GENERATIVE MODELING BASED MANIFOLD LEARNING FOR ADAPTIVE FILTERING GUIDANCE

Karim Helwani, Paris Smaragdis*,♯, and Michael M. Goodwin**

*Amazon Web Services, ♯University of Illinois at Urbana-Champaign

ABSTRACT

In most practical adaptive filtering problems, estimated filters are not arbitrary, but instead lie on a manifold that encapsulates characteristics of the problem at hand. Consequently, it is desirable to steer adaptation towards filters that lie on that manifold. In this paper, we propose a novel approach to learn the manifold of a set of impulse responses and subsequently employ that learned manifold in an adaptation algorithm for system identification. The presented approach is a practical adaptive filtering recipe for enforcing a data-driven search domain constraint, instead of using conventional constrained optimization methods.

Index Terms— Manifold learning, generative models, variational autoencoder, adaptive filtering.

1. INTRODUCTION

Real-time communications scenarios are often dependent on the success of the design of adaptive filters. For example, in a hands-free teleconferencing system, an acoustic echo canceler (AEC) is implemented to prevent the loudspeaker signal captured by the microphone to be sent back to the far-end causing disturbing echoes. An AEC creates a replica of the echo signal and subtracts it from the microphone signal before sending it back to the far-end. Microphone arrays often employ adaptive interference cancellation algorithms such as generalized sidelobe cancelers (GSC) which are in turn based on adaptive filters estimating the leakage path of the interferer into the desired beamformer direction and hence, minimizing the interference [1]. Tasks of prediction and inverse modeling as in dereverberation are also classes of adaptive filtering problems. Although data-driven approaches have taken the lead in research in recent years due to the overwhelming success of deep learning, system models derived from physics are indispensable in many applications [2] and hybrid combinations of data-driven and system based approaches have been shown to provide reliable and efficient solutions to problems of signal processing see, e.g., [3, 4].

The success of designing an adaptive filter depends on the adequacy of prior information about the statistics of the data to be processed and the system to be estimated [5]. For example, the main challenges of estimating a room impulse response in a linear acoustic echo cancellation scenario are solving an ill-conditioned linear system of equations due to the correlated excitement of the acoustic path and the contamination of the microphone signal with near-end noise. The availability of the prior knowledge about the system, e.g., impulse response sparsity, can help alleviate the ill-conditioning problem. Similarly, prior knowledge about the noise signal statistics can help to minimize bias from the estimation, e.g., [6].

The incorporation of prior knowledge into an optimization problem can be achieved either by means of adding constraints to the optimization cost function and keeping the search space the same, or by constraining the search space in the first place which often has favorable numerical properties as has been shown in the prominent example of natural gradient based algorithms e.g., for blind source separation tasks [7].

In this paper, we focus on an approach to constrain the search space of adaptive filtering optimization problems. The approach is a hybrid deep learning/DSP based solution where a deep neural network architecture is trained to generate impulse responses on a learned manifold. The generation network is then used to directly estimate the actual impulse response in real-time based on acoustic environment measurements. Hence, the adaptive filter refines the network-generated impulse response to the actual environment and the neural network aids the adaptive filter by constraining the search space of the online adaptation.

2. RELATION TO PRIOR WORK

Manifold learning for audio applications has been discussed in [8] and [9] focusing on localization and speech enhancement tasks. In the present paper we use a deep generative model to guide an adaptive filter on a learned manifold. Along these lines, a recent approach showed how to estimate the optimal update using a neural network [10]. In contrast, the current paper does not focus on the update itself, but on the estimated filter, ensuring that it conforms to the data geometry of the problem at hand. The foundations of optimization on manifolds foundation have been discussed in [11], in which most approaches assume the existence of a deterministic manifold characterization. In [12, 13, 14] approaches for unsupervised adaptation for audio tasks on apriori known manifolds were discussed. In [15] and [16] the relation between compressive domain adaptive filtering and manifold learning has been established for the special case of sparse signals. In this paper a generic manifold learning as well as the adaptation on the learned manifold are being introduced.

3. MANIFOLD LEARNING

In signal processing, it is often the case that high-dimensional data can be assumed to lie on a manifold globally isometric to a subset of a low-dimensional Euclidean space. This assumption allows the accurate and efficient low-dimensional parameterization of high-dimensional data [17]. A manifold is a topological space that is locally Euclidean [13], a local parameterization in the Euclidean tangent space is given using what is called a retraction [11]. A retraction on a manifold $\mathcal{M}$ is a smooth mapping ψ from the tangent bundle $T\mathcal{M}$ onto $\mathcal{M}$ with the following properties. Let $\psi_{\mathbf{h}}$ denote the restriction of ψ to $T_{\mathbf{h}}\mathcal{M}$. (i) $\psi_{\mathbf{h}}(0_{\mathbf{h}}) = \mathbf{h}$, where $0_{\mathbf{h}}$ denotes the zero element of $T_{\mathbf{h}}\mathcal{M}$. (ii) With the canonical identification $T_{0_{\mathbf{h}}}T_{\mathbf{h}}\mathcal{M} \simeq T_{\mathbf{h}}\mathcal{M}$, $\psi_{\mathbf{h}}$ satisfies

$$\mathrm{D}\psi_{\mathbf{h}}(0_{\mathbf{h}}) = \mathrm{id}_{T_{\mathbf{h}}\mathcal{M}}, \quad (1)$$

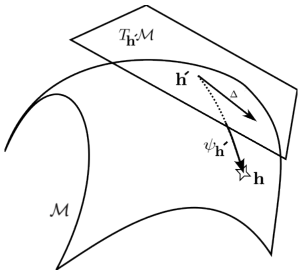

Fig. 1. The concept of retraction, mapping a point in the tangent space back to the manifold.

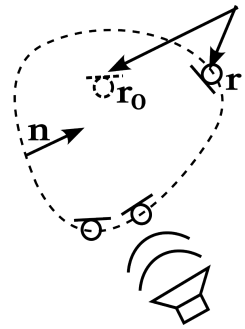

Fig. 2. Kirchhoff-Helmholtz integral setup. The pressure within an enclosed volume is fully determined by the pressure and its gradient at the boundary of this volume.

where $\mathrm{id}_{T_{\mathbf{h}}\mathcal{M}}$ denotes the identity mapping on $T_{\mathbf{h}}\mathcal{M}$. See Fig. 1 for illustrating our setup.

Data-driven approaches to obtain a manifold description mainly for non-linear dimensionality reduction have been discussed in the literature extensively, e.g., [18, 19, 17, 20].

Variational Auto-Encoders (VAEs) are one of the most prominent deep neural network models that achieves both a non-linear dimensionality reduction and generative modeling. In general, the latent space parameterization in a VAE doesn't reflect the same topological structure as the input data but as we will see in Sect. 4, this can be enforced by introducing a proper cost function. A VAE minimizes the following cost function which can be seen as the negative of the evidence lower bound (ELBO):

$$\mathcal{L}_r := \mathcal{E}_{\mathbf{h}}\Big[\mathcal{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{h})}\left[-\log p_\theta(\mathbf{h}|\mathbf{z})\right] + \mathrm{KL}\left(q_\phi(\mathbf{z}|\mathbf{h})\|p(\mathbf{z})\right)\Big] + D\left(q_\phi(\mathbf{z})\|p(\mathbf{z})\right), \quad (2)$$

with respect to θ, ϕ denoting the parameters of the decoder and encoder respectively and $\mathbf{z}$ is the latent variable. D is a regularization term that we chose to disentangle the latent space as proposed in the so-called DIP-VAE-II [21]. The motivation to favor the disentanglement in our use case is that we will be performing the adaptation in the latent space. Ensuring that the covariance of the latent representation is diagonal results in an approximately diagonal Hessian as we will see and hence, better numerical properties of the adaptation algorithm. Hence, we choose the regularization term to be

$$D\left(q_\phi(\mathbf{z})\|p(\mathbf{z})\right) := \lambda_{off}\sum_{i\neq j}\Big[\mathrm{Cov}_{q_\phi(\mathbf{z})}[\mathbf{z}]\Big]_{ij}^2 + \lambda_{diag}\sum_i\left(\Big[\mathrm{Cov}_{q_\phi(\mathbf{z})}[\mathbf{z}]\Big]_{ii} - 1\right)^2, \quad (3)$$

with λ_{off} a Lagrangian multiplier constraining the off diagonals of the covariance matrices, λ_{diag} another Lagrangian for the diagonal elements, and

$$\mathrm{Cov}_{q(\mathbf{z})}[\mathbf{z}] := \mathcal{E}_{q(\mathbf{z})}\left[\left(\mathbf{z} - \mathcal{E}_{q(\mathbf{z})}[\mathbf{z}]\right)\left(\mathbf{z} - \mathcal{E}_{q[\mathbf{z}]}(\mathbf{z})\right)^{\mathrm{T}}\right]. \quad (4)$$

4. TOPOLOGY AWARE AUTOENCODER

It has been shown that standard VAEs may not preserve the topology between the input and the latent space. Attempts to modify VAEs to preserve the topology for performing interpolations tasks in the latent space have been proposed in [22] and [23]. The core idea is to constrain the VAE to approximate a simplicial map satisfying the condition

$$\varphi\left(\sum_{j=1}^{k}\gamma_j\sigma_j\right) = \sum_{j=1}^{k}\gamma_j\varphi\left(\sigma_j\right), \quad (5)$$

where φ denoting the mapping performed by the encoder. σ is a k-simplex in a simplicial complex K, γ is a convex coefficient vector. Intuitively, this condition means that the vertices of a simplex in the input space spans a simplex in the latent space.

$$\mathcal{L}_t(\varphi, K, \alpha) = \sum_{\sigma\in K}\mathcal{E}_{\gamma_j\sim\mathrm{Dir}(\dim(\sigma),\alpha)}\mathcal{L}_t\left(\varphi\left(\sum_{j=1}^{\dim(\sigma)}\gamma_j\sigma_j\right), \sum_{j=1}^{\dim(\sigma)}\gamma_j\varphi\left(\sigma_j\right)\right). \quad (6)$$

The cost function of a VAE can be constrained to fulfill this condition, which results in the following optimization cost function for a topology aware VAE [23]

$$\mathcal{L} := \mathcal{L}_r + \lambda\mathcal{L}_t. \quad (7)$$

K is a simplicial complex built from the input space. σ is a simplex belonging to the simplicial complex K. σ_j is here the vertex j of the dim(σ)-simplex σ. A simplex σ has dim(σ) + 1 vertices. σ_j. $\mathcal{E}_{\gamma_j\sim\mathrm{Dir}(\dim(\sigma)+1,\alpha)}$ is the expectation for the $(\gamma_j)_{j=0,\ldots,\dim(\sigma)}$ following a symmetric Dirichlet distribution with the order dim(σ) + 1 and the concentration parameter α.

5. FROM MEASUREMENT POINTS TOPOLOGY TO IMPULSE RESPONSES TOPOLOGY

In this work, we are primarily interested in the topology of the impulse responses and not of the position at which the impulse responses have been measured. We relate the measured impulse responses and the microphone positions with the so-called Kirchhoff-Helmholtz integral.

$$P(\mathbf{r},\omega) = \oint\left(\frac{\partial}{\partial\mathbf{n}}\underline{h}\left(\mathbf{r}\mid\mathbf{r_0},\omega\right)P\left(\mathbf{r_0},\omega\right) - \frac{\partial}{\partial\mathbf{n}}P\left(\mathbf{r_0},\omega\right)\underline{h}\left(\mathbf{r}\mid\mathbf{r_0},\omega\right)\right)\mathrm{d}\mathbf{r_0}, \quad (8)$$

with $\underline{h}$ being the Green's function representation in the frequency domain due to a source at $\mathbf{r_0}$, $\mathbf{n}$ is the normal vector along the enclosing boundary, $P(\mathbf{r},\omega)$ the sound pressure at $\mathbf{r}$ and the frequency

ω. A high frequency approximation for this integral [24] allows deriving the pressure at any point within the boundary while taking into account only the second term in the integral, the setup is shown in Fig. 2. Please note that $\underline{h}\left(\mathbf{r} \mid \mathbf{r_0}, \omega\right)$ is the acoustic transfer function between the position $\mathbf{r}$ and $\mathbf{r_0}$.
Hence, from given impulse response measurement locations, we define a simplicial complex. The vertices of each simplex are understood as a discretized boundary for the Kirchhoff-Helmholtz integral. A combination of the vertices in a simplex gives a point $\mathbf{r}_0$ within the simplex (boundary). The latent space representation of the impulse response from a loudspeaker outside the simplex to a virtual microphone at $\mathbf{r}_0$ should be equal to the sum of the latent representation of the impulse responses from the randomly selected loudspeaker position to the vertices after being filtered by the transfer function between the respective vertex and $\mathbf{r}_0$.

6. ADAPTATION ON PARTLY SMOOTH MANIFOLDS

In the literature two main approaches to optimization on manifolds can be found; retraction mapping based approaches where only local information in the Euclidean tangent space and the mapping between the tangent space and the manifold are employed e.g., [13]. The second category covers retraction and transport-based approaches. In this second category, information from multiple tangent spaces and specifically a *Levi-Civita connection* is employed to ensure the optimization along geodesics of the manifold since a geodesic always parallel transports its tangent vector. See e.g., [25]. While transport-based approaches can offer better understanding of the behavior of the algorithm in certain applications, obtaining a parallel transport can be challenging in the general case especially, when the manifold is obtained in a data driven manner. In the present study, our approach to optimization can be classified as retraction-based. In a retraction based approach, the optimization is done in the Euclidean tangent space by translating the parameters Element by the vector Δ. The updated parameters are mapped back onto the manifold by the retraction mapping ψ of the tangent space at $\mathbf{h}'$

$$\Delta_{\text{opt}} = \arg \min_{\Delta \in \mathbb{R}^L} \mathcal{L}_a \left(\psi_{\text{h}'}(\Delta)\right), \tag{9}$$

with $\mathcal{L}_a$ denoting the loss function of the adaptive filter optimization problem. Finding an optimal point over time t iteratively, can be done by solving the ordinary differential equation

$$\frac{d\Delta}{dt} = \nabla_\Delta \mathcal{L}_a, \tag{10}$$

which can be solved using the Euler method until a steady state is reached.
The gradient of the loss w.r.t. the tangent space can be obtained using the chain rule

$$\nabla_\Delta \mathcal{L}_a \left(\psi_{\mathbf{h}'}(\Delta)\right)\big|_{\Delta=0} = \left.\frac{\partial\left(\psi_{\mathbf{h}'}^{\mathrm{T}}\right)}{\partial\left(\Delta\right)}\right|_{\Delta=0} \left.\frac{\partial \mathbf{h}'^{\mathrm{T}}}{\partial \psi_{\mathbf{h}'}}\right|_{\Delta=0} \frac{\partial \mathcal{L}_a}{\partial \mathbf{h}'}, \tag{11}$$

with the defining properties of the retraction function, it can be seen that at $\Delta = 0$, $\psi(\mathbf{0}) = \mathbf{h}'$, hence,

$$\left.\frac{\partial \mathbf{h}'^{\mathrm{T}}}{\partial \psi_{\mathbf{h}'}}\right|_{\Delta=0} = \mathbf{I}, \tag{12}$$

with $\mathbf{I}$ being the identity matrix. Hence, the gradient w.r.t. tangent space simplifies to

$$\nabla_\Delta \mathcal{L}_a \left(\psi_{\mathbf{h}'}(\Delta)\right)\big|_{\Delta=0} = \left.\frac{\partial\left(\psi_{\mathbf{h}'}^{\mathrm{T}}\right)}{\partial\left(\Delta\right)}\right|_{\Delta=0} \frac{\partial \mathcal{L}_a}{\partial \mathbf{h}'}. \tag{13}$$

This gives the update for the Euler method in the Euclidean tangent space with a step size μ,

$$\Delta(n) = \Delta(n-1) - \mu \cdot \boldsymbol{\Xi} \cdot \frac{\partial \mathcal{L}_a}{\partial \mathbf{h}'}, \tag{14}$$

with

$$\boldsymbol{\Xi} := \left.\frac{\partial\left(\psi_{\mathbf{h}'}^{\mathrm{T}}\right)}{\partial\left(\Delta\right)}\right|_{\Delta=0}. \tag{15}$$

A retraction onto the manifold gives the final parameters vector. Hence, the final update is given by the retraction mapping which is given by the VAE decoder

$$\mathbf{h} = \psi_{\mathbf{h}'}(\Delta). \tag{16}$$

Please note, that an Euclidean update can be seen as a special case where the retraction is a pure translation. Figure 3 depicts a block-diagram of the proposed adaptive filtering scheme.

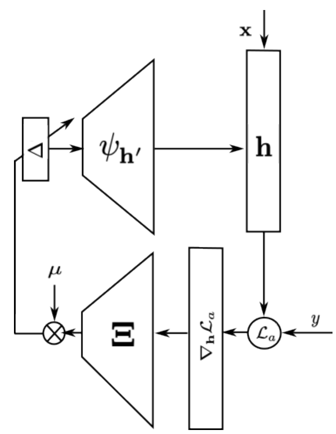

Fig. 3. An overview of the proposed adaptive filtering scheme. The input signal $\mathbf{x}$, filtered with the estimated filter $\mathbf{h}$, results in a replica of the target signal y. A loss $\mathcal{L}_a$ is estimated, then the gradient of the loss with respect to $\mathbf{h}$ is calculated. The Jacobi matrix $\boldsymbol{\Xi}$ of the retraction map, which is here the decoder of a topology-aware VAE, is calculated, $\boldsymbol{\Delta}$ is then given as the multiplication of the gradient of the loss and the Jacobi matrix of the retraction map and a step size μ. Finally, applying the retraction map $\psi_{\mathbf{h}'}$ from the tangent space at the old point $\mathbf{h}'$ onto the learned manifold gives the updated filter parameters.

7. LINEAR SYSTEM IDENTIFICATION AS ILLUSTRATING EXAMPLE

In an AEC scenario, the echo path from the a loudspeaker to a microphone can be identified by minimizing the following cost function[1]

$$\mathcal{L}_a = \mathcal{E}\left\{|e(n)|^2\right\} = \mathcal{E}\left\{\left|y(n) - \mathbf{h}^{\mathrm{H}}\mathbf{x}(n)\right|^2\right\}, \tag{17}$$

[1]This choice is for illustration. More robust cost functions that take explicitly the statistics of the near-end noise into account, e.g., weighted least-squares or Huber loss, fit equally into the framework we are presenting.

with

$$\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-L+1)]^{\mathrm{T}}, \qquad (18)$$

and

$$\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^{\mathrm{T}}. \qquad (19)$$

The gradient of the loss function w.r.t. $\mathbf{h}$ is,

$$\nabla_{\mathbf{h}} \mathcal{L}_a = -2\mathcal{E}\left\{\mathbf{x}(n)\left[y^*(k) - \mathbf{h}^{\mathrm{T}}\mathbf{x}^*(n)\right]\right\} \qquad (20)$$

Hence, the update of the adaptive filter in the manifold parameterization space is given Eq. (13) and is given by the Jacobian of the decoder $\Xi(n)$ multiplied by the gradient of the original adaptive filtering problem Eq. (20).

$$\mathbf{z}(n) = \mathbf{z}(n-1) + \mu\Xi(n)\mathbf{x}(n)e^*(n). \qquad (21)$$

Interestingly, even the adaptation is a first-order optimization, we have a structure for the update where a matrix is being multiplied with the gradient similar to preconditioned gradient descent methods or second order approaches. Finally, the filter given the retraction defined by the decoder ψ

$$\mathbf{h}(n) = \psi_{\mathbf{h}(n-1)}(\mathbf{z}(n)). \qquad (22)$$

For an actual second-order adaptation approach, a Newton-based update in the tangent space can be derived. To achieve this, the inverse of the Hessian needs to be calculated and it can be shown that the Hessian is given by

$$\boldsymbol{F}(n) := \Xi(n)\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\Xi^{\mathrm{T}}(n). \qquad (23)$$

It should be emphasized, that the disentanglement of the latent space mentioned above contributes to the compaction of the Hessian matrix along the diagonal. Finally, the second-order update is:

$$\mathbf{z}(n) = \mathbf{z}(n-1) + \mu\boldsymbol{F}^{-1}(n)\Xi(n)\mathbf{x}(n)e^*(n). \qquad (24)$$

8. EXPERIMENTAL RESULTS

To prove our concept we experiment with synthesized room impulse responses using the framework Pyroomacoustics [26]. We train the presented acoustic topology-aware variational autoencoder on 200 impulse responses from a shoebox room of the dimensions $6 \times 6 \times 2$ m. We use a three-layer convolutional VAE with sizes [2,4,8] and 5-tap filters using strides [1,2,2] and $\tanh$ as a nonlinearity for each layer respectively. Three and four dimensional simplices are selected, and per simplex a random position in that simplex is selected as focal point for the virtual microphone which is used later for the topological constraint as explained int Sect. 5. After training, unseen impulse responses from the same room are used to simulate an echo cancellation scenario under SNR condition of 20 dB. Figure 4 depicts the achieved system distance with the proposed 1^{st}- and 2^{nd}-order updates with $\mu = 0.15$ compared to a recursive least squares adaptive filter (RLS), and ℓ_1-regularized RLS based estimations with a forgetting factor of 0.95 and ℓ_1-norm constraint Lagrangian of 0.1 [27]. The effect of the reduction of the filter size in the latent space is manifested in the rapid drop of the system distance even before processing as many samples as filter taps, although the adaptive filter is not implemented in an order recursive manner. Figure 5 depicts snapshots of the estimated impulse responses during the adaptation process, Fig. 6 shows snapshots of the estimated impulse response for the proposed approach. It can be seen that after few iterations, the impulse response takes already the structure of an expected room impulse responses in contrast to the ℓ_1-regularized RLS.

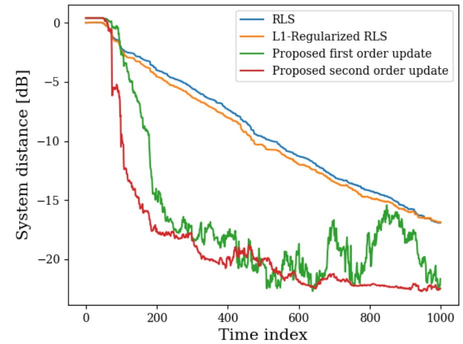

Fig. 4. Resulting system distance by the proposed 1^{st}- and 2^{nd}-order updates in contrast to uninformed RLS and ℓ_1-regularized RLS. In the initial phase, the system predicts an impulse response in the manifold even though the system hasn't been excited resulting in an initially relatively large system distance, once the system excitation allows for a complete latent space representation, a fast convergence can be observed.

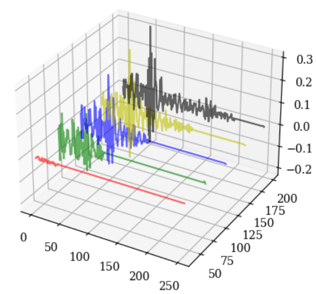

Fig. 5. The estimated impulse responses at different points of the first 200 adaptation steps for ℓ_1-regularized RLS. Note that at the beginning of the adaptation the estimated impulses do not look very much like shoebox room filters.

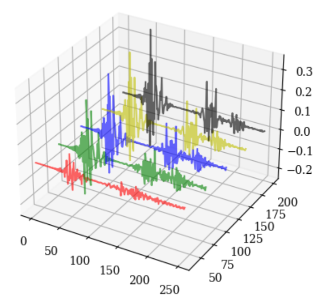

Fig. 6. The estimated impulse responses of the proposed method at different points of the first 200 adaptation steps. Note that in contrast to the filter estimates in figure 5, the filters look more like the room responses at hand early on in the adaptation process.

9. CONCLUSIONS AND OUTLOOK

An approach to learn the acoustic environment by means of manifold learning using a variational autoencoder has been presented. It has been shown, how adaptive filtering can be performed on the learned manifold. Preliminary results confirm the ability to estimate underdetermined acoustic systems by performing the adaptation on such a manifold. The present study outlines a path for future research on generating acoustically plausible impulse responses from a variety of acoustic environments given limited real-time measurements.

10. REFERENCES

[1] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 1, pp. 27–34, 1982.

[2] Holger Boche, Adalbert Fono, and Gitta Kutyniok, "Limitations of deep learning for inverse problems on digital hardware," *arXiv preprint arXiv:202.13490*, 2022.

[3] Jean-Marc Valin, Srikanth Tenneti, Karim Helwani, Umut Isik, and Arvindh Krishnaswamy, "Low-complexity, real-time joint neural echo control and speech enhancement based on percepnet," 2021.

[4] Jean-Marc Valin, "A hybrid dsp/deep learning approach to real-time full-band speech enhancement," 2017.

[5] Simon Haykin, *Adaptive filter theory*, Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002.

[6] J. Benesty and T. Gansler, "A robust fast recursive least squares adaptive algorithm," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, 2001, vol. 6, pp. 3785–3788 vol.6.

[7] Shun-ichi Amari, "Natural Gradient Works Efficiently in Learning," *Neural Computation*, vol. 10, no. 2, pp. 251–276, 02 1998.

[8] Ronen Talmon, Israel Cohen, Sharon Gannot, and Ronald R Coifman, "Diffusion maps for signal processing: A deeper look at manifold-learning techniques based on kernels and graphs," *IEEE signal processing magazine*, vol. 30, no. 4, pp. 75–86, 2013.

[9] Yuancheng Luo, Dmitry N Zotkin, and Ramani Duraiswami, "Gaussian process models for hrtf based sound-source localization and active-learning," *arXiv preprint arXiv:1502.03163*, 2015.

[10] Jonah Casebeer, Nicholas J Bryan, and Paris Smaragdis, "Meta-af: Meta-learning for adaptive filters," *arXiv preprint arXiv:2204.11942*, 2022.

[11] P-A Absil, Robert Mahony, and Rodolphe Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, 2009.

[12] Mark D. Plumbley, "Geometry and manifolds for independent component analysis," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, 2007, vol. 4, pp. IV–1397–IV–1400.

[13] Herbert Buchner, "A systematic approach to incorporate deterministic prior knowledge in broadband adaptive mimo systems," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*. IEEE, 2010, pp. 461–468.

[14] Herbert Buchner, Karim Helwani, and Simon Godsill, "Unsupervised bayesian estimation and tracking of time-varying convolutive multichannel systems," in *2019 22th International Conference on Information Fusion (FUSION)*. IEEE, 2019, pp. 1–8.

[15] Herbert Buchner, Karim Helwani, and Simon Godsill, "Adaptive dynamical systems in compressive domains as a manifold learning framework," in *SPARS Workshop*, 2015.

[16] Karim Helwani and Herbert Buchner, "Multichannel adaptive filtering in compressive domains," in *2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, 2014, pp. 174–177.

[17] David L. Donoho and Carrie Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proceedings of the National Academy of Sciences*, vol. 100, no. 10, pp. 5591–5596, 2003.

[18] Sam T Roweis and Lawrence K Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[19] Mikhail Belkin and Partha Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in neural information processing systems*, vol. 14, 2001.

[20] Ronald R. Coifman and Stéphane Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006, Special Issue: Diffusion Maps and Wavelets.

[21] Abhishek Kumar, Prasanna Sattigeri, and Avinash Balakrishnan, "Variational inference of disentangled latent concepts from unlabeled observations," 2017.

[22] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt, "Topological autoencoders," 2021.

[23] Jose Daniel Gallego Posada, "Simplicial autoencoders: A connection between algebraic topology and probabilistic modelling," 2018.

[24] Augustinus J. C. Berkhout, "A holographic approach to acoustic control," *Journal of The Audio Engineering Society*, vol. 36, pp. 977–995, 1988.

[25] Alan Edelman, T. A. Arias, and Steven T. Smith, "The geometry of algorithms with orthogonality constraints," 1998.

[26] Robin Scheibler, Eric Bezzam, and Ivan Dokmanic, "Pyroomacoustics: A python package for audio room simulation and array processing algorithms," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. apr 2018, IEEE.

[27] Karim Helwani, Herbert Buchner, and Sascha Spors, "Multichannel adaptive filtering with sparseness constraints," in *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*, 2012, pp. 1–4.