensor frame at the midpoints between the two cameras. For simplicity, we assume the two cameras are identical and in a fronto-parallel configuration, depicted in fig. 2.2. The *baseline*, *b*, is the distance between the origin of $\vec{\mathcal{F}}_{\ell}$ and $\vec{\mathcal{F}}_r$ (the two optical centers). The *focal length*, *f*, of the camera is the distance from the optical center to the image plane. Let f_u and f_v denote the *focal length* of the camera expressed in row and column pixel coordinates, respectively. Let c_u and c_v denote the pixel coordinates of the optical center of the camera sprojected onto their respective image planes. Following [1], y is related to a point \mathbf{p}_s (expressed in $\vec{\mathcal{F}}_s$) as:

$$\mathbf{y} = \mathbf{M} \frac{1}{\mathbf{e}_3^T \mathbf{p}_s} \mathbf{p}_s, \tag{2.3}$$

where **M** is the intrinsic parameter matrix, and $\mathbf{e}_i \in \mathbb{R}^4$ is the i^{th} column of the 4×4 identity matrix. We notice that $\mathbf{p}_s/(\mathbf{e}_3^T \mathbf{p}_s)$ is the projection of the point onto the plane $z_s = 1$. With basic trigonometry rules, one can derive the intrinsic parameter matrix of the stereo camera:

$$\mathbf{M} = \begin{bmatrix} f_u & 0 & c_u & f_u \frac{b}{2} \\ 0 & f_v & c_v & 0 \\ f_u & 0 & c_u & -f_u \frac{b}{2} \\ 0 & f_v & c_v & 0 \end{bmatrix}.$$
 (2.4)

If the point is expressed in $\vec{\mathcal{F}}_w$ as \mathbf{p}_w , we can write the forward stereo camera model as:

$$\mathbf{y} = \mathbf{M} \frac{1}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_w} \mathbf{T} \mathbf{p}_w.$$
(2.5)

2.1.3 Optimization

Localization methods are often formalized as maximum likelihood estimation problems (maximizing the likelihood of the data given unknown pose parameters), which rely on solving an underlying optimization problem [12]. The general form of an optimization problem is

$$\min_{\mathbf{x}} \qquad f_0(\mathbf{x}) \tag{2.6a}$$

subject to
$$f_i(\mathbf{x}) = 0, \quad \forall i \in \{1, \dots, K\}$$
 (2.6b)

$$h_i(\mathbf{x}) \le 0, \quad \forall i \in \{1, \dots, P\},$$

$$(2.6c)$$

where the set of x satisfying the constraints, \mathcal{D} , is nonempty [2]. Such a problem is called *convex* if its objective function f_0 and constraint set \mathcal{D} are *convex*. Convex problems have the desirable property that every local minimum is also a global minimum. They, thus, are far easier to solve than non-convex optimization problems, which may have multiple local minima.

2.1.4 Least-Squares Pose Optimization

By modeling forward sensor models with Gaussian distributions, many maximum-likelihood localization methods simplify to a non-linear least-squares optimization problem [26] [1]:

$$\min_{\mathbf{T}} \qquad J(\mathbf{T}) = \sum_{n} \mathbf{g}_{n} (\mathbf{T} \mathbf{p}_{n})^{T} \mathbf{W}_{n} \mathbf{g}_{n} (\mathbf{T} \mathbf{p}_{n})$$
(2.7a)

subject to
$$\mathbf{T} \in SE(3)$$
 (2.7b)

where $g_n(\cdot)$ are non-linear error functions, and W_n are symmetric weight matrices, usually taken as the inverse of the measurement covariance matrix. This section outlines the necessary background for local optimization methods of $J(\mathbf{T})$ subject to $\mathbf{T} \in SE(3)$.

First, we choose an unconstrained parameterization of **T** to make eq. (2.7) an unconstrained optimization problem. Following [1], a convenient parameterization uses the exponential map between SE(3) and its Lie algebra of: $\mathbf{T} = \exp(\boldsymbol{\xi}^{\wedge})$, where $\boldsymbol{\xi} \in \mathbf{R}^{6\times 1}$. The definition of the $(\cdot)^{\wedge}$ operator is

$$\boldsymbol{\xi}^{\wedge} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\psi} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\psi}^{\times} & \boldsymbol{\rho} \\ \mathbf{0}^{T} & 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} u \\ v \\ w \end{bmatrix}^{\times} = \begin{bmatrix} 0 & -w & v \\ w & 0 & -u \\ -v & u & 0 \end{bmatrix}, \quad (2.8)$$

where $\psi, \rho \in \mathbb{R}^{3 \times 1}$. Crucially, selecting such a parameterization turns our constrained optimization problem over the entries of **T** into an unconstrained problem over ξ .

Next, starting with an initial guess or the pose estimate from the previous iteration, \mathbf{T} , we perturb this guess on the left:

$$\mathbf{T} \leftarrow \exp(\boldsymbol{\epsilon}^{\wedge})\mathbf{T},$$
 (2.9)

where ϵ is the perturbation. Then we linearize g about ϵ :

$$\mathbf{g}_n(\exp(\boldsymbol{\epsilon})\mathbf{T}\mathbf{p}_n) \approx \mathbf{g}_n((1+\boldsymbol{\epsilon}^{\wedge})\mathbf{T}\mathbf{p}_n)$$
 (2.10)

$$\approx \mathbf{g}_n(\mathbf{T}\mathbf{p}_n) + \frac{\partial \mathbf{g}_n}{\partial(\mathbf{T}\mathbf{p}_n)} \boldsymbol{\epsilon}^{\wedge}(\mathbf{T}\mathbf{p}_n).$$
(2.11)

Now we use the identity

$$\boldsymbol{\xi}^{\wedge} \mathbf{q} = \mathbf{q}^{\odot} \boldsymbol{\xi}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^{6 \times 1}, \mathbf{q} \in \mathbb{R}^{4 \times 1},$$
(2.12)

where

$$\begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\eta} \end{bmatrix}^{\circ} = \begin{bmatrix} \eta \mathbf{I} & -\boldsymbol{\zeta}^{\times} \\ \mathbf{0}^{T} & \mathbf{0}^{T} \end{bmatrix}, \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^{3 \times 1}, \boldsymbol{\eta} \in \mathbb{R},$$
(2.13)

to rearrange:

$$\mathbf{g}_{n}(\exp(\boldsymbol{\epsilon})\mathbf{T}\mathbf{p}_{n}) \approx \mathbf{g}_{n}(\mathbf{T}\mathbf{p}_{n}) + \frac{\partial \mathbf{g}_{n}}{\partial(\mathbf{T}\mathbf{p}_{n})}(\mathbf{T}_{op}\mathbf{p}_{n})^{\odot}\boldsymbol{\epsilon}$$
(2.14)

$$=\boldsymbol{\beta}_n + \boldsymbol{\Delta}_n^T \boldsymbol{\epsilon}. \tag{2.15}$$

Then, the objective to be minimized at each iteration with respect to our perturbation parameter ϵ is

$$J(\boldsymbol{\epsilon}) = \sum_{n} (\boldsymbol{\beta}_{n} + \boldsymbol{\Delta}_{n}^{T} \boldsymbol{\epsilon})^{T} \mathbf{W}_{n} (\boldsymbol{\beta}_{n} + \boldsymbol{\Delta}_{n}^{T} \boldsymbol{\epsilon}).$$
(2.16)

We can differentiate with respect to ϵ to solve for the optimal update parameters ϵ^* :

$$\frac{\partial J}{\partial \boldsymbol{\epsilon}^{T}} = 2 \sum_{n} \boldsymbol{\Delta}_{n} \mathbf{W}_{n} (\boldsymbol{\beta}_{n} + \boldsymbol{\Delta}_{n}^{T} \boldsymbol{\epsilon}) = \mathbf{0}$$
(2.17)

$$\implies -\left(\sum_{n} \Delta_{n} \mathbf{W}_{n} \boldsymbol{\beta}_{n}\right) = \left(\sum_{n} \Delta_{n} \Delta_{n}^{T}\right) \boldsymbol{\epsilon}^{*}, \qquad (2.18)$$

which can be solved for the (locally) optimal ϵ^* . Finally, we update our pose estimate $\mathbf{T} \leftarrow \exp(\epsilon^{*\wedge}))\mathbf{T}$, and iterate. This will serve as a baseline method used for the comparison against the localization algorithms we develop. Further, while time-consuming, running this local solver for many different initial conditions and taking the best solution provides a likely means to find a globally optimal solution.

2.1.5 Stereo Localization Optimization Problem

With the stereo camera model and least-squares pose optimization in hand, we can introduce and solve stereo localization as a local optimization problem. Our optimization problem for stereo localization minimizes the *re-projection error*:

$$\min_{\mathbf{T}} \qquad J(\mathbf{T}) = \sum_{n=1}^{N} \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right)^T \mathbf{W}_n \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right) \quad (2.19a)$$

subject to $\mathbf{T} \in SE(3)$ (2.19b)

where:

- {p_n ∈ ℝ^{4×1} | ∀n ∈ {1,...,N}} are the known homogeneous coordinates of the land-mark points in the world,
- {y_n ∈ ℝ^{4×1} | ∀n ∈ {1,...,N}} are the known and noisy stereo camera measurements of the landmarks (eq. (2.5)),
- $\mathbf{M} \in \mathbb{R}^{4 \times 4}$ is the intrinsic stereo camera matrix,
- $\{\mathbf{W}_n \in \mathbb{R}^{4 \times 4} \mid \forall n \in \{1, \dots, N\}\}$ are known/chosen weight matrices,
- and T is the unknown world-to-sensor frame coordinate transformation matrix (eq. (2.1)).

This optimization problem seeks to minimize the error in *pixel space*, i.e., find the pose such that the pixel coordinates of the landmarks align with the measurements. The measurement error in pixel space is a Gaussian. Thus, W_n is usually taken of the inverse of the covariance matrix such that this optimization is a maximum-likelihood estimation problem.

Gauss-Newton Method For Stereo Localization

From our optimization problem in eq. (2.19), we can see

$$\mathbf{g}_{n}(\mathbf{T}\mathbf{p}_{n}) = \mathbf{g}_{n}(\exp\left(\boldsymbol{\epsilon}^{\wedge}\right)\mathbf{T}_{op}\mathbf{p}_{n}) = \mathbf{y}_{n} - \mathbf{M}\frac{1}{\mathbf{e}_{3}^{T}\mathbf{T}\mathbf{p}_{n}}\mathbf{T}\mathbf{p}_{n},$$
(2.20)

with

$$\frac{\partial \mathbf{g}_n}{\partial (\mathbf{T}\mathbf{p}_n)} = \frac{-1}{\mathbf{e}_3^T \mathbf{T}\mathbf{p}_n} \mathbf{M} + \frac{1}{\left(\mathbf{e}_3^T \mathbf{T}\mathbf{p}_n\right)^2} \mathbf{M} \mathbf{T}\mathbf{p}_n \mathbf{e}_3^T.$$
(2.21)

We have specified g_n and its derivative, meaning we can use the local optimization algorithm described in section 2.1.4.

This local optimization method is prone to local minima. We demonstrate this with an example problem, shown in fig. 2.3. We can see that the solution from the Gauss-Newton method in red is far from the global minima in blue. This motivates certifiable methods for the stereo localization problem.

2.1.6 Lagrangian Duality Theory

This section describes Lagrangian duality theory, the basis for certifiably optimal algorithms. As an example, we apply it to a QCQP, the results of which will be used in our methods.



Figure 2.3: An example problem demonstrating a local minimum in the stereo localization problem. The green dots are the landmarks, the blue frame is the global minimum, and the red frame is a local minima.

Consider the general optimization problem in eq. (2.6). The Lagrangian, L, is defined as:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{K} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{P} \nu_i h_i(\mathbf{x}).$$
(2.22)

Minimizing *L* with respect to \mathbf{x} , $\boldsymbol{\lambda}$, $\boldsymbol{\nu}$ turns eq. (2.6) into an unconstrained optimization problem. We call the optimization problem in eq. (2.6) the *primal problem*, and denote its optimal value with p^* .

The Lagrangian dual function is

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}).$$
(2.23)

As described by Boyd et al. [2], the dual function lower bounds the optimal value of the primal problem, i.e., for an $\lambda \ge 0$ and any ν :

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \le p^{\star}, \tag{2.24}$$

as depicted in fig. 2.4. To find the largest lower bound that can be obtained from the dual function, we set up the *dual problem*:

$$\max \qquad g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \qquad (2.25a)$$

subject to
$$\lambda \ge 0.$$
 (2.25b)

This problem is concave; thus, it is always possible to solve for the globally optimal dual variables λ , ν [2]. Let d^* be the maximal value from eq. (2.25). If the equality

$$d^{\star} = p^{\star} \tag{2.26}$$



Figure 2.4: An illustration of Lagrangian duality in the cases of weak and strong duality, assuming Slater's constraint qualifications. Note the curves are for illustration purposes only. Under weak duality, we cannot certify the problem nor solve the dual of the dual for the globally optimal solution. If strong duality holds, the dual problem provides us with a means to certify solutions, and the dual of the dual is a tight convex relaxation of the primal problem.

holds, then we say the *duality gap is zero* and *strong duality holds* (see fig. 2.4b). It is not guaranteed that strong duality holds in general. However, the dual problem allows us to check if a solution is globally optimal. If we solve the dual problem and find $d^* = p^*$ (fig. 2.4b), then we know our solution is globally optimal (but if $p^* \neq d^*$ we cannot say we are not at a global minimum in general). Further, if we know strong duality holds, then if $d^* \neq p^*$, we know our solution is not globally optimal.

Slater's Condition: If the primal problem eq. (2.6) is convex or concave, and it is strictly feasible, then strong duality holds [2]. Referring to eq. (2.6) strict feasibility means that $\exists x$ such that

$$f_i(\mathbf{x}) = 0, \forall i \in \{1, \dots, K\} \text{ and},$$
 (2.27)

$$h_i(\mathbf{x}) < 0, \forall i \in \{1, \dots, P\}.$$
 (2.28)

The combination of convexity and strict feasibility is called *Slater's condition* or *Slater's constraint qualification*. Empirically, for many optimization problems in robotics, Slater's condition holds. If strong duality also holds, then the dual of the dual is a tight convex relaxation of the primal problem, as shown in fig. 2.4b. In this case, we can solve this *primal relaxation* for the globally optimal solution.

2.1.7 QCQP

A QCQP is an optimization problem where the cost function is quadratic, and there are quadratic constraints on the variables [5]. This class of optimization problems is important because least-squares cost functions (eq. (2.7)), commonly used in localization, can often be re-written as a QCQP, and there is much theory about how to approach solving a QCQP globally optimally [5].

A convenient expression (for later derivations) of a QCQP is:

$$\min_{\mathbf{x}} \qquad \mathbf{x}^T \mathbf{Q} \mathbf{x} \tag{2.29a}$$

subject to $\mathbf{x}^T \mathbf{A}_k \mathbf{x} = b_k, \quad \forall k \in \{1, \dots, K\},$ (2.29b)

where $\mathbf{x} \in \mathbb{R}^{D}$ and $\mathbf{Q}, \mathbf{A}_{1}, \dots, \mathbf{A}_{K}$ are symmetric [2]. While problems of this form are nonconvex and NP-hard in general, the application of duality theory leads to convenient expressions for the certificate and convex relaxation, described below.

We can write the Lagrangian of eq. (2.29) as

$$L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \sum_{k=1}^K \lambda_k (b_k - \mathbf{x}^T \mathbf{A}_k \mathbf{x})$$
(2.30)

$$=\sum_{k}\lambda_{k}b_{k}+\mathbf{x}^{T}\left(\mathbf{Q}-\sum_{k=1}^{K}\lambda_{k}\mathbf{A}_{k}\right)\mathbf{x}$$
(2.31)

$$= \mathbf{b}^T \boldsymbol{\lambda} + \mathbf{x}^T \mathbf{H}(\boldsymbol{\lambda}) \mathbf{x}, \qquad (2.32)$$

where $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_K]^T \in \mathbb{R}^K$, $\mathbf{b} = [b_1, \dots, b_K]^T \in \mathbb{R}^K$, and $\mathbf{H}(\boldsymbol{\lambda}) = \mathbf{Q} - \sum_k \lambda_k \mathbf{A}_k \in \mathbb{R}^{D \times D}$. We can write the dual function as

$$g(\boldsymbol{\lambda}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} \mathbf{b}^{T} \boldsymbol{\lambda} & \text{if } \mathbf{H}(\boldsymbol{\lambda}) \ge 0, \\ -\infty & \text{otherwise.} \end{cases}$$
(2.33)

In the case that $H(\lambda) \ge 0$, then the maximum of the dual function equals the minimum of the primal problem because, from eq. (2.32):

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{0} \implies \mathbf{H}(\boldsymbol{\lambda})\mathbf{x} = \mathbf{0} \implies p^* = \mathbf{b}^T \boldsymbol{\lambda}, \qquad (2.34)$$

so strong duality holds. This brings us to the QCQP certification problem; given a candidate

solution $\hat{\mathbf{x}}$, if we can solve

find
$$\mathbf{H}, \boldsymbol{\lambda}$$
 (2.35a)

s.t
$$\mathbf{H} = \mathbf{Q} - \sum_{k=1}^{K} \lambda_k \mathbf{A}_k$$
 (2.35b)

$$\mathbf{H} \ge 0 \tag{2.35c}$$

$$\mathbf{H}\hat{\mathbf{x}} = 0, \tag{2.35d}$$

then strong duality holds as per eq. (2.33). Note that if Slater's constraint qualifiers holds, and there are more state variables than constraints $(D \ge K)$, we can solve for λ directly:

$$\mathbf{H}(\boldsymbol{\lambda})\hat{\mathbf{x}} = \mathbf{0} \implies \begin{bmatrix} \mathbf{A}_1 \hat{\mathbf{x}} & \dots & \mathbf{A}_K \hat{\mathbf{x}} \end{bmatrix} \boldsymbol{\lambda} = \mathbf{Q}\hat{\mathbf{x}}.$$
 (2.36)

Further, if $\mathbf{H}(\boldsymbol{\lambda})$ is not positive semi-definite, then strong duality does not hold.

Now we consider the dual of the dual problem. Considering eq. (2.33), we assume that $H(\lambda) \ge 0$ and we can write the Lagrangian as

$$L' = \mathbf{b}^T \boldsymbol{\lambda} + \operatorname{tr}\left(\mathbf{X}\left(\mathbf{Q} - \sum_{k=1}^K \lambda_k \mathbf{A}_k\right)\right), \qquad (2.37)$$

where $\mathbf{X} \ge 0$ is a symmetric matrix of Lagrange multipliers to enforce $\mathbf{H}(\boldsymbol{\lambda}) = \mathbf{Q} - \sum_{k=1}^{K} \lambda_k \mathbf{A}_k \ge 0$. We can re-write L' as

$$L' = \operatorname{tr}(\mathbf{QX}) + \begin{bmatrix} b_1 - \operatorname{tr}(\mathbf{A}_1\mathbf{X}) & \dots & b_K - \operatorname{tr}(\mathbf{A}_K\mathbf{X}) \end{bmatrix} \boldsymbol{\lambda}.$$
 (2.38)

Now the dual of the dual is

$$q(\mathbf{X}) = \sup_{\boldsymbol{\lambda}} L'(\boldsymbol{\lambda}, \mathbf{X})$$
(2.39)

$$= \begin{cases} \operatorname{tr}(\mathbf{Q}\mathbf{X}) & \text{if } \operatorname{tr}(\mathbf{A}_k\mathbf{X}) = b_k \ \forall k \\ \infty & \text{otherwise.} \end{cases}$$
(2.40)

We assume the first condition, leading to the following optimization problem

minimize
$$tr(QX)$$
 (2.41a)

s.t
$$\operatorname{tr}(\mathbf{A}_k \mathbf{X}) = b_k \quad \forall k \in \{1, \dots, K\}$$
 (2.41b)

$$\mathbf{X} \ge 0. \tag{2.41c}$$

This is a semi-definite program (SDP), which makes solving the primal relaxation amenable to the wealth of SDP solver machinery. We can also obtain this relaxation if we re-write eq. (2.29)



Figure 2.5: A summary of duality theory applied to a QCQP.

as

minimize
$$\operatorname{tr}(\mathbf{Q}\mathbf{x}\mathbf{x}^T)$$
 (2.42a)

s.t
$$\operatorname{tr}(\mathbf{A}_k \mathbf{x} \mathbf{x}^T) = b_k.$$
 (2.42b)

If we introduce $\mathbf{X} = \mathbf{x}\mathbf{x}^T \ge 0$ and relax the constraint that $rank(\mathbf{X}) = 1$, then we arrive back at eq. (2.41).

In summary, given a QCQP in eq. (2.29) and a candidate solution, we can certify that solution as globally optimal if we can solve the certification problem in eq. (2.35). If we know strong duality holds, we can solve the primal relaxation in eq. (2.41) for a globally optimal solution (we have to extract x from $xx^T = X$). Depending on the context, the certificate may be preferable to solving eq. (2.41) for computational efficiency (e.g., if x is high-dimensional).

2.2 Related Work

Motivation For Certifiable Algorithms

Yang et al. [15] provide a *manifesto on certifiable perception*, which defines and motivates certifiable algorithms: Given an optimization problem that depends on some data, an algorithm is *certifiable* if, after solving the optimization problem, it provides a certificate of the quality of the solution (e.g., a proof of optimality) [15]. State estimation algorithms may return an estimate

arbitrarily far from the optimal solution without an easy way to check if the answer is sub-optimal [15]. However, [27] [28] and [15] show that we cannot directly compute an optimal solution for outlier-robust state estimation problems in polynomial time. Yang et al. [15] describe how this motivates a paradigm shift towards certifiable algorithms that can perform well on typical instances and certify correctness while also detecting worst-case instances. This is important for detecting failures early in the autonomy stack, making safety-critical robotics systems less brittle [15].

Certifiable Algorithms for Robotics

With this motivation in mind, many prior works have investigated certifiable algorithms for robotics problems.

Rotation Averaging: [6]–[9] tackle the problem of *rotation averaging* (RA); determining a set of absolute orientations from estimated relative rotations between those orientations.

Pose-graph Optimization: [10], [11] investigate certifiable algorithms for *pose-graph optimization* (PGO), an extension of RA that solves for absolute orientation and position from relative rotations and translations. These works cast the PGO problem as a QCQP and relax it to an SDP, which was often tight in practice. [12] proved that this SDP relaxation provides globally optimal solutions for sufficiently low noise levels and develops a structure-exploiting algorithm to efficiently solve large problem instances with many poses (with a computational cost comparable to local methods).

Multiple Point Cloud Registration: [13], [14] study certifiable algorithms for *multiple point cloud registration* (MPCR) using SDP relaxations and Lagrangian duality. MPCR is the problem of estimating a set of transformations that align observed point sets with respect to a global coordinate frame.

Landmark-Based SLAM: [4] unifies the three SLAM sub-problems — RA, MPCR, and PGO — developing an efficient certifiable algorithm for the full SLAM problem.

Robust Certifiable Algorithms: The above problems estimate unknown transformations from sensor data. In practice, many of these measurements are outliers [16]. Robust certifiable algorithms seek to estimate these transformations and provide a certificate of this solution in the

presence of outliers. [15]–[21] present outlier-robust certifiable or globally-optimal algorithms for specific problems in polynomial time.

Extrinsic Calibration: [22]–[24] study certifiable optimization in the context of extrinsic sensor calibration (estimating the relative position between multiple sensors).

Image Segmentation: [25] presents a fast and certifiable algorithm for inference in Markov Random Fields, a technique popular in semantic segmentation.

Monocular PnP: [17], [29], [30] develop certifiable algorithms for the Perspective-n-Point (PnP) problem, a type of localization where the goal is to estimate the pose of the perspective camera given n 2D-3D point correspondences. These prior works use the back-projection error in the formulation of their optimization problem:

$$\min_{\mathbf{T}} \qquad J(\mathbf{T}) = \sum_{n=1}^{N} \left(\mathbf{y}_n \mathbf{e}_3^T \mathbf{T} \mathbf{p}_n - \mathbf{M} \mathbf{T} \mathbf{p}_n \right)^T \mathbf{W}_n \left(\mathbf{y}_n \mathbf{e}_3^T \mathbf{T} \mathbf{p}_n - \mathbf{M} \mathbf{T} \mathbf{p}_n \right)$$
(2.43a)

subject to $\mathbf{T} \in SE(3).$

(2.43b)

This cost is already quadratic in the optimization variables T, so it is easily manipulated into a QCQP, and strong duality holds in practice. We verify this finding in chapter A. Contrast this with the re-projection error problem we aim to solve (eq. (2.19)) where our optimization variables in T appear in the denominator.

Redundant Constraints and The Lasserre Hierarchy

[15], [31]–[34] describe how adding *redundant constraints* to an optimization problem can be used to decrease the duality gap ([32] calls this duality strengthening) or tighten the primal relaxation of a QCQP. [32] states that whenever a new scalar constraint f_{k+1} is added to the Lagrangian, a new dual variable λ_{k+1} is added to the domain of the dual problem, and the bound provided by the new dual problem d_{new} is at least as good as the one provided by the previous one $d^* \leq d^*_{new} \leq p^*$. For example, [32] devise 21 scalar rotation matrix constraints, 15 of which are redundant, to tighten their dual problem.

The Lasserre hierarchy provides a systematic method for adding redundant constraints to obtain increasingly tighter lower bounds on the optimal value of a polynomial optimization problem. Each level of the hierarchy α consists of all monomials of degree α formed from the variables in the previous levels in the hierarchy. All the constraints that accompany those monomials are also added to the problem. While we will not describe Lasserre's in detail here, we refer the interested reader to [17] for an approachable treatment and [35] for the original text. We will apply Lasserre's hierarchy to the primal relaxation of a QCQP, and rely on two properties:

- 1. As the order of the hierarchy α increases, the optimal value of the primal relaxation converges to the lower bound of the QCQP (at each level, the relaxation gets tighter).
- 2. For some problems, a finite α yields a tight relaxation where the minimum of the primal relaxation is exactly equal to the global minimum of the QCQP.

2.2.1 Research Gap

As of writing, no one has developed a certifiable algorithm for stereo localization using the reprojection error. The closest prior art is monocular localization using the back-projection error (eq. (2.43)). The reason we want to use re-projection error instead of back-projection error is two-fold:

- Firstly, the pixel-space measurements of landmarks are subject to additive Gaussian noise. As first discussed by Matthies and Shafer [36], this Gaussian pixel space distribution manifests as a non-gaussian distribution when projected back into 3D space. However, the back-projection error problem (eq. (3.47)) is formulated as maximum likelihood estimation in 3D space assuming a Gaussian distribution. This is less accurate than the re-projection problem (eq. (2.19)), where the maximum likelihood estimation is in pixel space where the noise is a Gaussian distribution. This inaccurate model of uncertainty inherent to back-projection error is illustrated in fig. 2.6, where the true distribution of back-projected measurements is in blue, and the Gaussian approximation is shown with the red ellipse.
- Secondly, much prior work in stereo visual odometry (VO), localization, and SLAM solve the re-projection error problem [37]–[40]. Thus, if we could find a certifiable algorithm to solve this problem, it would be widely applicable to many robotics systems.

Our research task may be difficult because the re-projection error is non-polynomial in the pose variables. Further, this work is the first step towards more complex optimization problems involving the stereo-camera measurement model, including SLAM (and its various sub-problems) and robust estimation (i.e., measurement/correspondence outliers).



Figure 2.6: An illustration of stereo camera triangulation (the area in blue) uncertainty and its Gaussian approximation (the ellipse in red), adapted from [36]. Observe that the Gaussian approximation does support the long tail of the true distribution.

3 Methods and Results

This section describes our attempts to tighten the stereo localization problem and our findings. We will begin by introducing the datasets used to evaluate the forthcoming methods. We will describe each attempted method and evaluate its performance using the datasets, which will motivate subsequent methods.

3.1 Datasets

This section introduces the datasets for the experimental assessment of our methods.

STEREOSIM: This dataset is created by a custom simulator used to assess stereo-localization algorithms. It allows for the random generation of *stereo localization problems* where the camera and landmarks in its field of view are placed randomly in the world (see fig. 3.1a). It also generates observations using the forward camera model in eq. (2.5) with added isotropic Gaussian noise on the pixel measurements, as illustrated in fig. 3.1b. Simulation allows us to quickly generate localization problems with specific characteristics: noise level, number of landmarks, camera parameters, etc.

STARRYNIGHT: This dataset was collected using a stereo camera tracked by a ten-camera motion capture system [41]. The dataset was collected using twenty reflective landmarks placed on a black background. The motion capture system measured the positions of the landmarks to within a few millimeters of accuracy, providing the locations of the known landmarks for our stereo localization problem. Stereo-image pairs were logged at 15 Hz. We filter images with less than three features or co-linear features (in 3D space).

Camera Model: STARRYNIGHT used a stereo camera with intrinsics $f_u = 484.5, f_v = 484.5, c_u = 322, c_v = 247, b = 0.24 \text{ m}$. We simulated the same stereo camera model for experiments with STEREOSIM.



Figure 3.1: StereoSim: Left is a depiction of the landmark points, the sensor frame \mathcal{F}_s , and the world frame \mathcal{F}_w .



Figure 3.2: A visualization of the camera trajectory and the landmark positions in the STAR-RYNIGHT dataset.

3.2 Stereo Localization as a QCQP

In this section, we re-write eq. (2.19) in the form of a QCQP presented in eq. (2.29). Introducing the variable $\mathbf{v}_n = \frac{1}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n$:

$$\min_{\mathbf{x}} \qquad J(\mathbf{x}) = \sum_{n=1}^{N} \left(\mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)^T \mathbf{W}_k \left(\mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)$$
(3.1a)

subject to

et to
$$\mathbf{T} \in SE(3)$$
 (3.1b)

$$(\mathbf{I} - \mathbf{v}_n \mathbf{e}_3^T) \mathbf{T} \mathbf{p}_n = 0 \quad \forall n \in \{1, \dots, N\}$$
 (3.1c)

where the last constraint follows from the definition of \mathbf{v}_n . To make the constraints and cost quadratic, we relax the SE(3) constraint to O(3) and remove any linear terms in the optimization variables by introducing a homogenization variable $\omega = 1$:

$$\min_{\mathbf{x}} \qquad J(\mathbf{x}) = \sum_{n=1}^{N} \left(\omega \mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)^T \mathbf{W}_n \left(\omega \mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)$$
(3.2a)

subject to
$$\mathbf{C}^T \mathbf{C} = \mathbf{I}$$

$$(\omega \mathbf{I} - \mathbf{v}_n \mathbf{e}_3^T) \mathbf{T} \mathbf{p}_n = 0 \quad \forall n \in \{1, \dots, N\}$$
 (3.2c)

$$\omega^2 = 1 \tag{3.2d}$$

where

$$\mathbf{T} = \begin{bmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \mathbf{r} \\ 0 & 0 & 0 & 1 \end{bmatrix} \in SE(3).$$
(3.3)

Note that if we solve for C and it is left-handed (det(C) = -1), then we can simply determine its right-handed equivalent. The definition of \mathbf{v}_n implies that its third entry is 1:

$$\mathbf{v}_{n} = \begin{bmatrix} v_{n1} \\ v_{n2} \\ 1 \\ v_{n4} \end{bmatrix} \in \mathbb{R}^{4 \times 1}.$$
(3.4)

Let

$$\mathbf{u}_{n} = \begin{bmatrix} v_{n1} \\ v_{n2} \\ v_{n4} \end{bmatrix} \in \mathbb{R}^{3 \times 1}, \tag{3.5}$$

then we define our optimization variables as

$$\mathbf{x} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \mathbf{c}_3 \\ \mathbf{r} \\ \mathbf{u}_1 \\ \cdots \\ \mathbf{u}_N \\ \omega \end{bmatrix} \in \mathbb{R}^{13+3N}.$$
(3.6)

Below, we derive the Q and A_k matrices such that our quadratic formulation (eq. (3.2)) matches the standard QCQP form (eq. (2.29)). We will use three convenient relationships:

$$\omega = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix} \mathbf{x} = \mathbf{e}_D^T \mathbf{x}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$
(3.7)

$$\mathbf{v}_{n} = \mathbf{u}_{n} + \mathbf{e}_{3} = \begin{bmatrix} 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{E}_{\mathbf{v}_{n}} \mathbf{x}$$
(3.8)

$$\mathbf{T}\mathbf{p}_{n} = \begin{bmatrix} p_{n1}\mathbf{c}_{1} + p_{n2}\mathbf{c}_{2} + p_{n3}\mathbf{c}_{3} + \mathbf{r} \\ 1 \end{bmatrix}$$
(3.9)

$$= \begin{bmatrix} p_{n1}\mathbf{I}_{3\times3} & p_{n2}\mathbf{I}_{3\times3} & p_{n3}\mathbf{I}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{4\times1} & \dots & \mathbf{0}_{4\times1} \\ \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & 0 & \dots & 1 \end{bmatrix} \mathbf{x}$$
(3.10)

$$=\mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}\mathbf{x}.$$
(3.11)

Cost: Substituting $\omega = \mathbf{e}_D^T \mathbf{x}$ and $\mathbf{v}_n = \mathbf{E}_{\mathbf{v}_n} \mathbf{x}$ into our QCQP cost eq. (3.2):

$$J = \sum_{n=1}^{N} \left(\mathbf{y}_n \mathbf{e}_D^T \mathbf{x} - \mathbf{M} \mathbf{E}_{\mathbf{v}_n} \mathbf{x} \right)^T \mathbf{W}_n \left(\mathbf{y}_n \mathbf{e}_D^T \mathbf{x} - \mathbf{M} \mathbf{E}_{\mathbf{v}_n} \mathbf{x} \right)$$
(3.12)

$$= \mathbf{x}^{T} \left(\sum_{n=1}^{N} \left(\mathbf{y}_{n} \mathbf{e}_{D}^{T} - \mathbf{M} \mathbf{E}_{\mathbf{v}_{n}} \right)^{T} \mathbf{W}_{n} \left(\mathbf{y}_{n} \mathbf{e}_{D}^{T} - \mathbf{M} \mathbf{E}_{\mathbf{v}_{n}} \right) \right) \mathbf{x}$$
(3.13)

$$=\mathbf{x}^{T}\mathbf{Q}\mathbf{x}$$
(3.14)

Therefore our cost matrix is

$$\mathbf{Q} = \sum_{n=1}^{N} \left(\mathbf{y}_{n} \mathbf{e}_{D}^{T} - \mathbf{M} \mathbf{E}_{\mathbf{v}_{n}} \right)^{T} \mathbf{W}_{n} \left(\mathbf{y}_{n} \mathbf{e}_{D}^{T} - \mathbf{M} \mathbf{E}_{\mathbf{v}_{n}} \right).$$
(3.15)

Rotation Matrix Constraints: From the constraint $\mathbf{C}^T \mathbf{C} = \mathbf{I}$, we get the nine constraints

$$\mathbf{c}_{i}^{T}\mathbf{c}_{j} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad \forall i, j \in \{1, 2, 3\}.$$
(3.16)

Due to symmetry, only 9 - 3 = 6 of these constraints are unique:

$$\mathbf{x}^{T} \mathbf{E}_{\mathbf{c}_{i}}^{T} \mathbf{E}_{\mathbf{c}_{j}} \mathbf{x}, \quad \forall i \in \{1, 2, 3\}, \ \forall j \in \{i, \dots, 3\},$$
(3.17)

where $\mathbf{E}_{\mathbf{c}_i} \in \mathbb{R}^{3 \times (13+3N)}$ has the 3×3 identity matrix $\mathbf{I}_{3 \times 3} \in \mathbb{R}^{3 \times 3}$ in columns 3i - 2 to 3i.

Constraints on \mathbf{v}_n **:** We can re-write eq. (3.2c) as

$$\left((\mathbf{x}^T \mathbf{e}_D) \mathbf{I} - \mathbf{E}_{\mathbf{v}_n} \mathbf{x} \mathbf{e}_3^T \right) \mathbf{E}_{\mathbf{T} \mathbf{v}_n} \mathbf{x} = \mathbf{E}_{\mathbf{T} \mathbf{v}_n} \mathbf{x} \mathbf{x}^T \mathbf{e}_D - \mathbf{E}_{\mathbf{v}_n} \mathbf{x} \mathbf{e}_3^T \mathbf{E}_{\mathbf{T} \mathbf{v}_n} \mathbf{x} = \mathbf{0} \in \mathbb{R}^{4 \times 1}.$$
(3.18)

This constitutes four scalar constraints (one per row). Extracting the i^{th} row:

$$\mathbf{e}_{i}^{T}\mathbf{E}_{\mathbf{T}\mathbf{v}_{n}}\mathbf{x}\mathbf{x}^{T}\mathbf{e}_{D} - \mathbf{e}_{i}^{T}\mathbf{E}_{\mathbf{v}_{n}}\mathbf{x}\mathbf{e}_{3}^{T}\mathbf{E}_{\mathbf{T}\mathbf{v}_{n}}\mathbf{x} = 0, \quad \forall i \in \{1, \dots, 4\}.$$
(3.19)

 $\mathbf{x}^T \mathbf{e}_D, \mathbf{e}_i^T \mathbf{E}_{\mathbf{T}\mathbf{v}_n} \mathbf{x}$, and $\mathbf{e}_i^T \mathbf{E}_{\mathbf{v}_n} \mathbf{x}$ are scalars, so we can re-arrange:

$$\mathbf{x}^{T}\mathbf{e}_{D}\mathbf{e}_{i}^{T}\mathbf{E}_{\mathbf{T}\mathbf{v}_{n}}\mathbf{x} - \mathbf{x}^{T}\mathbf{E}_{\mathbf{v}_{n}}^{T}\mathbf{e}_{i}\mathbf{e}_{3}^{T}\mathbf{E}_{\mathbf{T}\mathbf{v}_{n}}\mathbf{x} = 0, \quad \forall i \in \{1, \dots, 4\}.$$
(3.20)

Notice that if i = 3, then $\mathbf{E}_{\mathbf{v}_n}^T \mathbf{e}_i = \mathbf{e}_D$, so the equation above is trivial. Thus, we have 3N additional constraints from \mathbf{v}_n :

$$\mathbf{x}^{T} \left(\mathbf{E}_{\mathbf{v}_{n}}^{T} \mathbf{e}_{i} \mathbf{e}_{3}^{T} \mathbf{E}_{\mathbf{T}\mathbf{v}_{n}} - \mathbf{e}_{D} \mathbf{e}_{i}^{T} \mathbf{E}_{\mathbf{T}\mathbf{v}_{n}} \right) \mathbf{x} = 0, \quad \forall i \in \{1, 2, 4\}, \quad \forall n \in \{1, \dots, N\}.$$
(3.21)

Slack variable constraint: The condition $\omega^2 = 1$ adds one more constraint: $\omega^2 = 1 \implies \mathbf{x}^T \mathbf{e}_D \mathbf{e}_D^T \mathbf{x} = 1.$

Note that the total number of constraints is 7 + 3N, which is less than the number of unknowns in x: D = 12 + 3N.

3.3 Tightness of the Stereo Localization QCQP

Recall from section 2.1.7 that duality theory applied to a QCQP gives us the means to certify solutions and a tight convex relaxation of the primal problem if strong duality holds. In this section we assess if strong duality holds for the QCQP derived in section 3.2.

Certificate: In this section, we empirically assess the duality gap of the certificate. Using STEREOSIM, we generate a set of 32 {landmark config, ground truth camera pose} pairs — call these *problems* — ranging from 5 landmarks to 20. For realistic noise variance levels (in pixel space) of $\{0.1, 0.3, 0.5, 0.7, 1, 4\}$, we generated the simulated stereo camera measurements for each problem. We also generated 50 random different initial conditions for each problem. Then, for all noise levels, we solved every problem with the local solver (GAUSSNEWTON, from section 2.1.5) for each initial condition. We took the lowest cost solution on a given problem as the globally optimal solution. We discarded solutions that did not converge after 100 iterations of GAUSSNEWTON (i.e., did not meet the first-order optimality conditions). Using this, we can



Noise Standard Deviation [pixels]

Figure 3.3: Each dot represents one solution from GAUSSNEWTON. The blue dots are globally optimal solutions, and the red dots are local minima. We observe that while there is a gap between the minimum eigenvalues of $H(\lambda)$ for the local and global minima, this gap decreases as the noise level increases. This result indicates that strong duality does not hold in practice.

classify all other solutions to that problem as globally optimal or not. For each solution, we also generate the certificate matrix $H(\lambda)$ and save its minimum eigenvalue. See fig. 3.3 for the results, where the blue dots denote the solutions classified as globally optimal, and the red dots are the non-globally optimal solutions.

We see that for these realistic noise levels, there is a gap between the minimum eigenvalue of the globally optimal solutions and the non-globally optimal solutions. However, this gap decreases significantly as the noise level increases, indicating that strong duality does not hold. However, this certificate may be useful in practice if the thresholds on the minimum eigenvalues of the certificate matrix beyond which to expect local minima are known (e.g., from testing).

Tightness of Primal Relaxation: In this section we empirically assess the duality gap of the primal relaxation for the stereo localization QCQP. We use the same set of 32 problems described in the previous section. The results are presented in fig. 3.4a. Notice that the duality gap $p^* - d^* > 0$ in all instances, reaffirming that strong duality does not hold.

3.4 Redundant Constraints

If the primal relaxation is not tight, it may be tightened with redundant constraints, as discussed above. Below we derive some redundant constraints for the stereo localization QCQP formulation in section 3.2.

Parallel Constraint Note that $\mathbf{v}_n = \frac{1}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n$, \mathbf{v}_n and $\mathbf{T} \mathbf{p}_n$ are parallel. Further, for any parallel column vectors $\mathbf{u}, \mathbf{w}, \exists c \in \mathbb{R}$ such that

$$\mathbf{u} = c\mathbf{w} \implies \mathbf{u}\mathbf{w}^T = c\mathbf{w}\mathbf{w}^T = \mathbf{w}\mathbf{u}^T.$$
 (3.22)

Thus, we can write:

$$\mathbf{v}_n(\mathbf{T}\mathbf{p}_n)^T = (\mathbf{T}\mathbf{p}_n)\mathbf{v}_n^T \in \mathbb{R}^{4 \times 4}.$$
(3.23)

Now the i^{th}, j^{th} constraint in this 4×4 matrix is

$$\mathbf{e}_i^T \mathbf{v}_n (\mathbf{T} \mathbf{p}_n)^T \mathbf{e}_j = \mathbf{e}_i^T (\mathbf{T} \mathbf{p}_n) \mathbf{v}_n^T \mathbf{e}_j, \quad \forall i, j \in \{1, \dots, 4\}$$
(3.24)

$$\implies \mathbf{e}_i^T \mathbf{E}_{\mathbf{v}_n} \mathbf{x} \mathbf{x}^T \mathbf{E}_{\mathbf{T}\mathbf{p}_n}^T \mathbf{e}_j = \mathbf{e}_i^T \mathbf{E}_{\mathbf{T}\mathbf{p}_n} \mathbf{x} \mathbf{x}^T \mathbf{E}_{\mathbf{v}_n}^T \mathbf{e}_j.$$
(3.25)

Because $\mathbf{x}^T \mathbf{E}_{\mathbf{T}\mathbf{p}_n}^T \mathbf{e}_j$, $\mathbf{e}_i^T \mathbf{E}_{\mathbf{v}_n} \mathbf{x}$, $\mathbf{e}_i^T \mathbf{E}_{\mathbf{T}_{\mathbf{p}_n}} \mathbf{x}$, and $\mathbf{x}^T \mathbf{E}_{\mathbf{v}_n}^T \mathbf{e}_j$ are scalars, we can use the fact they commute to write:

$$\mathbf{x}^{T} \mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}^{T} \mathbf{e}_{j} \mathbf{e}_{i}^{T} \mathbf{E}_{\mathbf{v}_{n}} \mathbf{x} - \mathbf{x}^{T} \mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}^{T} \mathbf{e}_{j} \mathbf{e}_{i}^{T} \mathbf{E}_{\mathbf{T}\mathbf{p}_{n}} \mathbf{x} = 0$$
(3.26)

$$\implies \mathbf{x}^{T} (\mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}^{T} \mathbf{e}_{j} \mathbf{e}_{i}^{T} \mathbf{E}_{\mathbf{v}_{n}} - \mathbf{E}_{\mathbf{v}_{n}}^{T} \mathbf{e}_{j} \mathbf{e}_{i}^{T} \mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}) \mathbf{x} = 0$$
(3.27)

which is of the desired form (eq. (2.29)). Note that when i = j, $(\mathbf{E}_{\mathbf{T}\mathbf{p}_n}^T \mathbf{e}_j \mathbf{e}_i^T \mathbf{E}_{\mathbf{v}_n} - \mathbf{E}_{\mathbf{v}_n}^T \mathbf{e}_j \mathbf{e}_i^T \mathbf{E}_{\mathbf{T}\mathbf{p}_n})$ is skew symmetric, so equation 3.27 is trivially satisfied. Therefore, we will can add (16-4)N = 12N constraints to the QCQP:

$$\mathbf{x}^{T}(\mathbf{E}_{\mathbf{T}\mathbf{p}_{n}}^{T}\mathbf{e}_{j}\mathbf{e}_{i}^{T}\mathbf{E}_{\mathbf{v}_{n}} - \mathbf{E}_{\mathbf{v}_{n}}^{T}\mathbf{e}_{j}\mathbf{e}_{i}^{T}\mathbf{E}_{\mathbf{T}\mathbf{p}_{n}})\mathbf{x} = 0 \quad \text{else } \forall i, j \in \{1, \dots, 4\}|_{i \neq j}$$
(3.28)

Redundant SO(3) **Constraints** Apart from the 6 constraints that come from $\mathbf{C}^T \mathbf{C} = \mathbf{I}$ in eq. (3.17), we can add 15 more redundant constraints to characterize $\mathbf{C} \in SO(3)$ [33]. Let

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{bmatrix}.$$
(3.29)

Because $\mathbf{C}\mathbf{C}^T = \mathbf{I}$, we have the constraints

$$\mathbf{r}_i^T \mathbf{r}_j = \delta_{ij}.\tag{3.30}$$

We note that

$$\mathbf{r}_{i} = \mathbf{E}_{\mathbf{r}_{i}} \mathbf{x} = \begin{bmatrix} \mathbf{e}_{i}^{T} \\ \mathbf{e}_{i+3}^{T} \\ \mathbf{e}_{i+6}^{T} \end{bmatrix} \mathbf{x}, \quad \mathbf{E}_{\mathbf{r}_{i}} \in \mathbb{R}^{3 \times D},$$
(3.31)

so

$$\mathbf{x}^T \mathbf{E}_{\mathbf{r}_j}^T \mathbf{E}_{\mathbf{r}_j} \mathbf{x} = \delta_{ij} \quad \forall i \in \{1, 2, 3\}, \forall j \in \{i, \dots, 3\}.$$
(3.32)

This amounts to 6 additional constraints. Further, we can exploit the mutual orthogonality of the columns of C:

$$\mathbf{c}_i^{\times} \mathbf{c}_j - \omega \mathbf{c}_k = \mathbf{0}, \quad \forall i, j, k \in \operatorname{cyclic}(1, 2, 3),$$
(3.33)

which results in another 9 constraints:

$$\mathbf{e}_m^T \mathbf{c}_i^{\times} \mathbf{c}_j - \omega \mathbf{e}_m^T \mathbf{c}_k = 0, \quad i, j, k \in \operatorname{cyclic}(1, 2, 3), m \in \{1, 2, 3\}.$$
(3.34)

Let

$$\mathbf{c}_{i} = \begin{bmatrix} c_{1i} & c_{2i} & c_{3i} \end{bmatrix}^{T} \implies \mathbf{c}_{i}^{\times} = \begin{bmatrix} 0 & -c_{3i} & c_{2i} \\ c_{3i} & 0 & -c_{1i} \\ -c_{2i} & c_{1i} & 0 \end{bmatrix},$$
(3.35)

so

$$\mathbf{e}_m^T \mathbf{c}_i^{\times} = \mathbf{x}^T \mathbf{E}_{-\mathbf{c}_i^{\times} \mathbf{e}_m}^T, \qquad (3.36)$$

where $\mathbf{E}_{-\mathbf{c}_{i}^{\times}\mathbf{e}_{m}}^{T} \in \mathbb{R}^{D\times3}$ has $-\mathbf{e}_{m}^{\times}$ in rows 3i - 2 to 3i, with zeros elsewhere. Then we can write the constraints in eq. (3.34) as

$$\mathbf{x}^{T} \left(\mathbf{E}_{-\mathbf{c}_{i}^{\times}\mathbf{e}_{m}}^{T} \mathbf{E}_{\mathbf{c}_{j}} - \mathbf{e}_{\omega} \mathbf{e}_{D}^{T} \mathbf{E}_{\mathbf{c}_{k}} \right) \mathbf{x} = 0, \quad \forall i, j, k \in \operatorname{cyclic}(1, 2, 3), m \in \{1, 2, 3\}.$$
(3.37)

Cross-Coupling Constraints We also investigated adding new variables and redundant constraints associated with those variables. Consider adding the $\binom{N}{2}$ cross-coupling variables

$$q_{ij} = (\mathbf{e}_{3}^{T} \mathbf{T} \mathbf{p}_{i}) (\mathbf{e}_{3}^{T} \mathbf{T} \mathbf{p}_{j}), \forall (i, j) \in \{(i, j) | 1 \le i < j \le N\} = \mathcal{S} \text{ to our state, so}$$

$$\begin{bmatrix} \mathbf{c}_{1} \\ \mathbf{c}_{2} \\ \mathbf{c}_{3} \\ \mathbf{r} \\ \mathbf{u}_{1} \\ \cdots \\ \mathbf{u}_{N} \\ \mathbf{\omega} \\ q_{12} \\ q_{13} \\ \cdots \\ q_{1N} \\ q_{23} \\ \cdots \\ q_{N(N-1)} \end{bmatrix}.$$

$$(3.38)$$

Then we add the following quadratic constraints on x to enforce the definition of q_{ij} :

$$\omega q_{ij} = \left(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_i\right) \left(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_j\right).$$
(3.39)

which amounts to $\binom{N}{2}$ constraints. We also get the following expressions:

$$q_{ij}\mathbf{v}_i = \mathbf{T}\mathbf{p}_i(\mathbf{e}_3^T\mathbf{T}\mathbf{p}_j), \forall (i,j) \in \mathcal{S}$$
(3.40)

$$q_{ij}\mathbf{v}_j = \mathbf{T}\mathbf{p}_j(\mathbf{e}_3^T\mathbf{T}\mathbf{p}_i), \forall (i,j) \in \mathcal{S}$$
(3.41)

which add $6\binom{N}{2}$ constraints,

$$\frac{q_{ij}}{q_{im}} = \frac{(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_i)(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_j)}{(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_i)(\mathbf{e}_3^T \mathbf{T} \mathbf{p}_m)} = \frac{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_j}{\mathbf{e}_3^T \mathbf{T} \mathbf{p}_m}$$
(3.42)

$$\implies \mathbf{e}_3^T \mathbf{T} \mathbf{p}_m q_{ij} = \mathbf{e}_3^T \mathbf{T} \mathbf{p}_j q_{im}, \forall (i, j, m) \in \{(i, j, m) | 1 \le i < j < m \le N\}$$
(3.43)

which adds $\binom{N}{3}$ constraints and,

$$q_{ij}q_{km} = q_{im}q_{kj} = q_{jm}q_{ik}, (3.44)$$

$$\forall (i, j, k, m) = \{ (i, j, k, m) | 1 \le i < j < k < m \le N \},$$
(3.45)

which adds $3\binom{N}{4}$ constraints. In total we have $\binom{N}{2}$ new variables and $7\binom{N}{2} + \binom{N}{3} + 3\binom{N}{4}$ new constraints. These constraints are easy to re-write in the QCQP form (eq. (2.29)) by using eqs. (3.7), (3.8) and (3.11) and following expression $q_{ij} = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} \mathbf{x} = \mathbf{e}_{q_{ij}}^T \mathbf{x}$.



(a) Without redundant constraints

(b) With redundant constraints

Figure 3.4: A plot of the duality gap of the primal relaxation. The plot on the left has no redundant constraints, while the plot on the right includes all of our redundant constraints. Notice that in both cases, the gap is non-zero for all problems indicating that in both cases, the primal relaxation is not tight.

3.4.1 Tightness with Redundant Constraints

Figure 3.4a compares the cost gap for the primal relaxation before and after adding redundant constraints using the same problems described in section 3.3. We added all of the redundant constraints and variables described in the preceding sections because for each redundant constraint, the new dual cost d_{new}^{\star} will be at least as large as the dual cost without that redundant constraint [32]. The redundant constraints did not tighten the problem nor appreciably shrink the cost gap.

3.5 Iterative SDP

Motivated by the lack of strong duality for the primal relaxation, we investigate an iterative approach described below.

As outlined by [17], [29], [30], the back-projection error PnP problem, in eq. (2.43), has a tight primal relaxation (we verify this in appendix A). We observe that the back-projection error (eq. (2.43)) differs from the re-projection error (eq. (2.19)) by a per-point inverse depth squared weighting:

$$\sum_{n=1}^{N} \frac{1}{z_n^2} (z_n \omega \mathbf{y}_n - \mathbf{M} \mathbf{T} \mathbf{p}_n)^T \mathbf{W}_n (\omega \mathbf{y}_n - \mathbf{M} \mathbf{T} \mathbf{p}_n) \quad (\text{Re-projection error})$$
(3.46)

$$\sum_{n=1}^{N} (\omega z_n \mathbf{y}_n - \mathbf{MT} \mathbf{p}_n)^T \mathbf{W}_n (\omega z_n \mathbf{y}_n - \mathbf{MT} \mathbf{p}_n) \quad \text{(Back-projection error)}, \tag{3.47}$$

where $z_n = \mathbf{e}_3^T \mathbf{T} \mathbf{p}_n$. This motivates an iterative algorithm for stereo localization with re-

projection error, which aims to solve for the optimal back projection error at each iteration, and uses that intermediate solution to find depth estimates of each point to be used in the next iteration.

To derive this method, we start by re-writing g_n :

$$\mathbf{g}_n = \mathbf{y}_n - \frac{1}{z_n} \mathbf{M} \mathbf{T} \mathbf{p}_n \tag{3.48}$$

$$=\frac{1}{z_n}\left(z_n\mathbf{y}_n - \mathbf{MTp}_n\right) \tag{3.49}$$

$$=\frac{1}{z_n}\left(\mathbf{y}_n\mathbf{e}_3^T-\mathbf{M}\right)\mathbf{T}\mathbf{p}_n\tag{3.50}$$

$$= \frac{1}{z_n} \left(\mathbf{y}_n \mathbf{e}_3^T - \mathbf{M} \right) \left(\mathbf{p}_n^T \otimes \mathbf{I} \right) \operatorname{vec}(\mathbf{T}), \tag{3.51}$$

where \otimes is the Kronecker product operation, and vec(**T**) is a column vector of the columns of **T**, both defined in [1]. Thus, we can re-write our objective function in eq. (2.19) as

$$\min_{\mathbf{x}} \qquad J = \mathbf{x}^T \mathbf{Q} \mathbf{x} \tag{3.52a}$$

subject to
$$\mathbf{C}^T \mathbf{C} = \mathbf{I}$$
 (3.52b)

(relaxing the SE(3) constraint to O(3)), where

$$\mathbf{Q} = \sum_{n} \frac{1}{z_n^2} (\mathbf{p}_n^T \otimes \mathbf{I})^T (\mathbf{y}_n \mathbf{e}_3^T - \mathbf{M})^T \mathbf{W}_n (\mathbf{y}_n \mathbf{e}_3^T - \mathbf{M}) (\mathbf{p}_n^T \otimes \mathbf{I}),$$
(3.53)

and $\mathbf{x} = \text{vec}(\mathbf{T})$. Also, notice that

$$z_n^2 = (\mathbf{e}_3 \mathbf{T} \mathbf{p}_n)^2 \tag{3.54}$$

$$= \mathbf{e}_3^T \mathbf{T} \mathbf{p}_n \mathbf{p}_n^T \mathbf{T} \mathbf{e}_3 \tag{3.55}$$

$$= \mathbf{e}_{3}^{T} \left(\mathbf{p}_{n}^{T} \otimes \mathbf{T} \right) \operatorname{vec}(\mathbf{T}) \operatorname{vec}(\mathbf{T})^{T} \left(\mathbf{p}_{n}^{T} \otimes \mathbf{T} \right)^{T} \mathbf{e}_{3}$$
(3.56)

$$= \operatorname{tr}\left(\left(\mathbf{p}_{n} \otimes \mathbf{I}\right)^{T} \mathbf{e}_{3} \mathbf{e}_{3}^{T} \left(\mathbf{p}_{n} \otimes \mathbf{I}\right) \mathbf{X}\right)$$
(3.57)

where $\mathbf{X} = \operatorname{vec}(\mathbf{T})\operatorname{vec}(\mathbf{T})^T$, meaning we can update $\frac{1}{z_n^2}$ in eq. (3.53) without ever extracting \mathbf{T} from \mathbf{X} .

To summarize, we start with an initial guess T to initialize $\mathbf{X} = \text{vec}(\mathbf{T})\text{vec}(\mathbf{T})^T$ (we use identity in all of our experiments). Then we compute z_n^2 , $\forall n$ with eq. (3.57), and find Q with equation eq. (3.53). Finally, we solve eq. (3.52) as an SDP for X, and iterate.

Figure 3.6 plots the gap between the cost of the solution extracted from the iterative SDP method q_{iter}^{\star} and the globally optimal solution. We use the same 32 problems with 50 different initial conditions described in section 3.3. Notice that for most problems, the iterative solver



Figure 3.5: An example problem where the iterative SDP method does not find the globally optimal solution. In this case, one landmark is at a significantly larger depth than the others.

converges to within some reasonable tolerance of the globally optimal solution. However, it does not converge in all instances, as shown by the group of solutions with a large cost gap. We qualitatively notice a failure cause occurs when there is a large difference in depths between points in the problem (e.g., a point at a very large depth relative to the others). See fig. 3.5 for an example. The re-projection error heavily down-weights points with high z_n relative to the back-projection error, so the iterative back-projection solutions jump out of the basin of convergence on the first iteration.



Figure 3.6: For many problem instances, this figure plots the gap between the global minimum cost p^* and the cost obtained from the solution found by the iterative SDP algorithm denoted with q^* .

3.6 Solver Comparison

3.6.1 Global Optimality

This section compares GAUSSNEWTON to our proposed SDP methods. ITERSDP is the iterative SDP method discussed in section 3.5, where the solution is refined with GAUSSNEWTON. PRIMALRELAX is the primal relaxation discussed in section 3.2 with no redundant constraints, where the solution is refined with GAUSSNEWTON.

We used the same dataset discussed in section 3.3 to compare how often GAUSSNEWTON, ITERSDP, and PRIMALRELAX find the globally optimal solution across various noise levels. Figure 3.7 presents the results.

We observe that ITERSDP and PRIMALRELAX *always* find the globally optimal solution in practice, while the GAUSSNEWTON does not. These results indicate that ITERSDP and PRIMAL-RELAX can help avoid local minima in stereo localization problems.

Figure 3.8 shows this same comparison on STARRYNIGHT. We solved each problem with 50 initial guesses for each stereo-image pair in the dataset. We see that all the methods do well on this dataset, likely because the point configuration is on a plane (which may be an easy case). However, we still observe that PRIMALRELAX and ITERSDP methods outperform GAUSSNEWTON.



Noise Standard Deviation [pixels]

Figure 3.7: A comparison of the localization methods measuring the percentage of globally optimal solutions on StereoSIM at various pixel-space noise levels.

More investigation is required to determine how the QCQP primal relaxation brings the solution within the basin of convergence of GAUSSNEWTON, even though it is not tight.

3.6.2 Solver Time Efficiency

In this section, we assess the time efficiency of GAUSSNEWTON, ITERSDP, and PRIMALRELAX. Using STEREOSIM, we construct a dataset with problems containing 4 to 20 landmarks. For each number of landmarks, we add 50 problems to the dataset with various realistic noise levels. Then, we run GAUSSNEWTON, ITERSDP, and PRIMALRELAX on these problems and produced the plot in fig. 3.9, which shows the average solution time as a function of the number of landmarks. We observe that while the runtime of PRIMALRELAX scales exponentially with the number of landmarks, GAUSSNEWTON and ITERSDP have roughly constant and equal solution times. Thus, there is a strong case for employing ITERSDP for practical problems; empirically, it always finds the globally optimal solution and is equally as fast as GAUSSNEWTON.

We also run a practical test by measuring the average runtime across all problems in StereoSIM, as if we were trying to localize the camera at every measurement time. Table 3.1 presents the results, which supports our conclusion that ITERSDP is roughly as fast as GAUSSNEWTON and much faster than PRIMALRELAX.



Figure 3.8: A comparison of the localization methods measuring the percentage of globally optimal solution on STEREOSIM at various pixel-space noise levels.

Solver	Average Runtime (s)		
GaussNewton	0.23 ± 0.36		
IterSDP	0.20 ± 0.08		
PrimalRelax	1.60 ± 2.19		

Table 3.1: Average solver runtime on STARRYNIGHT.

3.7 Visualizing Solution Trajectories

To understand the behavior of GAUSSNEWTON, ITERSDP, and PRIMALRELAX, we plot the pose at each solution iteration in fig. 3.10. We see that GAUSSNEWTON requires many iterations to converge, and its pose estimate jumps around. On the other hand, ITERSDP finds the globally optimal solution in 1-3 iterations from the initial guess, while PRIMALRELAX takes just one iteration to get the pose roughly correct. Note that both ITERSDP and PRIMALRELAX show the iterations where their solution is refined with GAUSSNEWTON. This confirms the idea that the iterative back-projection solutions are a good proxy for the re-projection error problem because within just one iteration, ITERSDP is already near the globally optimal solution.



Figure 3.9: Solver runtime as a function of the number of landmarks.



Figure 3.10: Visualizing the solution trajectories of the GAUSSNEWTON, ITERSDP, and PRIMAL-RELAX. The solver pose at the i^{th} iteration is labeled \mathcal{F}_i . PRIMALRELAX has more than one pose in its trajectory because the output from the primal relaxation is refined with GAUSSNEWTON.

4 Challenges and Future Work

Unfortunately this set of redundant constraints and additional variables did not tighten the primal relaxation of the re-projection error stereo localization problem, which means the certificate is also not tight. This section outlines our attempts to tighten the primal relaxation and describes areas for future work.

4.1 1D Problem

We begin by attempting to tighten a one-dimensional version of the stereo localization problem, visualized in fig. 4.1. The robot is at a position x along the real number line. Landmarks are at positions a_1, \ldots, a_N . The forward measurement model is

$$y_n = \frac{1}{x - a_n},\tag{4.1}$$

and the cost function is

$$J(x) = \sum_{n} \left(y_n - \frac{1}{x - a_n} \right)^2.$$
 (4.2)

Notice that the state variable x appears in the denominator like the re-projection error. We make the substitution

$$z_n = \frac{1}{x - a_n},\tag{4.3}$$

to re-write the problem as a QCQP:

$$\min_{x} J(x) = \sum_{n} (y_n - z_n)^2 (4.4)$$

 $z_n(x-a_n) = 1 \tag{4.5}$



Figure 4.1: 1D problem.



Figure 4.2: Investigating tightness of the 1D problem. The blue dashed line denotes the global minimum cost for a single problem instance, and the bars indicate the cost of the primal relaxation both with and without redundant constraints.

To determine the globally optimal solution, we developed a local solver detailed in appendix B, and ran this with many different initial conditions to find the best solution. Using a test problem with 5 landmarks and random Gaussian noise added to the measurements, we observe that the primal relaxation of the 1D QCQP is not tight, as shown by the gap between the globally optimal solution and the primal cost in fig. 4.2. In an attempt to tighten this problem, we add the following redundant constraint:

$$z_j - z_i = \frac{1}{x - a_j} - \frac{1}{x - a_i} = \frac{a_j - a_i}{(x - a_j)(x - a_i)} = z_j z_i (a_j - a_i)$$
(4.6)

When added to the primal relaxation, this redundant constraint tightens the problem, as shown by the orange bar in fig. 4.2. This provided the insight that some cross-coupling constraints between substituted variables of different landmarks may be required to tighten the 3D problem. Unfortunately, these cross-coupling constraints did not tighten the 3D stereo localization problem, as detailed in section 3.4.



Figure 4.3: The 2D problem. T is the unknown transformation matrix, the solid gray arrows denote the sensor frame, the dashed gray indicates $x_s = 1$ plane, the black dots are the camera optical sensors, and the colored dots are the landmarks.

4.2 **2D Problem**

N

Having tightened the 1D problem, in this section, we attempt to tighten the 2D problem, given below:

 $\min_{\mathbf{T}}$

$$J(\mathbf{T}) = \sum_{n=1}^{T} \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_2^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right)^T \mathbf{W}_n \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_2^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right) \quad (4.7a)$$

et to $\mathbf{T} \in SE(2),$ (4.7b)

T

subject to

where M is chosen to emulate the intrinsic parameter matrix of a stereo camera (eq. (2.4))

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0.5 \\ 1 & 0 & -0.5 \end{bmatrix}.$$
 (4.8)

(4.7b)

This is scenario is visualized in fig. 4.3. Following a similar procedure to the 3D problem, we can re-write this problem as a QCQP. See appendix C for the details. We run a local solver for many different initial conditions to find the globally optimal solution and take the best result. The derivation of the local solver can be found in appendix C. On a test problem with three



Figure 4.4: Investigating tightness of the 2D problem. The blue dashed line denotes the global minimum cost for a single problem instance, and the bars indicate the cost found by the primal relaxation both with and without redundant constraints.

landmarks with added pixel-space noise, we observe that, like the 1D and 3D problems, the primal relaxation is not tight, as shown by the gap between the primal cost and the globally optimal cost in fig. 4.4. To tighten this problem, we applied the Lasserre hierarchy of order $\alpha = 2$. This adds all monomials of degree $\alpha = 2$ to the problem and their associated redundant constraints. This results in a tight primal relaxation, as shown in fig. 4.4.

Unfortunately, this straightforward application of the Lasserre hierarchy has a drawback: the SDP has many more variables and constraints, increasing the solution time. Table 4.1 quantifies this issue on the 2D and 3D stereo localization problems with just three landmarks. With no redundant constraints, both problems are tractable but not tight. When adding the redundant constraints for the Lasserre hierarchy at $\alpha = 2$, the 3D problem becomes too large to solve in a reasonable amount of time. Thus, we could not verify if the Lasserre hierarchy at $\alpha = 2$ could tighten the 3D stereo localization problem with the re-projection error. Nor did we find the subset of variables and redundant constraints required to tighten the 2D stereo localization problem.

	Problem Dimension	Variables	Constraints	Solution Time (s)
No Dodundont Constraints	2	13	10	0.103
No Redundant Constraints	3	22	16	0.650
$L_{accountrol} = 0$	2	104	4235	157.394
Lassene $\alpha = 2$	3	275	27692	NA

Table 4.1: Comparing the computational complexity of the primal relaxation with and without the Lasserre hierarchy. All problems in both 2D and 3D used three landmarks.

5 Conclusion

In summary, we have investigated globally optimal algorithms for the stereo camera localization problem with re-projection error. First, we formulated the re-projection error problem as a QCQP and found through experimentation that strong duality does not hold for this problem in practice. Second, we attempted to tighten the QCQP by adding various redundant constraints and additional variables, but these redundant constraints did not tighten the problem. Third, we pivoted to an iterative approach, ITERSDP, for solving the stereo localization problem, motivated by prior work on monocular localization with the back-projection error. Through extensive experimentation, we found that when ITERSDP is paired with a local solver, it was able to find the globally optimal solution across many problem instances with realistic noise levels. Further, ITERSDP is far more efficient than the primal SDP relaxation of the QCQP, and it is competitive with the local solver in solution time. Finally, we investigated how to tighten the stereo localization problem by tightening analogous 1D and 2D problems. Future work should try to tighten the stereo localization problem with re-projection error by

- manually finding redundant constraints or new QCQP formulations, or
- algorithmically searching or constructing a sparse set of the constraints and variables from the Lasserre hierarchy that tightens the problem while remaining computationally tractable. This procedure could be applicable to tightening other optimization problems.

This work in tightening the re-projection error problem will have applications in many areas of robotics, including stereo visual odometry, landmark-based SLAM, and stereo localization, moving us towards more reliable robotics systems.

Bibliography

- T. D. Barfoot, *State Estimation for Robotics*, 2nd. USA: Cambridge University Press, 2022, ISBN: 1107159393.
- [2] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [3] L. Carlone, T.-J. Chin, A. Eriksson, and H. Yang, Certifiable robot perception: From global optimization to safer robots, 2020. [Online]. Available: https://mit-spark. github.io/%5C%5C%7B%7DCertifiablePerception-RSS2020/.
- [4] C. Holmes and T. D. Barfoot, An efficient global optimality certificate for landmark-based slam, 2022. DOI: 10.48550/ARXIV.2206.12961. [Online]. Available: https://arxiv.org/abs/2206.12961.
- T. D. Barfoot, C. Holmes, and F. Dumbgen, A fine line: Total least-squares line fitting as qcqp optimization, 2022. DOI: 10.48550/ARXIV.2206.05082. [Online]. Available: https://arxiv.org/abs/2206.05082.
- [6] A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin, *Rotation averaging and strong duality*, 2017. DOI: 10.48550/ARXIV.1705.01362. [Online]. Available: https://arxiv.org/ abs/1705.01362.
- [7] F. Dellaert, D. M. Rosen, J. Wu, R. Mahony, and L. Carlone, "Shonan rotation averaging: Global optimality by surfing so(pn)," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds., Cham: Springer International Publishing, 2020, pp. 292–308, ISBN: 978-3-030-58539-6.
- [8] A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin, "Rotation averaging with the chordal distance: Global minimizers and strong duality," *IEEE Transactions on Pattern Analysis* and Machine Intelligence, vol. 43, no. 1, pp. 256–268, 2021. DOI: 10.1109/TPAMI. 2019.2930051.

- [9] K. Wilson, D. Bindel, and N. Snavely, "When is rotations averaging hard?" In *Computer Vision ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham: Springer International Publishing, 2016, pp. 255–270, ISBN: 978-3-319-46478-7.
- [10] L. Carlone, G. C. Calafiore, C. Tommolillo, and F. Dellaert, "Planar pose graph optimization: Duality, optimal solutions, and verification," *IEEE Transactions on Robotics*, vol. 32, no. 3, pp. 545–565, 2016. DOI: 10.1109/TRD.2016.2544304.
- [11] L. Carlone, D. M. Rosen, G. Calafiore, J. J. Leonard, and F. Dellaert, "Lagrangian duality in 3d slam: Verification techniques and optimal solutions," in 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015, pp. 125–132. DOI: 10.1109/IROS.2015.7353364.
- [12] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.
- [13] K. N. Chaudhury, Y. Khoo, and A. Singer, "Global registration of multiple point clouds using semidefinite programming," *SIAM Journal on Optimization*, vol. 25, no. 1, pp. 468–501, 2015. DOI: 10.1137/130935458. eprint: https://doi.org/10.1137/130935458.
 [Online]. Available: https://doi.org/10.1137/130935458.
- [14] J. P. Iglesias, C. Olsson, and F. Kahl, "Global optimality for point set registration using semidefinite programming," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020.
- [15] H. Yang, J. Shi, and L. Carlone, "Teaser: Fast and certifiable point cloud registration," 2020. DOI: 10.48550/ARXIV.2001.07715. [Online]. Available: https://arxiv.org/ abs/2001.07715.
- [16] L. Carlone, *Estimation contracts for outlier-robust geometric perception*, 2022. [Online].
 Available: https://arxiv.org/abs/2208.10521.
- [17] H. Yang and L. Carlone, "Certifiably optimal outlier-robust geometric perception: Semidefinite relaxations and scalable global optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2022. DOI: 10.1109/TPAMI.2022.3179463.
- [18] H. Yang and L. Carlone, "One ring to rule them all: Certifiably robust geometric perception with outliers," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc.,

2020, pp. 18846-18859. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/da6ea77475918a3d83c7e49223d453cc-Paper.pdf.

- [19] H. Yang and L. Carlone, "A quaternion-based certifiably optimal solution to the wahba problem with outliers," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2019.
- [20] P.-Y. Lajoie, S. Hu, G. Beltrame, and L. Carlone, "Modeling perceptual aliasing in slam via discrete–continuous graphical models," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1232–1239, 2019. DOI: 10.1109/LRA.2019.2894852.
- [21] L. Carlone and G. C. Calafiore, "Convex relaxations for pose graph optimization with outliers," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1160–1167, 2018. DOI: 10.1109/LRA.2018.2793352.
- [22] M. Giamou, Z. Ma, V. Peretroukhin, and J. Kelly, "Certifiably globally optimal extrinsic calibration from per-sensor egomotion," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 367–374, Apr. 2019. DOI: 10.1109/lra.2018.2890444. [Online]. Available: https://doi.org/10.1109%2Flra.2018.2890444.
- [23] E. Wise, M. Giamou, S. Khoubyarian, A. Grover, and J. Kelly, "Certifiably optimal monocular hand-eye calibration," in 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2020, pp. 271–278. DOI: 10.1109/MF149285.2020.9235219.
- [24] J. Heller, D. Henrion, and T. Pajdla, "Hand-eye and robot-world calibration by global polynomial optimization," in 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 3157–3164. DOI: 10.1109/ICRA.2014.6907313.
- [25] S. Hu and L. Carlone, "Accelerated inference in markov random fields via smooth riemannian optimization," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1295– 1302, 2019. DOI: 10.1109/LRA.2019.2895124.
- [26] T. D. Barfoot, Introduction to mobile robotics: Lecture 19: Stereo cameras, visual odometry, 2014.
- [27] V. Tzoumas, P. Antonante, and L. Carlone, "Outlier-robust spatial perception: Hardness, general-purpose algorithms, and guarantees," in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019, pp. 5383–5390. DOI: 10.1109/ IROS40897.2019.8968174.

- [28] T.-J. Chin, Z. Cai, and F. Neumann, "Robust fitting in computer vision: Easy or hard?" In *Proceedings of the European Conference on Computer Vision (ECCV)*, Sep. 2018.
- [29] L. Sun and Z. Deng, "Certifiably optimal and robust camera pose estimation from points and lines," *IEEE Access*, vol. 8, pp. 124032–124054, 2020.
- [30] G. Terzakis and M. Lourakis, "A consistently fast and globally optimal solution to the perspective-n-point problem," in *European Conference on Computer Vision*, Springer, 2020, pp. 478–494.
- [31] M. Garcia-Salguero, J. Briales, and J. Gonzalez-Jimenez, "Certifiable relative pose estimation," *Image and Vision Computing*, vol. 109, p. 104 142, 2021, ISSN: 0262-8856. DOI: https://doi.org/10.1016/j.imavis.2021.104142. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885621000470.
- [32] J. Briales, L. Kneip, and J. Gonzalez-Jimenez, "A certifiably globally optimal solution to the non-minimal relative pose problem," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 145–154. DOI: 10.1109/CVPR.2018.00023.
- [33] J. Briales and J. Gonzalez-Jimenez, "Convex global 3d registration with lagrangian duality," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 5612–5621. DOI: 10.1109/CVPR.2017.595.
- [34] Y. Wang, X. Wen, L. Yin, C. Xu, Y. Cao, and F. Gao, "Certifiably optimal mutual localization with anonymous bearing measurements," *arXiv preprint arXiv:2203.09312*, 2022.
- [35] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on optimization*, vol. 11, no. 3, pp. 796–817, 2001.
- [36] L. Matthies and S. Shafer, "Error modeling in stereo navigation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 3, pp. 239–248, 1987.
- [37] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—a modern synthesis," in *Vision Algorithms: Theory and Practice: International Workshop* on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings, Springer, 2000, pp. 298–372.
- [38] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE transactions on robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

- [39] P. Furgale and T. Barfoot, "Stereo mapping and localization for long-range path following on rough terrain," in 2010 IEEE International Conference on Robotics and Automation, IEEE, 2010, pp. 4410–4416.
- [40] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in 2008 IEEE/RSJ international conference on intelligent robots and systems, IEEE, 2008, pp. 3946– 3952.
- [41] T. D. Barfoot, State estimation for aerospace vehicles: Aer1513 course assignments a least-squares trilogy, AER1513-003, Rev: 2.2, 4925 Dufferin Street, Toronto, Ontario, Canada M3H 5T6, Aug. 2020.

Appendices

A Back Projection Error Tightness

In this section, we confirm that the primal relaxation of the back-projection error stereo localization problem is tight, as indicated by [17], [29], [30]. The back-projection error problem in eq. (2.43) is already a QCQP. On the set of 32 simulated problems in section 3.3, we solve the primal relaxation of this QCQP, and plot the duality gap $p^* - d^*$ across all problems in fig. A.1. Notice that the duality gap is (machine) zero within this realistic noise range, confirming that strong duality holds. We used the best solution from a local solver across many initial conditions to find the globally optimal cost p^* .



Figure A.1: A plot of the duality gap of the primal relaxation of the back projection error problem. d^* is the globally optimal value and q^* is from the cost from the primal relaxation of the back-projection error problem.

B 1D problem

B.1 QCQP Formulation

In this section, we write the 1D problem:

$$\min_{x} \qquad \qquad J(x) = \sum_{n} (\omega y_n - z_n)^2 \qquad (B.1)$$

subject to

$$z_n(x - \omega a_n) = 1, \tag{B.2}$$

$$\omega^2 = 1, \tag{B.3}$$

in the standard form of a QCQP in eq. (2.29). First, we define our optimization variables:

$$\mathbf{x} = \begin{bmatrix} x \\ z_1 \\ \dots \\ z_N \\ \omega \end{bmatrix}, \tag{B.4}$$

and some helpful matrices:

$$x = \mathbf{e}_x^T \mathbf{x} = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \mathbf{x}$$
(B.5)

$$z_n = \mathbf{e}_{z_n}^T \mathbf{x} = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} \mathbf{x}$$
(B.6)

$$\omega = \mathbf{e}_{N+1}^T \mathbf{x} = \begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix} \mathbf{x}.$$
 (B.7)

Then we can re-write our cost in the standard form:

$$J(x) = \sum_{n} \left(y_n \mathbf{e}_{N+1}^T \mathbf{x} - \mathbf{e}_{z_n}^T \mathbf{x} \right)^2$$
(B.8)

$$= \mathbf{x}^{T} \left(\sum_{n} \left(y_{n} \mathbf{e}_{N+1}^{T} - \mathbf{e}_{z_{n}}^{T} \right)^{T} \left(y_{n} \mathbf{e}_{N+1}^{T} - \mathbf{e}_{z_{n}}^{T} \right) \right) \mathbf{x}, \tag{B.9}$$

and re-write the homogenization constraint in the standard form:

$$\omega^2 = 1 \implies \mathbf{x}^T \mathbf{e}_{N+1} \mathbf{e}_{N+1}^T \mathbf{x} = 1.$$
 (B.10)

Finally, for the measurement constraint:

$$z_n (x - \omega a_n) = 1 \implies \mathbf{x}^T \left(\mathbf{e}_{z_n} \mathbf{e}_x - a_n \mathbf{e}_{z_n}^T \mathbf{e}_{N+1} \right) \mathbf{x} = 1.$$
(B.11)

B.2 Local Solver

In this section, we develop a local solver for the 1D stereo localization problem given in section 4.1.

The cost function for the 1D stereo localization problem is

$$J(x) = \sum_{n} \left(y_n - \frac{1}{x - a_n} \right)^2,$$
 (B.12)

so we can write the non-linear error functions as

$$u_n = y_n - \frac{1}{x - a_n}.\tag{B.13}$$

We assume that $x = x_{op} + \delta x$, and linearize g_n about x_{op} :

$$g_n \approx g_n(x_{op}) + \frac{\partial g_n}{\partial x}|_{x=x_{op}} \delta x = g_n(x_{op}) + \frac{1}{(x_{op} - a_n)^2} \delta x = a_n + b_n \delta x.$$
(B.14)

Finally, we substitute this linearized expression for g_n into J(x) and solve for the optimal update δx^* :

$$J(x) \approx \sum_{n} (a_n + b_n \delta x)^2 \tag{B.15}$$

$$\implies \frac{\partial J}{\partial \delta x} = \sum_{n} 2(a_n + b_n \delta x^*) b_n = 0 \tag{B.16}$$

$$\implies \delta x^* = \frac{\sum_n a_n b_n}{\sum_n b_n^2}.$$
 (B.17)

Therefore the local solver algorithm for the 1D problem is

- 1. Start with an initial guess or the result from the previous iteration x_k .
- 2. Solve for the optimal update $\delta x^* = \frac{\sum_n a_n b_n}{\sum_n b_n^2}$.
- 3. If δx^* is small or a maximum number of iterations has been reached, return.
- 4. Otherwise set $x_{k+1} = x_k + \delta x^*$ and repeat.

C 2D problem

C.1 QCQP Formulation

In this section, we rewrite the 2D problem in the standard form of a QCQP. We begin with the 2D version of the re-projection error, from eq. (4.7):

$$\min_{\mathbf{T}} \qquad J(\mathbf{T}) = \sum_{n=1}^{N} \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_2^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right)^T \mathbf{W}_n \left(\mathbf{y}_n - \mathbf{M} \frac{1}{\mathbf{e}_2^T \mathbf{T} \mathbf{p}_n} \mathbf{T} \mathbf{p}_n \right) \quad (C.1a)$$

subject to
$$\mathbf{T} \in SE(2)$$
. (C.1b)

We make the substitution

$$\mathbf{v}_{n} = \frac{\mathbf{T}\mathbf{p}_{n}}{\mathbf{e}_{2}^{T}\mathbf{T}\mathbf{p}_{n}} = \begin{bmatrix} u_{n1} \\ 1 \\ u_{n2} \end{bmatrix}.$$
 (C.2)

Let

$$\mathbf{u}_n = \begin{bmatrix} u_{n1} \\ u_{n2} \end{bmatrix}.$$
 (C.3)

Now we can re-write the optimization problem as

$$\min_{\mathbf{T}} \qquad J(\mathbf{T}) = \sum_{n=1}^{N} \left(\omega \mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)^T \mathbf{W}_n \left(\omega \mathbf{y}_n - \mathbf{M} \mathbf{v}_n \right)$$
(C.4a)

subject to $\mathbf{C}^T \mathbf{C} = \mathbf{I}$

$$\mathbf{C}^{T}\mathbf{C} = \mathbf{I}$$
(C.4b)
$$\left(\mathbf{y} \ \mathbf{e}_{2}^{T} - \mathbf{I}\right)\mathbf{T}\mathbf{p} = 0$$
(C.4c)

$$\left(\mathbf{v}_{n}\mathbf{e}_{2}-\mathbf{1}\right)\mathbf{1}\mathbf{p}_{n}=0$$
(C.4c)

$$\omega^2 = 1 \tag{C.4d}$$

where

$$\mathbf{T} = \begin{bmatrix} \mathbf{C} & \mathbf{r} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{r} \\ 0 & 0 & 1 \end{bmatrix} \in SE(2).$$
(C.5)

We define our optimization variable

$$\mathbf{x} = \begin{vmatrix} \mathbf{c}_{1} \\ \mathbf{c}_{2} \\ \mathbf{r} \\ \mathbf{u}_{1} \\ \dots \\ \mathbf{u}_{N} \\ \omega \end{vmatrix} \in \mathbb{R}^{2N+7}.$$
(C.6)

From this point, we follow a similar procedure to that described in section 3.2 to re-write the QCQP in the standard form.

C.2 Local Solver

In this section, we develop a local solver for the 2D stereo localization problem. We parameterize the 2D pose **T** with $\phi = \begin{bmatrix} x & y & \theta \end{bmatrix}^T$, where x and y are the position of the camera origin in the world frame, and θ is the yaw angle of the camera frame with respect to the world frame. Referring to eq. (4.7), we define the error function as

$$\mathbf{g}_n = \mathbf{y}_n - \mathbf{M} \frac{\mathbf{q}_n}{\mathbf{e}_2^T \mathbf{q}_n}, \quad \mathbf{q}_n = \mathbf{T} \mathbf{p}_n.$$
 (C.7)

Then we can write

$$\frac{\partial \mathbf{g}_n}{\partial \boldsymbol{\phi}} = \frac{\partial \mathbf{g}_n}{\partial \mathbf{q}_n} \frac{\partial \mathbf{q}_n}{\partial \boldsymbol{\phi}} \tag{C.8}$$

(C.9)

where

$$\frac{\partial \mathbf{g}_n}{\partial \mathbf{q}_n} = -\mathbf{M} \frac{1}{\mathbf{e}_2^T \mathbf{q}_n} + \mathbf{M} \frac{\mathbf{q}_n}{\left(\mathbf{e}_2^T \mathbf{q}_n\right)^2} \mathbf{e}_2^T$$
(C.10)

and

$$\frac{\partial \mathbf{q}_n}{\partial \boldsymbol{\phi}} = \begin{bmatrix} 1 & 0 & -\sin(\theta)p_{n,x} - \cos(\theta)p_{n,y} \\ 0 & 1 & \cos(\theta)p_{n,x} - \sin(\theta)p_{n,y} \\ 0 & 0 & 0 \end{bmatrix}, \quad (C.11)$$

where $\mathbf{p}_n = \begin{bmatrix} p_{n,x} & p_{n,y} & 1 \end{bmatrix}^T$. We assume $\boldsymbol{\phi} = \boldsymbol{\phi}_{op} + \delta \boldsymbol{\phi}$, where we want to solve for the optimal update $\delta \boldsymbol{\phi}$. We can linearize \mathbf{g}_n about $\boldsymbol{\phi}_{op}$:

$$\mathbf{g}_{n}(\boldsymbol{\phi}) \approx \mathbf{g}_{n}(\boldsymbol{\phi}_{op}) + \frac{\partial \mathbf{g}_{n}}{\partial \boldsymbol{\phi}}|_{\boldsymbol{\phi} = \boldsymbol{\phi}_{op}} = \boldsymbol{\beta}_{n} + \boldsymbol{\Delta}_{n}^{T} \delta \boldsymbol{\phi}.$$
(C.12)

From here, we can follow a similar procedure to the 3D problem, starting in eq. (2.16); we substitute this linearized expression into J to obtain a cost that is quadratic in the variable $\delta\phi$ and solve for the optimal update $\delta\phi^*$. Our final algorithm is:

- 1. Start with an initial guess or the result from the previous iteration ϕ_k .
- 2. Solve for the optimal update $\delta \phi^*$.
- 3. If $\delta \phi^*$ is small or a maximum number of iterations has been reached, return.
- 4. Otherwise set $\phi_{k+1} = \phi_k + \delta \phi^*$ and repeat.