# MANIFOLD DENOISING BASED ON SPECTRAL GRAPH WAVELETS

Shay Deutsch<sup>†</sup>, Antonio Ortega<sup>‡</sup> and Gérard Medioni<sup>†</sup>

<sup>†</sup> Department of Computer Science, <sup>‡</sup> Department of Electrical Engineering University of Southern California

# ABSTRACT

We propose a new framework for manifold denoising using the Spectral Graph Wavelet transform, which enables non-iterative denoising directly in the graph frequency domain, an approach inspired by conventional wavelet-based signal denoising methods. We theoretically justify our approach, based on the fact that for smooth manifolds the coordinate information tends to create energy in the low spectral graph wavelet coefficients, while the noise affects all frequency bands in a similar way. Experimental results show that our suggested manifold frequency denoising (MFD) approach significantly outperforms the state of the art manifold denosing methods, and is robust to a wide range of parameter selections, e.g., the choice of k nearest neighbor connectivity of the graph.

Index Terms— Manifold Learning, Denoising, Graph Signal Processing

# 1. INTRODUCTION

Manifold learning has been proposed to extend linear approaches such as PCA to the more general case where data lies on a nonlinear manifold embedded in a low-dimensional space. When the data lies strictly on the manifold, manifold learning techniques such as Isomap [1], locally linear embedding (LLE) [2], Laplacian eigenmaps (LE) [3], or local tangent space alignment (LTSA) [4], can provide effective tools to analyze high dimensional data with complex structure. However, in the presence of noise, i.e., when the observed data does not lie exactly on the manifold, the performance of these methods degrades significantly. Only a handful of methods have been suggested to handle noisy manifolds, e.g., [5], [6], and these tend to over-penalize either the local or the global structure of the manifold.

In this paper, we address the manifold denoising problem by proposing a new graph-frequency framework called Manifold Frequency Denoising (MFD). Our approach uses Spectral Graph Wavelets (SGW) [7], which, similar to wavelets for signals in regular domains, provide a trade-off between spectral and vertex domain localization. This allows us to overcome the limitations of existing manifold denoising methods by taking advantage of global smoothness characteristics (energy concentrated in lower frequency SGW bands) without over-smoothing at discontinuities (which correspond to large magnitude coefficients in the high frequency SGW bands). In our proposed framework we build a graph where each vertex corresponds to one of the noisy observations of the manifold, with edge weights between two vertices a function of the distance between the corresponding observations in the ambient space. Then, we apply the SGW to coordinate graph-signals, one per dimension, where in a coordinate graph signal each vertex is assigned the scalar coordinate of the vertex in the corresponding dimension. Thus, our graph is based on vector distances between observations, while denoising is applied to the observations (coordinates) in each dimension. In this paper, we theoretically justify our approach by showing that for smooth manifolds the coordinate signals also exhibit smoothness (i.e., the maximum variation across neighboring nodes is bounded by a term that decreases as the manifold smoothness increases). Our experimental study demonstrates that graph signal processing methods are effective for processing smooth manifolds, since in a graph signal defined based on these manifolds, most of the energy is concentrated in the low frequencies, making it easier to separate noise from signal information. While there has been a significant amount of recent interest in graph signal processing [8], to the best of our knowledge MFD is the first attempt to use these tools for manifold denoising. Our work is different from image denoising based on graph signal processing [9] in that our smoothness prior explicitly assumes that the original data lies on a smooth manifold. Another crucial aspect in manifold denoising is the efficiency and robustness of the process. Most of the current manifold denoising algorithms consist of iterative, global or semi global operations, which may also be sensitive to the parameter selection. In contrast, our denoising approach provides a fast and non-iterative process, with low computational complexity that scales linearly in the number of points for sparse data, is robust for a large range of parameter selections, and in particular, selection of k, the number of nearest neighbors used to construct the graph. In addition, our approach does not require knowledge of the intrinsic dimensionality of the manifold. Experimental results demonstrate that our framework significantly outperforms the state of the art, making it possible to use denoising as a pre-processing step before applying current manifold learning approaches. The paper is organized as follows: in Section 2 we summarize the related work. In Section 3 we introduce the notation and provide an overview of spectral graph wavelets. Section 4 presents our theoretical justification, and Section describes our proposed approach. The experimental results are provided in Section 6, and in Section 7 we conclude our work and suggest future work.

## 2. RELATED WORK

Current unsupervised state of the art manifold denoising methods include manifold denoising (MD) [5] and LLD [6]. Also related are statistical modeling approaches for manifold learning such as Probabilistic non-linear PCA with Gaussian process latent variable models (GP-LVM) [10] and its variants for manifold denoising [11]. The method most related to our work is MD, which develops an algorithm that applies a diffusion process on the graph Laplacian, using an iterative procedure that solves differential equations on the graph. It can be shown that each iteration in the diffusion process in MD is equivalent to the solution of the regularization problem on the graph [5] (the regularization problem on the graph solved by MD is also known as Tikhonov regularization). The main limitation of MD is over-smoothing of the data and the sensitivity to the choice of k nearest neighbor construction graph (as mentioned in [5]) especially at high noise levels. In contrast, our suggested framework for denoising can be viewed as a multi-resolution approach which exploit Spectral Wavelets to find a trade-off between localization in the frequency and vertex domains. Our work is inspired by a classical approach for denoising [12], and its many extensions to image denoising [13]. However, while denoising signals that lie on regular meshes using wavelets is a well studied problem, the more general case of irregular domains is much less understood. Some recent work [9] has explored graph-based techniques for image denoising, while here we focus on the manifold denoising case. Spectral Graph Wavelets (SGW) [7] (see Section 3) provide us with an efficient tool to select spectral- and vertex-domain localization and are key component of our method.

## **3. PRELIMINARIES**

Consider a set of points  $\mathbf{x} = {\mathbf{x}_i}$ , i = 1, ...N,  $\mathbf{x}_i \in R^D$  which are sampled from an unknown manifold M. An undirected, weighted graph G = (V, E), constructed over  $\mathbf{x}$ , where V corresponds to the nodes and E to the set of edges on the graph. The adjacency matrix  $\mathbf{W} = (w_{ij})$  consists of the weights  $w_{i,j}$  between node i and node j. In this work, the weights are chosen using the Gaussian kernel function

$$W_{ij} = \begin{cases} \exp(||\mathbf{x}_i - \mathbf{x}_j||_2^2)/2\sigma_d^2 & \text{if } ||\mathbf{x}_i - \mathbf{x}_j||_2^2 \le \delta \\ 0 & \text{otherwise} \end{cases}$$
(1)

where (|| ||) denotes the L2 distance between the points  $\mathbf{x_i}, \mathbf{x_j}$ . In order to characterize the global smoothness of a function  $f \in \mathbb{R}^N$ , we define its Graph Laplacian quadratic form with respect to the graph as:

$$|| \nabla \mathbf{f} ||^2 = \sum_{i \sim j} w_{ij} [f(i) - f(j)]^2 = \mathbf{f}^T \mathbf{L} \mathbf{f}, \qquad (2)$$

where  $i \sim j$ , if *i* and *j* are connected on the graph by an edge, and **L** denotes the combinatorial graph Laplacian, defined as  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ , with **D** the diagonal degree matrix with entries  $d_{ii} = d(i)$ . The degree d(i) of vertex *i* is defined as the sum of weights of edges that are connected to *i*. The eigenvalues and eigenvectors of **L** are  $\lambda_1, \ldots, \lambda_N$  and  $\phi_1, \ldots, \phi_N$ , respectively. The Graph Fourier Transform (GFT)  $\hat{f}$  is defined as the expansion of *f* in terms of the eigenvectors  $\phi$  of the Graph Laplacian  $\hat{f}(\lambda_l) = \sum_i f(i)\phi_l(i)$ . Spectral graph Wavelets (SGWs)[7] define a scaling operator in the Graph Fourier domain, based on the eigenvectors of the graph Laplacian **L**, which can be thought as an analog of the Fourier transform for functions on weighted graphs. SGWs are constructed using a kernel function operator  $T_g = g(\mathbf{L})$  which acts on a graph signal **f** by modulating each graph Fourier mode:

$$\overline{T_g}f(l) = g(\lambda_l)\widehat{f}(l).$$
(3)

Scaling is defined in the spectral domain by the operator  $T_g^s = g(s\mathbf{L})$ , and the wavelets are calculated by applying  $T_g^s$  to a single vertex operator:  $\psi_s^n = T_g^s \delta$ . Given a function f, the wavelet coefficients take the form  $\Psi_f(s,n) = (T_g^s f(n)) = \sum_{l=1}^N g(s\lambda_l) \widehat{f}(l) \phi_l(n)$ . SGWs can be computed with a fast algorithm based on approximating the scaled generating kernels by low order polynomials. The wavelet coefficients at each scale can then be computed as a polynomial of  $\mathbf{L}$  applied to the input data. When the graph is sparse, which is typically the case under the manifold learning model, the

computational complexity scales linearly with the number of points, leading to a computational complexity of O(N).

#### 4. THEORETICAL JUSTIFICATION

We first introduce some notation and recall the definition of the condition number  $1/\tau$  [14] which provides an efficient measure to capture both the local and global geometric properties of a manifold, and the geodesic covering regularity on the manifold [15].

**Definition 1** The condition number of a manifold M is the largest number  $\rho$  such that each point in  $M \oplus \rho$  has a unique projection onto M, where  $M \oplus \rho = \bigcup_{\mathbf{x} \in M} T_{\mathbf{x}}^{\perp} M \cap B_D(\mathbf{x}, \rho), T_{\mathbf{x}}^{\perp} M$  is the normal space of a point  $\mathbf{x} \in M$ , and  $B_D(\mathbf{x}, \rho)$  is an open ball in  $R^D$  centered at  $\mathbf{x}$  with radius  $\rho$ .

**Definition 2** Given T > 0, the covering number G(T) of a compact manifold M is defined as the smallest number G(T) = |A|, where |A| = CN denotes the number of points of the set A on M, such that for all  $\mathbf{x} \in M$ :  $\min_{a \in A} d_M(\mathbf{x}, a) \leq T$ .

In the following lemma we establish a connection between the smoothness of the manifold and the smoothness of the coordinate signal  $f_r(n)$ . We define the coordinate graph signals  $f_r$ ,  $r = 1, \ldots, D$ , one per dimension of the ambient space, so that for vertex *i*, corresponding to data point  $\mathbf{x}_i$ , we have that:  $f_r(i) = \mathbf{x}_r(i)$ , i.e., the coordinate of  $\mathbf{x}_i$  along the *r*-th dimension. This lemma motivates our choice of denoising each of the coordinate signals in the graph domain. By using a sufficiently high sampling rate that depends on the smoothness properties of the manifold and its condition number  $1/\tau$ , we obtain that the points that are connected on the graph belong to the local neighborhood on the manifold, and the corresponding coordinate signals vary smoothly.

**Lemma 1** Consider a manifold M, with a condition number  $1/\tau$ , which is sampled at a resolution of a geodesic covering number G(T) = CN. Let  $d_M(\mathbf{x}_m, \mathbf{x}_n)$  denote the geodesic distance on the manifold M. Then, for all  $m, n \in G$  such that  $d_M(\mathbf{x}_m, \mathbf{x}_n) < \delta$ , we have that:

$$|f_r(n) - f_r(m)| \le \frac{1}{\tau} - \frac{1}{\tau}\sqrt{1 - \frac{T}{C}},$$
 (4)

where C is a constant,  $C \ge \max\left\{\frac{T}{2\tau}, \frac{2T}{\tau^2}\right\}$ .

**Proof** First note that for each r we have

$$|f_r(n) - f_r(m)| \le ||\mathbf{x}_m - \mathbf{x}_n|| \le d_M(\mathbf{x}_m, \mathbf{x}_n)$$
(5)

By Proposition 6.3 in [14], we have that

$$d_M(\mathbf{x}_m, \mathbf{x}_n) \le \frac{1}{\tau} - \frac{1}{\tau} \sqrt{1 - 2||\mathbf{x}_m - \mathbf{x}_n||\tau}$$
(6)

for  $\mathbf{x}_m, \mathbf{x}_n$  which obey  $||\mathbf{x}_m - \mathbf{x}_n|| \le \tau/2$ . Taking  $\delta = \frac{T}{2\tau C}$ , where C obeys  $C \ge max\left\{\frac{T}{2\tau}, \frac{2T}{\tau^2}\right\}$  we have that since  $||\mathbf{x}_m - \mathbf{x}_n|| \le \delta$ , then  $d_M(\mathbf{x}_m, \mathbf{x}_n) \le \frac{1}{\tau} - \frac{1}{\tau}\sqrt{1 - \frac{T}{C}}$  and thus the inequality is obtained.  $\Box$ 

Lemma 1 shows that for a manifold which is sampled with a sufficient density, the manifold coordinate signals  $f_r$  change smoothly, i.e., their local variation is bounded. Note that the bound is such that  $\Delta(\frac{1}{\tau},T) = 1/\tau - 1/\tau \sqrt{1 - \frac{T}{C}} \le 1$  and decreases as  $1/\tau$ 

decreases. In the next lemma we bound the total variation of coordinate signals with respect to the graph as a function the smoothness properties of the manifold (the condition number and the geodesic cover regularity.) This lemma shows that smoother manifolds will lead to coordinate signals with lower variation and thus lower graph frequencies.

**Lemma 2** Given a manifold M with a condition number  $1/\tau$ , sampled at a resolution of a geodesic covering number G(T), with the conditions of Lemma 1 satisfied. Then the following inequality holds:

$$|| \nabla \mathbf{f}_r ||^2 \le \Delta(\frac{1}{\tau}, T) \frac{\lambda_N}{C_{f_r}} \tag{7}$$

where  $\Delta(\frac{1}{\tau},T) = 1/\tau - 1/\tau \sqrt{1 - \frac{T}{C}}$ , and  $C_{f_r} = \bar{f_r}^2$  is the square of the mean of the graph signal  $f_r$ .

Proof Using the definition of the graph Laplacian we have

$$|| \bigtriangledown \mathbf{f}_r ||^2 = f_r^T \mathbf{L} f_r = \sum_{i \sim j} w_{ij} [f_r(i) - f_r(j)]^2$$

next using normalization we obtain

$$\frac{\sum_{i\sim j} w_{ij} [f_r(i) - f_r(j)]^2}{||f_r||^2} \le \frac{\sum_{i\sim j} w_{ij} (f_r(i) - f_r(j))^2}{N\bar{f_r}^2} \quad (8)$$

where in the second inequality we used the Cauchy-Schwartz inequality. By applying Lemma 1 with T, C that obey its conditions, we can bound the coordinate signal difference terms in (8) for all vertices that are 1-hop neighborhoods on the graph:

$$\frac{\sum_{i \sim j} w_{ij} [f_r(i) - f_r(j)]^2}{N \bar{f_r}^2} \le \frac{\Delta (1/\tau, T) \sum_{i \sim j} w_{ij}}{N \bar{f_r}^2}$$

where we used  $\Delta(1/\tau, T) = 1/\tau - 1/\tau \sqrt{1 - \frac{T}{C}}$ . Next summing over all vertices we get  $\sum d_i < \sum d_{max}$ , where  $d_{max}$  is the maximum degree and since  $d_{max} < \lambda_N$  [16], the Lemma is obtained.  $\Box$ 

Essentially, the lemma states that if we had two manifolds with samples leading to the same Laplacian L, the graph with the smoothest characteristics  $(1/\tau, T)$  would lead to coordinate signals with less variation  $|| \bigtriangledown \mathbf{r}_r ||^2$  and thus more energy concentrated in the lower frequencies. This will be reflected in the SGW domain as well. Furthermore, for a given Laplacian L, a noise signal would have a relatively flat energy distribution across bands in the SGW domain. Thus, if we think of the observed points x as the original manifold points to which noise has been added, Lemma 1 provides some justification for our approach based on setting to zero the higher frequencies of the SGW representation of coordinate signals.

#### 5. PROPOSED APPROACH

We now describe our approach for manifold denoising. We assume that the noiseless points lie on a smooth or piecewise smooth manifold  $M \in \mathbf{R}^D$ . Denoising is performed independently for each  $f_r()$ . In the noisy case, we assume that we are given a set of noisy points  $\tilde{f}_r(n) = f_r(n) + \epsilon_r(n)$ , contaminated with Gaussian noise  $\epsilon_r(n) \sim N(0, \sigma^2)$  with zero mean and variance  $\sigma^2$ . We assume the noise to be i.i.d. at each position and for each dimension r. Following this noise model, the goal is to provide an estimate  $\hat{f}_r(i)$  of the original coordinates  $f_r(i)$  given  $\tilde{f}_r(i)$  for each r and for all i. The reconstructed manifold points can be found by constructing  $\hat{\mathbf{x}}_i$ , which is based on  $\hat{f}_r(i)$ . In what follows, we will describe the processing done for each of these signals and unless required for clarity we drop the subscript r and use f to denote the graph signal. Our proposed algorithm is motivated by the following properties of smooth manifold:

- (i) The energy of the manifold coordinate signals is concentrated in the low frequency spectral wavelets.
- (ii) The noise power is spread out equally across all wavelets bands.

Property (i) is illustrated in Figure 1. As can be seen (Figure 1(b)) most of the energy is concentrated in the GFT coefficients that correspond to the smallest eigenvalues, and similarly (Figure 1(c)) the energy in each of the wavelet frequency bands for a 6 scale spectral wavelet decomposition can be seen to be concentrated in the low frequency wavelet bands. It is also important to note the difference between our denosing strategy and shrinkage based methods commonly used in classical wavelet denosing algorithms. In the case of wavelet image denoising, the signals lie on regular grids that are independent of the signal, while in our case, the graph and the noise free signal are closely related by our graph construction. In wavelet denoising for regular signals we mainly deal with piecewise smooth signals, which lead to a predominantly low frequency signal with localized high frequency coefficients that correspond to discontinuities in the piecewise smooth signal. In contrast, in our graph construction both the domain and the observations depend on the smoothness of the manifold. This has significant implications. For example, if the sampling rate along the manifold varies with the degree of smoothness, we may expect locally smooth behavior of coordinate signals even in areas where the geometry is not as smooth. Thus we do not see SGW domain characteristics similar to what is observed in wavelet domain representation of piecewise smooth regular domain signals (isolated high frequency coefficients).

Based on these properties, denoising is performed directly in the spectral graph domain, by retaining all wavelet coefficients that correspond to the low frequency wavelets bands  $s \leq s'$ , and discarding all wavelet coefficients in high frequency bands above s > s'.



**Fig. 1**. Plot of the energy of a Swiss roll with a hole (a) in the graph Fourier transform (b) and in the spectral wavelet domain (c)

We summarize the proposed denoising algorithm for smooth manifolds as follows:

1. Construct an undirected affinity graph **W**, using Gaussian weights as in (1), and construct the Laplacian **L** from **W**. For each dimension, assign the corresponding coordinates values of each point to its corresponding vertex on the graph, and apply Steps 2, 3 and 4 to each of the dimensions independently.

- 2. Transform the noisy coordinate signal using SGW defined on L.
- 3. Retain all scaling coefficients and all wavelet coefficients below a low pass frequency  $s \leq s'$ , for which the total accumulated energy is above threshold  $E_{thresh}$ . Discard all wavelet coefficients above scales > s'.

4. Take the inverse spectral wavelet transform of the processed wavelet coefficients from Step 3.

This approach has several attractive features, in particular, it is (1) non-iterative, i.e., denoising is performed directly in the spectral graph wavelet domain in one step, (2) robust against a wide range of k values chosen for nearest neighbor assignment on the graph, (3) computationally efficient, as the computational complexity is O(ND).

## 6. EXPERIMENTAL RESULTS



**Fig. 2.** Experimental results on a fish bowl, helix, and a sinus embedded in dimension D=200. Ground truth is shown in blue color, denoised in red. In each row, from left to right: noisy points, results with MD, LLD, and MFD.



**Fig. 3.** RMSE reconstruction error of the noisy manifolds using different selection of k nearest neighbor

We present experimental results with a variety of manifolds, including ones with complex geometric structure such as fish bowl a Swiss roll with a hole and a sinus function that was embedded in high dimensional space of D = 200. All manifolds were sampled using a uniform distribution with N = 1000 samples, which were contaminated with isotropic Gaussian noise in all dimensions. In the results shown, the sinus and fish-bowl were contaminated with noise of variance of 0.2, and the circle, helix and Swiss roll with a hole with

Data/Method	MD	LLD	MFD
CMU MoCap	11.84	7.35	3.42
Frey face datasets	109.1	62.2	51.2

Table 1. RMSE average error results on MoCap and Frey datasets

variance 0.1. We used s = 5 wavelet decomposition levels, and retained all wavelet coefficients which correspond to the lowest s < s'scales above total accumulated energy threshold  $E_{thresh(var=0.1)}$ for noise variance equal to 0.1, and the lowest scales  $s \leq s'$  above total accumulated energy threshold  $E_{thresh(var=0.2)}$  with variance 0.2. The order of the Chebyshev polynomial approximation used was k/2 for a k nearest neighbor graph in order to process the manifold locally, via the approximation of the spectral wavelet coefficients. For comparison and evaluation with the state of the art in manifold denoising, we compared our approach to MD[5] and LLD [6]. The denoising results are shown in Figure 2. In all of these cases, our method significantly outperforms the state of the art, and produced a smooth reconstruction that is faithful to the true topological structure of the manifold. We also performed quantitative analysis and compared the reconstruction error in terms of the root mean square error (RMSE) of the denoised manifolds in comparison to the ground truth. The comparison results in Figure 3 shows that our method is robust against a wide range of k nearest neighbor graph selections, outperforming the competing methods by orders of magnitudes for a wide range of k nearest neighbors selection. For experiments with real data-sets, we tested our method on the CMU Motion capture data-set and the Frey faces data-set [17], [2]. For the Motion capture we test 10 sequences of subject 86, where each sequence has a dimension D=62. The Frey face dataset consists of low resolution faces with dimension D=560. For the Motion capture data-set, in order to perform evaluation in a strictly unsupervised framework, we remove the temporal information from the data. The data is contaminated using Gaussian noise of variance 0.1 in all dimensions. The experimental results evaluation in terms of the average RMSE error are shown in Table 1, where MFD shows significant improvement over LLD and MD.

#### 7. CONCLUSIONS AND FUTURE WORK

We have presented a new framework for manifold denoising which simultaneously operates in the vertex and frequency graph domains by using spectral graph wavelets. The advantage of such an approach is that it allows us to denoise the manifold locally, while taking into account the fine-grain regularity properties of the manifold. Our approach is based on the property that the energy of a smooth manifold is concentrated in the low frequencies of the graph, while the noise effects all frequency bands in a similar way.

The suggested MFD framework also possesses additional appealing properties: it is non-iterative, has low computational complexity, and it does not require the knowledge of the intrinsic dimensionality of the manifold. Experimental results on manifolds with complex geometric structure show that our approach significantly outperforms the state of the art, and is robust to a wide range of parameter selection of k nearest neighbors on the graph. Future work includes addressing the case of non-Gaussian noise, and further investigation of how the underlying graph construction affects the spectral transform properties.

Acknowledgments This work was supported in part by grant DE-NA0001683 from the U.S. Department of Energy.

#### 8. REFERENCES

- J. Tenenbaum, V. de Silva, and J. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [2] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *SCIENCE*, vol. 290, pp. 2323– 2326, 2000.
- [3] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [4] Z. Zhang and H. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2005.
- [5] M. Hein and M. Maier, "Manifold denoising," in Advances in Neural Information Processing Systems 19, Cambridge, MA, USA, Sept. 2007, Max-Planck-Gesellschaft, pp. 561–568, MIT Press.
- [6] Dian Gong, Fei Sha, and Gérard G. Medioni, "Locally linear denoising on image manifolds," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2010, Chia Laguna Resort, Sardinia, Italy, May 13-15, 2010, 2010, pp. 265–272.*
- [7] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129– 150, Mar. 2011.
- [8] D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, May 2013.
- [9] Jiahao Pang, Gene Cheung, Antonio Ortega, and Oscar C. Au, "Optimal graph laplacian regularization for natural image denoising," *IEEE International Conference on Acoustics, Speech* and Signal Processing, ICASSP, , no. 6, 2015.
- [10] Neil Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data," in *In NIPS*, 2003, p. 2004.
- [11] Yan Gao, Kap Luk Chan, and Wei-Yun Yau, "Manifold denoising with gaussian process latent variable models," in 19th International Conference on Pattern Recognition (ICPR 2008), December 8-11, 2008, Tampa, Florida, USA, 2008, pp. 1–4.
- [12] David L. Donoho and Iain M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.
- [13] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, "A review of image denoising algorithms, with a new one," *Multi-scale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [14] Partha Niyogi, Stephen Smale, and Shmuel Weinberger, "Finding the homology of submanifolds with high confidence from random samples," *Discrete & Computational Geometry*, vol. 39, no. 1, pp. 419–441, 2008.
- [15] Richard G. Baraniuk and Michael B. Wakin, "Random projections of smooth manifolds," in *Foundations of Computational Mathematics*, 2006, pp. 941–944.

- [16] A.E. Brouwer and W.H. Haemers, "A lower bound for the laplacian eigenvalues of a graph-proof of a conjecture by guo," Discussion Paper 2008-27, Tilburg University, Center for Economic Research, 2008.
- [17] Sam Roweis, Lawrence K. Saul, and Geoffrey E. Hinton, "Global coordination of local linear models," in *Advances in Neural Information Processing Systems* 14. 2002, pp. 889–896, MIT Press.