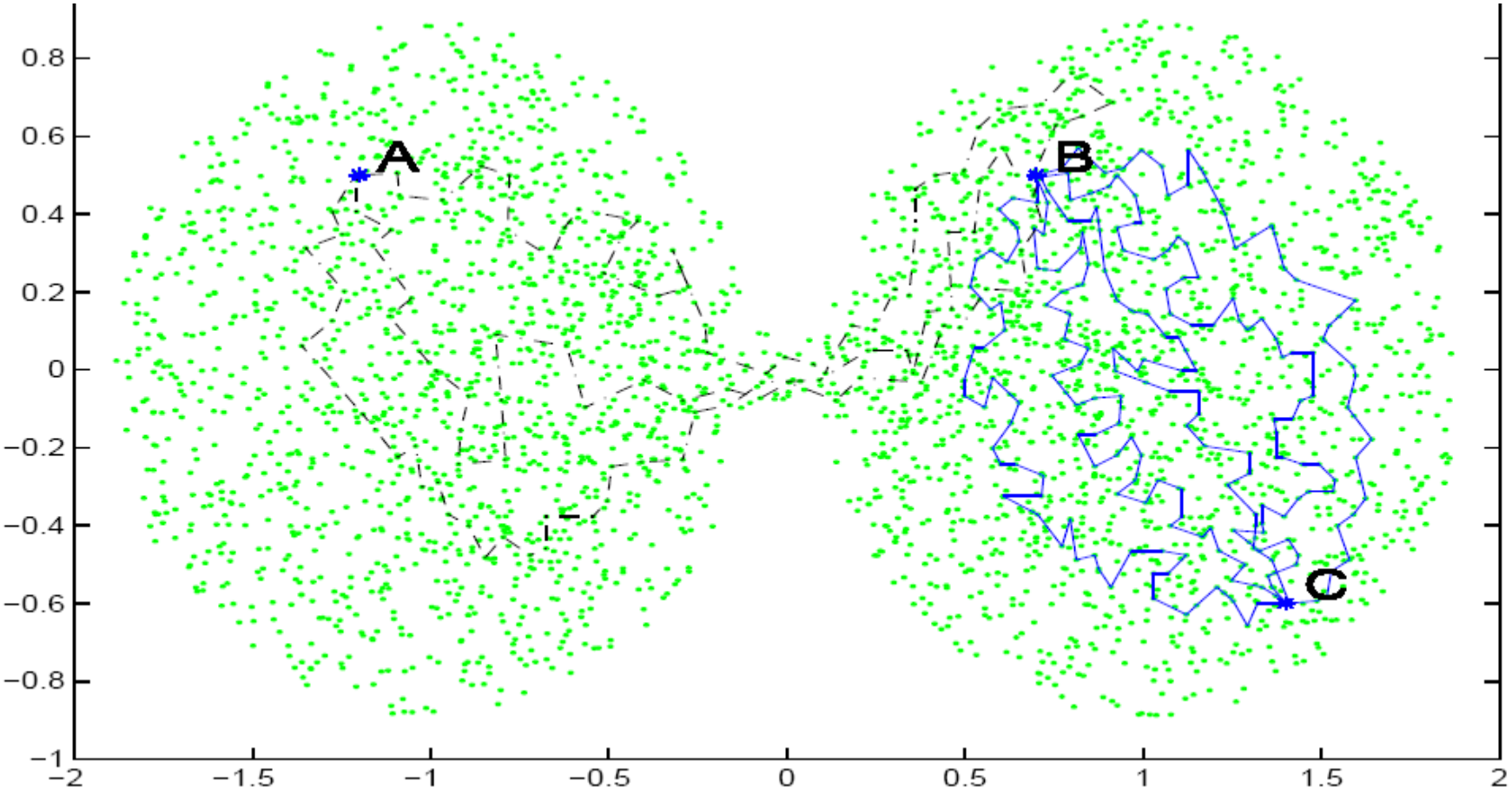
# Diffusion Geometries & Diffusion Wavelets

Extracted from the slides of
Ronald Coifman & Mauro Maggioni

December 19, 2008




Diffusions between A and B have to go through the bottleneck ,while C is easily reachable from B. The Markov matrix defining a diffusion could be given by a kernel , or by inference between neighboring nodes.

The diffusion distance accounts for preponderance of inference . The shortest path between A and C is roughly the same as between B and C . The diffusion distance however is larger since diffusion occurs through a bottleneck.

## *Diffusion Maps*

Let $\{\boldsymbol{x}_i\}_{i=1}^N$ denote a set of $N$ points in $\mathbb{R}^p$.

View collection of data as a graph with $N$ vertices and with connection strength between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ given by $k_\varepsilon(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where

$$k_\varepsilon(\boldsymbol{x}, \boldsymbol{y}) = \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{y}\|^2}{2\varepsilon}\right)$$

Construct a Markov chain random walk based on these weights:

$$M_{i,j} = \Pr\{\boldsymbol{x}(t+\varepsilon) = \boldsymbol{x}_i | \boldsymbol{x}(t) = \boldsymbol{x}_j\} = \frac{k_\varepsilon(\boldsymbol{x}_i, \boldsymbol{x}_j)}{p_\varepsilon(\boldsymbol{x}_j)}$$

where

$$p_\varepsilon(\boldsymbol{x}_j) = \sum_i k_\varepsilon(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

Claim: first few eigenvectors and eigenvalues of this matrix $\{\lambda_j, \phi_j\}$ contain useful geometric information.

## *Diffusion Maps / Normalized Graph Laplacian*

The diffusion map at time $t$ is defined as the non-linear embedding

$$\boldsymbol{x} \to \Phi_t(\boldsymbol{x}) = \left(\lambda_1^t \phi_1(\boldsymbol{x}), \lambda_2^t \phi_2(\boldsymbol{x}), ..., \lambda_k^t \phi_k(\boldsymbol{x})\right)$$

## *Diffusion Distances*

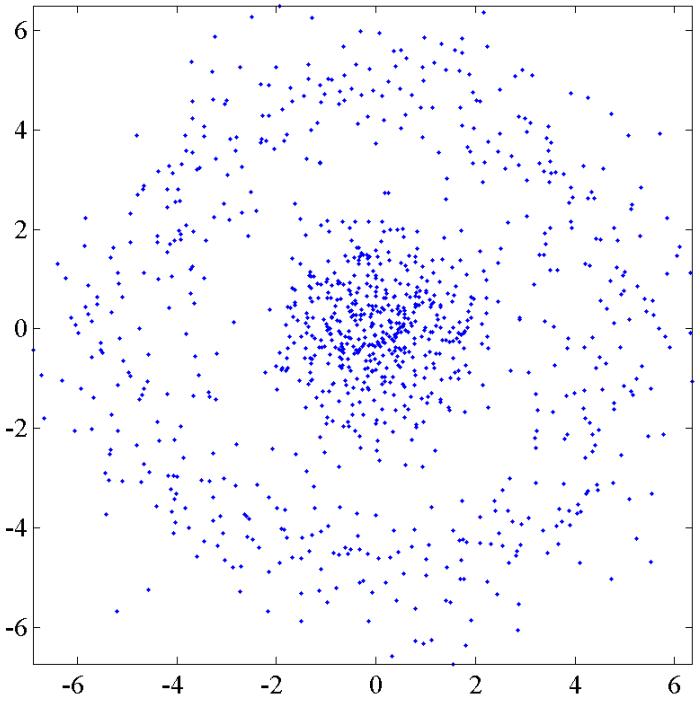
**Definition:** Diffusion distance at time $t$,

$$D_t^2(\boldsymbol{x}, \boldsymbol{y}) = M_s^t(\boldsymbol{x}, \boldsymbol{x}) + M_s^t(\boldsymbol{y}, \boldsymbol{y}) - 2M_s^t(\boldsymbol{x}, \boldsymbol{y})$$

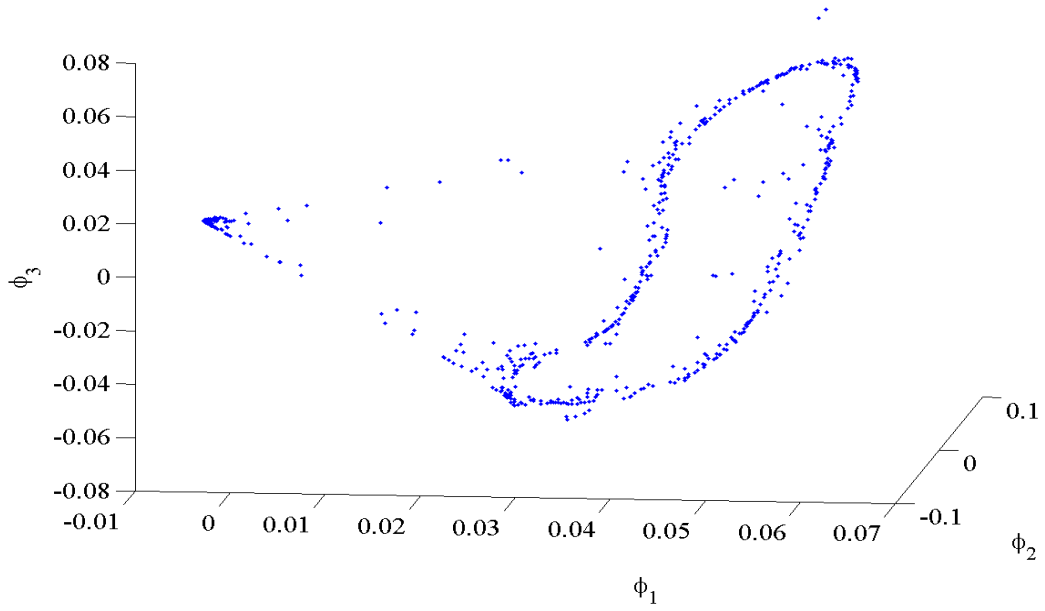
**Spectral theorem:**

$$D_t^2(\boldsymbol{x}, \boldsymbol{y}) = \sum_j \lambda_j^{2t} \left(\phi_j(\boldsymbol{x}) - \phi_j(\boldsymbol{y})\right)^2 = \|\Phi_t(\boldsymbol{x}) - \Phi_t(\boldsymbol{y})\|^2$$

Diffusion map converts the diffusion distance into Euclidian distance.

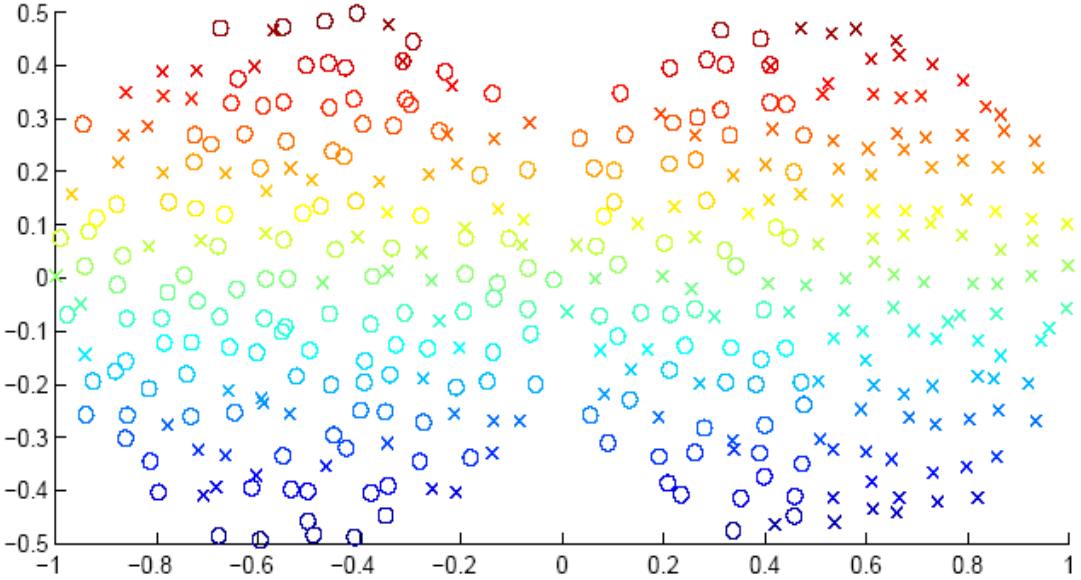


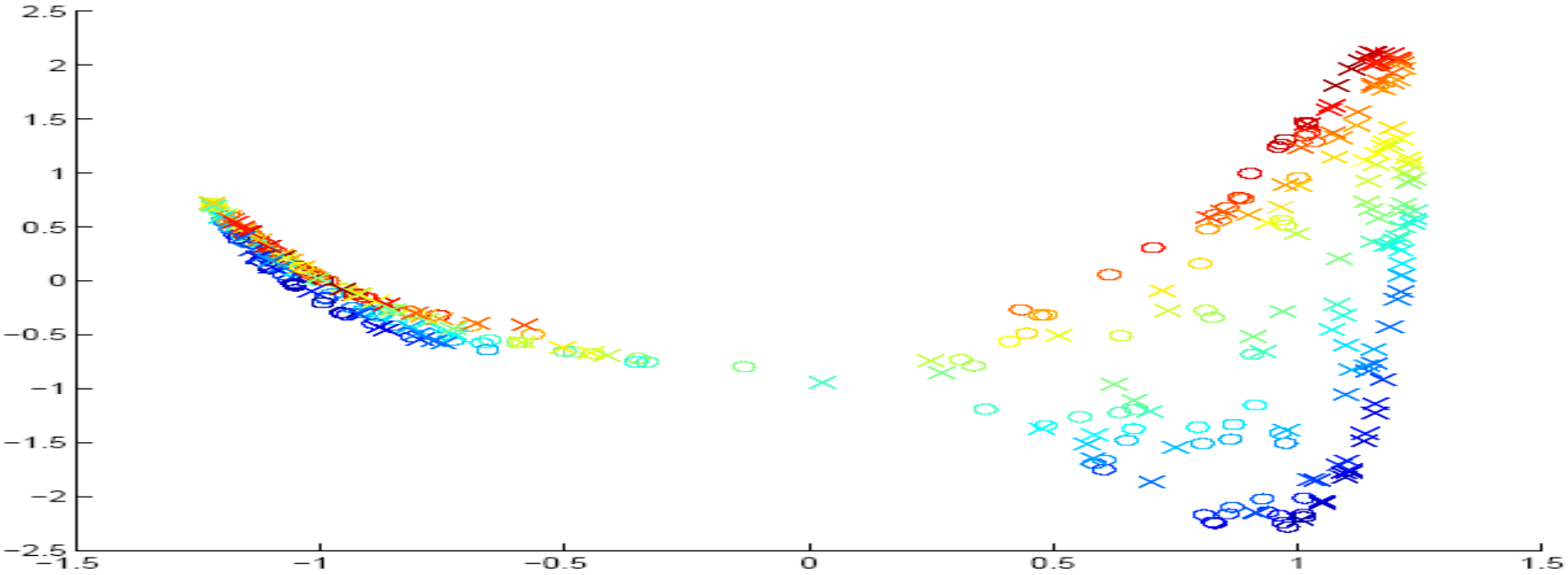

*Original data set*

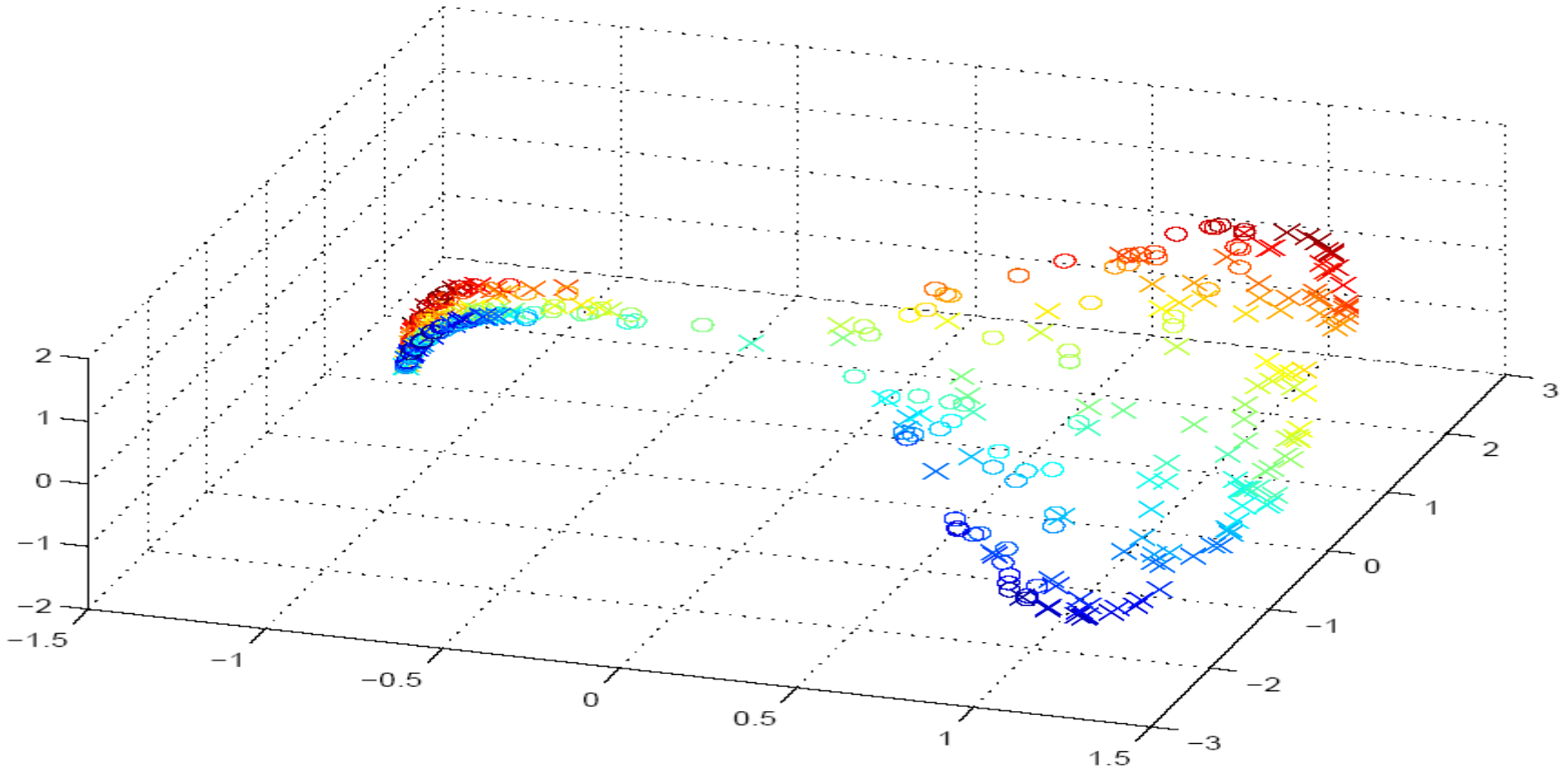



*Embedding of data into the first 3 diffusion coordinates*

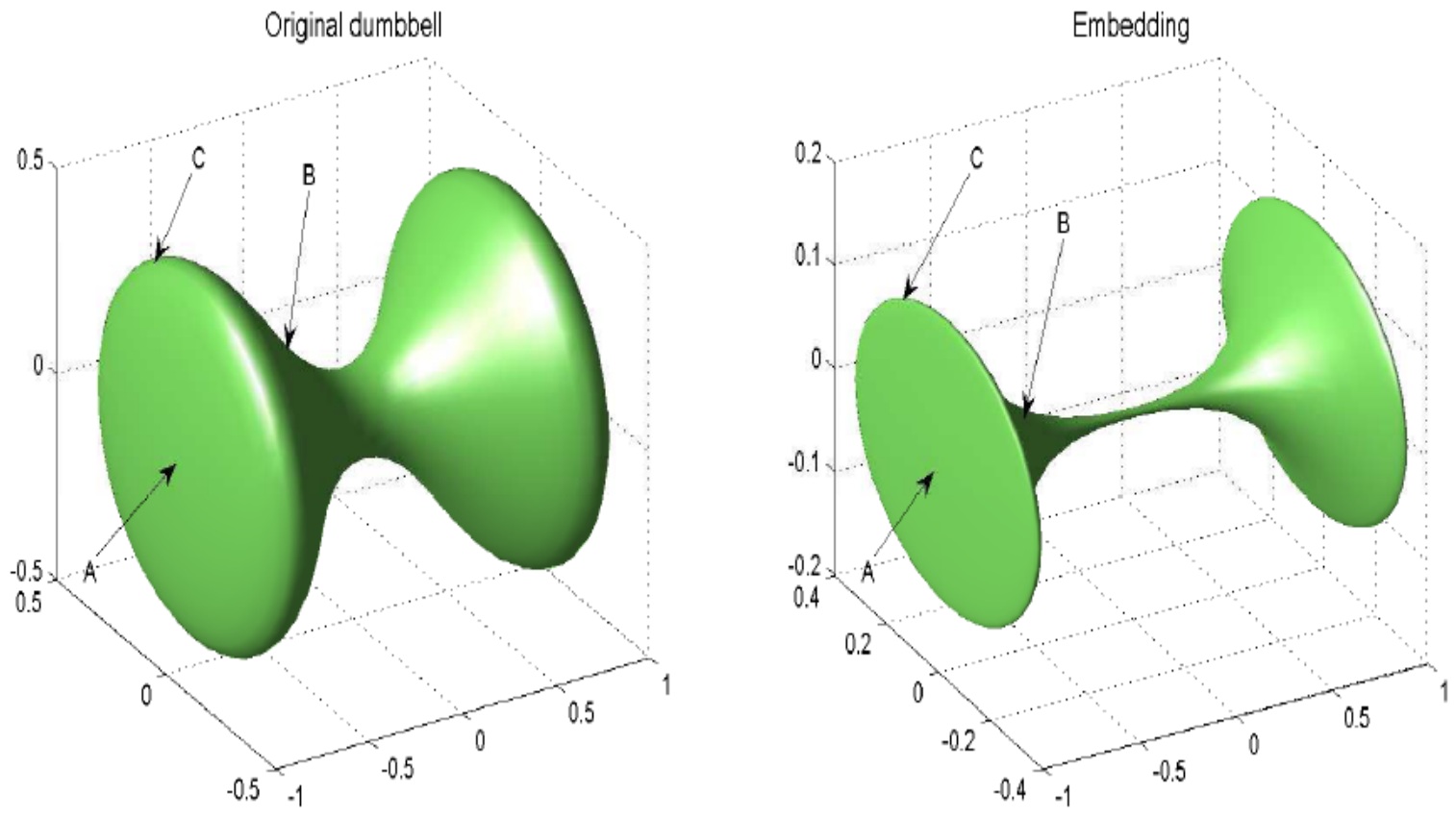
The long term diffusion of heterogeneous material is remapped below . The left side has a higher proportion of heat conducting material ,thereby reducing the diffusion distance among points , the bottle neck increases that distance






Diffusion map into 3 d of the heterogeneous graph
The distance between two points measures the diffusion between them.



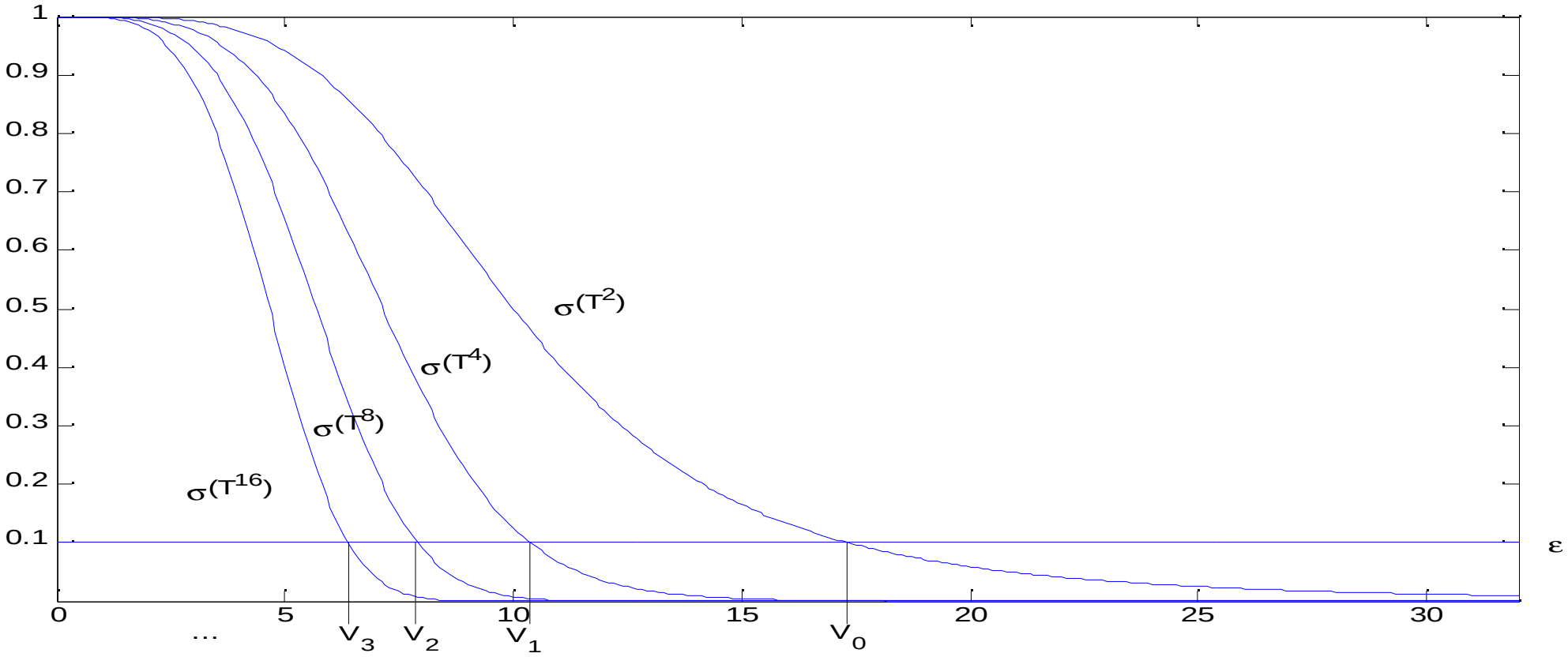

The natural diffusion on the surface of the dumbell is mapped out in the embedding . Observe that A is closer to B than to C ,and that the two lobes are well separated by the bottleneck.

# Diffusion Wavelets

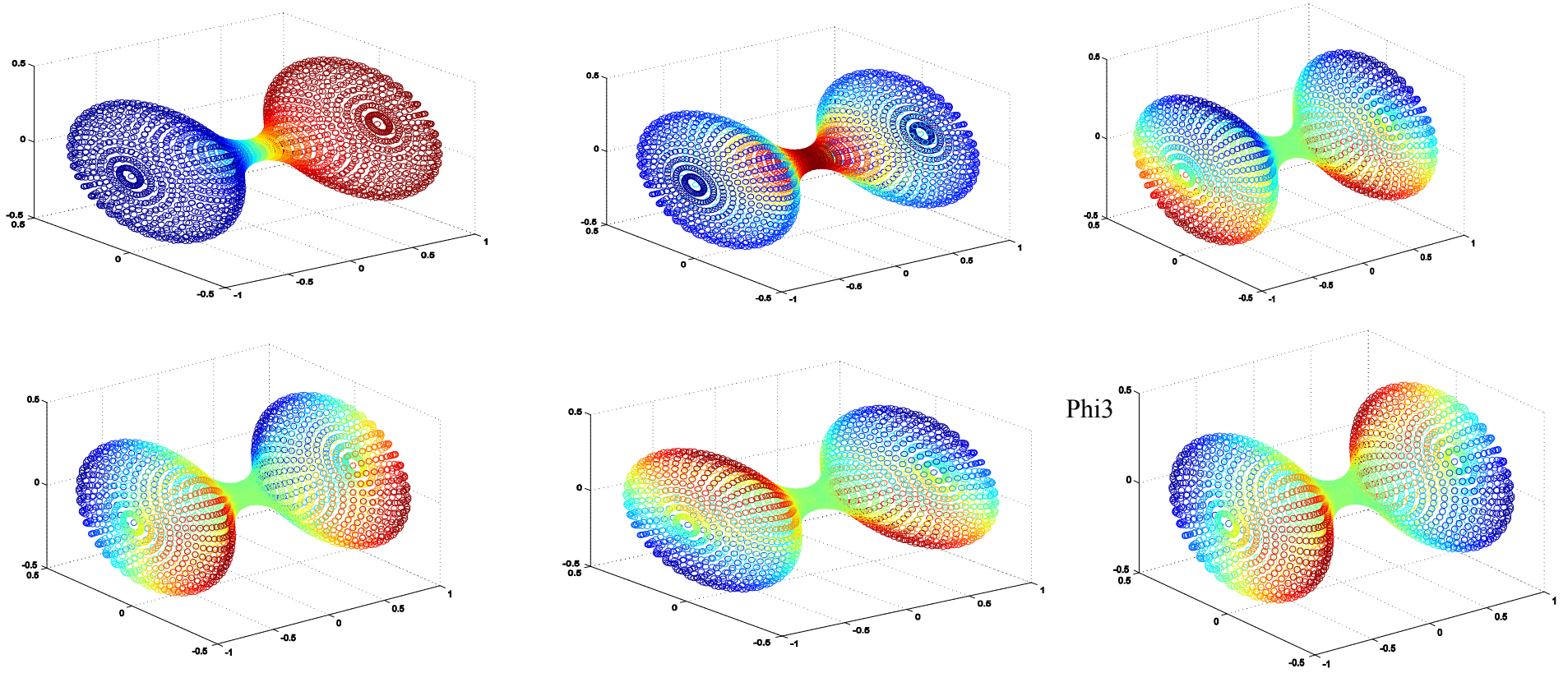
RR Coifman & MM

Eigenfunctions are like global Fourier Analysis on the data set, they live in different "frequency bands" but are not localized. We would like to have elements **localized** both in frequency and space (compatibly with Heisenberg principles), and **critically sampled** at the "rate" corresponding to the frequency band.

Where are the "frequencies"?



Eigenfunctions of the Laplacian are like global Fourier modes.




Drawbacks of global Fourier analysis for function approximation, are well-known even in 1D!

What about building **multiscale localized bases**, or wavelets?

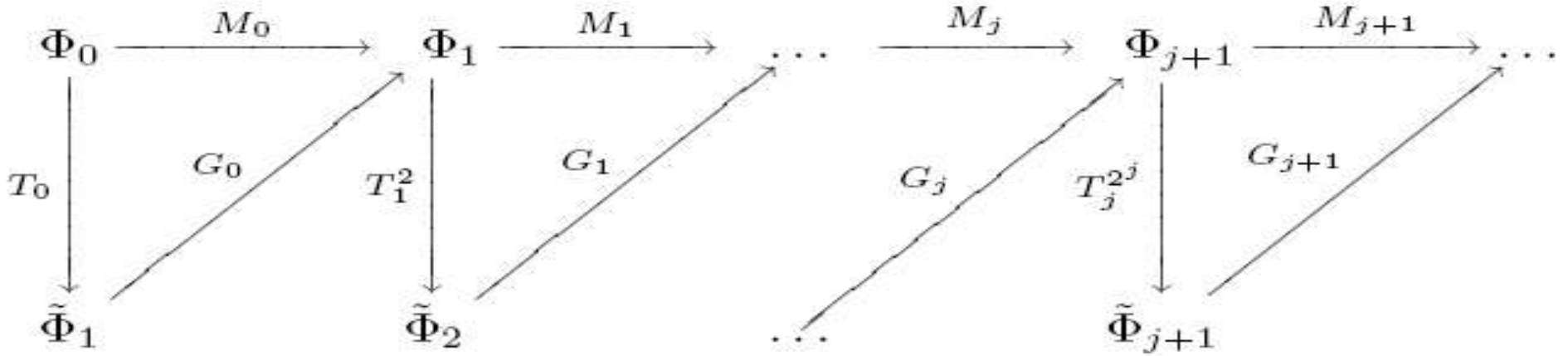
# Dilations, translations, downsampling

We have **frequencies**: the eigenvalues of the diffusion T.
What about dilations , translations , downsampling?
The geometry seems hard to handle, let's think in terms of functions on the set X.

Dilations:
Use the diffusion operator T and its dyadic powers as dilations.

Translations and downsampling:
Idea: diffusing a basis of “scaling functions” at a certain scale by a power of T should yield a redundant set of coarser “scaling functions” at the next coarser scale: reduce this set to a Riesz (i.e. well-conditioned)-basis. This is downsampling in the function space, and corresponds to finding a well-conditioned subset (downsampling) of “translates”.

# Construction of diffusion multiresolution




$$T_j^{2^j} = \left(\Phi_j T_{j-1}^{2^{j-1}} \Phi_j^T\right)^2 = M_{j-1} M_{j-1}^T \,.$$

$$\varphi_{j,l,k}(x) = M_{j-1}\varphi_{j-1,l,k}(x) \quad , x \in X_{j-1}$$

$$= M_{j-1}M_{j-2} \cdot \ldots \cdot M_0\, \varphi_{0,l,k}(x) = \left(\prod_{l=0}^{j-1} M_l\right) \varphi_{0,l,k}(x) \quad , x \in X_0$$

# Generalized Fast Multipoles

In Fast Multipole Methods [Rohklin-Greengard, Beylkin-Coifman-Rohklin], the Green's function is given and is compressed. This is not readily available in non-homogeneous problems.

We started from the local, differential operator, and "discover" its multiscale, multipole expansion through its iterates, and compress "on the go". Similar to multigrid [Brandt, Hackbusch...], but keeping coarse equations exact at each scale.

We can then apply the Green's function in O(n log(n)) to any vector by Schultz's method:

$$(I-T)^{-1} f = \sum_{k=1}^{+\infty} T^k f$$

and, if $S_K = \sum_{k=1}^{2^K} T^k$, we have

$$S_{K+1} = S_K + T^{2^K} S_K = \prod_{k=0}^{K} \left( I + T^{2^k} \right) f \, .$$

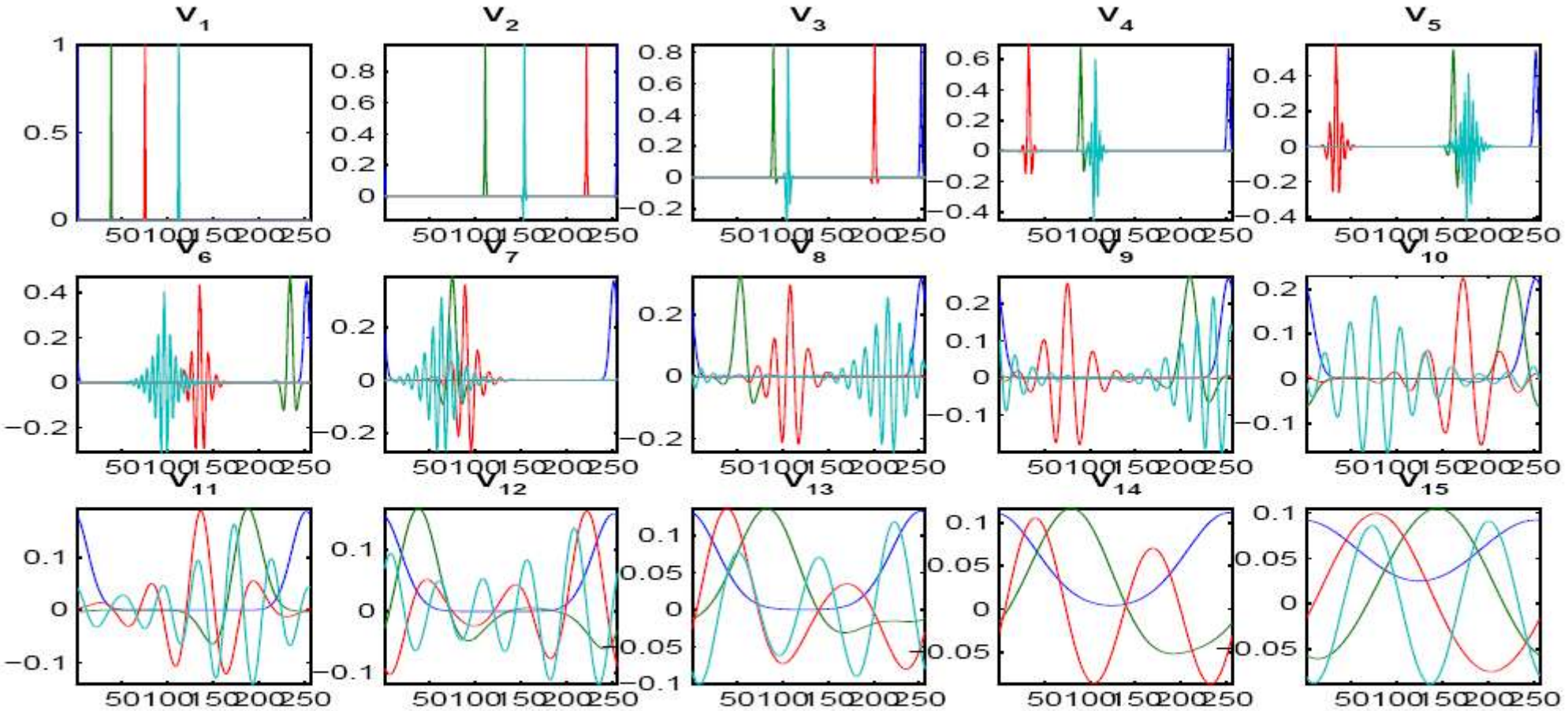



Fig. 3. Multiresolution Analysis on the circle. We consider 256 points on the unit circle, start with $\varphi_{0,k} = \delta_k$ and with the standard diffusion. We plot several scaling functions in each approximation space $V_j$.

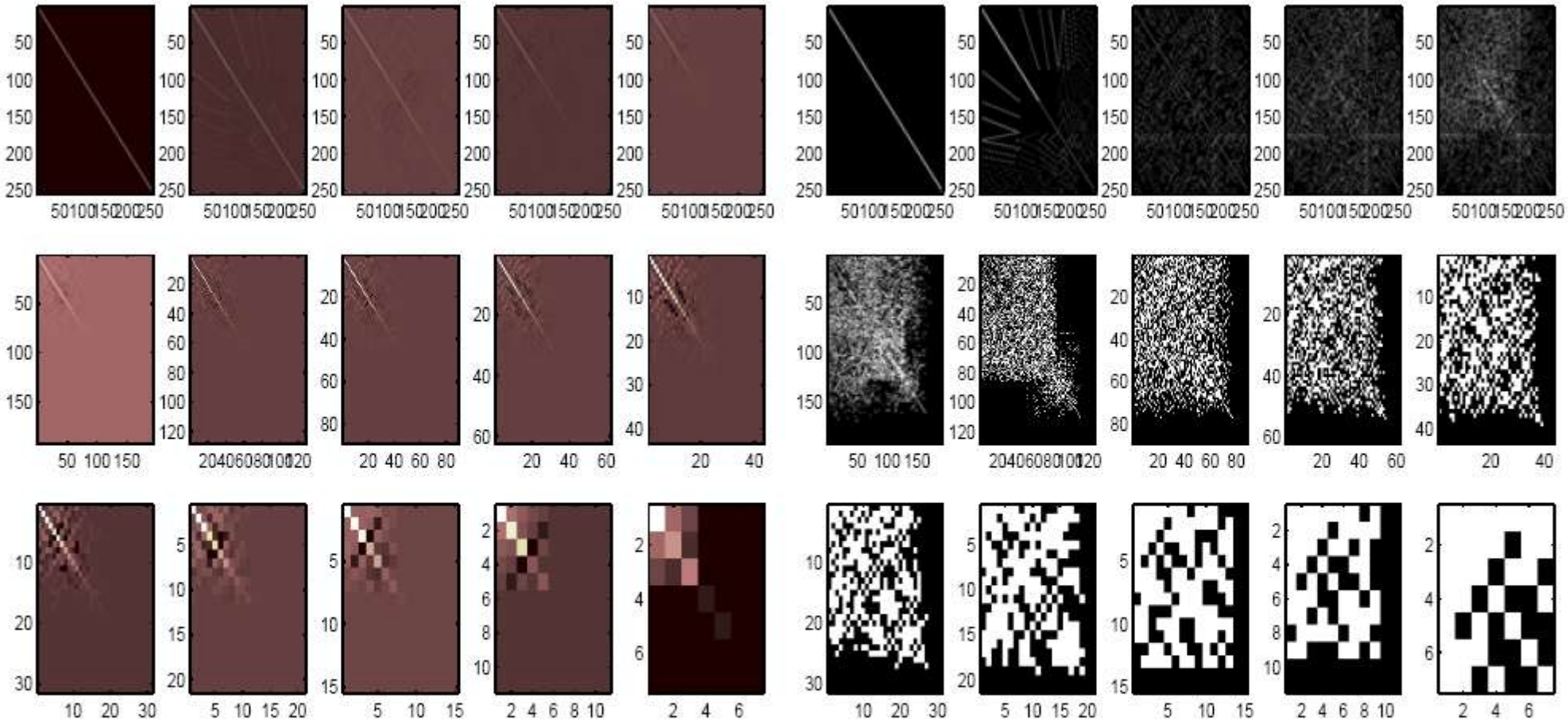


Fig. 4. Multiresolution Analysis on the circle: on the left we plot the compressed matrices representing powers of the diffusion operator, on the right we plot the entries of the same matrices which are above working precision.

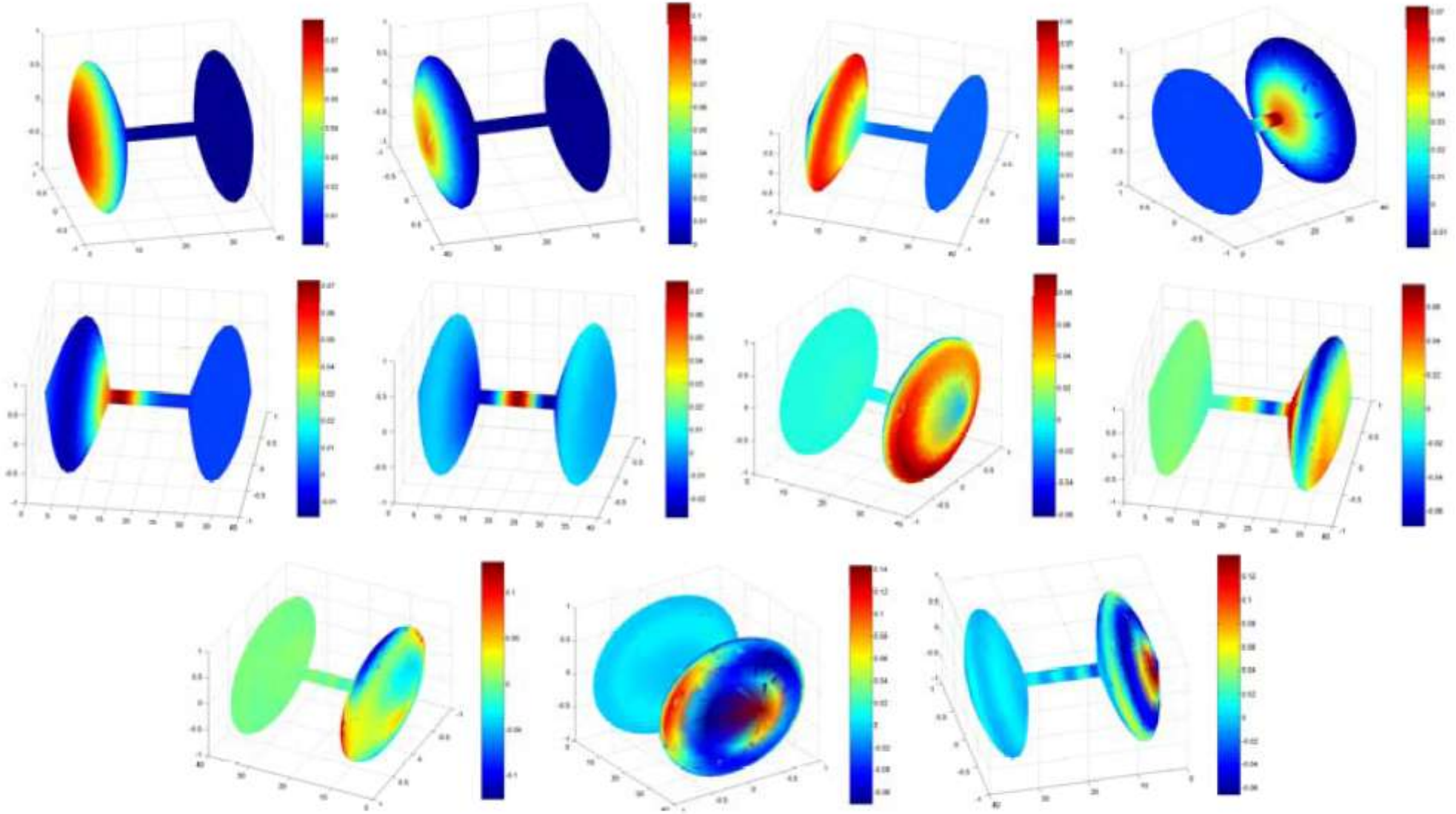


Fig. 8. Some diffusion scaling functions and wavelets at different scales on a dumb-bell-shaped manifold sampled at 1400 points.