Curve Reconstruction with Many Fewer Samples

S. Ohrhallinger¹, S.A. Mitchell² and M. Wimmer¹

¹Institut für Computergrafik und Algorithmen, TU Wien, Austria
²Center for Computing Research, Sandia National Laboratories, U.S.A.

Abstract

We consider the problem of sampling points from a collection of smooth curves in the plane, such that the CRUST family of proximity-based reconstruction algorithms can rebuild the curves. Reconstruction requires a dense sampling of local features, i.e., parts of the curve that are close in Euclidean distance but far apart geodesically. We show that \( \varepsilon < 0.47 \)-sampling is sufficient for our proposed HNN-CRUST variant, improving upon the state-of-the-art requirement of \( \varepsilon < \frac{1}{3} \)-sampling. Thus we may reconstruct curves with many fewer samples. We also present a new sampling scheme that reduces the required density even further than \( \varepsilon < 0.47 \)-sampling. We achieve this by better controlling the spacing between geodesically consecutive points. Our novel sampling condition is based on the reach, the minimum local feature size along intervals between samples. This is mathematically closer to the reconstruction density requirements, particularly near sharp-angled features. We prove lower and upper bounds on reach \( \rho \)-sampling density in terms of lfs \( \varepsilon \)-sampling and demonstrate that we typically reduce the required number of samples for reconstruction by more than half.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

The connect-the-dots game without numbers on the dots corresponds to the problem of reconstructing the connectivity of a planar curve from a set of unstructured points sampled on that curve.

More formally, our problem is to sample points from a curve, throw away the curve, then connect points to those nearby. For the reconstruction to be correct, the points should be connected in the same order as on the curve. A sparser sampling is valuable whenever placing points, storing them, or reconnecting them is expensive. But it must not be too sparse because the connectivity must be restorable from just the points.

The samples capture the essential shape information, topological and geometric. The Human Visual System is able to complete the connectivity based on the Gestalt principles of Proximity and Continuity. Familiar examples are planting flower bulbs to form a shape, or animating patterns in the night sky by lit drones. Recon-
construction algorithms are also built on these proximity and continuity principles. Potential applications include generating efficient shape descriptors based on points (as opposed to curve-based descriptions), and compressing or progressively streaming point sets. These can be used to decide whether to request additional samples from a sensor, or that the sample set is of sufficient quality.

If the curve is sampled densely, connecting nearby points will reconstruct the correct curve. The less dense the sampling, the more challenging it is to reconstruct the curve, especially at features where two intervals of the curve come close to each other, or where the curvature is high. Reconstruction algorithms require some sampling conditions on the input in order to guarantee a correct output. The particular algorithm determines the required density.

Sampling algorithms also guarantee some sampling conditions on the output. However, these are rarely of exactly the same form, and it is non-trivial to describe the reconstruction algorithm’s requirements in terms of the sampling algorithm’s guarantees. This leads to a mismatch between the minimum local density required for reconstruction, and the maximum local density a sampling algorithm produces. Typically we choose some local measure of a curve, and sample density is guaranteed to be some parameterized fraction of that measure. The closer the guarantees match the requirements, and the tighter we can describe the necessary and sufficient parameter values, the more efficient we can make our sampling. This leads to our goal: to sample curve features as sparsely as possible, yet still guarantee that the reconstructed curve is correct.

We describe the reconstruction algorithm HNN-CRUST, a variant of NN-CRUST [DK99]. Many sampling algorithms use the ε-sampling condition, which is based on comparing ε times the local feature size (lfs) at a point to the distance to its nearest sample [ABE98]. The known parameter bounds for this combination, ε < 1/3-sampling, appear weak, and we show a better one, ε < 0.47-sampling. Furthermore, we provide a better sampling condition based on a different measure of the curve, the reach, ρ. The reach is bounded by the minimum local feature size at all points between two samples. The reach is more suitable for HNN-CRUST, and we believe for proximity-based reconstruction in general.

Our first contribution is the tightening of ε < 1/3-sampling to ε < 0.47-sampling.

Our second and main contribution is the new reach-based p-sampling condition, with the following properties:

- p-sampling is simple, with a single parameter like ε-sampling.
- p < 0.9-sampling guarantees that HNN-CRUST correctly reconstructs the curve.
- The polygonal reconstruction geometrically approximates the original curve, similar to ε < 0.47-sampling.
- p < 0.9-sampling has only half the samples when lfs is constant, and never more than ε < 0.47-sampling.
- The same condition holds when limiting the Hausdorff distance from the polygonal reconstruction to the original curve.
- Thus, p < 0.9-sampling permits much sharper angles: up to 73°, compared to 120° for ε < 1/3-sampling.

Programs for sampling smooth curves under both sampling conditions are provided online as open source. One can explore varying ε and p parameters, as well as Hausdorff distance limits.

2. Related Work

We briefly review curve reconstruction algorithms and their associated sampling conditions. Early methods guaranteed curve reconstruction from uniformly dense samples, where the maximum distance between consecutive samples is a global constant [EKS83, KR85, FMG94, Att97]. However, since the sampling density is constant, it depends on the maximum curvature, which is inefficient for flat parts of the curve. Those methods work well for curves whose curvature is limited above by a global constant, such as for r-regular sets [DT14, DT15], for which guarantees are given for non-noisy [Ste08] and noisy point sets [ST09].

Sampling framework: To get rid of this over-sampling, the seminal paper by [ABE98] proposed the CRUST algorithm. It filters edges from the Delaunay triangulation. Sampling density varies according to both curvature and Euclidean distance between geodesically-far curve intervals. They also introduced a non-uniform sampling case based on local feature size, called ε-sampling, and proved that CRUST reconstructs a manifold boundary; [Dey06] proved ε < 0.2 is sufficient. Many subsequent methods use this sampling framework. [Go99] optimized and simplified CRUST to a single-step algorithm. This family of algorithms constrain their output to edges of the Delaunay triangulation.

Proximity-based algorithms: [DK99] introduced the simple proximity-based algorithm NN-CRUST for general dimensions. It guarantees reconstruction of closed curves for ε < 1/3. [Alt01] improved the condition to ε < 0.5, but required α > 151°. [Len06] claims a better bound for NN-CRUST: ε < 0.4, or ε < 0.48 with additional angle restrictions, but does not show proof. He also noted shortcomings of ε-sampling, e.g. for sharp corners, as open problems. These investigations show that there is still room for improvement. Without angle restrictions, the best proven bound is ε < 1/3-sampling, and this is not tight.

Extensions: NN-CRUST was extended to CONSERVATIVE-CRUST [DMR99] to handle open curves, and later to GATHAN-C [DW02], which modified the sampling condition to handle sharp corners, but requires α > 150° otherwise. [FR01] introduced the notion of curve reconstruction as requiring a homeomorphism between the polygonal reconstruction and the curve, but not geometric closeness. They also presented their own sampling condition, requiring several parameters, in order to reconstruct collections of open and closed curves with sharp corners. Other approaches proposed a sampling condition using a vision function based on human perception and some empirically established parameters [ZNYL08, NZ08]. [OM13] presented a three-step method which is able to reconstruct very sparsely-sampled features, for closed curves, by considering it as a global problem. The first step guarantees reconstruction for ε < 0.5, but in order to handle the sharp angles of 0°–60° it requires an additional constraint, slowly varying density as a maximum ratio between adjacent edge lengths.

Sampling: [LKV14] generate connect-the-dot puzzles from curves which vary in the criterion of connectivity, using different sampling criteria. Their variant connect-the-closest-dot corresponds closely to our problem, but our sampling condition neither requires encoding of topology indicators nor a minimum distance between points.
3. Overview

We describe a variant of NN-CRUST [DK99] that we call HNN-CRUST, which permits reconstruction of angles sharper than \(< 90^\circ\), as small as \(60^\circ\). While it improves the reconstruction, it is mostly a vehicle to compare our \(\rho\)-sampling condition to the widely used \(\varepsilon\)-sampling condition [ABE98]. We consider only these two sampling conditions for comparison because the others are highly tailored to specific reconstruction algorithms and require careful adjustment of many parameters.

The HNN-CRUST reconstruction algorithm implies that two edges meet at an angle of at least \(60^\circ\). The reconstruction is correct for \(\varepsilon < 0.47\). The angle between consecutive edges is related to curvature and sampling density: the flatter the curve and the denser the sampling, the larger the angle. (In the limit, for an infinite sampling of a regular curve, we get \(180^\circ\).)

The essence of our paper is a new sampling condition that samples more sparsely where possible, closer to the minimum tolerated by the reconstruction. The weakness of \(\varepsilon\)-lfs sampling is that, in essence, the sampling condition’s output guarantee is that the maximum distance between consecutive samples is limited by the lfs at a point half way between them. The sampling condition is less sensitive to the lfs at other points, and the lfs at the samples themselves are completely irrelevant. In contrast, the reconstruction algorithm’s input requirements are sensitive to small lfs at the samples themselves. This mismatch leads one to select an \(\varepsilon\) small enough that the algorithm is correct even when the lfs changes rapidly between the midpoint and the sample. The sampling density is driven by this worst case, and is much denser than necessary when the lfs is not changing rapidly. The strength of our new measure, the reach, is that it is sensitive to small lfs at the samples, and so the sampling condition is more closely matched to the reconstruction requirements.

The rest of the paper is organized as follows. In Section 4 we introduce the required background and definitions. We explain the reconstruction algorithm in Section 5 together with some properties. In Section 6 we give our improved \(\rho < 0.9\)-sampling condition based on the reach rather than local feature size. In Section 7 we prove that \(\rho < 0.9\)-sampling suffices. We also prove bounds relating \(\rho\)-sampling to \(\varepsilon\)-sampling, which indirectly proves \(\varepsilon < 0.47\)-sampling suffices. We compare the results of our reconstruction algorithm and sample density for our sampling condition in Section 8. In Section 9 we give our conclusions along with potential extensions.

4. Definitions

We give the following definitions, most of which have been introduced by [ABE98]:

The domain is a collection of smooth curves \(C\), by which we mean bounded 1-manifolds embedded in \(\mathbb{R}^2\), which are twice-differentiable everywhere except perhaps at boundaries. This permits \(C\) to consist of multiple connected components, such as a circle and a closed segment, but without crossings, T-intersections or sharp angles. The boundary of a closed segment consists of two terminus points. Note that each connected component of \(C\) induces a natural geodesic ordering of its points, which can be traversed in one of the two possible directions. Based on such a directed ordering, we say that a curve point lies before or after another, or between two curve points. The interval \(I(p) \equiv [s_0, s_1]\) is the set of points \(p \in C\) between \(s_0\) and \(s_1\). A chord is the straight edge between two points of an interval.

The set of samples is \(S\). Samples \(s_0\) and \(s_1\) are adjacent or consecutive if there is no other sample on their interval. Let \(\| \cdot \|\) denote the Euclidean \(L_2\)-norm. We measure distances in the Euclidean metric, except where we specifically denote geodesic distance.

The nearest neighbor \(s_0\) to sample point \(s_1\) is \(\arg\min_{s_i \in S \setminus \{s_1\}} \|s_1 - s_i\|\). The half neighbor \(s_2\) is the closest sample in the half-space \(H\) which is partitioned by the perpendicular bisector of the edge \(e_{s_0 s_1}\) and does not contain \(s_0\): \(\arg\min_{s_i \in S \setminus \{s_1\}, s_i \in H} \|s_1 - s_i\|\). We often order all neighbors by Euclidean distance: let \(n_j\) be the \(i\)-th nearest sample to \(s_1\).

We define the manifold boundary \(B\) as the correct piece-wise linear reconstruction of \(C\), which connects the samples of each connected component in the same order as on \(C\) and adds no other edges.

The medial axis \(M\) of \(C\) is the closure of all points in \(\mathbb{R}^2\) with two or more closest points in \(C\) [Blu67].

We define the local feature size \(lfs(p)\) for a point \(p \in C\) as the Euclidean distance from \(p\) to its closest point \(m\) of \(M\). This definition is loosely based on [Rup93], but simplified because we are only considering smooth curves. Note \(lfs(p)\) is slowly varying, 1-Lipschitz continuous with \(\|lfs(p_0) - lfs(p_1)\| < \|p_0 - p_1\|\).

Definition 1 is the widely used lfs sampling condition [ABE98]:

**Definition 1** A smooth curve \(C\) is \(\varepsilon\)-sampled by point set \(S\) if every point \(p \in C\) is closer to a sample than an \(\varepsilon\)-fraction of its local feature size: \(\forall p \in C\); \(\exists s \in S\); \(\|p - s\| < \varepsilon lfs(p)\).

In contrast, the reach [Fed59] for a set \(S\) is the largest “radius” \(r\) such that points closer than \(r\) to \(S\) have a unique closest point of \(S\). The reach is similar to the smallest distance to the medial axis. This inspires our definition of the reach of a curve interval \(I\) as \(inflfs(p) : p \in I\), where the lfs is defined by all of \(C\).

5. Our Improved Reconstruction Algorithm HNN-CRUST

HNN-CRUST simply connects each sample \(s \in S\) to its nearest and half neighbor. (If \(s\) is a terminus of a curve, then only the nearest neighbor gets an edge. If the terminus is not specifically marked, then the reconstruction will have an extra edge.) Let \(h\) be the perpendicular bisector of the nearest neighbor edge, and \(H\) its half-space containing \(s\). Then the half neighbor lives in \(H\) but outside the nearest-neighbor radius around \(s\); see Figure 2. In Figure 3 we show how CRUST and HNN-CRUST compare when consecutive samples make sharp angles.

5.1. HNN-CRUST Mimics the Human Vision System

Three Gestalt principles are implicitly present in HNN-CRUST. (Since our algorithm does not attempt to reproduce the Human Vision System, some reconstructions will not match typical human.
perception.) These principles can be observed in Figure 3, and are as follows:

- **Proximity** is enforced by always connecting the nearest neighbor, and for the second neighbor choosing the nearest neighbor inside the restricted halfspace.
- **Good Continuity** arises from requiring angles between incident edges to be more than 60°.
- **Closure** means we close the curve, unless excessive distance between points implies a hole or an open curve.

6. An Improved Sampling Condition

We will show that HNN-CRUST reconstructs a smooth curve for an ε-sampling with ε < 0.47. For higher values of ε, [ABE98] observed some interesting properties. Theorem 12 noted that for ε < 1, the reconstruction B ⊂ DT (Delaunay Triangulation). Theorem 13 showed that the distance from any point p ∈ C to the polygonal reconstruction B is bounded above by ε²IFS(p)/2. However, we have not seen any attempts to guarantee reconstruction for 0.47 ≤ ε < 1, so we will investigate why this is hard.

6.1. Large ε Do Not Keep Geodesically Distant Intervals Away

IFS ε-sampling (Definition 1) just requires a sample to be within an ε-fraction of the IFS at that point. Thus, as p ∈ C approaches a sample point, IFS(p) may be arbitrarily small, and the sampling condition is still satisfied. The only thing keeping geodesically distant curves sections separate is the ε-IFS condition at points farther away, such as the point x ∈ C midway between samples often used in proofs. Therefore, for an ε-sampling with 0.47 ≤ ε < 1, HNN-CRUST may connect non-adjacent samples and fail.

6.2. The Solution for Keeping Them at the Proper Distance

To sample more sparsely where samples are not needed, but still ensure samples are dense enough where the curve approaches itself, we must have a sampling condition that depends more strongly on the IFS near samples. Our sampling condition replaces IFS(p) by the reach, the minimum IFS on an interval.

**Definition 2** The reach [Fed59] of interval I is infp∈I IFS(p).

**Definition 3** A smooth curve C is ρ-sampled by point set S if every point p ∈ C is closer to a sample than a ρ-fraction of the reach of the interval I(s₀,s₁) of consecutive samples containing it. That is, ∀ p ∈ I = [s₀,s₁] with s₀,s₁ ∈ S: ||p,s₀|| < ρreach(I) or ||p,s₁|| < ρreach(I).

7. Correctness of HNN-CRUST for ρ < 0.9 and ε < 0.47

The goal of this section is to show reconstruction provides correct output for certain ρ. Indeed, we will show that every ε-sample is also a ρ-sample, so this implies correctness for certain ε. The idea is to show that consecutive samples are close together, that geodesically close samples are farther, and geodesically distant samples are farther as well. We establish a series of geometric preliminaries relating distances between samples, the curve, and its medial axis. Most are similar to previous observations, but in some cases we provide stronger results or more elegant proofs.

The first lemma is useful for geodesically close samples. Theorem 2 in [OM13] shows, amongst other things, that Euclidean chord length increases monotonically with geodesic distance, as long as chords do not intersect M. In particular, for I = [p₀,p₂], as x advances on C from p₀ to p₂, chord length ||p₀x|| is strictly increasing, and has no local maxima. Here we show something stronger, with a more elegant proof.

**Lemma 1** Let p₀,p₂ ∈ C. If the chord h ≡ p₀p₂ does not cross the medial axis M of C, the interval I = [p₀,p₂] lies inside the smallest circle O₀₂ containing p₀p₂. Moreover, for t ∈ I, distances ||p₀t|| and ||p₂t|| are strictly monotonic in t’s ordering on I.

**Proof** See Figure 5 left. For each point x on segment h, consider the largest radius disk O centered at x with no points of C in its interior. Let t be a point of C on the boundary of O. Then we have the function T(x) = t with r ∈ C and x ∈ h. Note T(p₀) = p₀ and T(p₂) = p₂, with radius zero. If T is discontinuous (multivalued) at some x, then O touches C at two or more points, and x ∈ M. Hence T(x) must be continuous. Thus {t} must lie on a single connected component of C, an interval, and h is a chord. Since O can never contain p₀ or p₂ in its interior, O lies inside the diameter disk, and hence so must all t. Observe O has strictly higher curvature (i.e. smaller radius) than O₀₂.

The continuity and curvature limit of T implies I cannot be perpendicular to h: if it were, then t⁺ = T(⟨x⟩) for some continuous range of x. Continuity of T at the boundary of this range implies the curvature of I at t⁺ is at most that of O₀₂, a contradiction. Hence the {x} where T(x) = t is a single point for all t. Hence T is monotonic. This leads to the range of T being I. For curves that are topological circles, the range might instead be I’ where I’ = [p₂,p₀]. Since here the orientation of I is arbitrary, we will label the enclosed interval “I’.

Besides t = T(x) being monotonically ordered on I, the distance ||p₀t|| is also monotonic. It two points t₁ and t₂ of I are equidistant from p₀, then they lie on a circle O₁₀ centered at p₀. Let t₁ be the
∃ curve varies continuously and monotonically with circle center $x$ along curve interval must lie in the diameter disk. Left, tangent point $t$ ples are close. We exploit the principle that an Open and closed curves. Left: Sample points. Center: Lemma 2

Let $s$ures that an adjacent sample $s$ s

contradiction. By symmetry, $∥O(h)−s∥ < 2$.


Figure 4: Open and closed curves. Left: Sample points. Center: CRUST reconstruction. Right: HNN-CRUST reconstruction.

Figure 5: If a chord does not cross a medial axis point, then the curve interval must lie in the diameter disk. Left, tangent point $t$ varies continuously and monotonically with circle center $x$ along $\overline{p_0p}$. Right, distance from $p_0$ to $t$ is strictly increasing, else $t_2$ is unreachable.

one closer to $h$. (They cannot be equidistant because $T$ is single-valued.) Then any circle in $O$ touching $t_2$ has $t_1$ in its interior, a contradiction. By symmetry, $∥p_1t∥$ is also monotonic in $x$. □

The next two lemmas quantify the fact that consecutive samples are close. We exploit the principle that an $\varepsilon$-sampling ensures that an adjacent sample $s_2$ is close to $s_1$ in terms of lfs. From [ABE98] Lemma 1’s proof:

Lemma 2 Let $s_1, s_2$ be adjacent samples in $C$. For an $\varepsilon$-sampled curve $\exists y \in I[s_1, s_2]$ such that

$$∥s_1s_2∥ \leq 2∥s_2y∥ = 2∥s_1y∥ < 2lfs(y)$$

$$lfs(s_1)/(1 + \varepsilon) < lfs(y) < lfs(s_1)/(1 - \varepsilon)$$

Proof $\varepsilon$-sampling ensures $2∥s_1y∥ < 2lfs(y)$ and the triangle inequality provides $∥s_1s_2∥ \leq 2∥s_1y∥$. Since lfs is 1-Lipschitz, lfs($y$) $\leq$ lfs($s_1$) + $∥s_1y∥$ replaced with above inequality for $∥s_1y∥$ yields lfs($y$) $< lfs(s_1)/(1 - \varepsilon)$. Also from the 1-Lipschitz property, lfs($y$) $\geq$ lfs($s_1$) $- ∥s_1y∥$ replaced with $∥s_1y∥$ from above provides lfs($y$) $> lfs(s_1)/(1 + \varepsilon)$. □

Lemma 3 For adjacent samples $s_0, s_1, s_2$, let $x \in I[s_0, s_1]$ with $∥s_0x∥ = ∥s_1x∥$ and $y \in I[s_1, s_2]$ with $∥s_2y∥ = ∥s_1y∥$. Then,

$$lfs(x) > \frac{1 - \varepsilon}{1 + \varepsilon} lfs(y).$$

For the reach, the situation is considerably simpler.

Lemma 4 For a $\rho$-sampled curve with consecutive samples $s_0$ and $s_1$, $∥s_0s_1∥ < 2\rho$ reach($l_0$) $\leq 2\rho$ lfs($s_1$). Moreover, for midpoint $x$,

$$lfs(s_1)/(1 + \rho) < lfs(x) < (1 + \rho)lfs(s_1).$$

Proof $\exists x \in I[s_0, s_1]$ such that $∥s_0x∥ = ∥x, s_1∥ < \rho$ reach($l_0$) $\leq \rho$ lfs($s_1$). The bound on lfs($x$) follows from reach($l_0$) $\geq lfs(x)$ and 1-Lipschitz. □

The next two lemmas show that geodesically distant samples are also far in Euclidean distance. We then relate $\rho$- and $\varepsilon$-sampling. Finally we provide additional restrictions on the interval between consecutive samples, quantifying how close it must be to a straight line, and additional lower bounds on Euclidean distance.

We call a disk with no point of $C$ in its interior “$C$-free”, and a disk with no point of $M$ in its interior “$M$-free” Recall [ABE98] Lemma 7:

Lemma 5 A disk tangent to a smooth curve $C$ at a point $p$ with radius at most $lfs(p)$ is $C$-free.

We generalize Lemma 5 to the following.

Lemma 6 A rolling tangent circle $Rty$ with center interior to circle $O(y,lfs(y))$ touches $C$ at a single point $p$ in interval $O(y,lfs(y))\cap C$.

Proof By definition, $O(y,lfs(y))$ is $M$-free. Following the proof of Lemma 5, growing a tangent disk at $y$ with continuously increasing radius cannot intersect another point of $C$ before the radius reaches lfs($y$), else the center would be a point of $M$. By the same argument,
Lemma 7

Ranges for $x$ reach $\geq (1 - \varepsilon)\text{ifs}(x)$.

Proof: The proof is the same as that of Lemma 2, combined with using Lemma 1 to show distances are monotonic along $I = [s_0, s_1]$. For any $\varepsilon$-sampled interval $I = [s_0, s_1]$, we first show $\text{reach}(I) \geq (1 - \varepsilon)\text{ifs}(x)$, then show the condition holds $\forall p \in I$. Let $x \in I$ be equidistant from $s_0$ and $s_1$. From Lemma 1, $\|p\| \geq \|x_0\|$. By 1-Lipschitz, $\text{ifs}(p) \geq \text{ifs}(x) - \|x_0\| > (1 - \varepsilon)\text{ifs}(x)$. Thus $\text{reach} = \inf_p \text{ifs}(p) \geq (1 - \varepsilon)\text{ifs}(x)$. Again by Lemma 1, $\forall p \in [s_0, x_2], |ps| \leq \|x_0\| \leq \varepsilon(1 - \varepsilon)\text{reach}$. The argument for $p \in [x, s_1]$ is the same. Thus, for $r < 1$, any $\varepsilon < r$-sampling is also a $\rho < r/(1 - r)$-sampling.

Corollary 1 $\rho < r/(1 - r)$-sampling does not require more samples than $\varepsilon < r$-sampling.

Lemma 8 For a $\rho < 1$-sampling, $\angle{s_0s_1s_2} \geq \pi - 4\arcsin\rho/2$ and $\angle{x_1y} \geq \pi - 2\arcsin\rho/2$. This is tight for constant curvature.

Proof: Consider the C-free tangent disk to $s_1$ of radius $\text{ifs}(s_1)$. The reach on each interval containing $s_1$ is at most $\text{ifs}(s_1)$. In Figure 7(a), this leads to $\|x_0\| \leq \rho\text{ifs}(s_1)$, then $\theta = 2\arcsin(\rho/2)$ and $\angle{s_0s_1s_2} \geq 2\theta = \pi - 2\theta$.  

Lemma 9 For an $\varepsilon$-sampled curve, with $\varepsilon < 0.5$, the angle spanned by three adjacent samples is at least $\pi - 4\arcsin(\varepsilon/(2 - 2\varepsilon))$.

Proof: Combine Lemma 8 with Theorem 1.

Figure 6: Forbidden regions: the red circles are C-free except for $I = [s_0, s_1]$, and $I$ lies in their lune-shaped intersection and $z \in I$ on the green line inside that lune. The lunes bound the extreme cases of constant curvature, where $\text{ifs} = \text{reach} = \text{ifs}(x)$. In (a), the black lines have length $0.5\text{ifs}$ and the blue lines $\text{ifs}$. In (b), the black lines have length $\text{ifs}$ and the blue triangles are equilateral.

We may now continuously vary the center within $O(y, \text{ifs}(y))$, keeping a continuous tangent at $p$ in an interval around $y$.

Combining the idea of growing a tangent ball at $y$ with the fact that local curvature is less than $1/\text{ifs}(y)$ results in the forbidden regions from [ABE98]. We summarize the properties we use in the following lemma.

Lemma 7 The two circles through consecutive samples $s_0$ and $s_1$ with the maximum curvature allowed by the sampling condition are C-free except for $I = [s_0, s_1]$. Moreover, $I$ lies in the lune of intersection of the two circles. See Figure 6.

In the following sense, our $\rho$-sampling is at least as good (i.e. as sparse) as $\varepsilon$-sampling:

Theorem 1 Any $\varepsilon < r$-sampling is also a $\rho < r/(1 - r)$-sampling, for $r < 1$. E.g. an $\varepsilon < 0.5$-sampling is also a $\rho < 1$-sampling, and an $\varepsilon < 1/3$-sampling is also a $\rho < 0.5$-sampling.

Proof: The proof is the same as the proof of Lemma 2, combined with using Lemma 1 to show distances are monotonic along $I = [s_0, s_1]$. For any $\varepsilon$-sampled interval $I = [s_0, s_1]$, we first show $\text{reach}(I) \geq (1 - \varepsilon)\text{ifs}(x)$, then show the condition holds $\forall p \in I$. Let $x \in I$ be equidistant from $s_0$ and $s_1$. From Lemma 1, $\|p\| \geq \|x_0\|$. By 1-Lipschitz, $\text{ifs}(p) \geq \text{ifs}(x) - \|x_0\| > (1 - \varepsilon)\text{ifs}(x)$. Thus $\text{reach} = \inf_p \text{ifs}(p) \geq (1 - \varepsilon)\text{ifs}(x)$. Again by Lemma 1, $\forall p \in [s_0, x_2], |ps| \leq \|x_0\| \leq \varepsilon(1 - \varepsilon)\text{reach}$. The argument for $p \in [x, s_1]$ is the same. Thus, for $r < 1$, any $\varepsilon < r$-sampling is also a $\rho < r/(1 - r)$-sampling.

Corollary 1 $\rho < r/(1 - r)$-sampling does not require more samples than $\varepsilon < r$-sampling.

Lemma 8 For a $\rho < 1$-sampling, $\angle{s_0s_1s_2} \geq \pi - 4\arcsin\rho/2$ and $\angle{x_1y} \geq \pi - 2\arcsin\rho/2$. This is tight for constant curvature.

Proof: Consider the C-free tangent disk to $s_1$ of radius $\text{ifs}(s_1)$. The reach on each interval containing $s_1$ is at most $\text{ifs}(s_1)$. In Figure 7(a), this leads to $\|x_0\| \leq \rho\text{ifs}(s_1)$, then $\theta = 2\arcsin(\rho/2)$ and $\angle{s_0s_1s_2} \geq 2\theta = \pi - 2\theta$.

Lemma 9 For an $\varepsilon$-sampled curve, with $\varepsilon < 0.5$, the angle spanned by three adjacent samples is at least $\pi - 4\arcsin(\varepsilon/(2 - 2\varepsilon))$.

Proof: Combine Lemma 8 with Theorem 1.

Figure 7: Ranges for $x, y, s_0, s_2$, angles and $H$ for $\rho$-sampling. The red circles are tangent to $C$ at $s_1$ with radius $\text{ifs}(s_1)$, and are C-free and exclude $x, y, s_1, s_2$ from their interior. In (b), the green circle is $O(s_1, \text{ifs}(s_1))$, and contains $x$ and $y$. Sample $s_0$ lies in the union of the three purple sectors and one green sector to the left of $s_1$. Hence $H$ contains the purple and green right sectors and $s_2$.
moves off a red circle, \( H \) just retreats farther from \( s_2 \)’s admissible region.

Thus, we need only show that no other sample \( z \) in \( H \) is closer. While showing that \( s_0 \) was the nearest neighbor, we already established that any \( z \) was farther than \( \|s_1s_2\| \) except perhaps when \( \mathbb{T}_1 \cap M \neq \emptyset \) and its closest point of \( I \) is \( q \in [s_0, s_1] \). For \( \rho < 0.5 \), the remainder is trivial because \( \|s_3\| \geq 15s(1) > 2\rho\|s_1s_2\| \). For larger \( \rho \), the main idea of the proof is to use rolling tangent balls to cover the part of \( O(s_1, [s_1s_2]) \) in \( H \). From Lemma 7, \( I = [s_0, s_2] \) is restricted to lie in the union of two lunes, which provides a lower bound on the radii of the rolling tangent balls from Lemma 6. Hence the balls are large and cover the portion of the circle \( O(s_1, [s_1s_2]) \)) in \( H \). Unfortunately, we do not have a closed-form algebraic description of this fact. Instead, we have a computer assisted proof. We consider the possible ranges of positions, with \( \|x, s_1\|/\|x, s_1\| \in [0, 1] \) and the tangent angles between \( \mathbb{T}_1 \) and \( \mathbb{T}_2 \in [0^\circ, 53.5^\circ] \) (Lemma 8). We divide each of these three ranges into small intervals. For all feasible combinations of intervals, we take the worst case value for each quantity independently when used. For all ranges we construct a collection of rolling tangent circles that covers \( O(s_1, [s_1s_2]) \)). Figure 8 provides a few representative examples. These figures and all other feasible combinations can be reproduced with a matlab script available online.

**Theorem 3** For an \( \varepsilon \)-sampled smooth curve \( C \), with \( \varepsilon < 0.47 \), HNN-CRUST outputs the manifold boundary \( B \).

**Proof** This follows immediately from Theorems 1 and 2.

8. Results

8.1. Comparison of HNN-CRUST

Figure 4 shows that unlike the CRUST \([ABE98]\), our proposed algorithm reconstructs sharp corners up to \( 60^\circ \) and handles close curves well. Our reconstruction algorithm is local and therefore scales well to large point sets. HNN-CRUST also handles open curves gracefully. It only outputs edges which are reconstructed bijectively, i.e. are consistent from both end points, in order to avoid catastrophic failure. We provide open source code for this algorithm that reproduces figures and tables of this paper; https://github.com/stefango74/hnn-crust-ssp16.

8.2. Comparison of \( \rho < 0.9 \)-sampling

![Figure 4](image4.png)

**Figure 4**: Comparison of HNN-CRUST and CRUST handles close curves well. Our reconstruction algorithm is local and therefore scales well to large point sets. HNN-CRUST also handles open curves gracefully.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sampling condition</th>
<th>Bound</th>
<th>min α</th>
<th>circle</th>
<th>par</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA/THANG</td>
<td>|p, y| &lt; |p, b|</td>
<td>( \varepsilon &lt; 0.5 )</td>
<td>( &gt; 150^\circ )</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>CRUST</td>
<td>|p, x| &lt; |p, b|</td>
<td>( \varepsilon &lt; 0.2 )</td>
<td>( &gt; 157^\circ )</td>
<td>15.5</td>
<td>5</td>
</tr>
<tr>
<td>NN-CRUST</td>
<td>|p, x| &lt; |p, b|</td>
<td>( \varepsilon &lt; \frac{1}{4} )</td>
<td>( &gt; 142^\circ )</td>
<td>9.4</td>
<td>3</td>
</tr>
<tr>
<td>NN-CRUST+</td>
<td>|p, x| &lt; |p, b|</td>
<td>( \varepsilon &lt; 0.4 )</td>
<td>( &gt; 134^\circ )</td>
<td>7.8</td>
<td>2.5</td>
</tr>
<tr>
<td>[Len06]+</td>
<td>--</td>
<td>( \varepsilon &lt; 0.48 )</td>
<td>( &gt; 124^\circ )</td>
<td>6.5</td>
<td>2.1</td>
</tr>
<tr>
<td>HNN-CRUST</td>
<td>--</td>
<td>( \varepsilon &lt; 0.47 )</td>
<td>( &gt; 126^\circ )</td>
<td>6.6</td>
<td>2.1</td>
</tr>
<tr>
<td>HNN-CRUST</td>
<td>|p, x| &lt; |p, b|</td>
<td>( \varepsilon &lt; 0.9 )</td>
<td>( &gt; 75^\circ )</td>
<td>3.4</td>
<td>1.1</td>
</tr>
</tbody>
</table>

**Table 1**: Bounds for differing sampling conditions (*=not proven), guaranteed minimum angles spanned between three adjacent samples for constant curvature and based on those the averaged number of points required to sample a circle and parallel lines with length equal to their distance. Here, \( p, x \in C \) is in the curve interval \( I(p) \) between adjacent samples \( s_0 \) and \( s_1 \), and \( s \) is any sample.
In Table 1 we compare sampling conditions w.r.t. their minimum angle and how many samples this represents on a circle or on parallel lines. Note that we derive the minimum angle for all conditions from the given bounds, except for GATHAN [DW02], which relies on additional conditions to handle sharp corners. Note especially that for constant curvature (circular arcs, parallel lines), our proposed \( \rho < 0.9 \)-sampling requires just little more than a third of the samples than \( \varepsilon < \frac{1}{3} \)-sampling.

We implemented a sampling algorithm which can apply both \( \varepsilon \)-sampling and \( \rho \)-sampling and outputs a number of samples on the input curve. The parameters \( \varepsilon, \rho \) and \( d \) (the Hausdorff distance between original curve and polygonal reconstruction) can be varied. To verify whether the edges in the reconstruction are correct, they are output as well (see Figures 9, 10 and 11). As input curves we use cubic Bezier curves and subsample them very densely to approximate the needed IFS closely at these curve points. The implementation is also available as open source online.

Figure 9 visualizes sampling different curves with \( \varepsilon < \frac{1}{3} \)-sampling and \( \rho < 0.9 \)-sampling. The number of respective samples together with \( \varepsilon < 0.47 \)-sampling are shown in Table 2.

Figure 10 shows the advantage of \( \rho < 0.9 \)-sampling over \( \varepsilon < \frac{1}{3} \)-sampling when the sampling must also ensure that the reconstructed polygon lies within Hausdorff distance \( d \) of the original curve.

Table 2 shows that a \( \rho < 0.9 \)-sampling requires many fewer samples than an \( \varepsilon < 0.47 \)-sampling, while still guaranteeing reconstruction with HNN-CRUST, approaching half of what \( \varepsilon < \frac{1}{3} \)-sampling produces. Since for curve intervals of constant local feature size the reach is equal to this IFS, circular arcs or parallel lines require only exactly half the samples in the limit. The lower bound of \( \rho < 0.9 \)-sampling is therefore \( \varepsilon < 0.9 \)-sampling, the upper bound \( \varepsilon < 0.47 \)-sampling as shown in Corollary 1.

The more drastically the IFS changes, the more samples have to be placed, approximating the limit of \( \varepsilon < 0.47 \)-sampling.

Table 3 shows how sample redundancy for \( \varepsilon \)-samplings decreases as the required Hausdorff distance between the reconstruction and original curve becomes smaller than the feature size. Note that for the BUNNY in Figure 9(b), the \( \rho < 0.9 \)-sampling requires just adding 2 samples to achieve the 1% reconstruction error (see Figure 10).

The limits of HNN-CRUST are shown in the lower half of Figure 11, where the sampling condition is violated by too close curves or too sharp corners, while its top half shows that GATHAN yields for such cases rather arbitrary results due to a lack of an intuitively understandable sampling condition. Those can be handled by specialized algorithms such as GATHAN [DW02], which rely on heuristics or global data structures such as Delaunay triangulation. Their disadvantage is that due to the heuristic criteria, they cannot give as good guarantees w.r.t. angles as ours. Also the required global data structures cannot be well partitioned for local construction, such as is possible for the kd-tree we use for determining nearest neighbors.

### Table 2: Number of samples required for the given sampling conditions (* = in the limit) for example curves and the % of redundant samples compared with \( \rho < 0.9 \) in brackets (see Figures 1 and 9).

<table>
<thead>
<tr>
<th>Model</th>
<th>( \rho &lt; 0.9 )</th>
<th>( \varepsilon &lt; 0.47 )</th>
<th>( \varepsilon &lt; \frac{1}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARALLEL</td>
<td>20</td>
<td>35 (75%)</td>
<td>48 (140%)</td>
</tr>
<tr>
<td>TEASER</td>
<td>26</td>
<td>43 (65%)</td>
<td>61 (135%)</td>
</tr>
<tr>
<td>BUNNY</td>
<td>58</td>
<td>94 (62%)</td>
<td>131 (126%)</td>
</tr>
<tr>
<td>CAT</td>
<td>180</td>
<td>254 (41%)</td>
<td>356 (98%)</td>
</tr>
</tbody>
</table>

### Table 3: Number of samples required for the given sampling conditions for the BUNNY curve and given Hausdorff distance limit in terms of maximum point set dimension, the % of redundant samples compared with \( \rho < 0.9 \) in brackets.

<table>
<thead>
<tr>
<th>Hausdorff distance</th>
<th>( \rho &lt; 0.9 )</th>
<th>( \varepsilon &lt; 0.47 )</th>
<th>( \varepsilon &lt; \frac{1}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \infty )</td>
<td>58</td>
<td>94 (62%)</td>
<td>131 (126%)</td>
</tr>
<tr>
<td>1%</td>
<td>60</td>
<td>94 (57%)</td>
<td>131 (118%)</td>
</tr>
<tr>
<td>0.3%</td>
<td>73</td>
<td>99 (36%)</td>
<td>133 (82%)</td>
</tr>
<tr>
<td>0.1%</td>
<td>105</td>
<td>123 (17%)</td>
<td>148 (41%)</td>
</tr>
<tr>
<td>0.03%</td>
<td>173</td>
<td>186 (8%)</td>
<td>204 (18%)</td>
</tr>
</tbody>
</table>
curves, \( \rho \)-sampling, has enabled us to prove a much tighter bound in terms of local sampling density. That new bound, \( \rho < 0.9 \), permits reconstruction of smooth curves with our proposed simple and fast algorithm HNN-CRUST. We believe that 0.9 is close to tight, based on Figure 8(d). The bound allows for much more sparse sampling while keeping the geometric approximation of the reconstructed polygon to the original curve. The improved \( \varepsilon \)-sampling bound already requires up to 45% fewer samples (in the limit, for constant curvature). Additionally, based on that new sampling condition, smooth curves can be reconstructed from even fewer points, typically half of the state-of-the-art bound, in the limit roughly one third. We are currently working on framing conditions to enhance our sampling framework to support non-smooth curves, as [OM13] shows they can be reconstructed for extremely sparse sampling.

Further we believe that it can be extended to handle noisy samples with outliers in the sense of [DS06]. Another work in progress is the extension of the reconstruction algorithm into \( \mathbb{R}^3 \) for surface reconstruction with a similar condition for the sampling required on a smooth boundary, together with the above enhancements. While the edge-pairs reconstructed at points in \( \mathbb{R}^2 \) correspond to closed triangle fans in \( \mathbb{R}^3 \), the output of the reconstruction algorithm does not match, as shown in [OMW13]. Flat tetrahedra can lie parallel to the surface (slivers) and so an additional condition is required to yield a unique triangulation.

Acknowledgements

We thank Tamal Dey and Marshall Bern for helpful discussions about edge angles. This work has been funded by FWF grant P24600-N23 and FP7-ICT project 323567 (HARVEST4D). San-
dia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

References


[ST09] STEELDINGER P., TCHERNIAVSKI L.: Provably correct reconstruction of surfaces from sparse noisy samples. Pattern Recognition 42, 8 (2009), 1650–1659. 2


© 2016 The Author(s)