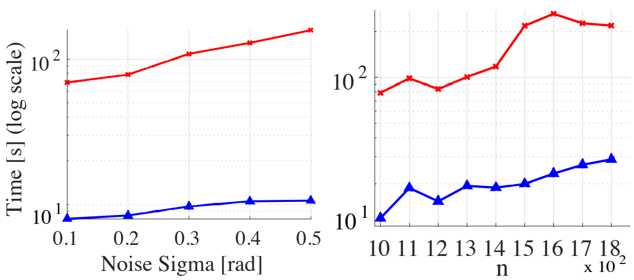
Rotation Coordinate Descent for Fast Globally Optimal Rotation Averaging

Álvaro Parra[1*] Shin-Fang Chng[1*] Tat-Jun Chin[1] Anders Eriksson[2] Ian Reid[1]
[1] School of Computer Science, The University of Adelaide
[2] School of Information Technology and Electrical Engineering, University of Queensland

A. Additional Results

A.1. Varying noise levels and number of cameras in SLAM graphs



(a) (b)

Figure 1. Runtime [s] (in log-scale) for SLAM camera graphs with $d_{\mathcal{G}} = 0.2$. (a) Varying σ in $[0.1, 0.5]$ rad. and $n = 1000$. (b) Varying n in $[1000, 1800]$ with $\sigma = 0.1$ rad. See Fig. 3 in the main text for the description of the legend.

A.2. Quantitative results for the SfM large scale real-world dataset (Error in °)

Table 1 provides the mean, median and maximum angular errors (in °) for the SfM large scale real-world data of Table 2 in the main text.

B. Further details

B.1. Conditions on the noise level for the strong duality of Eq. (5)

For the following rotation averaging problem (Eq. (5) in the main text)

$$\min_{R_1,\ldots,R_n \in SO(3)} \sum_{(i,j)\in\mathcal{E}} d_{\text{chordal}}(R_j R_i^T, \tilde{R}_{ij})^2, \tag{1}$$

we present a bound on the *angular* residual errors

$$\alpha_{ij} = d_\angle(R_j^* R_i^{*T}, \tilde{R}_{ij}) \tag{2}$$

such that its *strong duality* holds.

The main result of [Theorem 4.1, 11] is the proof of the strong duality of Problem (1) if

$$|\alpha_{ij}| \leq \alpha_{\max} \quad \forall (i,j) \in \mathcal{E}, \tag{3}$$

where

$$\alpha_{\max} = 2 \arcsin\left(\sqrt{\frac{1}{4} + \frac{\lambda_2(L_G)}{2d_{\max}}} - \frac{1}{2}\right). \tag{4}$$

$\lambda_2(L_G)$ and $d_{\max}$ in (4) are related to the structure of the camera graph. More precisely, $\alpha_{\max}$ depends on the connectivity of the camera graph represented by its Fiedler value $\lambda_2(L_G)$ (the second smallest eigenvalue of its Laplacian L_G), and its maximal vertex degree $d_{\max}$ (c.f. to [11] and [12] for more details).

From the dependency of $\alpha_{\max}$ on the structure of the camera graph, it can be established that the most favourable case (admitting the largest residuals) is the complete graph for which $\alpha_{\max} \approx 42.9°$. The other extreme case is a cycle with $\alpha_{\max} = \pi/n$, which induces a low angular bound for a large number of cameras although [11] suggested that this bound was "quite conservative".

Although conditions were presented in terms of the angular distance, we remark that a chordal bound can also be established for the chordal residuals $\{d_{\text{chordal}}(R_j^* R_i^{*T}, \tilde{R}_{ij})\}$ of Problem 1 as both distances are related [16]:

$$d_{\text{chordal}}(R, S) = 2\sqrt{2} \sin\left(\frac{d_\angle(R,S)}{2}\right). \tag{5}$$

B.2. Zero duality gap between (P) and (DD)

Eriksson *et al.* [11] have proven that under mild conditions on the noise level (see Sec. B.1), there is zero duality gap between their primal problem (P_{orig}) and their SDP relaxation (DD_{orig}). Since we defined our primal problem (P) and its SDP relaxation (DD) following a different convention for the relative rotation definition than [11], here we show that our (P) and (DD) problems are equivalent to their

[*] equal contribution

Dataset characteristics				Error [°]					
Name	n	m	$d_{\mathcal{G}}$	Mean		Median		Max	
				Init.	RCD	Init.	RCD	Init.	RCD
Alcatraz Tower	172	14706	1.00	4.65	0.63	4.16	0.63	20.57	0.63
Doge Palace	241	19753	0.68	5.95	0.56	5.18	0.52	23.07	1.82
King's College	328	41995	0.78	6.98	0.39	4.89	0.36	32.99	0.83
Alcatraz Garden	419	51635	0.59	7.79	0.92	6.62	0.92	25.65	1.79
Linkoping	538	34462	0.24	7.42	0.63	6.90	0.58	25.15	2.63
UWO	692	80301	0.33	7.20	0.55	6.42	0.52	24.43	2.19
Orebro Castle	761	116589	0.40	6.66	0.52	5.55	0.50	27.30	1.33
Spilled Blood	781	117814	0.39	11.41	6.81	7.61	6.46	42.82	12.35
Lund Cathedral	1207	177289	0.24	8.05	0.40	7.52	0.34	27.12	3.52
San Marco	1498	757037	0.67	5.03	0.22	4.45	0.21	26.59	0.94

Table 1. Quantitative results for the SfM large scale real-world dataset [25]. We provide the mean, median, and maximum angular errors of the initial solution (Init.) and RCD (in °).

counterparts in [11]. Hence the zero duality gap extends to them.

We defined our primal problem as follows. By rewriting the chordal distance using trace, (1) becomes (Eq. (8) in the main text)

$$\min_{R_1,\dots,R_n\in SO(3)} -\sum_{(i,j)\in\mathcal{E}} \operatorname{tr}(R_j^T\tilde{R}_{ij}R_i). \tag{6}$$

By the transpose invariance of the trace, (6) is equivalent to

$$\min_{R_1,\dots,R_n\in SO(3)} -\sum_{(i,j)\in\mathcal{E}} \operatorname{tr}(R_i^T\tilde{R}_{ij}^TR_j). \tag{7}$$

Our primal definition comes from rewriting (6) more compactly as

$$\min_{R\in SO(3)^n} -\operatorname{tr}(R^T\tilde{R}R) \tag{P}$$

using matrix notations, where

$$R = \left[R_1^T\, R_2^T \cdots R_n^T\right]^T \in SO(3)^n \tag{8}$$

contains the target variables, and $\tilde{R}$ encodes the transposes of the relative rotations. $\tilde{R}$ is then defined as

$$\tilde{R} = \begin{bmatrix} 0_3 & a_{12}R_{12}^T & \cdots & a_{1n}R_{1n}^T \\ a_{21}R_{21}^T & 0_3 & \cdots & a_{2n}R_{2n}^T \\ \vdots & 0_3 & \ddots & \vdots \\ a_{n1}R_{n2}^T & a_{n2}R_{n2}^T & \cdots & 0_3 \end{bmatrix}, \tag{9}$$

where a_{ij} are the elements of the adjacency matrix A of $\mathcal{G}$.

We now show that (P) is equivalent to the primal in [11], which is defined as (Eq. (11) in [11])

$$\min_{Q\in SO(3)^n} -\operatorname{tr}(Q\tilde{Q}Q^T), \tag{P$_{\text{orig}}$}$$

where Q is a "row" vector containing rotation matrices

$$Q = [Q_1,\dots,Q_n], \tag{10}$$

and $\tilde{Q}$ encodes the relative measurements as

$$\tilde{Q} = \begin{bmatrix} 0_3 & a_{12}Q_{12} & \cdots & a_{1n}Q_{1n} \\ a_{21}Q_{21} & 0_3 & \cdots & a_{2n}Q_{2n} \\ \vdots & 0_3 & \ddots & \vdots \\ a_{n1}Q_{n2} & a_{n2}Q_{n2} & \cdots & 0_3 \end{bmatrix}. \tag{11}$$

However, relative rotations Q_{ij} in [11] are defined such that (Eq. (4) in [11])

$$Q_{ij} = Q_i^TQ_j. \tag{12}$$

Contrast to our definition from Eq. (1) in the main text where we define relative rotations in the ideal case as

$$R_{ij} = R_jR_i^T. \tag{13}$$

The following equivalences can then be established:

$$R_i = Q_i^T \text{ and } R_{ij} = Q_{ij}^T, \tag{14}$$

which implies that $Q = R^T$, $\tilde{Q} = \tilde{R}$, and therefore (P) is equivalent to (P$_{\text{orig}}$) in the sense that their objective values are the same and their optimisers are related by a translation.

Similarly, our SDP relaxation

$$\min_{Y\in\mathbb{R}^{3n\times 3n}} \quad -\operatorname{tr}(\tilde{R}Y) \tag{DD}$$

$$\text{s.t.} \quad Y_{i,i} = I_3,\; i = 1,\dots,n. \tag{15a}$$

$$Y \succeq 0, \tag{15b}$$

is equivalent to its counterpart in [11]. In effect, they are the same as matrices encoding rotations are the same for both problems ($\tilde{Q} = \tilde{R}$).

B.3. Validity of Algorithm 1 as equivalent to BCD in Eriksson et al. [11]

Here we show that BCD as presented in Algorithm 1 in the main text is equivalent to the original BCD algorithm for rotation averaging proposed in [11]. To facilitate presentation, we call BCD-Ours to Algorithm 1 in the main text and BCD-Orig to Algorithm 1 in [11].

The improvement of BCD-Ours over BCD-Orig is that instead of creating a temporary large square matrix

$$B = \begin{bmatrix} Y^{(t)}_{(1:k-1);(1:k-1)} & Y^{(t)}_{(1:k-1);(k+1:n)} \\ Y^{(t)}_{(k+1:n);(1:k-1)} & Y^{(t)}_{(k+1:n);(k+1:n)} \end{bmatrix} \tag{16}$$

as in BCD-Orig, BCD-Ours creates a temporary vector which allows to operates directly on $Y^{(t)}$ as we will show next.

Note that B are the elements in $Y^{(t)}$ that are kept constant during the current iteration in BCD-Orig and BCD-Ours. On the other hand, the updated components for $Y^{(t)}$ in BCD-Orig are obtained from the optimiser X^* of an SDP problem (Problem (26) in the main text) which has the following explicit solution:

$$X^* = BC\left[\left(C^T BC\right)^{\frac{1}{2}}\right]^{\dagger}, \tag{17}$$

where $C \in \mathbb{R}^{3(n-1)\times 3}$ is the k-th column of $\tilde{R}$ without its k-th row, i.e.,

$$C = \begin{bmatrix} \tilde{R}^{(t)}_{(1:k-1);(k:k)} \\ \tilde{R}^{(t)}_{(k+1:n);(k:k)} \end{bmatrix}. \tag{18}$$

Instead of computing the updates from (17), BCD-Ours solves

$$S = Z\left[\left(W^T Z\right)^{\frac{1}{2}}\right]^{\dagger}, \tag{19}$$

where $W \in \mathbb{R}^{3n\times 3}$ is the k-th column of $\tilde{R}$, i.e.,

$$W = \tilde{R}_{:,k}\,, \tag{20}$$

and

$$Z = Y^{(t)}W \tag{21}$$

is a temporary vector.

We will show next that X^* es equal to S without its k-th element. Since BCD-Ours ignores the k-th element of S during the update (Line 7 in BCD-Ours), BCD-Ours and BCD-Orig produce the same output.

Note first that the pseudo-inverse parts of (17) and (19) are the same since

$$C^T BC = W^T Z \tag{22}$$

as the k-th element in W is zero (W is the k-th column of $\tilde{R}$ which has diagonal elements equal to 0_3). Similarly BC is equal to Z if removing the k-th element of Z. Hence (19) produces X^* after removing the k-th element of S.