Error in proof of Theorem 12 of [2]

Theorem 12, stated on p. 43, is proven on pp. 41-43. The final line of the proof says "combining inequalities (134), (140), (151), and (152), we obtain the following [Theorem 12]," but unfortunately (151) and (152) are not true in general. Nonetheless, the statement of Theorem 12 is true as stated, and I propose an alternative argument that restores the proof. In addition, we are able to prove a bound that is stronger as the measurement error goes to zero.

1 Recalling the notation

We recall that $\underline{R}, R^*$ are elements of $O(d)^n$, meaning

$$\underline{R} = \begin{bmatrix} \underline{R}_1 & \cdots & \underline{R}_n \end{bmatrix} \text{ and } R^* = \begin{bmatrix} R_1^* & \cdots & R_n^* \end{bmatrix},$$

where $\underline{R}_i, R_i^* \in \mathbb{R}^{d\times d}$ are orthogonal matrices. P is the orthogonal projection of R^* onto the orthogonal complement of $\underline{R}$. In $O(d)^n$, this means

$$P = R^* - \frac{1}{n} R^* \underline{R}^T \underline{R},$$

and in (143) we discover that

$$\|P\|_F^2 = dn - \frac{1}{n} \left\| \underline{R} R^{*T} \right\|_F^2.$$

We define the $O(d)$ orbital distance between $\underline{R}$ and R^* by

$$d_{\mathcal{O}}(\underline{R}, R^*) = \min_{G \in O(d)} \|\underline{R} - GR^*\|_F,$$

and from Theorem 5 (p. 35) we have that

$$d_{\mathcal{O}}(\underline{R}, R^*)^2 = 2dn - 2\|\underline{R} R^{*T}\|_*.$$

2 The argument for (150) is incorrect

The goal is to prove

$$\|P\|_F^2 \geq \frac{1}{2} d_{\mathcal{O}}(\underline{R}, R^*)^2, \tag{1}$$

which is essential in the proof of Theorem 12 (and Theorem 12 is quoted in the proof of the main result, Proposition 2, on p. 44!). However, the "optimizaton strategy" of (147)-(150) doesn't work. Setting $\delta = d_{\mathcal{O}}(\underline{R}, R^*)$, we can rephrase the optimization problem of (147) as

$$\begin{array}{rl} \max & \sum_{i=1}^{d} \sigma_i^2 \\ \text{s.t.} & \sum_{i=1}^{d} \sigma_i = a \\ & \sigma_i \geq 0 \end{array},$$

where $a = dn - \delta^2/2$. This, in turn, is obviously equivalent to

$$\max_{\|x\|_1 = a} \|x\|_2^2.$$

From standard norm inequalities, we have that $\frac{1}{\sqrt{d}}\|x\|_1 \leq \|x\|_2 \leq \|x\|_1$, with equality on the left when $x_i = t\mathbb{1}$ and equality on the right when $x = e_i$ for some $i \in [d]$. In this instance, the maximal value will be $\|x\|_1^2$, or

$$\epsilon^2 = \left(\sum_{i=1}^{d} \sigma_i\right)^2 = \left(dn - \frac{\delta^2}{2}\right)^2 = d\left(d\left(n - \frac{\delta^2}{2d}\right)^2\right),$$

which is greater than the value given in (150) by a factor of d. This leads to (151) becoming

$$\|P\|_F^2 \geq dn - \frac{d^2}{n}\left(n - \frac{\delta^2}{2d}\right)^2 = d\delta^2 + dn(1-d) - \frac{\delta^4}{4n}.$$

Following the argument of (152), we get

$$\|P\|_F^2 \geq \frac{d}{2}\delta^2 - dn(d-1),$$

which, if combined with (134) and (140), gives

$$\delta^2 \leq \frac{4n\|\Delta Q\|_2}{\lambda_{d+1}(Q)} + 2n(d-1).$$

In this state, the bound proves nothing; in particular, we don't have that $\lim_{\|\Delta Q\| \to 0}\|\Delta R\| = 0$, not to mention that Theorem 5 (p. 35) trivially bounds $\delta^2 \leq 2dn$.

3 An alternative argument

Consider that, from (143) and (144), $\|P\|_F^2 \geq \frac{1}{2} d_{\mathcal{O}}(\underline{R}, R^*)^2$ holds for $\underline{R}, R^* \in O(d)^n$ iff

$$dn - \frac{1}{n}\left\|\underline{R}R^{*T}\right\|_F^2 \geq dn - \left\|\underline{R}R^{*T}\right\|_* \iff \left\|\underline{R}R^{*T}\right\|_* \geq \frac{1}{n}\left\|\underline{R}R^{*T}\right\|_F^2,$$

so it suffices to prove this last inequality for all $\underline{R}, R^* \in O(d)^n$. Fix $\underline{R}, R^* \in O(d)^n$; taking $\sigma_1 \geq \ldots \geq \sigma_d$ to be the singular values of $\underline{R}R^{*T}$ and setting $(\sigma_1, \ldots, \sigma_n)^T =: \sigma \in \mathbb{R}^n$, this happens iff

$$\|\sigma\|_1 \geq \frac{1}{n}\|\sigma\|_2^2. \tag{2}$$

By Hölder's inequality, we have

$$\|\sigma\|_2^2 \leq \|\sigma\|_1 \|\sigma\|_\infty. \tag{3}$$

Now

$$\|\sigma\|_\infty = \left\|\underline{R}R^{*T}\right\| = \left\|\sum_{i=1}^n \underline{R}_i R_i^{*T}\right\| \leq \sum_{i=1}^n \left\|\underline{R}_i R_i^{*T}\right\| = n, \tag{4}$$

since $\underline{R}_i R_i^{*T}$ is orthogonal. Combining (3) and (4), we have

$$\frac{1}{n}\|\sigma\|_2^2 \leq \frac{1}{n}\|\sigma\|_1 \|\sigma\|_\infty \leq \frac{1}{n}\|\sigma\|_1 n = \|\sigma\|_1,$$

which yields (2) as needed.

4 Improvement to the bound of Theorem 12

In sections 2 and 3, we ignored the origin of $\underline{R}$ and R^*, but in the present section 4, we note that $\underline{R}$ and R^* arise as optima of certain constrained optimization problems, though for now we ignore the specifics and assume only the relevant identities. Given symmetric matrices $\tilde{Q}, \underline{Q} \succeq 0$ in $\mathbb{R}^{dn \times dn}$, we assume $\mathrm{Tr}(\underline{Q}\,\underline{R}^T \underline{R}) = 0$ and that $\mathrm{Tr}(\tilde{Q} R^{*T} R^*) \leq \mathrm{Tr}(\tilde{\underline{Q}}\,\underline{R}^T \underline{R})$. Setting $\Delta Q = \tilde{Q} - \underline{Q}$, these may be combined into

$$\begin{aligned}\mathrm{Tr}(\tilde{Q}\underline{R}^T\underline{R}) &= \mathrm{Tr}(\Delta Q \underline{R}^T \underline{R}) + \mathrm{Tr}(\underline{Q}\,\underline{R}^T\underline{R}) \\ &\qquad \geq \mathrm{Tr}(\Delta Q R^{*T} R^*) + \mathrm{Tr}(\underline{Q} R^{*T} R^*) = \mathrm{Tr}(\tilde{Q} R^{*T} R^*),\end{aligned}$$

which we rearrange to get

$$\mathrm{Tr}(\underline{Q} R^{*T} R^*) \leq \mathrm{Tr}(\Delta Q \underline{R}^T \underline{R}) - \mathrm{Tr}(\Delta Q R^{*T} R^*). \tag{5}$$

In the style of the proof of Lemma 4.1 in [1], we are able to achieve a notable improvement to the bound of Theorem 12 in [2]. From (5), and using $\mathrm{Tr}(\Delta Q \underline{R}^T \underline{R}) = \mathrm{vec}(\underline{R})^T (\Delta Q \otimes I_n) \, \mathrm{vec}(\underline{R})$, we get

$$\begin{aligned}
\mathrm{Tr}(\underline{Q} R^{*T} R^*) &\leq \mathrm{vec}(\underline{R} - R^*)^T (\Delta Q \otimes I_n) \, \mathrm{vec}(\underline{R} + R^*) \\
&\leq \|\mathrm{vec}(\underline{R} - R^*)\|_2 \|\Delta Q \otimes I_n\|_2 \|\mathrm{vec}(\underline{R} + R^*)\|_2 \\
&= \|\underline{R} - R^*\|_F \|\Delta Q\|_2 \|\underline{R} + R^*\|_F \\
&\leq 2\sqrt{dn} \|\underline{R} - R^*\|_F \|\Delta Q\|_2
\end{aligned}$$

Assuming, without loss of generality, that $\underline{R}$ and R^* are representatives of their orbits such that $\|\underline{R} - R^*\|_F = d_{\mathcal{O}}(\underline{R}, R^*)$ and combining this with (1) and (140) of [2], which gives

$$\mathrm{Tr}(\underline{Q} R^{*T} R^*) \geq \lambda_{d+1}(Q) \|P\|_F^2,$$

we have

$$d_{\mathcal{O}}(\underline{R}, R^*) \leq \frac{4\sqrt{dn} \|\Delta Q\|_2}{\lambda_{d+1}(Q)}. \tag{6}$$

We compare this to the original result, which gives

$$d_{\mathcal{O}}(\underline{R}, R^*) \leq \sqrt{\frac{4dn \|\tilde{Q} - \underline{Q}\|_2}{\lambda_{d+1}(\underline{Q})}}. \tag{7}$$

We remark that, as $\|\tilde{Q} - \underline{Q}\|_2$ goes to zero, the higher exponent in the new bound guarantees a faster rate of convergence of $R^* \to \underline{R}$. Since both bounds are valid, we may always use whichever is stronger: indeed, the square root bound of (7) is stronger exactly when $\|\tilde{Q} - \underline{Q}\|_2 / \lambda_{d+1}(\underline{Q}) > \frac{1}{4}$. We remark that (6) is trivial when $\|\tilde{Q} - \underline{Q}\|_2 / \lambda_{d+1}(\underline{Q}) \geq \frac{1}{2\sqrt{2}}$, while (7) is trivial when $\|\tilde{Q} - \underline{Q}\|_2 / \lambda_{d+1}(\underline{Q}) \geq \frac{1}{2}$, as $d_{\mathcal{O}}(\underline{R}, R^*) \leq \sqrt{2dn}$.

References

[1] A. S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*, 163(1):145–167, May 2017.

[2] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard. SE-sync: A certifiably correct algorithm for synchronization over the special euclidean group. *CoRR*, abs/1611.00128, 2016.