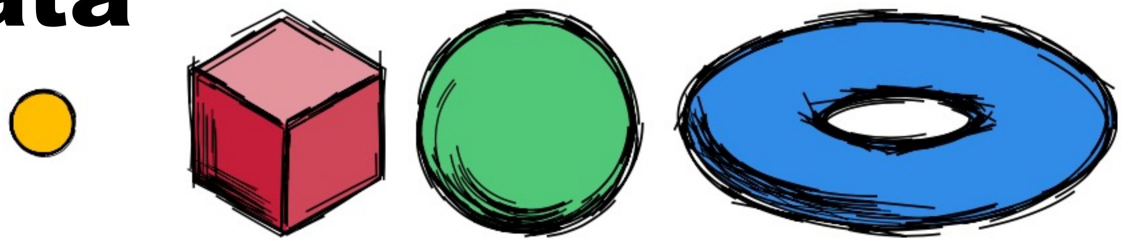
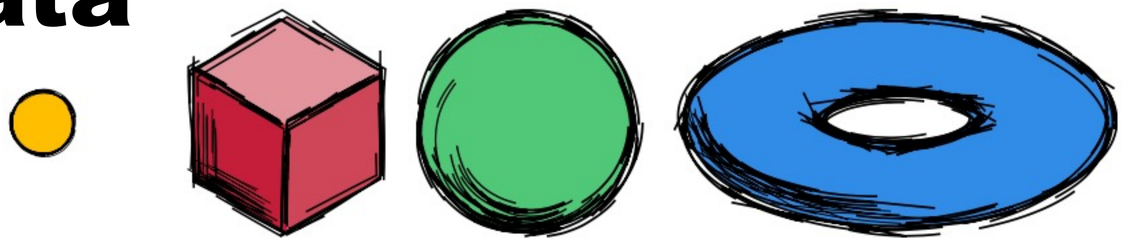
Shape of data

The shape of data



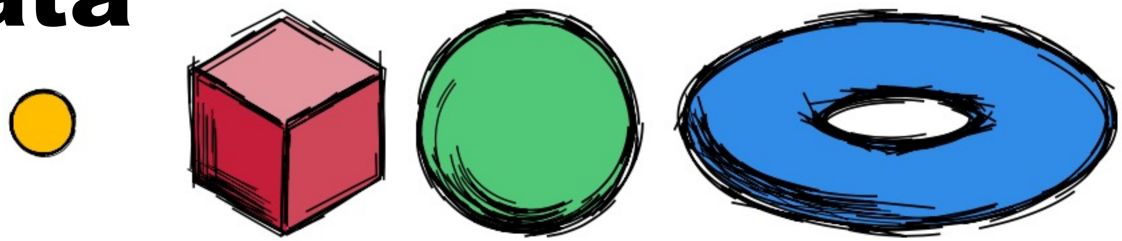
The shape of data



Betti numbers

β_0 Components
β_1 1-dim circular holes
β_2 voids

The shape of data



Betti numbers

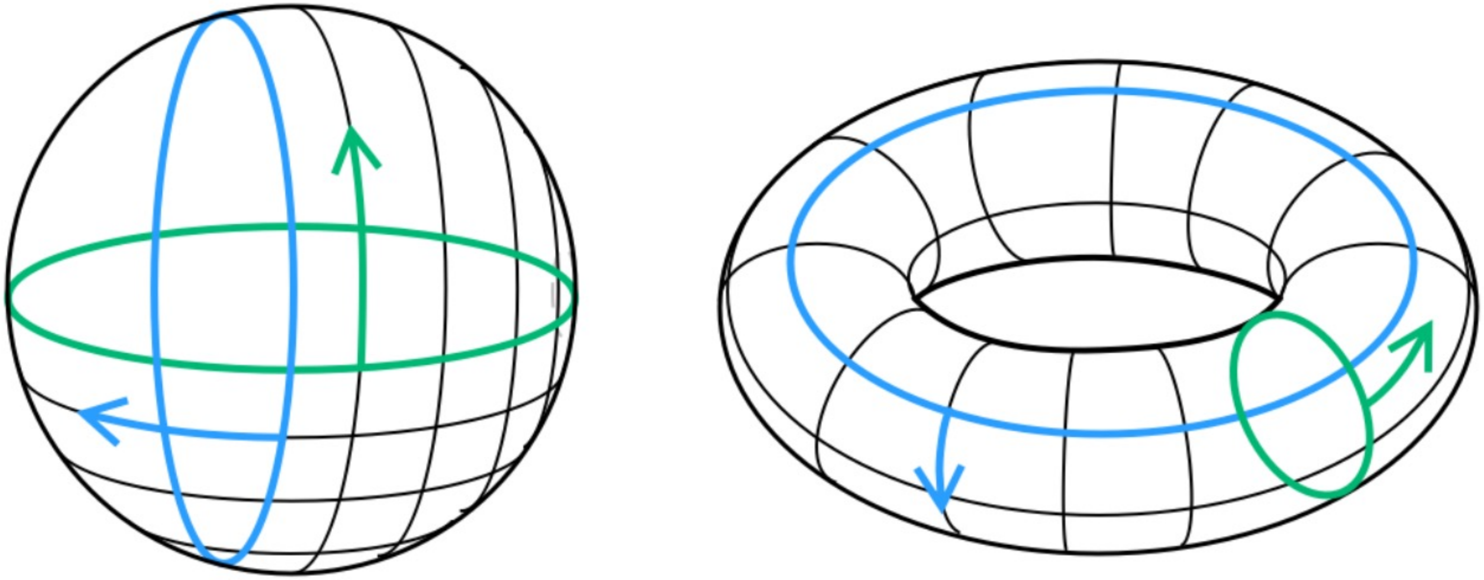
β_0 Components
β_1 1-dim circular holes
β_2 voids

Space	β_0	β_1	β_2
Point	1	0	0
Cube	1	0	1
Sphere	1	0	1
Torus	1	2	1

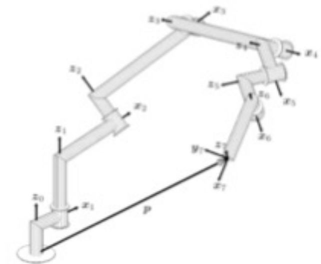
1895 Poincare
1925 E. Noether
2000 Data analysis

Why?

Space	β_0	β_1	β_2
Point	1	0	0
Cube	1	0	1
Sphere	1	0	1
Torus	1	2	1

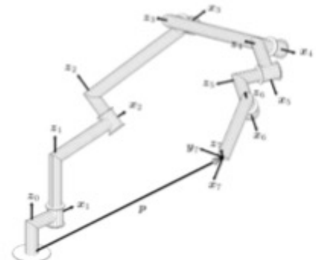


Algebraic models of data



6R IKP: general fiber of $\underbrace{\mathbb{P}^1 \times \ldots \times \mathbb{P}^1}_{\text{6 times}} \to Q \subset \mathbb{P}^7$

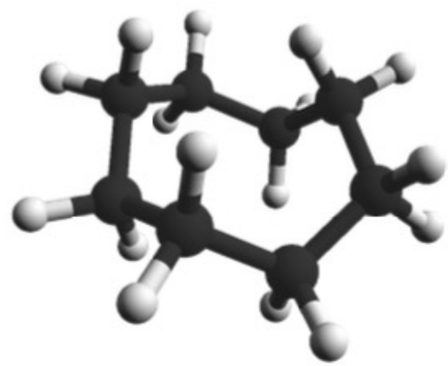
Algebraic models of data



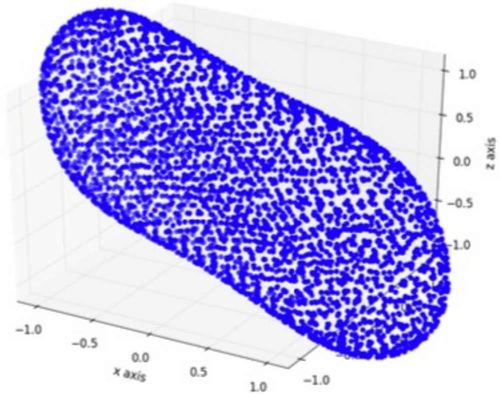
6R IKP: general fiber of $\underbrace{\mathbb{P}^1 \times \ldots \times \mathbb{P}^1}_{\text{6 times}} \to Q \subset \mathbb{P}^7$

Configuration space of Cyclooctane

13 equations in 15 varaiables

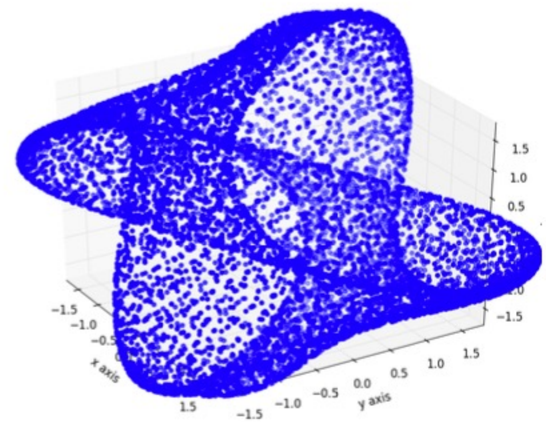


Cyclooctane solutions



Spherical component

β_0 1
β_1 0
β_2 1



Klein component

β_0 1
β_1 2
β_2 0

Two main questions

2019-2022

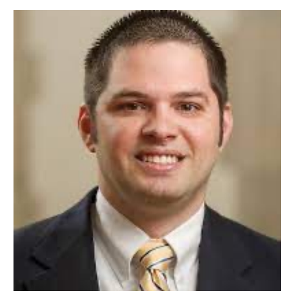
1. Provide criteria for **density** of sampling
2. Give algorithms for **sampling**

D. Eklund
RISE

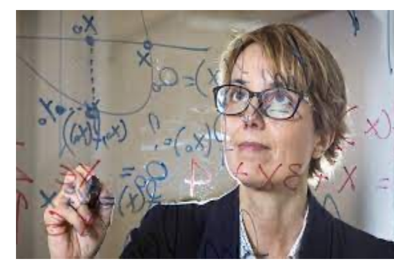
M. Weinstein
Standford

P. Edwards
Notre Dame

O. Gävfert
KTH->Oxford



Jon Hauenstein
Notre Dame



S. Di Rocco
KTH

Vetenskapsrådet

Knut och Alice Wallenbergs Stiftelse

Dansity

Definition
A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Dansity

Definition

A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Consider the Čech complex, $C(E, \varepsilon) = \mathscr{R}_\varepsilon$, given by the nerve of the union of all closed ε-balls $\overline{B}_e(\varepsilon)$.

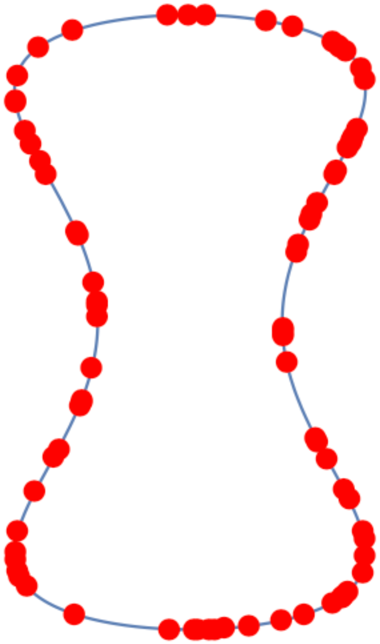
Dansity

Definition
A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Consider the Čech complex, $C(E,\varepsilon) = \mathscr{R}_\varepsilon$, given by the nerve of the union of all closed ε-balls $\overline{B}_e(\varepsilon)$.

The sample E gives a topological signature of the variety if the homology of X can be computed from the homology of $C(E,\varepsilon)$.

Dansity

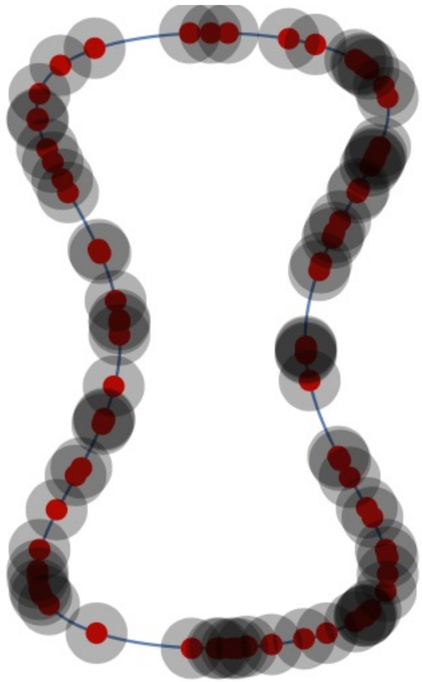


Definition
A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Consider the Čech complex, $C(E,\varepsilon) = \mathscr{R}_\varepsilon$, given by the nerve of the union of all closed ε-balls $\overline{B}_e(\varepsilon)$.

The sample E gives a topological signature of the variety if the homology of X can be computed from the homology of $C(E,\varepsilon)$.

Dansity

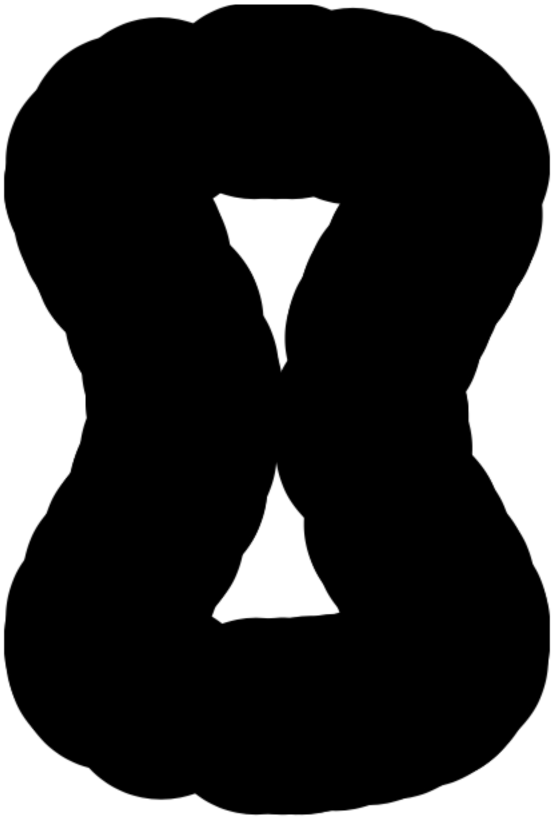


Definition
A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Consider the Čech complex, $C(E,\varepsilon) = \mathscr{R}_\varepsilon$, given by the nerve of the union of all closed ε-balls $\overline{B}_e(\varepsilon)$.

The sample E gives a topological signature of the variety if the homology of X can be computed from the homology of $C(E,\varepsilon)$.

Dansity



Definition

A *sample* of a variety $X \subset \mathbb{R}^n$ is a finite subset $E \subset X$. For $\varepsilon > 0$ a sample $E \subset X$ is called an ε-sample if for every $x \in X$ there is an element $e \in E$ such that $\|x - e\| < \varepsilon$.

Consider the Čech complex, $C(E, \varepsilon) = \mathscr{R}_\varepsilon$, given by the nerve of the union of all closed ε-balls $\overline{B}_e(\varepsilon)$.

The sample E gives a topological signature of the variety if the homology of X can be computed from the homology of $C(E, \varepsilon)$.

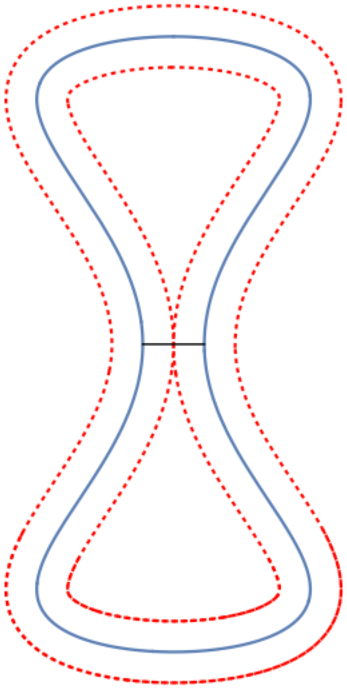
Dansity

Theorem
The homology groups of X can be recovered from an ε-sample if $\varepsilon < \tau_X$.

- Niyogi-Smale-Weinberger 2005 Discrete Comp. Geom.
- Chazal and Lieutier 2005
- Cohen-Steiner et al 2007
- Cucker-Krick-Shub, 2018 Fo.C.M
- Bürgisser-Cucker-Lairez 2019 J.ACM

Dansity

Bottleneck degree



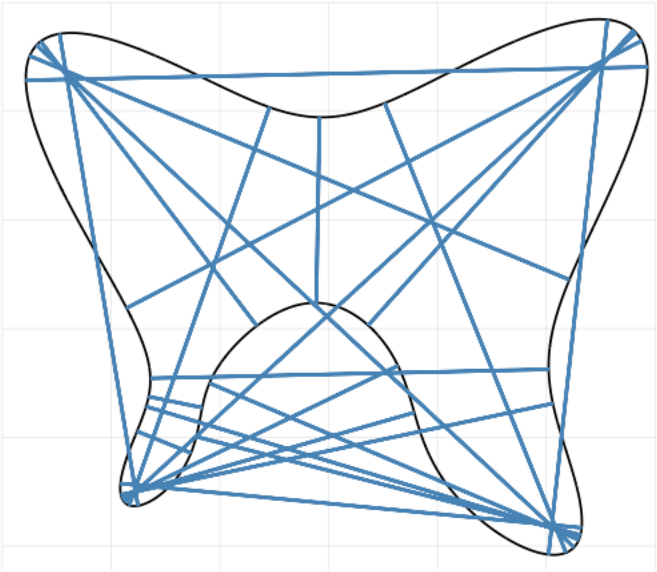
Theorem
The homology groups of X can be recovered from an ε-sample if $\varepsilon < \tau_X$.

- Niyogi-Smale-Weinberger 2005 Discrete Comp. Geom.
- Chazal and Lieutier 2005
- Cohen-Steiner et al 2007
- Cucker-Krick-Shub, 2018 Fo.C.M
- Bürgisser-Cucker-Lairez 2019 J.ACM

- Di Rocco-Eklund-Weinstein 2020 SIAM J. of A.G.
- Di Rocco-Eklund-Gäfvert 2022 J. of Math Comp
- Di Rocco-Edward-Eklund-Gäfvert-Hawenstein 2022

The bottleneck of a curve

Quartic curve $x^4 + y^4 + 1 - 4y - x^2y^2 - 4x^2 - x - 2y^2 = 0$ in $\mathbb{R}^2$ and its 22 bottleneck lines.



Theorem (DR-Eklund-Weinstein, 2020)

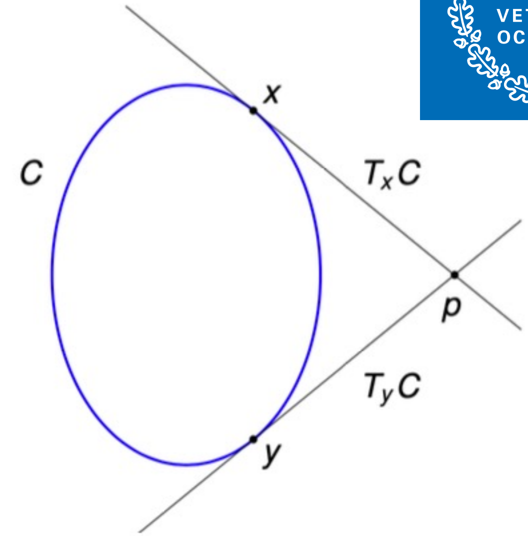
Let $X \subset \mathbb{P}^n$ be a smooth m-dimensional variety in **general position**. *Let* $k = \min\{\left\lfloor \frac{n-1}{2} \right\rfloor, m\}$ *and for* $i = 0, \ldots, k$, *put* $\varepsilon_i = \sum_{j=r_i}^{m-i} \deg p_j$ *where* $r_i = \max\{0, m-n+1+i\}$.

$$\mathbf{BND}(\overline{\mathbf{X}}) = \sum_{i=0}^{k} \varepsilon_i^2 - \deg B_{m,n},$$

for some polynomial $B_{m,n}$ in the polar classes and the hyperplane class of X.

Polar classes

- 1070-1910 Segre, Noether
- 1920-1940 Severi, Todd
- 1970- Kleiman, Fulton
- 2000- June Hu

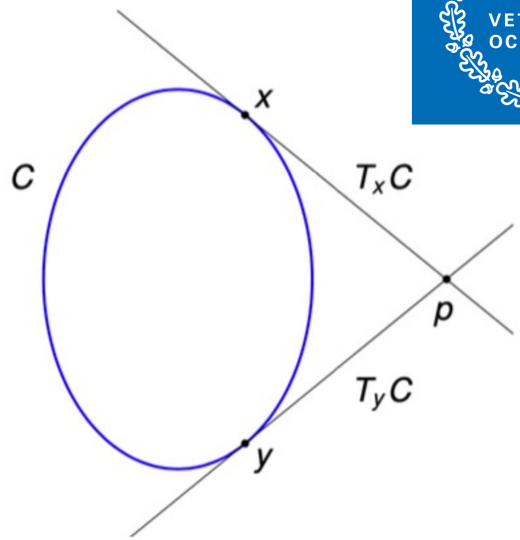


- $P_1(X,p) = \{x,y\}$ is the first polar locus of the ellipse.
- The polar locus depends on the choice of p but two polar loci $P_1(X,p)$ and $P_1(X,p')$ represent the same *polar class* p_1 on C.

Polar classes

- 1070-1910 Segre, Noether
- 1920-1940 Severi, Todd
- 1970- Kleiman, Fulton

- 2000- June Hu



- $P_1(X,p) = \{x,y\}$ is the first polar locus of the ellipse.
- The polar locus depends on the choice of p but two polar loci $P_1(X,p)$ and $P_1(X,p')$ represent the same *polar class* p_1 on C.

THANKS!