A Program of Research for Globally Optimal State Estimation

Frederike Dümbgen* Connor Holmes† Timothy D. Barfoot†

*Inria, École Normale Supérieure, PSL Research University, Paris, France
†Robotics Institute, University of Toronto, Toronto, Canada

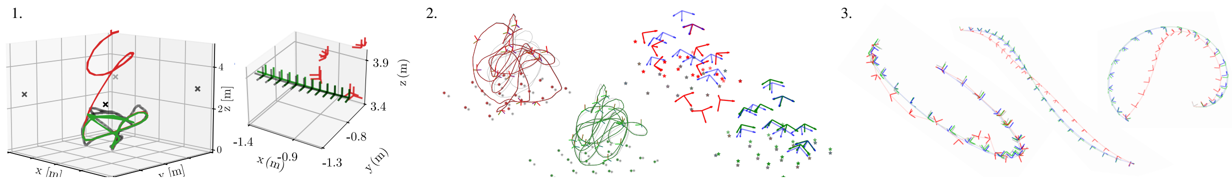

Fig. 1. Selection of classic estimation problems, depicting poor local minima (in red) and the corresponding, certified, global minima (in green). The ground truth is depicted in gray and black. Problems 1. involve range-only estimation, 2. involve landmark-based estimation, and 3. involve trajectory estimation. A concise overview of the formulations of these and other problems we have treated can be found in Table I.

***Abstract*—We provide a concise overview of our progress in identifying and certifying globally optimal solutions of state estimation problems in robotics. We give a summary of the theoretical background, providing pointers for novices in the field to learn about this topic. We then given an overview of the problems we have treated, putting them in a unified framework to make it easier to find, compare, and build upon them. We discuss our methods for simplifying the generation and analysis of these solutions, and finally, present advances to allow to embed them in real-world robotics pipelines. We conclude with a discussion of important open research problems.**

I. Introduction

Reliable state estimation is the foundation of most successful robotics applications. With the proliferation of batch estimation, where state estimation is generally posed as a nonlinear least squares problem, fast local solvers have become a frequent component of the robotics pipelines. They allow to solve problems in more than thousands of variables quickly and with little memory consumption, enabling their use in real-time applications and on embedded platforms with limited computing resources [1], [2]. However, by definition, these local solvers may converge to local minima that may be far from the globally optimal solution, as shown in Figure 1. Relying on such solutions without a verification mechanism may result in performance degradation and even catastrophic consequences. This paper summarizes our efforts from the last two years to address the risk of local minima in robotics by deriving *certifiably optimal* solutions to classic state estimation problems. We structure the paper along the three frontiers that we continue to explore.

a) Extending the catalogue: In Section IV-A, we give a coherent overview of the new problems that we have globally solved or certified to date, thus extending the catalogue of optimally solvable state estimation problems in the literature. We provide the chronological thread and put all problems in a unified notation to facilitate identifying their commonalities and differences. By exposing the thought process behind the developed solutions, we hope to enable not only a deeper understanding of the material, but also to empower researchers and practitioners to contribute to this important field.

b) Simplifying the onboarding process: Most sufficiently realistic problems that we have looked at are NP-hard to solve to global optimality. Therefore, by definition, the methodology used in Section IV-A does not always work. It does work when the semidefinite relaxations developed in the process are *tight*, a concept which we briefly explain in Section III. Understanding and simplifying the process of achieving tightness for a new problem are therefore essential, and our advances in this direction are summarized in Section IV-B.

c) Making solutions practical: Optimality is only one axis in the landscape of desiderata for solvers in robotic state estimation. Our third frontier revisits two of the other fundamental requirements: the solver speed and its integration within end-to-end learned pipelines. We outline our advances to make semidefinite program (SDP) solvers faster by exploiting sparsity, and our work on global optimality in differentiable optimization, in Section IV-C.

II. Problem Statement

The core subject of this line of work is the MAP estimator, a probabilistic approach that typically amounts to minimizing the weighted sum of squared errors between the output of a sensor forward model and the measurement [3], and (optionally) prior terms. We assume that all measurements and priors are recorded in a measurement graph with nodes corresponding to variables and known parameters and edges corresponding to measurements. All problems discussed here then take the form

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min} \sum_{(i,j)\in\mathcal{E}} \boldsymbol{e}_{ij}(\boldsymbol{\theta},\boldsymbol{\Psi})^\top \boldsymbol{W}_{ij}\boldsymbol{e}_{ij}(\boldsymbol{\theta},\boldsymbol{\Psi}), \tag{1}$$

where $\mathcal{E}$ is the edge set, which we will further split into subsets $\mathcal{R}\subset\mathcal{E}$ and $\mathcal{A}\subset\mathcal{E}$ for relative measurements (*i.e.*, inertial measurement unit (IMU) measurements or motion priors) and

This research was conducted while the first author was at University of Toronto, and it was funded in part by the Swiss National Science Foundation, Postdoc Mobility under Grant 206954 and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

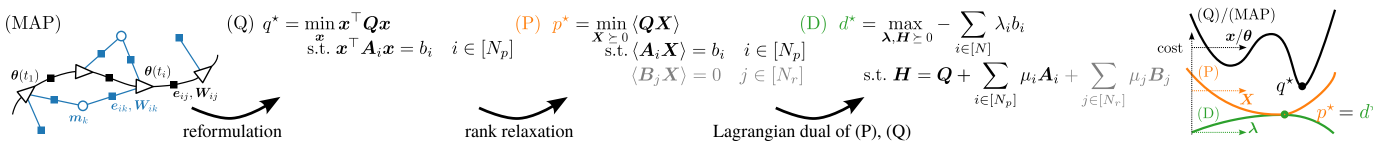

Fig. 2. Depiction of an example MAP problem (1) (in factor graph form), its QCQP reformulation (Q), its rank relaxation, also called primal (P) and the Lagrangian dual problem of (Q) and (P), denoted by (D). The graphs on the right show a simplified depiction of the relationship between these problems.

absolute measurements (*i.e.*, range measurements to landmarks), respectively. The variable $\boldsymbol{\theta}$ may be a continuous-time (CT) or discrete-time (DT) representation of the state, and Θ is its feasible set. All other problem parameters are combined in $\boldsymbol{\Psi}$, where we distinguish between $\boldsymbol{\Psi}_f$ for fixed parameters, such as known landmarks, and $\boldsymbol{\Psi}_m$ for changing entities such as measurements. The terms $\boldsymbol{e}$ and $\boldsymbol{W}$ are the error terms and weighting matrices derived from the measurement or prior models, respectively. A general measurement graph associated with this optimization problem is shown in Figure 2.

We use $SO(d)$ and $SE(d)$ to denote the special orthogonal and Euclidean group, respectively. We use $\boldsymbol{r} \in \mathbb{R}^d$ to denote translation in d dimensions, and $\boldsymbol{R} \in SO(d)$ for rotation. We introduce $\boldsymbol{T} \in SE(d)$, the associated transformation matrix, and use superscript i to denote the frame in which we operate, but dropping it when we operate in the inertial frame. In our conventions, a vector $\boldsymbol{m}_k$ in the inertial frame can be expressed in frame i by using: $\boldsymbol{m}_k^i = \boldsymbol{K}\boldsymbol{T}^i \bar{\boldsymbol{m}}_k = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}^i & \boldsymbol{r}^i \\ \boldsymbol{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{m}_k \\ 1 \end{bmatrix}$, where the identity matrix and vector of all zeros are denoted by $\boldsymbol{I}$ and $\boldsymbol{0}$, respectively. We let $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle$ be the standard matrix inner product, and $\boldsymbol{X} \succeq 0$ means matrix $\boldsymbol{X}$ is positive semidefinite (PSD).

Often, the sensor models in (1) are nonlinear, or Θ is a non-convex set, in which case (1) is non-convex and may be hard to solve to global optimality. In the next section, we see how to relax problems of the form (1) to SDPs and how to exploit those for global optimality certificates. Readers familiar with certifiable estimation may safely skip directly to Section IV.

III. Background on Semidefinite Relaxations

We give a brief overview of the standard machinery used in certifiably optimal estimation. All problem formulations that we refer to in this section are given in Figure 2.

a) Semidefinite relaxations: Our approach uses QCQPs and their relaxations, which encompasses numerous estimation problems. Indeed, many problems of the form (1) are in fact a polynomial optimization problem (POP) (polynomial cost and constraints). It is always possible to transform any POP into a QCQP by introducing so-called substitution variables. As a simple example, consider a *cubic* function of the form $p(\theta) = \theta^3$. By introducing a new variable, z, and associated *quadratic* constraint, $z = \theta^2$, we can express this function *quadratically* as $p(z, \theta) = z\theta$. In the final formulation, we typically collect all variables in a vector, $\boldsymbol{x}$, and formulate the cost and constraint accordingly. A comprehensive outline of this process is given in [4]. The general form of the thus obtained QCQP, which we call (Q), is given in Figure (2).

This non-convex QCQP reformulation is still as difficult to solve as the original problem (1). However, QCQPs have a well-studied rank relaxation and Lagrangian dual problem, which we denote by (P) (for primal) and (D), respectively, in what follows. They are both SDPs with standard forms given in Figure 2. Detailed derivations of these problems can be found in [5], for example. Problem (D) is a concave lower-bounding function of the original problem, and (P) is a convex lower-bounding function. Under mild assumptions, we have, for their respective optimal values, $d^\star = p^\star$, but the optimal value of (QCQP) may be higher in general. When we also have $d^\star = p^\star = q^\star$, we say that strong duality holds, or that we have *cost tightness*. If we additionally have that $\operatorname{rank}(\boldsymbol{X}^\star) = 1$ for the optimizer of (P), we say that we have *rank tightness*.

b) Certifying optimality: We use two different approaches to certify optimality. In the first approach, a standard local solver produces a candidate solution and the dual problem (D) is used to derive an optimality certificate based on properties of the dual matrix, $\boldsymbol{H}$, which is therefore referred to as the *certificate matrix*.

Cost certificate. *Let $\hat{\boldsymbol{x}}$ be a solution candidate (a second-order critical point) of (Q). If there exists $\boldsymbol{\lambda}^\star$ such that:*

$$\boldsymbol{H}(\boldsymbol{\lambda}^\star)\hat{\boldsymbol{x}} = \boldsymbol{0} \ \textit{(stationarity)}, \tag{2a}$$

$$\boldsymbol{H}(\boldsymbol{\lambda}^\star) \succeq 0 \ \textit{(dual feasibility)}, \tag{2b}$$

then $\hat{\boldsymbol{x}} = \boldsymbol{x}^\star$ is the globally optimal solution of (Q).

This is very powerful in practice: it means that, given a local solution of the original problem, we can certify its global optimality by solving the feasibility problem (2). As we will see, for certain problems, in particular when Linearly Independent Constraint Qualification (LICQ) holds, (2a) admits a closed-form solution for $\boldsymbol{\lambda}^\star$, which can then be plugged into (2b) to certify global optimality. When LICQ does not hold, the key challenge in certification involves solving for the optimal Lagrange multipliers. As we will see, this is the case when redundant constraints are used to tighten the relaxation.

The second approach consists of solving (P) using an SDP solver and checking the rank of the solution.

Rank certificate. *Let $\hat{\boldsymbol{X}}$ be a solution of (P). If* $\operatorname{rank}(\hat{\boldsymbol{X}}) = 1$ *then we can obtain the exact factorization $\hat{\boldsymbol{X}} = \boldsymbol{x}^\star \boldsymbol{x}^{\star\top}$, where $\boldsymbol{x}^\star$ is the globally optimal solution of (Q).*

c) Improving tightness: The ability to use the two above approaches depends on the tightness of the relaxations. We review three important concepts used to address tightness, that are fundamental to our contributions.

Redundant constraints: It is well known (see, *e.g.*, [12]) that adding constraints that are redundant in (Q) (but non-redundant in (P) and (D)) can improve the tightness of the relaxation. Intuitively speaking, such constraints can be used

TABLE I
OVERVIEW OF STATE ESTIMATION PROBLEMS THAT CAN BE WRITTEN AS POLYNOMIAL PROBLEMS WITH TIGHT SEMIDEFINITE RELAXATIONS.

	relative errors $\boldsymbol{e}_{ij}$	absolute errors $\boldsymbol{e}_{ik}$	weights	variables $\boldsymbol{\theta} \subset \Theta$	known $\boldsymbol{\Psi}_f$	measured $\boldsymbol{\Psi}_m$	reformulation	red.	reference
1.a	$\begin{bmatrix} \Delta t_i \boldsymbol{v}_i - (\boldsymbol{r}_i - \boldsymbol{r}_j) \\ \boldsymbol{v}_i - \boldsymbol{v}_j \end{bmatrix}$	$\tilde{d}_{ik}^2 - \lVert \boldsymbol{r}_i - \boldsymbol{m}_k \rVert^2$	aniso.	$\boldsymbol{r}(t) \in \mathbb{R}^d, \boldsymbol{v}(t) \in \mathbb{R}^d$	$\boldsymbol{m}_k$	$\tilde{d}_{ik}$	$z_i = \lVert \boldsymbol{r}_i \rVert^2$	no	[6] (WNOA)
1.b	—	$\tilde{d}_{ik}^2 - \lVert \boldsymbol{r}_i - \boldsymbol{m}_k \rVert^2$	iso.*	$\boldsymbol{r}_i \in \mathbb{R}^d$	$\boldsymbol{m}_k$	$\tilde{d}_{ik}$	$\boldsymbol{z}_i = \operatorname{vech}(\boldsymbol{r}_i \boldsymbol{r}_i^\top)$	yes	[7] (RO)
1.c	—	$\tilde{d}_{ik,j}^2 - \lVert \boldsymbol{K} \boldsymbol{T}_i \boldsymbol{p}_j - \boldsymbol{m}_k \rVert^2$	iso.*	$\boldsymbol{T}_i \in SE(d)$	$\boldsymbol{m}_k, \boldsymbol{p}_j$	$\tilde{d}_{ik,j}$	$\boldsymbol{\ell}_{ij} = \boldsymbol{K}\boldsymbol{T}_i \boldsymbol{p}_j,\ z_{ij} = \lVert \boldsymbol{\ell}_{ij} \rVert^2$	yes	[8] (Static)
1.d	—	$\tilde{d}_{ik,j}^2 - \left\lVert \boldsymbol{K}\boldsymbol{T} e^{\Delta t_i \boldsymbol{\omega}^\wedge} \boldsymbol{p}_j - \boldsymbol{m}_k \right\rVert^2$	iso.*	$\boldsymbol{T} \in SE(d), \boldsymbol{\omega} \in \mathbb{R}^{2d}$	$\boldsymbol{m}_k, \boldsymbol{p}_j$	$\tilde{d}_{ik,j}$	$\boldsymbol{\ell}_{ij} \approx \boldsymbol{K}(\boldsymbol{I} + \Delta t_i \boldsymbol{\omega}^\wedge)\boldsymbol{T}\boldsymbol{p}_j,$ $z_{ij} = \lVert \boldsymbol{\ell}_{ij} \rVert^2$	yes	[8] (Dynamic)
2.a	$\operatorname{vec}\left(\tilde{\boldsymbol{T}}_{ij}\boldsymbol{T}_j - \boldsymbol{T}_i\right)$	$\tilde{\boldsymbol{m}}_k - \boldsymbol{K}\boldsymbol{T}_i \bar{\boldsymbol{m}}_k$	iso.	$\boldsymbol{T}_i \in SE(d)$	$\boldsymbol{m}_k$	$\tilde{\boldsymbol{m}}_k, \tilde{\boldsymbol{T}}_{ij}$	—	no	[9] (SLAM)
2.b	—	$\tilde{\boldsymbol{m}}_k - \boldsymbol{K}\boldsymbol{T}_i \boldsymbol{m}_k$	aniso.	$\boldsymbol{T}_i \in SE(d)$	$\boldsymbol{m}_k$	$\tilde{\boldsymbol{m}}_k$	$\boldsymbol{z}_{ik} = \boldsymbol{K}\boldsymbol{T}_i \boldsymbol{m}_k$	yes	[10] (Wahba)
2.c	$\operatorname{vec}\left(\tilde{\boldsymbol{T}}_{ij}\boldsymbol{T}_j - \boldsymbol{T}_i\right)$	$\tilde{\boldsymbol{m}}_k - \boldsymbol{K}\boldsymbol{T}_i \boldsymbol{m}_k$	mixed	$\boldsymbol{T}_i \in SE(d)$	$\boldsymbol{m}_k$	$\tilde{\boldsymbol{m}}_k, \tilde{\boldsymbol{T}}_{ij}$	$\boldsymbol{z}_{ik} = \boldsymbol{K}\boldsymbol{T}_i \boldsymbol{m}_k$	yes	[10] (SLAM)
2.d	—	$\tilde{\boldsymbol{y}}_k - (\boldsymbol{T}_i \boldsymbol{m}_k)_d^{-1} \boldsymbol{C}\boldsymbol{T}_i \boldsymbol{m}_k$	iso.*	$\boldsymbol{T}_i \in SE(d)$	$\boldsymbol{m}_k, \boldsymbol{C}$	$\tilde{\boldsymbol{y}}_k$	$\boldsymbol{z}_{ik} = (\boldsymbol{T}_i \boldsymbol{m}_k)_d^{-1} \boldsymbol{C}\boldsymbol{T}_i \boldsymbol{m}_k$	yes	[7] (Stereo)
3.a	—	$\operatorname{cay}^{-1}\left(\boldsymbol{T}\tilde{\boldsymbol{T}}_i^{-1}\right)$	aniso.	$\boldsymbol{T} \in SE(d)$	—	$\tilde{\boldsymbol{T}}_i$	$\boldsymbol{z}_i = \operatorname{cay}^{-1}\left(\boldsymbol{T}\tilde{\boldsymbol{T}}_i^{-1}\right)$	yes	[11] (Pose avg.)
3.b	$\operatorname{cay}^{-1}\left(\boldsymbol{T}_i \boldsymbol{T}_j^{-1} \tilde{\boldsymbol{T}}_{ij}^{-1}\right)$	$\operatorname{cay}^{-1}\left(\boldsymbol{T}_i \tilde{\boldsymbol{T}}_i^{-1}\right)$	aniso.	$\boldsymbol{T}_i \in SE(d)$	—	$\tilde{\boldsymbol{T}}_i, \tilde{\boldsymbol{T}}_{ij}$	$\boldsymbol{z}_i = \operatorname{cay}^{-1}\left(\boldsymbol{T}_i \tilde{\boldsymbol{T}}_i^{-1}\right)$ $\boldsymbol{z}_{ij} = \operatorname{cay}^{-1}\left(\boldsymbol{T}_i \boldsymbol{T}_j^{-1} \tilde{\boldsymbol{T}}_{ij}^{-1}\right)$	yes	[11] (DT synch.)
3.c	$\begin{bmatrix} \Delta t_i \boldsymbol{\omega}_j - \operatorname{cay}^{-1}\left(\boldsymbol{T}_i \boldsymbol{T}_j^{-1}\right)^\vee \\ \boldsymbol{\omega}_i - \boldsymbol{\omega}_j \end{bmatrix}$	$\operatorname{cay}^{-1}\left(\boldsymbol{T}_i \tilde{\boldsymbol{T}}_i^{-1}\right)$	aniso.	$\boldsymbol{T}_i \in SE(d)$	—	$\tilde{\boldsymbol{T}}_i$	$\boldsymbol{z}_i = \operatorname{cay}^{-1}\left(\boldsymbol{T}_i \tilde{\boldsymbol{T}}_i^{-1}\right)$ $\boldsymbol{z}_{ij} = \operatorname{cay}^{-1}\left(\boldsymbol{T}_i \boldsymbol{T}_j^{-1}\right)$	yes	[11] (CT synch.)

Δt_i time interval $\boldsymbol{v}_i/\boldsymbol{w}_i$ linear/generalized velocity $\boldsymbol{C}$ stereo camera matrix $\tilde{d}_{ik}, \tilde{\boldsymbol{y}}_{ik}, \tilde{\boldsymbol{m}}_{ik}, \tilde{\boldsymbol{T}}_{ij}$ measurements
iso. isotropic, *aniso.* anisotropic, *iso.** easily extended to anisotropic vech $(\cdot)$ half-vectorization $^\wedge, ^\vee$ skew-symmetric operator and its inverse

to reintroduce structure in the primal (P) that is lost when relaxing the rank constraint.

Lasserre's hierarchy: Finally, it was shown that by adding enough higher-order terms to $\boldsymbol{x}$ in (Q), along with associated (primary and possibly redundant) constraints, the reformulation will eventually have a tight relaxation [12], [13].

SDP stability: Once tightness has been estabilished for a given problem, a key result presented by [14] allows us to extend this tightness results to other problems with similar parameters (*e.g.*, state-estimation problems that are tight for zero noise, remain tight at low noise levels).

IV. SUMMARY OF CONTRIBUTIONS

Over the last years, researchers have identified many problems that admit polynomial formulations and can thus be treated using the methodology outlined in Section III; examples include (range-aided) pose-graph optimization [15], rotation averaging [16], and robust estimation [4]. Taking a fresh look at classical state estimation problems, we have focused particularly on landmark-based, range-based, and rotation-based estimation, which we describe next. To facilitate their comparison, we write each considered problem in the form (1) and provide an overview of the different components for each problem in Table I.

A. Extending the catalogue

Our first work on **landmark-based estimation**, [9] (**2.a** in Table I) showed how to handle landmark variables efficiently in the framework proposed in [15], and introduced an efficient cost certificate. The certificate is applied to landmark-based simultaneous localization and mapping (SLAM) and its subgroups such as bundle adjustment and pose-graph optimization. In these first works, an isotropic noise model was used in the formulation of the MAP problem, which is crucial for certain simplifications of the cost function. In our follow-up work, [10] (**2.b** and **2.c**), we show that when using anisotropic noise models, we need to introduce higher-order substitutions and a relatively large number of redundant constraints for tightness. In methods 2.a to 2.c, we assume linear measurement models, so when dealing with camera images, pixel measurements of features are first projected to Euclidean space using the inverse camera model. In a later, independent line of work [7] (**2.d**), we studied a method that minimizes the reprojection error (in pixel space) instead, which better captures noise on pixel measurements. This relaxation turned out to be significantly less tight, and required adding bilinear terms and a significant number of redundant constraints to $\boldsymbol{x}$ (thus going up one level in Lassere's hierarchy [12]).

In parallel work, we investigated **range-based estimation** problems. In range-only localization, the trajectory of a device is to be determined based on possibly asynchronous distance measurements from known and fixed anchor points, for example based on ultra-wideband (UWB) measurements. When we can solve for one position at a time (*i.e.*, when we get more than $d + 1$ non-degenerate measurements at each timestep) globally optimal solvers exist [17], [18]. We instead treated the more realistic underdetermined case (**1.a**) [6] where we use a Gaussian process (GP) motion prior to regularize the problem and obtain a CT trajectory estimate. GP motion priors allow to incorporate physical assumptions such as the widely used white-noise-on-acceleration (WNOA) prior. Since the GP only consists of augmenting the state with temporal derivatives and adding quadratic cost terms, it is a suitable candidate for polynomial optimization and hence SDP relaxations and indeed, the relaxation was found to be tight without redundant constraints.[1] Finally, we derived certifiably optimal solvers for estimating a robot's pose, using multiple distance tags on the robot to determine its orientation (**1.c** and **1.d**). This problem is significantly higher-dimensional than problems 1.a and 1.b and requires redundant constraints for tightness, even when using the minimal substitution. Therefore, when trying to solve for multiple poses at once, we resort to a constant-velocity reparametrization of the trajectory (1.d).

A more permissive approach than the 'hard' constant-velocity reparametrization used in (1.d) is to use a 'soft' motion prior based on the constant-velocity assumption, for

[1] Later, we found that a less minimal substitution than in [6], [17] requires the addition of redundant constraints for tightness (**1.b**).

which the GP method [19], [20] is a suitable candidate, as done in (1.a). However, when rotation variables are involved in this prior, the formulation includes a matrix exponential, which leads to a non-polynomial optimization problem. As a remedy, we proposed to use the Cayley map instead of the exponential map, which is similar in terms of the noise characteristics it may capture, but leads to polynomial constraints instead. Applying this map to CT pose synchronization problems leads to method (**3.c**). In fact, the Cayley map presents an alternative way to globally solve many other categories of **trajectory estimation problems**, such as DT pose synchronization (**3.b**), pose averaging (**3.a**) and rotation averaging (omitted for brevity). Formulations 3.a to 3.c showcase a progression in complexity, accompanied by a growth of the number of redundant constraints required for rank tightness.

B. Simplifying the onboarding process

We saw that many of the studied problems used redundant constraints to tighten the relaxations. Finding these constraints for a new problem has been of paramount importance in our and prior works, but the search process usually requires good intuitions or lengthy manual work. To make this process more scalable and generalizable, we have developed a method to automatically determine a sufficient set of redundant constraints. This method, which we dubbed AUTOTIGHT is what allowed us to find the redundant constraints for all problems in Table I. The method proceeds by determining the nullspace of a number of rank-1 feasible samples of (P) for an example problem. Each nullspace vector corresponds to one learned redundant constraint. These learned constraints can then be generalized to new problem instances through a templating mechanism [7].

Finally, in [10], we show that we can better understand tightness by drawing the link between the certificate matrix and the uncertainty of the posterior estimate. Indeed, the posterior uncertainty is tightly coupled to the Fisher Information Matrix (FIM), defined as the second derivative of the negative log-posterior distribution, which, in turn, is linked to the (Riemannian) Hessian of the optimization problem. The (Riemannian) Hessian, on the other hand, has a direct link with the certificate matrix $\boldsymbol{H}$. We show in [10] that a higher posterior uncertainty affects the tightness of the semidefinite relaxation (SDR) — for example, when the landmarks are configured such that uncertainty is concentrated along a given axis, the relaxation becomes less tight.

C. Making solutions practical

Despite the progress mentioned in the previous sections, the question of whether certifiable solutions can keep up with the demands of real robotics applications remains. To be viable, these algorithms need to be fast enough for realtime operation while still remaining accurate. Moreover, certifiable techniques must integrate well in modern robotics pipelines, which increasingly leverage deep learning.

Most real-time solvers of MAP estimation problems rely on the efficient exploitation of sparse problem structure [2]. Similarly, many of our developed optimality certificates exploit sparsity explicitly. We design an efficient representation of the sparse cost matrix in landmark-based estimation [9], and propose a linear-time PSD test of the sparse certificate matrix in [6]. In our most recent work [21] we further show that in many problems, chordal sparsity can be exploited for a significant speedup, in particular in state estimation problems without loop closures, such as localization. An important finding in this work, particularly considering the number of problems that require redundant constraints for tightness, is that the *aggregate sparsity* is the determining factor — meaning, the sparsity considering both cost and constraint matrices.

Recent integration of classical optimization approaches into end-to-end learning pipelines has been enabled by research into *differentiable optimization*. However, when the optimization is non-convex, convergence to local minima can lead to incorrect gradient information that corrupts the training process. Our recent work shows that, by leveraging results from differentiable convex optimization, certifiable algorithms can be used directly in end-to-end robotics learning pipelines [22]. We showed that certifiable methods always provide correct gradient information and applied this approach to train and run a robotic, outdoor localization pipeline that uses deep-learning to train lighting-invariant features.

V. OUTLOOK

We have seen that certifiable algorithms have begun to be adopted in robotics, primarily in perception and state estimation, but also in other areas such as planning, where promising first advances have already been made [23]. However, we believe that for these algorithms to be used more widely in robotics in general, progress must be made in three key directions:

Theory: A deeper theoretical understanding of the mechanisms that lead to tight SDP relaxations is sorely needed, especially in terms of the cause of loss of tightness and a systematic approach to efficiently regain tightness. We have made some initial steps along both of these lines (see [7], [10]), but our approach thus far has been largely empirical. Investigation of certificates of *solution accuracy* in addition to global optimality are also an important avenue for ensuring the safe operation of robots in general [24].

Exploiting Sparsity: Particularly in state estimation, it is well known that many large-scale problems of interest exhibit significant sparsity. It is imperative that algorithmic progress is made to efficiently exploit this sparsity. In online applications, this involves addressing loop closures and enabling cheap incremental updates, similar to the approach by [2] for local optimization. Furthermore, finding sparsity-preserving redundant constraints is important to harvest the speedups described in Section IV-C.

Implementation: The theoretical background alone presents significant ‘barrier-to-entry’ for certifiable methods if practitioners are required to build solutions from scratch. Although some certifiable, plug-and-play, backend solutions for robotics problems exist (e.g., [15], [25]), more are needed to spur widespread adoption. In a similar vein, open-source, robotics-friendly, SDP solvers with key features such as warm starting, parallelization, and anytime solutions are also important.

REFERENCES

[1] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," *International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.

[2] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the bayes tree," *International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

[3] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2017.

[4] 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*, vol. 45, no. 3, pp. 2816–2834, 2023.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[6] F. Dümbgen, C. Holmes, and T. D. Barfoot, "Safe and Smooth: Certified Continuous-Time Range-Only Localization," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 1117–1124, 2023.

[7] F. Dümbgen, C. Holmes, B. Agro, and T. D. Barfoot, "Toward Globally Optimal State Estimation Using Automatically Tightened Semidefinite Relaxations," *arXiv:2308.05783 [cs]*, 2023.

[8] A. Goudar, F. Dümbgen, T. D. Barfoot, and A. P. Schoellig, "Optimal Initialization Strategies for Range-Only Trajectory Estimation," *IEEE Robotics and Automation Letters*, vol. 9, no. 3, pp. 2160–2167, 2024.

[9] C. Holmes and T. D. Barfoot, "An Efficient Global Optimality Certificate for Landmark-Based SLAM," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1539–1546, 2023.

[10] C. Holmes, F. Dümbgen, and T. D. Barfoot, "On Semidefinite Relaxations for Matrix-Weighted State-Estimation Problems in Robotics," *arXiv:2308.07275 [cs]*, 2024.

[11] T. D. Barfoot, C. Holmes, and F. Dümbgen, "Certifiably Optimal Rotation and Pose Estimation Based on the Cayley Map," *arXiv:2308.12418 [cs]*, 2023.

[12] J. B. Lasserre, "Global Optimization with Polynomials and the Problem of Moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[13] J. Nie, "Optimality conditions and finite convergence of Lasserre's hierarchy," *Mathematical Programming*, vol. 146, no. 1, pp. 97–121, 2014.

[14] D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, "On the local stability of semidefinite relaxations," *Mathematical Programming*, vol. 193, pp. 629–663, 2022.

[15] 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," *International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.

[16] A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin, "Rotation averaging and strong duality," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 127–135.

[17] A. Beck, P. Stoica, and J. Li, "Exact and Approximate Solutions of Source Localization Problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1770–1778, 2008.

[18] M. Larsson, V. Larsson, K. Astrom, and M. Oskarsson, "Optimal trilateration is an eigenvalue problem," in *IEEE ICASSP*, 2019, pp. 5586–5590.

[19] T. Barfoot, C. Hay Tong, and S. Sarkka, "Batch continuous-time trajectory estimation as exactly sparse Gaussian process regression," in *Robotics: Science and Systems*, 2014.

[20] S. Anderson, T. D. Barfoot, C. H. Tong, and S. Särkkä, "Batch Nonlinear Continuous-Time Trajectory Estimation as Exactly Sparse Gaussian Process Regression," *Autonomous Robots*, pp. 221–238, 2015.

[21] F. Dümbgen, C. Holmes, and T. D. Barfoot, "Exploiting chordal sparsity for fast global optimality with application to localization," *arXiv:2406.02365 [cs]*, 2024.

[22] C. Holmes, F. Dümbgen, and T. D. Barfoot, "SDPRLayers: Certifiable Backpropagation Through Polynomial Optimization Problems in Robotics," *arXiv:2405.19309*, May 2024. arXiv: 2405.19309 [cs].

[23] T. Marcucci, M. Petersen, D. von Wrangel, and R. Tedrake, "Motion planning around obstacles with convex optimization," *Science Robotics*, vol. 8, no. 84, 2023.

[24] L. Carlone, "Estimation Contracts for Outlier-Robust Geometric Perception," *Foundations and Trends in Robotics*, vol. 11, no. 2-3, pp. 90–224, 2023.

[25] H. Yang, J. Shi, and L. Carlone, "TEASER: Fast and Certifiable Point Cloud Registration," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 314–333, 2021.