Convex Global 3D Registration with Lagrangian Duality

Jesus Briales   Javier Gonzalez-Jimenez

MAPIR-UMA Group

University of Malaga, Spain

{jesusbriales,javiergonzalez}@uma.es

Abstract

The registration of 3D models by a Euclidean transformation is a fundamental task at the core of many application in computer vision. This problem is non-convex due to the presence of rotational constraints, making traditional local optimization methods prone to getting stuck in local minima. This paper addresses finding the globally optimal transformation in various 3D registration problems by a unified formulation that integrates common geometric registration modalities (namely point-to-point, point-to-line *and* point-to-plane*). This formulation renders the optimization problem independent of both the number and nature of the correspondences.*

The main novelty of our proposal is the introduction of a strengthened *Lagrangian dual relaxation for this problem, which surpasses previous similar approaches [32] in effectiveness. In fact, even though with* no theoretical guarantees*, exhaustive empirical evaluation in both synthetic and real experiments* always *resulted on a tight relaxation that allowed to recover a guaranteed globally optimal solution by exploiting duality theory.*

Thus, our approach allows for effectively solving the 3D registration with global optimality guarantees while running at a fraction of the time for the state-of-the-art alternative [34], based on a more computationally intensive Branch and Bound method.

1. Introduction

The problem of registering 3D geometric data is a classical problem in numerous fields, including computer vision, robotics, photogrammetry or medical imaging [23, 12, 37]. It seeks the transformation that brings closest together different surfaces according to some *meaningful distance function*. Consider the ubiquituous scenario in which a system (usually a sensor) returns 3D points $\{x_i\}_{i=1}^m$ of an object and a model of the same is available consisting of 3D primitives $\{P_i\}$ (typically points, lines and/or planes) [3, 14, 47, 28]. Assuming the correspondences between the
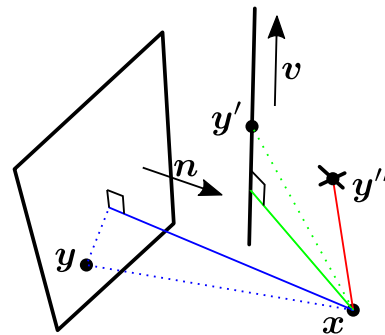

Figure 1. This paper presents a unified formulation for the 3D registration with *point-to-plane*, *point-to-line* and *point-to-point* correspondences, and then provides a certifiable globally optimal solution using Lagrangian duality.

sets are given, $x_i \leftrightarrow P_i$, the general problem reduces to finding the *optimal* roto-translation $T = (R, t) \in \mathrm{SE}(3)$ as

$$T^\star = \arg\min_{T \in \mathrm{SE}(3)} \sum_{i=1}^m \mathrm{d}_{P_i}(T \oplus x_i)^2. \qquad (1)$$

Here $\mathrm{SE}(3)$ stands for the usual Special Euclidean group in 3D, $T \oplus x_i$ denotes the Euclidean transformation of the point x_i and $\mathrm{d}_{P_i}(\cdot)$ is the distance to the primitive P_i [34]. Even with known correspondences, the registration problem (1) is tough to solve in a global fashion due to the non-convexity of the constraints in the rotation $R \in \mathrm{SO}(3)$. A closed-form solution exists only if all the correspondences are point-to-point, as given by Horn *et al.* [24]. For the other cases, most pipelines resort to local approximations, with the inherent risk of getting stuck in local minima [33, 34].

In this paper we address the *global optimization* of the general 3D registration problem (1). After providing a thorough overview of potential alternatives for this task in Section 2, we present in Section 3 a unified formulation integrating *point-to-point*, *point-to-line* and *point-to-plane* correspondences into a single quadratic objective that is a function of the rotation R only. Our main contribution is the development of a novel convex relaxation for this formulation under the usual framework of the Lagrangian dual problem,

which results on a small constant size *semidefinite program* (SDP). This novel relaxation, fully characterized in Section 4, turned tight for *all* the evaluations performed on synthetic and real data, under an extensive range of different problem conditions (Section 5). This allowed us to solve the original non-convex problem in a global fashion using its connection to the convex relaxation through duality theory [4].

In summary, the proposed algorithm Alg. 1 provides an iterative optimization framework for 3D registration problems that, unlike current local iterative approaches, is *able to* provide a certifiable global solution and, in fact, our *empirical* observation is that it always does so.

We remark that this global performance comes without any theoretical guarantees so far. Whereas we acknowledge the importance of the missing formal proof that justifies this behavior, we hope the empirical performance of the proposed relaxation is motivation enough to encourage further exploitation of this approach.

2. Related work

As a result of its widespread interest, geometric registration has been object of extensive research for decades now. The number of works aiming at guaranteing global optimality in registration problems is much smaller, though, and focuses mainly on the case of *point* cloud registration with unknown correspondences [27, 35, 50, 52].

Multimodal registration

Registration across diverse 3D geometric primitives arises in a variety of problems and applications. In the Iterative Closest Point (ICP) framework, since the original proposal using point-to-point correspondences by Besl and McKay [3], different types of correspondences have been introduced for improved performance: *e.g.* point-to-plane [14] or plane-to-plane [45, 46]. Multimodal correspondences play also an important role in the extrinsic calibration of sensors of different nature, such as camera and lidar [51, 49, 16, 18]. Numerous state-of-the-art SLAM (Simultaneous Localization and Mapping) solutions work with models that include plane and line primitives too [2, 47, 17, 43, 28, 19].

In the context of multimodal registration a significant effort has been devoted for decades to the resolution of minimal problems [26], *e.g.* for line-to-plane correspondences [14, 49, 6, 5] and point-to-plane correspondences [21, 39, 38]. These solutions find application mainly in random sample and consensus (RANSAC) frameworks, but they do not face the more general least squares problem (1). Instead, this optimization problem is traditionally addressed through local linear approximations on the rotation, assuming an initialization close to the solution is given [14], which can easily lead to suboptimal local solutions [33].

Global optimization

Next we review the main approaches available in the literature for global optimization [22] and connect them to the registration problem (1) at hand.

Analytical solutions A classical approach involves the computation of all the stationary points (among which there is the global minimum). This approach is used by Censi [13] for the global resolution of the *2D* registration problem, reducing the problem to solving a 4-th order polynomial equation. However, this approach does not generalize well to the 3D case due to the higher complexity of the rotation space, which produces an explosion in the complexity of the resulting polynomial system [29, 30].

Branch and Bound A commonly used tool for NP-hard optimization problems is Branch and Bound (BnB). This is used by Olsson *et al*. [33, 34] for solving the same problem addressed in this paper, yielding a provably global solution. However, the resolution time is notably high (in the order of seconds) due to the exploratory nature and exponential worst-case performance of BnB.

Convex relaxations Convex relaxation techniques consider approximate, simpler versions of the problem whose global optimum is much easier to reach. If the approximation is good, the solution of the relaxed problem may then provide valuable information about the original problem. Thus, the main task is to find an appropriate relaxation. For problems such as ours (1), which are affected by the nonconvexity of rotation constraints, a possible relaxation is the search into the *convex* hull of SO(3). This has proved to give good approximations in rotation synchronization [44], SLAM [41] and 2D/3D registration [25].

Another generic (and successful) tool for providing relaxations of difficult constrained problems is the Lagrangian dual relaxation [4]. This provides particularly good approximations for many problems that can be reformulated as a Quadratically Constrained Quadratic Program (QCQP), where the relaxed problem becomes a Semidefinite Program (SDP) [4, 15]. Some problems involving rotations can be characterized as QCQPs, such as Pose Graph Optimization, for which recent literature applying the Lagrangian dual relaxation has shown impressive results finding globally optimal solutions based solely on convex relaxations [11, 10, 7, 40, 8].

The Lagrangian dual relaxation has been applied before to the QCQP formulation of 3D registration [32]. In this case the approximation can be very good and even provides the global solution in a certain range of problems, but it deteriorates when the noise level increases or the number

of correspondences approaches the minimal cases, that is, when the problem becomes inherently more difficult [33].

Still in the context of the Lagrangian dual relaxation, the relaxation can be *strengthened* (improved) by introducing additional valid constraints [31, Chap. 13]. This has found applicability in the optimization literature [36, 42], and in QCQP problems involving orthonormal constraints, improving the obtained relaxation considerably [1].

To the best of our knowledge, this "trick" has not been considered before for relaxations involving rotations. In this paper, we extend the QCQP formulation of [32] with a whole set of valid quadratic rotational constraints to achieve an improved relaxation that (empirically) returned a globally optimal solution *in all problem instances*, regardless of the noise level or the number of correspondences.

3. Formulation of multimodal registration

In this section we revisit the formulation of the 3D registration problem (1), rewriting it in a suitable form for later applying the Lagrangian dual relaxation.

3.1. Generalized distance function

First, we provide a unified formulation of the distances for the *point-to-point*, *point-to-line* and *point-to-plane* correspondences on the basis of the results presented in [34].

In the registration problem (1), the square distance from a 3D point $\boldsymbol{x}$ to a 3D primitive P minimized in the registration problem (1) is typically that to the closest point in P, defined by

$$\mathrm{d}_P^2(\boldsymbol{x}) = \min_{\boldsymbol{y}' \in P} \|\boldsymbol{x} - \boldsymbol{y}'\|_2^2. \qquad (2)$$

For all the primitives considered here (see Fig. 1) the closest distance problem (2) has a simple closed-form solution that fits a generalized distance function of the form

$$\mathrm{d}_P(\boldsymbol{x})^2 = \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{C}}^2 = (\boldsymbol{x} - \boldsymbol{y})^\top \boldsymbol{C}(\boldsymbol{x} - \boldsymbol{y}), \qquad (3)$$

namely

$$\min_{\boldsymbol{y}' \in P} \|\boldsymbol{x} - \boldsymbol{y}'\|_2^2 = \qquad (4)$$

$$= \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 = \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{I}_3}^2, \qquad \text{(point)}$$

$$= \|(\boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}^\top)(\boldsymbol{x} - \boldsymbol{y})\|_2^2 = \|\boldsymbol{x} - \boldsymbol{y}\|_{(\boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}^\top)}^2, \quad \text{(line)}$$

$$= (\boldsymbol{n}^\top(\boldsymbol{x} - \boldsymbol{y}))^2 = \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{n}\boldsymbol{n}^\top}^2. \qquad \text{(plane)}$$

Here $\boldsymbol{y} \in \mathbb{R}^3$ is any point lying in the primitive[1], $\boldsymbol{v}$ is the unit direction vector for a line, $\boldsymbol{n}_i$ is the unit normal vector for a plane, and $\boldsymbol{C} \in \mathbb{S}^3$ is a symmetric matrix whose expression depends on the primitive as reflected by (4). These results stem from applying elementary algebra to each primitive [34], but we provide the full proof in the **supplementary material** for completeness.

---

[1] For the *point-to-point* case $\boldsymbol{y}_i$ *is the primitive*.

3.2. Quadratic formulation and marginalization

The distances minimized in the registration problem (1) depend on the transformation $\boldsymbol{T}$,

$$\mathrm{d}_{P_i}^2(\boldsymbol{T} \oplus \boldsymbol{x}_i) = (\boldsymbol{T} \oplus \boldsymbol{x}_i - \boldsymbol{y}_i)^\top \boldsymbol{C}_i(\boldsymbol{T} \oplus \boldsymbol{x}_i - \boldsymbol{y}_i). \quad (5)$$

If a matrix representation is chosen for the rotation $\boldsymbol{R}$, the expression of the transformed point $\boldsymbol{T} \oplus \boldsymbol{x}_i$ is linear in the elements of $\boldsymbol{R}$ and $\boldsymbol{t}$:

$$\boldsymbol{T} \oplus \boldsymbol{x}_i = \boldsymbol{R}\boldsymbol{x}_i + \boldsymbol{t} = \underbrace{(\tilde{\boldsymbol{x}}^\top \otimes \boldsymbol{I}_3)}_{\boldsymbol{X}_i} \mathrm{vec}(\boldsymbol{T}), \qquad (6)$$

where $\tilde{\boldsymbol{x}} = [\boldsymbol{x}^\top, 1]^\top$ refers to the homogeneized version of $\boldsymbol{x}$, $\otimes$ is the Kronecker product and $\mathrm{vec}(\boldsymbol{T})$ is the vectorization (applied column-wise) of the transformation matrix,

$$\mathrm{vec}(\boldsymbol{T}) = \begin{bmatrix} \mathrm{vec}(\boldsymbol{R}) \\ \boldsymbol{t} \end{bmatrix}. \qquad (7)$$

Proof. **Supplementary material** □

With this linear parameterization of the transformed point, it is easy to see that the generalized distance to minimize is a *quadratic function* of $\boldsymbol{\tau} = \mathrm{vec}(\boldsymbol{T})$, writable as

$$\mathrm{d}_{P_i}^2(\boldsymbol{T} \oplus \boldsymbol{x}_i) = \tilde{\boldsymbol{\tau}}^\top \underbrace{\boldsymbol{N}_i^\top \boldsymbol{C}_i \boldsymbol{N}_i}_{\tilde{\boldsymbol{M}}_i} \tilde{\boldsymbol{\tau}}, \qquad (8)$$

with $\boldsymbol{N}_i = [\tilde{\boldsymbol{x}}_i^\top \otimes \boldsymbol{I}_3 | -\boldsymbol{y}_i]$ and $\tilde{\boldsymbol{\tau}} = \begin{bmatrix} \mathrm{vec}(\boldsymbol{T}) \\ 1 \end{bmatrix}^\top$. Because of the quadratic nature of the cost (8) it is possible to accumulate the observations, *compressing all the data* into a single 13×13 matrix term $\tilde{\boldsymbol{M}}$:

$$f(\boldsymbol{T}) = \sum_{i=1}^m \mathrm{d}_{P_i}^2(\boldsymbol{T} \oplus \boldsymbol{x}_i) = \tilde{\boldsymbol{\tau}}^\top \underbrace{\left(\sum_{i=1}^m \tilde{\boldsymbol{M}}_i\right)}_{\tilde{\boldsymbol{M}}} \tilde{\boldsymbol{\tau}}. \qquad (9)$$

Thanks to this compression step, the size of the following reformulated problem is independent of m.

When minimizing the quadratic objective $f(\boldsymbol{T})$ in (9), the problem can be further reduced if we apply marginalization on the unconstrained part of the unknown $\boldsymbol{T}$, that is, in the translation $\boldsymbol{t}$. It is well known from previous work that $\boldsymbol{t}$ can be derived in terms of $\boldsymbol{R}$ [24, 33]. In the quadratic formulation this is straightforward:

Lemma 1. *The optimal translation for a fixed $\boldsymbol{R}$ is*

$$\boldsymbol{t}^\star(\boldsymbol{R}) = -\tilde{\boldsymbol{M}}_{\boldsymbol{t},\boldsymbol{t}}^{-1} \tilde{\boldsymbol{M}}_{\boldsymbol{t},!\boldsymbol{t}} \, \tilde{\boldsymbol{r}}, \quad \tilde{\boldsymbol{r}} = \begin{bmatrix} \mathrm{vec}(\boldsymbol{R}) \\ 1 \end{bmatrix}. \qquad (10)$$

Input: List of correspondences $\{x_i \leftrightarrow P_i\}_{i=1}^m$
Output: Global optimum $T^\star$
```
/* build compressed quadratic form        */
```
initialize $\tilde{M} \leftarrow$ zeros(13);
for $(x_i, P_i) \in \{x_i \leftrightarrow P_i\}_{i=1}^m$ **do**
  $\tilde{M}_i \leftarrow$ buildTerm(x_i, P_i) ;   `// depends on` P_i
  $\tilde{M} \leftarrow \tilde{M} + \tilde{M}_i$;
end
```
/* minimize quadratic objective           */
```
marginalize $\tilde{Q} = \tilde{M}/\tilde{M}_{t,t}$;   `// Schur complement`
solve optimal rotation $R^\star$ using Alg. 2;
get optimal translation $t^\star = t^\star(R^\star)$ from (10);
return $T^\star = (R^\star, t^\star)$

Algorithm 1: Globally optimal 3D registration

Here the subindex t stands for the set of indexes correspond-ing to translation variables, whereas !t is its complement. The marginalized *optimization problem is then*

$$f^\star = \min_{R \in SO(3)} \underbrace{\tilde{r}^\top \tilde{Q} \tilde{r}}_{q(\tilde{r})}, \quad \tilde{r} = \begin{bmatrix} \mathrm{vec}(R) \\ 1 \end{bmatrix}, \qquad (11)$$

where the marginalized *quadratic form* $\tilde{Q} = \tilde{M}/\tilde{M}_{t,t}$ *is the* Schur complement *of the block* $\tilde{M}_{t,t}$ *in the matrix* $\tilde{M}$.

Proof. **Supplementary material** □

The registration problem is addressed then following the pipeline depicted in Algorithm 1, where the main complex-ity remains in solving the marginalized problem (11). This is still a non-convex optimization problem in R, for which a convex relaxation is provided in the next session. This re-laxation *empirically* proves to be tight in all the evaluated cases, allowing us to recover a globally optimal solution for the marginalized problem (11).

4. Tight dual relaxation

The key ingredient to solve the non-convex problem (11) lies upon an adequate application of Lagrangian duality. We will apply fundamental results from duality theory [4, 9], so some basic properties and notions regarding the Lagrangian dual problem are provided in the **supplementary material** for completeness. Then we present a specific formulation $(\tilde{\mathcal{P}})$ of the constrained problem (11) in Section 4.1 and its corresponding dual problem $(\mathcal{D})$ in Section 4.2. Finally we show how to recover the globally optimal solution in Sec-tion 4.3, provided that strong duality holds. The experi-ments of Section 5 show that, *empirically*, this relaxation is always tight (strong duality holds), even in extreme condi-tions.

Once the whole dual framework has been developed, its implementation is straightforward and the resolution of the marginalized problem (11) is done following the relatively simple pipeline depicted in Algorithm 2.

4.1. Primal problem

We will address now the task of formulating the opti-mization problem (11) in such a manner that the approach described above produces successful results.

The constraint $R_i \in$ SO(3) in this problem states that the 3×3 rotation matrix R fulfills the *orthonormality* and *determinant* constraints, that is,

$$\mathrm{SO}(3) \equiv \{R \in \mathbb{R}^{3\times3} : R^\top R = I_3, \det(R) = +1\}. \quad (12)$$

In order to apply the Lagrangian dual relaxation, it is particularly appealing to formulate the primal problem as a Quadratically Constrained Quadratic Program (QCPQ). In the usual characterization of SO(3), the orthonormality constraints are all quadratic *but* the determinant constraint is cubic. Because of this, it has been customary in other prob-lems involving rotations to relax the constraints by dropping the determinant constraint $\det(R) = +1$ and keeping only the orthonormality constraints $R^\top R = I_3$, which amounts to performing the optimization in O(3) rather than in SO(3). This approach has provided tight relaxations for other prob-lems [11, 10, 7, 40]. For the registration problem however it works well only in a certain range of problems [32].

Duality strengthening Let us now make a quick anno-tation about an important fact concerning the Lagrangian duality that will be key for the success of our proposal: By construction, every time a new scalar constraint $c_{k+1}(\cdot)$ is introduced into the Lagrangian a new dual variable λ_{k+1} appears and the domain of the dual problem increases its dimension in one. As a result, the bound $d_{k+1}^\star$ provided by the new dual problem is at least as good as that of the pre-vious one, $d_k^\star \leq d_{k+1}^\star \leq f^\star$. As a consequece the dual problem *is not intrinsic* [36, 4]: it depends on the particular formulation of the primal problem. In particular, it depends on the specific characterization of the optimization domain or *feasible region*: Adding appropriate redundant *valid con-straints* has actually shown to be remarkably effective for improving the quality of the dual relaxations in other prob-lems [31, Chap. 13].

Following this idea, our approach is to characterize the feasible set SO(3) with the largest possible amount of *quadratic* constraints. The chosen constraints need to be linearly independent to introduce any potential improve-ment [36]. It is important that we keep the complexity of the constraints quadratic in order to maintain the Lagrangian dual problem simple. For the set of orthogonal matrices O(3) it has been shown [1] that a complete set of quadratic constraints is given by the combination of both column-based and row-based orthogonality constraints: $\{R^\top R = I_3, RR^\top = I_3\}$. However, Tron *et al.* show in [48] that the

Table 1. Table of constraints, Lagrange multipliers and penalizations for Problem ($\tilde{\mathcal{P}}$)

Constraint type	Constraint equation	Dual variable	Penalization term
Orthonormal rows	$y^2 \boldsymbol{I}_3 - \boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{0}$	$\boldsymbol{\Lambda}_\mathrm{r} = \begin{bmatrix} \lambda_1 & \lambda_6 & \lambda_5 \\ \lambda_6 & \lambda_2 & \lambda_4 \\ \lambda_5 & \lambda_4 & \lambda_3 \end{bmatrix} \in \mathbb{S}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_\mathrm{r}(\boldsymbol{\Lambda}_\mathrm{r})\tilde{\boldsymbol{r}}$
Orthonormal columns	$y^2 \boldsymbol{I}_3 - \boldsymbol{R}^\top \boldsymbol{R} = \boldsymbol{0}$	$\boldsymbol{\Lambda}_\mathrm{c} = \begin{bmatrix} \lambda_7 & \lambda_{12} & \lambda_{11} \\ \lambda_{12} & \lambda_8 & \lambda_{10} \\ \lambda_{11} & \lambda_{10} & \lambda_9 \end{bmatrix} \in \mathbb{S}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_\mathrm{c}(\boldsymbol{\Lambda}_\mathrm{c})\tilde{\boldsymbol{r}}$
Handedness	$\boldsymbol{R}^{(1)} \times \boldsymbol{R}^{(2)} - y\boldsymbol{R}^{(3)} = \boldsymbol{0}$ $\boldsymbol{R}^{(2)} \times \boldsymbol{R}^{(3)} - y\boldsymbol{R}^{(1)} = \boldsymbol{0}$ $\boldsymbol{R}^{(3)} \times \boldsymbol{R}^{(1)} - y\boldsymbol{R}^{(2)} = \boldsymbol{0}$	$\boldsymbol{\lambda}_{\mathrm{d}_{123}} = \begin{bmatrix} \lambda_{13} & \lambda_{14} & \lambda_{15} \end{bmatrix}^\top \in \mathbb{R}^3$ $\boldsymbol{\lambda}_{\mathrm{d}_{231}} = \begin{bmatrix} \lambda_{16} & \lambda_{17} & \lambda_{18} \end{bmatrix}^\top \in \mathbb{R}^3$ $\boldsymbol{\lambda}_{\mathrm{d}_{312}} = \begin{bmatrix} \lambda_{19} & \lambda_{20} & \lambda_{21} \end{bmatrix}^\top \in \mathbb{R}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}})\tilde{\boldsymbol{r}}$
Homogeneization	$1 - y^2 = 0$	$\gamma \equiv \lambda_{22} \in \mathbb{R}$	$\gamma + \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_\mathrm{h}(\gamma)\tilde{\boldsymbol{r}}$

(a) Orthonormal rows: $\tilde{\boldsymbol{P}}_\mathrm{r}(\boldsymbol{\Lambda}_\mathrm{r})$ (b) Orthonormal columns: $\tilde{\boldsymbol{P}}_\mathrm{c}(\boldsymbol{\Lambda}_\mathrm{c})$ (c) Handedness: $\tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}})$ (d) Homogeneization: $\tilde{\boldsymbol{P}}_\mathrm{h}(\gamma)$
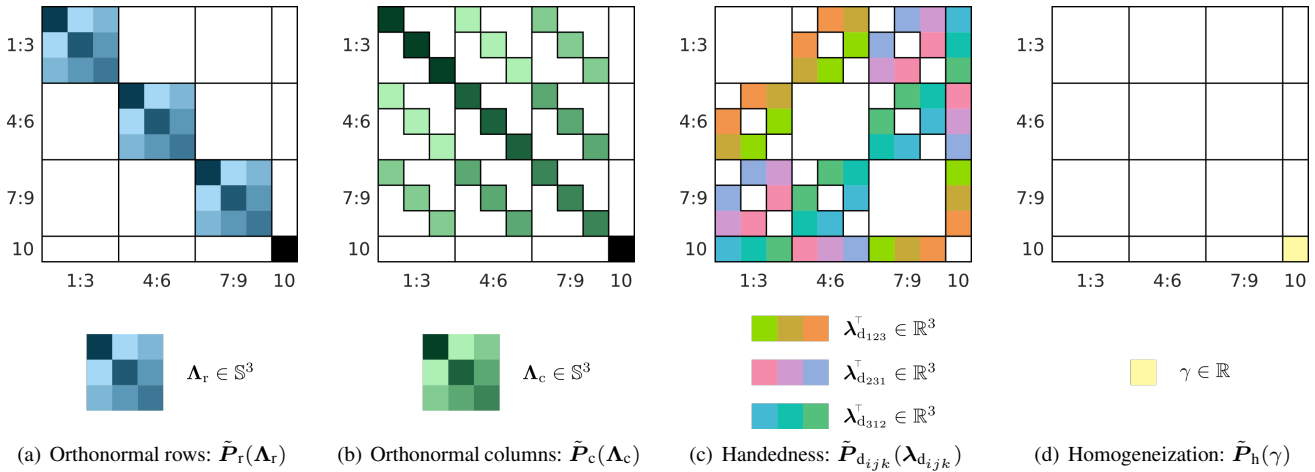
Figure 2. Pattern of the penalization matrices (top row) and dual variables (bottom row) for the different sets of constraints. A coloured cell indicates its value depends (linearly) only on the corresponding dual variable λ_i. Black cells in (a) and (b) stand for values involving a linear combination of several dual variables.

rotation space SO(3) has additional quadratic constraints due to the *handedness* property that forces the positive unit determinant: Since $\boldsymbol{R} \in \mathrm{O}(3) \Rightarrow \det(\boldsymbol{R}) = \pm 1$, provided that $\boldsymbol{R} \in \mathrm{O}(3)$ the positive sign is guaranteed if the matrix columns fulfill the well-known right-hand rule, $\boldsymbol{R}^{(1)} \times \boldsymbol{R}^{(2)} = \boldsymbol{R}^{(3)}$, where $\boldsymbol{R}^{(k)}$ is the k-th column of $\boldsymbol{R}$. Taking any of the three possible cyclic permutations of the column indexes for the right-hand rule provides exactly three independent quadratic constraints. Altogether we end up with $2 \cdot 6 + 3 \cdot 3 = 21$ scalar rotational constraints: 6 for each *symmetric* matrix constraint from orthonormality, and 3 for each vector constraint from the handedness constraint.

Problem homogeneization The optimization objective in (11) as well as the gathered constraints are all quadratic functions but in general not homogeneous. It is very convenient for simplifying the derivation of the dual problem to *homogeneize* the problem by introducing an auxiliary variable y with the constraint $y^2 = 1$ [11, 48]. We define the *equivalent*, *homogeneous*, *strengthened* primal problem:

$$\min_{\boldsymbol{R}} \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{Q}}\tilde{\boldsymbol{r}}, \quad \tilde{\boldsymbol{r}} = \begin{bmatrix} \mathrm{vec}(\boldsymbol{R}) \\ y \end{bmatrix}$$
$$\text{s.t. } \boldsymbol{R}^\top \boldsymbol{R} = y^2 \boldsymbol{I}_3,$$
$$\boldsymbol{R}\boldsymbol{R}^\top = y^2 \boldsymbol{I}_3, \quad\quad\quad (\tilde{\mathcal{P}})$$
$$\boldsymbol{R}^{(i)} \times \boldsymbol{R}^{(j)} = y\boldsymbol{R}^{(k)}, \ i, j, k = \mathrm{cyclic}(1, 2, 3)$$
$$y^2 = 1.$$

4.2. Dual problem

Once the primal problem has been clearly defined, the derivation of the dual problem is a mechanical work, basically reduced to the derivation of the penalization term corresponding to each constraint.

The primal problem ($\tilde{\mathcal{P}}$) is a QCQP, so the Lagrangian is

$$L(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{\lambda}}) = \gamma + \tilde{\boldsymbol{r}}^\top \underbrace{(\tilde{\boldsymbol{Q}} + \tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}}))}_{\tilde{\boldsymbol{Z}}} \tilde{\boldsymbol{r}}, \qquad (13)$$

where the "homogeneous" dual vector $\tilde{\boldsymbol{\lambda}} = \begin{bmatrix} \boldsymbol{\lambda}^\top, \gamma \end{bmatrix}^\top \in \mathbb{R}^{22}$ gathers the dual variables $\boldsymbol{\lambda}$ corresponding to all the rotation constraints (altogether 21) and the dual variable γ from the homogeneization constraint $y^2 = 1$, shown in Tab. 4. The *penalized matrix* $\tilde{\boldsymbol{Z}}$ is the sum of two terms: $\tilde{\boldsymbol{Q}}$ that contains all the data from the original problem, and $\tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}})$ that accumulates all the penalization terms corresponding to the different kinds of constraints:

$$\tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}}) = \tilde{\boldsymbol{P}}_r(\boldsymbol{\Lambda}_r) + \tilde{\boldsymbol{P}}_c(\boldsymbol{\Lambda}_c) + \tilde{\boldsymbol{P}}_d(\{\boldsymbol{\lambda}_{d_{ijk}}\}) + \tilde{\boldsymbol{P}}_h(\gamma).$$

This matrix is (by definition) a linear function of the dual variables, and the pattern of the different matrix components can be seen in Fig. 2. A detailed overview of the construction and formulae for $\tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}})$ is available in the accompanying **supplementary material**.

With this particularly simple expression for the Lagrangian function, the Lagrangian relaxation is an unconstrained problem which can be solved in closed-form as

$$d(\tilde{\boldsymbol{\lambda}}) = \min_{\tilde{\boldsymbol{r}}} L(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{\lambda}}) = \min_{\tilde{\boldsymbol{r}}} \gamma + \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{Z}} \tilde{\boldsymbol{r}} \qquad (14)$$

$$= \begin{cases} \gamma & \text{if } \tilde{\boldsymbol{Z}} \succeq \boldsymbol{0}, \\ -\infty & \text{otherwise.} \end{cases} \qquad (15)$$

The Lagrangian relaxation is unbounded below unless the penalized matrix $\tilde{\boldsymbol{Z}}$ is *positive semidefinite* (PSD). As a result, the *maximization* of the dual objective $d(\tilde{\boldsymbol{\lambda}})$ can be safely restricted to those vectors $\tilde{\boldsymbol{\lambda}}$ preserving the positive semidefiniteness of $\tilde{\boldsymbol{Z}}$. Thus, the *dual problem* corresponding to the *homogeneous* primal problem ($\tilde{\mathcal{P}}$) is a *Semidefinite Program* (SDP):

$$d^\star = \max_{\tilde{\boldsymbol{\lambda}}} \gamma, \quad \text{s.t. } \tilde{\boldsymbol{Z}}(\tilde{\boldsymbol{\lambda}}) = \tilde{\boldsymbol{Q}} + \tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}}) \succeq \boldsymbol{0}. \qquad (\mathcal{D})$$

This problem is *convex* and off-the-shelf specialized solvers exist for it [20].

4.3. Primal-via-dual resolution

In this section we begin by assuming that the duality gap for our primal-dual pair is zero (we will see in the experiments that this assumption always holds in practice). By duality theory [4], $\tilde{\boldsymbol{r}}^\star$ must be a minimizer of the Lagrangian (13) evaluated at $\tilde{\boldsymbol{\lambda}}^\star$,

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda}^\star) \Rightarrow (\tilde{\boldsymbol{r}}^\star)^\top \tilde{\boldsymbol{Z}}^\star \tilde{\boldsymbol{r}}^\star = 0. \qquad (16)$$

Since $\boldsymbol{Z}^\star \succeq \boldsymbol{0}$, this means that the primal optimum $\tilde{\boldsymbol{r}}^\star$ must lie in the nullspace of $\tilde{\boldsymbol{Z}}^\star$:

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda}^\star) \Rightarrow \tilde{\boldsymbol{r}}^\star \in \text{null}(\tilde{\boldsymbol{Z}}^\star). \qquad (17)$$

Input: Marginalized quadratic form $q(\tilde{\boldsymbol{r}})$ (11)
Output: Global optimum $\boldsymbol{R}^\star$ (if no error occurs)
build symbolic penalized matrix $\tilde{\boldsymbol{Z}}(\tilde{\boldsymbol{\lambda}}) \leftarrow \tilde{\boldsymbol{Q}} + \tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}})$;
solve the SDP problem ($\mathcal{D}$) to get $\tilde{\boldsymbol{\lambda}}^\star$ and $\tilde{\boldsymbol{Z}}^\star$; // CVX
get nullspace $\boldsymbol{V} = \text{null}(\tilde{\boldsymbol{Z}}^\star)$;
ASSERT(rank($\boldsymbol{V}$) == 1);
dehomogeneize $\tilde{\boldsymbol{r}}^\star \leftarrow \boldsymbol{V}/\boldsymbol{V}(end)$;
ASSERT($q(\tilde{\boldsymbol{r}}^\star) - d^\star \approx 0$); // strong duality
$\boldsymbol{R}^\star \hookleftarrow \tilde{\boldsymbol{r}}^\star$; // reshape
return $\boldsymbol{R}$

Algorithm 2: Convex global optimizer

If the nullspace has rank 1, the solution $\tilde{\boldsymbol{r}}^\star$ is recovered up to a scale factor. Then, since the solution must also fulfill the original constraints in the primal problem ($\tilde{\mathcal{P}}$), we fully determine the solution by setting $y = 1$, which in practice reduces to *dehomogeneizing* the solution $\tilde{\boldsymbol{r}}^\star$.

Then with the obtained primal solution $\boldsymbol{r}^\star$ we can check that our initial assumption holds and the duality gap is effectively zero, $d^\star = f(\boldsymbol{R}^\star)$. As we will see in the experiments, both conditions $\text{rank}(\text{null}(\tilde{\boldsymbol{Z}}^\star)) = 1$ and $d^\star = f^\star$ were fulfilled for absolutely all the experiments considered, even under the most extreme situations (in terms of noise and number of correspondences).

5. Experiments

In this section we show that in practice the *strong duality assumption* holds in our primal-dual formulation for any considered problem. As a result, using Algorithm 1 it is always possible to recover the primal optimal solution, and we do this at a fraction of the time necessary for the more complex exploratory techniques. The SDP ($\mathcal{D}$) is solved using CVX [20].

We assess the performance of our method, Ours, in both synthetic and real data and compare it to that of two different state-of-the-art approaches for solving the 3D registration problem globally: a provably optimal exploratory approach based on Branch and Bound, BnB [34], and a dual-based approach which provides a relaxation, referred to as Olsson [32].

But first of all, we define the main parameters that will characterize our general registration problem (1) as well as the metrics used in the assessment of performance.

***Effective* number of correspondences** The number of measurements has a notable impact in the complexity of the registration problem. Namely, Olsson and Eriksson [32] justify and illustrate that the problem becomes easier to solve as the number of measurements increase.

Following this intuition, the most difficult registration problems should be those close to minimal cases. In or-

der to measure how close a problem is to being minimal we consider the unifying framework presented by Ramalingam and Taguchi [38]: Point-to-point and point-to-line correspondences can be transformed into equivalent sets of 3 and 2 point-to-plane correspondences, respectively. We define then the *effective number of correspondences*, $\hat{m}$, as the equivalent number of independent point-to-plane correspondences

$$\hat{m} = 3m_{\text{point}} + 2m_L + 1m_\Pi. \tag{18}$$

The minimum value of $\hat{m}$ for which a general 3D registration problem *may* have a unique global minimum is 7 [38].

Geometry Even if the effective number of correspondences is higher than 6, degeneracies and symmetries with multiple global minima may still occur depending on the geometric distribution of the correspondences. These cases are identified as well in [38]. We took care during the evaluation on both synthetic and real data to discard these degenerate configurations where the true global minimum cannot be found from the data only.

Measurement noise This models the quality of the measurements (we consider no outliers). If there is no noise, a simple linear relaxation would provide the global solution. Then, as the noise level σ increases the problem becomes harder to solve [32].

Metrics Several metrics can be used to measure the effectiveness of the relaxation approaches, Ours and Olsson. Due to space issues we choose to show here the most significant metric, which is in our view the optimality ratio, that is, the percentage of cases in which a globally optimal solution was attained. A solution is considered *globally optimal* if the suboptimality gap $\Delta = f - f^\star$ is zero up to numerical precision. The global minimum $f^\star$ can be found from a provably global algorithm such as BnB or, as we will see next, also from our *tight* relaxation.

In order to measure the computational performance of the different methods we also plot the resolution times. In particular, we use *shaded error bars* to display the median values plus the 1st and 3rd quartiles reflecting the distribution of the values.

The statistics shown in the figures were generated from a population of 100 registration problems in each case.

Other interesting metrics, as well as other parameters ranges beyond those displayed in this document are available in the **supplementary material**. These additional results support the same conclusions about the methods reached in this document and previous works [32, 34].
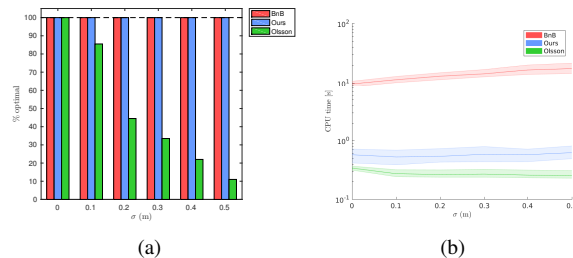

(a)   (b)

Figure 3. Results in the synthetic scenario for increasing level of noise σ, with $\hat{m} = 10$. (a) The effectivenes of the reference relaxation Olsson [32] drops steadily as the level of noise increases. Ours remained optimal in all cases. (b) Both convex relaxations Ours and Olsson take roughly the same time, whereas BnB gets *two orders of magnitude* slower.
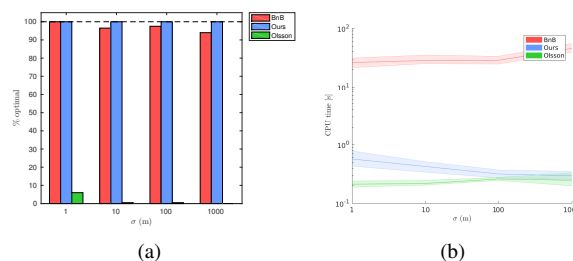

(a)   (b)

Figure 4. Results in the synthetic scenario for extremely severe σ noise conditions and reduced amount of effective number measurements $\hat{m} = 7$. These conditions are of reduced practical interest, but are used to show the general robustness of the tight condition for our relaxation compared to the reference relaxation Olsson [32]. In these regimes BnB suffers of numerical issues.

5.1. Synthetic data

For obtaining synthetic problems we generated a set of random model primitives that added up to $\hat{m}$ effective correspondences. Similarly to [32], each primitive was determined by randomly taking a point inside a sphere of radius 10 m (plus a random unit direction for the case of lines and planes). Then a "measured" point was randomly picked from the set defined by the primitive and we corrupted it with a Gaussian noise of standard deviation σ.

We show the behaviour w.r.t. the noise level in Fig. 3 in a challenging case with $\hat{m} = 10$. The results for a varying value of $\hat{m}$ were similar to those shown for the real data in Fig. 6(a). These and other evaluations for different ranges of parameters are shown in the **supplementary material**. To sum up, our method attained the globally optimal solution in *all* the considered cases, without exception, even in the most severe cases where the number of features $\hat{m}$ remained almost minimal and the measurement noise σ was raised way beyond any expectable value in real scenarios (see Fig. 4). The reference relaxation Olsson in contrast was rarely tight in the challenging scenarios and returned a suboptimal solution.
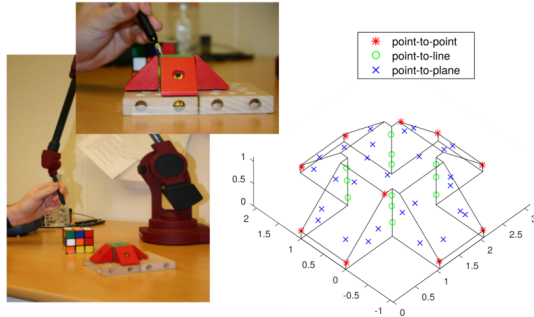
Figure 5. The experimental setup for obtaining real data, taken from [32], and optimal registration of the measurements to the model.

5.2. Real data

Thanks to the corresponding authors, it was possible to exploit the same real data employed in the references [32, 34]. Their experimental setup consisted of using a *MicroScribe-3DLX* 3D scanner to measure the 3D coordinates of some points on the real object, as shown in Fig. 5. For the *Space Station* model 49 points were measured on different primitives of the object, namely 27 on planes, 12 in lines and 10 on corner points. The registration of the complete set to the computer model is shown in 5. In [32] it was shown that for this complete problem Olsson works fine, attaining the global optimum as BnB but in much less time.

We use the same data to generate a more extensive set of challenging real problems. In this case, we can produce significantly more difficult problems by sampling a smaller set of measurements from the data: We choose different combinations of *point-to-point*, *point-to-line* and *point-to-plane* correspondences that result in a particular *effective number of measurements* $\hat{m}$. The precision σ of the measurements in this case is fixed by the sensor, with errors of about 0.5 millimeters according to the authors of the dataset.

The obtained results are consistent with those observed from the evaluation on synthetic data. The optimality ratio is displayed in Fig. 6(a). The behaviour for Olsson w.r.t. the parameter $\hat{m}$ was consistent with that predicted in the original work [32]: It hardly attained global optimality in near-minimal cases, and the performance improved steadily with the increase on $\hat{m}$. Meanwhile, our approach succeeded again in all the cases, always returning the globally optimal solution. Again, both convex relaxations Ours and Olsson took roughly the same time, whereas BnB is *two orders of magnitude* slower (see Fig. 6(b)).

In conclusion, our approach attained the same optimality ratio guaranteed so far only for the provably optimal BnB method, whereas Ours was *two* orders of magnitudes faster than BnB.
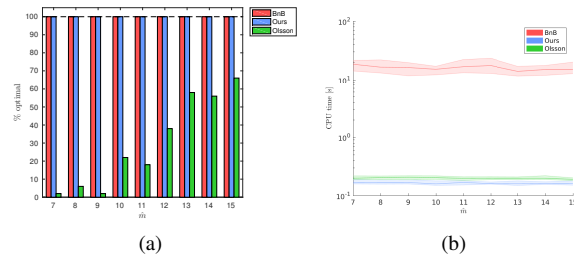

Figure 6. Results in the real scenario for a low effective number of measuremens $\hat{m}$. (a) The effectivenes of the reference relaxation Olsson [32] improves with the increase in $\hat{m}$, yet it hardly attains optimality in most cases whereas ours remained optimal in all cases. (b) Both convex relaxations Ours and Olsson take roughly the same time, whereas BnB is *two orders of magnitude* slower.

6. Conclusions

We have presented a unified formulation for the 3D registration problem involving point-to-point, point-to-line and point-to-plane correspondences that compresses the objective into a single quadratic function of the rotation. Thanks to its generality and flexibility, this formulation should have the potential to introduce further types of correspondences beyond those explored in this work.

The remaining optimization problem has then been characterized as a Quadratically Constrained Quadratic Program. Exploiting a full set of quadratic rotational constraints we obtain a Lagrangian dual relaxation from which a globally optimal solution could be recovered in 100% of the tested cases, although it remains open the theoretical question of why strong duality holds for this relaxation in virtually all cases.

Finally, even though the current approach is already 2 orders of magnitude faster than the competing BnB approach, we are just taking an off-the-shelf generic SDP solver so this performance could be improved further by using specialized solvers that exploit the low-rank structure of the underlying SDP problem.

Acknowledgements

References

[1] K. Anstreicher and H. Wolkowicz. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.*, 22(1):41–55, 2000.

[2] A. Bartoli and P. Sturm. Structure-from-motion using lines: Representation, triangulation, and bundle adjustment. *Comput. Vis. Image Underst.*, 100(3):416–441, 2005.

[3] P. J. Besl and N. D. McKay. Method for registration of 3-D shapes. In *Robot. Tentat.*, pages 586–606. International Society for Optics and Photonics, 1992.

[4] S. Boyd and L. Vandenberghe. *Convex optimization*. 2004.

[5] J. Briales and J. Gonzalez-Jimenez. A Minimal Closed-form Solution for the Perspective Three Orthogonal Angles (P3oA) Problem: Application To Visual Odometry. *J. Math. Imaging Vis.*, pages 1–18, 2015.

[6] J. Briales and J. Gonzalez-Jimenez. A minimal solution for the calibration of a 2D laser-rangefinder and a camera based on scene corners. In *Intell. Robot. Syst. (IROS), 2015 IEEE/RSJ Int. Conf.*, pages 1891–1896, sep 2015.

[7] J. Briales and J. González-Jiménez. Fast Global Optimality Verification in 3D SLAM. In *Int. Conf. Intell. Robot. Syst.* IEEE/RSJ, 2016.

[8] J. Briales and J. González-Jiménez. Initialization of 3D Pose Graph Optimization using Lagrangian duality. In *Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2017.

[9] G. C. Calafiore and L. El Ghaoui. *Optimization Models*. Cambridge university press, 2014.

[10] L. Carlone, G. C. Calafiore, C. Tommolillo, and F. Dellaert. Planar Pose Graph Optimization: Duality, Optimal Solutions, and Verification. *IEEE Trans. Robot.*, 32(3):545–565, 2016.

[11] L. Carlone, D. M. Rosen, G. Calafiore, J. J. Leonard, and F. Dellaert. Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions. In *Intell. Robot. Syst. (IROS), 2015 IEEE/RSJ Int. Conf.*, pages 125–132, sep 2015.

[12] U. Castellani and A. Bartoli. 3d shape registration. In *3D Imaging, Anal. Appl.*, pages 221–264. Springer, 2012.

[13] A. Censi. An ICP variant using a point-to-line metric. *Robot. Autom. 2008. ICRA 2008. IEEE*, 2008.

[14] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image Vis. Comput.*, 10(3):145–155, 1992.

[15] Y. Ding. On efficient semidefinite relaxations for quadratically constrained quadratic programming. 2007.

[16] E. Fernández-Moral, J. González-Jiménez, and V. Arévalo. Extrinsic calibration of 2D laser rangefinders from perpendicular plane observations. *Int. J. Rob. Res.*, 34(11):1401–1417, 2015.

[17] E. Fernández-Moral, W. Mayol-Cuevas, V. Arévalo, and J. González-Jiménez. Fast place recognition with plane-based maps. In *Robot. Autom. (ICRA), 2013 IEEE Int. Conf.*, pages 2719–2724. IEEE, 2013.

[18] R. Gomez-Ojeda, J. Briales, E. Fernández-Moral, and J. González-Jiménez. Extrinsic Calibration of a 2D Laser-Rangefinder and a Camera based on Scene Corners. In *IEEE Int. Conf. Robot. Autom.*, pages 3611–3616. IEEE, 2015.

[19] R. Gomez-ojeda, J. Briales, and J. Gonzalez-jimenez. PL-SVO: Semi-Direct Monocular Visual Odometry by Combining Points and Line Segments.

[20] M. Grant and S. Boyd. {CVX}: Matlab Software for Disciplined Convex Programming, version 2.1. \url{http://cvxr.com/cvx}, 2014.

[21] W. Grimson and T. Lozano-Perez. Model-based recognition and localization from sparse range or tactile data. *Int. J.*, 1984.

[22] R. Hartley and F. Kahl. Optimal Algorithms in Multiview Geometry. In *Comput. Vis. ACCV 2007*, pages 13–34. Springer Berlin Heidelberg, Berlin, Heidelberg.

[23] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003.

[24] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.

[25] Y. Khoo and A. Kapoor. Non-iterative rigid 2D/3D point-set registration using semidefinite programming. *IEEE Trans. Image Process.*, 25(7):2956–2970, 2016.

[26] Z. Kukelova, M. Bujnak, and T. Pajdla. Automatic generator of minimal problem solvers. *Comput. VisionECCV 2008*, 2008.

[27] H. Li and R. Hartley. The 3D-3D registration problem revisited. In *2007 IEEE 11th Int. Conf. Comput. Vis.*, pages 1–8. IEEE, 2007.

[28] L. Ma, C. Kerl, J. Stückler, and D. Cremers. CPA-SLAM: Consistent Plane-Model Alignment for Direct RGB-D SLAM.

[29] F. M. Mirzaei and S. I. Roumeliotis. Globally optimal pose estimation from line correspondences. In *IEEE Int. Conf. Robot. Autom. (ICRA), 2011*, pages 5581–5588. Ieee, may 2011.

[30] F. M. Mirzaei and S. I. Roumeliotis. Optimal estimation of vanishing points in a Manhattan world. In *IEEE Int. Conf. Comput. Vis. (ICCV), 2011*, pages 2454–2461. IEEE, 2011.

[31] Y. Nesterov, H. Wolkowicz, and Y. Ye. Semidefinite programming relaxations of nonconvex quadratic optimization. In *Handb. semidefinite Program.*, pages 361–419. Springer, 2000.

[32] C. Olsson and A. Eriksson. Solving quadratically constrained geometrical problems using lagrangian duality. In *Pattern Recognition, 2008. ICPR 2008. 19th Int. Conf.*, pages 1–5. IEEE, 2008.

[33] C. Olsson, F. Kahl, and M. Oskarsson. The registration problem revisited: Optimal solutions from points, lines and planes. In *Comput. Vis. Pattern Recognition, 2006 IEEE Comput. Soc. Conf.*, volume 1, pages 1206–1213. IEEE, 2006.

[34] C. Olsson, F. Kahl, and M. Oskarsson. Branch-and-Bound Methods for Euclidean Registration Problems. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(5):783–794, 2009.

[35] C. Papazov and D. Burschka. Stochastic global optimization for robust point set registration. *Comput. Vis. Image Underst.*, 115(12):1598–1609, 2011.

[36] P. A. Parrilo and S. Lall. Semidefinite programming relaxations and algebraic optimization in control. *Eur. J. Control*, 9(2):307–321, 2003.

[37] F. Pomerleau, F. Colas, and R. Siegwart. A review of point cloud registration algorithms for mobile robotics. *Found. Trends Robot.*, 4(1):1–104, 2015.

[38] S. Ramalingam and Y. Taguchi. A theory of minimal 3D point to 3D plane registration and its generalization. *Int. J. Comput. Vis.*, 2013.

[39] S. Ramalingam, Y. Taguchi, T. K. Marks, and O. Tuzel. P2Π: A minimal solution for registration of 3D points to 3D planes. In *Eur. Conf. Comput. Vis.*, pages 436–449. Springer, 2010.

[40] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard. A Certifiably Correct Algorithm for Synchronization over the Special Euclidean Group. nov 2016.

[41] D. M. Rosen, C. DuHadway, and J. J. Leonard. A convex relaxation for approximate global optimization in simultaneous localization and mapping. In *Robot. Autom. (ICRA), 2015 IEEE Int. Conf.*, pages 5822–5829. IEEE, 2015.

[42] J. P. Ruiz and I. E. Grossmann. Using redundancy to strengthen the relaxation for the global optimization of MINLP problems. *Comput. Chem. Eng.*, 35(12):2729–2740, 2011.

[43] R. F. Salas-Moreno, B. Glocken, P. H. J. Kelly, and A. J. Davison. Dense planar slam. In *Mix. Augment. Real. (IS-MAR), 2014 IEEE Int. Symp.*, pages 157–164. IEEE, 2014.

[44] J. Saunderson, P. A. Parrilo, and A. S. Willsky. Semidefinite descriptions of the convex hull of rotation matrices. *arXiv Prepr. arXiv1403.4914*, 2014.

[45] A. Segal, D. Haehnel, and S. Thrun. Generalized-ICP. In *Robot. Sci. Syst.*, volume 2, 2009.

[46] J. Serafin and G. Grisetti. NICP : Dense Normal Based Point Cloud Registration. *Prepr. Submitt. to 2015 IEEE Int. Conf. Robot. Autom.*, page 8, 2015.

[47] Y. Taguchi, Y.-D. Jian, S. Ramalingam, and C. Feng. Point-plane SLAM for hand-held 3D sensors. In *Robot. Autom. (ICRA), 2013 IEEE Int. Conf.*, pages 5182–5189. IEEE, 2013.

[48] R. Tron, D. M. Rosen, and L. Carlone. On the Inclusion of Determinant Constraints in Lagrangian Duality for 3D SLAM.

[49] F. Vasconcelos, J. P. Barreto, and U. Nunes. A Minimal Solution for the Extrinsic Calibration of a Camera and a Laser-Rangefinder. *Pattern Anal. Mach. Intell. IEEE Trans.*, 34(11):2097–2107, nov 2012.

[50] J. Yang, H. Li, and Y. Jia. Go-icp: Solving 3d registration efficiently and globally optimally. In *Proc. IEEE Int. Conf. Comput. Vis.*, pages 1457–1464, 2013.

[51] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration). In *IROS*, pages 2301–2306. IEEE, 2004.

[52] Q.-Y. Zhou, J. Park, and V. Koltun. Fast Global Registration. In *Eur. Conf. Comput. Vis.*, pages 766–782. Springer, 2016.

Supplementary material for
Convex Global 3D Registration with Lagrangian Duality

Jesus Briales   Javier Gonzalez-Jimenez

MAPIR-UMA Group

University of Malaga, Spain

`{jesusbriales,javiergonzalez}@uma.es`

Contents

Table 1. Symbols used in the paper

Matrix operators	
$\mathrm{vec}(\cdot)$	Column-wise vectorization
$\otimes$	Matrix Kronecker product
Problem variables	
$\boldsymbol{r} \in \mathbb{R}^9$	Column-wise vectorization of $\boldsymbol{R} \in \mathrm{SO}(3)$
$\boldsymbol{\tau} \in \mathbb{R}^{12}$	Vectorization of $\boldsymbol{T} \in \mathrm{SO}(3) \times \mathbb{R}^3$
Homogeneous variables	
$\tilde{\boldsymbol{a}} \equiv [\boldsymbol{a}; 1]/[\boldsymbol{a}; y]/[\boldsymbol{a}; \gamma]$	Augmented homogeneous vector
$\tilde{\boldsymbol{A}} \equiv \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{b}^\top & c \end{bmatrix}$	Augmented matrix (for hom. coordinate)

(a) Point-to-point distance   (b) Point-to-line distance   (c) Point-to-plane distance
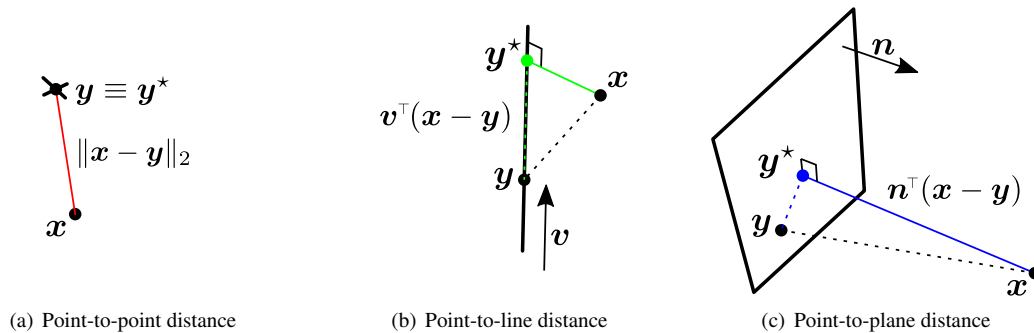
Figure 1. Given a 3D point $\boldsymbol{x}$, the distance to the closest point $\boldsymbol{y}^\star$ in the primitive is readily obtained using simple algebra for the case of a point (a), a line (b) or a plane (c).

1. Generalized distance function

This section provides the expression for the closest distance of a point to the different primitives considered in the paper (see Fig. 1), namely the square distance from a 3D point $\boldsymbol{x}$ to a 3D primitive P is typically defined by

$$\mathrm{d}_P^2(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{y}^\star\|_2^2 = \min_{\boldsymbol{y}' \in P} \|\boldsymbol{x} - \boldsymbol{y}'\|_2^2. \tag{1}$$

The considered primitives have simple closed-form solutions to (1) that are used in each case to characterize the parameters of a generalized distance function of the form

$$\mathrm{d}_P(\boldsymbol{x})^2 = \|\boldsymbol{x} - \boldsymbol{y}\|_C^2 = (\boldsymbol{x} - \boldsymbol{y})^\top \boldsymbol{C}(\boldsymbol{x} - \boldsymbol{y}). \tag{2}$$

A summary is available in Tab. 2.

Table 2. Generalized parameters (2) for each correspondence

Correspondence	$\boldsymbol{y}$	$\boldsymbol{C}$
Point-to-point	The point primitive	$\boldsymbol{I}_3$
Point-to-line	Any point in the line	$\boldsymbol{I}_3 - \boldsymbol{v}\boldsymbol{v}^\top$
Point-to-plane	Any point in the plane	$\boldsymbol{n}\boldsymbol{n}^\top$

1.1. Point-to-point

There is only one point, so the closest point $\boldsymbol{y}^\star$ is the proper primitive (Fig. 1(a)), and this is a trivial case with

$$\mathrm{d}_{\mathrm{point}}^2(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 = (\boldsymbol{x} - \boldsymbol{y})^\top(\boldsymbol{x} - \boldsymbol{y}). \tag{3}$$

This is a particularization of the generalized metric (2) with

$$\boldsymbol{C} = \boldsymbol{I}_3. \tag{4}$$

1.2. Point-to-line

Given any point $\boldsymbol{y}$ in the line and its direction described by the unit vector $\boldsymbol{v}$, the signed distance from $\boldsymbol{y}$ to the point $\boldsymbol{y}^\star$ closest to $\boldsymbol{x}$ is $\boldsymbol{v}^\top(\boldsymbol{x} - \boldsymbol{y})$ (see Fig. 1(b)). As a result the seeked distance can be computed as

$$\mathrm{d}_{\mathrm{line}}^2(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{y}^\star\|_2^2 \tag{5}$$

$$= \|\boldsymbol{x} - (\boldsymbol{y} + (\boldsymbol{v}^\top(\boldsymbol{x} - \boldsymbol{y}))\boldsymbol{v})\|_2^2 \tag{6}$$

$$= \|(\boldsymbol{x} - \boldsymbol{y}) - \boldsymbol{v}\boldsymbol{v}^\top(\boldsymbol{x} - \boldsymbol{y})\|_2^2 \tag{7}$$

$$= \|(\boldsymbol{I}_3 - \boldsymbol{v}\boldsymbol{v}^\top)(\boldsymbol{x} - \boldsymbol{y})\|_2^2. \tag{8}$$

The matrix $\boldsymbol{I} - \boldsymbol{vv}^\top$ is the orthogonal projection onto the line. An orthogonal projection matrix is symmetric,

$$\boldsymbol{I} - \boldsymbol{vv}^\top = (\boldsymbol{I} - \boldsymbol{vv}^\top)^\top \tag{9}$$

and idempotent,

$$(\boldsymbol{I} - \boldsymbol{vv}^\top)^2 = \boldsymbol{I} - \boldsymbol{vv}^\top. \tag{10}$$

As a result, the distance to the line (8) can be refactored as

$$\begin{aligned}
\mathrm{d}_{\text{line}}^2(\boldsymbol{x}) &= (\boldsymbol{x} - \boldsymbol{y})^\top (\boldsymbol{I} - \boldsymbol{vv}^\top)^\top (\boldsymbol{I} - \boldsymbol{vv}^\top)(\boldsymbol{x} - \boldsymbol{y}) \\
&= (\boldsymbol{x} - \boldsymbol{y})^\top (\boldsymbol{I} - \boldsymbol{vv}^\top)(\boldsymbol{x} - \boldsymbol{y}),
\end{aligned} \tag{11}$$

and comparing to the generalized distance (2) this settles for the line case

$$\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{vv}^\top. \tag{12}$$

1.3. Point-to-plane

Given any point $\boldsymbol{y}$ in the plane and its unit normal vector $\boldsymbol{n}$, it is well known from basic geometry that the signed distance from $\boldsymbol{x}$ to the plane is $\boldsymbol{n}^\top(\boldsymbol{x} - \boldsymbol{y})$ (see Fig. 1(c)). The square of this distance yields

$$\begin{aligned}
\mathrm{d}_{\text{plane}}^2(\boldsymbol{x}) &= (\boldsymbol{n}^\top(\boldsymbol{x} - \boldsymbol{y}))_2^2 \tag{13} \\
&= (\boldsymbol{x} - \boldsymbol{y})^\top (\boldsymbol{nn}^\top)(\boldsymbol{x} - \boldsymbol{y}), \tag{14}
\end{aligned}$$

so generalized distance (2) for the plane case is given by

$$\boldsymbol{C} = \boldsymbol{nn}^\top. \tag{15}$$

1.4. Discussion

Note that even though we have used here the 2-norm for the closest distance (1), the considered generalized distance could also cover other extended norms such as $\|\boldsymbol{x} - \boldsymbol{y}^\star\|_A$, with $\boldsymbol{A}$ a positive semidefinite matrix condensing uncertainty weights.

2. Linear transformation operator

The Euclidean transformation of a point $\boldsymbol{x}$ by $\boldsymbol{T}$ into $\boldsymbol{x}'$ is usually referred to symbolically as the operation $\boldsymbol{T} \oplus \boldsymbol{x}$. Depending on the specific parameterization chosen for the transformation $\boldsymbol{T} = (\boldsymbol{R}, \boldsymbol{t})$, this is a non-linear operator w.r.t. $\boldsymbol{T}$. However, if we choose a matrix representation for the rotation $\boldsymbol{R}$, the expression of the transformed point $\boldsymbol{x}'$ is linear in the elements of $\boldsymbol{R}$ and $\boldsymbol{t}$:

$$\boldsymbol{x}' = \boldsymbol{T} \oplus \boldsymbol{x} = \boldsymbol{Rx} + \boldsymbol{t} \tag{16}$$

If we take a different perspective on the operator, this can be seen as the mapping of the transformation $(\boldsymbol{R}, \boldsymbol{t})$ into a point $\boldsymbol{x}'$ according to a fixed $\boldsymbol{x}$:

$$(\cdot) \oplus \boldsymbol{x} : \text{SE}(3) \to \mathbb{R}^3 \tag{17}$$

$$\boldsymbol{T} \oplus \boldsymbol{x} \to \boldsymbol{x}' \tag{18}$$

It will be convenient for the formulation exploited in the paper to write this linear operator in a matrix form with the parameters in the transformation given in vector form:

$$(\cdot) \oplus \boldsymbol{x} : \text{vec}(\text{SE}(3)) \equiv \mathbb{R}^{12} \to \mathbb{R}^3 \tag{19}$$

Using the properties of the vectorization operator $\text{vec}(\cdot)$ [2], namely $\text{vec}(\boldsymbol{AXB}) = (\boldsymbol{B}^\top \otimes \boldsymbol{A})\,\text{vec}(\boldsymbol{X})$, we see that

$$\boldsymbol{Rx} = \text{vec}(\boldsymbol{Rx}) = (\boldsymbol{x}^\top \otimes \boldsymbol{I}_3)\,\text{vec}(\boldsymbol{R}). \tag{20}$$

Thus, with some simple block algebra the linear operator is

$$\boldsymbol{T} \oplus \boldsymbol{x} = \boldsymbol{Rx} + \boldsymbol{t} = \underbrace{\begin{bmatrix} \boldsymbol{x}^\top \otimes \boldsymbol{I}_3 & \boldsymbol{I}_3 \end{bmatrix}}_{\tilde{\boldsymbol{x}}^\top \otimes \boldsymbol{I}_3} \underbrace{\begin{bmatrix} \text{vec}(\boldsymbol{R}) \\ \boldsymbol{t} \end{bmatrix}}_{\boldsymbol{\tau}}. \tag{21}$$

Here $\tilde{\boldsymbol{x}}^\top \otimes \boldsymbol{I}_3$ becomes the matrix defining the linear operator, with $\tilde{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{x}^\top, 1 \end{bmatrix}^\top$ the homogeneous augmented version of $\boldsymbol{x}$, whereas $\boldsymbol{\tau}$ stands for a vectorization of the variables defining the transformation, $\boldsymbol{\tau} \equiv \text{vec}(\boldsymbol{T})$.

3. Marginalization in quadratic forms

This section provides a detailed overview on how the minimization of a quadratic objective can be marginalized w.r.t. the unconstrained variables. In our particular case the problem is

$$f^\star = \min_{\substack{\boldsymbol{R} \in \mathrm{SO}(3) \\ \boldsymbol{t}_i \in \mathbb{R}^3}} \tilde{\boldsymbol{\tau}}^\top \tilde{\boldsymbol{M}} \tilde{\boldsymbol{\tau}}, \quad \tilde{\boldsymbol{\tau}} = \left[\mathrm{vec}(\boldsymbol{R})^\top, \boldsymbol{t}^\top, 1 \right]^\top \tag{22}$$

From the method of Lagrange multipliers, the necessary conditions for a local minimum (and hence for a global mimum $(\boldsymbol{R}^\star, \boldsymbol{t}^\star)$) is that the gradient of the Lagrangian w.r.t. all the parameters vanishes:

$$\nabla \mathrm{L}(\boldsymbol{R}, \boldsymbol{t}, \boldsymbol{\lambda}) = \boldsymbol{0}. \tag{23}$$

For our problem (22) the Lagrangian takes the form

$$\mathrm{L}(\boldsymbol{R}, \boldsymbol{t}, \boldsymbol{\lambda}) = f(\boldsymbol{R}, \boldsymbol{t}) + \boldsymbol{\lambda} \cdot \boldsymbol{c}_{\mathrm{SO}(3)}(\boldsymbol{R}), \tag{24}$$

that is, the constraints in the problem affect only the rotation and are abstracted to some set of functions $\boldsymbol{c}_{\mathrm{SO}(3)}(\boldsymbol{R})$. These constraints are independent of $\boldsymbol{t}$ and taking the gradient w.r.t. $\boldsymbol{t}$ we get

$$\frac{\partial \mathrm{L}(\boldsymbol{R}, \boldsymbol{t}, \boldsymbol{\lambda})}{\partial \boldsymbol{t}} = \frac{\partial f(\tilde{\boldsymbol{\tau}})}{\partial \boldsymbol{t}} = \boldsymbol{0}_{3 \times 1}. \tag{25}$$

Let us reorder the quadratic form as

$$f(\tilde{\boldsymbol{\tau}}) = \begin{bmatrix} !t \\ t \end{bmatrix}^\top \begin{bmatrix} \tilde{\boldsymbol{M}}_{!t,!t} & \tilde{\boldsymbol{M}}_{!t,t} \\ \tilde{\boldsymbol{M}}_{t,!t} & \boldsymbol{M}_{t,t} \end{bmatrix} \begin{bmatrix} !t \\ t \end{bmatrix} \tag{26}$$

$$= !t^\top \tilde{\boldsymbol{M}}_{!t,!t} \, !t + 2t^\top \tilde{\boldsymbol{M}}_{t,!t} \, !t + t^\top \boldsymbol{M}_{t,t} \, t, \tag{27}$$

where $!t$ refers to the complementary set of variables in the quadratic form that are not $\boldsymbol{t}$:

$$!t = \tilde{\boldsymbol{\tau}} \backslash t = \begin{bmatrix} \mathrm{vec}(\boldsymbol{R}) \\ 1 \end{bmatrix} \equiv \tilde{\boldsymbol{r}}. \tag{28}$$

The zero gradient condition w.r.t. $\boldsymbol{t}$ (25) becomes

$$\frac{\partial f(\tilde{\boldsymbol{\tau}})}{\partial \boldsymbol{t}} = 2\tilde{\boldsymbol{M}}_{t,!t} \, \tilde{\boldsymbol{r}} + 2\boldsymbol{M}_{t,t} \, \boldsymbol{t} = \boldsymbol{0}, \tag{29}$$

which is a linear system from which the optimal $\boldsymbol{t}^\star$ can be obtained as

$$\boldsymbol{t}^\star(\tilde{\boldsymbol{r}}) = -\boldsymbol{M}_{t,t}^{-1} \tilde{\boldsymbol{M}}_{t,!t} \, \tilde{\boldsymbol{r}}. \tag{30}$$

Note that if $\boldsymbol{M}_{t,t}$ is not invertible, for any vector $\boldsymbol{u}$ in its nullspace we get $\boldsymbol{M}_{t,t}\boldsymbol{u} = \boldsymbol{0}$. If $\boldsymbol{t}^\star$ is a solution to the linear system (29), so is $\boldsymbol{t}^\star + \boldsymbol{u}$ as well:

$$\tilde{\boldsymbol{M}}_{t,!t} \, \tilde{\boldsymbol{r}} + \boldsymbol{M}_{t,t}(\boldsymbol{t}^\star + \boldsymbol{u}) \tag{31}$$

$$= (\tilde{\boldsymbol{M}}_{t,!t} \, \tilde{\boldsymbol{r}} + \boldsymbol{M}_{t,t} \, \boldsymbol{t}^\star) + \boldsymbol{M}_{t,t} \, \boldsymbol{u} = \boldsymbol{0}. \tag{32}$$

This case would leave us with infinitely many solutions and thus the problem would not be well-posed.

The substitution of $\boldsymbol{t}^\star(\tilde{\boldsymbol{r}})$ into the complete quadratic form (26) leaves us with a reduced quadratic form $q(\tilde{\boldsymbol{r}})$ function of $\tilde{\boldsymbol{r}}$ only:

$$q(\tilde{\boldsymbol{r}}) = \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{Q}} \tilde{\boldsymbol{r}}. \tag{33}$$

The *marginalized* quadratic form matrix $\tilde{\boldsymbol{Q}}$ has the expression

$$\tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{M}}_{!t,!t} - \tilde{\boldsymbol{M}}_{!t,t} \boldsymbol{M}_{t,t}^{-1} \tilde{\boldsymbol{M}}_{t,!t}, \tag{34}$$

which is equivalent to the Schur complement $\tilde{\boldsymbol{M}}/\boldsymbol{M}_{t,t}$ of the complete matrix $\tilde{\boldsymbol{M}}$ w.r.t. the translation diagonal subblock $\boldsymbol{M}_{t,t}$. Thus, after marginalization the *marginalized* optimization problem (35) reduces to finding the optimal value for the remaining variable $\boldsymbol{r}$ so that

$$\boldsymbol{r}^\star = \arg\min_{\boldsymbol{R} \in \mathrm{SO}(3)} \underbrace{\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{Q}} \tilde{\boldsymbol{r}}}_{q(\tilde{\boldsymbol{r}})}, \quad \tilde{\boldsymbol{r}} = \begin{bmatrix} \mathrm{vec}(\boldsymbol{R}) \\ 1 \end{bmatrix}. \tag{35}$$

4. Relaxations and Lagrangian duality

The marginalized problem (35) is difficult only because of the non-convex constraints that apply to the rotation $\boldsymbol{R}$. In fact, if we drop the constraint $\boldsymbol{R} \in \mathrm{SO}(3)$ and perform the optimization in $\boldsymbol{R} \in \mathbb{R}^{3 \times 3}$ the remaining *relaxed* problem

$$f_l = \min_{\boldsymbol{R} \in \mathbb{R}^{3 \times 3}} \tilde{\boldsymbol{r}}^{\top} \tilde{\boldsymbol{Q}} \tilde{\boldsymbol{r}}, \tag{36}$$

is quadratic and straightforward to solve. However, the optimization domain in the unconstrained problem is bigger, $\dim(\mathbb{R}^{3 \times 3}) = 9$ vs. $\dim(\mathrm{SO}(3)) = 3$, and its optimal solution will most probably lie out of $\mathrm{SO}(3)$, that is, it will not be a proper rotation. This "naive" relaxation can be improved if we penalize the violation of constraints in the optimized objective in order to favour solutions closer to the original feasible domain.

The Lagrangian function does this for any general constrained optimization problem of the form

$$f^{\star} = \inf_{\boldsymbol{x}} f(\boldsymbol{x}), \quad \text{s.t. } c_i(\boldsymbol{x}) = 0, \ \forall i \in \mathcal{C}, \tag{37}$$

(where $\mathcal{C}$ is a set indexing the constraints $c_i(\boldsymbol{x})$) by penalizing the original optimization objective with a weighted sum of all the constraints:

$$\mathrm{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i \in \mathcal{C}} \lambda_i c_i(\boldsymbol{x}). \tag{38}$$

Note the optimum $\boldsymbol{x}^{\star}$ for the original problem (37) fulfills all the constraints $c_i(\boldsymbol{x}^{\star}) = 0$, so $\mathrm{L}(\boldsymbol{x}^{\star}, \boldsymbol{\lambda}) = f^{\star}$ for any $\boldsymbol{\lambda}$. The unconstrained problem

$$d(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}} \mathrm{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \leq f^{\star} \tag{39}$$

provides a family of relaxations parameterized by the Lagrange multipliers $\boldsymbol{\lambda}$.

The key insight of Lagrangian duality is that, by choosing the parameter $\boldsymbol{\lambda}$ appropiately, we can obtain better relaxations ($d(\boldsymbol{\lambda})$ closer to $f^{\star}$). The (Lagrangian) *dual problem* seeks the best possible relaxation by choosing the weights λ_i, henceforth *dual variables*, that maximize the lower bound obtained from the Lagrangian relaxation:

$$d^{\star} = \max_{\boldsymbol{\lambda}} d(\boldsymbol{\lambda}). \tag{40}$$

For consistency with this terminology, the original problem, its unknowns and objective are referred to as *primal*. From the definitions above it is easy to prove that

$$d^{\star} \leq f^{\star}. \tag{41}$$

This fundamental result, referred to as *weak (Lagrangian) duality*, holds for any optimization problem and the positive value $f^{\star} - d^{\star} \geq 0$ is called the *duality gap*. If the duality gap is zero, that is, $d^{\star} = f^{\star}$, we say that there is *strong duality* and the relaxation is *tight*.

The dual problem has very appealing properties [1, Sec. 5.2]:

1. It is *always* convex, so it can be solved globally using local search techniques, which are inherently faster that global search methods.

2. If there is *strong duality*, the primal optimum $\boldsymbol{x}^{\star}$ is also the minimizer of the Lagrangian evaluated at the dual optimum:

$$\boldsymbol{x}^{\star} = \arg \min_{\boldsymbol{x}} \mathrm{L}(\boldsymbol{x}, \boldsymbol{\lambda}^{\star}). \tag{42}$$

This last property is highly useful, as in many cases these constraints suffice to fully recover the primal optimum $\boldsymbol{x}^{\star}$ in terms of the dual optimum $\boldsymbol{\lambda}^{\star}$. As a result, if there is strong duality we can avoid the non-convexity of the general primal problem and solve instead the dual problem, then recover the primal solution. This settles a very interesting course of action for global optimization.

5. Lagrangian derivation

This section shows the construction of the Lagrangian function for the *equivalent*, *homogeneous*, *strengthened* primal problem:

$$\min_{\boldsymbol{R},y} \underbrace{\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{Q}} \tilde{\boldsymbol{r}}}_{q(\tilde{\boldsymbol{r}})}, \quad \tilde{\boldsymbol{r}} = \begin{bmatrix} \operatorname{vec}(\boldsymbol{R}) \\ y \end{bmatrix} \tag{$\tilde{\mathcal{P}}$}$$

$$\text{s.t. } \boldsymbol{R}^\top \boldsymbol{R} = y^2 \boldsymbol{I}_3, \tag{43}$$

$$\boldsymbol{R}\boldsymbol{R}^\top = y^2 \boldsymbol{I}_3, \tag{44}$$

$$\boldsymbol{R}^{(i)} \times \boldsymbol{R}^{(j)} = y\boldsymbol{R}^{(k)}, \ (i,j,k) \in (123), \tag{45}$$

$$y^2 = 1. \tag{46}$$

We use the notation (123) to refer to the set of cyclic permutations: $(123) = \{(1,2,3),(2,3,1),(3,1,2)\}$. The Lagrangian of the primal problem ($\tilde{\mathcal{P}}$) is then built by introducing the equality constraints as penalization terms

$$\begin{align}
\mathrm{L}(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{\lambda}}) = q(\tilde{\boldsymbol{r}}) && \text{[objective]} \tag{47} \\
+ \langle \boldsymbol{\Lambda}_\mathrm{r}, y^2 \boldsymbol{I}_3 - \boldsymbol{R}\boldsymbol{R}^\top \rangle && \text{[orthonormal rows]} \tag{48} \\
+ \langle \boldsymbol{\Lambda}_\mathrm{c}, y^2 \boldsymbol{I}_3 - \boldsymbol{R}^\top \boldsymbol{R} \rangle && \text{[orthonormal columns]} \tag{49} \\
+ \sum_{(i,j,k)\in(123)} \langle \boldsymbol{\lambda}_{\mathrm{d}_{ijk}}, \boldsymbol{R}^{(i)} \times \boldsymbol{R}^{(j)} - y\boldsymbol{R}^{(k)} \rangle && \text{[handedness]} \tag{50} \\
+ \langle \gamma, 1 - y^2 \rangle. && \text{[homogeneization]} \tag{51}
\end{align}$$

Every penalization term is written as the inner product of a Lagrange multiplier with a set of constraints. If the constraint is scalar (51) the inner product is a simple scalar product $\langle a, b \rangle = ab$. For vector constraints (50) it becomes the dot product for vectors, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \boldsymbol{a}^\top \boldsymbol{b}$. Finally if the constraints form a matrix (48)(49), we use the inner product for matrices $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \operatorname{tr}(\boldsymbol{A}^\top \boldsymbol{B})$.

Since the constraints in the primal problem ($\tilde{\mathcal{P}}$) are (homogeneous) quadratic functions of $\boldsymbol{R}$, a linear combination of them will still be quadratic w.r.t. $\boldsymbol{R}$ and the Lagrangian can be refactored into

$$\mathrm{L}(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{\lambda}}) = \gamma + \tilde{\boldsymbol{r}}^\top (\tilde{\boldsymbol{Q}} + \tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}})) \tilde{\boldsymbol{r}}. \tag{52}$$

The offset γ originates from the homogeneization constraint (51), the only constraint with a non-homogeneous term. The augmented vector of dual variables $\tilde{\boldsymbol{\lambda}} \in \mathbb{R}^{22}$ is formed from the concatenation of all the dual variables, and the *penalization matrix* $\tilde{\boldsymbol{P}}$ is the sum of the different components coming from the constraints in ($\tilde{\mathcal{P}}$), namely orthonormality of rows (r) (48), orthonormality of columns (c) (49), positive determinant or handedness (d) (50) and homogeneization (h) (51):

$$\tilde{\boldsymbol{P}}(\tilde{\boldsymbol{\lambda}}) = \tilde{\boldsymbol{P}}_\mathrm{r}(\boldsymbol{\Lambda}_\mathrm{r}) + \tilde{\boldsymbol{P}}_\mathrm{c}(\boldsymbol{\Lambda}_\mathrm{c}) + \tilde{\boldsymbol{P}}_d(\underbrace{\{\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}\}}_{\boldsymbol{\lambda}_\mathrm{d}}) + \tilde{\boldsymbol{P}}_\mathrm{h}(\gamma). \tag{53}$$

Next we will fully characterize these penalization terms. A complete summary of the constraints, associated dual variables and penalizations can be seen in Tab. 5. A visualization of the pattern for the resulting matrices $\tilde{\boldsymbol{P}}_{(\cdot)}$ is available in Fig. 2.

5.1. Orthonormality of rotation rows (48)

The matrix constraint for the orthonormality of rows, $\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{I}_3$, provides the penalization term

$$\langle \boldsymbol{\Lambda}_\mathrm{r}, y^2 \boldsymbol{I}_3 - \boldsymbol{R}\boldsymbol{R}^\top \rangle = \operatorname{tr}(\boldsymbol{\Lambda}_\mathrm{r}^\top (y^2 \boldsymbol{I}_3 - \boldsymbol{R}\boldsymbol{R}^\top)) \tag{54}$$

$$[\text{using } \boldsymbol{\Lambda}_\mathrm{r} = \boldsymbol{\Lambda}_\mathrm{r}^\top \text{ and cyclic property of } \operatorname{tr}(\cdot)] = y^2 \operatorname{tr}(\boldsymbol{\Lambda}_\mathrm{r}) - \operatorname{tr}(\boldsymbol{R}^\top \boldsymbol{\Lambda}_\mathrm{r} \boldsymbol{R}) \tag{55}$$

$$[\text{using } \operatorname{tr}(\boldsymbol{I}\boldsymbol{X}^\top \boldsymbol{B}\boldsymbol{Y}) \text{ (81)}] = y^2 \operatorname{tr}(\boldsymbol{\Lambda}_\mathrm{r}) - \boldsymbol{r}^\top (\boldsymbol{I}_3 \otimes \boldsymbol{\Lambda}_\mathrm{r}) \boldsymbol{r} \tag{56}$$

$$[\text{reordering into block matrix}] = \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_\mathrm{r}(\boldsymbol{\Lambda}_\mathrm{r}) \tilde{\boldsymbol{r}}. \tag{57}$$

Table 3. Table of constraints and penalizations for $(\tilde{\mathcal{P}})$

Constraint type	Constraint equation	Dual variable	Penalization term
Orthonormal rows	$y^2 \boldsymbol{I}_3 - \boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{0}$	$\boldsymbol{\Lambda}_{\mathrm{r}} = \begin{bmatrix} \lambda_1 & \lambda_6 & \lambda_5 \\ \lambda_6 & \lambda_2 & \lambda_4 \\ \lambda_5 & \lambda_4 & \lambda_3 \end{bmatrix} \in \mathbb{S}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{r}}(\boldsymbol{\Lambda}_{\mathrm{r}})\tilde{\boldsymbol{r}}$, see (58)
Orthonormal columns	$y^2 \boldsymbol{I}_3 - \boldsymbol{R}^\top \boldsymbol{R} = \boldsymbol{0}$	$\boldsymbol{\Lambda}_{\mathrm{c}} = \begin{bmatrix} \lambda_7 & \lambda_{12} & \lambda_{11} \\ \lambda_{12} & \lambda_8 & \lambda_{10} \\ \lambda_{11} & \lambda_{10} & \lambda_9 \end{bmatrix} \in \mathbb{S}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{c}}(\boldsymbol{\Lambda}_{\mathrm{c}})\tilde{\boldsymbol{r}}$, see (63)
Handedness	$\boldsymbol{R}^{(1)} \times \boldsymbol{R}^{(2)} - y\boldsymbol{R}^{(3)} = \boldsymbol{0}$ $\boldsymbol{R}^{(2)} \times \boldsymbol{R}^{(3)} - y\boldsymbol{R}^{(1)} = \boldsymbol{0}$ $\boldsymbol{R}^{(3)} \times \boldsymbol{R}^{(1)} - y\boldsymbol{R}^{(2)} = \boldsymbol{0}$	$\boldsymbol{\lambda}_{\mathrm{d}_{123}} = \begin{bmatrix} \lambda_{13} & \lambda_{14} & \lambda_{15} \end{bmatrix}^\top \in \mathbb{R}^3$ $\boldsymbol{\lambda}_{\mathrm{d}_{231}} = \begin{bmatrix} \lambda_{16} & \lambda_{17} & \lambda_{18} \end{bmatrix}^\top \in \mathbb{R}^3$ $\boldsymbol{\lambda}_{\mathrm{d}_{312}} = \begin{bmatrix} \lambda_{19} & \lambda_{20} & \lambda_{21} \end{bmatrix}^\top \in \mathbb{R}^3$	$\tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}})\tilde{\boldsymbol{r}}$, see (74)
Homogeneization	$1 - y^2 = 0$	$\gamma \equiv \lambda_{22} \in \mathbb{R}$	$\gamma + \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{h}}(\gamma)\tilde{\boldsymbol{r}}$, see (77)

(a) Orthonormal rows, see (58) (b) Orthonormal columns, see (63) (c) Handedness, see (74) (d) Homogeneization, see (77)
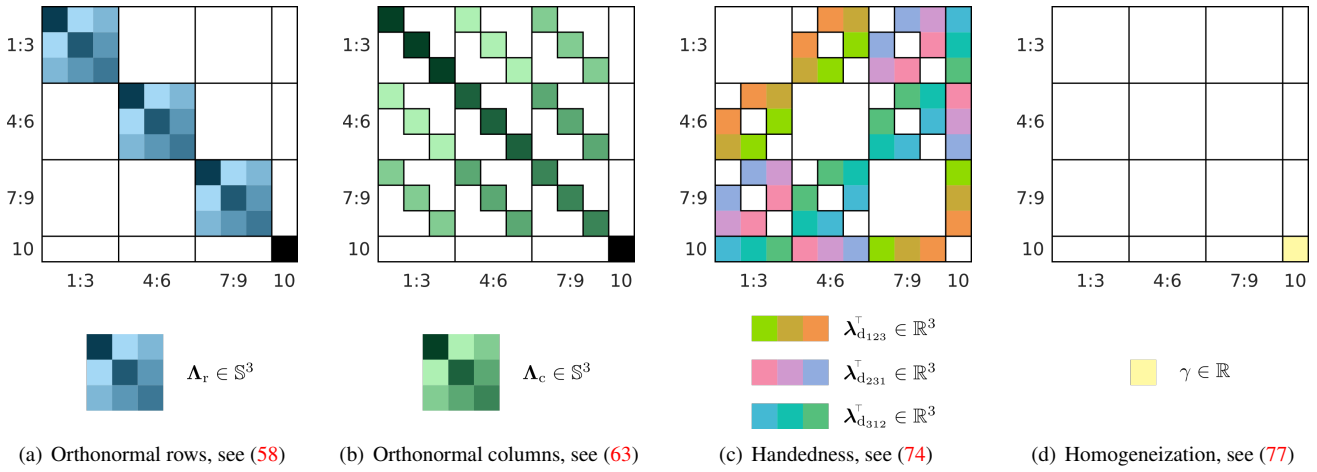
Figure 2. Sparsity pattern of the penalization matrices (top row) and dual variables (bottom row) due to the different sets of constraints. A coloured cell indicates its value depends (linearly) only on the corresponding dual variable λ_i. Black cells (in $(10, 10)$) stand for values involving a linear combination of several dual variables.

The matrix Lagrange multiplier $\boldsymbol{\Lambda}_{\mathrm{r}}$ is symmetric because only 6 effective different scalar constraints stand here. The quadratic penalization matrix $\tilde{\boldsymbol{P}}_{\mathrm{r}}$ is shown in Fig. 2(a) and takes the expression

$$\tilde{\boldsymbol{P}}_{\mathrm{r}}(\boldsymbol{\Lambda}_{\mathrm{r}}) = \left[\begin{array}{c|c} -\boldsymbol{I}_3 \otimes \boldsymbol{\Lambda}_{\mathrm{r}} & \boldsymbol{0}_{9\times 1} \\ \hline \boldsymbol{0}_{1\times 9} & \mathrm{tr}(\boldsymbol{\Lambda}_{\mathrm{r}}) \end{array} \right]. \tag{58}$$

5.2. Orthonormality of rotation columns (49)

Similar to the previous orthonormality constraint but with exchanged transposes, $\boldsymbol{R}^\top \boldsymbol{R} = \boldsymbol{I}_3$, the orthonormality of rotation columns provides the penalization

$$\langle \boldsymbol{\Lambda}_{\mathrm{c}}, y^2 \boldsymbol{I}_3 - \boldsymbol{R}^\top \boldsymbol{R} \rangle = \mathrm{tr}(\boldsymbol{\Lambda}_{\mathrm{c}}^\top (y^2 \boldsymbol{I}_3 - \boldsymbol{R}^\top \boldsymbol{R})) \tag{59}$$

$$= y^2 \mathrm{tr}(\boldsymbol{\Lambda}_{\mathrm{c}}) - \mathrm{tr}(\boldsymbol{\Lambda}_{\mathrm{c}}^\top \boldsymbol{R}^\top \boldsymbol{R}) \tag{60}$$

$$[\text{using } \mathrm{tr}(\boldsymbol{A}^\top \boldsymbol{X}^\top \boldsymbol{I}\boldsymbol{Y})\ (81)] = y^2 \mathrm{tr}(\boldsymbol{\Lambda}_{\mathrm{c}}) - \boldsymbol{r}^\top (\boldsymbol{\Lambda}_{\mathrm{c}} \otimes \boldsymbol{I}_3)\boldsymbol{r} \tag{61}$$

$$[\text{reordering into block matrix}] = \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{r}}(\boldsymbol{\Lambda}_{\mathrm{c}})\tilde{\boldsymbol{r}}. \tag{62}$$

Again, there are only 6 different constraints, so $\mathbf{\Lambda}_{\mathrm{c}} \in \mathbb{S}^3$. The quadratic penalization matrix $\tilde{\boldsymbol{P}}_{\mathrm{c}}$ is shown in Fig. 2(b) and takes the expression

$$\tilde{\boldsymbol{P}}_{\mathrm{c}}(\mathbf{\Lambda}_{\mathrm{c}}) = \left[\begin{array}{c|c} -\mathbf{\Lambda}_{\mathrm{c}} \otimes \boldsymbol{I}_3 & \mathbf{0}_{9\times1} \\ \hline \mathbf{0}_{1\times9} & \mathrm{tr}(\mathbf{\Lambda}_{\mathrm{c}}) \end{array} \right]. \tag{63}$$

5.3. Determinant constraints (50)

The right-hand rule applied to the columns of the rotation matrix $\boldsymbol{R}$ takes the form

$$\boldsymbol{R}^{(i)} \times \boldsymbol{R}^{(j)} = y\boldsymbol{R}^{(k)}, \tag{64}$$

where $\boldsymbol{R}^{(u)}$ refers to the u-th column in the matrix $\boldsymbol{R}$. Let us introduce the *canonical vector* $\boldsymbol{e}_u \in \mathbb{R}^3$, defined as the vector that is zero everywhere except for the u-th entry which is 1. It will be useful as well to define the *canonical matrix* $\boldsymbol{e}_{uv}$ as the matrix which has entry 1 at the position (u, v) and is zero everywhere else. Note $\boldsymbol{e}_{uv} = \boldsymbol{e}_u \boldsymbol{e}_v^\top$. Then the handedness constraint can be written in a more convenient form as

$$\langle \boldsymbol{\lambda}_{\mathrm{d}_{ijk}}, \boldsymbol{R}^{(i)} \times \boldsymbol{R}^{(j)} - y\boldsymbol{R}^{(k)} \rangle = \boldsymbol{\lambda}_{\mathrm{d}_{ijk}}^\top ((\boldsymbol{R}\boldsymbol{e}_i) \times (\boldsymbol{R}\boldsymbol{e}_j) - y(\boldsymbol{R}\boldsymbol{e}_k)) = \tag{65}$$

$$[\text{using } \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = -\boldsymbol{b} \cdot (\boldsymbol{a} \times \boldsymbol{c})] = -(\boldsymbol{R}\boldsymbol{e}_i)^\top (\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \times (\boldsymbol{R}\boldsymbol{e}_j)) - y\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}^\top (\boldsymbol{R}\boldsymbol{e}_k) \tag{66}$$

$$[\text{using } \boldsymbol{a} \times \boldsymbol{b} = [\boldsymbol{a}]_\times \boldsymbol{b}] = -\boldsymbol{e}_i^\top \boldsymbol{R}^\top \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times \boldsymbol{R}\boldsymbol{e}_j - y\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}^\top \boldsymbol{R}\boldsymbol{e}_k \tag{67}$$

$$[\text{using } \mathrm{tr}(a) = a, a \in \mathbb{R}] = -\mathrm{tr}(\boldsymbol{e}_i^\top \boldsymbol{R}^\top \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times \boldsymbol{R}\boldsymbol{e}_j) - y\,\mathrm{tr}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}^\top \boldsymbol{R}\boldsymbol{e}_k) \tag{68}$$

$$[\text{using cyclic property of } \mathrm{tr}(\cdot)] = -\mathrm{tr}(\boldsymbol{e}_j\boldsymbol{e}_i^\top \boldsymbol{R}^\top \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times \boldsymbol{R}) - y\,\mathrm{tr}(\boldsymbol{e}_k\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}^\top \boldsymbol{R}) \tag{69}$$

$$[\text{reordering transposes}] = -\mathrm{tr}(\boldsymbol{e}_{ij}^\top \boldsymbol{R}^\top \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times \boldsymbol{R}) - y\,\mathrm{tr}((\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}\boldsymbol{e}_k^\top)^\top \boldsymbol{R}) \tag{70}$$

$$[\text{using } \mathrm{tr}(\boldsymbol{A}^\top \boldsymbol{X}^\top \boldsymbol{I}\boldsymbol{Y}) \text{ (81) and } \mathrm{tr}(\boldsymbol{A}^\top \boldsymbol{B}) \text{ (80)}] = -\boldsymbol{r}^\top (\boldsymbol{e}_{ij} \otimes \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times)\boldsymbol{r} - y\,\mathrm{vec}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}\boldsymbol{e}_k^\top)^\top \boldsymbol{r} \tag{71}$$

$$[\text{using } \mathrm{vec}(\boldsymbol{ab}^\top) \text{ (79)}] = -\boldsymbol{r}^\top (\boldsymbol{e}_{ij} \otimes \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times)\boldsymbol{r} - y(\boldsymbol{e}_k \otimes \boldsymbol{\lambda}_{\mathrm{d}_{ijk}})^\top \boldsymbol{r} \tag{72}$$

$$[\text{reordering into block matrix}] = \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}}) \tilde{\boldsymbol{r}}. \tag{73}$$

The (symmetric) quadratic penalization matrix $\tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}}(\boldsymbol{\lambda}_{\mathrm{d}_{ijk}})$ takes the expression

$$\tilde{\boldsymbol{P}}_{\mathrm{d}_{ijk}} = \left[\begin{array}{c|c} -\boldsymbol{e}_{ij} \otimes \left[\boldsymbol{\lambda}_{\mathrm{d}_{ijk}} \right]_\times & -(\boldsymbol{e}_k \otimes \boldsymbol{\lambda}_{\mathrm{d}_{ijk}}) \\ \hline \mathbf{0}_{1\times9} & 0 \end{array} \right]. \tag{74}$$

Note this expression is generic for $(i, j, k) \in (123) = \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}$. The complete penalization matrix $\tilde{\boldsymbol{P}}_d(\boldsymbol{\lambda}_\mathrm{d}) = \tilde{\boldsymbol{P}}_{\mathrm{d}_{123}}(\boldsymbol{\lambda}_{\mathrm{d}_{123}}) + \tilde{\boldsymbol{P}}_{\mathrm{d}_{231}}(\boldsymbol{\lambda}_{\mathrm{d}_{231}}) + \tilde{\boldsymbol{P}}_{\mathrm{d}_{312}}(\boldsymbol{\lambda}_{\mathrm{d}_{312}})$ from all the determinant constraints, after taking an equivalent symmetric matrix for the representation, is shown in Fig. 2(c).

5.4. Homogeneization constraint (51)

Finally, the simple scalar constraint $y^2 = 1$ yields

$$\langle \gamma, 1 - y^2 \rangle = \gamma - \gamma y^2 \tag{75}$$

$$= \gamma + \tilde{\boldsymbol{r}}^\top \tilde{\boldsymbol{P}}_{\mathrm{h}}(\gamma)\tilde{\boldsymbol{r}} \tag{76}$$

For compatibility with the rest of quadratic expressions this has been written in matrix form too, where $\tilde{\boldsymbol{P}}_{\mathrm{h}}(\gamma) = -\gamma\boldsymbol{e}_{10,10}$ is the matrix with all zeros but the lower right corner, where it contains the value $-\gamma$:

$$\tilde{\boldsymbol{P}}_{\mathrm{h}}(\gamma) = \left[\begin{array}{c|c} \mathbf{0}_{9\times9} & \mathbf{0}_{9\times1} \\ \hline \mathbf{0}_{1\times9} & -\gamma \end{array} \right]. \tag{77}$$

6. Additional experimental results

Due to the lack of space, only the most relevant experimental results and metrics were shown in the paper. In this section we include all the obtained results comparing the `BnB` solution [4], and the dual-based relaxations, `Ours` and `Olsson` [3]. The results support the claims in the main document.

Evaluation metrics

We used several metrics to evaluate the performance of the methods. Next we describe them and give some notes on the representation used for the results. All the statistics were obtained from 200 samples.

Suboptimality gap The suboptimality gap is defined as $\Delta = f - f^\star$, where $f^\star$ is the minimum objective value attained in the global solution. Since our solution always provided a globally optimal solution (in a certifiable way, since $f = d^\star \Rightarrow f = f^\star$), it is possible to plot this value for all the methods giving a clear reference for the performance: A solution is optimal if $\Delta = 0$, suboptimal otherwise.

We displayed the suboptimality gap in the oncoming figures using boxplots. To show as much information as possible, we also superimposed a sorted set of points displaying the original values underlying each boxplot. Note that for the globally optimal methods, `Ours` and `BnB`, the boxplots and points degenerate to a single line in $\Delta = 0$.

In order to cope with the high variability of the suboptimality values for the different methods, we applied a custom scale in the Y axis, where the $\log(1 + x)$ of the values is displayed in order to better visualize different orders of magnitude while keeping the lower threshold $\Delta = 0$ in the origin.

Global optimality ratio The global optimality ratio provides a refined view of suboptimality metric, where we directly show a bar plot with the percentage of cases in which a globally optimal solution ($\Delta = 0$) was attained.

Time The computational performance of the different algorithms is evaluated by measuring the CPU time necessary to solve each problem. Note these times are just for orientation, as the implementations were not highly optimized. Nevertheless, the conclusions when comparing the different methods remain valid. The obtained values are represented with a shaded error bar plot displaying the median values plus a band defined by the 1st and 3rd quartiles to reflect the distribution of the values. Note that we always used a logarithmic scale for the time values, in order to better compare all the considered methods.

Problem parameters

As discussed in the paper, the main problem parameters defining the overall difficulty of solving the optimization problem globally are the *effective number of correspondences* $\hat{m}$ and the *level of the noise* σ affecting the measurements. A more detailed explanation of these is available in the main document.

6.1. Evaluation on synthetic data

6.1.1   Challenging range of $\hat{m}$

The evaluation in a challenging range of $\hat{m}$ values, going from the almost minimal case $\hat{m} = 7$ up to $\hat{m} = 15$, shows that as the number of correspondences increases the level of difficulty of the problem decreases (see Fig. 3). This is consistent with the claim in [3]. This explains why the looser relaxation `Olsson` has better performance as $\hat{m}$ grows. This trend holds for different levels of noise σ, although as expected the general difficulty keeps higher when the noise is larger, making the performance of `Olsson` worse. In all cases, as the provably optimal method `BnB`, our method provided a globally optimal solution. However, the computational burden for the `BnB` method was notably higher, and in the *easier* scenario ($\sigma = 0.1$) it increased with the decrease in the number of correspondences. This is consistent with the remarks in [4], where the near-minimal cases were tagged as harder to explore in the Branch and Bound framework. The time for the dual-based algorithms (`Ours` and `Olsson`) remained practically constant in all cases and was notably lower than for `BnB` (around 2 orders of magnitude).

(a) Suboptimality gap, $\sigma = 0.1$ m

(b) % of optimality, $\sigma = 0.1$ m

(c) Times, $\sigma = 0.1$ m

(d) Suboptimality gap, $\sigma = 0.5$ m

(e) % of optimality, $\sigma = 0.5$ m

(f) Times, $\sigma = 0.5$ m

(g) Suboptimality gap, $\sigma = 1$ m

(h) % of optimality, $\sigma = 1$ m
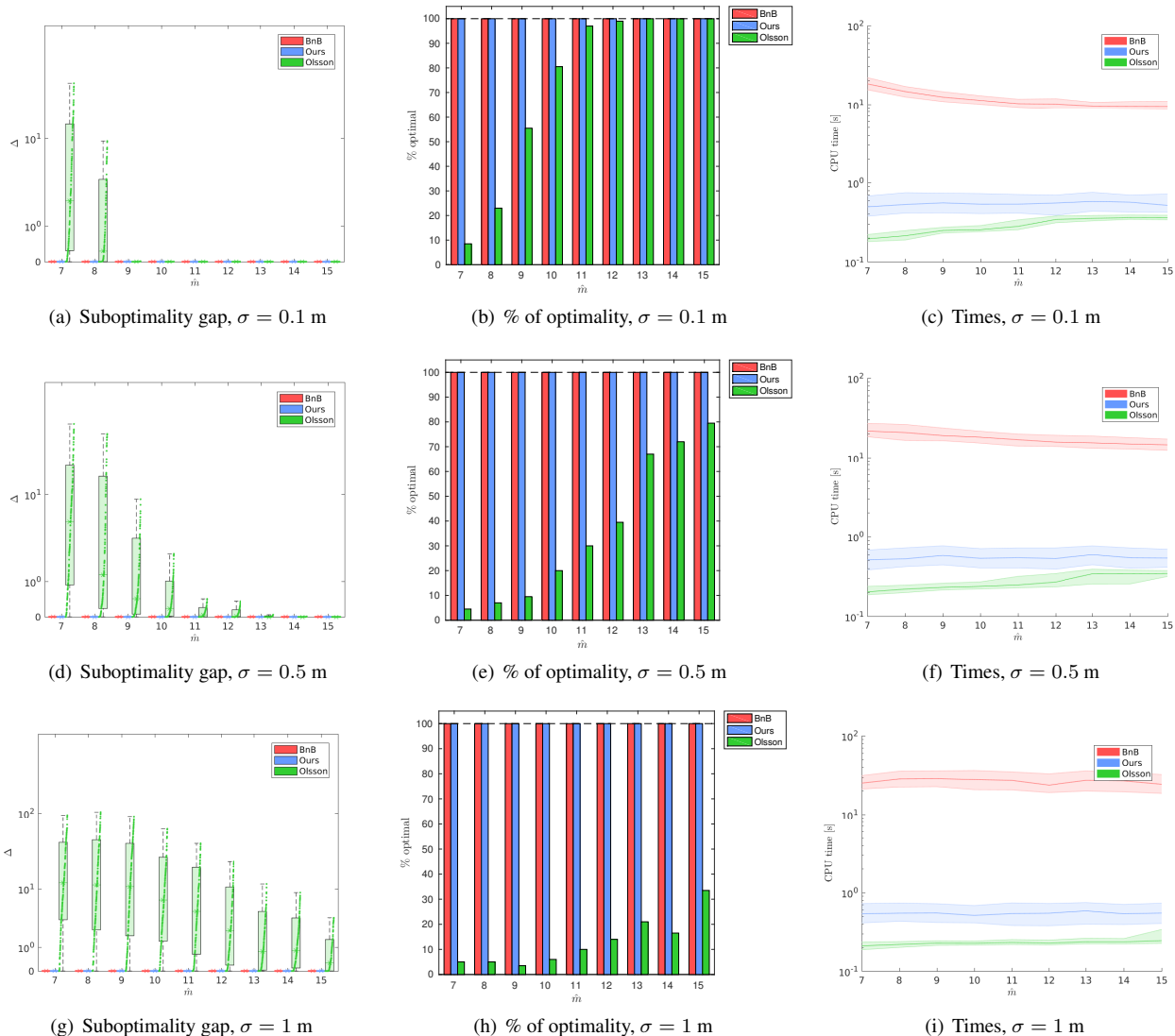
(i) Times, $\sigma = 1$ m

Figure 3. Evaluation in a challenging range of effective number of correspondences, from the almost minimal case $\hat{m} = 7$ to $\hat{m} = 15$. Different levels of noise were considered: A low value of $\sigma = 0.1$ m (top row), a middel value of $\sigma = 0.5$ m (middle row), and a more challenging scenario with $\sigma = 1$ m (bottom row).

6.1.2 Evaluation w.r.t. noise level

As in [3], we provide a continuous analysis on how the performance of the methods are affected for different levels of noise σ within usual application values, going from the ideal noise-free case with $\sigma = 0$ to the significantly high value of $\sigma = 0.5$ m. This evaluation is repeated several times under a fixed effective number of correspondences, namely $\hat{m} = \{7, 10, 14, 21\}$. The whole set of results can be seen in Fig. 4.

As expected, the general trend is that the difficulty of the problem increases with the noise level σ, making it harder for Olsson to reach the global solution, and also rendering the resolution with BnB slower. In all cases our solution provided the globally optimal solution, at a constant low resolution time.
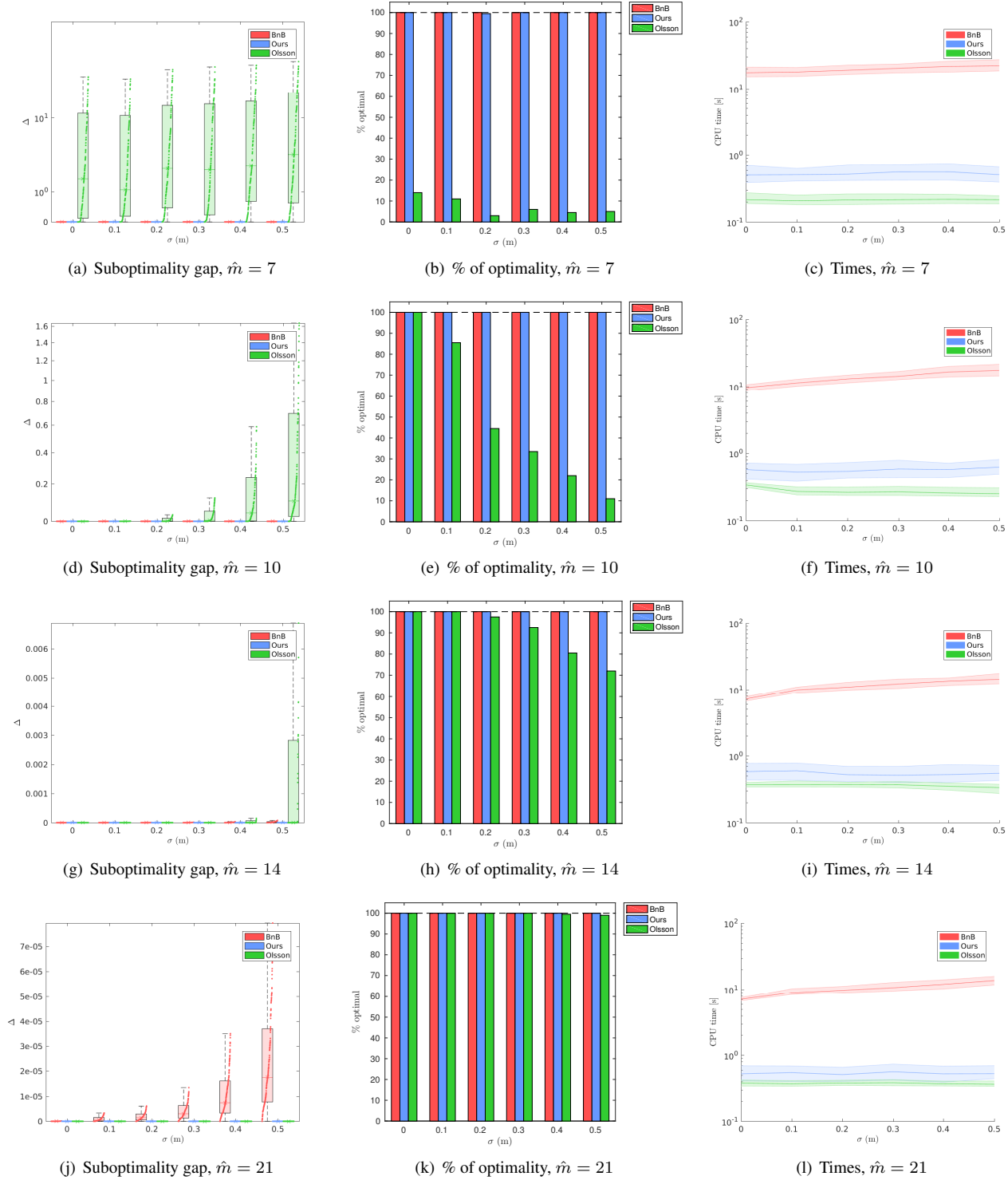
(a) Suboptimality gap, $\hat{m} = 7$

(b) % of optimality, $\hat{m} = 7$

(c) Times, $\hat{m} = 7$

(d) Suboptimality gap, $\hat{m} = 10$

(e) % of optimality, $\hat{m} = 10$

(f) Times, $\hat{m} = 10$

(g) Suboptimality gap, $\hat{m} = 14$

(h) % of optimality, $\hat{m} = 14$

(i) Times, $\hat{m} = 14$

(j) Suboptimality gap, $\hat{m} = 21$

(k) % of optimality, $\hat{m} = 21$

(l) Times, $\hat{m} = 21$

Figure 4. Evaluation w.r.t. noise level from $\sigma = 0$ to $\sigma = 0.5$ m, for different effective numbers of measurements: $\hat{m} = \{7, 10, 14, 21\}$. Note the decreasing scale in the suboptimality gap (first column of figures) as the effective number of correspondences $\hat{m}$ increases.

6.1.3   Evaluation in extreme conditions

Finally, we used the synthetic framework to test our method in a substantially difficult scenario. For that purpose, we took the most challenging conditions for the registration problem, that is almost minimal number of correspondences $\hat{m} = 7$ and *very* high noise level (far beyond any expectable value found in a real problem). Specifically, we increased the noise level σ exponentially from $\sigma = 1$ m to $\sigma = 1000$ m. With such a high level of noise the results of the problem would of little use in a real application, but again we do this just with the main aim of testing the behaviour and robustness of all the algorithms from a purely mathematical point of view.

The results, displayed in Fig. 5, show that under these severe conditions the relaxation used by `Olsson` never worked in the higher regimes of noise, and even the `BnB` method incurred in some suboptimalities probably due to numerical issues and exit conditions in the algorithm. Meanwhile, `Ours` worked in 100% of the cases without exception.

(a) Suboptimality gap, $\hat{m} = 7$ (b) % of optimality, $\hat{m} = 7$ (c) Times, $\hat{m} = 7$
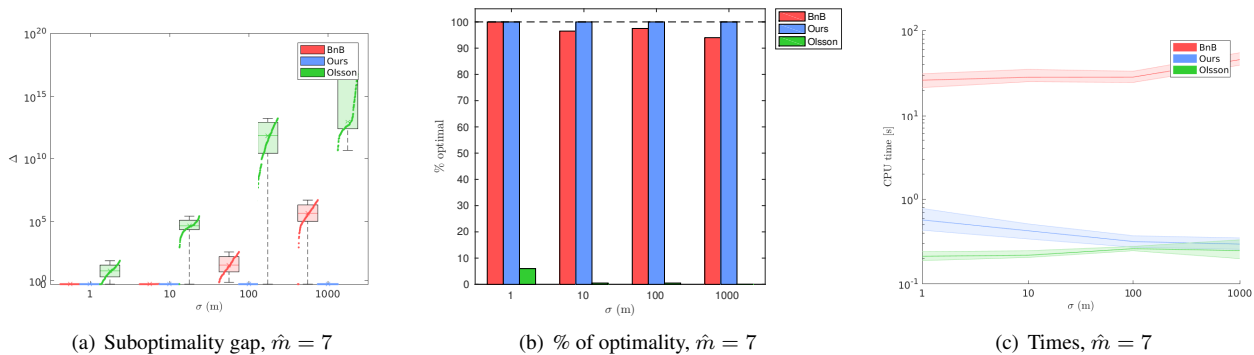
Figure 5.   Evaluation in the most severe regime of parameters: Lowest possible number of correspondences $\hat{m} = 7$ and significantly high measurement noise. Note the different scales in the axes.

6.2. Evaluation on real data

We use the same real dataset used in [3]. Following the procedure explained in the main document, we were able to generate a large set of different registration problems by taking different effective numbers of correspondences $\hat{m}$. The noise level was intrinsic to the sensor used for the measurements, so it is out of our control.

As a result, as in Section 6.1.1, the evaluation on real data considers an increasing value for $\hat{m}$. Note however that, even though the sensor noise should be well below the synthetic noise considered in the top row of Fig. 3, $\sigma = 0.1$ m, the results are closer to those of a much higher synthetic noise (compare to $\sigma = 0.5$ m in Fig. 3). This may be a consequence of the non-Gaussian noise in the real measurements, which might be turning the problem even harder.

In any case, our method still reached a globally optimal solution in all cases, regardless of the effective number of measurements $\hat{m}$, and at a fraction of the time for `BnB`.

(a) Suboptimality gap   (b) % of optimality   (c) Times
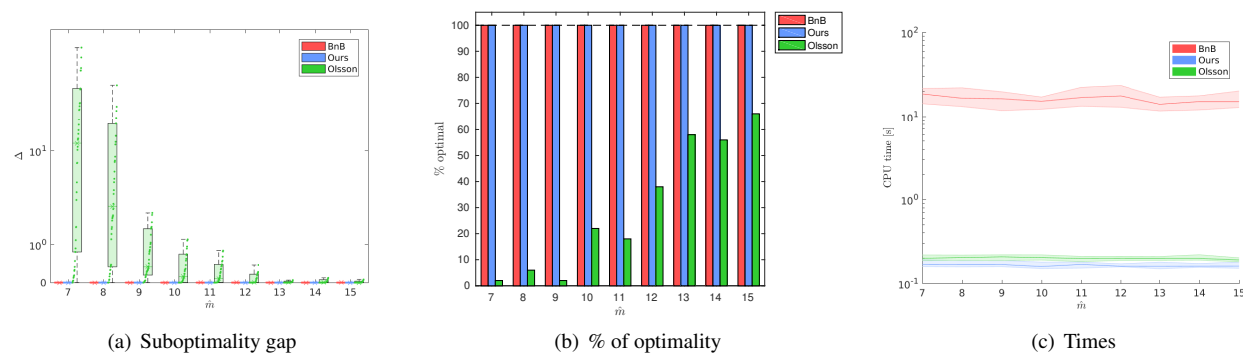
Figure 6.   Evaluation in real data taken from the Space Station dataset [3]. Different subsets of the complete dataset are sampled to get more challenging registration problems with a variable effective number of correspondences $\hat{m}$.

A. Some matrix calculus

Next we provide some properties related to the vectorization of matrix expressions. These relations were taken from [2]:

$$\text{vec}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}) = (\boldsymbol{B}^\top \otimes \boldsymbol{A})\,\text{vec}(\boldsymbol{X}) \tag{78}$$

$$\text{vec}(\boldsymbol{a}\boldsymbol{b}^\top) = \boldsymbol{b} \otimes \boldsymbol{a} \tag{79}$$

$$\text{tr}(\boldsymbol{A}^\top \boldsymbol{B}) = \text{vec}(\boldsymbol{A})^\top \text{vec}(\boldsymbol{B}) \tag{80}$$

$$\text{tr}(\boldsymbol{A}^\top \boldsymbol{X}^\top \boldsymbol{B}\boldsymbol{Y}) = \text{vec}(\boldsymbol{X})^\top (\boldsymbol{A} \otimes \boldsymbol{B})\,\text{vec}(\boldsymbol{Y}) \tag{81}$$

References

[1] S. Boyd and L. Vandenberghe. *Convex optimization*. 2004. 5

[2] P. L. Fackler. Notes on Matrix Calculus. pages 1–14, 2005. 3, 13

[3] C. Olsson and A. Eriksson. Solving quadratically constrained geometrical problems using lagrangian duality. In *Pattern Recognition, 2008. ICPR 2008. 19th Int. Conf.*, pages 1–5. IEEE, 2008. 8, 9, 10, 12

[4] C. Olsson, F. Kahl, and M. Oskarsson. Branch-and-Bound Methods for Euclidean Registration Problems. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(5):783–794, 2009. 8, 9