Hyperbolic geometry of the olfactory space

Yuansheng Zhou, Brian H. Smith, Tatyana O. Sharpee

Daiyuan Li

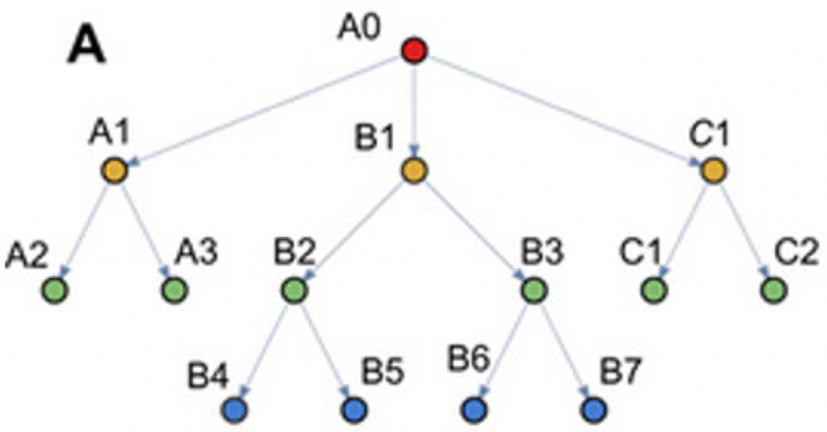
Week 13

April 12, 2022

- **Biological background**
- A statistical approach to generate Betti curves
- A topological method for finding geometric spaces consistent with correlation matrices
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- Visualization of the natural olfactory space using nonmetric MDS
- What are the axes of this olfactory space?
- The Betti curves of human olfactory perception

Background

- Biological data are often represented by hierarchical tree structures
- These data can be equivalently represented using Venn diagrams



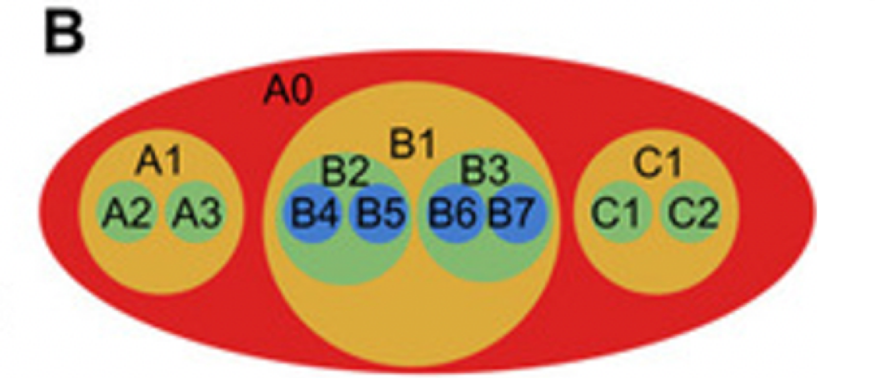
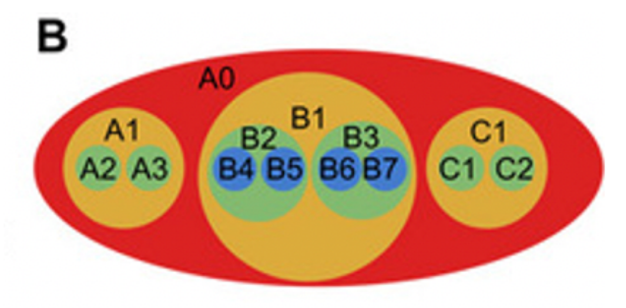


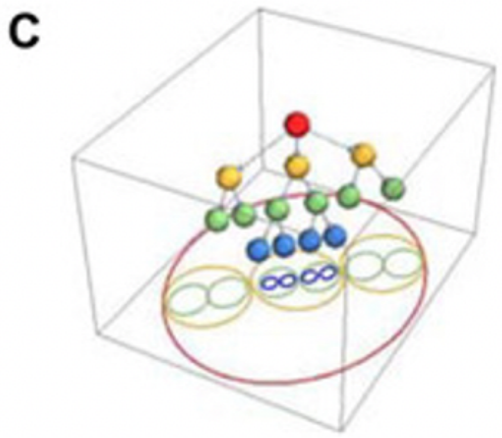
Figure1. (A) example hierarchical description of data (B) equivalent representation using Venn diagrams

Background

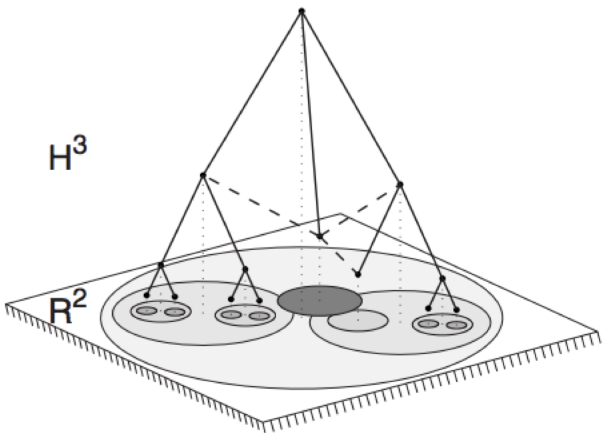
- Venn diagrams can be mapped onto points in a 3D space, **approximately** forming a tree
- Map center coordinates of the Venn circles to x and y coordinates in the 3D space, and map the circle radius to the vertical coordinate
- The metric in the resulting 3D space is hyperbolic and the resulting 3D space can be described by the Poincare half-space model



(B) Venn diagrams



(C) Mapping Venn onto 3D points



Structure with loops

Poincare half-space model

- Poincare half-space model for the hyperbolic space

Point sets: $\{(x, y, z): z > 0\}$

Metric: $\dfrac{dx^2+dy^2+dz^2}{z^2}$

- Non-Euclidean metric: rejects the Euclid's fifth validity, the “parallel”
  - Red dashed line: shortest path between two points in a 2D half-space model
  - Red solid line: discrete approximation

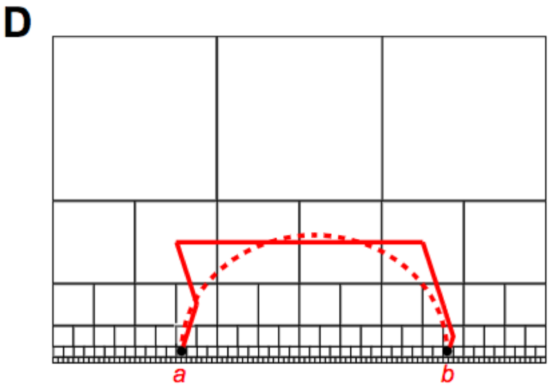


Fig. 1 (D) Discrete and continuous shortest paths to a half-space model of the hyperbolic space in 2D.

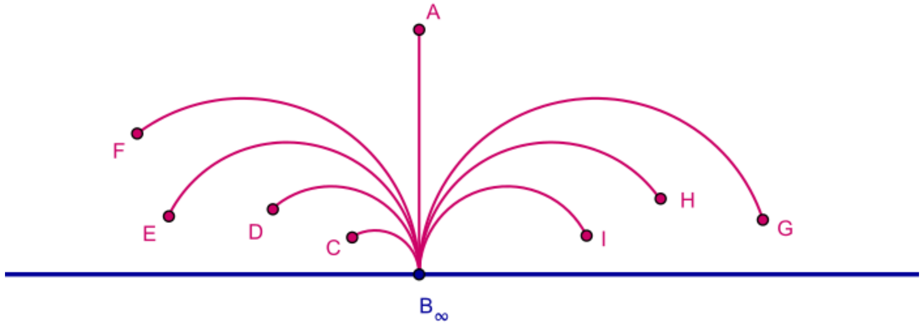


Fig.1 (Extra) Parallel rays in Poincare half-plane model of hyperbolic geometry

Data sets

1. From statistics:

- **Single natural odors** (monomolecular odors data sets)
- Examples: alcohol, ester, ketone, furan
- **Natural odor mixture** (individual samples)
- Examples: tomato, strawberry, blueberry, mouse urine

2. From previous research:

- **Human olfactory perception**(underlying set of coordinates that can represent olfactory mixtures)
- The Dravnieks database is used in this paper

Odors reflect leaves of the tree network.

- Biological background
- **A statistical approach to generate Betti curves**
- A topological method for finding geometric spaces consistent with correlation matrices
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- Visualization of the natural olfactory space using nonmetric MDS
- What are the axes of this olfactory space?
- The Betti curves of human olfactory perception

Statistical method

- Statistical data sets: concentrations of individual odors
- 4 odors data sets:

	mouse urine	tomato	blueberry	strawberry
# samples	50	79	101	54
# odors	69	66	45	78

- Step1: compute correlations between the concentrations of odor across samples to construct a correlation matrix
  - Correlation coefficient between two odors x, y:

$$corr(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

  - x_i & y_i:concentrations of odor x and odor y in $i-th$ sample
  - $\bar{x}$ & $\bar{y}$: the mean values of concentrations across all samples
  - n: the number of samples

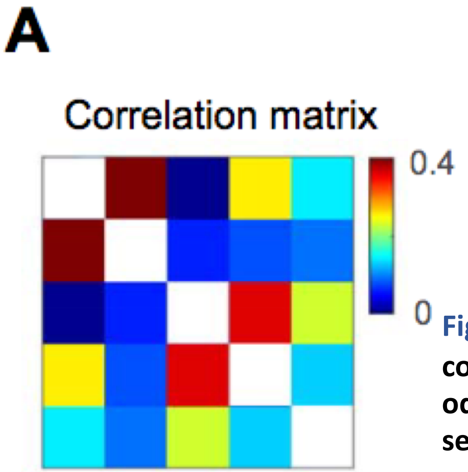


Fig. 2: (A) Example correlation matrix for five odors in strawberry data set

Statistical method

- Step2: set different thresholds. The correlation coefficients (in absolute value) above the threshold are set to 1, and the rest of values are set to 0
- Step3: visualize the matrix as a topological graph where all unit values represent links between the corresponding odors. The graph can be characterized by Betti numbers (# cycles)

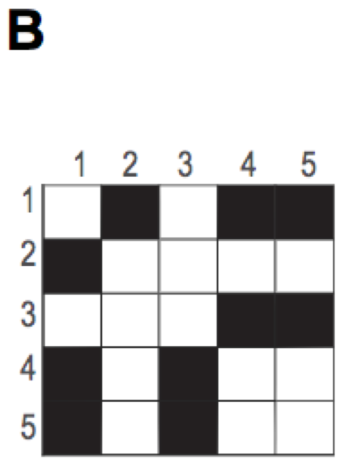
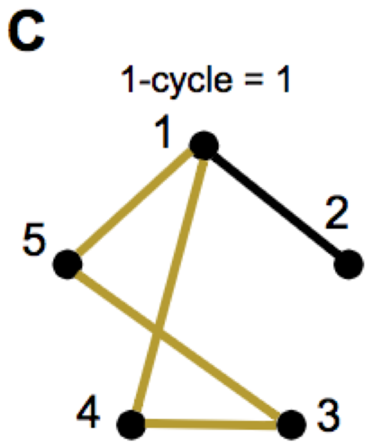


Fig. 2 (B) Correlation matrix after applying a threshold of 0.25 of Fig. 2(A)



(C) Resulting complex

Statistical method

- Step4: plot the Betti numbers as a function of edge density to generate the Betti curves
  - Density of edges ρ: The ratio of the actual connected nodes (after applying the threshold) to all possible connected nodes(all odor pairs)
  - For a $\mathrm{n}\times\mathrm{n}$ correlation matrix, $\rho = \dfrac{\#\,\text{connected nodes}}{\mathrm{n}(\mathrm{n}-1)/2}$

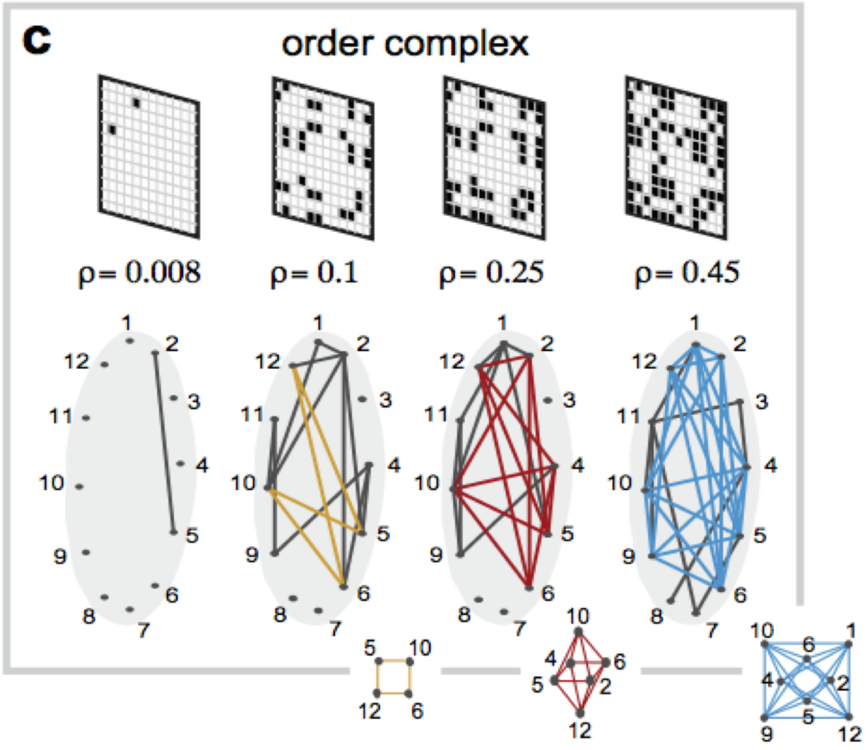


Fig.2 (Extra) A sequence of matrices with different edge densities

Resulting Betti curves

- Yellow, red and blue curves represent the 1st, 2nd, and 3rd Betti curves, respectively
- Dashed lines are Betti curves derived from our data sets
- Solid curves in (D)&(E) are predicting Betti curves using model geometry of 3D hyperbolic space in or Euclidean space, respectively
-

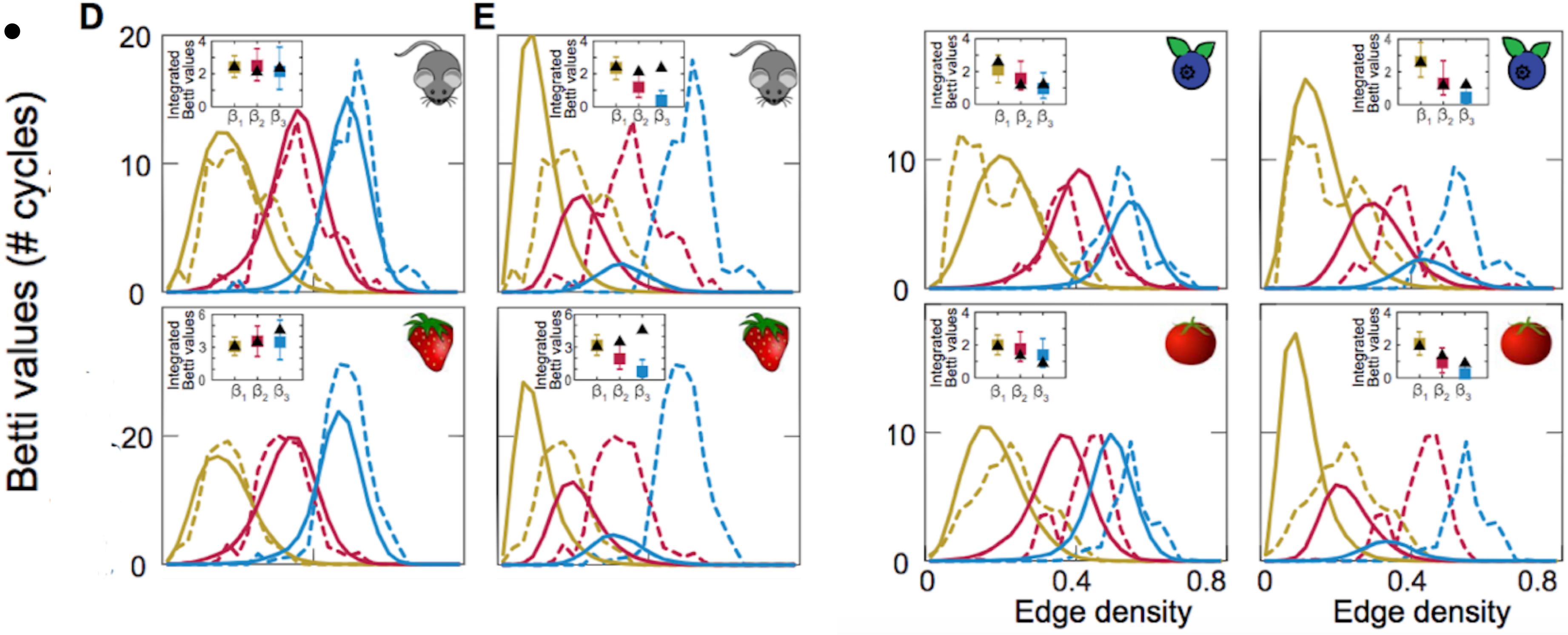


Fig.2 **(D)&(E) Betti curves with the number of cycles different dimensions plotted as a function of edge densities.**

- Biological background
- A statistical approach to generate Betti curves
- **A topological method for finding geometric spaces consistent with correlation matrices**
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- Visualization of the natural olfactory space using nonmetric MDS
- What are the axes of this olfactory space?
- The Betti curves of human olfactory perception

Uniformly distribution and distance

- In each space, uniformly sample points(# points = # odors in each data set) based on the metric of the space
- For a d-D Euclidean space, uniformly distribute points in a d-cube with Euclidean distance
- For the 3D hyperbolic space, use partial space by setting the minimal radius R_{min} and maximal radius R_{max} of the hyperbolic sphere
- Take the angles of points randomly and selected radii r within $[R_{min}, R_{max}]$ following the distribution:

$$\rho(r) \sim \sinh((d-1)r)$$

- The distance between two points for hyperbolic geometry (derived from hyperbolic law of cosines

$$\cosh x = \cosh(r)\cosh(r') - \sinh(r)\sinh(r')\cos\Delta\theta$$

  - r and r' are the radial distances of the two points
  - $\Delta\theta$ is the angle between two points

Multiplicative noise

- Considering that noise may exist in the statistical data sets and the amount of noise may differ between different data sets, before plotting the Betti curves, add multiplicative noise to the distance matrices for both Euclidean model and hyperbolic model to make the prediction data sets more fitting the statistical data sets
- Finally, there is the topological distance matrices D of the sampled points in geometric spaces

$$D = D_{geo} \cdot (1 + \varepsilon \cdot N(0,1))$$

  - D_{geo} is geometric distance matrices
  - ε is the noise level

Betti curves for single odor data sets

- Dashed lines are Betti curves derived from our data sets
- Solid curves in (D)&(E) are predicting Betti curves using model geometry of 3D hyperbolic space in or Euclidean space, respectively
- The first three Betti curves are the most sensitive measures of the distance matrices

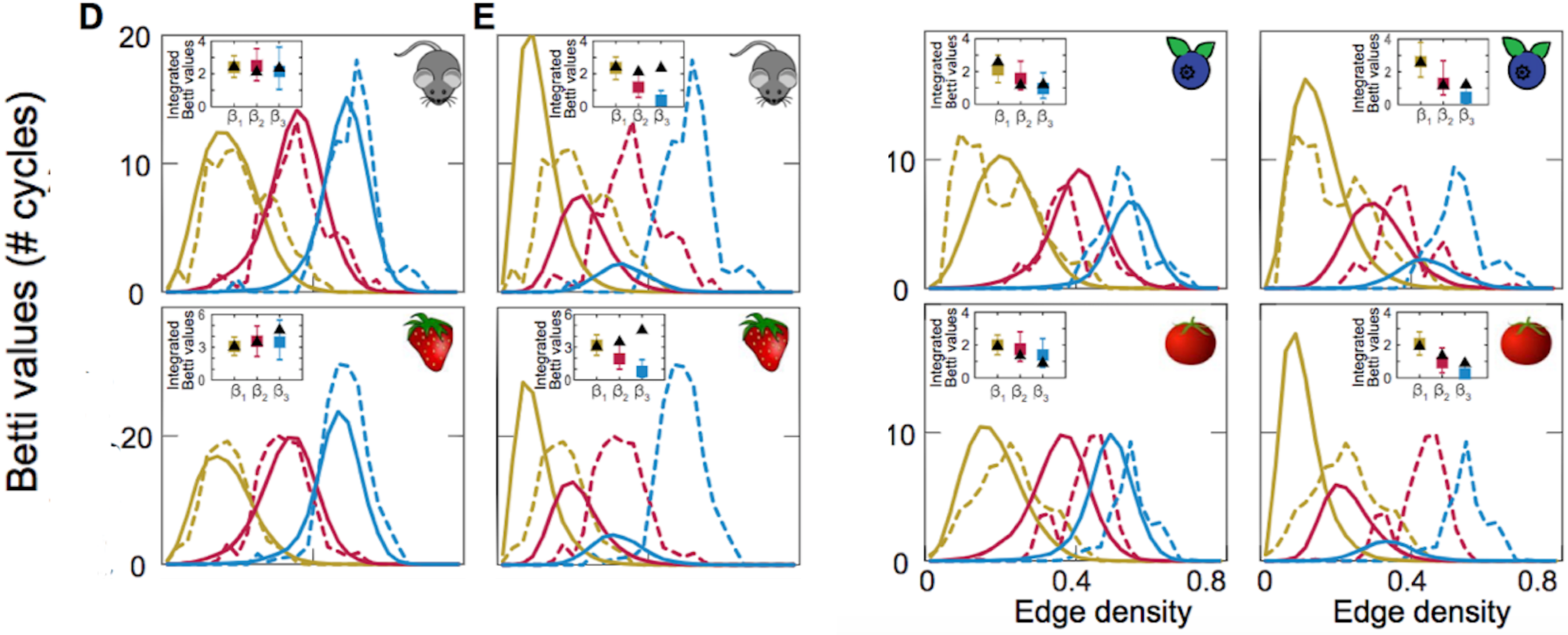


Fig.2 (D)&(E) Betti curves with the number of cycles different dimensions plotted as a function of edge densities.

- Biological background
- A statistical approach to generate Betti curves
- A topological method for finding geometric spaces consistent with correlation matrices
- **Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere**
- Visualization of the natural olfactory space using nonmetric MDS
- What are the axes of this olfactory space?
- The Betti curves of human olfactory perception

Nonmetric MDS

- The nonmetric MDS(metric multidimensional scaling) algorithm embeds a set of points within a prespecified space(hyperbolic) while attempting to preserve rank-ordered distances between points
- In the Euclidean space, use the Euclidean-based MDS algorithm in MATLAB
- In the 3D hyperbolic space, replace the Euclidean distance with hyperbolic distance in MATLAB to modify the metric MDS

$$\cosh x = \cosh(r)\cosh(r') - \sinh(r)\sinh(r')\cos\Delta\theta$$

- Biological background
- A statistical approach to generate Betti curves
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- A topological method for finding geometric spaces consistent with correlation matrices
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- **Visualization of the natural olfactory space using nonmetric MDS**
- What are the axes of this olfactory space?
- The Betti curves of human olfactory perception

Visualization by nonmetric MDS

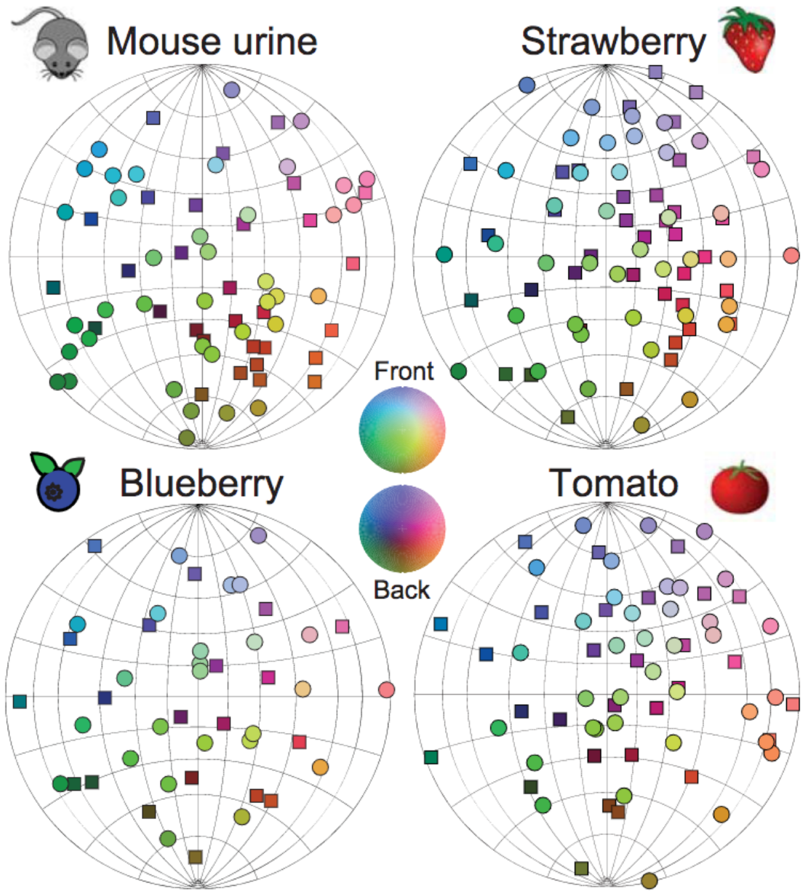


Fig.3 uniform sampling in all four data sets

- Use nonmetric MDS to visualize how the **points consistent with correlation statistics** distribute in the hyperbolic space
- The points are located near a surface of a sphere
- Range:$R_{max} = 7$, $R_{min} = 0.9\ R_{max}$
- Present the points on a sphere using the two angles of latitude and longitude. The RGB color scales show the location of each point
- Circles and squares show the front and back side, respectively
- All four data sets are approximately uniform sampling

- Biological background
- A statistical approach to generate Betti curves
- A topological method for finding geometric spaces consistent with correlation matrices
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- Visualization of the natural olfactory space using nonmetric MDS
- **What are the axes of this olfactory space?**
- The Betti curves of human olfactory perception

Previous findings and data sets choice

- There are axes correlated with perceptual odor pleasantness and physicochemical properties such as boiling point and acidity exist
- Points corresponding to pleasant and unpleasant odors occupied different parts of the space
- For mixture odors (individual samples) analysis, use only tomato and strawberry samples; while use all individual monomolecular odors for the individual odor components analysis

Mixtures (individual fruit samples) pleasantness analysis

- This analysis (mixture analysis) was based only on tomato and strawberry data sets (108 fruit samples: 54 tomatoes and 54 strawberries)
- For odor mixtures produced by individual fruit samples, use the "overall-liking" rating assigned by humans to fruit samples as a measure of pleasantness

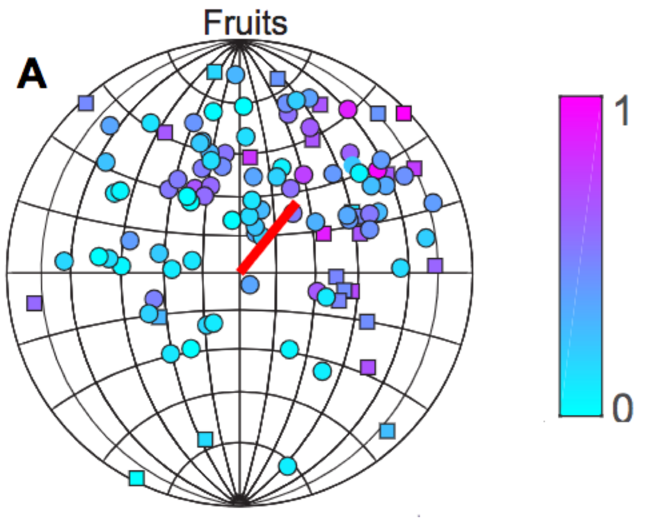


Fig. 4 (A) Axes associated with pleasantness for mixture samples

- For all 108 fruit samples, using normalized linear combinations of odor coordinates in the space. The red line shows the direction most associated with the pleasantness ratings.

Mixtures (individual fruit samples) pleasantness analysis

- To test if the identified pleasantness axis works well, if it can predict measured pleasantness rankings for novel samples, regenerate this axis using only strawberry samples and use it to predict pleasantness ranking for tomato samples. The correlation was significant

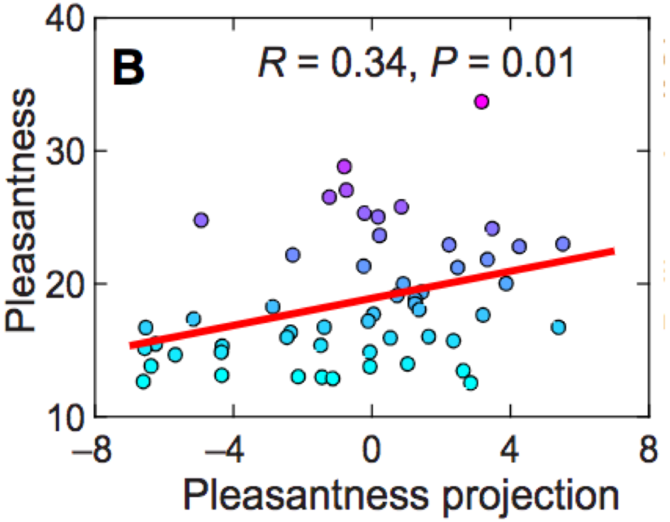
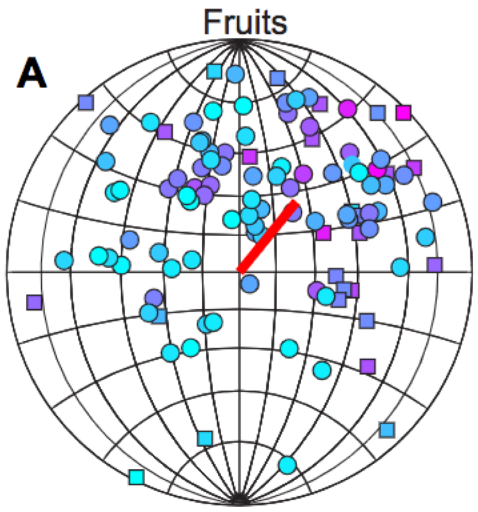
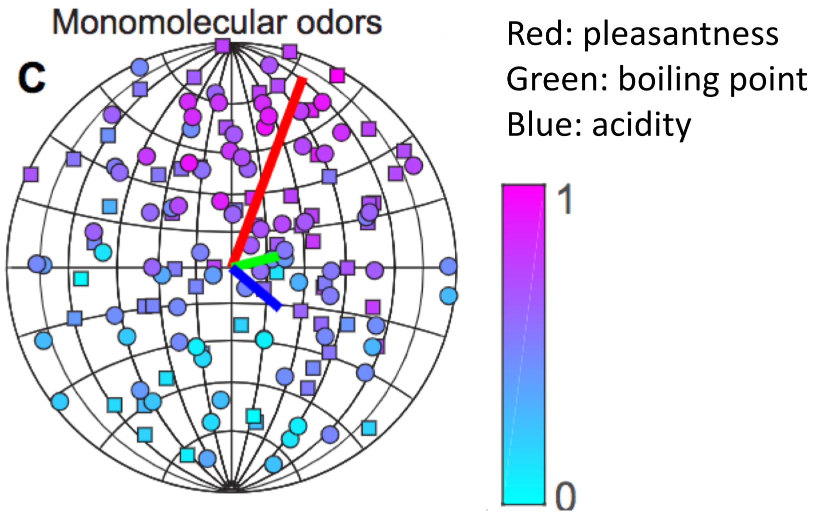


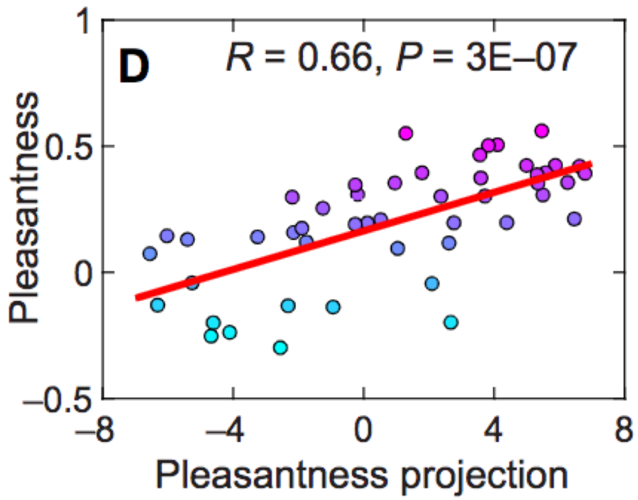
Fig. 4 (B) The correlation between the projection of tomato samples onto the pleasant axis of strawberry samples and the pleasantness rankings of tomato samples



Individual odor components pleasantness analysis

- The pleasantness values could also be assigned to individual odor components based on the correlation between the odor concentration in a mixture and mixture pleasantness computed over samples
- Stronger correlation between pleasantness and odor coordinates



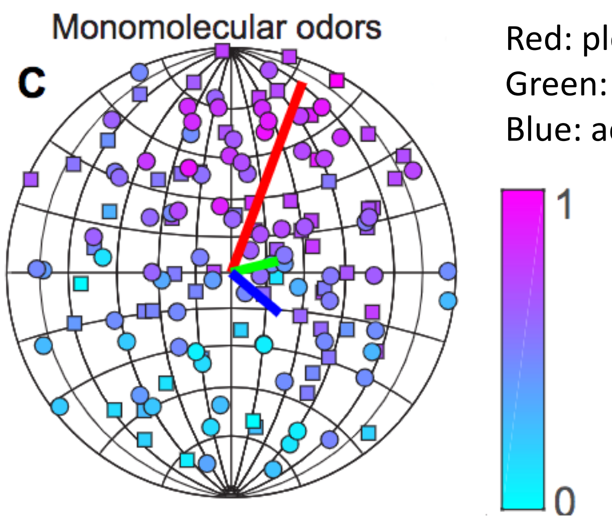


2/3 monomolecular odors were used to determine the direction of the pleasantness;
1/3 were projected onto the directions as validation sets to evaluate the correlations.

Fig. 4 (C) Visualization of 144 individual monomolecular odors. (D) Correlation between odor pleasantness with projection onto the pleasantness axes

Individual odor components boiling point and acidity analysis

- Boiling point: a measure of odor volatility
- Acidity: the correlation coefficient between its concentration and fruit sample acidity measurement



Red: pleasantness
Green: boiling point
Blue: acidity

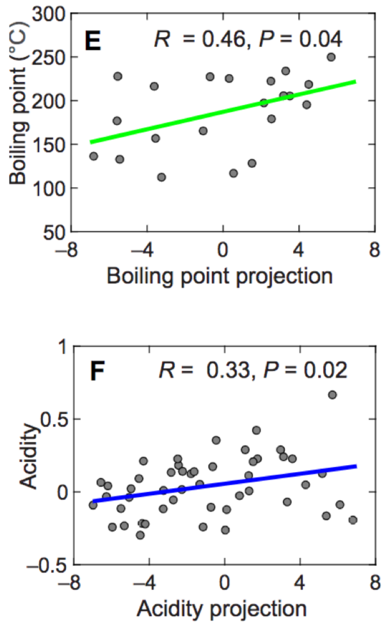


Fig. 4 **(C) Visualization of 144 individual monomolecular odors.**

(E) Correlation between molecular boiling point and the associated axes.
(F) Correlation between acidity value and the associated axes.

- Biological background
- A statistical approach to generate Betti curves
- A topological method for finding geometric spaces consistent with correlation matrices
- Nonmetric MDS embedding of odors on the surface of a 3D hyperbolic sphere
- Visualization of the natural olfactory space using nonmetric MDS
- What are the axes of this olfactory space?
- **The Betti curves of human olfactory perception**

Signature of the hyperbolic space

- Previous analysis of the Dravnieks database containing human perceptual descriptions of >120 monomolecular odors showed that the perceptual space is likely to be curved as a “potato-chip” surface
- Potato-chip surface have a negative curvature and serve as an example of hyperbolic surfaces

Property	Hyperbolic
Curvature K	<0
Parallel lines	∞
Triangles are	Thin
Shape of triangles	
Sum of angles in triangles	$<\pi$
Circle length	$2\pi \sinh \zeta r$
Disk area	$2\pi(\cosh \zeta r - 1)$

Betti curves

- To test whether the perceptual space is described by hyperbolic geometry, apply the Betti curve method to the Dravnieks database

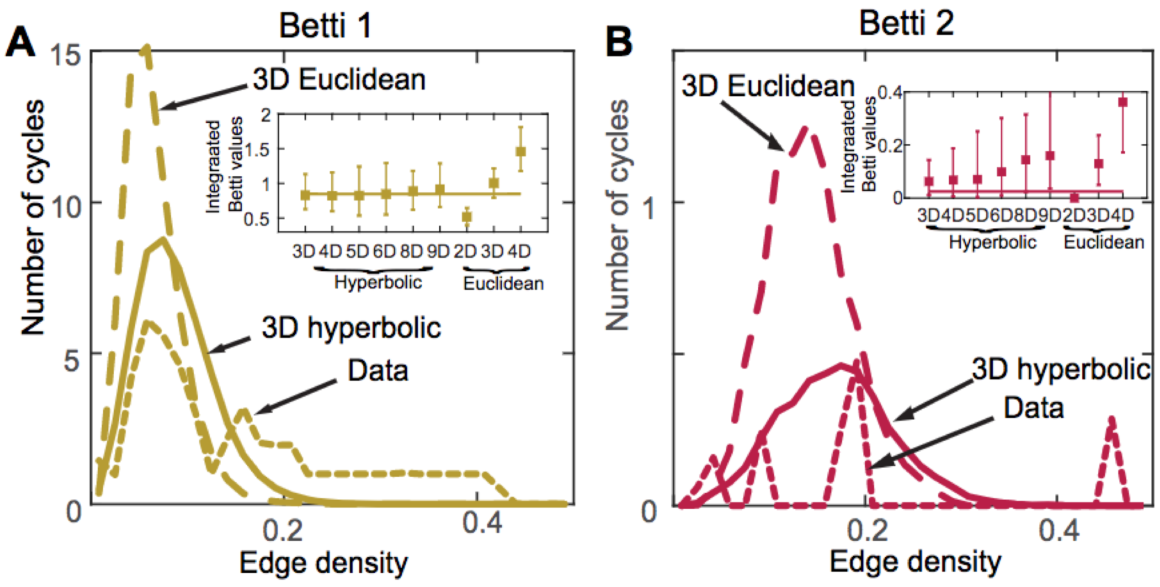
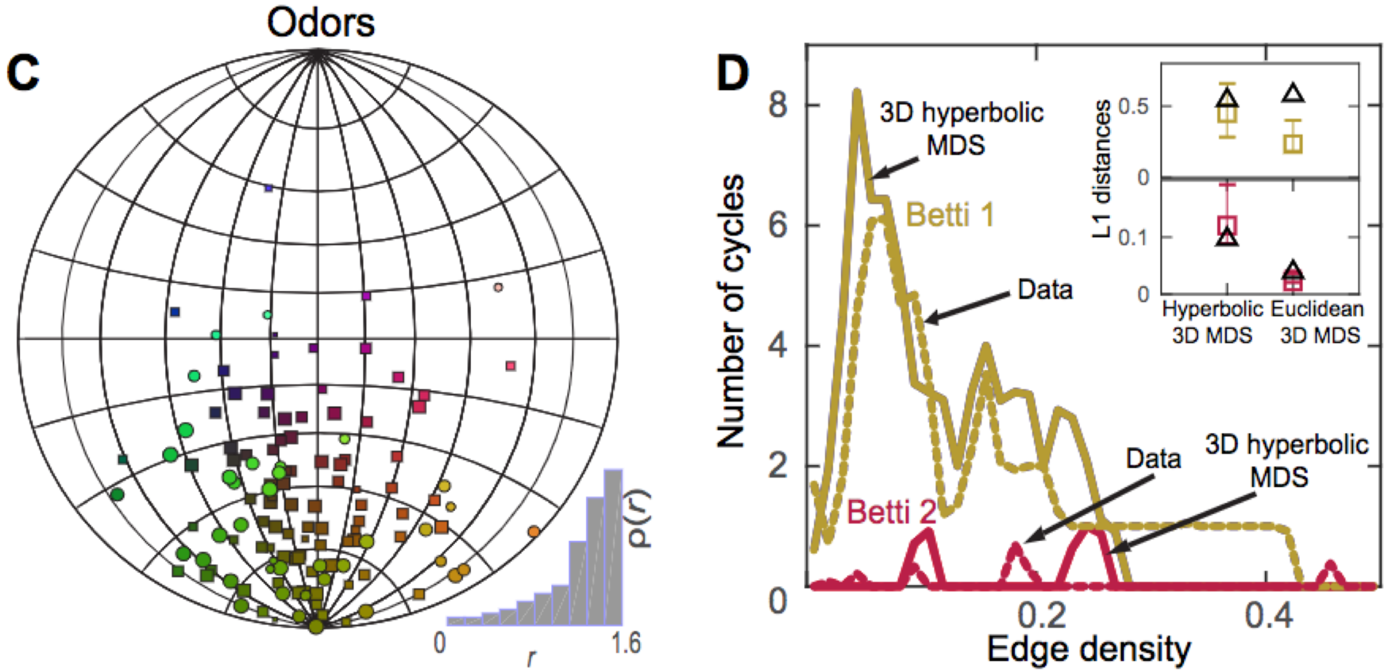


Fig. 5 **(A)&(B)Euclidean spaces were not consistent with measured Betti curves; hyperbolic spaces only fit the first Betti curve**

- Dotted line: perceptual data set
- Solid line: 3D hyperbolic space
- Dashed line: 3D Euclidean space
- Euclidean and hyperbolic spaces of other dimensions provided a worse fit
- The 1st Betti curve could not be matched to the Euclidean space, but the hyperbolic space could match. However, hyperbolic spaces could not match the 2nd Betti curve well
- The 1st and 2nd Betti curves were not as regular as the odorant case and contained multiple peaks

Nonuniform distribution by MDS

- Visualize the human olfactory perception space using nonmetric MDS
- Unlike the olfactory space, which are sampled uniformly, the distribution of points obtained using MDS is not uniform and clusters in one-half of the sphere
- The Betti curves generated by sampling points from this embedding match with the Betti curves derived from perceptual data



- The sizes of points are proportional to their radii, indicate location.

Fig. 5 (C) Visualization of human olfactory perception space using nonmetric MDS
(D) The first and second Betti curves fit with the curves derived from nonuniform distribution of points within the 3D hyperbolic space

Summary

- Find that both natural odors and human perception can be described using a 3D hyperbolic space.
- The axes of the olfactory space are pleasantness and 2 physicochemical properties including boiling point and acidity.
- Most natural odors in the statistical data sets fall near the sphere of a 3D hyperbolic space and are nearly uniformly distributed. Although the perceptual database also has a hyperbolic organization, the distribution concentrate on one side of the space. These analyses thus suggest perceptual coordinates that are yet to be explored

References

- Yuansheng Zhou, Brian H. Smith and Tatyana O. Sharpee, Hyperbolic geometry of the olfactory space. Science Advances. 2375-2548 (2018)
- D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá, Hyperbolic geometry of complex networks. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 82,036106 (2010)
- Giusti, E. Pastalkova, C. Curto, V. Itskov, Clique topology reveals intrinsic geometric structure in neural correlations. Proc. Natl. Acad. Sci. U.S.A.112, 13455–13460 (2015)
- .M. Khan, C. H. Luk, A. Flinker, A. Aggarwal, H. Lapid, R. Haddad, N. Sobel, Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world.J. Neurosci.27, 10015–10023 (2007).