Stochastic Convergence of Persistence Landscapes and Silhouettes

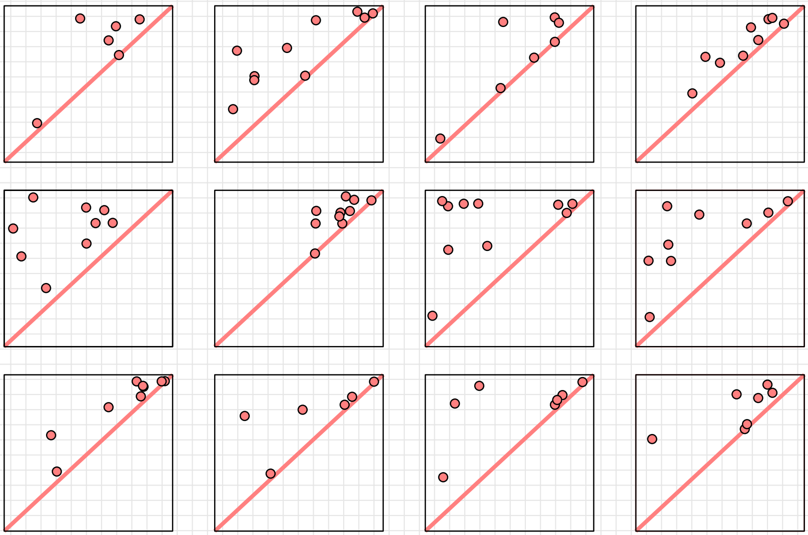
Brittany Terese Fasy
brittany.fasy@alumni.duke.edu

joint work with F. Chazal, F. Lecci,
A. Rinaldo, L. Wasserman

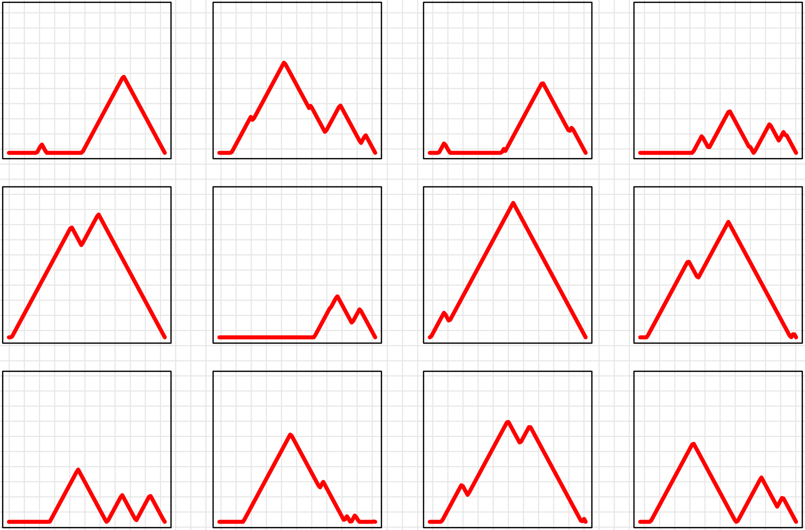
Postdoctoral Researcher, Tulane University

11 June 2014

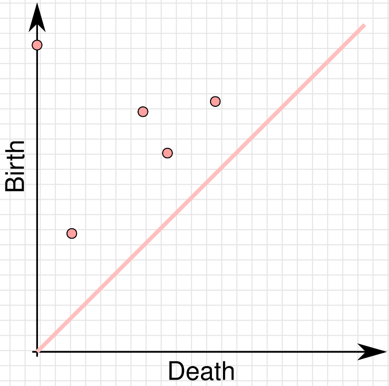
Sample (Bounded) Persistence Diagrams



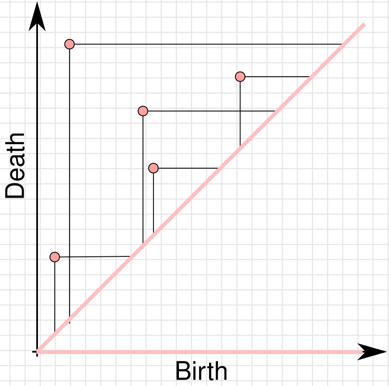
Sample (Bounded) Persistence Landscapes



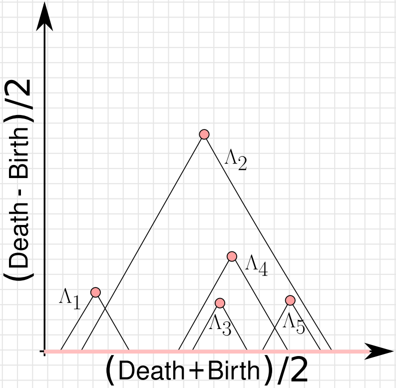
Persistence Landscapes



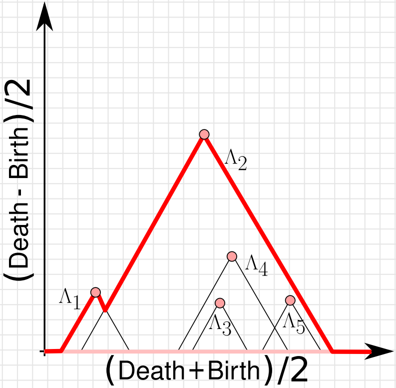
Persistence Landscapes



Persistence Landscapes



Persistence Landscapes



[B-2012] Statistical Topology Using Persistent Homology. ArXiv 1207.6437

Motivation

Where Do the Diagrams Come From?

Senario 1

Draw n functions from a probabilty distribution over the set of (Morse) functions. This induces a sample of persistence landscapes. Ex: each function is the distance to a compact set embedded in $\mathbb{R}^d$.

Motivation

Where Do the Diagrams Come From?

Senario 1

Draw n functions from a probabilty distribution over the set of (Morse) functions. This induces a sample of persistence landscapes. Ex: each function is the distance to a compact set embedded in $\mathbb{R}^d$.

Senario 2

Given a large dataset with N points, it is very expensive to compute the persistence landscape λ exactly. Instead, we use subsampling to compute approximations $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then, we can upper bound $\mathbb{E}[\lambda_i - \lambda]$ with

$$\mathbb{E}[\lambda_i - \mu] + \mathbb{E}[\mu - \lambda]$$

Motivation

Where Do the Diagrams Come From?

Senario 1

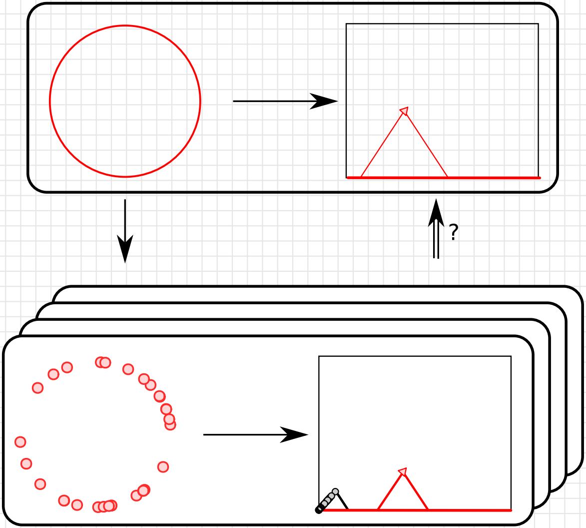
Draw n functions from a probabilty distribution over the set of (Morse) functions. This induces a sample of persistence landscapes. Ex: each function is the distance to a compact set embedded in $\mathbb{R}^d$.

Senario 2

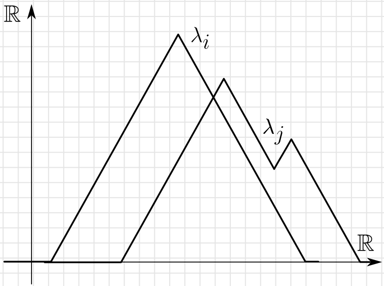
Given a large dataset with N points, it is very expensive to compute the persistence landscape λ exactly. Instead, we use subsampling to compute approximations $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then, we can upper bound $\mathbb{E}[\lambda_i - \lambda]$ with

$$\mathbb{E}[\lambda_i - \mu]$$

Topological Inference

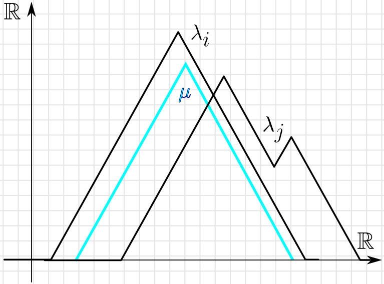


Pointwise Convergence of Landscapes



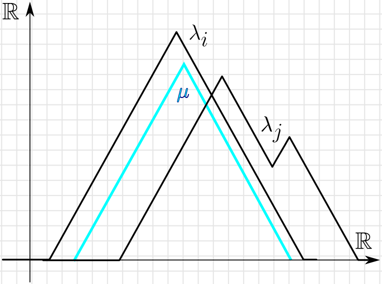
Let $\lambda_1, \ldots, \lambda_n \overset{\text{iid}}{\sim} \mathcal{L}_T$.

Pointwise Convergence of Landscapes



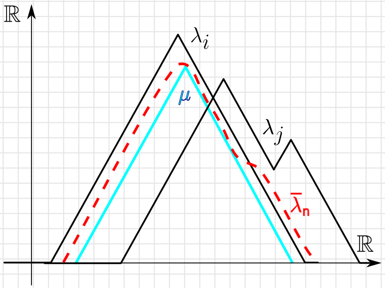
Let $\lambda_1, \ldots, \lambda_n \overset{\text{iid}}{\sim} \mathcal{L}_T$.
$\mu = \mathbb{E}(\lambda_i)$

Pointwise Convergence of Landscapes



Let $\lambda_1, \ldots, \lambda_n \overset{\text{iid}}{\sim} \mathcal{L}_T$.
$\mu = \mathbb{E}(\lambda_i)$
$\bar{\lambda}_n$: empirical average landscape

Pointwise Convergence of Landscapes

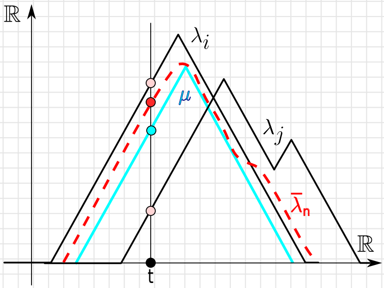


Let $\lambda_1, \ldots, \lambda_n \overset{\text{iid}}{\sim} \mathcal{L}_T$.
$\mu = \mathbb{E}(\lambda_i)$
$\bar{\lambda}_n$: empirical average landscape

Pointwise Convergence [B-2012].

$\bar{\lambda}_n$ converges pointwise to μ.

Pointwise Convergence of Landscapes



Let $\lambda_1, \ldots, \lambda_n \overset{\text{iid}}{\sim} \mathcal{L}_T$.
$\mu = \mathbb{E}(\lambda_i)$
$\bar{\lambda}_n$: empirical average landscape

Pointwise Convergence [B-2012].

$\bar{\lambda}_n$ converges pointwise to μ.

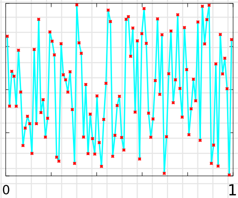
Key Properties

Landscapes are $(T/2)$-bounded and one-Lipschitz!

Gaussian Process

Gaussian Process

A GP over I is a set of independent random variables associated to each $t \in I$ such that every finite collection of random variables has a multi-variate normal distribution.



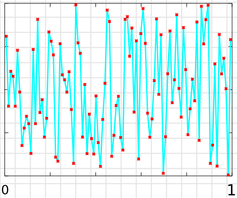
Gaussian Process

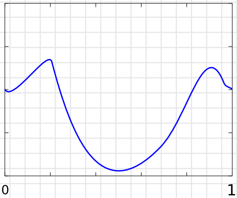
Gaussian Process

A GP over I is a set of independent random variables associated to each $t \in I$ such that every finite collection of random variables has a multi-variate normal distribution.

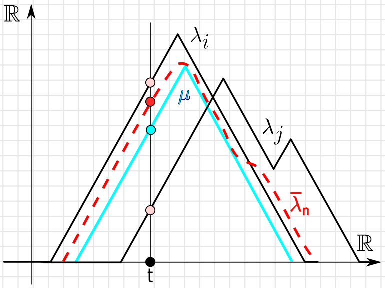
Brownian Bridge

A Brownian Bridge B defined over I is a continuous GP over I with a nice covariance structure such that $B(0) = B(1) = \mathbb{E}[B(i)] = 0$.



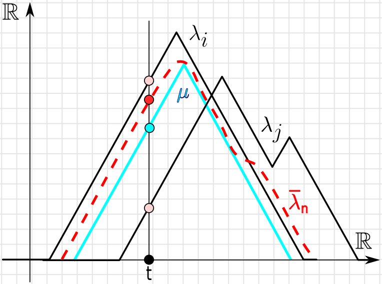


Emperical Process



$$f_t \quad : \quad \mathcal{L}_T \to \mathbb{R}$$
$$\lambda \mapsto \lambda(t)$$

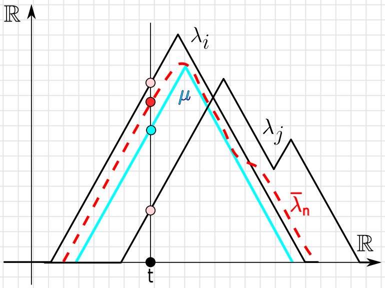
Emperical Process



$$f_t \quad : \quad \mathcal{L}_T \to \mathbb{R}$$
$$\lambda \mapsto \lambda(t)$$

$$\bar{\lambda}_n(t) - \mu(t)$$

Emperical Process



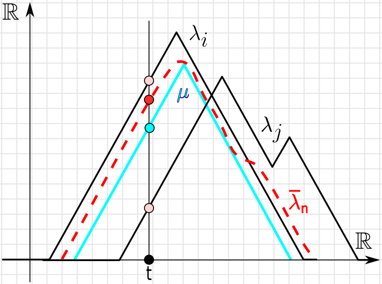
$$f_t \quad : \quad \mathcal{L}_T \to \mathbb{R}$$
$$\lambda \mapsto \lambda(t)$$

$$\bar{\lambda}_n(t) - \mu(t)$$

Emperical Process on $[0, T]$

$$f_t(\bar{\lambda}_n) - f_t(\mu)$$

Emperical Process



$$f_t \;:\; \mathcal{L}_T \to \mathbb{R}$$
$$\lambda \mapsto \lambda(t)$$

$$\bar{\lambda}_n(t) - \mu(t)$$

Emperical Process on $[0, T]$

For $t \in [0, T]$, we define $\mathbb{G}_n(f_t) = \mathbb{G}_n(t) := \frac{1}{\sqrt{n}} \left(f_t(\bar{\lambda}_n) - f_t(\mu) \right)$.

Weak Convergence

Weak Convergence

$\mathbb{G}_n(t) = \frac{1}{\sqrt{n}}(\bar{\lambda}_n(t) - \mu(t))$ converges weakly to the Brownian bridge $\mathbb{G}$ with covariance function

$$\kappa(f, g) = \int f(u)g(u)dP(u) - (\int f(u)dP(u))(\int g(u)dP(u)).$$

Uniform Convergence

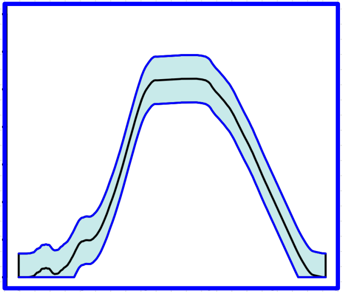
Let $\sigma(t) = \sqrt{n \, \mathrm{Var}\bar{\lambda}_n(t)}$.
Assume $\sigma(t) > 0$ on $[t_*, t^*] \subset [0, T]$.

Uniform CLT

There exists a random variable $W \overset{d}{=} \sup_{t \in [t_*, t^*]} |\mathbb{G}(f_t)|$ such that

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}\Big(\sup_{t \in [t_*, t^*]} |\mathbb{G}_n(t)| \leq z \Big) - \mathbb{P}\left(W \leq z \right) \right| = O\left(\frac{(\log n)^{\frac{7}{8}}}{n^{\frac{1}{8}}} \right).$$

Confidence Bands



Confidence Band

A $(1-\alpha)$-confidence band for μ is a pair of functions $\ell_n, u_n \colon [0,T] \to \mathbb{R}$ such that

$$\mathbb{P}(\ell_n(t) \leq \mu(t) \leq u_n(t) \text{ for all } t) \geq 1-\alpha.$$

The Multiplier Bootstrap

Let $\xi_1, \ldots, \xi_n \sim N(0,1)$. Then,

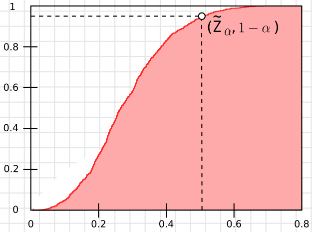
$$\widetilde{\mathbb{G}}_n(f_t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i(\lambda_i(t) - \bar{\lambda}_n(t))$$

is the multiplier bootstrap version of $\mathbb{G}_n(f_t)$.

α-Quantile

$\widetilde{Z}_\alpha$ is the unique value such that

$$\mathbb{P}\left(\sup_t |\widetilde{\mathbb{G}}_n(f_t)| > \widetilde{Z}_\alpha \;\middle|\; \{\lambda_i\} \right) = \alpha$$



The Multiplier Bootstrap

Let $\xi_1, \ldots, \xi_n \sim N(0,1)$. Then,

$$\widetilde{\mathbb{G}}_n(f_t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i(\lambda_i(t) - \bar{\lambda}_n(t))$$

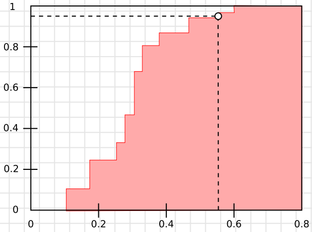
is the multiplier bootstrap version of $\mathbb{G}_n(f_t)$.

α-Quantile

$\widetilde{Z}_\alpha$ is the unique value such that

$$\mathbb{P}\left(\sup_t |\widetilde{\mathbb{G}}_n(f_t)| > \widetilde{Z}_\alpha \,\middle|\, \{\lambda_i\} \right) = \alpha$$

*Approx. $\widetilde{Z}_\alpha$ by MC simulation



Confidence Bands for Landscapes

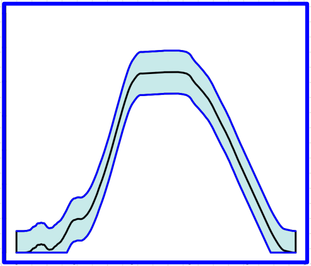
The Multiplier Bootstrap

Recalling

$\mathbb{G}_n(t) = \frac{1}{\sqrt{n}}(\bar{\lambda}_n(t) - \mu(t))$, let

$$\ell_n = \bar{\lambda}_n(t) - \frac{\tilde{Z}(\alpha)}{\sqrt{n}}$$

$$u_n = \bar{\lambda}_n(t) + \frac{\tilde{Z}(\alpha)}{\sqrt{n}}$$



Uniform Band

$$\mathbb{P}\left(\ell_n(t) \leq \mu(t) \leq u_n(t) \text{ for all } t\right) \geq 1 - \alpha - O\left(\frac{(\log n)^{\frac{7}{8}}}{n^{\frac{1}{8}}}\right).$$

Variable Width Confidence Bands

$$\mathbb{H}_n(f_t) := \mathbb{G}_n(t)/\sigma(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\lambda_i(t) - \mu(t)}{\sigma(t)}$$

$$\widetilde{\mathbb{H}}_n(f_t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \frac{\lambda_i(t) - \bar{\lambda}_n(t)}{\hat{\sigma}_n(t)}$$

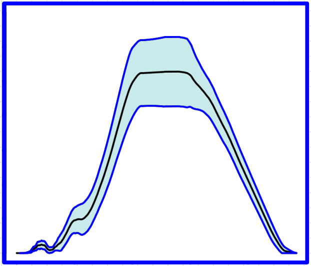
$\widetilde{Q}_\alpha$ is the unique value such that

$$\mathbb{P}\left(\sup_t |\widetilde{\mathbb{H}}_n(f_t)| > \widetilde{Q}_\alpha \;\middle|\; \{\lambda_i\} \right) = \alpha$$

Variable Width Confidence Bands

$$\ell_n = \bar{\lambda}_n(t) - \frac{\widetilde{Q}(\alpha)\hat{\sigma}_n(t)}{\sqrt{n}}$$

$$u_n = \bar{\lambda}_n(t) + \frac{\widetilde{Q}(\alpha)\hat{\sigma}_n(t)}{\sqrt{n}}$$



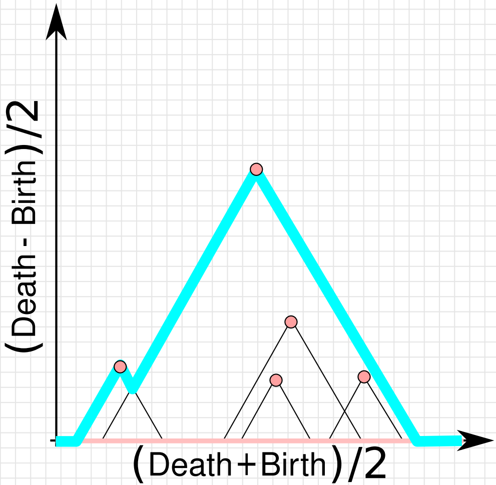
Variable Band

$$\mathbb{P}\left(\ell_n(t) \leq \mu(t) \leq u_n(t) \text{ for all } t\right) \geq 1 - \alpha - O\left(\frac{(\log n)^{\frac{7}{8}}}{n^{\frac{1}{8}}}\right).$$

Relax, that was the most technical part.

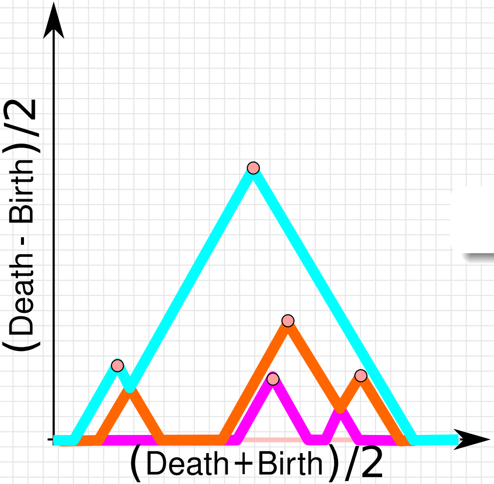
Persistence Silhouettes

Definitions



Persistence Silhouettes

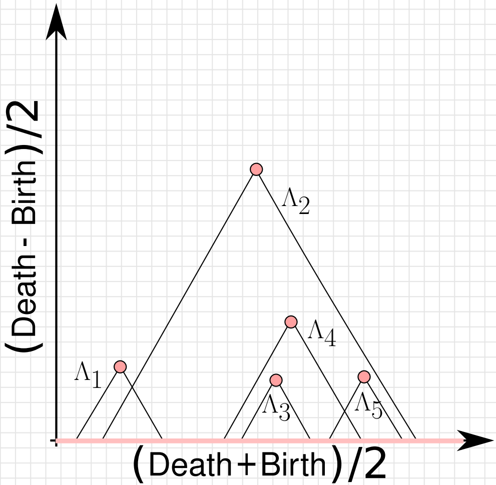
Definitions



$$\Lambda \colon \mathbb{R} \times \mathbb{Z}_+ \to \mathbb{R}$$

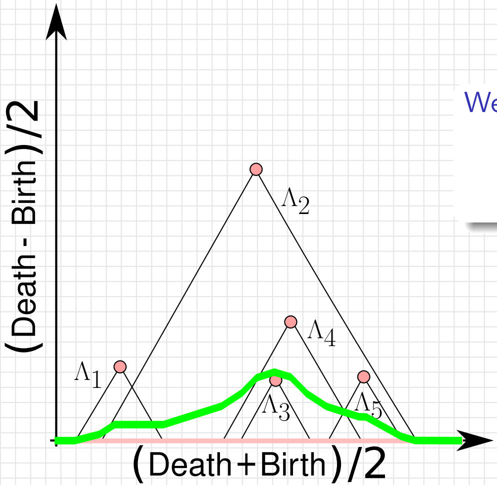
Persistence Silhouettes

Definitions



Persistence Silhouettes

Definitions

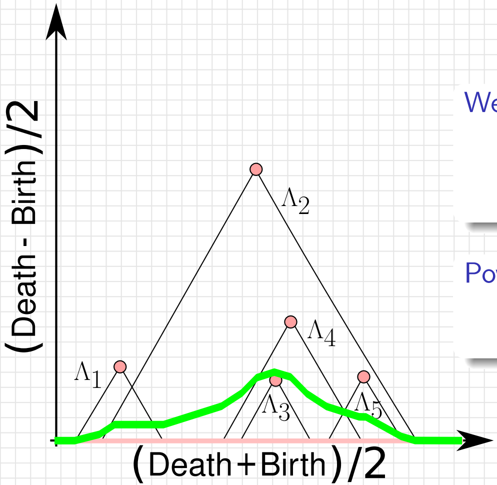


Weighted Silhouette

$$\phi(t) = \frac{\sum_{i=1}^{n} w_i \Lambda_i(t)}{\sum_{j=1}^{n} w_j}$$

Persistence Silhouettes

Definitions



Weighted Silhouette

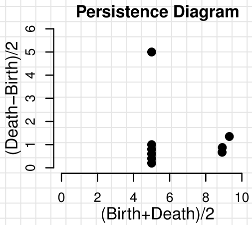
$$\phi(t) = \frac{\sum_{i=1}^{n} w_i \Lambda_i(t)}{\sum_{j=1}^{n} w_j}$$

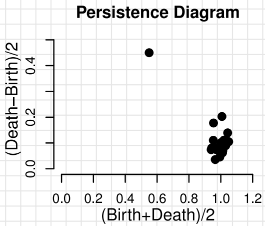
Power-Weighted Silhouette

$$w_i = |d_i - b_i|^p$$

Power-Weighted Silhouettes

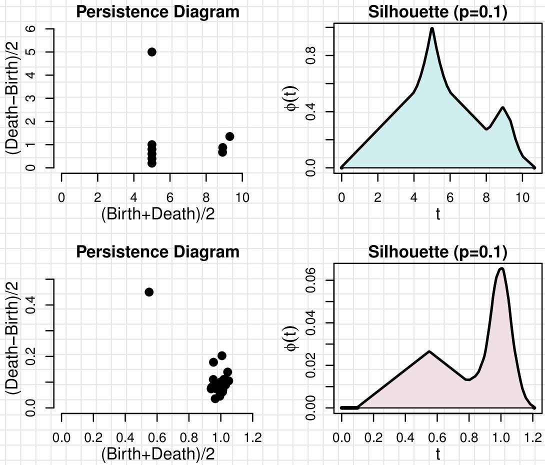
Two Examples





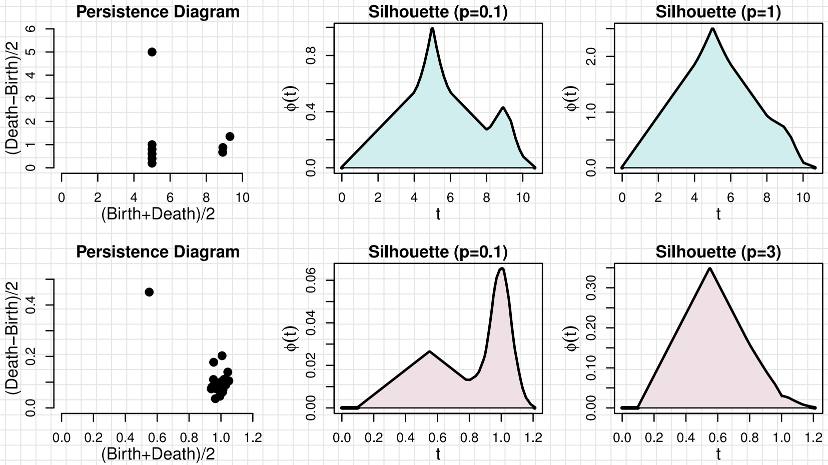
Power-Weighted Silhouettes

Two Examples



Power-Weighted Silhouettes

Two Examples



Persistence Silhouettes

Results

Since ϕ is one-Lipschitz for non-negative weights w_j ...

Convergence of Empirical Process

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\phi_i(t)-\mathbb{E}[\phi(t)]\right)$$

converges weakly to a Brownian bridge, with known rate of convergence.

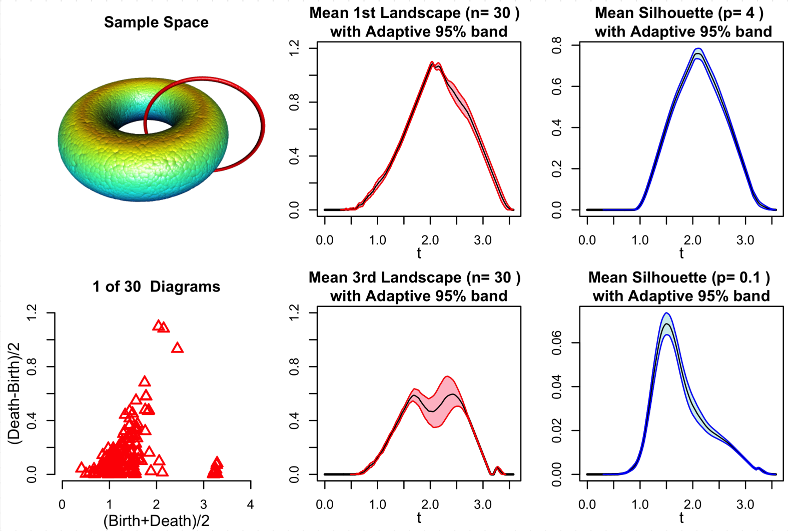
Confidence Bands

We can use the multiplier bootstrap to create a uniform (or a variable width) confidence band defined by ℓ_n^{sil} and u_n^{sil} such that

$$\lim_{n\to\infty}\mathbb{P}\left(\ell_n^{sil}(t)\leq\mu(t)\leq u_n^{sil}(t)\text{ for all }t\right)=1-\alpha.$$

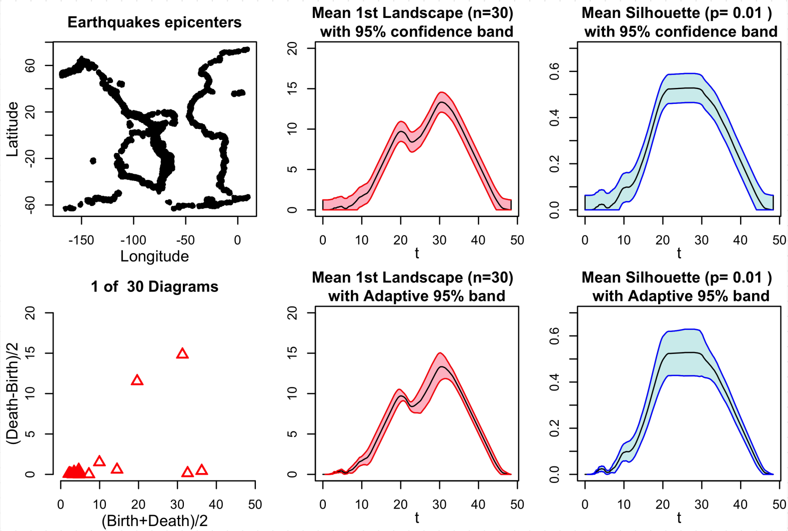
Example I

A Toy Example

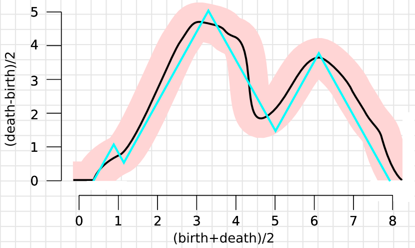


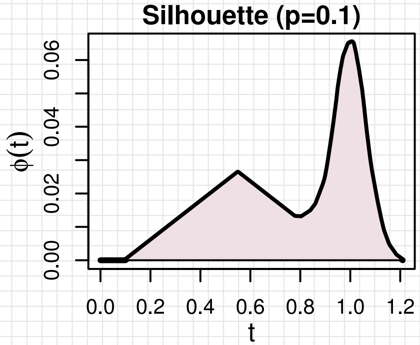
Example II

Earthquake Epicenters

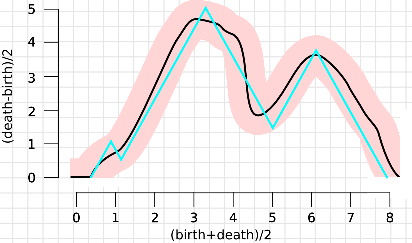


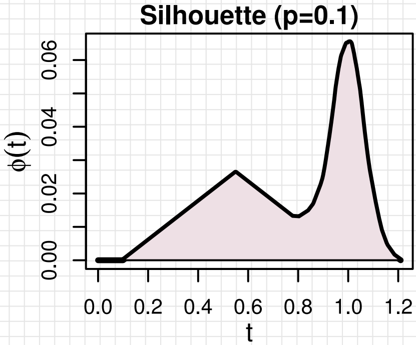
Summary


(death-birth)/2
(birth+death)/2
0
1
2
3
4
5
6
7
8


Silhouette (p=0.1)
φ(t)
t
0.00
0.02
0.04
0.06
0.0
0.2
0.4
0.6
0.8
1.0
1.2

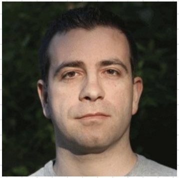
Summary





- First real use of statistical data analysis in TDA.
- Not just theoretical: we've implemented these techniques!
- Develops the theory of subsampling techinques. (Did you see the ArXiv this week?)

Coauthors



Stochastic Convergence of Persistence Landscapes and Silhouettes

Brittany Terese Fasy
brittany.fasy@alumni.duke.edu

joint work with F. Chazal, F. Lecci,
A. Rinaldo, L. Wasserman

Postdoctoral Researcher, Tulane University

11 June 2014