A Proofs

We provide the proofs for the following statements from Section 4.2:

1. Given an optimal reference set, the error of CID is proportional to the maximal distance of adjacent points in the reference set.
2. With the increasing number of points provided in an optimal reference set, the CID's error tends to zero.

A.1 Proof Statement 1

We start with defining two reference sets G and F, obtained by performing sampling within the domain $D \subset \mathbb{R}^n$ on a coarse and fine grids respectively. Let be S the number of all separated regions of the DOI, then, since the reference sets are optimal, both reference sets witness every separated continuous region $r_s,\ s = 1, \dots, S$.

From the Definition of CID, the following holds:

$$CID(A, Z) = \frac{1}{|Z|} \sum_{z \in Z} d_z = \frac{1}{|Z|} \sum_{s=1}^{S} \sum_{z \in (Z \cap r_s)} d_z \tag{1}$$

, i.e.,

$$CID(A, Z) = \sum_{s=1}^{S} CID(A, Z \cap r_s) \tag{2}$$

Computing CID for the test set A, using the reference sets G and F, we obtain:

$$CID_G = \frac{1}{|G|} \sum_{i=1}^{|G|} d_i^G \tag{3}$$

$$CID_F = \frac{1}{|F|} \sum_{j=1}^{|F|} d_i^F \tag{4}$$

From Equation 2, w.l.o.g., let's assume that $S = 1$, and let's denote r_1 as r. The difference between the CID values for the corresponding reference sets is then:

$$CID_F - CID_G = \frac{1}{|G|} \sum_{i=1}^{|G|} d_i^G - \frac{1}{|F|} \sum_{j=1}^{|F|} d_i^F \tag{5}$$

We partition the continuous region r using a Voronoi diagram built on G into $|G|$ Voronoi cells c_i, $i = 1, \dots, |G|$. Then, we assign to each reference point $f_j \in F$ a Voronoi cell c_i if the reference point lies within the cell, and denote the reference point as f_k^i. If f_j lies on a boundary of a cell, we assign it to any of the neighbouring cells. Let be R^* the maximal distance of adjacent points in G. Each of the cells lies within a ball $B(g_i; R_i)$, and the radius of each of the balls is bounded by R^*, i.e. $R_i \leq R^*$. As f_k^i belongs to the cell c_i with the center g_i, the distance between them is bounded: $\|\vec{\boldsymbol{\xi_{ik}}}\| \leq R^*$, where $\vec{\boldsymbol{\xi_{ik}}}$ is the relative position between f_k^i, and g_i.

By applying the triangle equality of vector addition for the relative positions from the reference points $f_j \equiv f_i^k$ and g_i to the test point yields:

$$\vec{\boldsymbol{d_j^F}} = \vec{\boldsymbol{d_i^G}} + \vec{\boldsymbol{\xi_{ik}}} \tag{6}$$

Using Equation 5, CID_F from Equation 4 can be rewritten as:

$$CID_F = \frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k=1}^{K_i} \|\vec{\boldsymbol{d_i^G}} + \vec{\boldsymbol{\xi_{ik}}}\|, \tag{7}$$

where K_i is the number of points f_k^i belonging to c_i.

With Equations 7 and 5, the following holds:

$$CID_F - CID_G = \frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k=1}^{K_i} \|\vec{\boldsymbol{d_i^G}} + \vec{\boldsymbol{\xi_{ik}}}\| - \frac{1}{|G|} \sum_{i=1}^{|G|} \|\vec{\boldsymbol{d_i^G}}\| \tag{8}$$

The norm $\|\cdot\|$ satisfies the following inequalities:

- Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$
- Reverse triangle inequality: $\|x + y\| \geq |\|x\| - \|y\||$

Applying the triangle inequality to Equation 8 yields:

$$\begin{aligned} CID_F - CID_G \leq \frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k=1}^{K_i} \|\vec{\boldsymbol{d_i^G}}\| - \frac{1}{|G|} \sum_{i=1}^{|G|} \|\vec{\boldsymbol{d_i^G}}\| + \frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k=1}^{K_i} \|\vec{\boldsymbol{\xi_{ik}}}\| \underset{*}{=} \\ \frac{1}{|F|} \sum_{i=1}^{|G|} (K_i - K^*) \|\vec{\boldsymbol{d_i^G}}\| + \frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k=1}^{K_i} \|\vec{\boldsymbol{\xi_{ik}}}\| \overset{|F| \to \infty}{\underset{**}{\leq}} C(|G|) + R^*. \end{aligned} \tag{9}$$

Note:

- $*$: For the equality, K^* is the average number of points f_k^i within a cell c_i, for $i = 1, \dots |G|$, i.e. $K^* = |F|/|G|$.

** : For the inequality, presuming we have a uniform reference set F of an infinite size, the first sum represents the deviation from average of the volume of each Voronoi cell. As shown in [**empty citation**], the deviation reduces with the increasing number of points. Furthermore, the reduction is even higher for higher dimensions [**empty citation**]. Note, that for a reference set obtained using a structured grid, the variance of the volume is equal to zero, as the volumes are equal for each cell, which means that the first sum is equal to zero. In other words, the first sum represents the uniform properties of the reference sets, and the more uniform the reference set, the smaller the variance is. We presume that the set of sampled points, and therefore the reference set G are uniform enough to neglect the first sum. The second sum is bounded, as the size of each cell is bounded by the respective radius of a ball.

Applying the reverse triangle inequality for Equation 8, the following holds:

$$CID_F - CID_G \geq \frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k=1}^{K_i}|d_i^G - \|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\|| - \frac{1}{|G|}\sum_{i=1}^{|G|} d_i^G \underset{*}{=}$$
$$\frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k=1}^{K_i^+}(d_i^G - \|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\|) + \frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k=1}^{K_i^-}(\|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\| - d_i^G) - \frac{1}{|F|}\sum_{i=1}^{|G|} K^* d_i^G = \tag{10}$$
$$\frac{1}{|F|}\sum_{i=1}^{|G|}(K_i^+ - K^*)d_i^G - \frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k}^{K_i^+}\|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\| + \frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k}^{K_i^-}(\|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\| - d_i^G)$$

Note:

* : For the equality, we separate the sum of the moduli over the reference points within the cell into two sums: where the expression under the modulus is positive (K_i^+ terms), and negative (K_i^- terms).

Now, we consider each of the sums from Equation 10 separately. For the second sum holds:

$$\frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k}^{K_i^+}\|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\| \leq \frac{1}{|F|}\sum_{i=1}^{|G|}\sum_{k}^{K_i}\|\vec{\boldsymbol{\xi}_{\boldsymbol{ik}}}\| \leq R^*, \tag{11}$$

as the size of each cell is bounded by the respective radius of a ball.

To assess the rest of the sums, let us study the dependence of K_i^- and K_i^+ on R^*. For a fixed test set, the following holds:

$$\exists G_0 : \forall G : |G| > |G_0| \implies K_i^- \leq R^* \cdot K^*. \tag{12}$$

As $K_i^+ + K_i^- = K_i$, and for a uniform structured grid, $K_i = K^*$, the following holds:

$$K_i^+ = K^* - K_i^- \geq K^* - R^* \cdot K^* = (1 - R^*) \cdot K^* \tag{13}$$

Then, from Equation 12, for the third sum holds:

$$\frac{1}{|F|} \sum_{i=1}^{|G|} \sum_{k}^{K_i^-} (\|\vec{\boldsymbol{\xi_{ik}}}\| - d_i^G) \leq C' \cdot \frac{1}{K^*} \cdot R^* \cdot K^* = C^{'} \cdot R^*;\ C' > 0. \tag{14}$$

Similarly, for the first sum in the Equation 13 holds:

$$\left| \frac{1}{|F|} \sum_{i=1}^{|G|} (K_i^+ - K^*) \cdot d_i^G \right| \leq C'' \cdot \frac{1}{K^*} \cdot R^* \cdot K^* = C'' \cdot R^*;\ C'' > 0. \tag{15}$$

Then for the Equation 10 holds:

$$CID_F - CID_G \geq -(C' + C'' + 1) \cdot R^* \geq -C''' \cdot R^*;\ C''' > 0. \tag{16}$$

Finally, from Equation 16 and Equation 9 follows:

$$|CID_F - CID_G| \leq C^* \cdot R^*,\ C^* = \max(C''', C). \tag{17}$$

Equation 17 implies that the error decreases in a linear fashion when the maximal distance between adjacent points in the reference set decreases, which proofs the Statement 1.

A.2 Proof Statement 2

Given the fact, that for an optimal reference set, the maximum distance between adjacent points R^* decreases with the increasing number of provided points in the reference set (s. Equation 17), the Statement 2 follows.