

## A. Proofs

We begin by proving that  $D$ -optimizable functions are Lipschitz on the cube  $\mathcal{C}^0$  as we claimed in the beginning of Subsection 3.2. In the remainder of the Section we restate and prove the theorems and lemmas stated in the main text.

**Lemma 5.** *If  $F$  is  $D$ -quasi optimizable in  $\mathcal{C}^0$ , then there exists  $L > 0$  such that.*

$$F(x_1) - F(x_2) \leq L\|x_1 - x_2\|, \forall x_1, x_2 \in \mathcal{C}^0 \quad (\text{A.1})$$

*Proof.* If  $F(x_1) \leq F(x_2)$  then (A.1) holds for all  $L \geq 0$ . Now assume  $F(x_1) > F(x_2)$  and let  $y_2$  be the minimizer of  $E(x_2, \cdot)$ . Then due to the differentiability of  $E(\cdot, y_2)$ ,

$$F(x_1) - F(x_2) \leq E(x_1, y_2) - E(x_2, y_2) \leq L\|x_1 - x_2\|$$

where  $L$  is the maximum of the norm of the gradient of  $E(\cdot, y)$  over all  $x \in \mathcal{C}^0$  and  $y \in \mathcal{Y}$ .  $\square$

**Theorem 2.** *There exist positive constants  $C_1, \dots, C_4$ , such that*

$$C_1\epsilon^{-D/2} \leq n_{\text{BnB}} \leq C_2\epsilon^{-D}. \quad (14)$$

$$n_{\text{qBnB}} \leq C_3\epsilon^{-D/2}. \quad (15)$$

Furthermore, if  $E$  has a finite number of minimizers  $(x_\ell^*, y_\ell^*)_{\ell=1}^N$ , and the Hessian of  $E(\cdot, y_\ell^*)$  is strictly positive definite for all  $\ell$ , then

$$n_{\text{qBnB}} \leq C_4 \log_2(1/\epsilon). \quad (16)$$

*Proof. Part 1.* We begin with a general discussion of the complexity of Algorithm 1 in both the BnB and quasi-BnB version, and prove the upper bound on  $n_{\text{BnB}}$  and  $n_{\text{qBnB}}$  simultaneously. To do so we denote

$$\Delta_\alpha(\delta) = M\delta^\alpha, \alpha = 1, 2. \quad (\text{A.2})$$

We denote by  $n_\alpha$  the number of  $F$ -evaluations used by the algorithm in each case, so

$$n_1 = n_{\text{BnB}}, n_2 = n_{\text{qBnB}}.$$

For simplicity of notation we will use  $\text{lb}_i$  in the first part of the proof to denote both lower bounds and quasi lower bounds (in contrast with the  $\text{qlb}_i$  notation used elsewhere). We also note that while our bounds (18),(21) include besides  $M\delta^2$  also higher terms in  $\delta$  these terms can be absorbed into  $M\delta^2$  with a larger value of  $M$ .

In the proof we will call  $g$  from Algorithm 1 the *generation* of the algorithm. We begin with showing that the algorithm necessarily terminates, and bounding the final value of  $g$  which we denote by  $g_f$ . First recall from Subsection 3.2 that if  $\mathcal{C}_{h_i}(x_i) \in L_g$  contains a global minimum  $x_*$ , then its sub-cubes will always be added to  $L_{g+1}$ . Recall also that the global lower bound  $\text{lb}$  obtained by minimizing over all  $\text{lb}_i$  obtained from the cubes in  $L_g$ , is a global lower bound for  $F$  over  $\mathcal{C}^0$ . It follows that if the algorithm does terminate, then the output  $x_*$  is an  $\epsilon$ -optimal solution as its difference from the minimum  $F^*$  satisfies

$$F(x_*) - F^* = \text{ub} - F^* \leq \text{ub} - \text{lb} \leq \epsilon. \quad (\text{A.3})$$

Now to bound  $g_f$  note that cubes in  $L_g$  have half-edge length  $h(g) \equiv h_0 2^{-g}$ . The algorithm must terminate once it visits all cubes of generation  $g$  whose edge length  $h(g)$  satisfies

$$\Delta_\alpha(\sqrt{D}h(g)) \leq \epsilon. \quad (\text{A.4})$$

This is because for all cubes  $\mathcal{C}_{h_j}(x_j)$  in this generation,

$$\text{ub} - \text{lb}_j \leq \text{ub}_j - \text{lb}_j = \Delta_\alpha(\sqrt{D}h_j) \leq \epsilon.$$

and by taking the minimum over  $j$  we obtain that  $\text{ub} - \text{lb} \leq \epsilon$ .

Some algebraic manipulation shows that (A.4) occurs when

$$g_f = g_f(\epsilon) = \lceil 1/\alpha \log \frac{\bar{C}_\alpha}{\epsilon} \rceil \text{ where } \bar{C}_\alpha = M(\sqrt{D}h_0)^\alpha \quad (\text{A.5})$$

and we use  $\log = \log_2$  throughout this proof. The number of  $F$  evaluations  $n_\alpha$  is bounded by the worst case scenario where all cubes need to be divided in all generations

$$n_\alpha \leq \sum_{g=0}^{g_f} 2^{Dg} = 2^{Dg_f} \sum_{g=0}^{g_f} 2^{-Dg} \leq 2^{Dg_f} \sum_{g=0}^{\infty} 2^{-Dg} = 2^{Dg_f} \frac{1}{1-2^{-D}} \stackrel{\text{(A.5)}}{\leq} \left[ \frac{2^D}{1-2^{-D}} \bar{C}_\alpha^{D/\alpha} \right] \epsilon^{-D/\alpha}.$$

this proves the upper bound on  $n_1, n_2$ .

**Part 2.** We now show a lower bound on  $n_{\text{BnB}}$ . Our first step is to show that the sub-cubes of a given cube have better=larger lower bounds. To see this let  $\mathcal{C}^1 = \mathcal{C}_{h_1}(x_1)$  be a cube, and let  $\mathcal{C}^2 = \mathcal{C}_{h_2}(x_2)$  be one of its sub-cubes. Then  $h_2 = h_1/2$  and  $|x_1 - x_2| = \sqrt{D}h_1/2$ . It follows that, using the notation of (A.2) with  $\alpha = 1$ ,

$$F(x_2) \geq F(x_1) - M\sqrt{D}h_1/2 \quad (\text{A.6})$$

Now denoting by  $\text{lb}_1$  and  $\text{lb}_2$  the lower bounds computed for the cubes  $\mathcal{C}^1$  and  $\mathcal{C}^2$  respectively, we have

$$\text{lb}_2 = F(x_2) - M\sqrt{D}h_1/2 \stackrel{\text{(A.6)}}{\geq} F(x_1) - M\sqrt{D}h_1 = \text{lb}_1,$$

and so we have  $\text{lb}_2 \geq \text{lb}_1$  as we stated.

Next note that by the quadratic bound (9) we have that for a global minimum  $x_*$  and  $\eta = \sqrt{\epsilon/C}$ ,

$$F(x) - F(x_*) \leq \epsilon, \text{ for all } x \in B_\eta(x_*). \quad (\text{A.7})$$

Now, let  $g_F$  denote the value of  $g_f(2\epsilon)$  from (A.5), for the case  $\alpha = 1$ . Recall that  $g_f(2\epsilon)$  is defined as the first integer for which (A.4) holds, where  $\epsilon$  is replaced by  $2\epsilon$ , and  $\alpha = 1$ . Thus for  $g < g_F$  we have that

$$M\sqrt{D}h(g) > 2\epsilon \quad (\text{A.8})$$

Let  $\mathcal{C}_{h_i}(x_i)$  be a cube of generation  $g_0$  containing  $x_*$ , where  $g_0$  is large enough so that the diameter of the cube is smaller than  $\eta$  and thus it is contained in  $B_\eta(x_*)$ . This occurs for

$$g_0 = \lceil \log(\bar{C}/\sqrt{\epsilon}) \rceil, \text{ where } \bar{C} = 2h_0\sqrt{CD}.$$

Every sub-cube  $\mathcal{C}_{h_j}(x_j)$  of  $\mathcal{C}_{h_i}(x_i)$ , from any generation  $g_0 \leq g < g_F$ , satisfies

$$\text{lb}_j = F(x_j) - M\sqrt{D}h_j < \stackrel{\text{(A.8)}}{<} F(x_j) - 2\epsilon \stackrel{\text{(A.7)}}{\leq} F(x_*) - \epsilon. \quad (\text{A.9})$$

In particular it follows that the cube  $\mathcal{C}_{h_i}(x_i)$ , and all its sub-cubes, will be visited during the BnB search. This is because we saw that lower bounds improve by refinement, and so by (A.9) any cube from the earlier generations  $g < g_0$  which contains  $\mathcal{C}_{h_i}(x_i)$ , also has lower bounds which are lower than the global minimum (by at least  $\epsilon$ ) and so such a cube would not be removed from the search..

We can now bound  $n_{\text{BnB}}$  by the number of subcubes of  $\mathcal{C}_{h_i}(x_i)$  at the  $g_F - 1$  generation alone:

$$n_{\text{BnB}} \geq 2^{D(g_F-1-g_0)} \quad (\text{A.10})$$

Now

$$g_F - 1 - g_0 = -1 + \lceil \log \frac{\bar{C}_1}{2\epsilon} \rceil - \lceil \log(\bar{C}/\sqrt{\epsilon}) \rceil \geq \log(\bar{C}_1/(2\bar{C}\sqrt{\epsilon})) - 2 = \log(\bar{C}_1/(8\bar{C}\sqrt{\epsilon})).$$

So returning to (A.10) we obtain

$$n_{\text{BnB}} \geq 2^{D(g_F-1-g_0)} \geq \left( \frac{\bar{C}_1}{8\bar{C}} \right)^D \epsilon^{-D/2}.$$

**Part 3.** We now turn to prove the last part of the theorem. Let  $\mathcal{J}(\ell)$  denote the set of indices  $k$  for which  $(x_\ell^*, y_k^*)$  is a minimizer. Note that we always have that  $\ell$  is in  $\mathcal{J}(\ell)$ , and if  $k \in \mathcal{J}(\ell)$  then  $x_\ell^* = x_k^*$ . Let  $m$  be half of the minimum over the minimal eigenvalue of the hessian of  $E(\cdot, y_\ell^*)$  at  $x_\ell^*$  for all  $\ell$ . The assumption that  $E$  has a finite number  $N$  of minimizers  $x_\ell^*, y_\ell^*$ , with strictly positive definite hessian, implies that  $m > 0$ , and so for small enough positive  $\eta$ ,

$$F(x) - F(x_\ell^*) = \min_{k \in \mathcal{J}(\ell)} E(x, y_k^*) - E(x_k^*, y_k^*) \geq m \|x - x_\ell^*\|^2, \forall 1 \leq \ell \leq N \text{ and } \forall x \in B_\eta(x_\ell^*). \quad (\text{A.11})$$

The minimum of  $F$  on  $\mathcal{C}^0 \setminus \bigcup_i B_{\eta/2}(x_\ell^*)$  is strictly larger than  $F^*$ . Therefore there exists some  $g_0$  independent of  $\epsilon$ , such that all cubes of generation  $g_0$  which are not contained in one of the balls  $B_\eta(x_\ell^*)$  will be removed in the  $g_0$ -th stage.

We now claim that for  $g \geq g_0$ ,  $g$ -th generation cubes  $\mathcal{C}_{h_i}(x_i)$  contained in one of the balls  $B_\eta(x_\ell^*)$  will be removed if

$$\|x_i - x_\ell^*\|_\infty > \sqrt{\frac{2MD}{m}} h_i = \sqrt{\frac{2\Delta_*(\sqrt{D}h_i)}{m}}. \quad (\text{A.12})$$

This is because

$$\begin{aligned} \text{qlb}_j &= F(x_i) - \Delta_*(\sqrt{D}h_i) \stackrel{(\text{A.11})}{\geq} F(x_\ell^*) + m\|x_i - x_\ell^*\|^2 - \Delta_*(\sqrt{D}h_i) \\ &= \text{ub} + (F(x_\ell^*) - \text{ub}) + m\|x_i - x_\ell^*\|^2 - \Delta_*(\sqrt{D}h_i) \stackrel{(*)}{\geq} \text{ub} + m\|x_i - x_\ell^*\|^2 - 2\Delta_*(\sqrt{D}h_i) \\ &\geq \text{ub} + m\|x_i - x_\ell^*\|_\infty^2 - 2\Delta_*(\sqrt{D}h_i) \stackrel{(\text{A.12})}{>} \text{ub} \end{aligned}$$

where  $(*)$  follows from the fact that if  $\mathcal{C}_{h_i}(x_i)$  is the  $g$ -th generation cube containing  $x_\ell^*$ , then

$$F(x_\ell^*) - \text{ub} \geq F(x_\ell^*) - \text{ub}_i = F(x_\ell^*) - F(x_i) \geq -\Delta_*(\sqrt{D}h_i).$$

Now for  $g \geq g_0$ , the condition (A.12) is not fulfilled in at most  $\bar{C} = (\sqrt{\frac{2MD}{m}} + 2)^D$  cubes surrounding each minimizer, and so in total only  $N\bar{C}$  cubes can survive each generation  $g > g_0$ . The important point is that this number is independent of  $\epsilon$ . So the total number of iterations is bounded by the sum of the total number of cubes in all generations  $g \leq g_0$ , which is some constant independent of  $\epsilon$  which we denote by  $b$ , and the constant  $N\bar{C}$  multiplied by the remaining number of iterations  $g_f - g_0$ , that is

$$n_2 \leq b + (g_f - g_0)N\bar{C} \leq n_2 \leq b + g_f N\bar{C} \stackrel{(\text{A.5})}{\leq} b + N\bar{C}(1/2 \log \frac{C_2}{\epsilon} + 1)$$

This bound can be replaced with a bound of the form (16) with an appropriate constant.  $\square$

**Theorem 3.** Let  $\delta > 0, r \in \mathbb{R}^D$  and  $r_*$  be a global minimizer of  $F_{\text{bi}}$ , and assume  $\|r - r_*\| \leq \delta$ . Let  $\sigma_{\mathcal{P}}, \sigma_{\mathcal{Q}}$  denote the Frobenius norm of the matrices whose columns are the points in  $\mathcal{P}$  and  $\mathcal{Q}$  respectively. Then  $\Delta_*(\delta)$  is given by

$$F_{\text{bi}}(r) - F_{\text{bi}}(r_*) \leq \Delta_*(\delta) \equiv \frac{2}{n} \sigma_{\mathcal{P}} \sigma_{\mathcal{Q}} \psi_2(\delta) \quad (\text{18})$$

*Proof.* To conclude the proof of the theorem for the case  $r_* = 0$  we need to show

**Lemma 6.** For all  $r \in \mathbb{R}^D$ ,

$$\|[r]\|_{\text{op}} \leq \|r\| \quad (\text{A.13})$$

*Proof.* The non-zero eigenvalues  $\lambda_j$  of a skew-symmetric real matrix  $[r]$  can be written as

$$a_1 i, -a_1 i, a_2 i, -a_2 i, \dots$$

where  $a_1 \geq a_2 \dots > 0$ . Therefore

$$\|[r]\|_{\text{op}}^2 = a_1^2 \leq 1/2 \sum_i |\lambda_i|^2 = 1/2 \|[r]\|_F^2 = \|r\|^2.$$

$\square$

For the general case  $r_* \neq 0$ , we define a change of variable  $\tilde{p}_i = R_{r_*} p_i$  and denote by  $\tilde{E}_{\text{bi}}$  the energy resulting by replacing  $p_i$  with  $\tilde{p}_i$  in the definition of  $E_{\text{bi}}$ . Then for all  $R_0, \pi$  we have

$$\tilde{E}_{\text{bi}}(R_0 R_{r_*}^T, \pi) = E_{\text{bi}}(R_0, \pi).$$

In particular  $\tilde{r}_* = 0$  is a minimizer of  $\tilde{F}_{\text{bi}}$  which is defined by replacing  $E_{\text{bi}}$  with  $\tilde{E}_{\text{bi}}$  in the definition of  $F_{\text{bi}}$ . We claim that there exists  $r_1$  such that

$$R_{r_1} = R_r R_{r_*}^T \text{ and } \|r_1\| \leq \|r - r_*\| \quad (\text{A.14})$$

In the case  $d = 2$  we can identify  $R_r R_{r_*}^T$  with  $e^{i(r-r_*)}$  and so we can simply choose  $r_1 = r - r_*$ . For  $d = 3$  it is proven in Lemma 3.2 in [18] that the angular distance between  $R_r$  and  $R_{r_*}$  is smaller or equal to  $\|r - r_*\|$ . As the angular distance is invariant to multiplication by rotations this means that the angular distance between  $R_r R_{r_*}^T$  and the identity is less than  $\|r - r_*\|$ . Since the exponential map is a radial isometry in  $B_\pi(0)$  this implies the existence of  $r_1$  satisfying (A.14).

Now for every  $\delta$  such that  $\|r - r_*\| \leq \delta$ , we using the bound from (18) for  $\tilde{F}$  which is minimized at zero, and satisfies  $\|r_1 - 0\| \leq \delta$ , to obtain

$$F_{\text{bi}}(r) - F_{\text{bi}}(r_*) = \tilde{F}_{\text{bi}}(r_1) - \tilde{F}_{\text{bi}}(0) \leq \frac{2}{n} \psi_2(\delta) \sigma_{\tilde{\mathcal{P}}} \sigma_{\mathcal{Q}} = \frac{2}{n} \psi_2(\delta) \sigma_{\mathcal{P}} \sigma_{\mathcal{Q}} .$$

□

**Theorem 4.** *Let  $(r_*, t_*)$  be a minimizer of  $F_{\text{CP}}$ , and let  $(r, t) \in \mathbb{R}^s \times \mathbb{R}^d$ , and  $\delta_1, \delta_2 > 0$  which satisfy  $\|r - r_*\| \leq \delta_1$  and  $\|t - t_*\| \leq \delta_2$ . Let  $f_*$  be some upper bound for the global minimum of  $F_{\text{CP}}$ . Then*

$$F_{\text{CP}}(r, t) - F_{\text{CP}}(r_*, t_*) \leq \Delta_*(\delta_1, \delta_2) \quad (20)$$

where

$$\begin{aligned} \Delta_*(\delta_1, \delta_2) = & \frac{1}{n} \left[ 2\psi_2(\delta_1) (\sigma_{\tilde{\mathcal{P}}}^2 + \sigma_{\mathcal{P}} \sqrt{nf_*}) \right. \\ & \left. + 2\delta_2 \psi_1(\delta_1) \sum_i \|p_i\| + n\delta_2^2 \right] \end{aligned} \quad (21)$$

*Proof.* The proof is very similar to the proof of Theorem 3. Let us first consider the case  $(r_*, t_*) = (0, 0)$ , and let  $\pi_*$  be the corresponding mapping so that  $(I_d, 0, \pi_*)$  minimizes  $E_{\text{CP}}$ . Then

$$\begin{aligned} F_{\text{CP}}(r, t) - F_{\text{CP}}(0, 0) & \leq E_{\text{CP}}(r, t, \pi_*) - E_{\text{CP}}(0, 0, \pi_*) \\ & = \frac{1}{n} \sum_{i=1}^n [2\langle (I_d - R_r)p_i, q_{\pi_*(i)} \rangle + 2\langle R_r p_i, t \rangle - 2\langle t, q_{\pi_*(i)} \rangle + \|t\|^2] \\ & =^{(*)} \frac{1}{n} \sum_{i=1}^n \left[ 2 \sum_{k=2}^{\infty} \frac{1}{k!} \langle [r]^k p_i, q_{\pi_*(i)} \rangle + 2 \sum_{k=1}^{\infty} \frac{1}{k!} \langle [r]^k p_i, t \rangle + \|t\|^2 \right] \\ & \leq \frac{1}{n} \sum_{i=1}^n \left[ 2 \sum_{k=2}^{\infty} \frac{1}{k!} \|r\|^k \|p_i\| (\|p_i\| + \|q_{\pi_*(i)} - p_i\|) + 2 \sum_{k=1}^{\infty} \frac{1}{k!} \|r\|^k \|p_i\| \|t\| + \|t\|^2 \right] \\ & \leq \frac{1}{n} \left[ 2\psi_2(\delta_1) (\sigma_{\tilde{\mathcal{P}}}^2 + \sigma_{\mathcal{P}} [\sum_i \|q_{\pi_*(i)} - p_i\|^2]^{1/2}) + 2\psi_1(\delta_1) \delta_2 \sum_i \|p_i\| + n\delta_2^2 \right] \\ & \leq \frac{1}{n} \left[ 2\psi_2(\delta_1) (\sigma_{\tilde{\mathcal{P}}}^2 + \sigma_{\mathcal{P}} \sqrt{nf_*}) + 2\psi_1(\delta_1) \delta_2 \sum_i \|p_i\| + n\delta_2^2 \right] \end{aligned}$$

Here (\*) follows from the fact that  $E_{\text{CP}}(\cdot, \cdot, \pi_*)$  is minimized at the origin and so the first order terms cancel out, and the next inequalities follow from the Cauchy-Schwarz inequality and from Lemma 6.

For general  $(r_*, t_*)$ , we use a change of variables  $\tilde{p}_i = R_* p_i$ ,  $\tilde{q}_i = q_i - t_*$ , and denote by  $\tilde{E}_{\text{CP}}$  and  $\tilde{F}_{\text{CP}}$  the functions obtained by replacing  $p_i, q_i$  by  $\tilde{p}_i, \tilde{q}_i$  in the definition of these functions. For given  $(r, t, \pi)$  we have

$$E_{\text{CP}}(R_r, t, \pi) = \tilde{E}_{\text{CP}}(R_r R_*^T, t - t_*, \pi)$$

For  $r, r_*, t, t_*$  satisfying  $\|r - r_*\| \leq \delta_1$  and  $\|t - t_*\| \leq \delta_2$ , we choose  $r_1 \in \mathbb{R}^D$  satisfying (A.14), and so we can applying the theorem to  $\tilde{F}_{\text{CP}}$  which is minimized at  $(0, 0)$  to obtain

$$\begin{aligned} F_{\text{CP}}(r, t) - F_{\text{CP}}(r_*, t_*) & = \tilde{F}_{\text{CP}}(r_1, t - t_*) - \tilde{F}_{\text{CP}}(0, 0) \\ & \leq \frac{1}{n} \left[ 2\psi_2(\delta_1) (\sigma_{\tilde{\mathcal{P}}}^2 + \sigma_{\mathcal{P}} \sqrt{nf_*}) + 2\psi_1(\delta_1) \delta_2 \sum_i \|p_i\| + n\delta_2^2 \right] \end{aligned}$$

□

## B. BnB for rigid closest point

In the following we explain how we construct a quasi-BnB framework for the rigid CP problem based on the BnB architecture proposed in Go-ICP [33]. Following Go-ICP, we use a nested BnB structure: We perform an “outer” BnB search on the rotation space, wherein the upper and lower bounds are functions of the translation component  $t$ ; in turn, to compute these bounds we perform an “inner” BnB search over the variable  $t$ . Namely, for given  $r_i$ , we define an upper bound for the outer BnB by

$$\bar{E}_{\text{CP}}(r_i) = \min_{t \in \mathcal{C}_1(0), \pi \in \Pi_{\text{CP}}} E_{\text{CP}}(r_i, t, \pi). \quad (\text{B.1})$$

To compute a quasi-lower bound for the outer BnB we note that if  $(r_*, t_*)$  minimizes  $F_{\text{CP}}$  and  $r_* \in \mathcal{C}_h(r_i)$ , then by using (21) where we set  $r = r_i$ , take  $\delta_1$  to be the maximal distance of a point in the cube from the center,  $t = t_*$  and  $\delta_2 = 0$  we obtain

$$F_{\text{CP}}(r_*, t_*) \geq F_{\text{CP}}(r_i, t_*) - \frac{2}{n} \left( 1 + \sqrt{\frac{f_*}{\sigma_{\mathcal{P}}}} \right) \sigma_{\mathcal{P}} \psi_2(\sqrt{D}h), \quad (\text{B.2})$$

and since  $F_{\text{CP}}(r_i, t_*) \geq \bar{E}_{\text{CP}}(r_i)$  it follows that if  $r_* \in \mathcal{C}_h(r_i)$  then

$$F_{\text{CP}}(r_*, t_*) \geq \bar{E}_{\text{CP}}(r_i) - \frac{2}{n} \left( 1 + \sqrt{\frac{f_*}{\sigma_{\mathcal{P}}}} \right) \sigma_{\mathcal{P}} \psi_2(\sqrt{D}h). \quad (\text{B.3})$$

The RHS of the equation above gives us our quasi lower bound for the rotation quasi BnB. To compute  $\bar{E}_{\text{CP}}(r_i)$  we compute a BnB in translation space, where throughout the translation BnB the rotation coordinate  $r_i$  is fixed. For a given translation cube  $\mathcal{C}_h(t_j)$  an upper bound for the value of  $\bar{E}_{\text{CP}}(r_i)$  is given by evaluation of  $F_{\text{CP}}(r_i, t_j)$ . If  $t_*$  is a minimizer of  $F_{\text{CP}}(r_i, \cdot)$  then a quasi-lower bound in the cube is given by

$$E_{\text{CP}}(r_i, t_*) \geq E_{\text{CP}}(r_i, t_j) - \frac{dh^2}{n}. \quad (\text{B.4})$$

We note that this bound is similar to what we would get by setting  $\delta_1 = 0$  and  $\delta_2$  to be the maximal distance in the cube from  $t_j$  in (21). Although the bound does not follow directly from this equation the derivation is similar, and can be obtained by studying the behavior of a minimizer of  $E(r_i, \cdot, \pi_*)$ , so we do not go into the details. Finally we note that when the quasi-lower bounds in the outer or inner BnB is lower than zero we replace it with zero.

The rest of the architecture of the BnB is also borrowed from Go-ICP. We use best-first-search, where the cube with the lowest lower bound is visited first. Every time the upper bound is improved, an ICP algorithm is run to improve the resolution of the solution. For more details see [33].

## C. Morphological data

The morphological data for the experiment shown in Figure 5 comes from the MorphoSource dataset. The figure shows ten different second mandibular molars of spider monkeys (*Ateles*), which come from three different taxonomical groups. More details on the data are shown in the table below.

ID	specimen #	specimen taxonomy	ark ID
M782-661	AMNH:M:67102	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M661">http://n2t.net/ark:/87602/m4/M661</a>
M783-663	AMNH:M:71787	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M663">http://n2t.net/ark:/87602/m4/M663</a>
M785-665	AMNH:M:76882	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M665">http://n2t.net/ark:/87602/m4/M665</a>
M787-666	USNM:241384	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M666">http://n2t.net/ark:/87602/m4/M666</a>
M788-668	USNM:406674	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M668">http://n2t.net/ark:/87602/m4/M668</a>
M790-669	USNM:406675	Ateles belzebuth	<a href="http://n2t.net/ark:/87602/m4/M669">http://n2t.net/ark:/87602/m4/M669</a>
M791-671	MCZ:34320	Ateles geoffroyi	<a href="http://n2t.net/ark:/87602/m4/M671">http://n2t.net/ark:/87602/m4/M671</a>
M793-673	MCZ:mamm:bom-5344	Ateles geoffroyi	<a href="http://n2t.net/ark:/87602/m4/M673">http://n2t.net/ark:/87602/m4/M673</a>
M795-675	USNM:mammals:336204	Ateles geoffroyi	<a href="http://n2t.net/ark:/87602/m4/M675">http://n2t.net/ark:/87602/m4/M675</a>
M797-677	MCZ:31759	Ateles paniscus	<a href="http://n2t.net/ark:/87602/m4/M677">http://n2t.net/ark:/87602/m4/M677</a>