Diffusion maps and radar data analysis

Y. S. Bhat[a] and Gregory Arnold[b]

[a]Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, Minnesota 55455, USA;
[b]Air Force Research Lab, AFRL/SNAT, Bldg. 620, 2241 Avionics Circle, Dayton, Ohio 45433, USA

ABSTRACT

Understanding and organizing data, in particular understanding the key modes of variation in the data, is a first step toward exploiting and evaluating sensor phenomenology. Spectral theory and manifold learning methods have been recently shown to offer sever powerful tools for many parts of the exploitation problem. We will describe the method of diffusion maps and give some examples with radar (backhoe data dome) data. The so-called diffusion coordinates are kernel based dimensionality reduction techniques that can, for example, organize random data and yield explicit insight into the type and relative importance of the data variation. We will provide sufficient background for others to adopt these tools and apply them to other aspects of exploitation and evaluation.

Keywords: manifold learning, diffusion geometry, laplacian eigenmap, spectral graph theory, radar

1. INTRODUCTION

We wish to explore the implementation of a dimension reduction algorithm which utilizes spectral graph theory, *i.e.* the "diffusion map" approach.[1] We use this approach to explore the manifold geometry of a diffusion coordinate representation of radar data. Our assumption is that the data set satisfies the "manifold model", *i.e.* the data set can be represented as points on a manifold in Euclidean space. Data collected by sensors often have a high-dimensional representation but may in fact be the result of a low-dimensional phenomenon. Furthermore, data collected by sensors often reflect the distribution of the information sampling process. We seek a representation of the data that is invariant to the sampling distribution. The procedure discussed here can be shown[2] to represent the data independently of the sampling distribution. While diffusion maps have been used for various applications,[3] to the best of our knowledge radar data has not been analyzed via diffusion maps. We wish to determine if diffusion maps are a viable strategy for use in radar, *i.e.* will radar data represented by diffusion coordinates have some type of geometry? Furthermore, what role will polarization play in this representation? We explore these issues.

In the course of implementing the algorithm we will explore various choices for certain parameters that occur in the algorithm. One of the fundamental assumptions of the spectral graph theory approach is that the pertinent information occurs locally. However, this notion of "local information" needs to be clarified with respect to parameter values. We specifically look at the space of distances between the data points and use different descriptive statistics of this space to guide our choice of parameter values. The paper is organized as follows. In Section 2 we explicitly discuss the diffusion map algorithm. The full algorithm is presented in Algorithm 1. In Section 3 we perform numerical experiments on SAR Backhoe Data Dome Version 1.0.[4] In Section 3.1 we represent range histories as diffusion coordinates and in Section 3.2 we consider the squiggle path data. In Section 4 we discuss the results of our experiments.

Further author information: (Send correspondence to G.A.)
Y.S.B.: E-mail: bhat@ima.umn.edu, Telephone: 1 612 624 6066
G.A.: E-mail: Gregory.Arnold@wpafb.af.mil, Telephone: 1 937 904 9873

2. THE METHOD OF DIFFUSION MAPS

In this section we discuss the method of representing a given data set $X = \{x_1, x_2, \ldots, x_N\}$ in terms of *diffusion coordinates*. The basis of this approach is in graph theory and we will accordingly first begin by discussing some basic notions of graph theory.

A *graph* is defined as a set of vertices V and a set of edges E. In what is to follow we will only consider finite graphs, *i.e.* the set of vertices $V = \{v_1, v_2, \ldots, v_n\}$ is finite. The set of edges E is a subset of the set of all unordered pairs of points of V. Any pair of points $(v_i, v_j) \in E$ are said to be *adjacent* vertices of the graph $G = (V, E)$. A *weighted graph* is a graph with an associated weight function $w : V \times V \to \mathbb{R}$. Furthermore the weight function is required to satisfy the following conditions:

(a) Symmetry: $w(v_i, v_j) = w(v_j, v_i)$ for any $v_i, v_j \in V$.

(b) Non-negativity: $w(v_i, v_j) \geq 0$ for any $v_i, v_j \in V$.

The *degree of the vertex* v_i, denoted by d_{v_i} is defined to be

$$d_{v_i} = \sum_{v_j \in V} w(v_i, v_j).$$

Define the $n \times n$ matrix T to be the diagonal matrix with $(T)_{ii} = d_{v_i}$ and define the $n \times n$ matrix L such that

$$(L)_{ij} = L(v_i, v_j) = \begin{cases} d_{v_i} - w(v_i, v_i) & \text{if } v_i = v_j, \\ -w(v_i, v_j) & \text{if } v_i \text{ and } v_j \text{ are adjacent}, \\ 0 & \text{otherwise}. \end{cases}$$

then the *Laplacian* of the graph G is defined to be

$$\mathcal{L} = T^{-1/2} L T^{-1/2}$$

where $(T^{-1})_{ii} \equiv 0$ if $d_{v_i} = 0$. Defining the transition probability $P(v_i, v_j)$ as

$$P(v_i, v_j) = \frac{w(v_i, v_j)}{d_{v_i}} \text{ for any } v_i, v_j \in V,$$

determines a random walk on the weighted graph. The random walk has probability $P(v_i, v_j)$ of moving from v_i to v_j. Using the transition probability we can construct an $n \times n$ matrix of transition, $\mathcal{P}$, where $(\mathcal{P})_{ij} = P(v_i, v_j)$. The eigenvectors and eigenvalues of $\mathcal{P}$ have been used for dimension reduction.[5] In the case of unweighted graphs we have[6] that $\mathcal{P} = T^{-1/2}(I - \mathcal{L})T^{1/2}$. We refer the reader to Chung[6] and Bollobás[7] for a more detailed discussion of graph theory.

For our purposes, the data set $X = \{x_1, x_2, \ldots, x_N\}$ is considered to be the set of vertices of a weighted graph with weight $w(\cdot, \cdot)$. The goal is to represent each $x_i \in \mathbb{R}^d$ as a point $y_i \in \mathbb{R}^m$ where $m \ll d$ but so that the set $Y = \{y_1, y_2, \ldots, y_N\}$ still retains all the intrinsic geometric information of the original data set.

Assuming the "manifold model", *i.e.* assuming that the original data set is in fact intrinsically low-dimensional and assuming that the points lie on a submanifold $\mathcal{M}$ of $\mathbb{R}^d$; Belkin and Niyogi[5] show that the eigenmaps of the Laplace Beltrami operator are, in some sense, a "good" choice to use for the embedding of the original data set. The geometry of the manifold is taken into account by using the Laplace Beltrami operator. In particular, the local information is well preserved through the use of these eigenmaps and furthermore there is an intrinsic connection between the Laplacian of a graph and the Laplace Beltrami operator, Δ, on manifolds. The Laplace Beltrami operator can be approximated in the discrete setting by the Laplacian of an appropriately weighted graph. The connection between the Laplace Beltrami operator on differentiable functions on a manifold and heat flow lead us to the heat kernel as a choice for an appropriate weight function. Belkin and Niyogi[5] approximate the heat kernel by the Gaussian weight $w_\epsilon(x_i, x_j) = \exp(-\|x_i - x_j\|^2/\epsilon)$ where $\|\cdot\|$ is the standard Euclidean metric and ϵ is a scale parameter. See Belkin and Niyogi[5] for a detailed discussion of Laplacian eigenmaps.

Coifmann, Keller and Lafon[2] show that using the Gaussian weight yields a representation of the data that is strongly correlated to the sampling distribution of the data points. A strategy for finding a distribution-independent representation is provided in Lafon.[1] The strategy is to normalize the Gaussian weight function w_ϵ. The normalized weight function $\tilde{w}_\epsilon$ is defined as

$$\tilde{w}_\epsilon(x_i, x_j) = \frac{w_\epsilon(x_i, x_j)}{q_\epsilon(x_i)q_\epsilon(x_j)},$$

where

$$q_\epsilon(x_i) = \sum_{x_k \in X} w_\epsilon(x_i, x_k).$$

We present the complete algorithm in Algorithm 1 (adapted from the algorithm presented by Lafon[1]). The normalization of the weight function is performed in steps 2 and 3, then the matrix of transition is created in steps 4 and 5. Finally, in steps 6 and 7, the eigenvectors and eigenvalues of the Laplace Beltrami operator are approximated.

Algorithm 1: Diffusion Map Algorithm

Input: $X = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^d, \epsilon, t, m$

Output: $Y = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^m$

1. Construct the matrix $\mathbf{W}_\epsilon$ with entries $(\mathbf{W}_\epsilon)_{ij} = \exp(-\|x_i - x_j\|^2/\epsilon)$.
2. Let $\mathbf{q} = \mathbf{W}_\epsilon * \mathbf{1}$ where $\mathbf{1} = (1, 1, \ldots, 1)'$.
3. Define $\tilde{\mathbf{W}}_\epsilon = \mathbf{W}_\epsilon./(\mathbf{q} * \mathbf{q}')$.
4. Let $\mathbf{d} = \text{sqrt}(\tilde{\mathbf{W}}_\epsilon * \mathbf{1})$.
5. Define $\mathbf{P} = \tilde{\mathbf{W}}_\epsilon./(\mathbf{d} * \mathbf{d}')$.
6. Diagonalize $\mathbf{P}$ by $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{P})$.
7. The eigenvalues of Δ are approximated by the diagonal entries of $\mathbf{S}$ and the eigenfunctions ϕ_j are approximated by $\mathbf{U}(:, j)./\mathbf{U}(:, 1)$.
8. The diffusion map is defined as $x_i \to y_i = (\lambda_2^{t/2}\phi_2(i), \lambda_3^{t/2}\phi_3(i), \ldots, \lambda_m^{t/2}\phi_m(i))'$ where $\phi_j(i)$ is the i-th component of the j-th eigenvector ϕ_j and $\lambda_j = s_{jj}$.

  NOTE: We are using MATLAB notation; "./" and "sqrt" denote element-wise operations and "svd" indicates we have performed a singular value decomposition.

Let us now formally define a *diffusion map*. Let $(X, \mathcal{X}, \mu)$ be a measure space, where X is the set of finitely many data points, and μ is the counting measure. Let W be defined as the *diffusion operator* with kernel $\tilde{w}_\epsilon$, *i.e.* for any $f \in L^2(X, d\mu)$

$$W f(x) \equiv \int_X \tilde{w}_\epsilon(x, y)f(y)d\mu(y).$$

Then since W is self-adjoint and assuming W is bounded and compact, by Mercer's Theorem[8] we have

$$\tilde{w}_\epsilon(x, y) = \sum_{j \geq 0} \lambda_j \phi_j(x)\phi_j(y)$$

where

$$W\phi_j(x) = \lambda_j \phi_j(x).$$

One can show that the eigenvalues are non-negative and furthermore

$$\lambda_0 = 1 \geq \lambda_1 \geq \lambda_2 \geq \ldots.$$

We can now implicitly define the family of *diffusion distances* $\{D_t\}_{t \geq 1}$ by

$$D_t^2(x, y) = \sum_{j \geq 0} \lambda_j^t(\phi_j(x) - \phi_j(y))^2 \tag{1}$$

where the scale parameter t controls the sensitivity of the diffusion metric D_t to the eigenvalues ϕ_j, *i.e.* since $0 \leq \lambda_j \leq 1$, a large value of t will cause the spectrum to decay rapidly thereby causing D_t to be largely dependent on only a few eigenvectors ϕ_j. The framework discussed above can be viewed from the point of view of diffusion processes. According to Coifman et. al,[9] $D_t(x,y)$ "measures the rate of connectivity of the points x and y by paths of length..." t and "unlike the geodesic distance, this metric is robust to perturbations on the data."

Finally, let us define the family of *diffusion maps* $\{\Phi_t\}_{t \geq 1}$ by

$$x \rightarrow \Phi_t(x) = (\lambda_0^{t/2}\phi_0(x), \lambda_1^{t/2}\phi_1(x), \lambda_2^{t/2}\phi_2(x), \dots)' \tag{2}$$

but note that it can be shown that ϕ_0 is a constant and so in practice the diffusion map used is

$$x \rightarrow (\lambda_1^{t/2}\phi_1(x), \lambda_2^{t/2}\phi_2(x), \dots, \lambda_{m(t)}^{t/2}\phi_{m(t)}(x))'$$

where $m(t)$ is that number for which the eigenvalues $\{\lambda_j^{t/2}\}_{j > m(t)}$ are numerically insignificant. Note that in light of equation (2), the diffusion distance (1) can be defined as

$$D_t(x,y) = \|\Phi_t(x) - \Phi_t(y)\| \tag{3}$$

where $\|\cdot\|$ is the standard Euclidean norm. We refer the interested reader to Lafon[1] for a detailed construction of the diffusion map. Note that the second eigenvector ϕ_1 is known as the *Fiedler vector* (see Lévy[10]) and can be used to order the underlying data set X (see Higgs, Solka and Weller[11] for example). Also, the terms ϵ, t, and $m(t)$ are all parameters of Algorithm 1. In Section 3 and 4 some strategies for selecting the parameter ϵ will be discussed. In the next section we apply this diffusion map approach to the Backhoe Data Dome.

3. NUMERICAL EXPERIMENTS

In this section we wish to provide some numerical experiments of the dimension reduction strategy outlined in Algorithm 1. The goal of this section is to highlight the implementation of the algorithm to radar data and to experiment with the parameter ϵ. The Backhoe Data Dome is simulated wideband ($7-13$ Ghz), full polarization, complex backscatter data from a backhoe vehicle. The data is public release #ASC 04–0273 and is available from the AFRL/SNA Sensor Data Management System. The k-space data for each square degree is given for $66° <$ azimuth $< 114°$ and $18° <$ elevation $< 42°$. For each square degree the data is a 3 dimensional complex array of size $512 \times 14 \times 14$ where the first dimension is the frequency range history evenly spaced between 7 and 13 GHz and the second and third dimensions are Elevation and Azimuth respectively. The data set in both our experiments is a collection of range histories over some range of (az, el) pairs, *i.e.* if $X = \{x_1, x_2, \dots, x_N\}$ is the data set, then each x_i is a range history represented as a vector in $\mathbb{C}^{512}$. For both experiments we use the Gaussian kernel $w_\epsilon(x_i, x_j) = \exp(-\|x_i - x_j\|_{L^2}^2/\epsilon)$ and we consider different choices of ϵ. We use $t = 2$ and we use $(\lambda_2\phi_2, \lambda_3\phi_3, \lambda_4\phi_4)$ as our embedding. For our choice of ϵ we employ several different strategies but in all cases we wish to pick ϵ so that the data points are numerically connected, *i.e.* we pick ϵ such that for each $i \in \{1, 2, \dots, N\}$ there exists at least one $j \in \{1, 2, \dots, N\}$, $j \neq i$, such that $w_\epsilon(x_i, x_j) \neq 0$. Accordingly, we consider the set

$$\{\|x_i - x_j\|_{L^2}^2 \big| i, j = 1, 2, \dots, N\}$$

and use certain statistics of this set or of subsets of this set as our choice of values for ϵ. Lafon[1] suggests

$$\epsilon = \frac{1}{N}\sum_{i=1}^{N}\min\{\|x_i - x_j\|_{L^2}^2 \big| \|x_i - x_j\|_{L^2} > 0, \ j = 1, 2, \dots, N\}, \tag{4}$$

while Higgs, Solka and Weller[11] use

$$\epsilon = \min\{\|x_i - x_j\|_{L^2}^2 \big| \|x_i - x_j\|_{L^2} > 0, \ i, j = 1, 2, \dots, N\}. \tag{5}$$

3.1. Radar Data

We wish to represent radar data using diffusion maps. In Figures 1–6 we consider the k-space data for $88° \leq$ az $< 92°$ and $28° \leq$ el $< 32°$ where only the vertical polarization data is used. Then with the Gaussian kernel and the procedure outlined in Algorithm 1, we embed the range history by diffusion coordinates. In Figure 3 and 4 we use color to correspond the geo-spatial location of the data collected to points on the manifold. In all the other figures of manifolds that appear, the horizontal polarization data is plotted as a darker shaded point while the vertical polarization data is plotted as a lighter shaded point. In Figure 1 we use

$$\epsilon = \frac{1}{N} \sum_{i=1}^{N} \max\{\|x_i - x_j\|_{L^2}^2 \big| j = 1, 2, \ldots, N\}, \tag{6}$$

and in Figure 2 we graph the projection of the manifold in two dimensions. In Figure 5 we use

$$\epsilon = \max\left\{ \min\{\|x_i - x_j\|_{L^2}^2 \big| \|x_i - x_j\|_{L^2} > 0, j = 1, 2, \ldots, N\}\big| i = 1, 2, \ldots, N\right\}, \tag{7}$$

with $t = 2$. In Figure 6 we use (4) with $t = 2$.

In Figure 8–10 we consider all the k-space data measurements for $106° \leq$ az $< 109°$ and $18° \leq$ el $< 21°$ where both the horizontal and vertical polarization data is used. In Figure 7 we graph the spectrum decay associated to diffusion coordinates for both data sets respectively. In Figure 8 we use (6) with $t = 2$. In Figure 8 on the left we plot the vertical polarization data separately. In Figures 9 we use (4) with $t = 2$. In Figure 10 we use (5) and $t = 2$. Notice in Figure 10 we see that the vertical and horizontal polarization data account for distinct portions of the diffusion embedding. In Figure 11 graph the manifold that results if we consider the VV data only for $106° \leq$ az $< 109°$ and $18° \leq$ el $< 21°$ for $\epsilon = 205$ thousand. In Figure 11, on the left, we graph the portion of the diffusion coordinates of the embeding seen in Figure 10 due to VV polarization data. On the right we graph the embedding that results if we restrict our initial data set to the VV polarization data only.
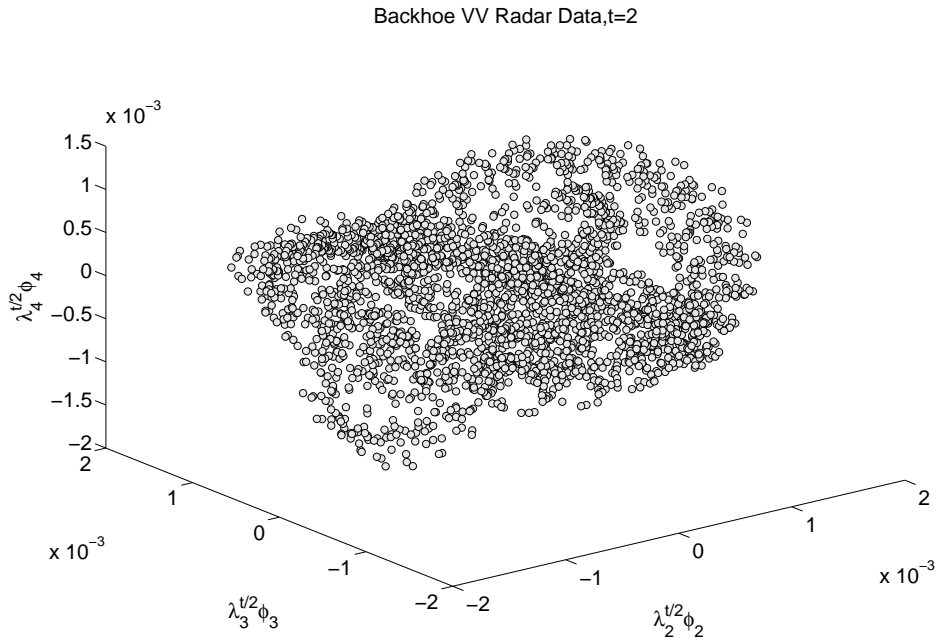

Backhoe VV Radar Data,t=2

Figure 1. The diffusion embedding that results when we use $\epsilon \approx 84$ million.
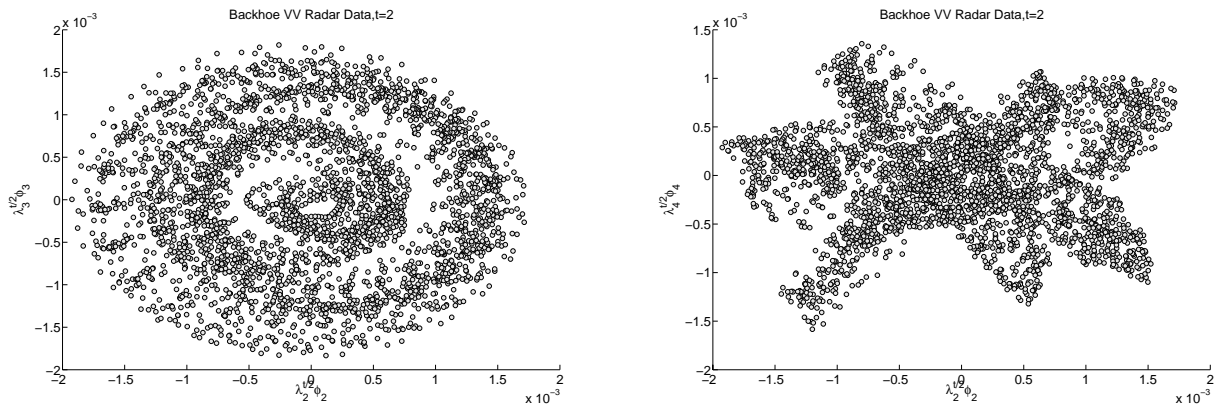
Figure 2. *Left:* Projection of the diffusion embedding in two dimensions. *Right:* Projection of the diffusion embedding in two dimensions
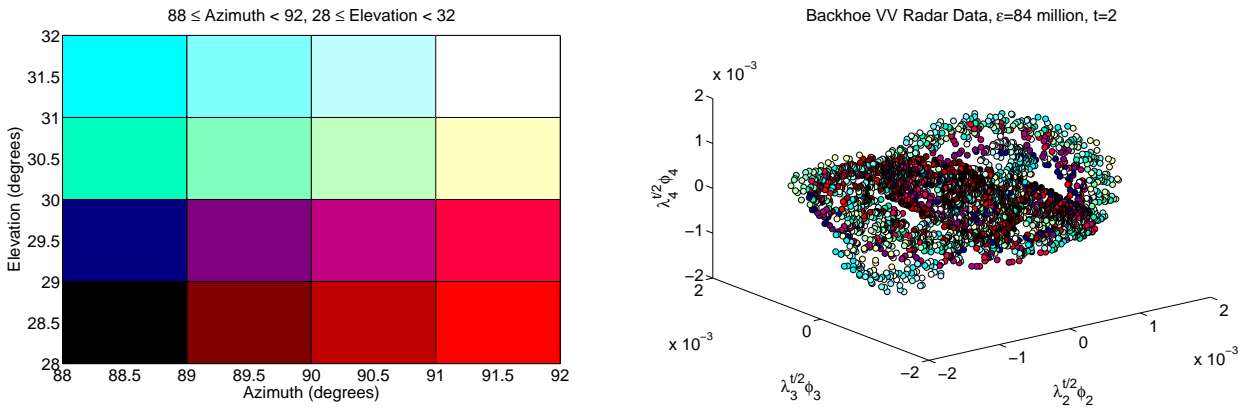

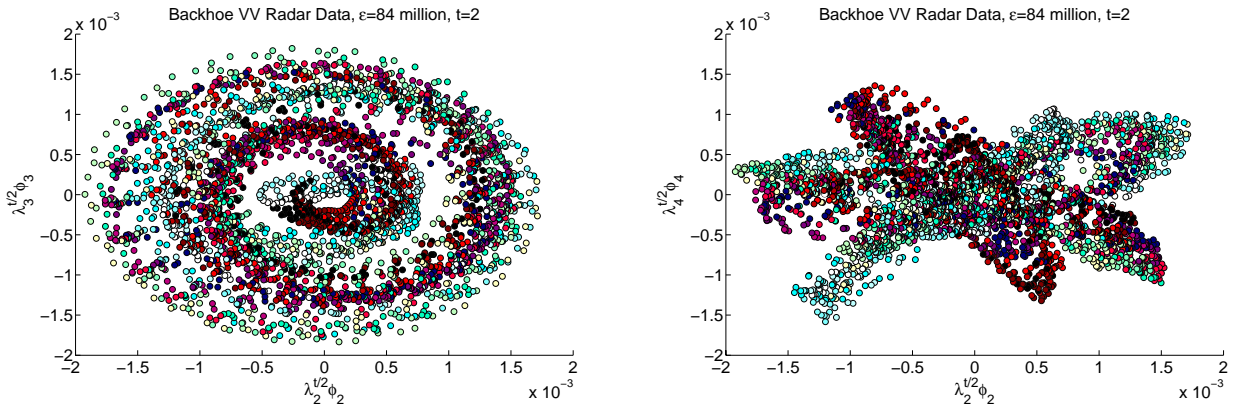**Figure 3.** We use color to correspond the Azimuth and Elevation information (left) to points on the embedding (right).

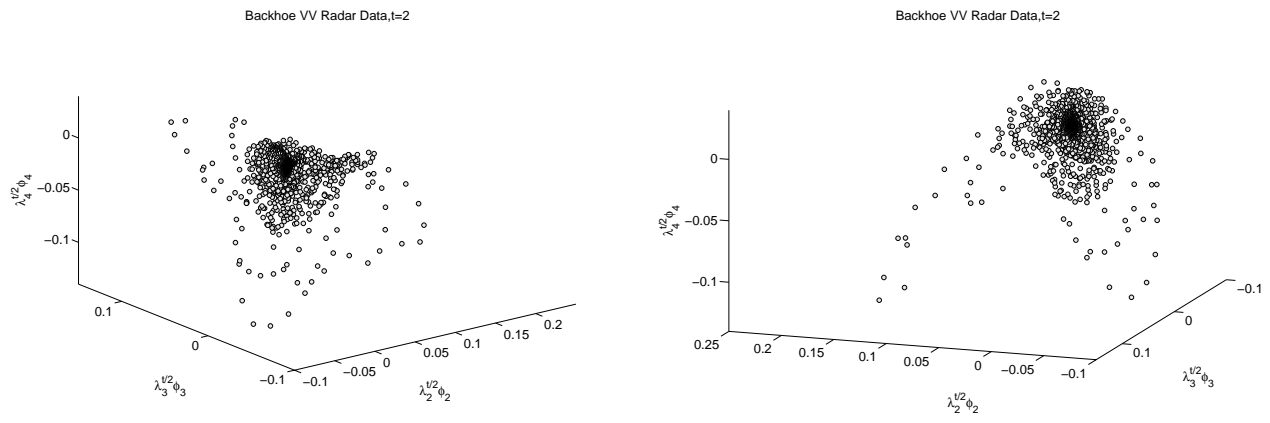**Figure 4.** Colors correspond to Azimuth and Elevation information, see Figure 3.

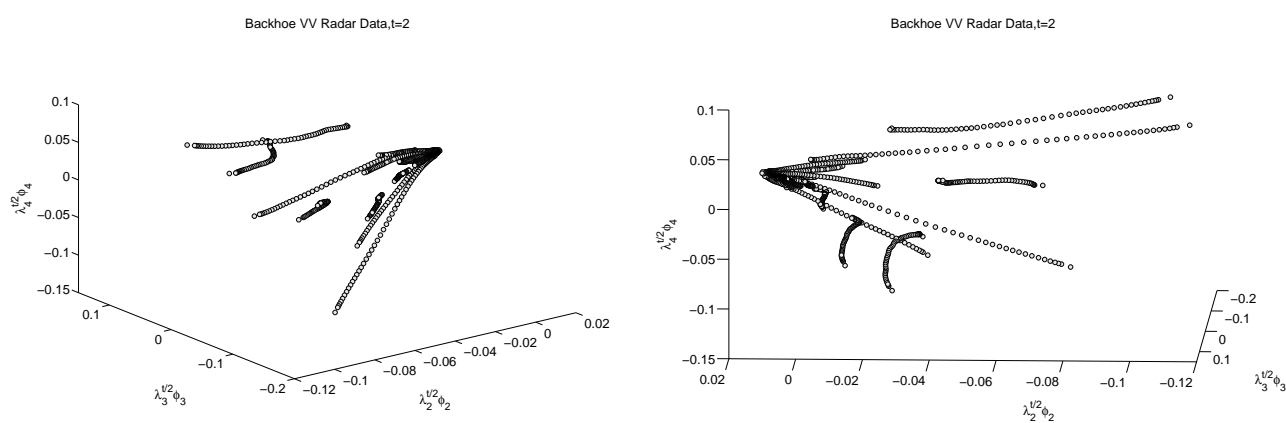**Figure 5.** The diffusion embedding that results when we use $\epsilon \approx 7.7$ million.

Figure 6. The diffusion embedding that results when we use $\epsilon \approx 2.7$ million.
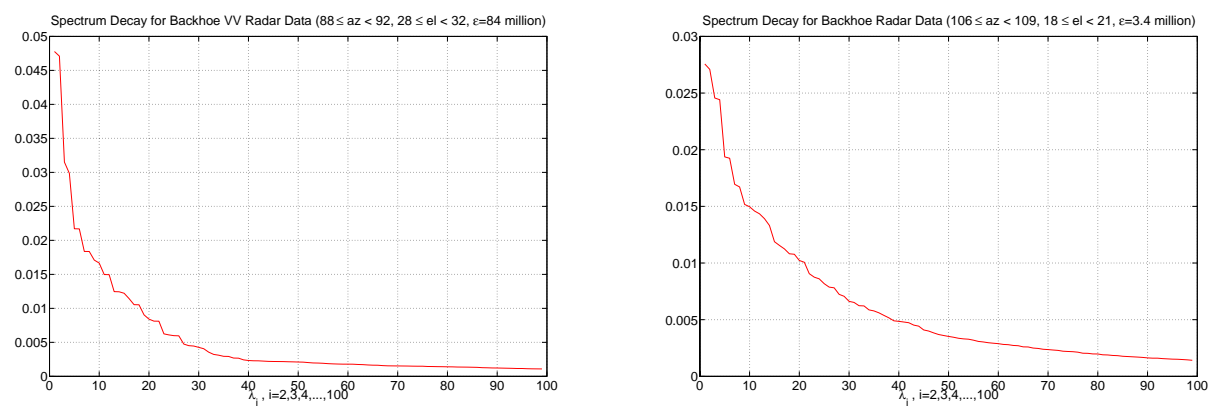
Figure 7. The spectrum decay of diffusion eigenvalues corresponding to the first (left) and second (right) data sets used for $\epsilon = 84$ million and $\epsilon = 3.4$ million respectively.
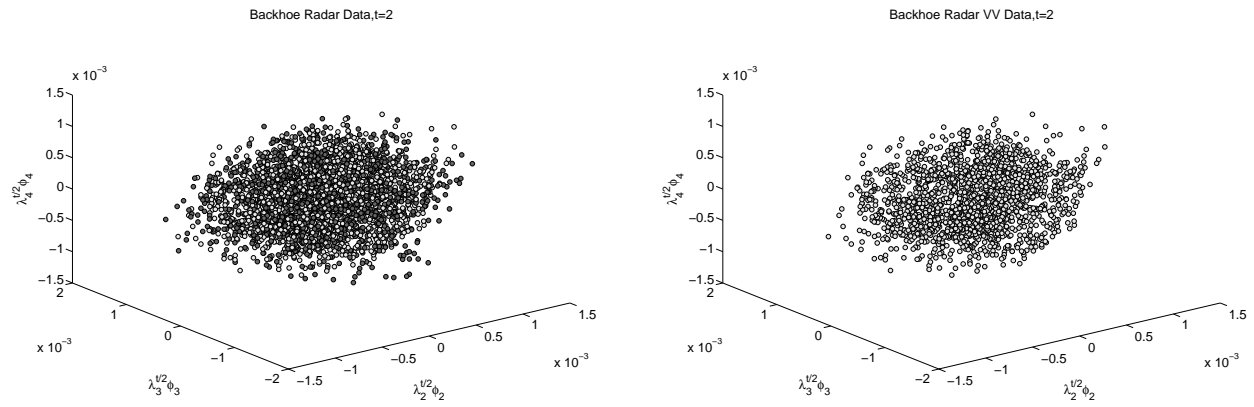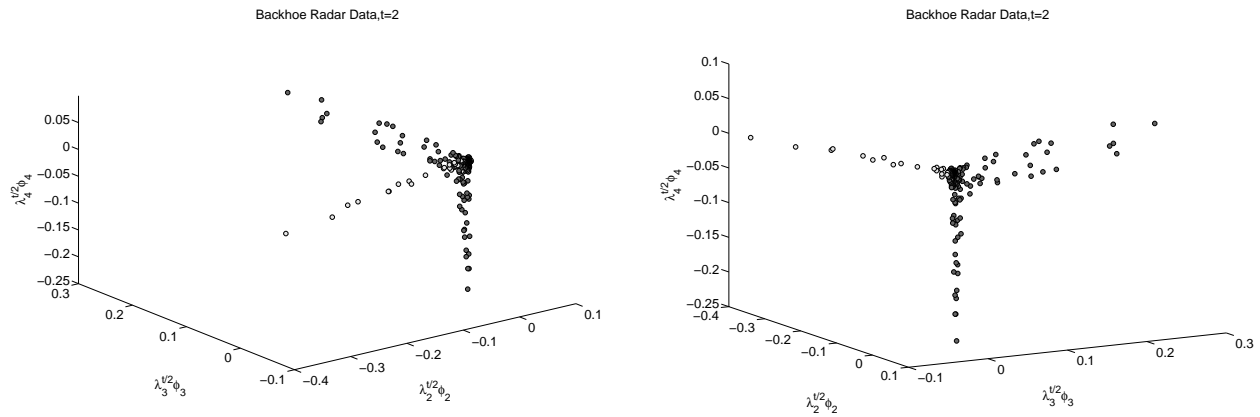
Figure 8. *Left:* Embedding that results when we use $\epsilon \approx 3.4$ million. *Right:* Portion of embedding due to VV polarization.

Figure 9. The diffusion embedding that results when we use $\epsilon \approx 350$ thousand.
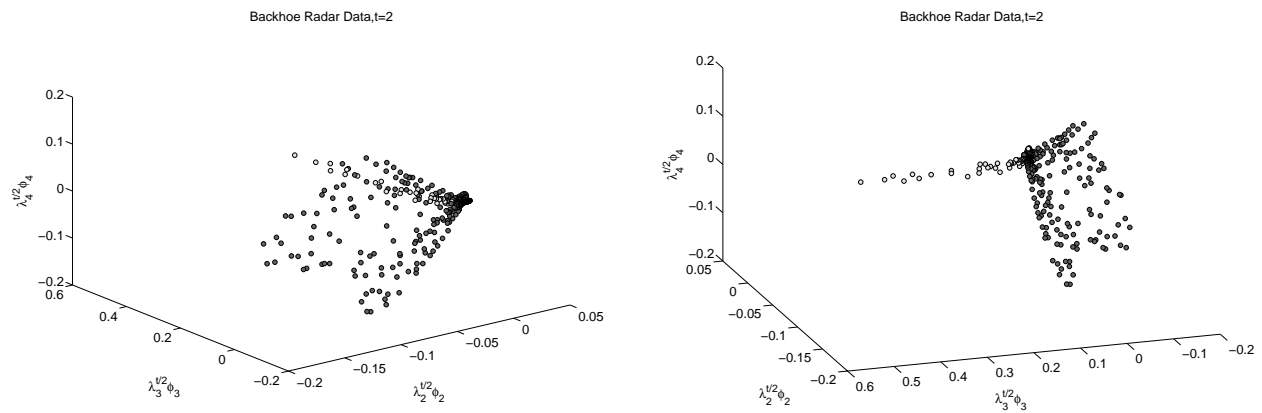

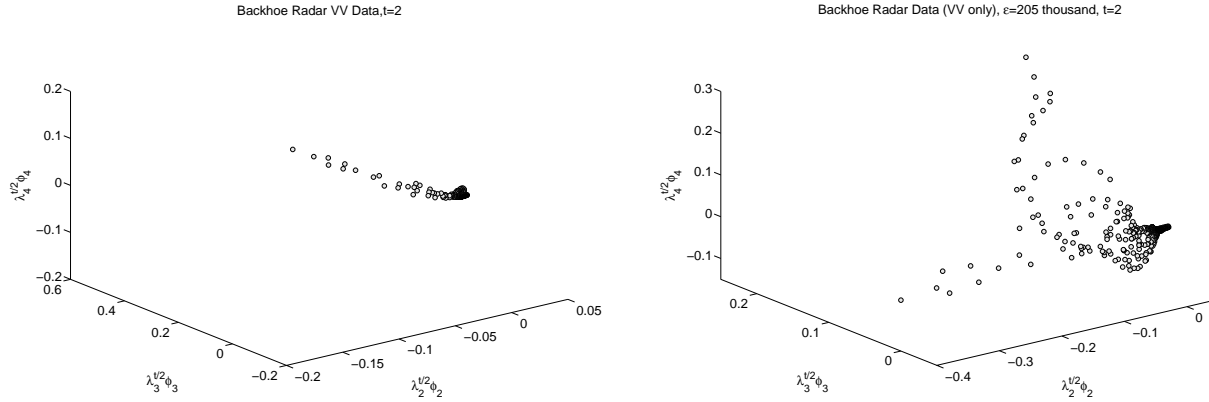**Figure 10.** The diffusion embedding that results when we use $\epsilon \approx 205$ thousand.

Figure 11. *Left:* Portion of Figure 10 due to VV data. *Right:* Embedding due to VV data only, $\epsilon = 205$ thousand.

3.2. Squiggle Path Data

Now we consider a manifold representation of the squiggle path. The (az, el) pairs of a randomly generated path are also available on the Backhoe Data Dome, Version 1.0. Interpolation is not needed since the (az, el) coordinates of the path coincide with the given (az, el) pairs. The original squiggle path consists of 3082 points however we use the horizontal and vertical polarization data of the first 1500 points (see the right side of Figure 15.) Thus our data set in this section contains 3000 points that correspond to the vertical and horizontal polarization range histories of each point on the squiggle path. Using Algorithm 1 we represent the squiggle path in terms of diffusion coordinates. As before, we plot the vertical polarization data as darker shaded point and the horizontal polarization data as a lighter shaded point. In Figures 12 and 13 we use (7) with $t = 2$. In Figure 14 on the left we see the manifold that results if we use

$$\epsilon = \frac{1}{k} \sum_{i,j=1}^{N} \|x_i - x_j\|_{L^2}^2, \tag{8}$$

where k is the cardinality of the set $\{\|x_i - x_j\|_{L^2} \big| \|x_i - x_j\|_{L^2} > 0, \ i, j = 1, 2, \ldots, N\}$. In Figure 14 on the right we see the manifold that results when we use (5) with $t = 2$. In Figure 15, on the left, we graph the spectrum decay associated to the diffusion coordinates for $\epsilon = 28$ million.
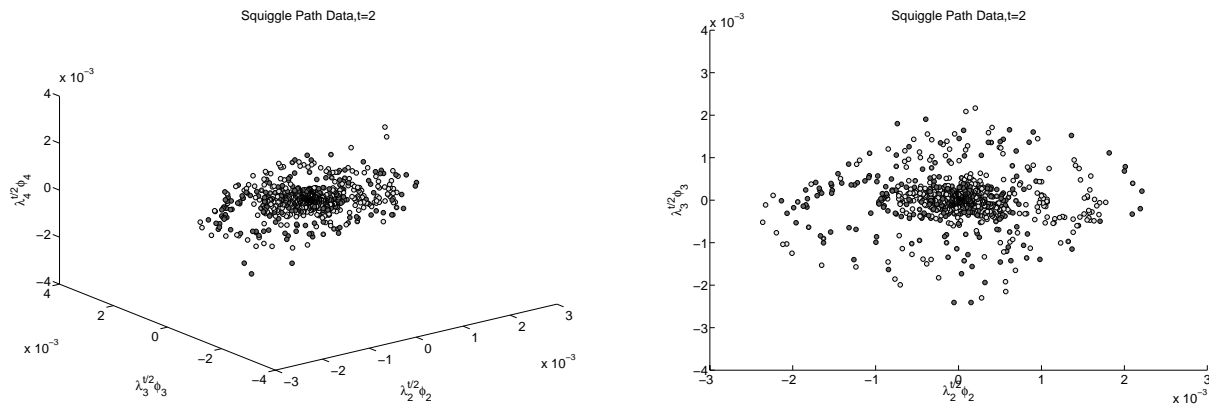

Figure 12. *Left:* The diffusion embedding that results when we use $\epsilon \approx 28$ million. *Right:* Projection of the diffusion embedding in two dimensions.
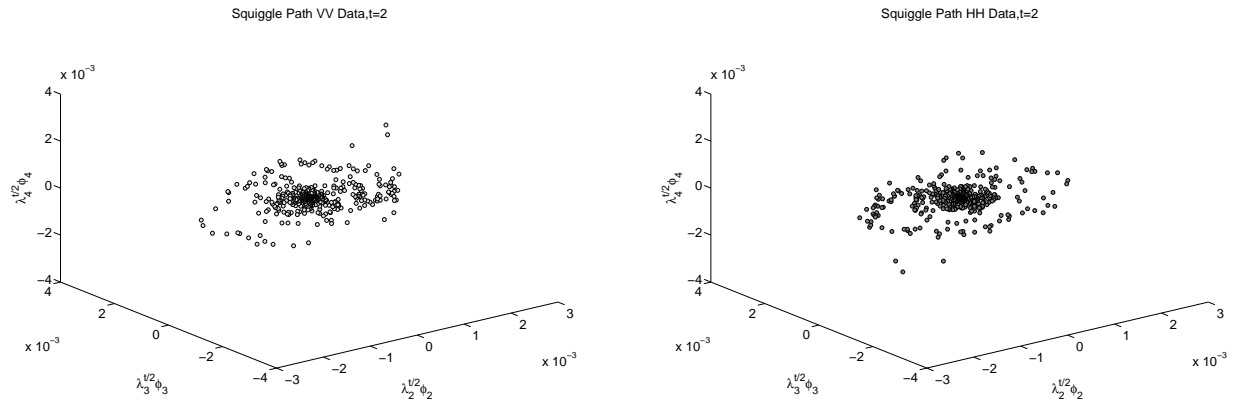
Figure 13. *Left:* Portion of diffusion embedding due to VV polarization. *Right:* Portion of diffusion embedding due to HH polarization.
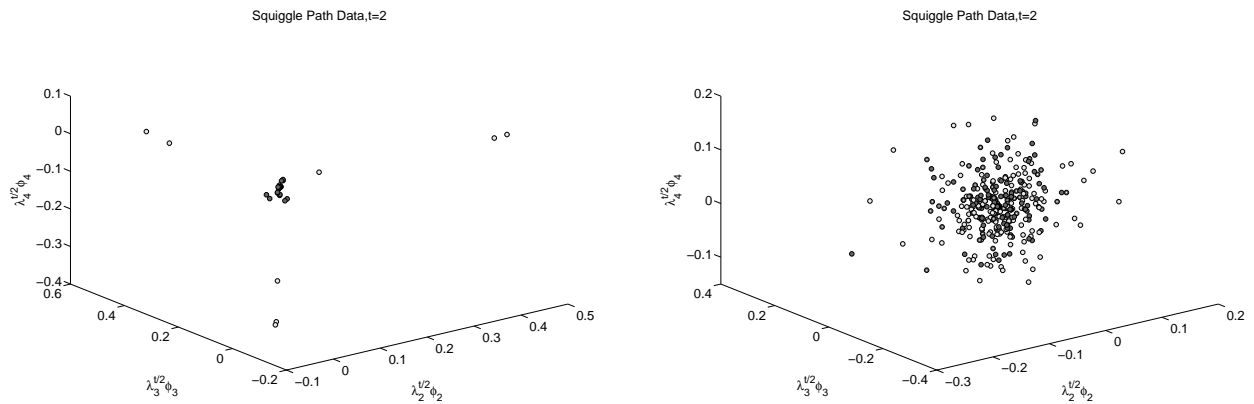

Figure 14. *Left:* Squiggle path diffusion embedding, $\epsilon \approx 9.1$ million. *Right:* The diffusion embedding that results when we use $\epsilon \approx 217$ thousand.
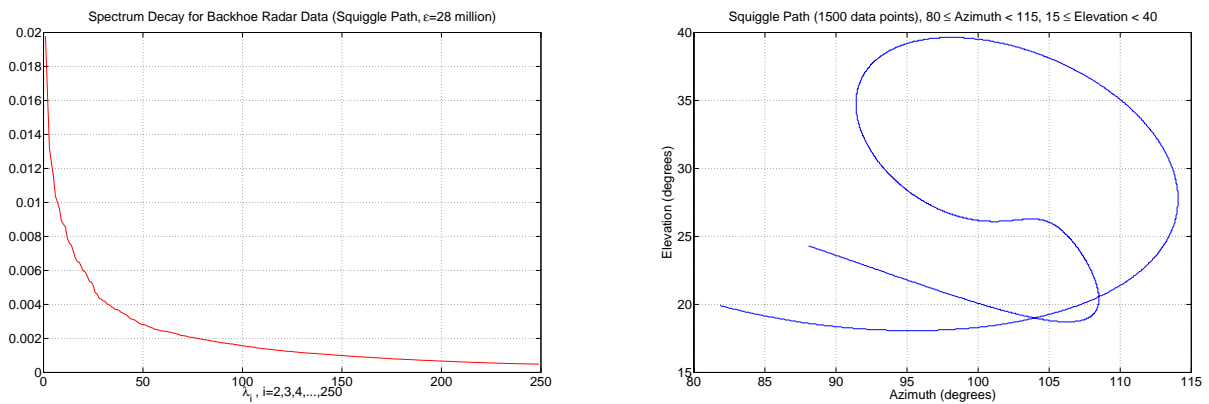

Figure 15. *Left:* Spectrum decay for diffusion coordinates when $\epsilon \approx 28$ million. *Right:* The geo-spatial plot of the squiggle path used here.

4. CONCLUSIONS

From the numerical experiments conducted we can conclude that there are certain "threshold" values for the parameter ϵ which drastically affect the shape and topological properties of the manifold. For example compare Figures 1, 5 and 6. The distinct embeddings are directly due to the ϵ value. When ϵ is large, close data points (with respect to the distance metric on the data set) will be identified with each other. As ϵ decreases there will be more distinction between close data points. Finding the appropriate ϵ may be dependent upon how much local information is needed. Alternately, Perona and Zelnik-Manor[12] propose the notion of "local scaling", *i.e.* they sugest the following weight function,

$$w(x_i, x_j) = \exp(-d^2(x_i, x_j)/\epsilon_i\epsilon_j)$$

where $d(\cdot, \cdot)$ is some distance metric and $\epsilon_i = d(x_i, x_K)$ for some fixed $K \in \{1, 2, \ldots, N\}$. Taking this idea further, we suggest

$$\epsilon_i = \min\{\|x_i - x_j\|_{L^2} \big| \|x_i - x_j\|_{L^2} > 0, \ j = 1, 2, \ldots, N\} \tag{9}$$

in the hope that we can construct the manifold in a local, piecewise fashion. For example, the resulting diffusion embedding of VV data for $88° \leq \text{az} < 92°$ and $28° \leq \text{el} < 32°$ when we use (9) is shown in Figure 16. The colors that appear in Figure 16 are in reference to the geo-spatial color scheme that appears in Figure 3 on the left. Further experiments utilizing local scaling would be the topic of future research.
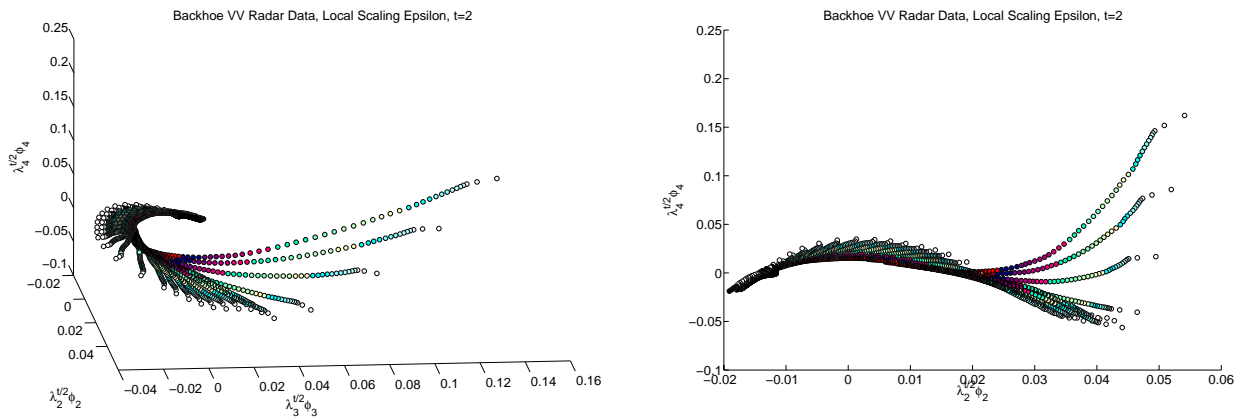

Figure 16. The diffusion embedding that results when we use a local scaling ϵ. Colors refer to Figure 3.

Note that from our results in Section 3 we see that our sampling of the Backhoe Data Dome seems to have some structure in diffusion coordinates. This is especially evident in Figures 1 and 2. However, we have not determined the reason for this apparent structure in the data and exploiting and exploring this structural information would be the topic of future work. The squiggle path data also seems to have some type of geometric structure in diffusion coordinates as can be seen in Figures 12. However, the relevance of this geometric structure to our analysis of the squiggle path data is key work that is the topic of future research. Also note that the vertical and horizontal polarization data play a key and distinct role in the embedding of data by diffusion coordinates as evidenced by Figures 8, 10 and 13. Again the exact relevance of the apparent structure in the diffusion embeddings of the vertical and horizontal polarization data need to be studied further. Note that we have not used any cross polarization data in our experiments. The effect of cross polarization data on the geometry of the manifold remains to be seen. In Figures 3 and 4 we clearly see that the geo-spatial location of the data is reflected in the structure of the resulting diffusion embedding. However, the relevance of the physical location to the manifold structure remains to be determined. In Figures 7 and 15(left) we see that the decay of the eigenvalues is quite rapid and numerically we require much less than 512 components to adequately capture the diffusion embedding of radar data. Note that in our experiments we have used $t = 2$ however increasing this value will cause the spectrum to decay at an even greater rate.

In general, the choice of weight function should be chosen to suit the intended application. Here, we approximate the heat kernel with the Gaussian and use a discrete L^2 metric. However, there may exist more appropriate approximations and metrics that may be used for radar data.

One can argue that the diffusion coordinate representation has organized the data with some type of geometry. Perhaps this geometry can be used to perform general ATR applications. However, at this point the key work lies in determining the relevance and connection between the underlying range histories and the apparent structure displayed by the data in diffusion coordinates.

ACKNOWLEDGMENTS

REFERENCES

1. S. Lafon, *Diffusion maps and geometric harmonics*, Ph. D. dissertation, Yale University, 2004.
2. R. R. Coifman, Y. Keller and S. Lafon, "Data Fusion and Multicue Data Matching by Diffusion Maps", *IEEE Trans. Pattern Anal. Mach. Intell.*, **Vol. 28**, no. 11, pp. 1784–1797, 2006.
3. R. R. Coifman, Y. Keller, S. Lafon, M. Maggioni, A. D. Szlam, F. J. Warner, and S. W. Zucker, "Geometries of sensor outputs, inference and information processing", in *Proc. SPIE*, edited by Intelligent Integrated Microsystems; Ravindra A. Athale, John C. Zolper, Eds., **Vol. 6232**, p. 623209, 2006.
4. K. Naidu and L. Lin, "Data Dome: Full k-Space sampling data for high frequency radar research", *Proceedings of SPIE, Algorithms for synthetic aperture radar imagery XI*, **Vol. 5427**, Edmund G. Zelnio, Frederick D. Garber, Eds., SPIE, Bellingham, WA, pp. 200–207, 2004.
5. M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation", *Neural Computation*, **Vol. 6**, no. 15, pp. 1373–1396, 2003.
6. F. Chung, *Spectral Graph Theory*, CBMS-AMS, Providence, R.I., no. 92, 1997.
7. B. Bollobás, *Graph Theory: An Introductory Course*, Springer-Verlag, New York, N.Y., 1979.
8. B. Schölkopf and A. J. Smola, *Learning with Kernels*, MIT Press, Cambridge, Massachusetts, 2002.
9. R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps", *Proceedings of the National Academy of Sciences*, **Vol. 102**, no. 21, pp. 7426–7431, 2005.
10. B. Lévy, "Laplace-Beltrami Eigenfunctions: Towards an algorithm that 'understands' geometry", appears in *IEEE International Conference on Shape Modeling and Applications 2006*, Matsushima, Japan, p. 13, 2006.
11. B. W. Higgs, J. L. Solka, J. Weller, "Spectral embedding finds meaningful (relevant) structure in image and microarray data", *BMC Bioinformatics*, **vol. 7**, p. 74, 2006.
12. P. Perona and L. Zelnik-Manor, "Self-Tuning Spectral Clustering", *Advances in Neural Information Processing Systems 17 (NIPS'04)*, Lawrence K. Saul, Yair Weiss, Léon Bottou, Eds., MIT Press, Cambridge, MA, pp. 1601–1608, 2005.