Goodness-of-fit tests for metallic microstructures

Alessandra Cipriani

w/ Christian Hirsch (U Aarhus) & Martina Vittorietti (U Palermo)

UCL

November 9, 2022

Microstructures

Metallic microstructures

Study (micro)properties of metals including

- grain size/number of grains
- grain shape
- grain clustering

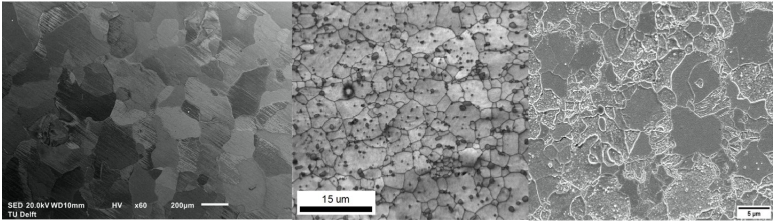


Figure: (a) Single-phase steel microstructure, (b) AISI stainless steel with M23C6 carbides precipitation (two phases), (c) ODS (Oxide Dispersion Strengthened) Eurofer steel (by W. Li, J. Hidalgo, V. Marques Serra Pereira).

Stereology

Stereology: sampling

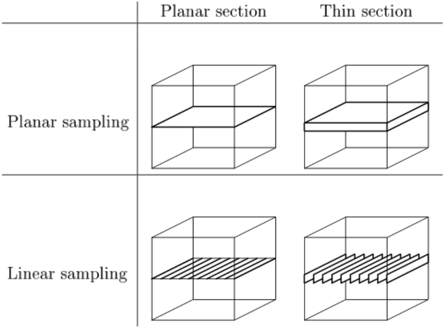
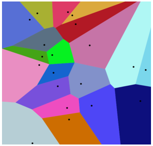
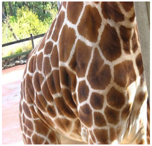


Figure: Survey of sectioning and sampling used for stereological estimation of particle size distribution. Using planar sampling design, a microstructure is observed in a planar window while linear sampling design uses test segments (commonly a system of parallel equidistant segments) (Ohser, Sandau 2000).

Poisson-Voronoi diagrams

The null model





Definition

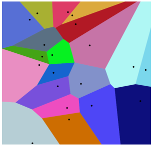
Given a set of distinct points $\Phi = \{x_i,\, i \in \mathbb{N}\}$ in $\mathbb{R}^d$ the Voronoi diagram of $\mathbb{R}^d$ with nuclei x_i is a partition of $\mathbb{R}^d$ consisting of cells

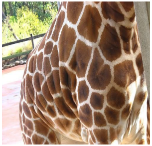
$$C_i = \{y \in \mathbb{R}^d : \|y - x_i\| \leq \|y - x_j\| \text{ for all } j \neq i\}, \quad i \in \mathbb{N}$$

where $\|\cdot\|$ is the Euclidean distance.

Poisson-Voronoi diagrams

The null model





Definition

Given a set of distinct points $\Phi = \{x_i,\ i \in \mathbb{N}\}$ in $\mathbb{R}^d$ the Voronoi diagram of $\mathbb{R}^d$ with nuclei x_i is a partition of $\mathbb{R}^d$ consisting of cells

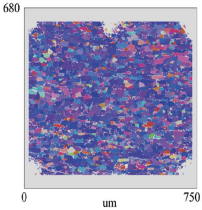
$$C_i = \{y \in \mathbb{R}^d : \|y - x_i\| \leq \|y - x_j\| \text{ for all } j \neq i\}, \quad i \in \mathbb{N}$$

where $\|\cdot\|$ is the Euclidean distance.

If X is the realization of a homogeneous Poisson point process the resulting structure is the Poisson-Voronoi diagram V_Φ.

Question

Given multiple 2D material sections equally spaced, could a Poisson-Voronoi diagram be a model for approximating the 3D material microstructure?



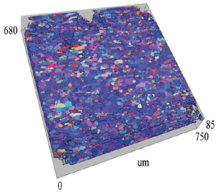


Figure: EBSD scans of extra low carbon strip steel (by J. Gálan López)

Question

2D Sectional Poisson-Voronoi Diagrams

Question: "For integers $2 \leq t \leq d-1$, is the intersection between an arbitrary but fixed t-dimensional linear affine subspace of $\mathbb{R}^d$ and the d-dimensional Voronoi tessellation generated by a point process X a t-dimensional Voronoi tessellation?"

Question

2D Sectional Poisson-Voronoi Diagrams

Question: "For integers $2 \leq t \leq d - 1$, is the intersection between an arbitrary but fixed t-dimensional linear affine subspace of $\mathbb{R}^d$ and the d-dimensional Voronoi tessellation generated by a point process X a t-dimensional Voronoi tessellation?"
Answer: NO (Chiu et al.,1996)

Topological data analysis

… and persistence diagrams

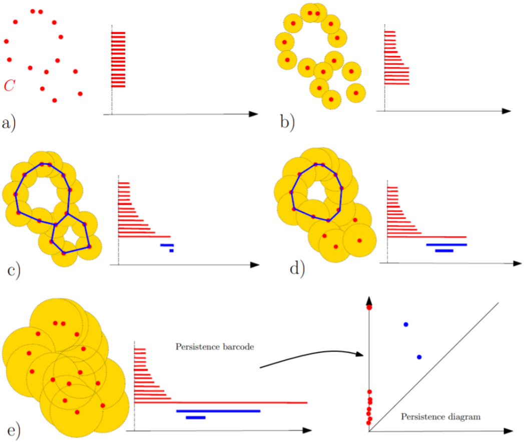


Figure: From *An introduction to TDA*, F. Chazal, B. Michel

Our contribution

Our vines

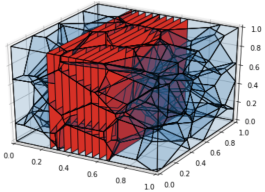


Figure: 3D Voronoi cells

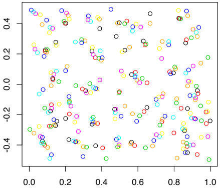


Figure: Nuclei over sections

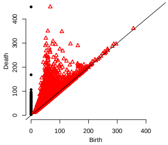


Figure: Persistence diagram, one section

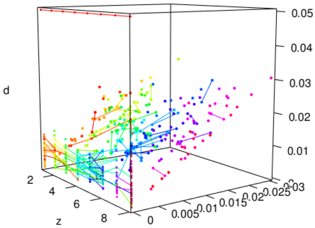


Figure: Stacked diagrams

Our contribution

Stacking persistence diagrams

At each section h we compute the persistence diagram given by the collection

$$\{(B_i^q(h), D_i^q(h))_i\}$$

of birth and death times of q features. We propose to "stack" persistence diagrams over sections

Topological data analysis

Persistence vineyards

Extension of persistence diagrams (Cohen-Steiner et al., 2006): time slices

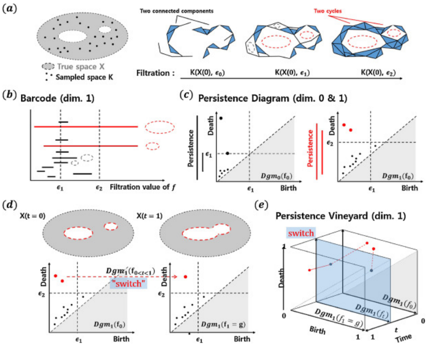


Figure: Persistence vineyards (Yoo et al., 2016)

Our contribution

Trajectories

$Q_n := [-n/2, n/2]^2$, metal$= Q_n \times [0, 1]$

Our contribution

Trajectories

$Q_n := [-n/2,\ n/2]^2$, metal$= Q_n \times [0,\ 1]$

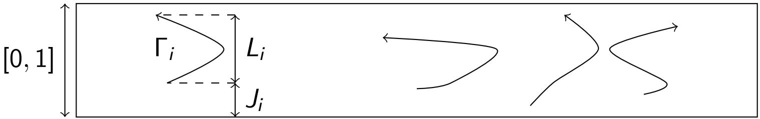


Figure: Point process of trajectories within the window $Q_n \times [0, 1]$

- Let $\{(J_{i,n}, L_{i,n}, G_{i,n})\}_{i\geq 1} \subseteq [0, 1] \times [0, 1] \times C([0, 1], Q_n)$ be a point process

Our contribution

Trajectories

$Q_n := [-n/2,\ n/2]^2$, metal$=$ $Q_n \times [0,\ 1]$

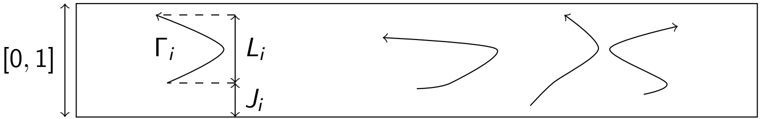


Figure: Point process of trajectories within the window $Q_n \times [0, 1]$

- Let $\{(J_{i,n}, L_{i,n}, G_{i,n})\}_{i\geq 1} \subseteq [0, 1] \times [0, 1] \times C([0, 1], Q_n)$ be a point process
- Set

$$X_{i,n}(h) := \left\{ \left(\Gamma_{i,n}((h - J_{i,n})/L_{i,n}), h\right) :\ J_{i,n} \leq h \leq J_{i,n} + L_{i,n} \right\}.$$

Our contribution

Trajectories

$Q_n := [-n/2, \ n/2]^2$, metal= $Q_n \times [0, \ 1]$

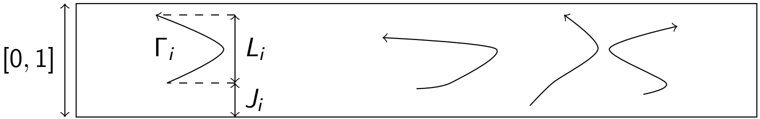


Figure: Point process of trajectories within the window $Q_n \times [0, 1]$

- Let $\{(J_{i,n}, L_{i,n}, G_{i,n})\}_{i \geq 1} \subseteq [0, 1] \times [0, 1] \times C([0, 1], Q_n)$ be a point process
- Set

$$X_{i,n}(h) := \left\{ \left(\Gamma_{i,n}((h - J_{i,n})/L_{i,n}), h \right) : \ J_{i,n} \leq h \leq J_{i,n} + L_{i,n} \right\}.$$

- $\mathrm{X}_n = \{X_{i,n}(\cdot)\}_{i \geq 1}$ is a process of trajectories with *offset* $J_{i,n} \in [0, 1]$, *length* $L_{i,n} \in [0, 1]$ and *shape* $\Gamma_{i,n} \in C([0, 1], Q_n)$

Example

Let $\mathcal{P} = (P_i)_i$ be a homogeneous Poisson process on $Q_n \times [0, 1]$ with associated P-V tessellation.
Let $G_{i,\,n}(h)$ be the centroid of the "sectional" cell i at level h. We consider

$$X_{i,\,n}(\cdot) = G_{i,n}(\cdot)$$

Our contribution

Test statistics

- Longitudinal test statistics (with feature tracking)

$$T_n = T(\mathrm{X}_n) := \sum_{i \text{ feature}} \xi\big(\{(B_i(h), D_i(h))\}_h\big).$$

Eg: ξ=average number of slices where feature i is present.

Our contribution

Test statistics

- Longitudinal test statistics (with feature tracking)

$$T_n = T(X_n) := \sum_{i \text{ feature}} \xi\big(\{(B_i(h), D_i(h))\}_h\big).$$

Eg: ξ=average number of slices where feature i is present.

- Cross-sectional statistics (without feature tracking)

$$T_n = \frac{1}{H}\sum_{h=1}^{H}\sum_{i} \xi'(B_i(h), D_i(h)).$$

Eg: averaged persistent Betti numbers, and

$$T^q_{TP} = \frac{1}{H}\frac{1}{|W|}\sum_{h=1}^{H}\sum_{i \text{ feature in } h}(D^q_i(h) - B^q_i(h)).$$

where H is the total number of sections, $|W|$ the size of the domain

Our contribution: asymptotic regime

Asymptotic normality: scalar level

$T_n = T(\mathrm{X}_n)$, X_n trajectories.

Our contribution: asymptotic regime

Asymptotic normality: scalar level

$T_n = T(X_n)$, X_n trajectories.

Assumption: exponential stabilization

X_n is exponentially stabilizing. Roughly, changing the point process far away from $x \in Q_n$ does not change the trajectories in the vicinity of x.

Our contribution: asymptotic regime

Asymptotic normality: scalar level

$T_n = T(X_n)$, X_n trajectories.

Assumption: exponential stabilization

X_n is exponentially stabilizing. Roughly, changing the point process far away from $x \in Q_n$ does not change the trajectories in the vicinity of x.

Theorem (C, Hirsch, Vittorietti, 2022)

Assume further that X_n emerges from a Poisson point cloud, and that ξ is bounded. Then, the statistic

$$\frac{T_n - E[T_n]}{n}$$

converges in distribution to a normal random variable.

Proof ideas

- Martingale CLT applied to T_n

Proof ideas

- Martingale CLT applied to T_n
- Important edge correction

Proof ideas

- Martingale CLT applied to T_n
- Important edge correction
- Exponential decay of stabilization radius

Our contribution: asymptotic regime

Asymptotic normality: functional level

Theorem (C, Hirsch, Vittorietti, 2022)

Assume further that the factorial moments of $\{X_{i,n}(h)\}_i$ *are uniformly bounded. Then, as a function on* $[0,1]^2$*, in the Skorohod topology the recentered and rescaled (M-bounded) persistent Betti numbers*

$$\frac{\beta_n - E[\beta_n]}{n}$$

converge in distribution to a centered Gaussian process.

Proof ideas

- Martingale CLT decomposition

Proof ideas

- Martingale CLT decomposition
- Restriction to sub-boxes in a grid

Proof ideas

- Martingale CLT decomposition
- Restriction to sub-boxes in a grid
- Exponential decay of correlations (via cumulant bounds)

Our contribution: finite-window regime

Simulation results

Simulation on $170 \times 170 \times 85$, 9 slices

- Null model: Poisson-Voronoi diagram with sites generated according to a Poisson process with intensity $\lambda = 2.18 * 10^{-4}$ ($\Rightarrow$ real data)

Our contribution: finite-window regime

Simulation results

Simulation on $170 \times 170 \times 85$, 9 slices

- Null model: Poisson-Voronoi diagram with sites generated according to a Poisson process with intensity $\lambda = 2.18 * 10^{-4}$ ($\Rightarrow$ real data)
- Matérn hard-core Voronoi diagrams

Our contribution: finite-window regime

Simulation results

Simulation on $170 \times 170 \times 85$, 9 slices

- Null model: Poisson-Voronoi diagram with sites generated according to a Poisson process with intensity $\lambda = 2.18 * 10^{-4}$ ($\Rightarrow$ real data)
- Matérn hard-core Voronoi diagrams
- Matérn cluster Voronoi diagrams with cluster number n_{cl}, intensity d_{cl} and cluster radius R

Our contribution: finite-window regime

Simulation results

Simulation on $170 \times 170 \times 85$, 9 slices

- Null model: Poisson-Voronoi diagram with sites generated according to a Poisson process with intensity $\lambda = 2.18 * 10^{-4}$ ($\Rightarrow$ real data)
- Matérn hard-core Voronoi diagrams
- Matérn cluster Voronoi diagrams with cluster number n_{cl}, intensity d_{cl} and cluster radius R

Table: Parameters

	HC_1	HC_2	CL_1	CL_2	CL_3
R	5.25	5.95	42.5	42.5	42.5
n_{cl}	/	/	10	5	4
λ_{cl}	/	/	10	20	25

Simulation results

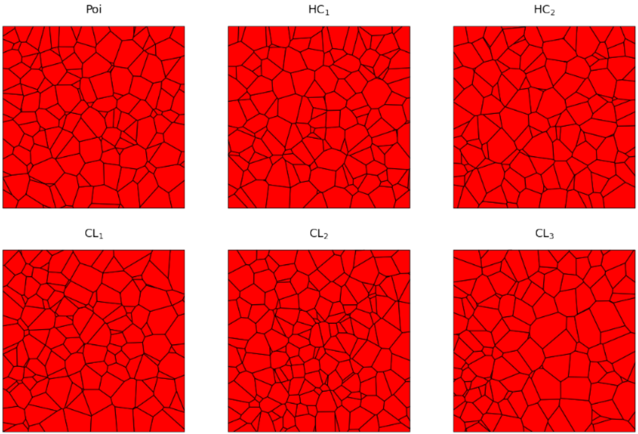


Figure: 2D slices of different Voronoi diagrams

Preliminary results

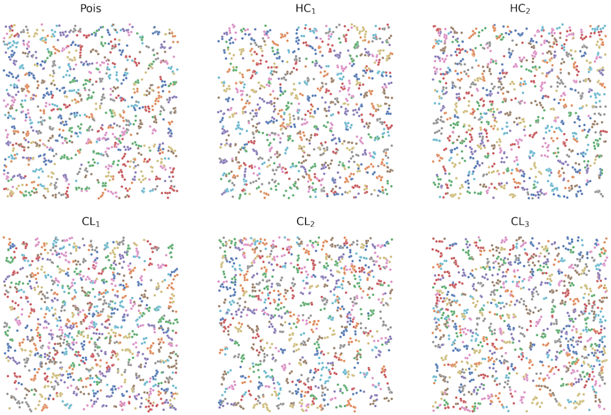


Figure: $x - y$ coordinates of the centers of gravity of the 2D sectional grains of 3D Voronoi diagrams. Different colors represent different sections

Simulation results

Our statistics

- Cross-sectional total persistence:

$$T^q_{TP} = \frac{1}{H} \sum_h \frac{1}{|W|} \sum_i (D_i(h) - B_i(h))$$

Simulation results

Our statistics

- Cross-sectional total persistence:

$$T^q_{TP} = \frac{1}{H}\sum_h \frac{1}{|W|}\sum_i (D_i(h) - B_i(h))$$

- Vine-based persistence:

$$T^q_M := \frac{1}{|W|}\sum_{i\le N^q} \frac{1}{n^q_{h(i)}}\sum_h (D_i(h) - B_i(h))$$

where N^q is the total number of features of dimension q observed in all the slices and $n^q_{h(i)}$ is the number of slices in which the ith q-feature is visible.

Simulation results

Our statistics

- Cross-sectional total persistence:

$$T^q_{TP} = \frac{1}{H}\sum_h \frac{1}{|W|}\sum_i (D_i(h) - B_i(h))$$

- Vine-based persistence:

$$T^q_M := \frac{1}{|W|}\sum_{i\leq N^q} \frac{1}{n^q_{h(i)}}\sum_h (D_i(h) - B_i(h))$$

  where N^q is the total number of features of dimension q observed in all the slices and $n^q_{h(i)}$ is the number of slices in which the ith q-feature is visible.

- (Pooled) Ripley K-function:

$$T_{Rip} := \int_0^{r_{Rip}} \widehat{K}_{pool}(r)dr$$

Exploratory analysis

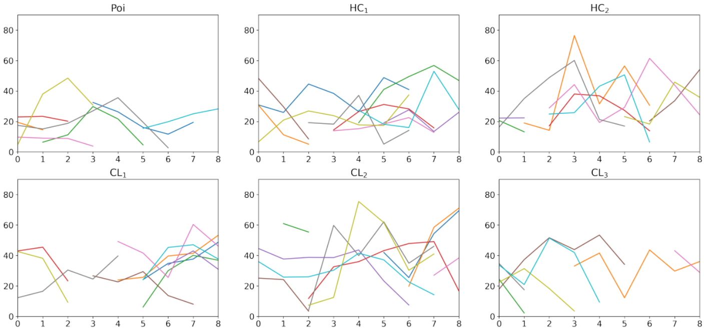


Figure 6: Samples of vines in dimension 0. Lines indicating the same persistence point observed in more than one slice. The x-axis denotes the slice in which the diagram is computed. The y-axis denotes the death time of the considered feature.

Preliminary results

Power of the tests

$\alpha = 0.05$, 5000 realisations, 9 sections, H_0 : $Poi(\lambda)$

T	PV	HC_1	HC_2	CL_1	CL_2	CL3
T^0_{TP}	5.08%	11.96%	15.94%	10.70%	16.58%	24.50 %
T^1_{TP}	5.30%	5.10%	5.64%	18.24%	28.77%	38.01 %
T^0_{M}	5.00%	8.30%	9.46%	18.84%	27.47%	34.85 %
T^1_{M}	4.74%	5.10%	5.02%	19.92%	27.83%	37.15 %
T_{Rip}	4.94%	7.82%	9.76%	52.01%	70.17%	79.50%

Figure: Rejection rates for the test statistics under the Poisson–Voronoi diagram null model and the alternatives

Preliminary results

Power of the tests

$\alpha = 0.05$, 5000 realisations, 9 sections, H_0 : $Poi(\lambda)$

T	PV	HC_1	HC_2	CL_1	CL_2	CL3
T^0_{TP}	5.08%	11.96%	15.94%	10.70%	16.58%	24.50 %
T^1_{TP}	5.30%	5.10%	5.64%	18.24%	28.77%	38.01 %
T^0_{M}	5.00%	8.30%	9.46%	18.84%	27.47%	34.85 %
T^1_{M}	4.74%	5.10%	5.02%	19.92%	27.83%	37.15 %
T_{Rip}	4.94%	7.82%	9.76%	52.01%	70.17%	79.50%

Figure: Rejection rates for the test statistics under the Poisson–Voronoi diagram null model and the alternatives

Reconstruction algorithm

Real data

z-scores

	T^0_{TP}	T^1_{TP}	T^0_{M}	T^1_{M}	T_{Rip}
z-score	23.00	5.40	94.75	30.05	81.8

Figure: z-scores associated with the different test statistics when the slices are tested against a Poisson-Voronoi null model

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics
  - More statistics to be evaluated

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics
  - More statistics to be evaluated
- Test on (more) real data

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics
  - More statistics to be evaluated
- Test on (more) real data
  - matching algorithm for the centers of gravity among the different sections

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics
  - More statistics to be evaluated
- Test on (more) real data
  - matching algorithm for the centers of gravity among the different sections
- Use of other observables to construct cells (vertices)

Conclusions

- Asymptotic normality of longitudinal and cross-sectional statistics
  - More statistics to be evaluated
- Test on (more) real data
  - matching algorithm for the centers of gravity among the different sections
- Use of other observables to construct cells (vertices)
- Other nulls (power-law correlations)...

Thank you for your attention!