Geometric Inference on Kernel Density Estimates

Bei Wang\textsuperscript{1} \hspace{1cm} Jeff M. Phillips\textsuperscript{2} \hspace{1cm} Yan Zheng\textsuperscript{2}

\textsuperscript{1}Scientific Computing and Imaging Institute, University of Utah
\textsuperscript{2}School of Computing, University of Utah

Feb 6, 2014
Take home message

- **Geometric inference** from a point cloud can be calculated by examining its **kernel density estimate (KDE)** of Gaussians.
- Such an inference is made possible with provable properties through the vehicle of **kernel distance**.
- Such an inference is **robust** to noise and **scalable**.
- We provide an algorithm to estimate the topology of kernel distance using **weighted Vietoris-Rips complexes**.
Geometric inference using the \textbf{kernel distance}, in place of the \textbf{distance to a measure} [Chazal Cohen-Steiner Merigot 2011].

1. \textbf{Robustness} Kernel distance is distance-like: 1-Lipschitz, 1-semiconcave, proper and stable.

2. \textbf{Scalability} Kernel distance has a small coreset, making efficient inference possible on 100 million points.

3. \textbf{Relation to KDE} Geometric inference based on kernel distance works naturally via superlevel sets of KDE: sublevel sets of the kernel distance are superlevel sets of KDE.

4. \textbf{Algorithm} to approximate the sublevel set filtration of kernel distance from a point cloud sample.
People love and are familiar with KDE, especially with Gaussian kernel.

Kernel distance provides a proper way to relate KDE with properties that are crucial for geometric inference.

We could approximate the topology of kernel distance via point cloud samples.
Background on kernels, KDE and kernel distance
A kernel is a similarity measure, more similar points have higher value,

\[ K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^+ \]

We focus on the Gaussian kernel (positive definite):

\[ K(p, x) = \sigma^2 \exp(-\|p - x\|^2/2\sigma^2) \]
A kernel density estimate represents a continuous distribution function over $\mathbb{R}^d$ for point set $P \subset \mathbb{R}^d$:

$$\text{KDE}_P(x) = \frac{1}{|P|} \sum_{p \in P} K(p, x)$$

More generally, it can be applied to any measure $\mu$ (on $\mathbb{R}^d$) as

$$\text{KDE}_\mu(x) = \int_{p \in \mathbb{R}^d} K(p, x) \mu(p) \, dp$$
For two point sets $P$ and $Q$, define similarity

$$\kappa(P, Q) = \frac{1}{|P| |Q|} \sum_{p \in P} \sum_{q \in Q} K(p, q)$$

If $Q = \{x\}$, $\kappa(P, x) = \text{KDE}_P(x)$.

The kernel distance (a metric between $P$ and $Q$):

$$D_K(P, Q) = \sqrt{\kappa(P, P) + \kappa(Q, Q) - 2\kappa(P, Q)}$$

Self similarity minus cross similarity... [Phillips, Venkatasubramanian 2011]
Kernel distance

For two point sets $P$ and $Q$, define similarity

$$\kappa(P, Q) = \frac{1}{|P|} \frac{1}{|Q|} \sum_{p \in P} \sum_{q \in Q} K(p, q)$$

If $Q = \{x\}$, $\kappa(P, x) = \text{KDE}_P(x)$.

The kernel distance (a metric between $P$ and $Q$):

$$D_K(P, Q) = \sqrt{\kappa(P, P) + \kappa(Q, Q) - 2\kappa(P, Q)}$$

Self similarity minus cross similarity... [Phillips, Venkatasubramanian 2011]
For two point sets $P$ and $Q$, define similarity

$$\kappa(P, Q) = \frac{1}{|P|} \frac{1}{|Q|} \sum_{p \in P} \sum_{q \in Q} K(p, q)$$

If $Q = \{x\}$, $\kappa(P, x) = \text{KDE}_P(x)$.

The kernel distance (a metric between $P$ and $Q$):

$$D_K(P, Q) = \sqrt{\kappa(P, P) + \kappa(Q, Q) - 2\kappa(P, Q)}$$

Self similarity minus cross similarity... [Phillips, Venkatasubramanian 2011]
For $D_K(\mu, \nu)$ between two measures $\mu$ and $\nu$, define similarity

$$\kappa(\mu, \nu) = \int_{p \in \mathbb{R}^d} \int_{q \in \mathbb{R}^d} K(p, q) \mu(p) \mu(q) dp dq$$

The kernel distance (a metric between $\mu$ and $\nu$):

$$D_K(\mu, \nu) = \sqrt{\kappa(\mu, \mu) + \kappa(\nu, \nu) - 2\kappa(\mu, \nu)}$$

If $\nu = \text{unit Dirac mass at } x$, $\kappa(\mu, x) = \text{KDE}_\mu(x)$,

$$D_K(\mu, x) = \sqrt{\kappa(\mu, \mu) + \kappa(x, x) - 2\kappa(\mu, x)}$$

$$= \sqrt{c_\mu - 2\text{KDE}_\mu(x)}$$

Kernel distance (current distance or maximum mean discrepancy) is a metric, if the kernel $K$ is characteristic (a slight restriction of being positive definite, e.g. Gaussian and Laplace kernels).
Geometric inference and distance to a measure: a review

[Lieutier 2004] [Chazal, Cohen-Steiner, Lieutier 2009] [Merigot, Ovsjanikov, Guibas 2009] [Chazal, Cohen-Steiner, Merigot 2010] [Biau, Chazal, Cohen-Steiner, Devroye and Rodriguez 2011] [Chazal Cohen-Steiner Merigot 2011]...
Given:

- An unknown object (e.g. a compact set) $S \subset \mathbb{R}^d$
- A finite point cloud $P \subset \mathbb{R}^d$ that comes from $S$ under some process

Aim: Recover topological and geometric properties of $S$ from $P$, e.g. # of components, dimension, curvature...

E.g. preserve homeomorphism, homotopy type, or homology of $S$ from $P$. 

Geometric inference
Reconstructs an approximation of $S$ by offsets from $P$.

Distance function: $f_P(x) = \inf_{y \in P} \|x - y\|$

Offset: $(P)^r = f_P^{-1}([0, r])$

Hausdorff distance: $d_H(S, P) := \|f_S - f_P\|_\infty = \inf_{x \in \mathbb{R}^d} |f_S(x) - f_P(x)|$
Distance function based geometric inference: the intuition

[Hausdorff stability w.r.t. distance functions]
If $d_H(S, P)$ is small, thus $f_S$ and $f_P$ are close, and subsequently, $S$, $(S)^r$ and $(P)^r$ carry the same topology for an appropriate scale $r$.

Theorem (Reconstruction from $f_P$)

Let $S, P \subset \mathbb{R}^d$ be compact sets such that $\text{reach}(S) > R$ and $\varepsilon := d_H(S, P) \leq R/7$. Then $S$ and $(P)^r$ are homotopy equivalent (and even isotopic) if $4\varepsilon \leq r \leq R - 3\varepsilon$.

[Chazal Cohen-Steiner Lieutier 2009] [Chazal Cohen-Steiner Merigot 2011]

$R$ ensures topological properties of $S$ and $(S)^r$ are the same; $\varepsilon$ ensures $(S)^r$ and $(P)^r$ are close, $\varepsilon \approx$ density of the sample.
Not robust to outliers.

If $S' = S \cup x$ and $f_S(x) > R$, then $|f_S - f_{S'}|_\infty > R$: offset-based inference methods fail...
Desirable properties for $f_S$ to be useful in geometric inference:

**(F1)** $f_S$ is 1-Lipschitz: for all $x, y \in \mathbb{R}^d$, $|f_S(x) - f_S(y)| \leq \|x - y\|$.  

**(F2)** $f_S^2$ is 1-semiconcave: $x \in \mathbb{R}^d \mapsto (f_S(x))^2 - \|x\|^2$ is concave.

**(F1)** ensures that $f_S$ is differentiable almost everywhere and the medial axis of $S$ has zero $d$-volume;  
**(F2)** is crucial, e.g. in proving the existence of the flow of the gradient of the distance function for topological inference.
Intuition: $W_2$ distance to $m_0$ fraction of the space.

$\mu$: probability measure on $\mathbb{R}^d$

$m_0 > 0$: a parameter smaller than the total mass of $\mu$

The distance to a measure $d_{\mu,m_0}^{\text{CCM}} : \mathbb{R}^n \rightarrow \mathbb{R}^+, \forall x \in \mathbb{R}^d$, 

$$d_{\mu,m_0}^{\text{CCM}}(x) = \left( \frac{1}{m_0} \int_{m=0}^{m_0} (\delta_{\mu,m}(x))^2 \, dm \right)^{1/2}$$

where $\delta_{\mu,m}(x) = \inf \{ r > 0 : \mu(\overline{B}_r(x)) \leq m \}$.

Wasserstein-$2$ distance $W_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} ||x - y||^2 \, d\pi(x,y) \right)^{1/2}$
Distance to a measure $d_{\mu,m_0}^{CCM}$ is distance-like

(D1) 1-Lipschitz
(D2) 1-semiconcave
(D3) [Stability] For probability measures $\mu$ and $\nu$ on $\mathbb{R}^d$ and $m_0 > 0$, then $\|d_{\mu,m_0}^{CCM} - d_{\nu,m_0}^{CCM}\|_{\infty} \leq \frac{1}{\sqrt{m_0}} W_2(\mu, \nu)$.
(D4) Proper (for Groves Isotopy Lemma).

[Chazal Cohen-Steiner Merigot 2011]
Our Main Results
Our results

Similar properties hold for the kernel distance defined as

\[ d^K_{\mu}(x) = D_K(\mu, x) = \sqrt{\kappa(\mu, \mu) + \kappa(x, x) - 2\kappa(\mu, x)} \]

\[ = \sqrt{c^2_{\mu} - 2KDE_{\mu}(x)} \]

Specifically, the following properties of \( d^K_{\mu} \) allow it to inherit the reconstruction properties of \( d^{\text{CCM}}_{\mu, m_0} \).

(K1) \( d^K_{\mu} \) is 1-Lipschitz.

(K2) \( (d^K_{\mu})^2 \) is 1-semiconvave: the map \( x \mapsto (d^K_{\mu}(x))^2 - \|x\|^2 \) is concave.

(K3) \( d^K_{\mu} \) is proper.

(K4) [Stability] \( \|d^K_{\mu} - d^K_{\nu}\|_\infty \leq D_K(\mu, \nu) \).

For the point cloud setting,

\[ d^K_P(x) = D_K(P, x) = \sqrt{\kappa(P, P) + \kappa(x, x) - 2\kappa(P, x)} \]

\[ = \sqrt{c^2_P - 2\text{KDE}_P(x)} \]
Properness of $d^K_\mu$

A continuous map $f : \mathbb{X} \rightarrow \mathbb{Y}$ between two topological spaces is proper if and only if the inverse image of every compact subset in $\mathbb{Y}$ is compact in $\mathbb{X}$.

Lemma (K3)

$d^K_\mu$ is proper (when its range is restricted to be less than $c_\mu$).

Corollary

The superlevel sets of $\text{KDE}_\mu$ for all ranges whose lower bound $\alpha > 0$ are compact.
Theorem (Isotopy lemma on $d^K_\mu$)

Let $r_1 < r_2$ be two positive numbers such that $d^K_\mu$ has no critical points in $(d^K_\mu)[r_1,r_2]$. Then all the sublevel sets $(d^K_\mu)^r := \{x \in \mathbb{R}^d \mid d^K_\mu(x) \leq r\}$ are isotopic for $r \in [r_1, r_2]$.

Theorem (Reconstruction on $d^K_\mu$, simple version)

Let $d^K_\mu$ and $d^K_\nu$ be two kernel distance functions such that $\|d^K_\mu - d^K_\nu\|_\infty \leq \varepsilon$. Suppose $\text{reach}(d^K_\mu) \geq R$. Then $\forall r \in [4\varepsilon, R - 3\varepsilon]$, and $\forall \eta \in (0, R)$, the sublevel sets $(d^K_\mu)^\eta$ and $(d^K_\nu)^r$ are homotopy equivalent for $\varepsilon \leq R/9$. 
For two probability measures $\mu$ and $\nu$ on $\mathbb{R}^d$,

**Lemma (K4)**

$$\|d_{K\mu}^K - d_{K\nu}^K\|_{\infty} \leq D_K(\mu, \nu).$$

Define $\bar{\mu} = \int_p p \cdot \mu(p)dp$.

We have $\lim_{\sigma \to \infty} D_K(\mu, \nu) = \|\bar{\mu} - \bar{\nu}\|$ and $\|\bar{\mu} - \bar{\nu}\| \leq W_2(\mu, \nu)$.

**Theorem**

$$\lim_{\sigma \to \infty} D_K(\mu, \nu) \leq W_2(\mu, \nu).$$

Open problem: remove need for limit in $\sigma$. 
Stability properties for $d_{\mu}^K$: Comparing $D_K$ to $W_2$

**Lemma (One-sided bound)**

There is no Lipschitz constant $\gamma$ s.t. for any $\mu$ and $\nu$, we have $W_2(\mu, \nu) \leq \gamma D_K(\mu, \nu)$.

Intuition, $W_2$ can grow arbitrarily big, there are some measures from which our bound is tighter.

**Lemma (Special case)**

Consider $\mu$ and $\nu$ where $\nu$ is represented by a Dirac mass at a point $x \in \mathbb{R}^d$. Then $d_{\mu}^K(x) = D_K(\mu, \nu) \leq W_2(\mu, \nu)$ for any $\sigma > 0$, where the equality only holds when $\mu$ is also a Dirac mass at $x$, and the difference decreases for any other $x$ as $\sigma$ increases.
Stability properties for $d^K_{\mu}$: Stability with Respect to $\sigma$

Theorem

For any $x$, $d^K_{\mu}(x)$ is $\ell$-Lipschitz with respect to $\sigma$, for

$$\ell = \frac{18}{e^3} + \frac{8}{e} + 2 < 6.$$ 

$\sigma \approx$ geometric notion of an outlier parameter
Algorithm:
Approximate the persistence diagram of sublevel sets filtration of kernel distance using weighted Rips filtration
Metric space \((\mathbb{X}, d_{\mathbb{X}}(\cdot, \cdot))\), a set \(P \subseteq \mathbb{X}\) and a function \(w : P \to \mathbb{R}\), the (general) power distance \(f\) associated with \((P, w)\) is

\[
f(x) = \sqrt{\min_{p \in P} (d_{\mathbb{X}}(p, x)^2 + w(p)^2)}.
\]

Persistence diagram of \(w\) can be approximated by weight Rips filtration based on \((P, w)\) and \(f\).
Sublevel set of $f$,

$$f^{-1}((-\infty, \alpha])$$

is the union of balls centered at points $p \in P$ with radius

$$r_p(\alpha) = \sqrt{\alpha^2 - w(p)^2}$$

for each $p$. 
Sublevel set of $f$,

$$f^{-1}((-\infty, \alpha])$$

is the union of balls centered at points $p \in P$ with radius $r_p(\alpha) = \sqrt{\alpha^2 - w(p)^2}$ for each $p$.

Weighted Čech complex $C_\alpha(P, w)$ for parameter $\alpha$ is the union of simplices $s$ such that $\bigcap_{p \in s} B(p, r_p(\alpha)) \neq 0$. 

Weighted Rips complex $R_\alpha(P, w)$ for parameter $\alpha$ is the maximal complex whose $1$-skeleton is the same as $C_\alpha(P, w)$.
Sublevel set of $f$,

$$f^{-1}((-\infty, \alpha])$$

is the union of balls centered at points $p \in P$ with radius $r_p(\alpha) = \sqrt{\alpha^2 - w(p)^2}$ for each $p$.

Weighted Čech complex $C_\alpha(P, w)$ for parameter $\alpha$ is the union of simplices $s$ such that $\bigcap_{p \in s} B(p, r_p(\alpha)) \neq 0$.

Weighted Rips complex $R_\alpha(P, w)$ for parameter $\alpha$ is the maximal complex whose 1-skeleton is the same as $C_\alpha(P, w)$. 
Sublevel set of $f$,

\[ f^{-1}((-\infty, \alpha]) \]

is the union of balls centered at points $p \in P$ with radius

\[ r_p(\alpha) = \sqrt{\alpha^2 - w(p)^2} \]

for each $p$.

Weighted Čech complex $C_\alpha(P, w)$ for parameter $\alpha$ is the union of simplices $s$ such that $\bigcap_{p \in s} B(p, r_p(\alpha)) \neq 0$.

Weighted Rips complex $R_\alpha(P, w)$ for parameter $\alpha$ is the maximal complex whose 1-skeleton is the same as $C_\alpha(P, w)$.

Weighted Rips filtration: $\{R_\alpha(P, w)\}$.
A power distance using $d^K_\mu$ for a measure $\mu$ is defined with a point set $P \subset \mathbb{R}^d$ and a metric $d(\cdot, \cdot)$ on $\mathbb{R}^d$, 

$$f_P(\mu, x) = \sqrt{\min_{p \in P} \left( d(p, x)^2 + d^K_\mu(p)^2 \right)}.$$
A power distance using $d^K_\mu$ for a measure $\mu$ is defined with a point set $P \subset \mathbb{R}^d$ and a metric $d(\cdot, \cdot)$ on $\mathbb{R}^d$, 

$$f_P(\mu, x) = \sqrt{\min_{p \in P} (d(p, x)^2 + d^K_\mu(p)^2)}.$$ 

We consider $d(p, x) := D_K(p, x)$.

Let $p_+ = \arg \max_{q \in \mathbb{R}^d} \kappa(\mu, q)$ and $P_+ = P \cup p_+$. We have 

$$\frac{1}{\sqrt{2}} d^K_\mu(x) \leq f^*_P(\mu, x) \leq \sqrt{14} d^K_\mu(x).$$
A power distance using $d^K_{\mu}$ for a measure $\mu$ is defined with a point set $P \subset \mathbb{R}^d$ and a metric $d(\cdot, \cdot)$ on $\mathbb{R}^d$,

$$f_P(\mu, x) = \sqrt{\min_{p\in P} \left( d(p, x)^2 + d^K_{\mu}(p)^2 \right)}.$$ 

We consider $d(p, x) := D_K(p, x)$.

Let $p_+ = \arg \max_{q \in \mathbb{R}^d} \kappa(\mu, q)$ and $P_+ = P \cup p_+$. We have

$$\frac{1}{\sqrt{2}} d^K_{\mu}(x) \leq f^K_{P_+}(\mu, x) \leq \sqrt{14} d^K_{\mu}(x).$$

However, constructing $p_+$ exactly seems quite difficult. We use its approximation $\hat{p}_+$ s.t. $D_K(P, \hat{p}_+) \leq (1 + \delta)D_K(P, p_+)$ (obtained through an algorithm that is polynomial in $n$ and $1/\delta$ under reasonable conditions). Define $\hat{P}_+ = P \cup \hat{p}_+$. 
Constructive topological estimation using $d^K_{\mu}$

**Theorem**

The weighted Rips filtration $\{R_\alpha(\hat{P}_+, d^K_{\mu_P})\}$ can be used to approximate the persistence diagram of $d^K_{\mu_P}$ such that $d_{ln}^B(\text{Dgm}(d^K_{\mu_P}), \text{Dgm}(\{R_\alpha(\hat{P}_+, d^K_{\mu_P})\})) \leq \ln(16)$.

**Proof** based on the power distance mechanisms [Buchet, Chazal, Oudot and Sheehy 2013], $\epsilon$-interleaving of persistence modules.

**Open problems:**

- Let $d(p, x) = ||x - p||$
- Calculate $\hat{p}_+$ without restrictions, efficiently (current works, but messy)
- Tighten bounds
Comments on relating $d^K_\mu$ with $f_S$

Figure: Showing that $\|x - 0\|/2 \leq D_K(x, 0) \leq \|x - 0\|$, where the second inequality holds for $\|x\| \leq \sqrt{3}\sigma$. The kernel distance $D_K(x, 0)$ is shown for $\sigma = \{1/2, 1, 2\}$ in purple, blue, and red, respectively.
Algorithmic and Approximation Observations: Advantages
Advantages of the kernel distance summary

(I) Small coreset representation for sparse representation and efficient, scalable computation.

(II) Its inference is easily interpretable and computable through the superlevel sets of a KDE.

(III) It is Lipschitz with respect to the outlier parameter $\sigma$ when the input $x$ is fixed.

(IV) As $\sigma \to \infty$, the kernel distance is bounded by the Wasserstein distance: $\lim_{\sigma \to \infty} D_K(\mu, \nu) \leq W_2(\mu, \nu)$. 
Small coreset

- There exists a small $\epsilon$-coreset $Q \subset P$ s.t. $\|d^K_P - d^K_Q\|_\infty \leq \epsilon$ and $\|\text{KDE}_P - \text{KDE}_Q\|_\infty \leq \epsilon$ with probability at least $1 - \delta$.
- Size $O(((1/\epsilon)^{(2d/(d+2))} \sqrt{\log(1/\epsilon \delta)})$ [Phillips 2013].
- The same holds under a random sample of size $O((1/\epsilon^2)(d + \log(1/\delta)))$ [Joshi Kommaraju Phillips 2011].
- Operate with $|P| = 100,000,000$ [Zheng Jestes Phillips Li 2013].
- Stability of persistence diagram is preserved: $d_B(\text{Dgm(\text{KDE}_P)}, \text{Dgm(\text{KDE}_Q)}) \leq \epsilon$. 
There exists a small $\epsilon$-coreset $Q \subset P$ s.t. $\|d_P^K - d_Q^K\|_\infty \leq \epsilon$ and $\|KDE_P - KDE_Q\|_\infty \leq \epsilon$ with probability at least $1 - \delta$.

Size $O(((1/\epsilon)\sqrt{\log(1/\epsilon\delta)})^{2d/(d+2)})$ [Phillips 2013].

The same holds under a random sample of size $O((1/\epsilon^2)(d + \log(1/\delta)))$ [Joshi Kommaraju Phillips 2011].

Operate with $|P| = 100,000,000$ [Zheng Jestes Phillips Li 2013].

Stability of persistence diagram is preserved: $d_B(Dgm(KDE_P), Dgm(KDE_Q)) \leq \epsilon$. 

\[ 
\begin{array}{c}
\text{Black dots represent data points.}
\end{array} 
\]
Recall $d^K_P(x) = \sqrt{c_P^2 - 2\text{KDE}_P(x)}$ where $c_P^2$ is a constant that depends only on $P$. Perform geometric inference on noisy $P$ by considering the super-level sets of $\text{KDE}_P$,

$$\{x \in \mathbb{R}^d \mid \text{KDE}_P(x) \geq \tau\}$$

Key:

- $d^K_P(\cdot)$ is monotonic with $\text{KDE}_P(\cdot)$; as $d^K_P(x)$ gets smaller, $\text{KDE}_P(x)$ gets larger.

- A clean and natural interpretation of the reconstruction problem through the well-studied lens of KDE. Geometric inference with sublevel sets of $d^K_P$ (superlevel sets of $\text{KDE}_P$).
Experiments

An example with $25\%$ of $P$ as noise, $\sigma = 0.05$
Experiments

An example with 25% of $P$ as noise, $\sigma = 0.003$
Experiments

An example with 25% of $P$ as noise, $\sigma = 0.001$
Comparison with distance to a measure

Figure: Sublevel sets of kernel distance (a), and distance to a measure (b) on a data set with two circles with different scales and densities. These plots fix a level set $\gamma$ and vary (a) $\sigma$ for $d^K_\mu$ and (b) $m_0$ for $d^{CCM}_{\mu,m_0}$. The parameters are chosen to provide similar appearing images.
Future directions
More general theory for KDE with systematic understanding of family of kernels: distance to a measure (KNN kernel), kernel distance (a larger class of kernels, e.g. Gaussian, Laplace; triangle kernel may work OK in practice with less perfect properties).
Laplace kernel $K(p, x) = \exp(-2\|x - y\|/\sigma)$
Triangle kernel: 

\[ K(x, p) = \max \left\{ 0, 1 - \frac{\|p-x\|}{\sigma=0.05} \right\} \]
Epanechnikov kernel: (reconstruction)

\[ K(x, p) = \max \left\{ 0, 1 - \frac{\|p-x\|^2}{(\sigma=0.05)^2} \right\} \]
Ball kernel: 
\[ K(x, p) = \begin{cases} 1 & \text{if } \|p - x\| \leq \sigma = 0.05 \\ 0 & \text{otherwise} \end{cases} \]

\(\alpha\)-shape can be viewed as using the ball kernel with \(\sigma = \alpha\) and \(r = 1/n\).
Alternative KDEs
Two parameters $r$ (isolevel) and $\sigma$ (outlier/bandwidth) that control the scale.

Figure: Sublevel sets for the kernel distance. Left: fix $\sigma$, vary $r$. Right: fix $r$, vary $\sigma$. The values of $\sigma$ and $r$ are chosen to make the plots similar.
The authors thank D. Sheehy, F. Chazal and the rest of the Geometrica group at INRIA-Saclay for enlightening discussions on geometric and topological reconstruction. We also thank D. Sheehy for personal communications regarding the power distance constructions.
Paper available at: http://arxiv.org/abs/1307.7760

jeffp@cs.utah.edu
http://www.cs.utah.edu/~jeffp/
beiwang@sci.utah.edu
http://www.sci.utah.edu/~beiwang/