Pose Graph Optimization in the Complex Domain: Duality, Optimal Solutions, and Verification

Luca Carlone, Giuseppe C. Calafiore, and Frank Dellaert

Abstract-Pose Graph Optimization (PGO) is the problem of estimating a set of poses from pairwise relative measurements. PGO is a nonconvex problem, and currently no known technique can guarantee the computation of a global optimal solution. In this paper, we show that Lagrangian duality allows computing a globally optimal solution, and enables to certify optimality of a given estimate. Our first contribution is to frame PGO in the complex domain. This makes analysis easier and allows drawing connections with existing literature on unit gain graphs. The second contribution is to formulate and analyze the properties of the Lagrangian dual problem in the complex domain. Our analysis shows that the duality gap is connected to the number of eigenvalues of the penalized pose graph matrix, which arises from the solution of the dual. We prove that if this matrix has a single eigenvalue in zero, then (i) the duality gap is zero, (ii) the primal PGO problem has a unique solution, and (iii) the primal solution can be computed by scaling an eigenvector of the penalized pose graph matrix. The third contribution is algorithmic: we leverage duality to devise and algorithm that computes the optimal solution when the penalized matrix has a single eigenvalue in zero. We also propose a suboptimal variant when the eigenvalues in zero are multiple. Finally, we show that duality provides computational tools to verify if a given estimate (e.g., computed using iterative solvers) is globally optimal. We conclude the paper with an extensive numerical analysis. Empirical evidence shows that in the vast majority of cases (100%) of the tests under noise regimes of practical robotics applications) the penalized pose graph matrix has a single eigenvalue in zero, hence our approach allows computing (or verifying) the optimal solution.

I. INTRODUCTION

Pose graph optimization (PGO) consists in the estimation of a set of poses (positions and orientations) from relative pose measurements. The problem can be formulated as a nonconvex minimization, and can be visualized as a graph, in which a (to-be-estimated) pose is attached to each vertex, and a given relative pose measurement is associated to each edge.

PGO is a key problem in many application endeavours. In robotics, it lies at the core of state-of-the-art algorithms for localization and mapping in both single robot [44], [18], [22], [47], [26], [9], [17], [33], [10], [11] and multi robot [36], [1], [37], [32] systems. In the single robot case, the to-be-estimated poses are sampled along the robot trajectory and the relative measurements are obtained by processing raw sensor data; examples of trajectories for robotic benchmarking datasets are shown in Fig. 1. In computer vision and control, problems

G.C. Calafiore is with the Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy. E-mail: giuseppe.calafiore@polito.it

F. Dellaert is with the College of Computing, Georgia Institute of Technology, Atlanta, GA, USA. E-mail: dellaert@cc.gatech.edu



Figure 1. We propose to use Lagrangian duality to compute optimal solution and to verify optimality of a given estimate in pose graph optimization. For instance, our verification techniques allows us to certify optimal solutions (right column) and discern them from sup-optimal ones (left column), a task that has been traditionally entrusted to visual inspection of a human operator.

closely related to PGO need to be solved for structure from motion [23], [45], [2], [24], [55], [28], [21], attitude synchronization [58], [29], [46], camera network calibration [61], [59], sensor network localization [49], [48], and distributed consensus on manifolds [52], [60].

Related work in robotics. Since the seminal paper [44], PGO attracted large attention from the robotics community. Most state-of-the-art techniques currently rely on iterative

L. Carlone is with the College of Computing, Georgia Institute of Technology, Atlanta, GA, USA. E-mail: luca.carlone@gatech.edu

nonlinear optimization, which refines a given initial guess (usually, the odometric estimate from dead-reckoning). The Gauss-Newton method is a popular choice [40], [35], [34], as it converges quickly when the initialization is close to a minimum of the cost function. Trust region methods (e.g., the Levenberg-Marquart method, or the Powell's Dog-Leg method) have also been applied successfully to PGO [51]; the gradient method has been shown to have a large convergence basin, while suffering from long convergence tails [47], [26]. A large body of literature focuses on speeding up computation. This includes exploiting sparsity [35], [20], using reduction schemes to limit the number of poses [39], [7], faster linear solvers [22], [16], or approximate solutions [10].

PGO is a nonconvex problem and iterative optimization can only guarantee local convergence. State-of-the-art iterative solvers fail to converge to a global minimum of the cost for relatively small noise levels [11], [13]. This fact recently triggered efforts towards the design of more robust techniques, together with a theoretical analysis of PGO. Huang et. al [31] discuss the number of minima in small PGO problems. Knuth and Barooah [38] investigate the growth of the error in absence of loop closures. Carlone [8] provides conservative estimates of the basin of convergence for the Gauss-Newton method. Huang et. al [30] and Wang et. al [62] discuss the nonlinearities in PGO. In order to improve global convergence, a successful strategy consists in solving for the rotations first, and then using the resulting estimate to bootstrap iterative solvers [9], [10], [11], [13]. This is convenient since the rotation subproblem¹ has a guaranteed solution in 2D [11]. and many techniques for rotation estimation also perform well in 3D [45], [23], [21], [13]. Despite the empirical success of state-of-the-art techniques, no approach can guarantee global convergence. It is not even known if the global optimizer is unique in general instances (while it is known that it is unique with probability one in the rotation subproblem [11]). The lack of guarantees promoted a recent interest in verification techniques for PGO. Carlone and Dellaert [12] use duality to evaluate the quality of a candidate solution in planar PGO.

Related work in other fields. Variations of the PGO problem appear in different research fields. In computer vision, a somehow more difficult variant of the problem is known as *bundle adjustment* [23], [45], [2], [24], [55], [28], [21]. Contrarily to PGO, in bundle adjustment the relative measurements between the (camera) poses are only known up to scale. While no closed-form solution is known for bundle adjustment, many authors focused on the solution of the rotation subproblem [23], [45], [2], [24], [55], [21], [28]. The corresponding algorithms have excellent performance in practice (see [13] for an empirical evaluation), but they come with little guarantees, as they are based on relaxation. Fredriksson and Olsson [21] use duality theory to evaluate convergence of quaternion-based rotation estimation.

Related work in multi robot systems and sensor networks also includes contributions on rotation estimation (also known as *attitude synchronization* [58], [29], [46], [15], [63]). Borra *et. al* [4] propose a distributed algorithm for planar rotation estimation. Tron and Vidal [61], [59] provide convergence results for distributed attitude consensus using gradient descent. Piovan *et. al* [49] provide observability conditions and discuss iterative algorithms for PGO. Peters *et. al* [48] propose an estimation algorithm, based on the limit of a set of continuous-time differential equations.

Excellent contributions on rotation estimation have been also proposed in the context of *cryo-electron microscopy* [56], [57]. Singer and Shkolnisky [56], [57] provide two approaches for rotation estimation, based on relaxation and semidefinite programming (SDP). Bandeira *et. al* [3] provide a Cheegerlike inequality that establishes performance bounds for the SDP relaxation. Saunderson *et. al* [53] propose a tighter SDP relaxation, based on a spectrahedral representation of the convex hull of the rotation group.

Contribution. This paper shows that Lagrangian duality allows computing a globally optimal solution for PGO, and enables to verify optimality of a given estimate.

Section II recalls preliminary concepts, and discusses the properties of a particular set of 2×2 matrices, which are scalar multiples of a planar rotation matrix. These matrices are omnipresent in planar PGO and acknowledging this fact allows reformulating the problem over complex variables.

Section III frames PGO as a problem in complex variables. This makes analysis easier and allows drawing connections with the recent literature on *unit gain graphs* [50]. Exploiting this connection we prove nontrival results about the spectrum of the matrix underlying the problem (the *pose graph matrix*).

Section IV formulates and analyzes the Lagrangian dual problem in the complex domain. The dual PGO problem is a semidefinite program (SDP). We show that the *duality gap* is connected to the zero eigenvalues of the penalized pose graph matrix, which arises from the solution of the dual problem. We prove that if this matrix has a single eigenvalue in zero (a condition that we call the single zero eigenvalue property, SZEP), then (i) the duality gap is zero, (ii) the primal PGO problem has a unique solution (up to an arbitrary rototranslation), and (iii) the primal solution can be computed by scaling the eigenvector of the penalized pose graph matrix corresponding to the zero eigenvalue. To the best of our knowledge, this is the first work that discusses the uniqueness of the PGO solution for general graphs and provides a provably optimal solution. Section IV also presents an SDP relaxation of PGO, interpreting the relaxation as the dual of the dual problem. Our SDP relaxation is related to the one of [56], [21], but we deal with 2D poses, rather than rotations; moreover, we only use the SDP relaxation to complement our discussion on duality and to support some of the proofs.

Section V exploits our analysis of the dual problem to devise computational approaches for PGO. We propose an algorithm that computes a guaranteed optimal solution when the penalized pose graph matrix satisfies the SZEP. We also propose a variant that deals with the case in which the SZEP is not satisfied. This variant, while possibly suboptimal, is shown to perform well in practice. Moreover, we show that duality provides tools to *verify* if a given estimate (e.g., computed using iterative solvers) is globally optimal.

¹We use the term "rotation subproblem" to denote the problem of associating a rotation to each node in the graph, using relative rotation measurements. This corresponds to disregarding the translations in PGO.

Section VI presents a numerical evaluation on simulated and real datasets. In practical regimes of operation (rotation noise < 0.3 rad), our Monte Carlo runs always produced a penalized pose graph matrix satisfying the SZEP. Hence, in this regime, our approach enables the computation (and the verification) of the optimal solution. For larger noise (e.g., 1rad standard deviation for rotations), we observed cases in which the penalized pose graph matrix has multiple zero eigenvalues.

This paper extends our initial proposal [12] in many directions: the formulation in the complex domain, all results involving the SZEP, and the optimal solution are novel. While we advertised our results in the workshop paper [42], the verification techniques and the experimental results presented in this paper are new and unpublished. Moreover, we provide extra results, proofs, and a toy example in which the duality gap is nonzero in the technical report [6].

II. NOTATION AND PRELIMINARY CONCEPTS

Section II-A introduces our notation. Section II-B recalls standard concepts from graph theory, and can be safely skipped by the expert reader. Section II-C, instead, discusses the properties of the set of 2×2 matrices that are multiples of a planar rotation matrix. We denote this set with the symbol α SO(2). α SO(2) is of interest in this paper since the action of any matrix $Z \in \alpha$ SO(2) can be conveniently represented as a multiplication between complex numbers (Section III-C). Table I summarizes the main symbols used in this paper.

A. Notation

The cardinality of a set \mathcal{V} is written as $|\mathcal{V}|$. The sets of real and complex numbers are denoted with \mathbb{R} and \mathbb{C} , respectively. I_n denotes the $n \times n$ identity matrix, $\mathbf{1}_n$ denotes the (column) vector of all ones of dimension n, $\mathbf{0}_{n \times m}$ denotes the $n \times m$ matrix of all zeros (we also use $\mathbf{0}_n \doteq \mathbf{0}_{n \times 1}$). For a matrix M, M_{ij} denotes the element of M in row i and column j. For matrices with a block structure we use $[M]_{ij}$ to denote the $d \times d$ block of M at the block row i and block column j. In this paper we only deal with matrices that have 2×2 blocks, i.e., d = 2, hence the notation $[M]_{ij}$ is unambiguous.

B. Graph terminology

A directed graph \mathcal{G} is a pair $(\mathcal{V}, \mathcal{E})$, where the vertices or nodes \mathcal{V} are a finite set of elements, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. Each edge is an ordered pair e = (i, j). We say that eis incident on nodes i and j, leaves node i, called *tail*, and is directed towards node j, called *head*. The number of nodes and edges is denoted with $n \doteq |\mathcal{V}|$ and $m \doteq |\mathcal{E}|$, respectively.

The *incidence matrix* \mathcal{A} of a directed graph is a $m \times n$ matrix with elements in $\{-1, 0, +1\}$ that exhaustively describes the graph topology. Each row of \mathcal{A} corresponds to an edge e = (i, j) and has exactly two non-zero elements, a -1 on the *i*-th column and a +1 on the *j*-th column.

C. The set $\alpha SO(2)$

The set $\alpha SO(2)$ is defined as

$$\alpha \mathrm{SO}(2) \doteq \{ \alpha \mathbf{R} : \alpha \in \mathbb{R}, \, \mathbf{R} \in \mathrm{SO}(2) \},\$$

where SO(2) is the set of 2D rotation matrices. Recall that SO(2) can be parametrized by an angle $\theta \in (-\pi, +\pi]$, and any matrix $\mathbf{R} \in SO(2)$ is in the form:

$$\boldsymbol{R} = \boldsymbol{R}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$
 (1)

Clearly, $SO(2) \subset \alpha SO(2)$. The set $\alpha SO(2)$ is closed under standard matrix *multiplication*, i.e., for any $Z_1, Z_2 \in \alpha SO(2)$, also the product $Z_1Z_2 \in \alpha SO(2)$. In full analogy with SO(2), it is trivial to show that the multiplication is commutative, i.e., for any $Z_1, Z_2 \in \alpha SO(2)$ it holds that $Z_1Z_2 = Z_2Z_1$. Moreover, for $Z = \alpha R$ with $R \in SO(2)$ it holds that $Z^{\top}Z =$ $|\alpha|^2 I_2$. The set $\alpha SO(2)$ is also closed under matrix *addition*: for $R_1, R_2 \in SO(2)$, we have that

$$\alpha_1 \mathbf{R}_1 + \alpha_2 \mathbf{R}_2 = \alpha_1 \begin{bmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{bmatrix} + \alpha_2 \begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha_1 c_1 + \alpha_2 c_2 & -(\alpha_1 s_1 + \alpha_2 s_2) \\ \alpha_1 s_1 + \alpha_2 s_2 & \alpha_1 c_1 + \alpha_2 c_2 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \alpha_3 \mathbf{R}_3,$$

$$(2)$$

where we used the shorthands c_i and s_i for $\cos(\theta_i)$ and $\sin(\theta_i)$, and we defined $a \doteq \alpha_1 c_1 + \alpha_2 c_2$ and $b \doteq \alpha_1 s_1 + \alpha_2 s_2$. In (2), the scalar $\alpha_3 \doteq \pm \sqrt{a^2 + b^2}$ (if nonzero) normalizes $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, such that $\mathbf{R}_3 \doteq \begin{bmatrix} a/\alpha_3 & -b/\alpha_3 \\ b/\alpha_3 & a/\alpha_3 \end{bmatrix}$ is a rotation matrix; if $\alpha_3 = 0$, then $\alpha_1 \mathbf{R}_1 + \alpha_2 \mathbf{R}_2 = \mathbf{0}_{2\times 2}$, which also falls in our definition of α SO(2). From this reasoning, it is clear that an alternative definition of α SO(2) is

$$\alpha \mathrm{SO}(2) \doteq \left\{ \left[\begin{array}{cc} a & -b \\ b & a \end{array} \right] : a, b \in \mathbb{R} \right\}.$$
(3)

The set α SO(2) is tightly coupled with the set of complex numbers \mathbb{C} . Indeed, a matrix in the form (3) is also known as a *matrix representation* of a complex number [19]. We explore the implications of this fact for PGO in Section III-C.

III. POSE GRAPH OPTIMIZATION IN THE COMPLEX DOMAIN

Sections III-A-III-B recall a standard statement of the pose graph optimization problem. Section III-C frames the problem in the complex domain. Section III-D discusses properties of the matrices involved in the real and complex formulations.

A. Standard PGO

PGO estimates n poses from m relative pose measurements. We focus on the planar case, in which the *i*-th pose x_i is described by the pair $x_i \doteq (p_i, R_i)$, where $p_i \in \mathbb{R}^2$ is a position in the plane, and $R_i \in SO(2)$. The pose measurement between two nodes, say *i* and *j*, is described by the pair (Δ_{ij}, R_{ij}) , where $\Delta_{ij} \in \mathbb{R}^2$ and $R_{ij} \in SO(2)$ are the relative position and rotation measurements, respectively.

The problem can be visualized as a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where an unknown pose is attached to each node in the set \mathcal{V} , and each edge $(i, j) \in \mathcal{E}$ corresponds to a relative pose measurement between nodes i and j (Fig. 2).

In a noiseless case, the measurements satisfy:

$$\boldsymbol{\Delta}_{ij} = \boldsymbol{R}_i^{\top} \left(\boldsymbol{p}_j - \boldsymbol{p}_i \right), \qquad \boldsymbol{R}_{ij} = \boldsymbol{R}_i^{\top} \boldsymbol{R}_j , \qquad (4)$$

and we can compute the unknown rotations $\{R_1, \ldots, R_n\}$ and positions $\{p_1, \ldots, p_n\}$ by solving a set of linear equations

Table ISYMBOLS USED IN THIS PAPER

*	
$\vec{k} = (\mathcal{V}, \mathcal{E})$	Directed graph
n	Number of edges/measurements
ļ,	Number of nodes/poses
,	Vertex set; $ \mathcal{V} = n$
	Edge set; $ \mathcal{E} = m$
$=(i,j)\in\mathcal{E}$	Edge between nodes i and j
$4 \in \mathbb{R}^{m imes n}$	Incidence matrix of \mathcal{G}
$\mathbf{A} \in \mathbb{R}^{m \times (n-1)}$	Anchored incidence matrix of \mathcal{G}
$\mathcal{L} = \mathcal{A}^{ op} \mathcal{A}$	Laplacian matrix of \mathcal{G}
$\mathbf{L} = \mathbf{A}^{ op} \mathbf{A}$	Anchored Laplacian matrix of \mathcal{G}
$egin{aligned} & \mathcal{E} \ & = (i,j) \in \mathcal{E} \ & \mathbf{A} \in \mathbb{R}^{m imes n} \ & \mathbf{A} \in \mathbb{R}^{m imes (n-1)} \ & \mathbf{L} = \mathbf{A}^{ op} \mathbf{A} \ & \mathbf{A} \in \mathbb{R}^{T} \mathbf{A} \end{aligned}$	Vertex set; $ \mathcal{V} = n$ Edge set; $ \mathcal{E} = m$ Edge between nodes <i>i</i> and <i>j</i> Incidence matrix of \mathcal{G} Anchored incidence matrix of \mathcal{G} Laplacian matrix of \mathcal{G}

Real PGO formulation

Graph

$ar{oldsymbol{\mathcal{A}}} = oldsymbol{\mathcal{A}} \otimes oldsymbol{I}_2$	Augmented incidence matrix
$ar{m{A}}=m{A}\otimesm{I}_2$	Augmented anchored incidence matrix
$ar{\mathcal{L}} = \mathcal{L} \otimes I_2$	Augmented Laplacian matrix
$oldsymbol{\mathcal{W}} \in \mathbb{R}^{4n imes 4n}$	Real pose graph matrix
$\boldsymbol{W} \in \mathbb{R}^{(4n-2) \times (4n-2)}$	Real anchored pose graph matrix
$oldsymbol{p} \in \mathbb{R}^{2n}$	Node positions
$oldsymbol{ ho} \in \mathbb{R}^{2(n-1)}$	Anchored node positions
$oldsymbol{r} \in \mathbb{R}^{2n}$	Node rotations

Complex PGO formulation

$\tilde{\boldsymbol{W}} \in \mathbb{C}^{(2n-1) \times (2n-1)}$	Complex anchored pose graph matrix
$\tilde{oldsymbol{ ho}}\in\mathbb{C}^{n-1}$	Anchored complex node positions
$ ilde{m{r}}\in\mathbb{C}^n$	Complex node rotations

Miscellanea

SO(2)	2D rotation matrices
$\alpha SO(2)$	Scalar multiple of a 2D rotation matrix
$ \mathcal{V} $	Cardinality of the set \mathcal{V}
I_n	$n \times n$ identity matrix
$0_n (1_n)$	Vector of zeros (ones) of dimension n
$Tr(\mathbf{X})$	Trace of the matrix X



Figure 2. Schematic representation of Pose Graph Optimization: the objective is to associate a pose x_i to each node of a directed graph, given relative pose measurements (Δ_{ij}, R_{ij}) for each edge (i, j) in the graph.

(relations (4) become linear after rearranging the rotation R_i to the left-hand side). In absence of noise, the problem admits a unique solution as long as one fixes the pose of a node (say $p_1 = 0_2$ and $R_1 = I_2$) and the underling graph is connected.

In this work we focus on connected graphs, as these are the ones of practical interest in PGO (a graph with k connected components can be split in k subproblems, which can be solved and analyzed independently).

Assumption 1 (Connected Pose Graph): The graph \mathcal{G} , underlying pose graph optimization, is connected.

In presence of noise, the relations (4) cannot be met exactly and pose graph optimization looks for a set of positions $\{p_1, \ldots, p_n\}$ and rotations $\{R_1, \ldots, R_n\}$ that minimize the mismatch with respect to the measurements. This mismatch can be quantified by different cost functions. We adopt the formulation proposed in [12]:

$$\min_{\substack{\{\boldsymbol{p}_i\},\\\{\boldsymbol{R}_i\in\mathrm{SO}(2)\}}}\sum_{(i,j)\in\mathcal{E}}\omega_{ij}^{\Delta}\|\boldsymbol{\Delta}_{ij}-\boldsymbol{R}_i^{\top}(\boldsymbol{p}_j-\boldsymbol{p}_i)\|_2^2 + \frac{\omega_{ij}^R}{2}\|\boldsymbol{R}_{ij}-\boldsymbol{R}_i^{\top}\boldsymbol{R}_j\|_{\mathrm{F}}^2,$$
(5)

where $\|\cdot\|_2$ is the standard Euclidean distance and $\|\cdot\|_F$ is the Frobenius norm. The Frobenius norm $\|\mathbf{R}_a - \mathbf{R}_b\|_F$ is a standard measure of distance between two rotations \mathbf{R}_a and \mathbf{R}_b , and it is commonly called *chordal* distance, see, e.g., [28]. In (5), we used the notation $\{\mathbf{p}_i\}$ (resp. $\{\mathbf{R}_i\}$) to denote the set of unknown positions $\{\mathbf{p}_1 \dots, \mathbf{p}_n\}$ (resp. rotations). The weights ω_{ij}^{Δ} and ω_{ij}^R allow accommodating measurement uncertainty; to simplify notation, in the following we assume $\omega_{ij}^{\Delta} = \omega_{ij}^R =$ 1: including these terms in the derivation is straightforward and they are indeed taken into account in our experiments.

Rearranging the terms, problem (5) can be rewritten as:

$$\min_{\substack{\{\boldsymbol{p}_i\},\\\{\boldsymbol{R}_i\in\mathrm{SO}(2)\}}}\sum_{(i,j)\in\mathcal{E}}\|(\boldsymbol{p}_j-\boldsymbol{p}_i)-\boldsymbol{R}_i\boldsymbol{\Delta}_{ij}\|_2^2+\frac{1}{2}\|\boldsymbol{R}_j-\boldsymbol{R}_i\boldsymbol{R}_{ij}\|_{\mathrm{F}}^2,$$
(6)

where we exploited the fact that the 2-norm is invariant to rotation, i.e., for any vector v and any rotation matrix R it holds $||Rv||_2 = ||v||_2$. Eq. (6) highlights that the objective is a quadratic function of the unknowns.

The complexity of the problem stems from the fact that the constraint $\mathbf{R}_i \in SO(2)$ is nonconvex, see, e.g., [53]. To make this more explicit, we follow [12], and use a more convenient representation for nodes' rotations. A planar rotation \mathbf{R}_i can be written as in (1), and is fully defined by the vector

$$\boldsymbol{r}_{i} = \begin{bmatrix} \cos(\theta_{i}) \\ \sin(\theta_{i}) \end{bmatrix}.$$
 (7)

Using this parametrization and with simple matrix manipulation, Eq. (6) becomes (*cf.* with Eq. (11) in [12]):

$$\min_{\{\boldsymbol{p}_i\},\{\boldsymbol{r}_i\}} \sum_{(i,j)\in\mathcal{E}} \|(\boldsymbol{p}_j - \boldsymbol{p}_i) - \boldsymbol{D}_{ij}\boldsymbol{r}_i\|_2^2 + \|\boldsymbol{r}_j - \boldsymbol{R}_{ij}\boldsymbol{r}_i\|_2^2$$
(8)
s.t.: $\|\boldsymbol{r}_i\|_2^2 = 1, \quad i = 1, \dots, n$

where we defined:

$$\boldsymbol{D}_{ij} = \begin{bmatrix} \Delta_{ij}^x & -\Delta_{ij}^y \\ \Delta_{ij}^y & \Delta_{ij}^x \end{bmatrix}, \quad (\text{with } \boldsymbol{\Delta}_{ij} \doteq [\Delta_{ij}^x \Delta_{ij}^y]^\top) , \quad (9)$$

and where the constraints $||\mathbf{r}_i||_2^2 = 1$ specify that we look for vectors \mathbf{r}_i that represent admissible rotations (i.e., such that $\cos(\theta_i)^2 + \sin(\theta_i)^2 = 1$).

Problem (8) is a quadratic problem with quadratic equality constraints. The latter are nonconvex, hence computing a global minimum of (8) is hard in general. There are two problem instances, however, for which it is easy to compute a global minimizer, which attains zero cost. These two cases are recalled in Propositions 1-2.

Proposition 1 (Zero cost in trees): An optimal solution for a PGO problem in the form (8) whose underlying graph is a tree attains zero cost. \blacksquare

This is a well known fact in robotics. The interested reader can find a formal proof in [6, Appendix 8.1]. Roughly speaking, in a tree, we can build an optimal solution by concatenating the relative pose measurements, and this solution annihilates the cost function. However, also in graphs with chords, it is possible to attain the zero cost.

Definition 1 (Balanced pose graph): A pose graph is balanced if the pose measurements compose to the identity along f each cycle in the graph^{2,3}.

In a balanced pose graph, there exists a configuration that explains exactly the measurements, as stated below.

Proposition 2 (Zero cost in balanced pose graphs): An optimal solution for a balanced pose graph optimization problem attains zero cost. ■

The proof is given in [6, Appendix 8.2]. The concept of balanced graph describes a noiseless setup, while in real problem instances the measurements do not compose to the identity along cycles, because of the presence of noise.

The following fact will be useful in Section III-B.

Proposition 3 (Coefficient matrices in PGO): The matrices $D_{ij}, I_2, -I_2, R_{ij}$ appearing in (8) belong to $\alpha SO(2)$. The proof is trivial, since $R_{ij}, I_2 \in SO(2) \subset \alpha SO(2)$ (the latter also implies $-I_2 \in \alpha SO(2)$). Moreover, D_{ij} in (9) clearly falls in the definition of matrices in $\alpha SO(2)$ in (3).

B. Matrix formulation and anchoring

In this section we rewrite the cost function (8) in a more convenient matrix form. The original cost is:

$$f(\boldsymbol{p}, \boldsymbol{r}) \doteq \sum_{(i,j) \in \mathcal{E}} \|(\boldsymbol{p}_j - \boldsymbol{p}_i) - \boldsymbol{D}_{ij} \boldsymbol{r}_i\|_2^2 + \|\boldsymbol{r}_j - \boldsymbol{R}_{ij} \boldsymbol{r}_i\|_2^2$$
(10)

where we denote with $\boldsymbol{p} \in \mathbb{R}^{2n}$ and $\boldsymbol{r} \in \mathbb{R}^{2n}$ the vectors stacking all nodes positions and rotations, respectively. Now, let $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{m \times n}$ denote the *incidence matrix* of the graph underlying the problem: if (i, j) is the k-th edge, then $\boldsymbol{\mathcal{A}}_{ki} = -1$, $\boldsymbol{\mathcal{A}}_{kj} = +1$. Let $\bar{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{A}} \otimes \boldsymbol{I}_2 \in \mathbb{R}^{2m \times 2n}$, and denote with $\bar{\boldsymbol{\mathcal{A}}}_k \in \mathbb{R}^{2 \times 2n}$ the k-th block row of $\bar{\boldsymbol{\mathcal{A}}}$. From the structure of $\bar{\boldsymbol{\mathcal{A}}}$, it follows that $\bar{\boldsymbol{\mathcal{A}}}_k \boldsymbol{p} = \boldsymbol{p}_j - \boldsymbol{p}_i$. Also, we define $\bar{\boldsymbol{D}}_k \in \mathbb{R}^{2m \times 2n}$ corresponding to the k-th edge (i, j) is all zeros, except for a 2×2 block $-\boldsymbol{D}_{ij}$ in the *i*-th block column. Using the matrices $\bar{\boldsymbol{\mathcal{A}}}$ and $\bar{\boldsymbol{D}}$, the first sum in (10) can be written as:

$$\sum_{(i,j)\in\mathcal{E}} \|(\boldsymbol{p}_j - \boldsymbol{p}_i) - \boldsymbol{D}_{ij}\boldsymbol{r}_i\|_2^2 = \sum_{k=1}^m \|\bar{\boldsymbol{\mathcal{A}}}_k \boldsymbol{p} + \bar{\boldsymbol{D}}_k \boldsymbol{r}\|_2^2 = \|\bar{\boldsymbol{\mathcal{A}}} \boldsymbol{p} + \bar{\boldsymbol{D}}\boldsymbol{r}\|_2^2$$
(11)

Similarly, we define $\bar{U} \in \mathbb{R}^{2m \times 2n}$ as a block matrix where the *k*-th block row $\bar{U}_k \in \mathbb{R}^{2 \times 2n}$ corresponding to the *k*-th edge (i, j) is all zeros, except for 2×2 blocks in the *i*-th and *j*-th block columns, which are equal to $-R_{ij}$ and I_2 , respectively. Using \bar{U} , the second sum in (10) becomes:

$$\sum_{(i,j)\in\mathcal{E}} \|\boldsymbol{r}_j - \boldsymbol{R}_{ij}\boldsymbol{r}_i\|_2^2 = \sum_{k=1}^m \|\bar{\boldsymbol{U}}_k\boldsymbol{r}\|_2^2 = \|\bar{\boldsymbol{U}}\boldsymbol{r}\|_2^2 \qquad (12)$$

²We use the term "composition" to denote the group operation for SE(2). For two poses $T_1 \doteq (p_1, R_1)$ and $T_2 \doteq (p_2, R_2)$, the composition is $T_1 \cdot T_2 = (p_1 + R_1 p_2, R_1 R_2)$ [14]. Similarly, the *identity* element is $(0_2, I_2)$.

³When composing measurements along a loop, edge direction is important: for two consecutive edges (i, k) and (k, j) in the loop, the composition is $T_{ij} = T_{ik} \cdot T_{kj}$, while if the second edge is in the form (j, k), the composition becomes $T_{ij} = T_{ik} \cdot T_{jk}^{-1}$. Combining (11) and (12), and defining $\bar{Q} \doteq \bar{D}^{\top}\bar{D} + \bar{U}^{\top}\bar{U}$ and $\bar{\mathcal{L}} \doteq \bar{\mathcal{A}}^{\top}\bar{\mathcal{A}}$, the cost in (10) becomes:

$$(\boldsymbol{p},\boldsymbol{r}) = \left\| \begin{bmatrix} \bar{\boldsymbol{\mathcal{A}}} & \bar{\boldsymbol{D}} \\ \boldsymbol{0} & \bar{\boldsymbol{U}} \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{r} \end{bmatrix} \right\|_{2}^{2} = \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{r} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} & \bar{\boldsymbol{\mathcal{A}}}^{\mathsf{T}} \bar{\boldsymbol{D}} \\ \bar{\boldsymbol{D}}^{\mathsf{T}} \bar{\boldsymbol{\mathcal{A}}} & \bar{\boldsymbol{Q}} \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{r} \end{bmatrix}$$
(13)

Since $\bar{\mathcal{A}} \doteq \mathcal{A} \otimes I_2$, it follows that $\bar{\mathcal{L}} = \mathcal{L} \otimes I_2$, where $\mathcal{L} \doteq \mathcal{A}^{\top} \mathcal{A}$ is the Laplacian matrix of the graph underlying the problem. A pose graph optimization instance is thus completely defined by the matrix

$$\boldsymbol{\mathcal{W}} \doteq \begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} & \bar{\boldsymbol{\mathcal{A}}}^{\top} \bar{\boldsymbol{D}} \\ \bar{\boldsymbol{D}}^{\top} \bar{\boldsymbol{\mathcal{A}}} & \bar{\boldsymbol{Q}} \end{bmatrix} \in \mathbb{R}^{4n \times 4n}$$
(14)

From (13), \mathcal{W} can be easily seen to be symmetric and positive semidefinite. Other useful properties of \mathcal{W} are stated below.

Proposition 4 (Properties of \mathcal{W}): The positive semidefinite matrix \mathcal{W} in (14) (i) has at least two eigenvalues in zero, and (ii) is composed by 2×2 blocks $[\mathcal{W}]_{ij}$, where each block is a multiple of a rotation matrix, i.e., $[\mathcal{W}]_{ij} \in \alpha \text{SO}(2), \forall i, j = 1, ..., 2n$. Moreover, the diagonal blocks of \mathcal{W} are nonnegative multiples of the identity matrix, i.e., $[\mathcal{W}]_{ii} = \alpha_{ii} I_2, \alpha_{ii} \geq 0.$

A formal proof of Proposition 4 is given in [6, Appendix 8.3]. An intuitive explanation of the second claim follows from the fact that (i) \mathcal{W} contains sums and products of the matrices in the original formulation (8) (which are in α SO(2) according to Lemma 3), and (ii) the set α SO(2) is closed under matrix sum and product (Section II-C).

The presence of two eigenvalues in zero has a natural geometric interpretation: the cost function encodes inter-nodal measurements, hence it is invariant to global translations of node positions, i.e., $f(\boldsymbol{p}, \boldsymbol{r}) = f(\boldsymbol{p} + \boldsymbol{p}_a, \boldsymbol{r})$, where $\boldsymbol{p}_a \doteq (\mathbf{1}_n \otimes \boldsymbol{I}_2)\boldsymbol{a} = [\boldsymbol{a}^\top \dots \boldsymbol{a}^\top]^\top$ (*n* copies of \boldsymbol{a}), with $\boldsymbol{a} \in \mathbb{R}^2$. Algebraically, this translates to the fact that $(\mathbf{1}_n \otimes \boldsymbol{I}_2) \in \mathbb{R}^{2n \times 2}$ is in the null space of the augmented incidence matrix $\bar{\boldsymbol{A}}$, which also implies a two dimensional null space for $\boldsymbol{\mathcal{W}}$.

Position anchoring. In this paper we show that the duality properties in pose graph optimization are tightly coupled with the spectrum of the matrix \mathcal{W} . We are particularly interested in the eigenvalues at zero, and from this perspective it is not convenient to carry on the two null eigenvalues of \mathcal{W} (claim (i) of Proposition 4), which are always present, and are due to an intrinsic observability issue.

We remove the translation ambiguity by fixing the position of an arbitrary node. Without loss of generality, we fix the position p_1 of the first node to the origin, i.e., $p_1 = 0_2$. This process is commonly called *anchoring*. Setting $p_1 = 0_2$ is equivalent to removing the corresponding columns and rows from \mathcal{W} , leading to the following "anchored" PGO problem:

$$f(\boldsymbol{r},\boldsymbol{\rho}) = \begin{bmatrix} \boldsymbol{0}_2 \\ \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix}^{\top} \boldsymbol{\mathcal{W}} \begin{bmatrix} \boldsymbol{0}_2 \\ \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix}^{\top} \boldsymbol{W} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix}$$
(15)

where ρ is the vector p without its first two-elements vector p_1 , and W is obtained from \mathcal{W} by removing the rows and the columns corresponding to p_1 . The structure of W is:

$$W = \begin{bmatrix} \bar{A}^{\top}\bar{A} & \bar{A}^{\top}\bar{D} \\ \bar{D}^{\top}\bar{A} & \bar{Q} \end{bmatrix} \doteq \begin{bmatrix} \bar{L} & \bar{S} \\ \bar{S}^{\top} & \bar{Q} \end{bmatrix}$$
(16)

where $\bar{A} = A \otimes I_2$, and A is the *anchored* (or *reduced*) incidence matrix, obtained by removing the first column from A, see, e.g., [10]. On the right-hand-side of (16) we defined $\bar{S} \doteq \bar{A}^{\top} \bar{D}$ and $\bar{L} \doteq \bar{A}^{\top} \bar{A}$.

We call W the *real (anchored) pose graph matrix.* W is still symmetric and positive semidefinite (it is a principal submatrix of a positive semidefinite matrix). Moreover, since W is obtained by removing a $2 \times 4n$ block row and a $4n \times 2$ block column from W, it is still composed by 2×2 matrices in α SO(2), as specified in the following remark.

Remark 1 (Properties of W): The positive semidefinite matrix W in (16) is composed by 2×2 blocks $[W]_{ij}$, that are such that $[W]_{ij} \in \alpha SO(2), \forall i, j = 1, ..., 2n - 1$. Moreover, the diagonal blocks of W are nonnegative multiples of the identity matrix, i.e., $[W]_{ii} = \alpha_{ii}I_2, \alpha \ge 0$. After anchoring, our PGO problem becomes:

$$f^{\star} = \min_{\boldsymbol{\rho}, \boldsymbol{r}} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix} \boldsymbol{W} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix}$$
(17)
s.t.:
$$\|\boldsymbol{r}_i\|_2^2 = 1, \qquad i = 1, \dots, n$$

C. To complex domain

In this section we reformulate (17), in which the decision variables are real vectors, into a problem in complex variables. The main motivation for this choice is that the real representation (17) is somehow redundant: as we show in Proposition 7, each eigenvalue of W is repeated twice (multiplicity 2), while the complex representation does not have this redundancy, making analysis easier. In the rest of the paper, quantities marked with a tilde $(\overline{\)}$ live in the complex domain \mathbb{C} .

Any real vector $v \in \mathbb{R}^2$ can be represented by a complex number $\tilde{v} = \eta e^{j\varphi}$, where $j^2 = -1$ is the *imaginary unit*, $\eta = ||v||_2$ and φ is the angle that v forms with the horizontal axis. We use the operator $(\cdot)^{\vee}$ to map a 2-vector to the corresponding complex number, $\tilde{v} = v^{\vee}$.

The action of a real 2×2 matrix Z on a vector $v \in \mathbb{R}^2$ cannot be represented, in general, as a scalar multiplication between complex numbers. However, if $Z \in \alpha SO(2)$, this is possible. To show this, assume that $Z = \alpha R(\theta)$, where $R(\theta)$ is a counter-clockwise rotation of angle θ . Then,

$$(\boldsymbol{Z}\boldsymbol{v})^{\vee} = (\alpha \boldsymbol{R}(\theta)\boldsymbol{v})^{\vee} = \tilde{\boldsymbol{z}} \; \tilde{\boldsymbol{v}}, \text{ where } \tilde{\boldsymbol{z}} = \alpha \mathrm{e}^{\mathrm{j}\theta}.$$
 (18)

With slight abuse of notation we extend the operator $(\cdot)^{\vee}$ to $\alpha SO(2)$, such that, given $\mathbf{Z} = \alpha \mathbf{R}(\theta) \in \alpha SO(2)$, then $\mathbf{Z}^{\vee} = \alpha e^{j\theta} \in \mathbb{C}$. By inspection, one can verify the following relations between the sum and product of two matrices $\mathbf{Z}_1, \mathbf{Z}_2 \in \alpha SO(2)$ and their complex representations $\mathbf{Z}_1^{\vee}, \mathbf{Z}_2^{\vee} \in \mathbb{C}$:

$$(Z_1 \ Z_2)^{\vee} = \tilde{Z}_1^{\vee} \ Z_2^{\vee}, \quad (Z_1 + Z_2)^{\vee} = Z_1^{\vee} + Z_2^{\vee}.$$
 (19)

We next discuss how to apply the machinery introduced so far to reformulate problem (17) in the complex domain. The variables in problem (17) are the vectors $\rho \in \mathbb{R}^{2(n-1)}$ and $r \in \mathbb{R}^{2n}$ that are composed by 2-vectors, i.e., $\rho = [\rho_1^\top, \ldots, \rho_{n-1}^\top]^\top$ and $r = [r_1^\top, \ldots, r_n^\top]^\top$, where $\rho_i, r_i \in \mathbb{R}^2$. We define the *complex positions* and the *complex rotations*:

$$\tilde{\boldsymbol{\rho}} = [\tilde{\rho}_1, \dots, \tilde{\rho}_{n-1}] \in \mathbb{C}^{n-1}, \text{ where: } \tilde{\rho}_i = \boldsymbol{\rho}_i^{\vee} \\ \tilde{\boldsymbol{r}} = [\tilde{r}_1, \dots, \tilde{r}_n] \in \mathbb{C}^n, \text{ where: } \tilde{r}_i = \boldsymbol{r}_i^{\vee}$$

$$(20)$$

Using the parametrization (20), the constraints in (17) become:

$$\tilde{r}_i|^2 = 1, \quad i = 1, \dots, n.$$
 (21)

Similarly, we would like to rewrite the objective as a function of $\tilde{\rho}$ and \tilde{r} . This re-parametrization is formalized in the following proposition, whose proof is given in Appendix VIII-A.

Proposition 5 (Cost in the complex domain): For any pair (ρ, r) , the cost function in (17) is such that:

$$f(\boldsymbol{\rho}, \boldsymbol{r}) = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix}^{\top} \boldsymbol{W} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{r} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{\rho}} \\ \tilde{\boldsymbol{r}} \end{bmatrix}^{\top} \tilde{\boldsymbol{W}} \begin{bmatrix} \tilde{\boldsymbol{\rho}} \\ \tilde{\boldsymbol{r}} \end{bmatrix}$$
(22)

where the vectors $\tilde{\rho}$ and \tilde{r} are built from ρ and r as in (20), and the matrix $\tilde{W} \in \mathbb{C}^{(2n-1)\times(2n-1)}$ is such that $\tilde{W}_{ij} = [W]_{ij}^{\vee}$, with i, j = 1, ..., 2n - 1.

Remark 2 (Real diagonal entries for \tilde{W}): According to Remark 1, the diagonal blocks of W are multiples of the identity matrix, i.e., $[W]_{ii} = \alpha_{ii}I_2$. Therefore, the diagonal elements of \tilde{W} are real, i.e., $\tilde{W}_{ii} = [W]_{ii}^{\vee} = \alpha_{ii} \in \mathbb{R}$. Proposition 5 enables us to rewrite problem (17) as:

$$f^{\star} = \min_{\tilde{\rho}, \tilde{r}} \begin{bmatrix} \tilde{\rho} \\ \tilde{r} \end{bmatrix}^{\top} \tilde{W} \begin{bmatrix} \tilde{\rho} \\ \tilde{r} \end{bmatrix}$$
s.t.: $|\tilde{r}_i|^2 = 1, \quad i = 1, \dots, n.$
(23)

We call \tilde{W} the *complex (anchored) pose graph matrix*. Clearly, the matrix \tilde{W} preserves the same block structure of W in (16):

$$\tilde{\boldsymbol{W}} \doteq \begin{bmatrix} \boldsymbol{L} & \tilde{\boldsymbol{S}} \\ \tilde{\boldsymbol{S}}^* & \tilde{\boldsymbol{Q}} \end{bmatrix}$$
(24)

where \tilde{S}^* is the conjugate transpose of \tilde{S} , and $L \doteq A^{\top}A$ where A is the anchored incidence matrix. In Section IV we apply Lagrangian duality to (23). Before that, we provide results to characterize the spectrum of W and \tilde{W} , drawing connections with the literature on unit gain graphs [50].

D. Spectrum of the real and complex pose graph matrices

In this section we take a closer look at the structure and the properties of the real and the complex pose graph matrices W and \tilde{W} . In analogy with (13) and (16), we write \tilde{W} as

$$\tilde{W} = \begin{bmatrix} A^{\top}A & A^{\top}\tilde{D} \\ (A^{\top}\tilde{D})^{*} & \tilde{U}^{*}\tilde{U} + \tilde{D}^{*}\tilde{D} \end{bmatrix} = \begin{bmatrix} A & \tilde{D} \\ 0 & \tilde{U} \end{bmatrix}^{*} \begin{bmatrix} A & \tilde{D} \\ 0 & \tilde{U} \\ (25) \end{bmatrix}$$

where $\tilde{U} \in \mathbb{C}^{m \times n}$ and $\tilde{D} \in \mathbb{C}^{m \times n}$ are the "complex versions" of \bar{U} and \bar{D} in (13), i.e., they are obtained as $\tilde{U}_{ij} = [\bar{U}]_{ij}^{\vee}$ and $\tilde{D}_{ij} = [\bar{D}]_{ij}^{\vee}, \forall i, j$.

The factorization (25) is interesting, as it allows identifying two important matrices that compose \tilde{W} : the first is A, the anchored incidence matrix that we introduced earlier; the second is \tilde{U} which is a generalization of the incidence matrix, known as the *complex incidence matrix* of a *unit gain graph* (see, e.g., [50]). A unit gain graph is a graph in which to each edge is associated a complex weight, having unit norm. For space reasons, we refer the reader to [6] and [50] for an introduction on unit gain graphs, while we provide an intuitive example of *complex incidence matrix* in Fig. 3.

Using existing results on the spectrum of the complex incidence matrix [50], we can characterize the presence of



Figure 3. Example of incidence matrix, anchored incidence matrix, and complex incidence matrix, for the toy PGO problem on the top left. If $\mathbf{R}_{ij} = \mathbf{R}(\theta_{ij})$ is the relative rotation measurement associated to edge (i, j), then the matrix $\tilde{\mathbf{U}}$ can be seen as the incidence matrix of a unit gain graph with gain $\mathrm{e}^{\mathrm{J}\theta_{ij}}$ associated to edge (i, j).

eigenvalues in zero for the matrix W, as specified in the following proposition (proof in [6, Appendix 8.5]).

Proposition 6 (Zero eigenvalues in \tilde{W}): The complex anchored pose graph matrix \tilde{W} has a single eigenvalue in zero if and only if the pose graph is balanced or is a tree.

Besides analyzing the spectrum of \tilde{W} , it is of interest to understand how the complex matrix \tilde{W} relates to its real counterpart W. The following proposition states that there is a tight correspondence between the eigenvalues of W and \tilde{W} (proof in Appendix VIII-B).

Proposition 7 (Spectrum of complex graph matrices): The 2(2n-1) eigenvalues of W are the 2n-1 eigenvalues of \tilde{W} , repeated twice.

The previous result is the main reason why we prefer to work in the complex domain.

IV. LAGRANGIAN DUALITY IN PGO

In the previous section we wrote the PGO problem in complex variables as per eq. (23). In the following, we refer to this problem as the *primal* PGO problem, that, defining $\tilde{x} \doteq [\tilde{\rho} \tilde{r}]$ (as a column vector), can be written succinctly as

$$f^{\star} = \min_{\tilde{\boldsymbol{x}}} \quad \tilde{\boldsymbol{x}}^{\star} \boldsymbol{W} \tilde{\boldsymbol{x}} \quad (\text{Primal problem}) \\ \text{s.t.:} \quad |\tilde{x}_i|^2 = 1, \quad i = n, \dots, 2n - 1.$$
(26)

In Section IV-A we derive the Lagrangian dual of (26). Then, in Section IV-B, we discuss an SDP relaxation of (26), that can be interpreted as the dual of the dual problem. Finally, in Section IV-C we analyze the properties of the dual problem, and discuss how it relates with the primal PGO problem.

A. The dual problem

The Lagrangian of the primal problem (26) is

$$\mathbb{L}(\tilde{\boldsymbol{x}},\boldsymbol{\lambda}) = \tilde{\boldsymbol{x}}^* \tilde{\boldsymbol{W}} \tilde{\boldsymbol{x}} + \sum_{i=1}^n \lambda_i (1 - |\tilde{\boldsymbol{x}}_{n+i-1}|^2)$$

where $\lambda_i \in \mathbb{R}$, i = 1, ..., n, are the Lagrange multipliers (or *dual variables*). Recalling the structure of \tilde{W} from (24), the Lagrangian becomes:

$$\mathbb{L}(ilde{x}, oldsymbol{\lambda}) \!=\! ilde{x}^* \! \left[egin{array}{c} L & ilde{S} \ ilde{S}^* & ilde{Q}(oldsymbol{\lambda}) \end{array}
ight] \! ilde{x} \!+\! \sum_{i=1}^n \lambda_i \!=\! ilde{x}^* ilde{W}(oldsymbol{\lambda}) ilde{x} \!+\! \sum_{i=1}^n \lambda_i,$$

where for notational convenience we defined

$$ilde{m{Q}}(m{\lambda}) \doteq ilde{m{Q}} - ext{diag}(\lambda_1, \dots, \lambda_n), \quad ilde{m{W}}(m{\lambda}) \doteq \left[egin{array}{cc} m{L} & m{S} \\ m{ ilde{m{S}}^*} & m{ ilde{m{Q}}}(m{\lambda}) \end{bmatrix} . \ (27)$$

The *dual function* $d : \mathbb{R}^n \to \mathbb{R}$ is the infimum of the Lagrangian with respect to \tilde{x} :

$$d(\boldsymbol{\lambda}) = \inf_{\tilde{\boldsymbol{x}}} \mathbb{L}(\tilde{\boldsymbol{x}}, \boldsymbol{\lambda}) = \inf_{\tilde{\boldsymbol{x}}} \tilde{\boldsymbol{x}}^* \tilde{\boldsymbol{W}}(\boldsymbol{\lambda}) \tilde{\boldsymbol{x}} + \sum_{i=1}^n \lambda_i.$$
(28)

For any choice of λ the dual function provides a lower bound on the optimal value of the primal problem [5, Section 5.1.3]. Therefore, the *Lagrangian dual problem* looks for a *maximum* of the dual function over λ :

$$d^{\star} \doteq \max_{\lambda} d(\lambda) = \max_{\lambda} \inf_{\tilde{x}} \tilde{x}^{*} \tilde{W}(\lambda) \tilde{x} + \sum_{i=1}^{n} \lambda_{i}.$$
 (29)

The infimum over \tilde{x} of $\tilde{x}^* \tilde{W}(\lambda) \tilde{x}$ drifts to $-\infty$ unless $\tilde{W}(\lambda) \succeq 0$. Therefore we can safely restrict the maximization to vectors λ that are such that $\tilde{W}(\lambda) \succeq 0$; these are called *dual feasible*. Moreover, at any dual-feasible λ , the \tilde{x} minimizing the Lagrangian are those that make $\tilde{x}^* \tilde{W}(\lambda) \tilde{x} = 0$. Therefore, (29) reduces to the following *dual problem*

$$d^{\star} = \max_{\substack{\lambda \\ \text{s.t.:}}} \sum_{i} \lambda_{i}, \qquad \text{(Dual problem)} \\ \text{s.t.:} \quad \tilde{W}(\lambda) \succeq 0. \qquad (30)$$

The importance of the dual problem is twofold. First, it holds

$$d^{\star} \le f^{\star} \tag{31}$$

This property is called *weak duality*, see, e.g., [5, Section 5.2.2]. For particular problems the inequality (31) becomes an equality, and in such cases we say that *strong duality* holds. Second, since $d(\lambda)$ is concave (minimum of affine functions), the dual problem (30) is always convex in λ , regardless the convexity properties of the primal problem. The dual PGO problem (30) is a semidefinite program (SDP).

B. SDP relaxation and the dual of the dual

We have seen that a lower bound d^* on the optimal value f^* of the primal (26) can be obtained by solving the Lagrangian dual problem (30). Here, we outline another, direct, relaxation method to obtain such bound.

Observing that $\tilde{x}^* \tilde{W} \tilde{x} = \text{Tr} (\tilde{W} \tilde{x} \tilde{x}^*)$, we rewrite (26) equivalently as

$$f^{\star} = \min_{\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{x}}} \quad \text{Tr} (\tilde{\boldsymbol{W}} \tilde{\boldsymbol{X}})$$
(32)
s.t.:
$$\text{Tr} (\boldsymbol{E}_{i} \tilde{\boldsymbol{X}}) = 1, \quad i = n, \dots, 2n - 1,$$
$$\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{x}} \tilde{\boldsymbol{x}}^{*}.$$

where E_i is a matrix that is zero everywhere, except for the *i*-th diagonal element, which is one. The condition $\tilde{X} = \tilde{x}\tilde{x}^*$ is equivalent to (i) $\tilde{X} \succeq 0$ and (ii) \tilde{X} has rank one. Thus, (32) is rewritten by eliminating \tilde{x} as

$$f^{\star} = \min_{\tilde{X}} \operatorname{Tr} (\tilde{W}\tilde{X})$$
(33)
s.t.: $\operatorname{Tr} (E_{i}\tilde{X}) = 1, \quad i = n, \dots, 2n-1,$
 $\tilde{X} \succeq 0$
 $\operatorname{rank}(\tilde{X}) = 1.$

Dropping the rank constraint, which is non-convex, we obtain the following SDP relaxation of the primal problem:

$$s^{\star} = \min_{\tilde{\boldsymbol{X}}} \quad \operatorname{Tr} (\tilde{\boldsymbol{W}}\tilde{\boldsymbol{X}})$$

s.t.:
$$\operatorname{Tr} (\boldsymbol{E}_{i}\tilde{\boldsymbol{X}}) = 1, \quad i = n, \dots, 2n-1,$$

$$\tilde{\boldsymbol{X}} \succeq 0$$
(34)

which we can also rewrite as

$$s^{\star} = \min_{\tilde{\boldsymbol{X}}} \operatorname{Tr} (\tilde{\boldsymbol{W}}\tilde{\boldsymbol{X}}) \qquad \text{(SDP relaxation)}$$

s.t.: $\tilde{\boldsymbol{X}}_{ii} = 1, \quad i = n, \dots, 2n-1, \qquad \overset{(35)}{\tilde{\boldsymbol{X}} \succeq 0}$

where \tilde{X}_{ii} denotes the *i*-th diagonal entry in \tilde{X} . Obviously, $s^* \leq f^*$, since the feasible set of (35) contains that of (33). One may then ask what is the relation between the Lagrangian dual and the SDP relaxation of problem (35): the answer is that the former is the dual of the latter hence, under constraint qualification, it holds that $s^* = d^*$, i.e., the SDP relaxation and the Lagrangian dual approach yield the *same* lower bound on f^* . This is formalized in the following proposition.

Proposition 8: The Lagrangian dual of problem (35) is problem (30), and vice-versa. Strong duality holds between these two problems, i.e., $d^* = s^*$. Moreover, if the optimal solution \tilde{X}^* of (35) has rank one, then $d^* = s^* = f^*$.

Proof. The fact that the SDPs (35) and (30) are related by duality can be found in standard textbooks (e.g. [5, Example 5.13]); since these are convex programs, under constraint qualification, the duality gap is zero, i.e., $d^* = s^*$. To prove that rank $(\tilde{X}^*) = 1 \Rightarrow s^* = d^* = f^*$, we observe that (i) Tr $\tilde{W}\tilde{X}^* \doteq s^* \leq f^*$ since (35) is a relaxation of (33). However, when rank $(\tilde{X}^*) = 1$, \tilde{X}^* is feasible for problem (35), hence, by optimality of f^* , it holds (ii) $f^* \leq f(\tilde{X}^*) = \text{Tr } \tilde{W}\tilde{X}^*$. Combining (i) and (ii) we prove that, when rank $(\tilde{X}^*) = 1$, then $f^* = s^*$, which also implies $f^* = d^*$.

To the best of our knowledge this is the first time in which this SDP relaxation has been proposed to solve PGO; in the context of SLAM, anther SDP relaxation has been proposed by Liu *et. al* [43], but it does not use the chordal distance and approximates the expression of the relative rotation measurements. For the rotation subproblem, SDP relaxations have been proposed in [57], [53], [21]. According to Proposition 8, one advantage of the SDP relaxation approach is that we can a-posteriori check if the duality (or, in this case, the relaxation) gap is zero, from the optimal solution \tilde{X}^* . Indeed, if we solve (35) and find that the optimal \tilde{X}^* has rank one, then we actually solved (26), hence the relaxation gap is zero. Moreover, in this case, from spectral decomposition of \tilde{X}^* we can get a vector \tilde{x}^* such that $\tilde{X}^* = (\tilde{x}^*)(\tilde{x}^*)^*$, and this vector is an optimal solution to the primal problem.

In the following section we derive similar a-posteriori condition for the dual problem (30). This condition enables the computation of a primal optimal solution. Moreover, it allows discussing the uniqueness of such solution.

C. Analysis of the dual problem

In this section we provide conditions under which the duality gap is zero. These conditions depend on the spectrum of $\tilde{W}(\lambda^*)$, which arises from the solution of (30). We refer to $\tilde{W}(\lambda^*)$ as the *penalized pose graph matrix*. Moreover, we discuss the relation between the dual and the primal problem.

A first proposition establishes that (30) attains an optimal solution (Proof in Appendix VIII-C).

Proposition 9: The optimal value d^* in (30) is attained at a finite λ^* . Moreover, the penalized pose graph matrix $\tilde{W}(\lambda^*)$ has an eigenvalue in 0.

The previous proposition guarantees that $\hat{W}(\lambda^*)$ always has an eigenvalue in zero. The following result states that when this eigenvalue is unique, then the duality gap must be zero (proof in Appendix VIII-D).

Proposition 10 (No duality gap): If the zero eigenvalue of the penalized pose graph matrix $\tilde{W}(\lambda^*)$ is simple then the duality gap is zero, i.e., $d^* = f^*$.

In the following we say that $\tilde{W}(\lambda^*)$ satisfies the *single zero* eigenvalue property (SZEP) if its zero eigenvalue is simple.

When the SZEP holds we are able to establish a precise relation between the solution of the dual and the solution of the primal problem (Corollary 1 below). Towards this goal we need some more notation. For a given λ , we denote by $\mathcal{N}(\lambda)$ the set of \tilde{x} that attain the optimal value in problem (28):

$$\mathcal{N}(m{\lambda}) \doteq \{ ilde{x}_{\lambda} \in \mathbb{C}^{2n-1} : \mathbb{L}(ilde{x}_{\lambda},m{\lambda}) = \min_{ ilde{x}} \mathbb{L}(ilde{x},m{\lambda}) = \min_{ ilde{x}} ilde{x}^* ilde{W}(m{\lambda}) ilde{x} \}$$

Since we already observed that for any dual-feasible λ the points \tilde{x} that minimize the Lagrangian are such that $\tilde{x}^* \tilde{W}(\lambda) \tilde{x} = 0$, it follows that for any dual-feasible λ :

$$\mathcal{N}(\boldsymbol{\lambda}) = \{ \tilde{\boldsymbol{x}} \in \mathbb{C}^{2n-1} : \tilde{\boldsymbol{W}}(\boldsymbol{\lambda}) \tilde{\boldsymbol{x}} = 0 \} = \mathrm{Kernel}(\tilde{\mathrm{W}}(\boldsymbol{\lambda})).$$
 (36)

The following result ensures that if a vector in $\mathcal{N}(\lambda)$ is feasible for the primal problem, then it is also an optimal solution for the PGO problem (proof in Appendix VIII-E).

Theorem 1: Given $\lambda \in \mathbb{R}^n$, if an $\tilde{x}_{\lambda} \in \mathcal{N}(\lambda)$ is primal feasible, then \tilde{x}_{λ} is primal optimal; moreover, λ is dual optimal, and the duality gap is zero.

We are now ready to characterize the relation between the solution of the primal and the dual problem when $\tilde{W}(\lambda^*)$ satisfies the SZEP. This is one of the key results of this paper.

Corollary 1 (SZEP $\Rightarrow \tilde{x}^* \in \mathcal{N}(\lambda^*)$): If the zero eigenvalue of $\tilde{W}(\lambda^*)$ is simple (SZEP), then the set $\mathcal{N}(\lambda^*)$ contains a primal optimal solution. Moreover, the primal optimal solution is unique, up to an arbitrary rotation.

Proof. Let \tilde{x}^* be a primal optimal solution, and let $f^* = (\tilde{x}^*)^* \tilde{W}(\tilde{x}^*)$ be the corresponding optimal value. From Proposition 10 we know that the SZEP implies that the duality gap is zero, i.e., $d^* = f^*$, hence

$$\sum_{i=1}^{n} \lambda_i^{\star} = (\tilde{\boldsymbol{x}}^{\star})^* \tilde{\boldsymbol{W}}(\tilde{\boldsymbol{x}}^{\star}).$$
(37)

Since \tilde{x}^* is a solution of the primal, it must be feasible, hence $|\tilde{x}_i^*|^2 = 1$, $i = n, \ldots, 2n - 1$. Therefore, the following equalities hold:

$$\sum_{i=1}^{n} \lambda_{i}^{\star} = \sum_{i=1}^{n} \lambda_{i}^{\star} |\tilde{x}_{n+i-1}^{\star}|^{2} = (\tilde{x}^{\star})^{*} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \operatorname{diag}(\boldsymbol{\lambda}^{\star}) \end{bmatrix} (\tilde{x}^{\star})$$
(38)

Plugging (38) back into (37):

$$(\tilde{\boldsymbol{x}}^{\star})^{*} \left[\tilde{\boldsymbol{W}} - \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \text{diag}(\boldsymbol{\lambda}^{\star}) \end{bmatrix} \right] (\tilde{\boldsymbol{x}}^{\star}) = 0 \Leftrightarrow (\tilde{\boldsymbol{x}}^{\star})^{*} \tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) (\tilde{\boldsymbol{x}}^{\star}) = 0$$
(39)

which proves that $\tilde{x}^* \in \text{Kernel}(\tilde{W}(\lambda^*))$, which coincides with $\mathcal{N}(\lambda^*)$ as per eq. (36), proving the first claim.

Let us prove the second claim. From the first claim we know that SZEP $\Rightarrow \tilde{x}^* \in \mathcal{N}(\lambda^*)$. Moreover, when $\tilde{W}(\lambda^*)$ has a single eigenvalue in zero, then $\mathcal{N}(\lambda^*) = \text{Kernel}(\tilde{W}(\lambda^*))$ is 1dimensional and can be written as $\mathcal{N}(\lambda^*) = \{\tilde{\gamma} \, \tilde{x}^* : \tilde{\gamma} \in \mathbb{C}\}$, or, using the polar form for $\tilde{\gamma}$:

$$\mathcal{N}(\boldsymbol{\lambda}^{\star}) = \{ \eta \mathrm{e}^{\mathrm{j}\varphi} \; \tilde{\boldsymbol{x}}^{\star} : \eta, \varphi \in \mathbb{R} \}$$
(40)

From (40) we see that any $\eta \neq 1$ would alter the norm of \tilde{x}^* , leading to a solution that is not primal feasible. On the other hand any vector $e^{j\varphi}\tilde{x}^*$ is primal feasible $(|e^{j\varphi}\tilde{x}_i^*| = |\tilde{x}_i^*|)$, and primal optimal:

$$(\mathrm{e}^{\mathrm{j}\varphi}\tilde{\boldsymbol{x}}^{\star})^{*}\tilde{\boldsymbol{W}}(\mathrm{e}^{\mathrm{j}\varphi}\tilde{\boldsymbol{x}}^{\star}) = \mathrm{e}^{\mathrm{j}\varphi}\mathrm{e}^{-\mathrm{j}\varphi}(\tilde{\boldsymbol{x}}^{\star})^{*}\tilde{\boldsymbol{W}}(\tilde{\boldsymbol{x}}^{\star}) = f^{\star} \qquad (41)$$

We conclude the proof by noting that the multiplication by $e^{j\varphi}$ corresponds to a global rotation of the pose estimate \tilde{x}^* :

$$\mathrm{e}^{\mathrm{j}\varphi}\tilde{\boldsymbol{x}}^{\star} = \mathrm{e}^{\mathrm{j}\varphi}[\tilde{\rho}_{1}^{\star}\dots\tilde{\rho}_{n-1}^{\star}\tilde{r}_{1}^{\star}\dots\tilde{r}_{n}^{\star}]$$
(42)

and:
$$\begin{cases} e^{j\varphi}\rho_i^{\star} = (\boldsymbol{R}(\varphi)\boldsymbol{p}_i)^{\vee} \\ e^{j\varphi}\tilde{r}_i^{\star} = (\boldsymbol{R}(\varphi)\boldsymbol{R}_i)^{\vee} \end{cases}$$
(43)

where in (42) we observed that \tilde{x}^* only stacks position and rotation estimates, and in (43) we used the properties (18) and (19) introduced in Section III-C.

Proposition 10 provides an *a-posteriori* condition on the duality gap, that requires solving the dual problem; while Sections V-VI show that this condition is very useful in practice, it is also interesting to devise a-priori conditions, that can be assessed from \tilde{W} , without solving the dual problem. A first step in this direction is the following proposition.

Proposition 11 (No gap in trees and balanced graphs): The duality gap is zero $(d^* = f^*)$ for any balanced pose graph optimization problem, and for any pose graph whose underlying graph is a tree.

Proof. Balanced pose graphs and trees have in common the fact that they attain $f^* = 0$ (Propositions 1-2). By weak duality we know that $d^* \leq 0$. However, $\lambda = 0_n$ is feasible (as $\tilde{W} \succeq 0$) and attains $d(\lambda) = 0$, hence $\lambda = 0_n$ is feasible and dual optimal, proving $d^* = f^*$.

V. ALGORITHMS: SOLUTIONS AND VERIFICATION

In this section we exploit the results presented so far to devise an algorithm to solve PGO (Section V-A) and to design verification techniques to assess optimality (Section V-B).

A. Optimal and Suboptimal Solutions

This section shows how to use the solution λ^* of the (convex) dual problem to compute the solution of the (nonconvex) primal problem. We split the presentation in two sections: Section V-A1 discusses the case in which $\tilde{W}(\lambda^*)$ satisfies the SZEP (in this case our algorithm computes the optimal solution), and Section V-A2 discusses the case in which $\tilde{W}(\lambda^*)$ has multiple eigenvalues in zero (in this case our algorithm has no optimality guarantees). 1) Case 1: $\tilde{W}(\lambda^*)$ satisfies the SZEP: According to Corollary 1, if $\tilde{W}(\lambda^*)$ has a single zero eigenvalue, then the optimal solution of the primal problem \tilde{x}^* is in $\mathcal{N}(\lambda^*) = \text{Kernel}(\tilde{W}(\lambda^*))$. Moreover, the null space $\text{Kernel}(\tilde{W}(\lambda^*))$ is 1-dimensional, hence it can be written explicitly as:

$$\mathcal{N}(\boldsymbol{\lambda}^{\star}) = \operatorname{Kernel}(\tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star})) = \{ \tilde{\boldsymbol{v}} \in \mathbb{C}^{2n-1} : \ \tilde{\boldsymbol{v}} = \gamma \tilde{\boldsymbol{x}}^{\star}, \gamma \in \mathbb{C} \}$$
(44)

which means that any vector in the null space is a scalar multiple of the primal optimal solution \tilde{x}^* . This observation suggests an approach to compute \tilde{x}^* . We first compute an eigenvector \tilde{v} corresponding to the single zero eigenvalue of $\tilde{W}(\lambda^*)$ (this is a vector in the null space of $\tilde{W}(\lambda^*)$). Then, since \tilde{x}^* must be primal feasible (i.e., $|\tilde{x}_n| = \ldots = |\tilde{x}_{2n-1}| = 1$), we compute a suitable scalar γ that makes $\frac{1}{\gamma}\tilde{v}$ primal feasible. This scalar is clearly $\gamma = |\tilde{v}_n| = \ldots = |\tilde{v}_{2n-1}|$ (we essentially need to normalize the norm of the last n entries of \tilde{v}). The existence of a suitable γ , and hence the fact that $|\tilde{v}_n| = \ldots = |\tilde{v}_{2n-1}| > 0$, is guaranteed by Corollary 1. As a result we get the optimal solution $\tilde{x}^* = \frac{1}{2}\tilde{v}$.

2) Case 2: $\tilde{W}(\lambda^*)$ does not satisfy the SZEP: Currently we are not able to compute a guaranteed optimal solution for PGO, when $\tilde{W}(\lambda^*)$ has multiple eigenvalues in zero. Nevertheless, it is interesting to exploit the solution of the dual problem to find a (possibly suboptimal) estimate, which can be used, for instance, as initial guess for an iterative technique.

We propose an approach based on the insight of Theorem 1: if there is a primal feasible $\tilde{x} \in \mathcal{N}(\lambda^*) = \text{Kernel}(\tilde{W}(\lambda^*))$, then \tilde{x} must be primal optimal. Therefore we look for a vector $\tilde{x} \in \text{Kernel}(\tilde{W}(\lambda^*))$ that is "close" to the feasible set. Let us denote with $\tilde{V} \in \mathbb{C}^{(2n-1)\times q}$ a basis of the null space of $\tilde{W}(\lambda^*)$, where q is the number of zero eigenvalues of $\tilde{W}(\lambda^*)$.⁴ Any vector \tilde{x} in the null space of $\tilde{W}(\lambda^*)$ can be written as $\tilde{x} = \tilde{V}\tilde{z}$, for some vector $\tilde{z} \in \mathbb{C}^q$. Therefore we propose to compute a possibly suboptimal estimate $\tilde{x} = \tilde{V}\tilde{z}^*$, where \tilde{z}^* solves the following optimization problem:

$$\max_{\tilde{\boldsymbol{z}}} \sum_{i=1}^{2n-1} \operatorname{real}(\tilde{\boldsymbol{V}}_{i}\tilde{\boldsymbol{z}}) + \operatorname{imag}(\tilde{\boldsymbol{V}}_{i}\tilde{\boldsymbol{z}})$$
(45)
s.t.: $|\tilde{\boldsymbol{V}}_{i}\tilde{\boldsymbol{z}}|^{2} \leq 1,$ $i = n, \dots, 2n-1$

where \tilde{V}_i denotes the *i*-th row of \tilde{V} , and real(·) and imag(·) return the real and the imaginary part of a complex number, respectively. For an intuitive explanation of problem (45), we notice that the feasible set of the primal problem (26) is described by $|\tilde{x}_i|^2 = 1$, for i = n, ..., 2n-1. In problem (45) we relax the equality constraints to convex inequality constraints $|\tilde{x}_i|^2 \leq 1$, for i = n, ..., 2n-1; these can be written as $|\tilde{V}_i \tilde{z}|^2 \leq 1$, recalling that we are searching in the null space of $\tilde{W}(\lambda^*)$, which is spanned by $\tilde{V}\tilde{z}$. Then, the objective function in (45) encourages "large" elements $\tilde{V}_i \tilde{z}$, hence pushing the inequality $|\tilde{V}_i \tilde{z}|^2 \leq 1$ to be tight. While other metrics can force large entries $\tilde{V}_i \tilde{z}$, we preferred the linear metric (45) to preserve convexity. Note that $\tilde{x} = \tilde{V}\tilde{z}^*$, in general, is neither optimal nor feasible for our PGO problem (26), hence we need to normalize it to get a feasible estimate.

 $^{{}^4} ilde{V}$ can be computed from singular value decomposition of $ilde{W}(\lambda^{\star})$.

```
 \begin{array}{l} \mbox{input} : \mbox{Complex PGO matrix } \tilde{W} \\ \mbox{output}: \mbox{Primal solution } \tilde{x}^{\star} \mbox{ and optimality certificate } \mbox{isOpt} \\ \mbox{solve the dual problem (30) and get } \lambda^{\star} \ ; \\ \mbox{if } \tilde{W}(\lambda^{\star}) \mbox{ has a single eigenvalue } \mu_{\min} \mbox{ at zero then} \\ \mbox{ compute the eigenvector } \tilde{v} \mbox{ of } \tilde{W}(\lambda^{\star}) \mbox{ corresponding to } \mu_{\min}; \\ \mbox{ compute } \tilde{x}^{\star} = \frac{1}{\gamma} \tilde{v}, \mbox{ where } \gamma = |\tilde{v}_j|, \mbox{ for any } j \in \{n, \dots, 2n-1\}; \\ \mbox{ set isOpt = true;} \\ \mbox{ else } \\ \mbox{ compute a basis } \tilde{V} \mbox{ for the null space of } \tilde{W}(\lambda^{\star}) \mbox{ using SVD;} \\ \mbox{ compute } \tilde{z}^{\star} \mbox{ by solving the convex problem (45);} \\ \mbox{ set } \tilde{x}^{\star} = \tilde{V}\tilde{z}^{\star} \mbox{ and normalize } |\tilde{x}_i| \mbox{ to 1, for all } i = n, \dots, 2n-1; \\ \mbox{ set isOpt = unknown;} \\ \mbox{ end } \\ \mbox{ return } (\tilde{x}^{\star}, \mbox{ isOpt}) \end{array}
```



The experimental section provides empirical evidence that, despite being heuristic in nature, this method performs well in practice. Algorithm 1 wraps up the optimal solution of Section V-A1 and the method proposed in this section.

B. Verification

In this section we consider the case in which we are given an estimate for the poses in the pose graph; we call this estimate $(\mathbf{R}_i^\circ, \mathbf{p}_i^\circ)$, i = 1, ..., n; for instance, this estimate is the one returned by a state-of-the-art iterative solver, e.g., iSAM2 [34] or g2o [40]. Then we want to answer a basic question: *can we quantify how far from optimality this estimate is, and possibly certify that it is globally optimal*?

We call this candidate solution $\tilde{x}^{\circ} \in \mathbb{C}^{2n-1}$, assuming that the poses $(\mathbf{R}_{i}^{\circ}, \mathbf{p}_{i}^{\circ})$, i = 1, ..., n, have been written as a complex vector, as in Section III-C. The following result provides a first tool for verification.

Corollary 2 (Verification of Primal Objective): Given a candidate solution \tilde{x}° for the primal problem (26), and calling d^{\star} the optimal objective of the dual problem (30), if $f(\tilde{x}^{\circ}) = d^{\star}$, then the duality gap is zero and \tilde{x}° is an optimal solution of (26). Moreover, even if the duality gap is nonzero, $f(\tilde{x}^{\circ}) - f^{\star} \leq f(\tilde{x}^{\circ}) - d^{\star}$, meaning that $f(\tilde{x}^{\circ}) - d^{\star}$ is an upper-bound for the sub-optimality gap of \tilde{x}° .

Proof. The first claim follows from weak duality $(d^* \leq f^*)$ and from optimality of f^* $(f^* \leq f(\tilde{x}^\circ))$. These two facts imply that whenever $d^* = f(\tilde{x}^\circ)$, then $d^* = f^* = f(\tilde{x}^\circ)$ (zero duality gap, optimality of \tilde{x}°). Similarly, by weak duality we have $f(\tilde{x}^\circ) - f^* \leq f(\tilde{x}^\circ) - d^*$, proving the second claim. \Box

Corollary 2 ensures that the candidate \tilde{x}° is optimal when $f(\tilde{x}^{\circ}) = d^*$. Moreover, even in the case in which we get $f(\tilde{x}^{\circ}) > d^*$, the quantity $f(\tilde{x}^{\circ}) - d^*$ can be used as an indicator of how far \tilde{x}° is from the global optimum. This first verification technique is summarized in Algorithm 2. Note that whenever $\tilde{W}(\lambda^*)$ satisfies the SZEP, we have $d^* = f^*$, hence we can classify as suboptimal any candidate solution that is such that $f(\tilde{x}^{\circ}) < d^*$; however, when the SZEP is not satisfied and $f(\tilde{x}^{\circ}) < d^*$, the algorithm is inconclusive (but it can still provide indications on the sub-optimality gap subOpt).

While SDPs are convex problems (hence can be solved in polynomial time), current SDP solvers are fairly slow and do not scale to large problems, as the ones usually encountered

$$\begin{array}{l} \mbox{input} : \mbox{Complex PGO matrix } \tilde{\boldsymbol{W}}, \mbox{ candidate solution } \tilde{\boldsymbol{x}}^{\circ} \\ \mbox{output: Optimality certificate isOpt and bound subOpt } \geq f(\tilde{\boldsymbol{x}}^{\circ}) - f^{\star} \\ \mbox{solve the dual problem (30) and get } \boldsymbol{\lambda}^{\star} \mbox{ and } d^{\star} \ ; \\ \mbox{if } f(\tilde{\boldsymbol{x}}^{\circ}) == d^{\star} \mbox{ then} \\ | \mbox{ set isOpt = true;} \\ \mbox{else} \\ | \mbox{ set isOpt = false;} \\ \mbox{else} \\ | \mbox{ set isOpt = false;} \\ \mbox{else} \\ | \mbox{ set isOpt = unknown;} \\ \mbox{end} \\ \mbox{subOpt = f}(\tilde{\boldsymbol{x}}^{\circ}) - d^{\star} \ ; \\ \mbox{return (isOpt, subOpt)} \end{array}$$



in PGO. Our derivation also enables a more sophisticated verification technique, which does not require solving an SDP:

Corollary 3 (Verification of Primal Solution): Given a candidate solution \tilde{x}° for the primal problem (26), if the following linear system admits a solution

$$W(\lambda^{\circ})\tilde{x}^{\circ} = 0$$
 (to be solved w.r.t. λ°) (46)

and such solution, called λ° , is such that $\tilde{W}(\lambda^{\circ}) \succeq 0$, then the duality gap is zero and \tilde{x}° is a primal optimal solution. **Proof.** Assume that the linear system (46) admits a solution λ° , and that this solution is dual feasible ($\tilde{W}(\lambda^{\circ}) \succeq 0$). Then, $\tilde{W}(\lambda^{\circ})\tilde{x}^{\circ} = 0$ implies $(\tilde{x}^{\circ})^*\tilde{W}(\lambda^{\circ})(\tilde{x}^{\circ}) = 0$, which, recalling the structure of $\tilde{W}(\lambda^{\circ})$ in (27), also implies:

$$(\tilde{\boldsymbol{x}}^{\circ})^{*} \begin{bmatrix} \tilde{\boldsymbol{W}} - \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \text{diag}(\boldsymbol{\lambda}^{\circ}) \end{bmatrix}] (\tilde{\boldsymbol{x}}^{\circ}) = 0 \Leftrightarrow$$
$$(\tilde{\boldsymbol{x}}^{\circ})^{*} \tilde{\boldsymbol{W}}(\tilde{\boldsymbol{x}}^{\circ})(\tilde{\boldsymbol{x}}^{\circ}) = \sum_{i=1}^{n} |\tilde{\boldsymbol{x}}_{i}^{\circ}|^{2} \lambda_{i}^{\circ} \Leftrightarrow f(\tilde{\boldsymbol{x}}^{\circ}) = d(\boldsymbol{\lambda}^{\circ}) \quad (47)$$

Since λ° is dual feasible, it must hold $d(\lambda^{\circ}) \leq d^{\star}$ (recall that d^{\star} is the maximum over λ). Thus, by weak duality and optimality of f^{\star} , it holds $d(\lambda^{\circ}) \leq d^{\star} \leq f^{\star} \leq f(\tilde{x}^{\circ})$. However, from eq. (47) we known that $d(\lambda^{\circ}) = f(\tilde{x}^{\circ})$, hence the chain of inequalities becomes tight, $d(\lambda^{\circ}) = d^{\star} = f^{\star} = f(\tilde{x}^{\circ})$, which implies that \tilde{x}° attains the optimal objective f^{\star} and the duality gap is zero.

This second verification technique is more convenient in practice, since it does not require solving the SDP (30), but only requires solving a sparse linear system and then verifying that the sparse matrix $\tilde{W}(\lambda^{\circ})$ is positive definite. The following remark clarifies the structure of the linear system $\tilde{W}(\lambda^{\circ})\tilde{x}^{\circ} = 0$ (recall that \tilde{x}° is a given vector and the system is solved with respect to λ°).

Remark 3 (Feasibility and uniqueness of $\tilde{W}(\lambda^{\circ})\tilde{x}^{\circ} = 0$): Recalling the structure of $\tilde{W}(\lambda^{\circ})$ from (27), we rewrite

$$\tilde{W}(\lambda^{\circ})\tilde{x}^{\circ} = \mathbf{0} \Leftrightarrow \begin{bmatrix} \tilde{W} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \operatorname{diag}(\lambda^{\circ}) \end{bmatrix} \end{bmatrix} \tilde{x}^{\circ} = 0$$
$$\Leftrightarrow \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \operatorname{diag}(\tilde{x}_{n}^{\circ}, \dots, \tilde{x}_{2n-1}^{\circ}) \end{bmatrix} \begin{bmatrix} \mathbf{0}_{n-1} \\ \lambda^{\circ} \end{bmatrix} = \tilde{W}\tilde{x}^{\circ} \qquad (48)$$

Defining the matrix $\boldsymbol{H} \doteq \text{diag}(\tilde{x}_n^{\circ}, \dots, \tilde{x}_{2n-1}^{\circ})$, and splitting $\tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{\circ}$ into two vectors $\boldsymbol{e}_p \in \mathbb{C}^{n-1}$ and $\boldsymbol{e}_r \in \mathbb{C}^n$, such that $\tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{\circ} = [\boldsymbol{e}_p \ \boldsymbol{e}_r]$, the linear system (48) is equivalent to:

$$\boldsymbol{e}_p = \boldsymbol{0}_{n-1} \qquad \boldsymbol{H}\boldsymbol{\lambda}^\circ = \boldsymbol{e}_r \tag{49}$$

Since H is diagonal and its entries are different from zero (\tilde{x}°) is primal feasible, hence it satisfies $|\tilde{x}_i^{\circ}| = 1$, $i = n, \ldots, 2n-1$), the system $H\lambda^{\circ} = e_r$ always admits a unique solution. Therefore, the linear system (46) is feasible and admits a unique solution if and only if $e_p = \mathbf{0}_{n-1}$.

The second verification technique of Corollary 3 is summarized in Algorithm 3. Note that in the algorithm we wrote the conditions $e_p = \mathbf{0}_{n-1}$ as $||e_p||/n \leq \tau_e$, where τ_e is a small (positive) threshold $(10^{-3}$ in our experiments); this allows accommodating the presence of numerical errors. Similarly, the condition $\tilde{W}(\lambda^{\circ}) \succeq 0$ is rewritten as $\mu_{\min}(\tilde{W}(\lambda^{\circ})) \geq \tau_{\mu}$, where $\mu_{\min}(\tilde{W}(\lambda^{\circ}))$ is the smallest eigenvalue of $\tilde{W}(\lambda^{\circ})$ and τ_{μ} is a small (negative) threshold $(-10^{-3}$ in our experiments).



Algorithm 3: Solution verification using duality.

VI. NUMERICAL ANALYSIS AND DISCUSSION

This section shows that the SZEP is satisfied in the vast majority of PGO instances. Moreover, it demonstrates the effectiveness of Algorithm 1 to solve PGO, and the effectiveness of Algorithms 2-3 to check optimality of a given solution. Section VI-A presents a Monte Carlo analysis on simulated datasets, while Section VI-B reports results on real datasets.

A. Monte Carlo Analysis

Simulation setup. We consider two different simulated scenarios, called random and grid. To create a random pose graph we draw the position of n poses from a uniform distribution in a $10m \times 10m$ square. Similarly, ground truth node rotations are randomly selected in $(-\pi, +\pi]$. Then we create a set of edges defining a spanning *path* of the graph (the *odometric edges*); moreover, we add further edges to the edge set, by connecting random pairs of nodes with probability $P_c = 0.1$ (the *loop closures*). From the randomly selected *true* poses, and for each edge (i, j) in the edge set, we generate the relative pose measurement using the following model:

$$\begin{aligned}
\boldsymbol{\Delta}_{ij} &= \boldsymbol{R}_i^{\top} \left(\boldsymbol{p}_j - \boldsymbol{p}_i \right) + \boldsymbol{\epsilon}_{\Delta}, & \boldsymbol{\epsilon}_{\Delta} \sim N(\boldsymbol{0}_2, \sigma_{\Delta}^2 \boldsymbol{I}_2) \\
\boldsymbol{R}_{ij} &= \boldsymbol{R}_i^{\top} \boldsymbol{R}_j \boldsymbol{R}(\boldsymbol{\epsilon}_R), & \boldsymbol{\epsilon}_R \sim N(\boldsymbol{0}, \sigma_R^2) \end{aligned}$$
(50)

where $\epsilon_{\Delta} \in \mathbb{R}^2$ and $\epsilon_R \in \mathbb{R}$ are zero-mean Normally distributed random variables, with standard deviation σ_{Δ} and σ_R , respectively, and $\mathbf{R}(\epsilon_R)$ is a random planar rotation of an angle ϵ_R .

In the grid scenario, the trajectory is simulated as the robot is traversing the rows of a regular grid (*cf.* Fig. 3 in [12]),

and loop closures are added with probability $P_c = 0.1$ only between nearby nodes. Measurements are generated as in (50).

Statistics are computed over 50 runs for the random dataset and over 20 runs over the larger grid dataset. We use CVX [25] as parser/solver to compute the solution of the SDP in eq. (30).

Duality gap is zero in most cases. This paragraph shows that for the levels of measurement noise of practical interest, the matrix $\tilde{W}(\lambda^*)$ satisfies the Single Zero Eigenvalue Property (SZEP), hence the duality gap is zero (Proposition 10). In both the random and grid scenario we choose the weights in the cost function (5) as $\omega_{ij}^{\Delta} = 1/\sigma_{\Delta}^2$ and $\omega_{ij}^{R} = 1/\sigma_{R}^2$.

Fig. 4(a1) shows, for the random scenario, the percentage of the tests in which the matrix $\tilde{W}(\lambda^*)$ has a single zero eigenvalue, for different values of rotation noise σ_R , and keeping fixed the translation noise to $\sigma_{\Delta} = 0.1$ m. For $\sigma_R \leq 0.5$ rad, $\tilde{W}(\lambda^*)$ satisfies the SZEP in all tests. This means that, in this range of operation, Algorithm 1 is guaranteed to compute a globally optimal solution for PGO, and Algorithms 2-3 discern exactly optimal from suboptimal solutions. For $\sigma_R = 1$ rad, the percentage of successful tests drops, while still remaining larger than 90%. Note that $\sigma_R = 1$ rad is a very large rotation noise and is not far from the case in which rotation measurements are uninformative (uniformly distributed in $(-\pi, +\pi]$). Fig. 4(a2) reports the same statistics for the grid dataset, confirming that the SZEP is satisfied for reasonable noise.

Fig. 4(b1) shows, for the random scenario, the percentage of tests in which $\tilde{W}(\lambda^*)$ satisfied the SZEP, for different values of translation noise σ_{Δ} , and keeping fixed the rotation noise to $\sigma_R = 0.1$ rad. In this case, the SZEP is satisfied, regardless the level of translation noise. Similar results can be observed in Fig. 4(b2) for the grid scenario.

We also tested the percentage of experiments satisfying the SZEP for different levels of connectivity of the graph, controlled by the parameter P_c . For the random scenario we observed 100% successful experiments, independently on the choice of P_c , for $\sigma_R = \sigma_\Delta = 0.1$ and $\sigma_R = \sigma_\Delta = 0.5$. A more interesting case if shown in Fig. 4(c1) and corresponds to the case $\sigma_R = \sigma_\Delta = 1$. The SZEP is always satisfied for $P_c = 0$: this is natural as $P_c = 0$ always produces trees, for which we are guaranteed to satisfy the SZEP (Proposition 11). For $P_c = 0.1$ the SZEP fails in few runs. Finally, increasing the connectivity beyond $P_c = 0.3$ re-establishes 100% of successful tests. This would suggest that the connectivity level of the graph influences the duality gap, and better connected graphs have more changes to satisfy the SZEP. The same trend can be observed for the grid scenario, see Fig. 4(c2).

Finally, we tested the percentage of tests satisfying the SZEP for different number of nodes n. For the random scenario, considering $\sigma_R = \sigma_{\Delta} = 0.1$ and $\sigma_R = \sigma_{\Delta} = 0.5$, the SZEP was satisfied in 100% of the tests, and we omit the results for brevity. The more challenging case $\sigma_R = \sigma_{\Delta} = 1$ is shown in Fig. 4(d1). Fig. 4(d2) reports the percentage of tests with SZEP for the grid dataset, choosing $\sigma_R = \sigma_{\Delta} = 0.5$. Note that we are already considering noise levels that are above the ones encountered in practical applications (usually, $\sigma_R \ll 0.3$ rad).

We remark that current SDP solvers do not scale well to large problems, hence a Monte Carlo analysis over larger problems becomes prohibitive. The CPU time required to solve



Figure 4. Percentage of tests in which $\tilde{W}(\lambda^*)$ satisfied the SZEP property, for different (a) rotation measurement noise σ_R , (b) translation measurement noises σ_{Δ} , (c) loop closure probability P_c , (d) number of nodes n. Bars attaining 100% are shown in green (**\Box**); bars in the range [80%, 100%) are shown in yellow (-); bars lower than 80% are shown in red (-).

1

49

the SDP in eq. (30) is shown in Fig. 5.

Performance of Algorithm 1. This paragraph shows that Algorithm 1 provides an effective solution for PGO. When $W(\lambda^{\star})$ satisfies the SZEP, the algorithm is provably optimal, and it enables to solve problems that are already challenging for iterative solvers. When the $W(\lambda^*)$ does *not* satisfy the SZEP, we show that the proposed approach, while not providing performance guarantees, largely outperforms competitors.

We compare Algorithm 1 (label: A1) against three techniques: (i) a Gauss-Newton method initialized at the odometric guess (label: GNO), (ii) a Gauss-Newton method initialized at the rotation estimate of [10] (label: Lago), (iii) a Gauss-Newton method initialized at the rotation estimate produced by the eigenvector method of Singer [56] (label: EigR).

Case 1: $\hat{W}(\lambda^*)$ *satisfies the SZEP.* Fig. 6 compares the cost attained by each technique for increasing levels of rotation



CPU time (in seconds) required to solve the dual problem in Figure 5. eq. (30) for increasing number of nodes. The transparent band around the solid line corresponds to 1 sigma standard deviation.



Figure 6. Primal cost (in log scale) attained by the proposed approach A1 and by the competitors Lago, EigR, and GNO, for increasing level of rotation noise in (a) the random scenario, and (b) the grid scenario. The figure only considers problem instances in which the SZEP is satisfied. The dual optimal value $d^{\star} = f^{\star}$ is shown for comparison.

noise (this is the most challenging test case). When $W(\lambda^{\star})$ satisfies the SZEP, A1 attains f^* , hence the cyan and greed bars in Fig. 6 have the same height (up to numerical errors). The techniques Lago, EigR, and GNO attain suboptimal solutions when the noise is large. The same considerations hold for both the random (Fig. 6(a)) and the grid dataset (Fig. 6(b)). Our algorithm provides a guaranteed optimal solution in a regime that is already challenging, and in which iterative approaches fail to converge even from a good initialization.

Case 2: $W(\lambda^*)$ does not satisfy the SZEP. In this case, Algorithm 1 computes an estimate, according to the approach of Section V-A2; results are then topped-off using Gauss-Newton. To evaluate the proposed approach, we considered 50 instances in which the SZEP was not satisfied and we compared our approach against the competitors mentioned above. We also report the dual optimal value d^* for comparison (by weak duality $d^{\star} \leq f^{\star}$). For the random scenario we considered the challenging case of $\sigma_R = \sigma_\Delta = 1$ and we plot the corresponding results in Fig. 7(a1). Besides the techniques Lago, EigR, and GNO, we included in the comparison the SDP relaxation of the primal given in eq. (35) (label: SDPr). Since all techniques are suboptimal, they are topped-off with Gauss-Newton. The proposed approach A1 provides the smallest average cost (green bar). A1 also outperforms competitors in the grid scenario shown in Fig. 7(a2). It is also interesting to look at the performance of these approaches without the Gauss-Newton refinement. We show this comparison in the box plot of Fig. 7(b1) (random) and Fig. 7(b2) (grid). Note that we excluded the techniques Lago and EigR, which, without the Gauss-Newton refinement, do not provide an estimate for



Figure 7. Statistics on tests where the SZEP is *not* satisfied: (a1)-(a2): Average cost for each technique, after a Gauss-Newton refinement. (b1)-(b2): Box plot of the cost attained by a subset of the techniques, with *no* refinement.



Figure 8. Percentage of estimates classified as optimal by Algorithm 3, for different techniques and for increasing levels of rotation noise in (a) the random and (b) the grid dataset.

the positions. The proposed approach provides the best initial guess, which attains a cost close to the lower bound d^* .

Verification: Algorithm 2. Given a candidate estimate \tilde{x}° , Algorithm 2 looks at the mismatch between $f(\tilde{x}^{\circ})$ and d^{*} to evaluate the sub-optimality of \tilde{x}° , or certify optimality when $f(\tilde{x}^{\circ}) = d^{*}$. We implicitly used Algorithm 2 when comparing the techniques in Fig. 6 and Fig. 7, where we used d^{*} as a benchmark to evaluate the performance of each technique.

Verification: Algorithm 3. Algorithm 3 allows us to classify a given estimate as optimal or suboptimal. To exemplify the use of this algorithm, we use Algorithm 3 to compute the percentage of tests in which the techniques A1, Lago, EigR, and GNO converge to the optimal solution. These results are shown in Fig. 8. Algorithm 3 correctly classifies as optimal all solutions produced by A1 (A1 is guaranteed to attain f^* when the SZEP holds). Moreover, we can see that the percentage of converging tests is in agreement with the performance of the techniques in Fig. 6: for instance, for the random scenario and $\sigma_R = 0.1$, Fig. 6 shows a larger average cost of EigR is larger, and, consistently, the percentage of converging tests in Fig. 7 drops. More details on the performance of Algorithm 3 in real

	d^{\star}	f^{\star}	$f(ilde{m{x}}^\circ)$	Time
INTEL	$7.89\cdot 10^{-1}$	$7.89\cdot 10^{-1}$	$7.89\cdot 10^{-1}$	3573.3
INTEL-a	$3.92\cdot10^{+2}$	$3.92\cdot10^{+2}$	$5.96\cdot10^{+2}$	3321.9
FR079	$7.19\cdot10^{-2}$	$7.19\cdot 10^{-2}$	$7.19\cdot 10^{-2}$	3121.5
FR079-a	$2.93\cdot10^{+2}$	$2.93\cdot10^{+2}$	$3.26\cdot 10^{+3}$	1663.6
CSAIL	$1.07 \cdot 10^{-1}$	$1.07\cdot 10^{-1}$	$1.07\cdot 10^{-1}$	2297.8
CSAIL-a	$1.51 \cdot 10^{+2}$	$1.51 \cdot 10^{+2}$	$1.51 \cdot 10^{+2}$	2389.9
Table II				

DUAL OPTIMAL VALUE (d^*) , PRIMAL OPTIMAL COST (f^*) , AND COST OBTAINED BY **g2o** $(f(\tilde{x}^\circ))$, IN DIFFERENT BENCHMARKING DATASETS. LAST COLUMN REPORTS CPU TIME IN SECONDS.

benchmarking datasets is discussed in the following.

B. Tests on benchmarking datasets

In this section we test the verification techniques of Algorithms 2-3 in practical problem instances. Besides considering standard benchmarking datasets, we introduce new challenging datasets to test the limits of applicability of our techniques.

- We consider the following datasets:
 INTEL: Intel Research Lab [41] (n = 1228, m = 1505);
- FR079: University of Freiburg, building 079 [41] (n = 989, m = 1217);
- CSAIL: *MIT*, CSAIL building [41] (n = 1045, m = 1172);
- M3500: Manhattan world [47] (n=3500, m=5453);
- INTEL-a, FR079-a, CSAIL-a: variants of the INTEL, FR079, and CSAIL datasets with extra additive noise on rotation measurements (std: 0.1 rad);
- M3500-a, M3500-b, M3500-c: variants of the M3500 dataset with extra additive noise on rotation measurements (std: 0.1, 0.2, and 0.3 rad, respectively) [11].

CVX was not able to solve the large SDPs arising from these datasets and it run out of memory in all tests. Therefore, for the tests in this section we used NEOS [27], an online service designed to solve large optimization problems. We chose sdpt3 as SDP solver in NEOS.

Table II reports the dual optimal value computed by sdpt3 and two primal costs, f^* and $f(\tilde{x}^\circ)$. f^* is obtained by bootstrapping a Gauss-Newton method with the algorithm presented in [11]. While this technique is not guaranteed to converge to the global minimum, we can verify optimality a posteriori, observing from the table that $d^{\star} = f^{\star}$ in all problem instances. The equality $d^{\star} = f^{\star}$ also shows that the duality gap is zero in practical problems instances, confirming the results of our Monte Carlo analysis. The cost $f(\tilde{x}^{\circ})$ is the one produced by g2o [40], a state-of-the-art iterative solver which refines the odometric guess. In the easy scenarios (INTEL, FR079, CSAIL) g20 converged to the correct solution and using d^{\star} we can correctly certify optimality of the resulting estimates. In some of the noisier scenarios (INTELa, FR079-a) g2o converges to a wrong estimate (Fig. 1(a1)-(a2)); consistently, Table II reveals a large mismatch between d^{\star} and $f(\tilde{x}^{\circ})$. Also sdpt3 run into memory issues for larger PGO instances, hence we were not able to solve the SDP for the M3500 scenarios. The last column in Table II reports the time required to solve the dual problem. The time is prohibitive and currently limits the use of Algorithms 1-2 in

		μ_{\min}	$\ oldsymbol{e}_p\ /n$	Time
INTEL-a	$f^{\star} = 3.92 \cdot 10^{+2}$	$4.9\cdot 10^{-8}$	$6.7\cdot 10^{-7}$	9.7
	$f(\tilde{x}^{\circ}) = 5.96 \cdot 10^{+2}$	$-7.1\cdot 10^0$	$1.3 \cdot 10^{-3}$	0.1
FR079-a	$f^{\star} = 2.93 \cdot 10^{+2}$	$6.3 \cdot 10^{-7}$	$5.8\cdot 10^{-6}$	5.4
	$f(\tilde{x}^{\circ}) = 3.26 \cdot 10^{+3}$	$-4.7 \cdot 10^{+1}$	$6.5 \cdot 10^{-1}$	0.4
CSAIL-a	$f^{\star} = 1.51 \cdot 10^{+2}$	$1.3\cdot 10^{-6}$	$5.5\cdot 10^{-5}$	4.6
	$f(\tilde{x}^{\circ}) = 1.51 \cdot 10^{+2}$	$-7.7 \cdot 10^{-8}$	$2.3\cdot\mathbf{10^{-4}}$	4.0
M3500-a	$f^{\star} = 9.12 \cdot 10^{+2}$	$6.1 \cdot 10^{-7}$	$2.8\cdot 10^{-5}$	100
	$f(\tilde{x}^{\circ}) = 1.13 \cdot 10^{+6}$	$-6.8 \cdot 10^{+2}$	$3.6\cdot 10^0$	130
M3500-b	$f^{\star} = 2.05 \cdot 10^{+3}$	$2.6\cdot\mathbf{10^{-7}}$	$3.2\cdot 10^{-4}$	187
	$f(\tilde{x}^{\circ}) = 1.50 \cdot 10^{+6}$	$-1.1 \cdot 10^{+3}$	$4.5\cdot 10^0$	101
М3500-с	$f^{\star} = 2.55 \cdot 10^{+3}$	$1.3\cdot 10^{-6}$	$2.7\cdot\mathbf{10^{-4}}$	187
	$f(\tilde{x}^{\circ}) = 1.08 \cdot 10^{+9}$	$-3.0 \cdot 10^{+4}$	$1.1\cdot10^{+2}$	101

Table III

Performance of Algorithm 3 on Benchmarking datasets. μ_{\min} is the smallest eigenvalue of $\tilde{W}(\lambda^{\circ})$. e_p is the residual error discussed in Remark 3. Time is expressed in seconds.

large problems; fortunately, as shown below, we can perform optimality verification in large scenarios using Algorithm 3.

Table III discusses the performance of Algorithm 3. For space reasons, we only focus on the most challenging scenarios (the ones in which we inflated extra noise). Each row of the table corresponds to a candidate solution that we want to verify. The rows starting with f^* correspond to "good" estimates (computed as in [11]): these estimates are visually correct, and for some of them (INTEL, FR079, CSAIL, CSAIL-a) we already have a certificate of optimality from Table II. On the other hand, the rows starting with $f(\tilde{x}^{\circ})$ contain the estimate returned by g20. We already know from Table II and from previous work [11], [12] that g2o is not able to converge to the global optimum when the noise in the odometric estimate is large. Indeed, we see that in most of the rows starting with $f(\tilde{x}^{\circ})$, g2o attains a suboptimal cost, shown in red (only in CSAIL-a, g20 attains f^*). The columns labelled with μ_{\min} and $\|\boldsymbol{e}_p\|/n$ report the two quantities used to verify optimality in Algorithm 3; $\mu_{\min} \doteq \mu_{\min}(\boldsymbol{W}(\boldsymbol{\lambda}^{\circ}))$ is the smallest eigenvalue of $\widehat{W}(\lambda^{\circ})$; e_p is the residual error discussed in Remark 3. Our algorithm certifies a solution as optimal if $||e_p||/n \le \tau_e$ and $\mu_{\min} \ge \tau_{\mu}$; in our tests we set $\tau_e = 10^{-3}$ and $\tau_{\mu} = -10^{-3}$. For each row in Table III, we show in green the entries that are within these bounds: these are certified as optimal by Algorithm 3. The entries in red are the ones that exceed the thresholds: Algorithm 3 correctly identifies as suboptimal those estimates. Fig. 1 shows the trajectory estimates corresponding to the "good" and the "bad" rows of Table III. Table III suggests that the value of μ_{\min} is very discriminative in distinguishing optimal from suboptimal solutions. Moreover, Algorithm 3 is effective in identifying sub-optimal solutions even when these only imply small artifacts, e.g., the small incorrect loop on the bottom right of Fig. 1(a1). The last column of Table III reports the CPU time required by Algorithm 3; this is essentially the time to compute the smallest eigenvalue of $W(\lambda^{\circ})$. For the results in this paper we used the Matlab command "eig", and we leave computational aspects (e.g., use algorithms that leverage sparsity) for future work.

14

VII. CONCLUSION

Lagrangian duality offers an appealing approach to compute a globally optimal solution for PGO and to verify optimality of the estimates returned by iterative solvers. We propose three contributions. First, we rephrase PGO as a problem in complex variables. This allows drawing connection with the recent literature on unit gain graphs, and enables results on the spectrum of the pose graph matrix. Second, we formulate the Lagrangian dual problem and we analyze the relations between the primal and the dual solutions. Our key result proves that the duality gap is connected to the number of eigenvalues of the penalized pose graph matrix. In particular, if this matrix has a single eigenvalue in zero (SZEP), then (i) the duality gap is zero, (ii) the primal PGO problem has a unique solution (up to an arbitrary roto-translation), and (iii) the primal solution can be computed by *scaling* an eigenvector of the penalized pose graph matrix. The third contribution is algorithmic: we propose an algorithm that returns a guaranteed optimal solution when the SZEP is satisfied, and (empirically) provides a very good estimate when the SZEP fails. Moreover, we provide two algorithms to verify the quality of a given estimate. One of the two algorithms does not require solving the dual problem, hence it can be applied to large-scale problems. Numerical results show that the SZEP holds for noise levels of practical applications, and confirms the effectiveness of the proposed algorithms in both simulations and real datasets.

VIII. APPENDIX

A. Proof of Proposition 5: Cost in the Complex Domain

Let us prove the equivalence between the complex cost and its real counterpart, as stated in Proposition 5.

We first observe that the dot product between two 2-vectors $x_1, x_2 \in \mathbb{R}^2$, can be written in terms of their complex representation $\tilde{x}_1 \doteq x_1^{\vee}$, and $\tilde{x}_2 \doteq x_2^{\vee}$, as follows:

$$\boldsymbol{x}_{1}^{\top}\boldsymbol{x}_{2} = \frac{\tilde{x}_{1}^{*}\tilde{x}_{2} + \tilde{x}_{1}\tilde{x}_{2}^{*}}{2}$$
(51)

Moreover, we know that the action of a matrix $Z \in \alpha SO(2)$ can be written as the product of complex numbers, see (18). Combining (51) and (18) we get:

$$(\boldsymbol{x}_{1}^{\top}\boldsymbol{Z}\boldsymbol{x}_{2})^{\vee} = \frac{\tilde{x}_{1}^{*} \, \tilde{z} \, \tilde{x}_{2} + \tilde{x}_{1} \, \tilde{z}^{*} \, \tilde{x}_{2}^{*}}{2} \tag{52}$$

where $\tilde{z} = Z^{\vee}$. Furthermore, when Z is multiple of the identity matrix, it easy to see that $z = Z^{\vee}$ is actually a real number, and Eq. (52) becomes:

$$(\boldsymbol{x}_1^{\top} \boldsymbol{Z} \boldsymbol{x}_1)^{\vee} = \tilde{\boldsymbol{x}}_1^* \ \boldsymbol{z} \ \tilde{\boldsymbol{x}}_1 \tag{53}$$

Defining the real vector $\boldsymbol{x} \doteq [\boldsymbol{\rho} \ \boldsymbol{r}]$ we write the left-hand side of (22) as $\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}$. With the machinery introduced so far we are ready to express $\boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x}$ in complex form. Since \boldsymbol{W} is symmetric, the product becomes:

$$\boldsymbol{x}^{\top} \boldsymbol{W} \boldsymbol{x} = \sum_{i=1}^{2n-1} \left[\boldsymbol{x}_i^{\top} [\boldsymbol{W}]_{ii} \boldsymbol{x}_i + \sum_{j=i+1}^{2n-1} 2 \boldsymbol{x}_i^{\top} [\boldsymbol{W}]_{ij} \boldsymbol{x}_j \right]$$
(54)

where $x_i \in \mathbb{R}^2$ is the subvector of x at indices (2i - 1, 2i). Using the fact that $[W]_{ii}$ is a multiple of the identity matrix, $\tilde{W}_{ii} \doteq [W]_{ii}^{\vee} \in \mathbb{R}$, and using (53) we conclude $x_i^{\top}[W]_{ii}x_i = \tilde{x}_i^*\tilde{W}_{ii}\tilde{x}_i$. Moreover, defining $\tilde{W}_{ij} \doteq [W]_{ij}^{\vee}$ (these will be complex numbers, in general), and using (52), eq. (54) becomes:

$$\boldsymbol{x}^{\top} \boldsymbol{W} \boldsymbol{x} = \sum_{i=1}^{2n-1} \left[\tilde{x}_{i}^{*} \tilde{\boldsymbol{W}}_{ii} \tilde{x}_{i} + \sum_{j=i+1}^{2n-1} (\tilde{x}_{i}^{*} \tilde{\boldsymbol{W}}_{ij} \tilde{x}_{j} + \tilde{x}_{i} \tilde{\boldsymbol{W}}_{ij}^{*} \tilde{x}_{j}^{*}) \right]$$
$$= \sum_{i=1}^{2n-1} \left[\tilde{x}_{i}^{*} \tilde{\boldsymbol{W}}_{ii} \tilde{x}_{i} + \sum_{j \neq i} \tilde{x}_{i}^{*} \tilde{\boldsymbol{W}}_{ij} \tilde{x}_{j} \right] = \tilde{\boldsymbol{x}}^{*} \tilde{\boldsymbol{W}} \tilde{\boldsymbol{x}} \quad (55)$$

where we set the lower triangular entries of \tilde{W} to $\tilde{W}_{ji} = \tilde{W}_{ij}^*$.

B. Proof of Proposition 7: Spectrum of Complex and Real Pose Graph Matrices

Recall that any Hermitian matrix has real eigenvalues, and possibly complex eigenvectors. Let $\mu \in \mathbb{R}$ be an eigenvalue of \tilde{W} , associated with an eigenvector $\tilde{v} \in \mathbb{C}^{2n-1}$, i.e.,

$$\tilde{W}\tilde{v} = \mu\tilde{v}$$
 (56)

From equation (56) we have, for $i = 1, \ldots, 2n - 1$,

$$\sum_{j=1}^{2n-1} \tilde{\boldsymbol{W}}_{ij} \tilde{\boldsymbol{v}}_j = \mu \tilde{\boldsymbol{v}}_i \Leftrightarrow \sum_{j=1}^{2n-1} [\boldsymbol{W}]_{ij} \boldsymbol{v}_j = \mu \boldsymbol{v}_i$$
(57)

where $v_i \in \mathbb{R}^2$ is such that $v_i^{\vee} = \tilde{v}_i$. Since eq. (57) holds for all i = 1, ..., 2n - 1, it can be written in compact form as:

$$\boldsymbol{W}\boldsymbol{v} = \boldsymbol{\mu}\boldsymbol{v} \tag{58}$$

hence v is an eigenvector of the real anchored pose graph matrix W, associated with the eigenvalue μ . This proves that any eigenvalue of \tilde{W} is also an eigenvalue of W.

To prove that the eigenvalue μ is actually repeated twice in W, consider now equation (56) and multiply both members by the complex number $e^{j\frac{\pi}{2}}$:

$$\tilde{\boldsymbol{W}}\tilde{\boldsymbol{v}}\mathrm{e}^{\mathrm{J}\frac{\pi}{2}} = \mu \tilde{\boldsymbol{v}}\mathrm{e}^{\mathrm{J}\frac{\pi}{2}}$$
(59)

For i = 1, ..., 2n - 1, we have:

$$\sum_{j=1}^{2n-1} \tilde{W}_{ij}^* \tilde{v}_j \mathrm{e}^{\mathrm{J}\frac{\pi}{2}} = \mu \tilde{v}_i \mathrm{e}^{\mathrm{J}\frac{\pi}{2}} \Leftrightarrow \sum_{j=1}^{2n-1} [W]_{ij} w_j = \mu w_i \qquad (60)$$

where w_i is such that $w_i^{\vee} = \tilde{v}_j e^{j\frac{\pi}{2}}$. Since eq. (60) holds for all i = 1, ..., 2n-1, it can be written in compact form as:

$$\boldsymbol{W}\boldsymbol{w} = \boldsymbol{\mu}\boldsymbol{w} \tag{61}$$

hence also w is an eigenvector of W associated with the eigenvalue μ .

Now it only remains to demonstrate that v and w are linearly independent. One can readily check that, if \tilde{v}_i is in the form $\tilde{v}_i = \eta_i e^{j\theta_i}$, then

$$\boldsymbol{v}_i = \eta_i \begin{bmatrix} \cos(\theta_i) \\ \sin(\theta_i) \end{bmatrix}.$$
 (62)

Moreover, observing that $\tilde{v}_j e^{j\frac{\pi}{2}} = \eta_i e^{j(\theta_i + \pi/2)}$, then

$$\boldsymbol{w}_{i} = \eta_{i} \begin{bmatrix} \cos(\theta_{i} + \pi/2) \\ \sin(\theta_{i} + \pi/2) \end{bmatrix} = \eta_{i} \begin{bmatrix} -\sin(\theta_{i}) \\ \cos(\theta_{i}) \end{bmatrix}$$
(63)

From (62) and (63) is it easy to see that $v^{\top}w = 0$, thus v, w are orthogonal, hence independent. To each eigenvalue μ of \tilde{W} there thus correspond an identical eigenvalue of W, of geometric multiplicity at least two. Since \tilde{W} has 2n - 1 eigenvalues and W has 2(2n - 1) eigenvalues, we conclude that to each eigenvalue μ of \tilde{W} there correspond exactly two eigenvalues of W in μ . The previous proof also shows how the set of orthogonal eigenvectors of W is related to the set of eigenvectors of \tilde{W} .

C. Proof of Proposition 9: Dual Optimal Solution

Let us first prove that the dual optimal value is attained at a finite λ^* . Since $\tilde{W}(\lambda) \succeq 0$ implies that the diagonal entries are nonnegative, the feasible set of (30) is contained in the set $\{\lambda : \tilde{Q}_{ii} - \lambda_i \ge 0, i = 1, ..., n\}$.⁵ On the other hand, $\bar{\lambda} = \mathbf{0}_{2n-1}$ is feasible and all vectors in the set $\{\lambda : \lambda_i \ge 0\}$ yield an objective that is at least as good as the objective at $\bar{\lambda}$. Therefore, problem (30) is equivalent to $\max_{\lambda} \sum_i \lambda_i$ subject to the original constraint, plus a box constraint $\lambda \in \{0 \le \lambda_i \le \tilde{Q}_{ii}, i = 1, ..., n\}$. Thus we maximize a linear function over a compact set, hence a finite solution λ^* must be attained.

Now let us prove that $W(\lambda^*)$ has an eigenvalue in zero. Assume by contradiction that $\tilde{W}(\lambda^*) \succ 0$. From the Schur complement rule applied to $\tilde{W}(\lambda^*)$ (cf. (27)) we know:

$$\tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) \succ 0 \Leftrightarrow \begin{cases} \boldsymbol{L} \succ 0\\ \tilde{\boldsymbol{Q}}(\boldsymbol{\lambda}^{\star}) - \tilde{\boldsymbol{S}}^{\star} \boldsymbol{L}^{-1} \tilde{\boldsymbol{S}} \succ 0 \end{cases}$$
(64)

The condition $L \succ 0$ is always satisfied for a connected graph, since $L = A^{\top}A$, and the anchored incidence matrix A, obtained by removing a node from the original incidence matrix, is always full-rank for connected graphs [54, Section 19.3]. Therefore, our assumption $\tilde{W}(\lambda^*) \succ 0$ implies that

$$\tilde{\boldsymbol{Q}}(\boldsymbol{\lambda}^{\star}) - \tilde{\boldsymbol{S}}^{\star} \boldsymbol{L}^{-1} \tilde{\boldsymbol{S}} = \tilde{\boldsymbol{Q}} - \tilde{\boldsymbol{S}}^{\star} \boldsymbol{L}^{-1} \tilde{\boldsymbol{S}} - \operatorname{diag}(\boldsymbol{\lambda}^{\star}) \succ 0 \quad (65)$$

Now, let

$$\epsilon = \lambda_{\min}(\tilde{Q}(\lambda^{\star}) - \tilde{S}^{\star}L^{-1}\tilde{S}) > 0.$$

which is positive by the assumption $\tilde{W}(\lambda^*) \succ 0$. Consider $\lambda = \lambda^* + \epsilon \mathbf{1}$, then

$$\tilde{\boldsymbol{Q}}(\boldsymbol{\lambda}) - \tilde{\boldsymbol{S}}^* \boldsymbol{L}^{-1} \tilde{\boldsymbol{S}} = \tilde{\boldsymbol{Q}}(\boldsymbol{\lambda}) - \tilde{\boldsymbol{S}}^* \boldsymbol{L}^{-1} \tilde{\boldsymbol{S}} - \epsilon \boldsymbol{I} \succeq \boldsymbol{0},$$

thus λ is dual feasible, and $\sum_i \lambda_i > \sum_i \lambda_i^*$, which would contradict optimality of λ^* . We thus proved that $\tilde{Q}(\lambda^*)$ must have a zero eigenvalue.

D. Proof of Proposition 10: No Duality Gap when SZEP holds

We have already observed in Proposition 8 that (35) is the dual problem of (30), therefore, we can interpret \tilde{X} as a Lagrange multiplier for the constraint $\tilde{W}(\lambda) \succeq 0$. If we consider the optimal solutions \tilde{X}^* and λ^* of (35) and (30), respectively, the *complementary slackness* condition ensures

⁵We recall that \tilde{Q}_{ii} is the bottom right block of \tilde{W} as per (24), and that the diagonal terms of \tilde{W} (and hence of \tilde{Q}) are reals according to Remark 2.

that Tr $(\tilde{W}(\lambda^*)\tilde{X}^*) = 0$ (see [5, Example 5.13]). Let us parametrize $\tilde{X}^* \succeq 0$ as

$$\tilde{\boldsymbol{X}}^{\star} = \sum_{i=1}^{2n-1} \mu_i \tilde{\boldsymbol{v}}_i \tilde{\boldsymbol{v}}_i^*,$$

where $0 \leq \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{2n-1}$ are the eigenvalues of \tilde{X} , and \tilde{v}_i form a unitary set of eigenvectors. Then, the complementary slackness condition becomes

$$\operatorname{Tr} \left(\tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) \tilde{\boldsymbol{X}}^{\star} \right) = \operatorname{Tr} \left(\tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) \sum_{i=1}^{2n-1} \mu_i \tilde{\boldsymbol{v}}_i \tilde{\boldsymbol{v}}_i^{\star} \right)$$
$$= \sum_{i=1}^{2n-1} \mu_i \operatorname{Tr} \left(\tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) \tilde{\boldsymbol{v}}_i \tilde{\boldsymbol{v}}_i^{\star} \right) = \sum_{i=1}^{2n-1} \mu_i \tilde{\boldsymbol{v}}_i^{\star} \tilde{\boldsymbol{W}}(\boldsymbol{\lambda}^{\star}) \tilde{\boldsymbol{v}}_i = 0.$$

Since $\tilde{W}(\lambda^*) \succeq 0$, the above quantity is zero at a nonzero \tilde{X}^* (\tilde{X}^* cannot be zero since it needs to satisfy the constraints $\tilde{X}_{ii} = 1$) if and only if $\mu_i = 0$ for i = m + 1, ..., 2n - 1, and $\tilde{W}(\lambda^*)\tilde{v}_i = 0$ for i = 1, ..., m, where m is the multiplicity of 0 as an eigenvalue of $\tilde{W}(\lambda^*)$. Hence \tilde{X}^* has the form

$$\tilde{\boldsymbol{X}}^{\star} = \sum_{i=1}^{m} \mu_i \tilde{\boldsymbol{v}}_i \tilde{\boldsymbol{v}}_i^*, \tag{66}$$

where \tilde{v}_i , i = 1, ..., m, form a unitary basis of the nullspace of $\tilde{W}(\lambda^*)$. Now, if m = 1, then the solution \tilde{X}^* to problem (35) has rank one, but according to Proposition 8 this implies $d^* = f^*$, proving the claim.

E. Proof of Theorem 1: Primal-dual Optimal Pairs

We prove that, given $\lambda \in \mathbb{R}^n$, if an $\tilde{x}_{\lambda} \in \mathcal{N}(\lambda)$ is primal feasible, then \tilde{x}_{λ} is primal optimal; moreover, λ is dual optimal, and the duality gap is zero.

By weak duality we know that for any λ :

$$\mathcal{L}(\tilde{\boldsymbol{x}}_{\lambda}, \boldsymbol{\lambda}) \le f^{\star} \tag{67}$$

However, if \tilde{x}_{λ} is primal feasible, by optimality of f^{\star} , it must also hold

$$f^{\star} \le f(\tilde{\boldsymbol{x}}_{\lambda}) \tag{68}$$

Now we observe that for a feasible \tilde{x}_{λ} , the terms in the Lagrangian associated to the constraints disappear and $\mathcal{L}(\tilde{x}_{\lambda}, \lambda) = f(\tilde{x}_{\lambda})$. Using the latter equality and the inequalities (67) and (68) we get:

$$f^{\star} \leq f(\tilde{\boldsymbol{x}}_{\lambda}) = \mathcal{L}(\tilde{\boldsymbol{x}}_{\lambda}, \boldsymbol{\lambda}) \leq f^{\star}$$
(69)

which implies $f(\tilde{x}_{\lambda}) = f^{\star}$, i.e., \tilde{x}_{λ} is primal optimal.

Further, we have that

$$d^{\star} \geq \min_{\tilde{\boldsymbol{x}}} \mathcal{L}(\tilde{\boldsymbol{x}}, \boldsymbol{\lambda}) = \mathcal{L}(\tilde{\boldsymbol{x}}_{\lambda}, \boldsymbol{\lambda}) = f(\tilde{\boldsymbol{x}}_{\lambda}) = f^{\star},$$

which, combined with weak duality $(d^* \leq f^*)$, implies that $d^* = f^*$ and that λ attains the dual optimal value.

REFERENCES

- R. Aragues, L. Carlone, G. Calafiore, and C. Sagues. Multi-agent localization from noisy relative pose measurements. In *IEEE Intl. Conf.* on Robotics and Automation (ICRA), pages 364–369, Shanghai, China, 2011.
- [2] M. Arie-Nachimson, S. Z. Kovalsky, I. Kemelmacher-Shlizerman, A. Singer, and R. Basri. Global motion estimation from point matches. In *3DIMPVT*, 2012.
- [3] A.S. Bandeira, A. Singer, and D.A. Spielman. A Cheeger inequality for the graph connection laplacian. *SIAM. J. Matrix Anal. & Appl.*, 34(4):1611–1630, 2013.
- [4] D. Borra, R. Carli, E. Lovisari, F. Fagnani, and S. Zampieri. Autonomous calibration algorithms for planar networks of cameras. *Automatica*, 2012.
- [5] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- [6] G. Calafiore, L. Carlone, and F. Dellaert. Pose graph optimization in the complex domain: Lagrangian duality, conditions for zero duality gap, and optimal solutions. In arXiv preprint: 1505.03437, 2015.
- [7] Nicholas Carlevaris-Bianco and Ryan M. Eustice. Generic factor-based node marginalization and edge sparsification for pose-graph SLAM. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 5728–5735, 2013.
- [8] L. Carlone. Convergence analysis of pose graph optimization via Gauss-Newton methods. In *IEEE Intl. Conf. on Robotics and Automation* (*ICRA*), pages 965–972, 2013.
- [9] L. Carlone, R. Aragues, J.A. Castellanos, and B. Bona. A linear approximation for graph-based simultaneous localization and mapping. In *Robotics: Science and Systems (RSS)*, 2011.
- [10] L. Carlone, R. Aragues, J.A. Castellanos, and B. Bona. A fast and accurate approximation for planar pose graph optimization. *Intl. J. of Robotics Research*, 2014.
- [11] L. Carlone and A. Censi. From angular manifolds to the integer lattice: Guaranteed orientation estimation with application to pose graph optimization. *IEEE Trans. Robotics*, 2014.
- [12] L. Carlone and F. Dellaert. Duality-based verification techniques for 2D SLAM. In Intl. Conf. on Robotics and Automation (ICRA), accepted, 2015.
- [13] L. Carlone, R. Tron, K. Daniilidis, and F. Dellaert. Initialization techniques for 3D SLAM: a survey on rotation estimation and its use in pose graph optimization. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2015.
- [14] G. S. Chirikjian. Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications (Applied and Numerical Harmonic Analysis). Birkhauser, 2012.
- [15] A. Chiuso, G. Picci, and S. Soatto. Wide-sense estimation on the special orthogonal group. *Commun. Inf. Syst*, 8:185–200, 2008.
- [16] F. Dellaert, J. Carlson, V. Ila, K. Ni, and C.E. Thorpe. Subgraphpreconditioned conjugate gradient for large scale slam. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2010.
- [17] G. Dubbelman, P. Hansen, B. Browning, and M.B. Dias. Orientation only loop-closing with closed-form trajectory bending. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2012.
- [18] T. Duckett, S. Marsland, and J. Shapiro. Fast, on-line learning of globally consistent maps. *Autonomous Robots*, 12(3):287–300, 2002.
- [19] M. Hazewinkel ed. Complex number, Encyclopedia of Mathematics. Springer, 2001.
- [20] R.M. Eustice, H. Singh, J.J. Leonard, and M.R. Walter. Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters. *Intl. J. of Robotics Research*, 25(12):1223–1242, Dec 2006.
- [21] J. Fredriksson and C. Olsson. Simultaneous multiple rotation averaging using lagrangian duality. In Asian Conf. on Computer Vision (ACCV), 2012.
- [22] U. Frese, P. Larsson, and T. Duckett. A multilevel relaxation algorithm for simultaneous localisation and mapping. *IEEE Trans. Robotics*, 21(2):196–207, April 2005.
- [23] V. M. Govindu. Combining two-view constraints for motion estimation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 218–225, 2001.
- [24] V.M. Govindu. Lie-algebraic averaging for globally consistent motion estimation. In *IEEE Conf. on Computer Vision and Pattern Recognition* (CVPR), 2004.
- [25] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming.

- [26] G. Grisetti, C. Stachniss, and W. Burgard. Non-linear constraint network optimization for efficient map learning. *Trans. on Intelligent Transportation systems*, 10(3):428–439, 2009.
- [27] W. Gropp and J. J. Moré. Optimization Environments and the NEOS Server. Approximation Theory and Optimization, M. D. Buhmann and A. Iserles, eds., Cambridge University Press, 1997.
- [28] R. Hartley, J. Trumpf, Y. Dai, and H. Li. Rotation averaging. *IJCV*, 103(3):267–305, 2013.
- [29] T. Hatanaka, M. Fujita, and F. Bullo. Vision-based cooperative estimation via multi-agent optimization. In *IEEE Conf. on Decision and Control*, 2010.
- [30] S. Huang, Y. Lai, U. Frese, and G. Dissanayake. How far is SLAM from a linear least squares problem? In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 3011–3016, 2010.
- [31] S. Huang, H. Wang, U. Frese, and G. Dissanayake. On the number of local minima to the point feature based SLAM problem. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 2074–2079, 2012.
- [32] V. Indelman, E. Nelson, N. Michael, and F. Dellaert. Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization. In *IEEE Intl. Conf. on Robotics* and Automation (ICRA), 2014.
- [33] H. Johannsson, M. Kaess, M. Fallon, and J.J. Leonard. Temporally scalable visual SLAM using a reduced pose graph. In *IEEE Intl. Conf.* on Robotics and Automation (ICRA), pages 54–61, 2013.
- [34] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *Intl. J. of Robotics Research*, 31:217–236, Feb 2012.
- [35] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental smoothing and mapping. *IEEE Trans. Robotics*, 24(6):1365–1378, Dec 2008.
- [36] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy, and S. Teller. Multiple relative pose graphs for robust cooperative mapping. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 3185–3192, Anchorage, Alaska, May 2010.
- [37] J. Knuth and P. Barooah. Collaborative 3D localization of robots from relative pose measurements using gradient descent on manifolds. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1101–1106, 2012.
- [38] J. Knuth and P. Barooah. Error growth in position estimation from noisy relative pose measurements. *Robotics and Autonomous Systems*, 61(3):229–224, 2013.
- [39] K. Konolige. Large-scale map-making. In Proc. 21th AAAI National Conference on AI, San Jose, CA, 2004.
- [40] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. g20: A general framework for graph optimization. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, Shanghai, China, May 2011.
- [41] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, and A. Kleiner. Slam benchmarking webpage, 2009.
- [42] F. Dellaert L. Carlone, G. Calafiore. Pose graph optimization in the complex domain: Lagrangian duality and optimal solutions. In *submitted* to the RSS workshop "Reviewing the review process", available online at http://openreview.informatik.uni-freiburg.de/node/41, 2015.
- [43] M. Liu, S. Huang, G. Dissanayake, and H. Wang. A convex optimization based approach for pose SLAM problems. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 1898–1903, 2012.
- [44] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. Autonomous Robots, pages 333–349, Apr 1997.
- [45] D. Martinec and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.
- [46] R. Olfati-Saber. Swarms on sphere: A programmable swarm with synchronous behaviors like oscillator networks. In *IEEE Conf. on Decision and Control*, pages 5060–5066, 2006.
- [47] E. Olson, J. Leonard, and S. Teller. Fast iterative alignment of pose graphs with poor initial estimates. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 2262–2269, May 2006.
- [48] J. R. Peters, D. Borra, B. Paden, and F. Bullo. Sensor network localization on the group of 3D displacements. *SIAM Journal on Control* and Optimization, submitted, 2014.
- [49] G. Piovan, I. Shames, B. Fidan, F. Bullo, and B. Anderson. On frame and orientation localization for relative sensing networks. *Automatica*, 49(1):206–213, 2013.
- [50] N. Reff. Spectral properties of complex unit gain graphs. *Linear Algebra and its Applications*, 436(9):3165–3176, 2012.

- [51] D.M. Rosen, M. Kaess, and J.J. Leonard. RISE: An incremental trustregion method for robust online sparse least-squares estimation. *IEEE Trans. Robotics*, 2014.
- [52] A. Sarlette and R. Sepulchre. Consensus optimization on manifolds. SIAM J. Control and Optimization, 48(1):56–76, 2009.
- [53] J. Saunderson, P.A. Parrilo, and A. Willsky. Semidefinite relaxations for optimization problems over rotation matrices. In *IEEE Conf. on Decision and Control*, May 2014.
- [54] Al. Schrijver. Theory of Linear and Integer Programming. Wiley, 1998.
- [55] G.C. Sharp, S.W. Lee, and D.K. Wehe. Multiview registration of 3D scenes by minimizing error between coordinate frames. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(8):1037–1050, 2004.
- [56] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. Appl. Comput. Harmon. Anal., 30:20–36, 2010.
- [57] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in Cryo-EM by eigenvectors and semidefinite programming. SIAM J. Imaging Sciences, pages 543–572, 2011.
- [58] J. Thunberg, E. Montijano, and X. Hu. Distributed attitude synchronization control. In *IEEE Conf. on Decision and Control*, 2011.
- [59] R. Tron, B. Afsari, and R. Vidal. Intrinsic consensus on SO(3) with almost global convergence. In *IEEE Conf. on Decision and Control*, 2012.
- [60] R. Tron, B. Afsari, and R. Vidal. Riemannian consensus for manifolds with bounded curvature. *IEEE Trans. on Automatic Control*, 2012.
- [61] R. Tron and R. Vidal. Distributed image-based 3-D localization in camera networks. In *IEEE Conf. on Decision and Control*, 2009.
- [62] H. Wang, G. Hu, S. Huang, and G. Dissanayake. On the structure of nonlinearities in pose graph SLAM. In *Robotics: Science and Systems* (*RSS*), 2012.
- [63] L. Wang and A. Singer. Exact and stable recovery of rotations for robust synchronization. *Information and Inference: A Journal of the IMA*, 30, 2013.