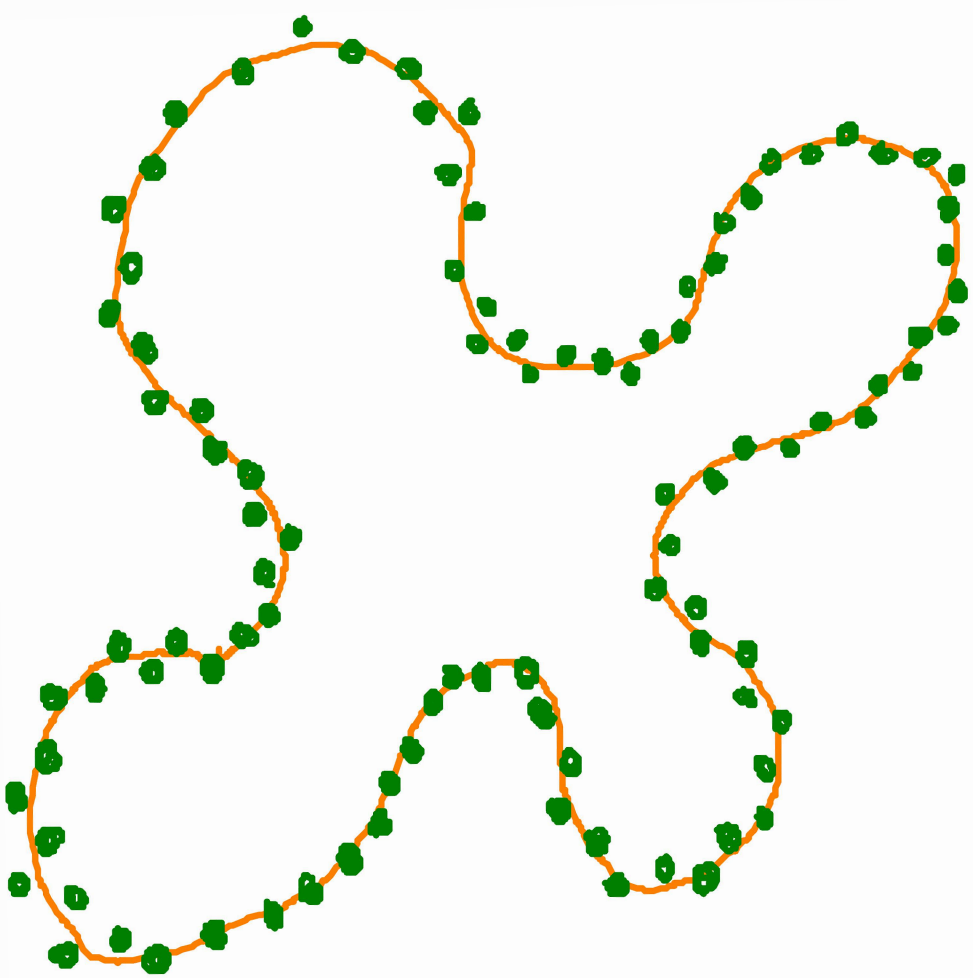
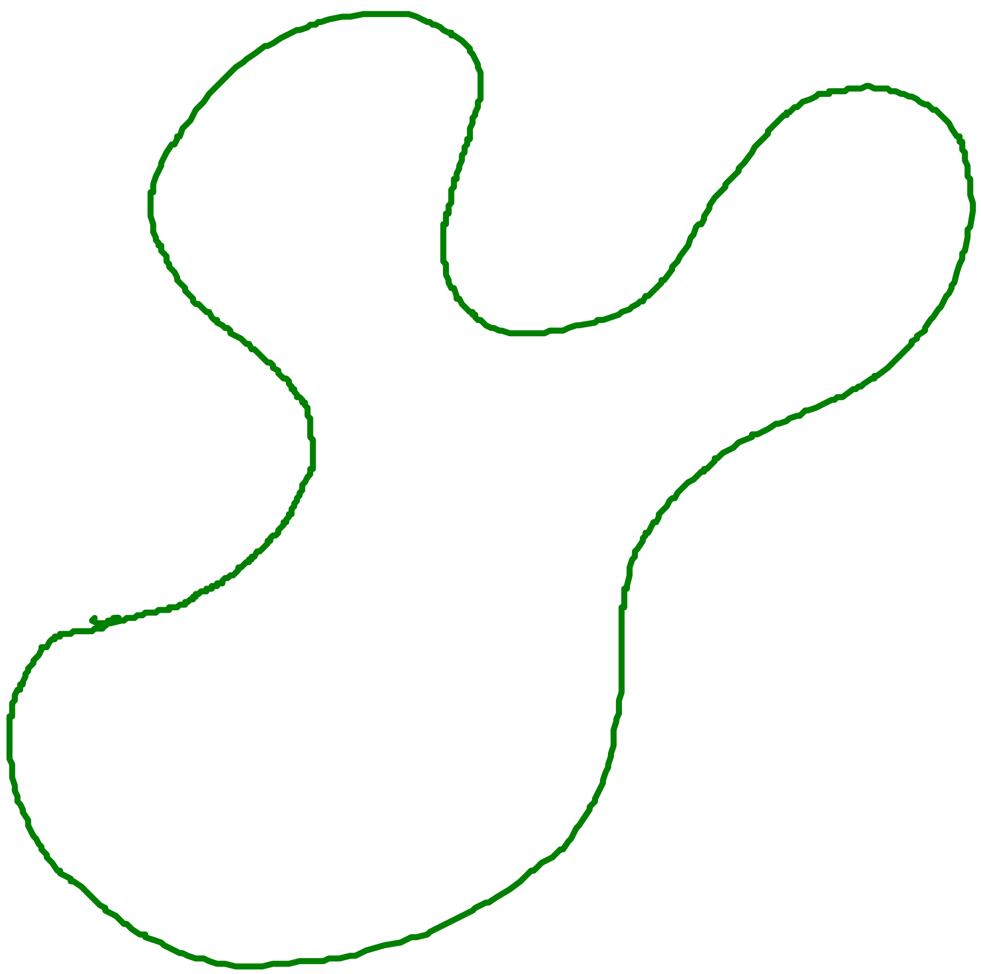
The Niyogi-Smale-Weinberger Theorem & its relatives



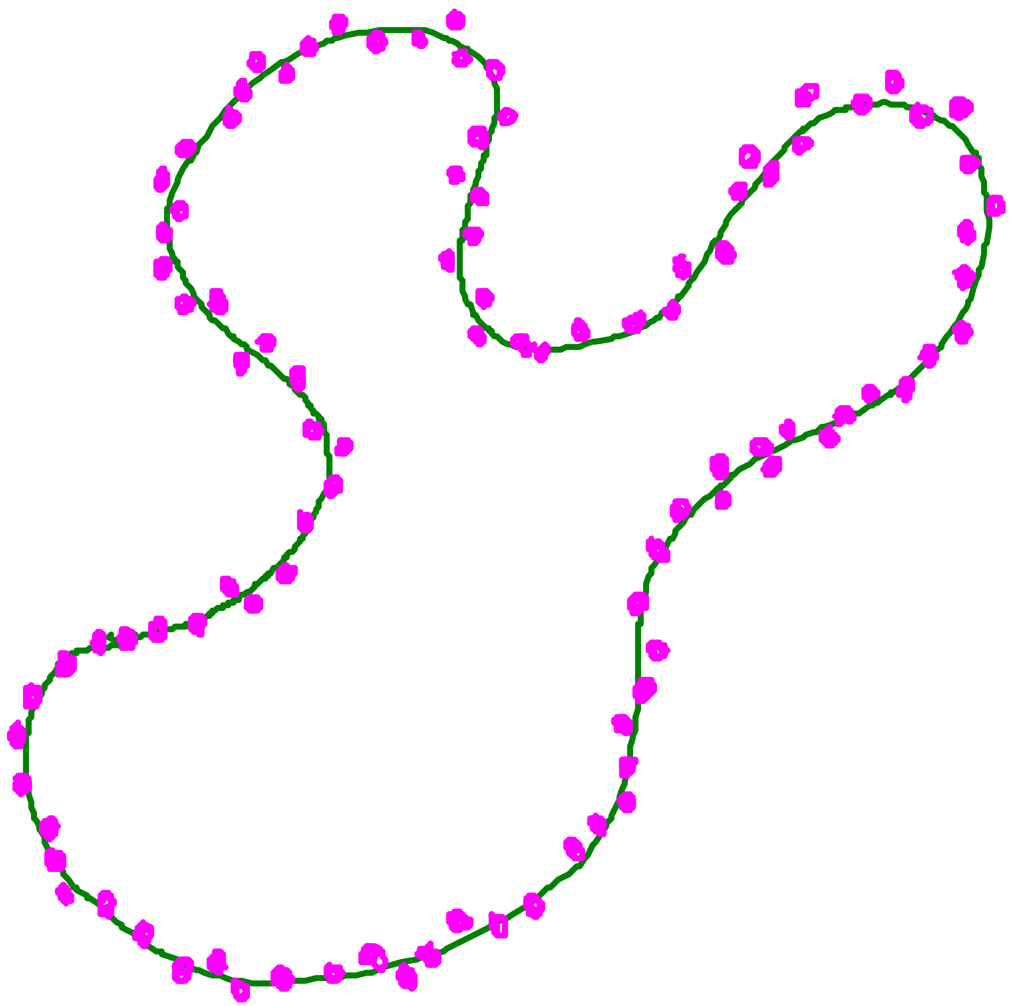
Josué TONELLI-CUETO

We want to get the topology of X...



(i) We sample points

We want to get the topology of X...



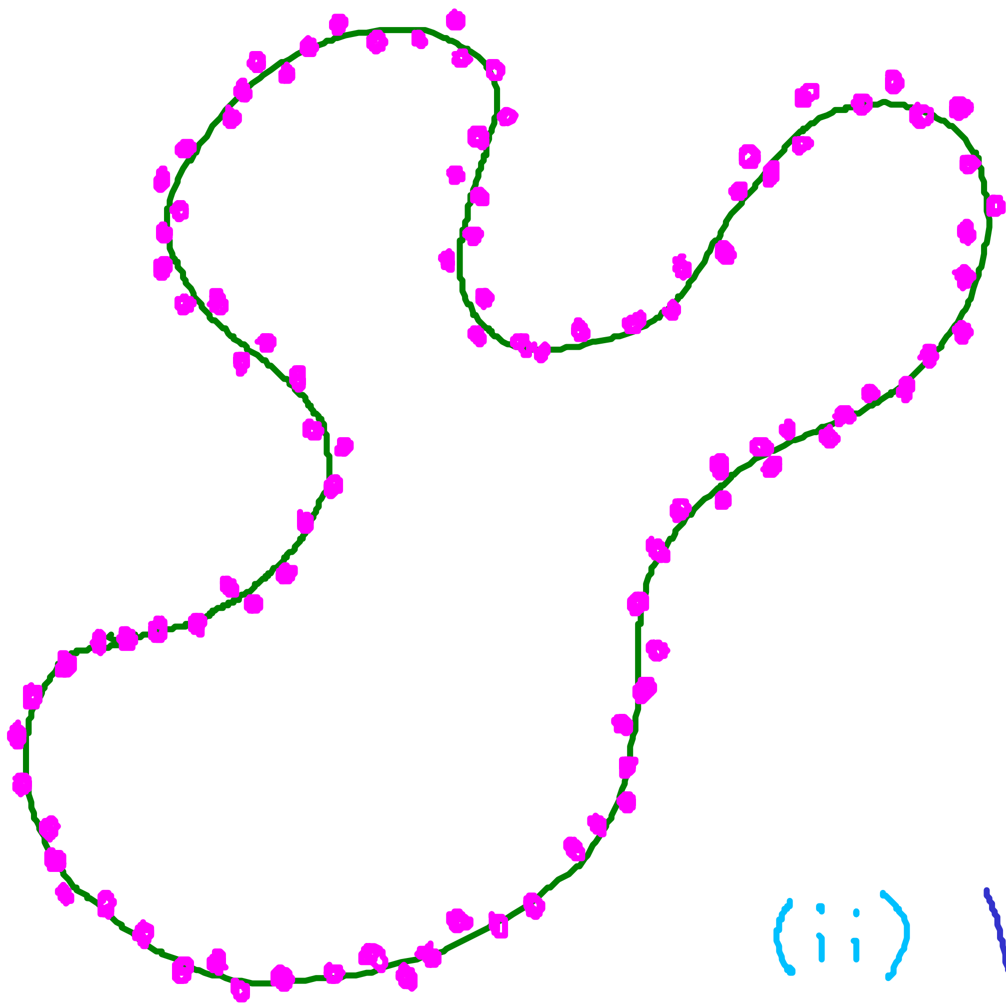
(i) We sample points

Maybe, even, more points!

How many points?

When do I have a good sample?

We want to get the topology of X...



(i) We sample points

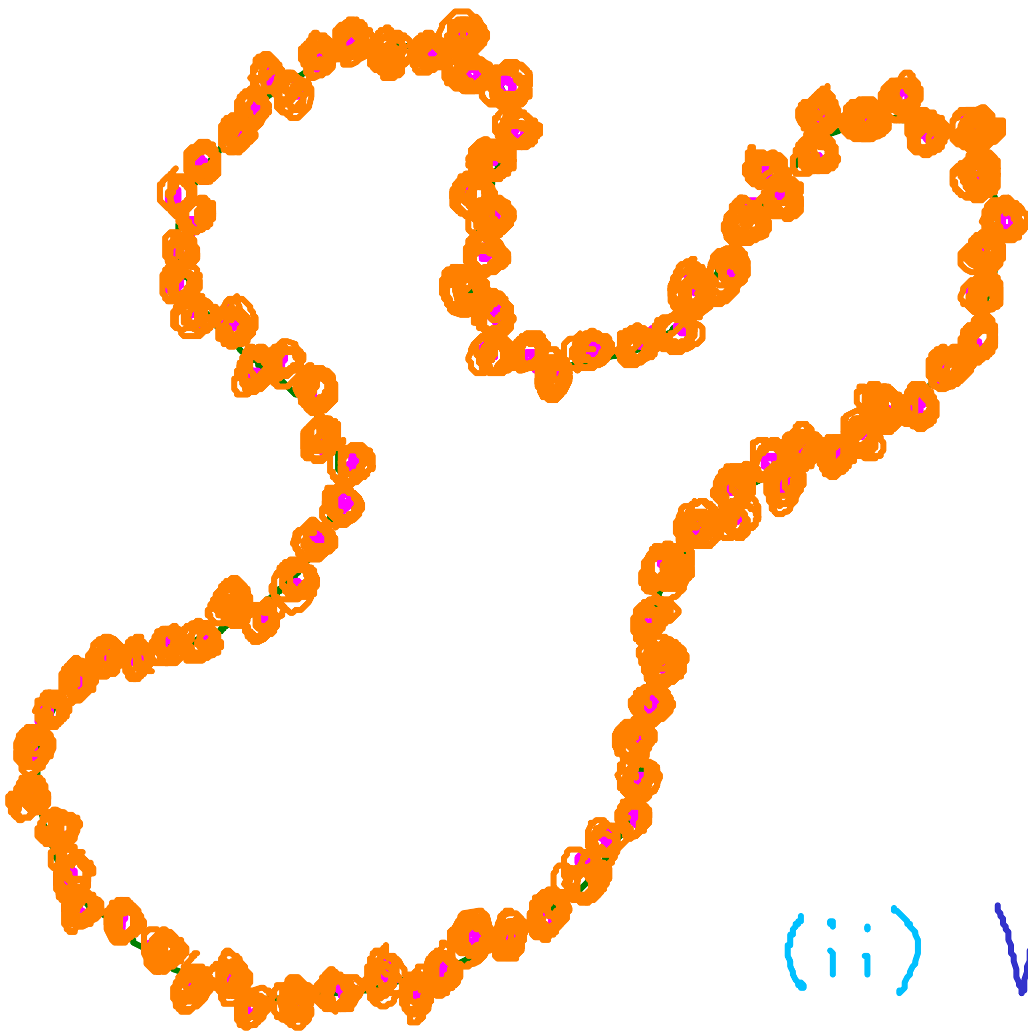
Maybe, even, more points!

How many points?

When do I have a good sample?

(ii) We fatten the points!

We want to get the topology of X...



(i) We sample points

Maybe, even, more points!

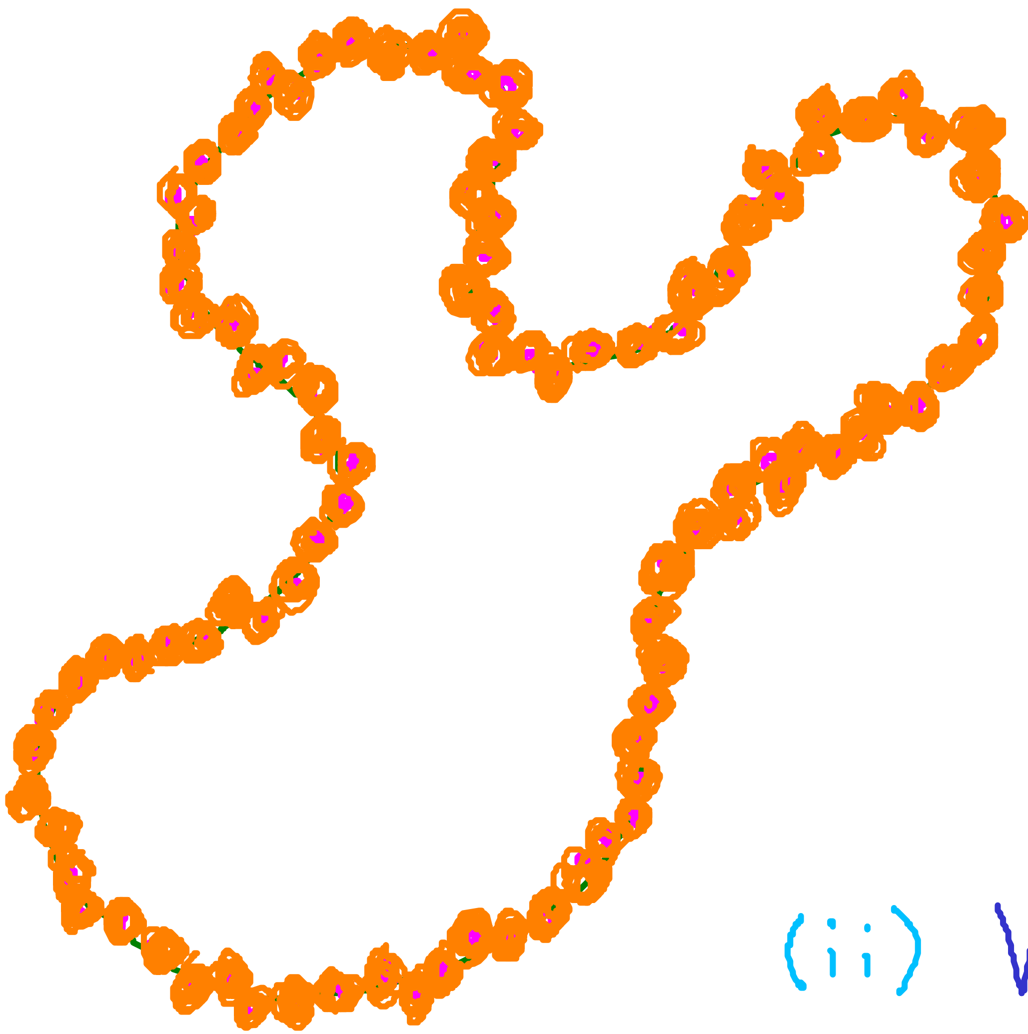
How many points?

When do I have a good sample?

(ii) We fatten the points!

How much should we fatten them?

We want to get the topology of X...



(i) We sample points

Maybe, even, more points!

How many points?

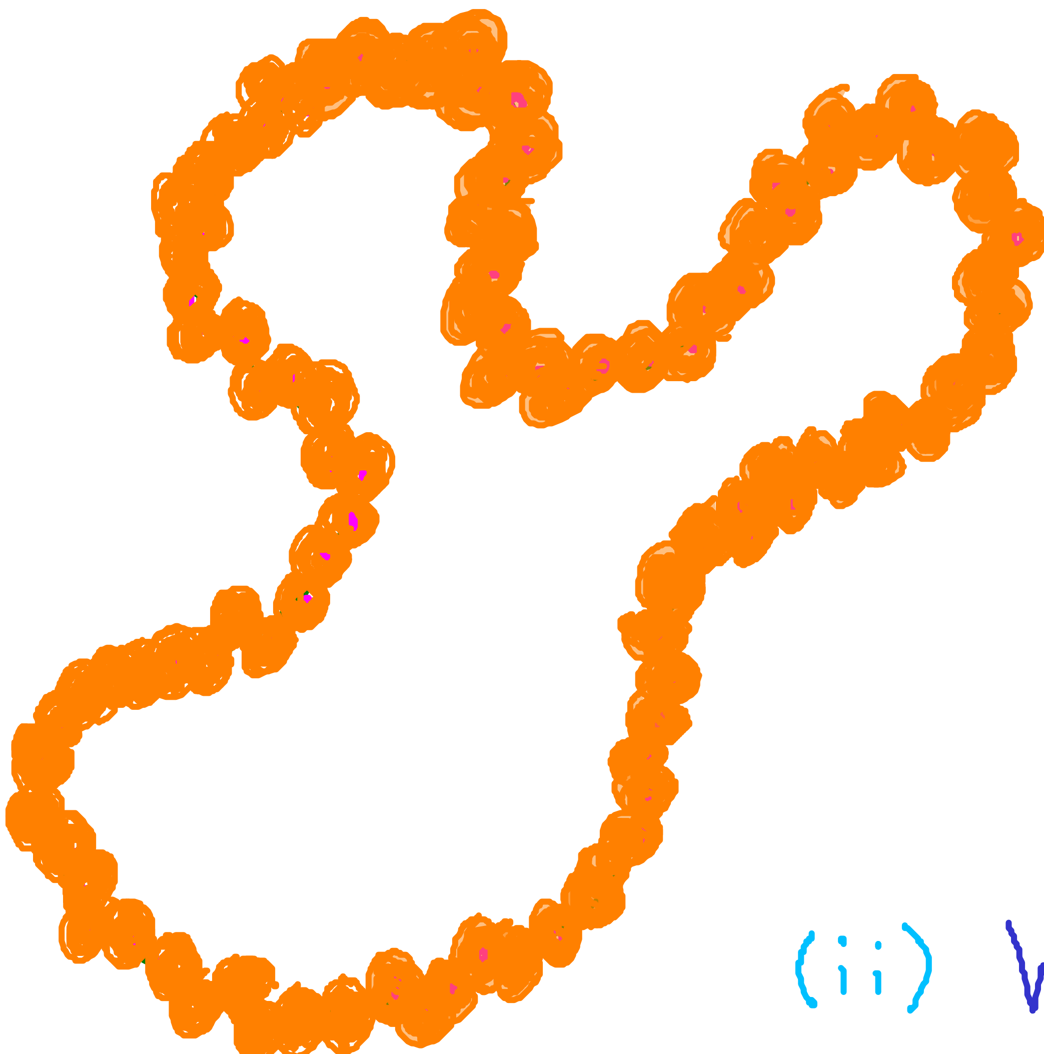
When do I have a good sample?

(ii) We fatten the points!

How much should we fatten them?

Maybe more?

We want to get the topology of X...



(i) We sample points

Maybe, even,
more points!

How many points?

When do I have a
good sample?

(ii) We fatten the points!

How much should we fatten them?

Maybe more? More?

OK, this was too much!

FORMAL QUESTION:

Given compact $X \subseteq \mathbb{R}^m$,

a finite set $S \subseteq \mathbb{R}^m$ and $\varepsilon > 0$,

under which conditions do X and

$$B(S, \varepsilon) := \{ x \in \mathbb{R}^m \mid \mathrm{dist}(x, S) \leq \varepsilon \}$$

"have the same topology"?

(i.e. are of the same homotopy type?)

⚠ The "topology" of $B(S, \varepsilon)$ is that of the Čech complex of S and ε and it can be computed...

See other tutorials for more!

Smale

That's what the NSW theorem is about!

Niyogi

Weinberger

How 'good' is the sample?

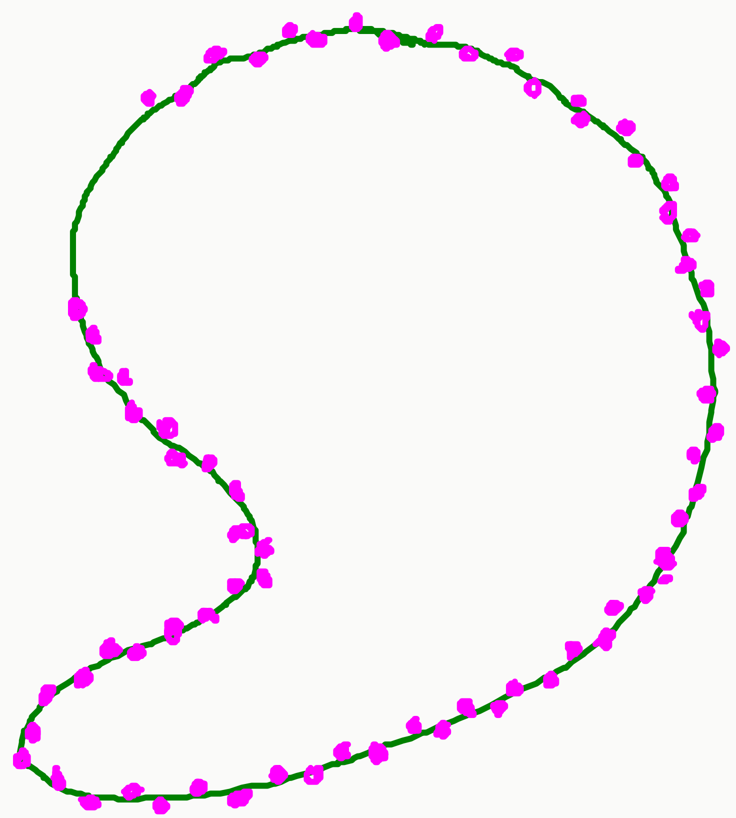
2 ingredients: 1. Hausdorff distance

2. Reach

(a.k.a. local feature size)

Is the sample 'good enough'?

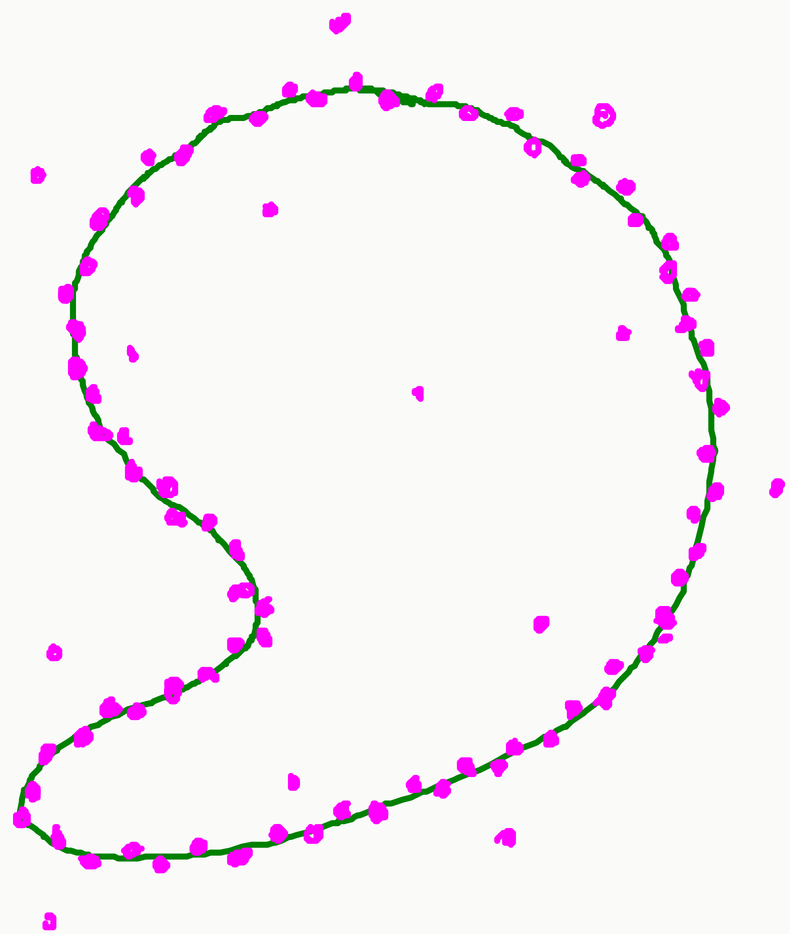
1st Ingredient: Hausdorff distance



Is this S a good sample of X?

No! Some points of X are 'too far' from S

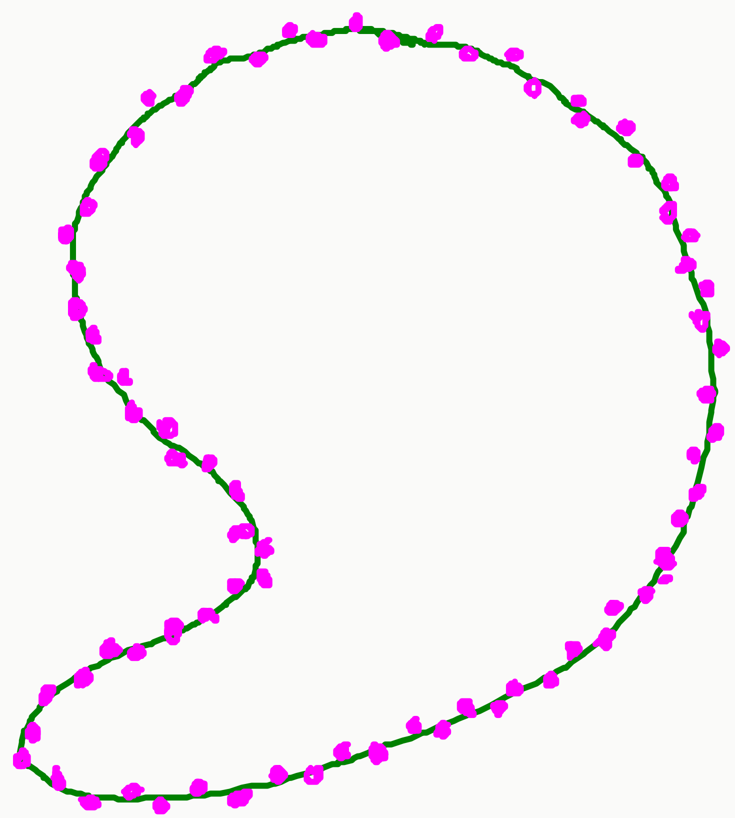
1st Ingredient: Hausdorff distance



Is this S a 'good' sample of X?

No! Some points of S are 'too far' from X

1st Ingredient: Hausdorff distance



Is this S a 'good' sample of X ?

Maybe? Every point of X is 'near' S
& every point of S is 'near' X

1st Ingredient: Hausdorff distance

$$\mathrm{dist}_H(A,B) := \max\Big\{\underbrace{\sup_{a\in A} \mathrm{dist}(a,B)}_{\text{How 'far' are the points of } A \text{ from } B\text{?}},\ \underbrace{\sup_{b\in B} \mathrm{dist}(b,A)}_{\text{How 'far' are the points of } B \text{ from } A\text{?}}\Big\}$$

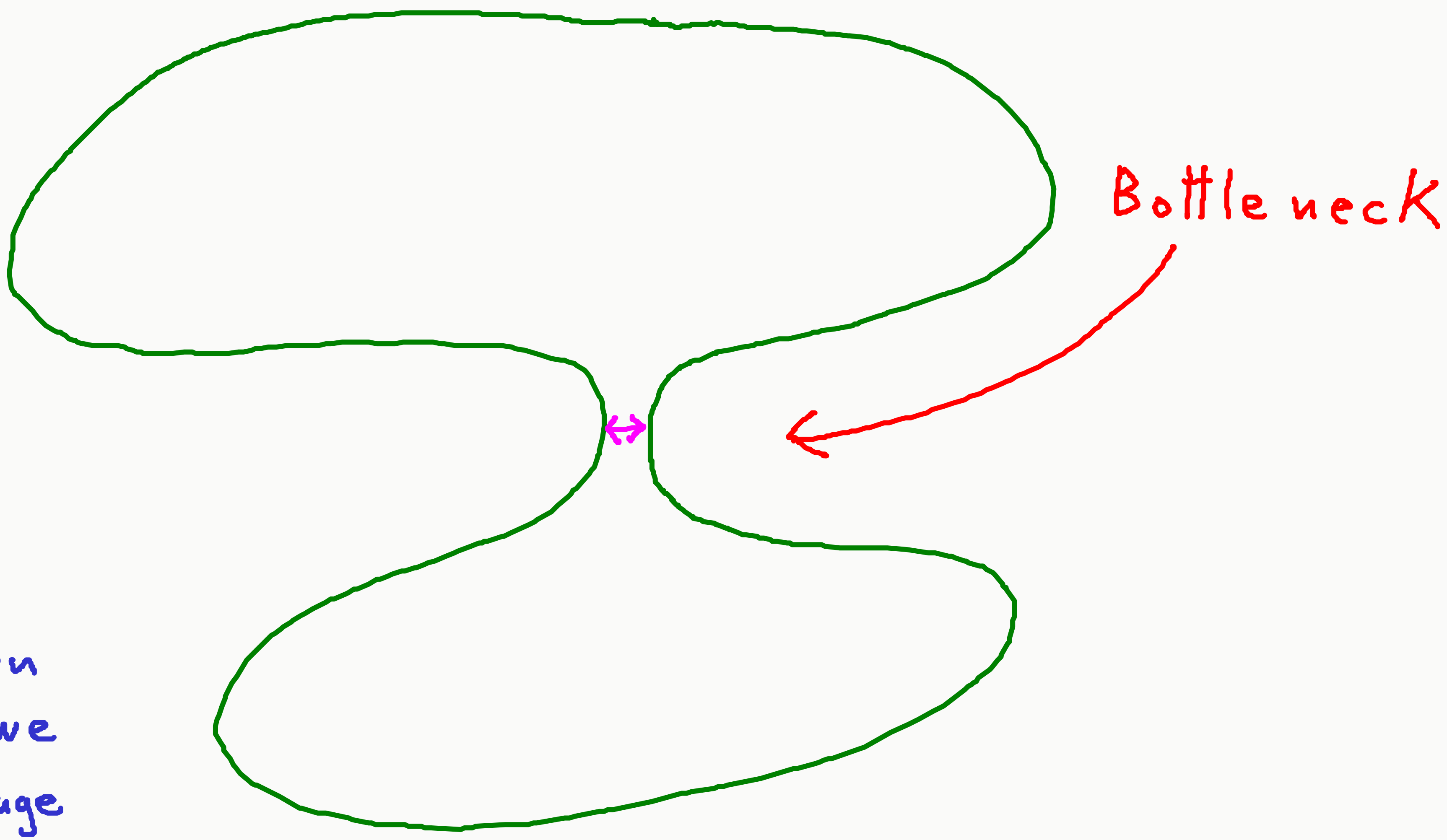
THM. dist_H is a metric on the set of non-empty compact subsets of $\mathbb{R}^m$

dist_H captures our intuitive notion of 'good sample':

$\mathrm{dist}_H(S,X)$ small '$\Leftrightarrow$' S is a 'good sample' of X

2nd Ingredient: Reach

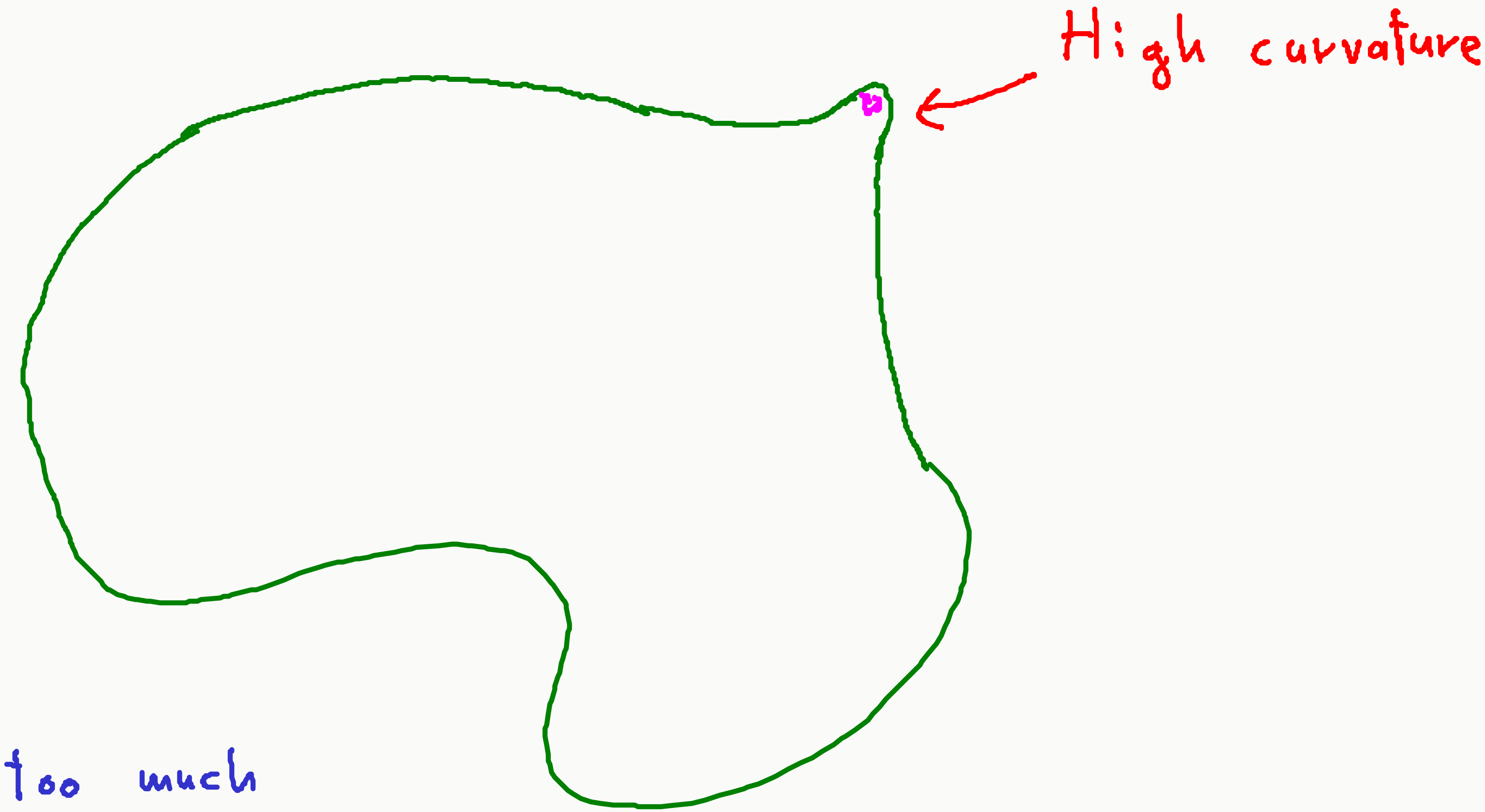
What can go wrong
when we fatten the sample S of X?



If we fatten
too much we
might change
topology!

2nd Ingredient: Reach

What can go wrong
when we fatten the sample S of X?



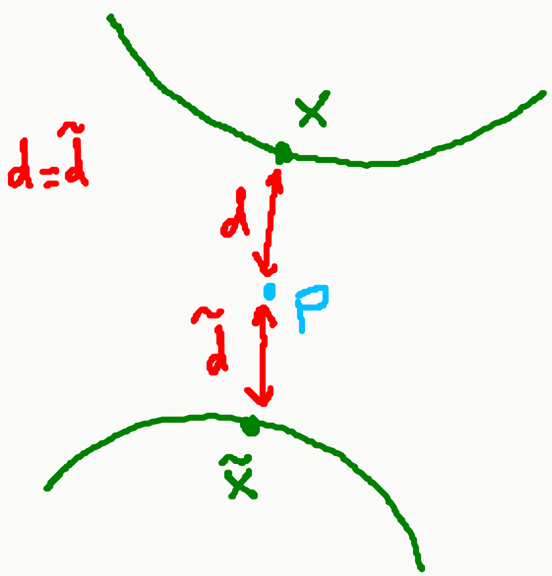
If we fatten too much
we might change topology!

2nd Ingredient: Reach

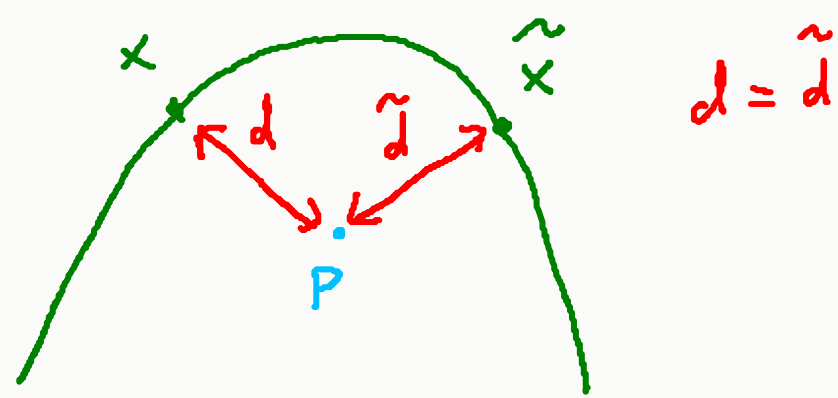
What can go wrong
when we fatten the sample S of X ?

Medial axis:

$$\Delta_X := \{ p \in \mathbb{R}^m \mid \exists\, x, \tilde{x} \in X : \underbrace{x \neq \tilde{x},\ \mathrm{dist}(p,X) = \mathrm{dist}(p,x) = \mathrm{dist}(p,\tilde{x})}_{\text{More than one nearest point in } X} \}$$



Bottleneck



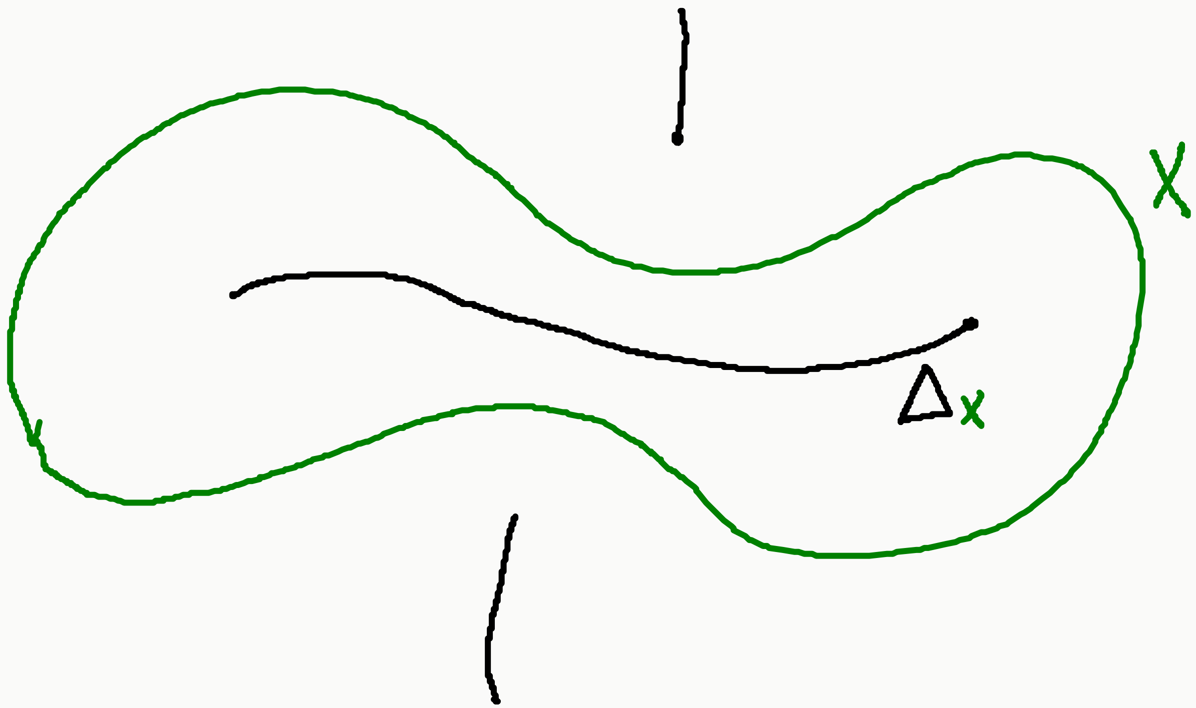
Curvature

2nd Ingredient: Reach

What can go wrong when we fatten the sample S of X?

Medial axis:

$$\Delta_X := \{p \in \mathbb{R}^m \mid \exists\, x, \tilde{x} \in X : \underbrace{x \neq \tilde{x},\ \mathrm{dist}(p,X) = \mathrm{dist}(p,x) = \mathrm{dist}(p,\tilde{x})}_{\text{More than one nearest point in } X}\}$$



2nd Ingredient: Reach

What can go wrong
when we fatten the sample S of X ?

Medial axis:

$$\Delta_X := \{ p \in \mathbb{R}^m \mid \exists\, x, \tilde{x} \in X : \underbrace{x \neq \tilde{x},\ \mathrm{dist}(p,X) = \mathrm{dist}(p,x) = \mathrm{dist}(p,\tilde{x})}_{\text{More than one nearest point in } X} \}$$

Reach:

$$\tau(X) := \inf_{x \in X} \mathrm{dist}(x, \Delta_X)$$

$\tau(X)$ measures how 'hard' is to sample X:

$\tau(X)$ 'small' $\Leftrightarrow$ X is 'hard' to sample

NSW theorem:

compact set X

finite set S

$\varepsilon > 0$

If (i) $\mathrm{dist}_H(S, X) \leq (\sqrt{9} - \sqrt{8})\,\tau(X)$ and

(ii) $$\frac{\mathrm{dist}_H(S,X) + \tau(X) - \sqrt{\mathrm{dist}_H(S,X)^2 + \tau(X)^2 - 6\,\mathrm{dist}_H(S,X)\,\tau(X)}}{2} < \varepsilon < \frac{\mathrm{dist}_H(S,X) + \tau(X) + \sqrt{\mathrm{dist}_H(S,X)^2 + \tau(X)^2 - 6\,\mathrm{dist}_H(S,X)\,\tau(X)}}{2}$$

Then

$B(S, \varepsilon)$ and X are of the same homotopy type.

NSW theorem (Easier to read version)

compact set X

finite set S

$\varepsilon > 0$

If

$$3\,\mathrm{dist}_H(S, X) < \varepsilon < \frac{1}{2}\tau(X)$$

Then

$B(S, \varepsilon)$ and X are of the same homotopy type.

Can we do more?

Yes...

+ Weak Reach (Chazal, Lieutier; 2005)
+ Balls of different radii
  (Chazal, Lieutier; 2007) (Han; 2019) (Eckhardt; 2020)
+ Vietoris-Rips complex
  (Attali, Lieutier, Salinas; 2012)
+ Ellipsoids (Kališnik, Lešnik; 2020+)

... and much more!

D. Attali, A. Lieutier, and D. Salinas.
Vietoris-Rips complexes also provide topologically correct reconstructions of sampled shapes.
Comput. Geom., 46(4):448–465, 2013.

F. Chazal and A. Lieutier.
Weak feature size and persistant homology: computing homology of solids in $\mathbb{R}^n$ from noisy data samples.
In *Computational geometry (SCG'05)*, pages 255–262. ACM, New York, 2005.

F. Chazal and A. Lieutier.
Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees.
Comput. Geom., 40(2):156–170, 2008.

A. Eckhardt.
An Adaptive Algorithm for Computing the Homology of Semialgebraic Sets.
Master's thesis, Technische Universität Berlin, 2020.

J. Han.
An Adaptive Grid Algorithm for Computing the Homology Group of Semialgebraic Set.
Master's thesis, Université Paris Sud, 2018.

S. Kališnik and D. Lešnik.
Finding the homology of manifolds using ellipsoids, 2020.
arXiv:2006.09194.

P. Niyogi, S. Smale, and S. Weinberger.
Finding the homology of submanifolds with high confidence from random samples.
Discrete Comput. Geom., 39(1-3):419–441, 2008.