kth Nearest Neighbor Sampling in the Plane*

Kirk P. Gardner
University of Connecticut
kirk.gardner@uconn.edu

Donald R. Sheehy
University of Connecticut
don.r.sheehy@gmail.com

Abstract

Let B be a square region in the plane. We give an efficient algorithm that takes a set P of n points from B, and produces a set $M \subset B$ with the property that the distance to the second nearest point in M approximates the distance to the kth nearest point of P. That is, there are constants $\alpha, \beta \in \mathbb{R}$ such that for all $x \in B$, we have $\alpha \mathbf{d}_{P,k}(x) \leq \mathbf{d}_{M,2}(x) \leq \beta \mathbf{d}_{P,k}(x)$, where $\mathbf{d}_{M,2}$ and $\mathbf{d}_{P,k}$ denote the second nearest and kth nearest neighbor distance functions to M and P respectively. The algorithm is based on Delaunay refinement. The output set M also has the property that its Delaunay triangulation has a lower bound on the size of the smallest angle. The primary application is in statistical density estimation and robust geometric inference.

1 Robust Sizing Functions

Since the pioneering work of Chew [3] and Ruppert [9], Delaunay refinement has remained an important approach to mesh generation (see for example the book [2]). The algorithm: Starting from the Delaunay triangulation of the input points P (restricted to a bounding box B), repeatedly add the circumcenter of any triangle whose circumradius is more than a constant times larger than the length if its shortest edge.[1] Such a triangulation is said to have *bounded radius-edge ratio* and will be referred to as a *quality mesh* and will necessarily also have a lower bound on the size of the smallest angle. Ruppert showed that not only does this remarkably simple algorithm terminate, it produces a point set that is asymptotically optimal in size [9].

The key to Ruppert's analysis is the so-called *feature size function*, which for a point set P is the distance to the second nearest point of P, denoted $\mathbf{d}_{P,2} : \mathbb{R}^2 \to \mathbb{R}$. There is a constant γ such that output set M has the property that

$$\gamma \mathbf{d}_{P,2} \leq \mathbf{d}_{M,2} \leq \mathbf{d}_{P,2}.$$

The optimality of the approach comes from proof that if M is the vertex set of a quality mesh containing P, then

$$|M| = \Theta\left(\int_{x\in B} \frac{dx}{\mathbf{d}_{P,2}(x)^2}\right).$$

The preceding integral defines a measure, whose density is related to the *second nearest neighbor density estimator* used in statistics. It is a useful feature of Delaunay refinement that it reveals this function without explicitly computing or estimating it. However, for the convergence of such estimators, one must generally choose k as a function of $n = |P|$ so that $k(n)/n \to 0$ and $k(n)/log(n) \to \infty$ as $n \to \infty$ [5]. So, for example, taking $k(n) = \sqrt{n}$ is reasonable and sufficient. Motivated by this relationship between mesh generation and density estimation, we address the following problem.

Problem 1 *Given a set P in a bounding box $B \subset \mathbb{R}^2$, find a quality mesh with vertex set $M \subset B$ such that*

$$\alpha \mathbf{d}_{P,k} \leq \mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$$

for some constants α and β.

We show how to solve this problem using a variation of Delaunay refinement.

2 Background

We will denote the Euclidean distance between points $a, b \in \mathbb{R}^2$ as $\|a - b\|$. For any set $S \subset \mathbb{R}^2$ and integer $k \geq 1$, define $\mathbf{d}_{S,k} : \mathbb{R}^2 \to \mathbb{R}$ so that $\mathbf{d}_{S,k}(x)$ is the distance to the kth nearest point of S to $x \in \mathbb{R}^2$. Formally, letting $\binom{S}{k}$ denote the set of k element subsets of S,

$$\mathbf{d}_{S,k}(x) := \min_{U \in \binom{S}{k}} \max_{y \in U} \|x - y\|.$$

The distance from a point $x \in \mathbb{R}^2$ to a set S is $\mathbf{d}_{S,1}(x)$ and will be denoted $\mathbf{d}_S(x)$.

We define the ball centered at a point $x \in \mathbb{R}^2$ with radius r as

$$\mathtt{ball}(x, r) := \{y \in \mathbb{R}^2 \mid \|x - y\| \leq r\}.$$

The smallest ball centered at a point $x \in \mathbb{R}^2$ containing k points in a set S will be denoted

$$\mathtt{ball}_{x,S_k} := \mathtt{ball}(x, \mathbf{d}_{S,k}(x)).$$

*This work was partially supported under grants CCF-1464379 and CCF-1525978.

[1] Ruppert's analysis also works for more general inputs including line segments as well.

For a set $S \subset \mathbb{R}^2$ we will define a triangle with vertices $u, v, w \in S$ as the convex closure of the points u, v, w:

$$t_{u,v,w} := \{au + bv + cw \mid a + b + c = 1,\ a, b, c \in \mathbb{R}_{\geq 0}\}$$

The **circumcenter** of a triangle $t = t_{u,v,w}$ is the unique point that is equidistant to each vertex u, v, w, and will be denoted $\mathtt{cc}(t)$. This distance to each vertex u, v, w is the **circumradius** of t, and will be denoted $\mathtt{rad}(t)$ so that

$$\mathtt{rad}(t) := \|\mathtt{cc}(t) - u\| = \|\mathtt{cc}(t) - v\| = \|\mathtt{cc}(t) - w\|$$

The **circumcircle** of t is the smallest ball containing the points u, v, w, and will be denoted $\mathtt{cball}(t)$. Formally,

$$\mathtt{cball}(t) := \mathtt{ball}(\mathtt{cc}(t), \mathtt{rad}(t)).$$

2.1 Delaunay Triangulations and Voronoi Diagrams

Definition 1 *For a set $S \subset \mathbb{R}^2$ the **Delaunay triangulation** of S is the set of triangles $t = t_{u,v,w}$ such that no point $p \in S \backslash \{u, v, w\}$ is contained in the circumcircle of t and is denoted*

$$\mathtt{Del}_S := \{t_{u,v,w} \mid \mathtt{cball}(t_{u,v,w}) \cap S = \{u, v, w\}\}$$

The set of Delaunay vertices of $\mathtt{Del}_S$ is the set S itself.

The **Voronoi cell** of a point $u \in S$ is the set of points $x \in \mathbb{R}^2$ that are closer to u than any other point in S.

$$\mathtt{Vor}_S(u) := \{x \in \mathbb{R}^2 \mid \mathbf{d}_S(x) = \|u - x\|\}.$$

Definition 2 *For a set $S \subset \mathbb{R}^2$ the **Voronoi diagram** of S is the set of all Voronoi cells of the points in S and is denoted*

$$\mathtt{Vor}_S := \{\mathtt{Vor}_S(u) \mid u \in S\}.$$

For a set $S \subset \mathbb{R}^2$ the Voronoi diagram $\mathtt{Vor}_S$ is dual to the Delaunay triangulation $\mathtt{Del}_S$. That is, Delaunay vertices $u \in S$ correspond to Voronoi cells (faces) $\mathtt{Vor}_S(u)$ and the Delaunay triangles $t \in \mathtt{Del}_S$ correspond to **Voronoi vertices**, defined to be the circumcenters $\mathtt{cc}(t)$. The set of Voronoi vertices corresponding to a point $u \in S$ will be denoted

$$\mathtt{Vor}^0_S(u) := \{\mathtt{cc}(t_{u,v,w}) \in \mathtt{Vor}_S(u) \mid t_{u,v,w} \in \mathtt{Del}_S\}.$$

The **Voronoi edge** corresponding to points $u, v \in S$ is the intersection of the Voronoi cells of u and v and will be denoted

$$\mathtt{Vor}_S(u, v) := \mathtt{Vor}_S(u) \cap \mathtt{Vor}_S(v).$$

For any $u, v \in M$ such that $\mathtt{Vor}_S(u, v) \neq \emptyset$ the Voronoi edge $\mathtt{Vor}_S(u, v)$ corresponds to an edge of the Delaunay triangulation, defined to be the convex closure of the points u and v.

The **outradius** of a Voronoi cell $\mathtt{Vor}_S(u)$ is the radius of the smallest ball centered at u containing $\mathtt{Vor}^0_S(u)$, and will be denoted

$$\mathtt{R}(u) := \max_{c \in \mathtt{Vor}^0_S(u)} \|c - u\|$$

The outradius of u is the distance to the farthest Voronoi vertex of $\mathtt{Vor}_S(u)$, denoted

$$\mathtt{farCorner}(\mathtt{Vor}_S(u)) := \operatorname*{argmax}_{c \in \mathtt{Vor}^0_S(u)} \|c - u\|.$$

Definition 3 *The **aspect ratio** of a Voronoi cell $\mathtt{Vor}_S(u)$ is the ratio of the distance to its farthest corner to the distance to its nearest edge and is denoted*

$$\mathtt{aspect}(\mathtt{Vor}_S(u)) := \frac{2\mathtt{R}(u)}{\mathbf{d}_{S,2}(u)}.$$

A set $S \subset \mathbb{R}^2$ is said to be **τ-well spaced** if

$$\mathtt{aspect}(\mathtt{Vor}_S(u)) \leq \tau$$

for all $u \in S$.

2.2 Periodic Point Sets

To avoid additional boundary conditions we will work in a covering space of the **flat torus**, which can be defined as $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (see [1]). That is, we will restrict ourselves to a **bounding box** $B = [0, 1)^2$ and use copies of a set $S \subset B$ to simulate periodicity.

For any finite point set $S \subset B$ we will define the corresponding **periodic point set** as the set $S + \mathbb{Z}^2$ imbued with an equivalence relation so that $x \sim y$ if there exists some $a \in \mathbb{Z}^2$ such that $y = x + a$. Noting that $s \in P$ for all $s \in S$ we may therefore refer to the set of periodic copies of s by the equivalence class

$$[s] := \left\{x \in S + \mathbb{Z}^2 \mid \exists a \in \mathbb{Z}^2 : x = s + a\right\}.$$

3 Algorithm

For a bounding box $B = [0, 1)^2$ the following algorithm will use periodic points $x \in B + \mathbb{Z}^2$ to denote equivalence classes $[x]$ of which $x \in B$ are representative. The use of periodic points is a technical requirement of the algorithm in order to avoid additional points along the edge of the bounding box. As our analysis does not depend on the use of periodic points we will return to the original sets in the following sections.

Algorithm 1 takes as input a finite point set $P \subset B$, a set of initial mesh vertices $M_0 \subset B$ which we will take as a set of points arranged along a square in B, and constants $\tau \geq 2$, $k \geq 1$. The algorithm constructs a periodic set of mesh vertices $M \subset B + \mathbb{Z}^2$ satisfying

$$\mathtt{aspect}(\mathtt{Vor}_M(v)) \leq \tau$$

for all $v \in M$ and

$$|\texttt{Vor}_M(v) \cap P| < k, \ \mathbf{d}_{P,k}(\texttt{cc}(t)) > \texttt{rad}(t)$$

for all $v \in M$, $t \in \texttt{Del}_M$.

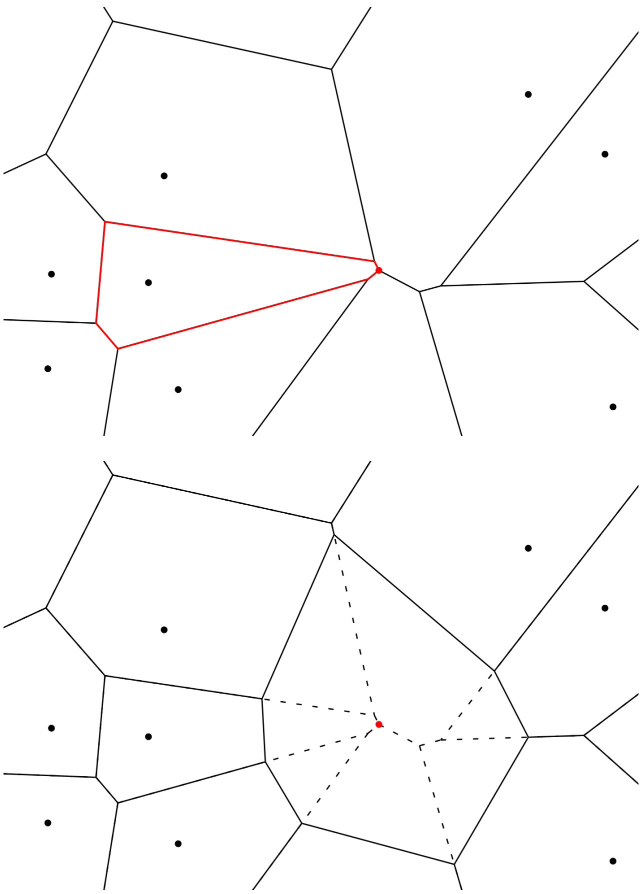

Figure 1: An application of the CLEAN procedure which chooses a point $v \in M$ with $\texttt{aspect}(\texttt{Vor}_M(v)) > \tau$ (red cell, top) and adds $\texttt{farCorner}(\texttt{Vor}_M(v))$ (red point, bottom).

Algorithm 1 $\text{KNNREFINE}(P, M_0, \tau, k)$

```
M_0 ← M_0 + Z², P ← P + Z², M ← M_0
while there is a v ∈ M or t ∈ Del_M such that
    d_{P,k}(cc(t)) ≤ rad(t) or |Vor_M(v) ∩ P| ≥ k do
    M ← BREAK(M, P)
    while ∃v ∈ M with aspect(v) > τ do
        M ← CLEAN(M, v)
procedure BREAK(M, P)
    if ∃t ∈ Del_M with d_{P,k}(cc(t)) ≤ rad(t) then
        M ← M ∪ {cc(t)}
    else if ∃v ∈ M with |Vor_M(v) ∩ P| ≥ k then
        M ← M ∪ {farCorner(Vor_M(v))}
procedure CLEAN(M, v)
    M ← M ∪ {farCorner(Vor_M(v))}
```

Figure 1 depicts an application of the CLEAN procedure to a point $v \in M$ such that $\texttt{aspect}(\texttt{Vor}_M(v)) > \tau$, resulting in the insertion of $\texttt{farCorner}(\texttt{Vor}_M(v))$.

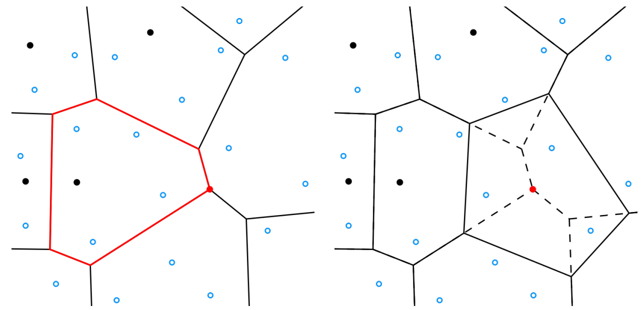

Figure 2: An example of the BREAK procedure applied to an instance with $v \in M$ such that $\texttt{Vor}_M(v)$ (red cell, left) containts at least $k = 4$ points in P (blue rings) initiating the insertion of $\texttt{farCorner}(\texttt{Vor}_M(v))$ (red point, right).

Figure 2 illustrates an application of the BREAK procedure to a Voronoi cell containing at least $k = 4$ points in P, initiating the insertion its farthest corner. Similarly, Figure 3 depicts an application of the BREAK procedure to a configuration in which a Delaunay circumcircle contains at least $k = 4$ points in P, initiating the insertion its circumcenter.

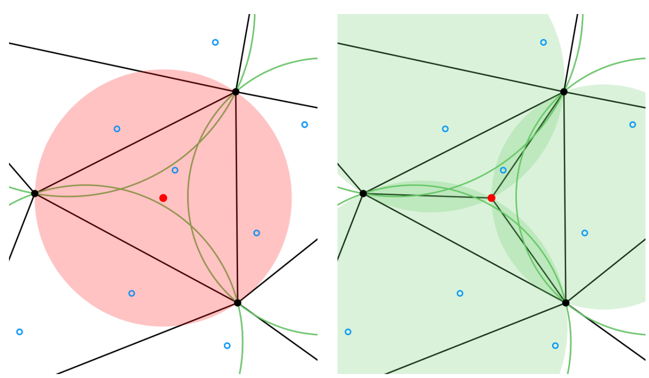

Figure 3: An example of the BREAK procedure applied to an instance with $t \in \texttt{Del}_M$ such that $\texttt{cball}(t)$ (red disk, left) contains at least $k = 4$ points in P (blue rings) initiating the insertion of $\texttt{cc}(t)$ (red point, right).

Restricting ourselves to the bounding box $B = [0, 1)^2$ the remainder of this section will provide upper and lower bounds on the second nearest neighbor function $\mathbf{d}_{M,2}$ in terms of the kth nearest neighbor function $\mathbf{d}_{P,k}$ in order to prove Theorem 6 stated below.

Theorem 6 (Main Theorem) *Let $P \subset B$ be a finite point set and $\tau \geq 2$, $k \geq 1$ be constants. Let $M_0 \subset B$ be a set of initial mesh vertices.*

If $\text{KNNREFINE}(P, M_0, \tau, k)$ terminates the resulting *set of mesh vertices M is τ-well spaced,*

$$|\mathtt{Vor}_M(v) \cap P| < k,\ \mathbf{d}_{P,k}(\mathtt{cc}(t)) > \mathtt{rad}(t)$$

for all $v \in M$, $t \in \mathtt{Del}_M$, and

$$\alpha \mathbf{d}_{P,k} \leq \mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$$

where $\alpha = \frac{\tau-2}{5\tau-2}$, $\beta = \frac{3-\vartheta_\tau}{1-\vartheta_\tau}$ and $\vartheta_\tau = \sqrt{1-\frac{1}{\tau^2}}$.

3.1 Lower Bound

We will first show that the second nearest neighbor function $\mathbf{d}_{M,2}$ to the output set $M \subset B$ is not too small compared to the kth-nearest neighbor function $\mathbf{d}_{P,k}$ to the input set $P \subset B$.

Let $M_0 \subseteq M$ be a set of input vertices and v_i be the ith circumcenter added to M. We define an order relation $\prec$ on M such that for all $i > 0$

$$u_0 \prec v_{i-1} \prec v_i$$

where $u_0 \in M_0$.

Definition 4 *For a set $M \subset B$ imbued with the order relation $\prec$ the* **insertion radius** *of a $v \in M$ is the distance from v to its nearest predecessor and is denoted*

$$\lambda_v := \min_{u \prec v} \|u - v\|.$$

The insertion radius λ_{v_i} of the ith vertex is the distance to the closest point in $M_i \subset M$ where $M_i = \{v_j \mid 0 < j < i\}$ is the ordered set of mesh vertices in M before the point v_i is added. Note that the insertion radius of a point v_i will be at most $\|v_i - v_j\|$ for all $v_j \in M_i$.

Lemma 1 shows that it suffices to bound the insertion radius for each $v \in M$ in order to bound $\mathbf{d}_{M,2}$ by $\mathbf{d}_{P,k}$ over M.

Lemma 1 *Let $K > 0$ be a constant.*
If $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ then

$$\mathbf{d}_{P,k}(v) \leq K\mathbf{d}_{M,2}(v)$$

for all $v \in M$.

Proof. Let $v \in M$ and $u \in M \setminus \{v\}$ be such that $\|u - v\| = \mathbf{d}_{M,2}(v)$. If $u \prec v$ then by assumption

$$\begin{aligned}\mathbf{d}_{P,k}(v) &\leq (K-1)\lambda_v < K\lambda_v \\ &\leq K\|u-v\| = K\mathbf{d}_{M,2}(v).\end{aligned}$$

So, we may assume $v \prec u$, which implies $\lambda_u \leq \|u - v\| = \mathbf{d}_{M,2}(v)$. Thus,

$$\begin{aligned}\mathbf{d}_{P,k}(v) &\leq \mathbf{d}_{P,k}(u) + \|u-v\| && [\mathbf{d}_{P,k} \text{ is Lipschitz}] \\ &\leq (K-1)\lambda_u + \|u-v\| && [\mathbf{d}_{P,k}(u) \leq (K-1)\lambda_u] \\ &\leq K\mathbf{d}_{M,2}(v). && [\lambda_u \leq d_{M,2}(v)]\end{aligned}$$

□

We now apply Lemma 1 to the kth-nearest neighbor function to provide a lower bound for $\mathbf{d}_{M,2}$ by induction on the set of mesh vertices M produced by Algorithm 1. We will set $\lambda_{u_0} \geq \mathbf{d}_{P,k}(u_0)$ for all $u_0 \in M_0$.

Lemma 2 *Let $M_0 \subset B$ be a set of initial mesh vertices such that $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$. For constants $\tau \geq 2$, $k \geq 1$ let $M \subset B$ be a set of mesh vertices imbued with the order relation $\prec$ resulting from* kNNRefine(P, M_0, τ, k).
If M is τ-well spaced then for all $v \in M$

$$\mathbf{d}_{P,k}(v) \leq K\mathbf{d}_{M,2}(v)$$

where $K = \frac{2\tau}{\tau-2}$.

Proof. Lemma 1 implies that it suffices to show that $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ for all $v \in M$. We will show this by induction on the number of circumcenters added.

Let M_i denote the set of mesh vertices in M before the ith circumcenter is added. In the base case we require $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$. It follows $\mathbf{d}_{P,k}(u_0) \leq (K-1)\lambda_{u_0}$ as $K \geq 2$.

Assume inductively that $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ for all $v \in M_i$, and note that Lemma 1 implies

$$\mathbf{d}_{P,k}(v) \leq K\mathbf{d}_{M_i,2}(v)$$

for all $v \in M_i$.

Let v_i be the ith circumcenter added and let $u \in M_i$ be the vertex whose Voronoi cell had poor quality, initiating the insertion of v_i. Letting $w \in M_i$ be the second nearest neighbor of u so that $w \neq u$ and $\|u - w\| = \mathbf{d}_{M_i,2}$ we have

$$\begin{aligned}\mathbf{d}_{P,k}(v_i) &\leq \mathbf{d}_{P,k}(u) + \|u - v_i\| && [\mathbf{d}_{P,k} \text{ is 1-Lipschitz}] \\ &\leq K\mathbf{d}_{M_i,2}(u) + \|u - v_i\| && [\text{Lemma 1 and hypothesis}] \\ &\leq K\|u-w\| + \|u-v_i\| && [\text{Definition of } w] \\ &\leq \|u-v_i\|\left(\frac{2K}{\tau}+1\right) && [\mathtt{aspect}(\mathtt{Vor}_{M_i}(u)) > \tau] \\ &\leq \|u-v_i\|(K-1) && \left[K = \frac{2\tau}{\tau-2}\right] \\ &\leq \lambda_{v_i}(K-1). && [\text{Definition of } u]\end{aligned}$$

As $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ for all $v \in M_i$ by our inductive hypothesis, $M_{i+1} = M_i \cup \{v_i\}$, and $\mathbf{d}_{P,k}(v_i) \leq \lambda_{v_i}(K-1)$ it follows that $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ for all $v \in M_{i+1}$. It follows by induction that $\mathbf{d}_{P,k}(v) \leq (K-1)\lambda_v$ for all $v \in M = \bigcup M_i$ by induction, and we may therefore conclude

$$\mathbf{d}_{P,k}(p) \leq K\mathbf{d}_{M,2}(p)$$

for all $p \in M$ by Lemma 1. □

Figure 4 illustrates the proof of Lemma 2, depicting the Voronoi cell $\mathtt{Vor}_{M_i}(u)$ with bad aspect ratio in the top figure, and the resulting Voronoi diagram $\mathtt{Vor}_{M_{i+1}}$ after the insertion of v_i on the bottom. Note, the insertion radius $\lambda_{v_i} = \|v_i - u\|$ of v_i satisfies $\min_{v \prec v_i} \|v - v_i\|$

as the closest point to v_i is the point u initiating its insertion. This fact allows for the inductive proof of Lemma 2 as for all i such that $\texttt{aspect}(\texttt{Vor}_{M_i}(u)) > \tau$ for some $u \in M_i$ the ith mesh vertex added to M is $v_i = \texttt{farCorner}(\texttt{Vor}_{M_i}(u))$ and $\lambda_{v_i} = \|v_i - u\|$.

Theorem 3 extends the bound on $\mathbf{d}_{M,2}$ over M provided by Lemma 2 to all points in B for τ-well spaced sets M.

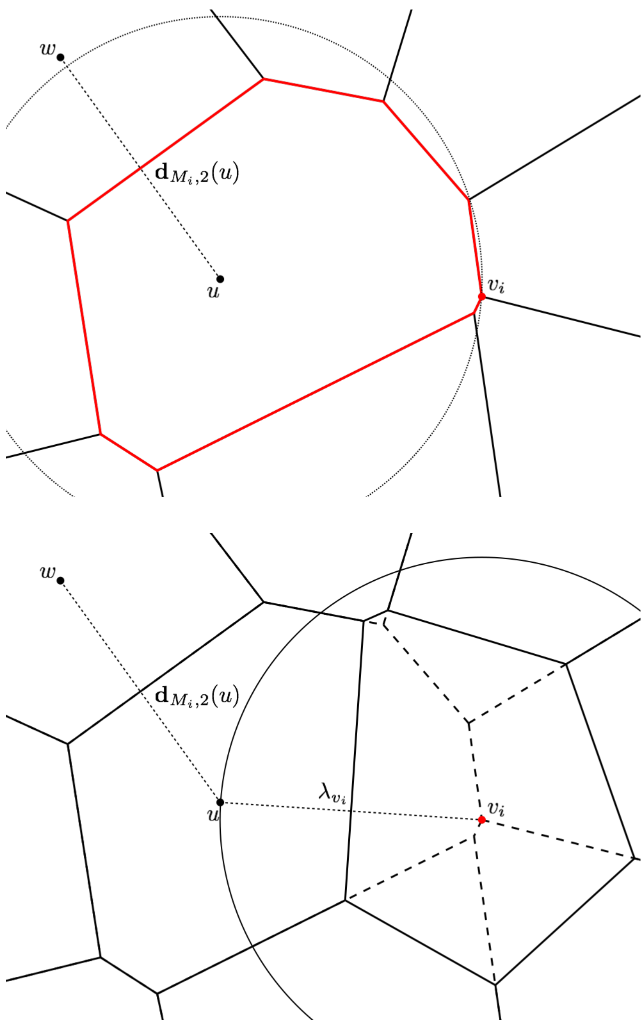

Figure 4: An illustration of Lemma 2 in which $\texttt{aspect}(\texttt{Vor}_{M_i}(u)) > \tau$ in $\texttt{Vor}_{M_i}$ (red cell, top) and the insertion radius λ_{v_i} of $v_i = \texttt{farCorner}(\texttt{Vor}_{M_i}(u))$ is the distance from u to v_i (red point, bottom).

Theorem 3 *Let $P \subset B$ be a finite point set and let $M_0 \subset B$ be a set of initial mesh vertices such that $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$. For constants $\tau \geq 2$, $k \geq 1$ let $M \subset B$ be a set of mesh vertices imbued with the order relation $\prec$ resulting from* kNNRefine(P, M_0, τ, k).

If M is τ-well spaced then for all $x \in B$

$$\alpha \mathbf{d}_{P,k}(x) \leq \mathbf{d}_{M,2}(x).$$

where $\alpha = \frac{\tau-2}{5\tau-2}$.

Proof. First note that because M is a ordered and $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$ Lemma 2 implies

$$\mathbf{d}_{P,k}(v) \leq K\mathbf{d}_{M,2}(v)$$

for all $v \in M$.

Let $x \in B$ and $v \in M$ be such that $x \in \texttt{Vor}_M(v)$. Because $\mathbf{d}_{M,2}$ and $\mathbf{d}_{P,k}$ are 1-Lipschitz, it follows

$$\begin{aligned}\mathbf{d}_{P,k}(x) - \|x - v\| &\leq \mathbf{d}_{P,k}(v)\\ &\leq K(\mathbf{d}_{M,2}(x) + \|x - v\|),\end{aligned}$$

therefore, because $\|x - v\| \leq \mathbf{d}_{M,2}(x)$ we have

$$\begin{aligned}\mathbf{d}_{P,k}(x) &\leq K\mathbf{d}_{M,2}(x) + \|x - v\|(K + 1)\\ &\leq \mathbf{d}_{M,2}(x)(2K + 1)\end{aligned}$$

where $K = \frac{2\tau}{\tau-2}$, which implies $\alpha\mathbf{d}_{P,k}(x) \leq \mathbf{d}_{M,2}(x)$ for all $x \in B$. $\square$

3.2 Upper Bound

We now must show that the second nearest neighbor function to the mesh vertices $M \subset B$ is not too large compared with the kth-nearest neighbor function to the input set P. We will first show that the distance from any point $x \in \texttt{Vor}_M(p)$ to two points in M is within a constant factor of the distance from p to the circumcenter $\texttt{cc}(t)$ of some Delaunay triangle $t \in \texttt{Del}_M$.

Lemma 4 *Let M be a τ-well spaced such that $|\texttt{Vor}_M(v) \cap P| < k$ for all $v \in M$ and let $\beta > 1$ be a constant. Let $p \in M$ be such that $x \in \texttt{Vor}_M(p)$ for any $x \in B$.*

If $\mathbf{d}_{P,k}(x) < \frac{1}{\beta}\mathbf{d}_{M,2}(x)$ then there exists a Delaunay triangle $t \in \texttt{Del}_M$ with $\texttt{cc}(t) \in \texttt{Vor}^0_M(p)$ such that

$$\mathbf{d}_{M,2}(x) \leq \frac{\beta}{\beta - 1}\|p - \texttt{cc}(t)\|.$$

Proof. Note that because $|\texttt{Vor}_M(v) \cap P| < k$ for all $v \in M$ we have that $\texttt{ball}_{x,P_k} \not\subset \texttt{Vor}_M(v)$, and there therefore must exist some $q \in M$ such that $\texttt{ball}_{x,P_k} \cap \texttt{Vor}_M(q) \neq \emptyset$ and $\texttt{Vor}_M(p, q) \neq \emptyset$. It follows that there exists some $x' \in \texttt{Vor}_M(p, q) \cap \texttt{ball}_{x,P_k}$ with $\mathbf{d}_{M,2}(x') = \|p - x'\| = \|q - x'\|$ such that

$$\begin{aligned}\mathbf{d}_{M,2}(x) &\leq \mathbf{d}_{M,2}(x') + \|x - x'\| && [\mathbf{d}_{M,2} \text{ is 1-Lipschitz}]\\ &= \|p - x'\| + \|x - x'\| && [\|p - x'\| = \mathbf{d}_{M,2}(x')]\\ &\leq \|p - x'\| + \mathbf{d}_{P,k}(x) && [\|x - x'\| \leq \mathbf{d}_{P,k}(x)]\\ &\leq \|p - x'\| + \frac{1}{\beta}\mathbf{d}_{M,2}(x) && \left[\mathbf{d}_{P,k}(x) < \frac{1}{\beta}\mathbf{d}_{M,2}(x)\right]\end{aligned}$$

Thus, $\mathbf{d}_{M,2}(x) \leq \frac{\beta}{\beta-1}\|p - x'\|$.

Because the point x' lies on the Voronoi boundary $\texttt{Vor}_M(p, q)$ there must exist some $t \in \texttt{Del}_M$ with $\texttt{cc}(t) \in$

$\texttt{Vor}^0_M(p) \cap \texttt{Vor}^0_M(q)$ such that $\|p - x'\| \leq \|p - \texttt{cc}(t)\|$. It follows that

$$\begin{aligned}\mathbf{d}_{M,2}(x) &\leq \frac{\beta}{\beta-1}\|p - x'\| \\ &\leq \frac{\beta}{\beta-1}\|p - \texttt{cc}(t)\|\end{aligned}$$

□

Theorem 5 states that when Algorithm 1 terminates $\mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$. We will draw a contradiction, depicted in Figure 5, in which there must be some Voronoi cell or Delaunay circumcircle containing at least k points in P whenever there exists some $x \in B$ such that $\mathbf{d}_{P,k}(x)$ is within a constant factor *less* than $\mathbf{d}_{M,2}(x)$, as in Lemma 4.

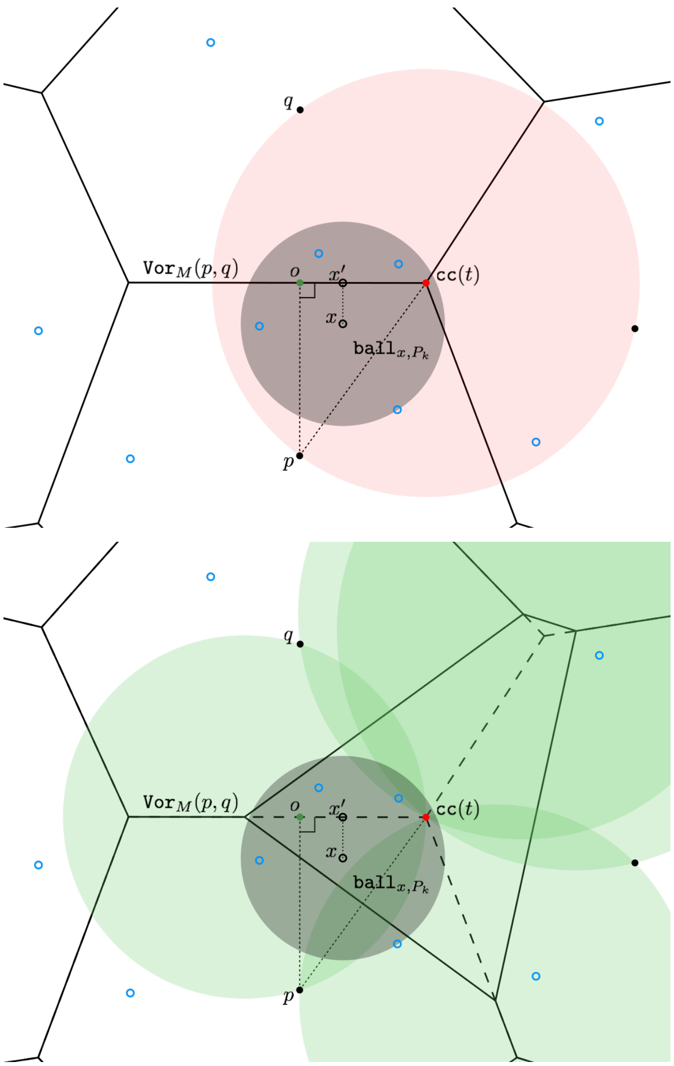

Figure 5: An illustration of the contradiction drawn in Theorem 5 (top) in which $\texttt{ball}_{x,P_k}$ (gray disk) is contained in the circumcircle of t (red disk). In this case, Algorithm 1 would not have terminated, as another Break move could be performed (bottom).

Theorem 5 *Let $P \subset B$ be a finite point set and $\tau \geq 2$, $k \geq 1$ be constants. Let $M \subset B$ be a τ-well spaced set such that $|\texttt{Vor}_M(v) \cap P| < k$ for all $v \in M$.*

If $\mathbf{d}_{P,k}(\texttt{cc}(t)) > \texttt{rad}(t)$ for all $t \in \texttt{Del}_M$ then

$$\mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$$

for a constant

$$\beta = \frac{3 - \vartheta_\tau}{1 - \vartheta_\tau} \text{ where } \vartheta_\tau = \sqrt{1 - \frac{1}{\tau^2}}.$$

Proof. Suppose, for the sake of contradiction, there exists a point $x \in B$ such that $\mathbf{d}_{P,k}(x) < \frac{1}{\beta}\mathbf{d}_{M,2}(x)$. Letting $p, q \in M$, $t \in \texttt{Del}_M$, and $x' \in \texttt{ball}_{x,P_k} \cap \texttt{Vor}_M(p,q)$ be such that $x \in \texttt{Vor}_M(p)$ and $\texttt{cc}(t) \in \texttt{Vor}^0_M(p) \cap \texttt{Vor}^0_M(q)$ as in Lemma 4 we have

$$\mathbf{d}_{M,2}(x) \leq \frac{\beta}{\beta-1}\|p - \texttt{cc}(t)\|.$$

To simplify notation we will set the point $o = \frac{p-q}{2}$ at the origin. Because the Delaunay edge of t containing the points p, q and the Voronoi edge $\texttt{Vor}_M(p,q)$ are orthogonal we can bound the distance from x' to the circumcenter of t by the Pythagorean Theorem as follows.

$$\begin{aligned}\|x' - \texttt{cc}(t)\|^2 &\leq \|\texttt{cc}(t) - o\|^2 \\ &\leq \|p - \texttt{cc}(t)\|^2 - \|p - o\|^2 \\ &\leq \left(1 - \frac{1}{\tau^2}\right)\|p - \texttt{cc}(t)\|^2. \quad \left[\frac{\|p - \texttt{cc}(t)\|}{\|p\|} \leq \tau\right]\end{aligned}$$

Noting that $\texttt{rad}(t) = \|p - \texttt{cc}(t)\|$ we can now prove that there are at least k points in P is the circumcircle of t.

$$\begin{aligned}\mathbf{d}_{P,k}(\texttt{cc}(t)) &\leq \mathbf{d}_{P,k}(x') + \|x' - \texttt{cc}(t)\| && [\mathbf{d}_{P,k} \text{ is 1-Lipschitz}] \\ &\leq \mathbf{d}_{P,k}(x') + \|p - \texttt{cc}(t)\|\vartheta_\tau && [\|x' - \texttt{cc}(t)\| \leq \vartheta_\tau] \\ &\leq 2\mathbf{d}_{P,k}(x) + \|p - \texttt{cc}(t)\|\vartheta_\tau && [\mathbf{d}_{P,k}(x') \leq 2\mathbf{d}_{P,k}(x)] \\ &< \frac{2}{\beta}\mathbf{d}_{M,2}(x) + \|p - \texttt{cc}(t)\|\vartheta_\tau && \left[\mathbf{d}_{P,k}(x) < \frac{1}{\beta}\mathbf{d}_{M,2}(x)\right] \\ &\leq \left(\frac{2}{\beta-1} + \vartheta_\tau\right)\|p - \texttt{cc}(t)\| && [\text{Lemma 4}] \\ &\leq \|p - \texttt{cc}(t)\|. && \left[\beta \geq \frac{3-\vartheta_\tau}{1-\vartheta_\tau}\right]\end{aligned}$$

It follows that $\mathbf{d}_{P,k}(\texttt{cc}(t)) \leq \|p - \texttt{cc}(t)\| = \texttt{rad}(t)$, a contradiction, as we assumed $\mathbf{d}_{P,k}(\texttt{cc}(t)) > \texttt{rad}(t)$ for all $t \in \texttt{Del}_M$. We may therefore conclude that if $M \subset B$ is τ-well spaced, $|\texttt{Vor}_M(v) \cap P| < k$ for all $v \in M$, and $\mathbf{d}_{P,k}(\texttt{cc}(t)) > \texttt{rad}(t)$ for all $t \in \texttt{Del}_M$, then $\mathbf{d}_{M,2} \leq \beta\, \mathbf{d}_{P,k}$. □

3.3 Main Theorem

Our final Theorem 6 states that when Algorithm 1 terminates the output set M is τ-well spaced with strictly less than k points from P in each Voronoi cell and Delaunay circumcircle. It will then follow from Theorems 3 and 5 that the second nearest neighbor distance function $\mathbf{d}_{M,2}$ defined on the bounding box $B = [0,1)^2$ is bounded above and below by the kth-nearest neighbor function of the input set $P \subset B$.

Theorem 6 (Main Theorem) *Let $P \subset B$ be a finite point set and $\tau \geq 2$, $k \geq 1$ be constants. Let M_0 be a set of initial mesh vertices with $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$.*

When $\textsc{kNNRefine}(P, M_0, \tau, k)$ *terminates the resulting ordered set of mesh vertices $M \subset B$ is τ-well spaced,*

$$|\mathtt{Vor}_M(v) \cap P| < k,\ \mathbf{d}_{P,k}(\mathtt{cc}(t)) > \mathtt{rad}(t)$$

for all $v \in M$, $t \in \mathtt{Del}_M$, and

$$\alpha \mathbf{d}_{P,k} \leq \mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$$

where $\alpha = \frac{\tau-2}{5\tau-2}$, $\beta = \frac{3-\vartheta_\tau}{1-\vartheta_\tau}$ and $\vartheta_\tau = \sqrt{1 - \frac{1}{\tau^2}}$.

Proof. First note that each $v \in M$, $t \in \mathtt{Del}_M$ must satisfy $\mathtt{aspect}(\mathtt{Vor}_M(v)) \leq \tau$ in order for each internal $\textsc{Clean}$ loop to complete. That is, if $\mathtt{aspect}(\mathtt{Vor}_M(v)) > \tau$ $\textsc{kNNRefine}(P, M_0, \tau, k)$ will not have terminated, as another $\textsc{Clean}$ procedure can be performed.

In order for each outer $\textsc{Break}$ loop to complete we must have that $|\mathtt{Vor}_M(v) \cap P| < k$ and $\mathbf{d}_{P,k}(\mathtt{cc}(t)) > \mathtt{rad}(t)$ for each $v \in M$, $t \in \mathtt{Del}_M$. Otherwise, $\textsc{kNNRefine}(P, M_0, \tau, k)$ will not have terminated, as an additional $\textsc{Break}$ procedure can be performed.

So, we may assume $\mathtt{aspect}(\mathtt{Vor}_M(v)) \leq \tau$, $|\mathtt{Vor}_M(v) \cap P| < k$, and $\mathbf{d}_{P,k}(\mathtt{cc}(t)) > \mathtt{rad}(t)$ for all $v \in M$, $t \in \mathtt{Del}_M$. As M is an indexed set, $\mathbf{d}_{P,k}(u_0) \leq \lambda_{u_0}$ for all $u_0 \in M_0$, and $\mathtt{aspect}(\mathtt{Vor}_M(v)) \leq \tau$ for all $v \in M$ it follows from Theorem 3 that

$$\alpha \mathbf{d}_{P,k} \leq \mathbf{d}_{M,2}.$$

Moreover, because $|\mathtt{Vor}_M(v) \cap P| < k$, and $\mathbf{d}_{P,k}(\mathtt{cc}(t)) > \mathtt{rad}(t)$ for all $v \in M, t \in \mathtt{Del}_M$ it follows from Theorem 5 that

$$\mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}.$$

We may therefore conclude that in order for Algorithm 1 to terminate the set M of mesh vertices constructed by $\textsc{kNNRefine}(P, M_0, \tau, k)$ must satisfy

$$\alpha \mathbf{d}_{P,k} \leq \mathbf{d}_{M,2} \leq \beta \mathbf{d}_{P,k}$$

for constants $\alpha = \frac{\tau-2}{5\tau-2}$ and $\beta = \frac{3-\vartheta_\tau}{1-\vartheta_\tau}$. □

4 Point Location

At the heart of any Delaunay refinement algorithm is an incremental Delaunay triangulation algorithm, constructing the Delaunay triangulation one vertex at a time. The standard approach in computational geometry for bounding the running time is to use randomization, yielding the randomized incremental algorithm. However, Delaunay refinement does not permit such arbitrary reordering of the points, because the points to be added are discovered in the course of running the algorithm. Thus, the original Chew and Ruppert algorithms could have $O(n^2)$ running times. This was improved by Miller [6] and later by Hudson et al. [4] who developed the so-called *Sparse Refinement* approach with their Sparse Voronoi Refinement (SVR) algorithm.

As the algorithm needs to know the number of input points contained in every Voronoi cell as well as the number of points contained in every Delaunay circumball, we will maintain two different point location data structures. The first is a 2-way mapping between points of P and the Voronoi cells (points of M). The second is a 2-way association between the points of P and the Delaunay circumballs. For each local update to the Delaunay triangulation induced by a single insertion, some Voronoi cells are affected as well as some circumballs. The sparse refinement approach always maintains some guarantee on the quality of the underlying triangulation.

We adopt the vocabulary used by Hudson et al. [4] when describing the algorithm in terms of $\textsc{Break}$ and $\textsc{Clean}$ moves. In fact, we will do a strict subset of the operations that would usually be performed by SVR. The difference is that in SVR, if any input point from P has not yet been added, then the algorithm will continue, whereas our algorithm will halt early if every Voronoi cell and every Delaunay circumball contains fewer than k points. Thus, it follows immediately that our algorithm will also achieve a running time (and output size) of $O(n \log \Delta)$. Here, Δ denotes the *spread* of the input defined as the ratio of the largest to smallest pairwise distances. We only need to observe that counting the points in a Voronoi cell or Delaunay cicumball is not more expensive than iterating through the list of these points which happens anyway each time a Voronoi cell changes or a Delaunay circumball is created or destroyed.

5 Conclusions and future work

We have shown how a simple modification of Delaunay refinement solves the kth nearest neighbor sampling problem. Several interesting open problems remain.

1. Does the algorithm work in $\mathbb{R}^d$ for $d > 2$? We believe the answer is yes.

2. Is it possible to eliminate the dependence on $\log(\Delta)$ as was done for Voronoi refinement of points [7, 8]?

3. Is the size of the sample we produce asymptotically optimal? Specifically, we would like to extend the optimality theory of Ruppert [9].

References

[1] M. Bogdanov, M. Teillaud, and G. Vegter. Covering spaces and delaunay triangulations of the 2d flat torus. In *28th European Workshop on Computational Geometry*, 2012.

[2] S.-W. Cheng, T. K. Dey, and J. R. Shewchuk. *Delaunay Mesh Generation*. CRC Press, 2012.

[3] L. P. Chew. Guaranteed-quality triangular meshes. Technical Report TR-89-983, Department of Computer Science, Cornell University, 1989.

[4] B. Hudson, G. Miller, and T. Phillips. Sparse Voronoi Refinement. In *Proceedings of the 15th International Meshing Roundtable*, pages 339–356, Birmingham, Alabama, 2006. Long version available as Carnegie Mellon University Technical Report CMU-CS-06-132.

[5] D. O. Loftsgaarden and C. P. Quesenberry. A nonparametric estimate of a multivariate density function. *Ann. Math. Statist.*, 36(3):1049–1051, 1965.

[6] G. L. Miller. A time efficient Delaunay refinement algorithm. In J. I. Munro, editor, *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11-14, 2004*, pages 400–409. SIAM, 2004.

[7] G. L. Miller, T. Phillips, and D. R. Sheehy. Beating the spread: Time-optimal point meshing. In *Proceedings of the 26th ACM Symposium on Computational Geometry*, pages 321–330, 2011.

[8] G. L. Miller, D. R. Sheehy, and A. Velingker. A fast algorithm for well-spaced points and approximate delaunay graphs. In *Proceedings of the 29th annual Symposium on Computational Geometry*, pages 289–298, 2013.

[9] J. Ruppert. A Delaunay refinement algorithm for quality 2-dimensional mesh generation. *J. Algorithms*, 18(3):548–585, 1995.