Adaptive Dynamical Systems in Compressive Domains as a Manifold Learning Framework

Herbert Buchner
University of Cambridge,
Cambridge, UK

Karim Helwani
Huawei European Research Center
Munich, Germany

Simon Godsill
University of Cambridge
Cambridge, UK

Adaptive dynamical systems based on finite impulse response (FIR) models constitute an important part in many signal processing applications [1]. In this contribution, we consider the case of *sparse systems* for increasing the convergence speed and/or reducing the computational complexity, i.e., in our scenarios only a small percentage of the system components has significant magnitude (temporal coefficients and/or spatial components (in multichannel case)). Note that this sparseness can also be emphasized by working in suitable transform domains, e.g., [2], [3]. In any case, the exact knowledge about the relevant dimensions is typically not available a priori. One well known way of taking into account the sparseness of systems is the use of certain regularization techniques (e.g., [4] and references therein).

In contrast to these techniques, our aim here is to go one step further and reduce the number of adaptive coefficients by performing the adaptation in *compressive domains* based on insights from the field of compressed sensing (CS), e.g., [5], [6], [7], and our recent work [8]. We present a simple, yet efficient novel CS-based adaptation scheme which allows us to express the resulting adaptation algorithm for sparse dynamical systems as some kind of transform-domain adaptive filtering operating in a lower-dimensional space by means of a certain self-adapting transformation matrix.

Using this formulation, we further show how the adaptive filtering in compressive domains can be expressed as a specific case of a very general, previously introduced framework for adaptive filtering on manifolds. This novel relation between CS and manifold learning has a number of significant practical benefits. Besides a new illustrative interpretation, it immediately shows us how to obtain novel practical algorithms in compressive domains for essentially all possible classes of adaptive filtering problems for supervised as well as for blind and semi-blind adaptation.

I. Adaptive Filtering in Compressive Domains

To begin with, we consider supervised single-channel system identification as shown in Fig. 1, where $\mathbf{h}$ denotes the length-L vector of FIR coefficients of the unknown system, $\hat{\mathbf{h}}$ is its estimate, and quantities in the compressed domain are underlined. Following the theory of compressed sensing, it should be possible to perfectly reconstruct the sparse vector $\hat{\mathbf{h}}$ from its undersampled version $\underline{\hat{\mathbf{h}}} := \mathbf{\Phi}\hat{\mathbf{h}}$ with a random $K \times L$ observation matrix $\mathbf{\Phi}$ ($K \ll L$). To perform a unique and simultaneous estimation of $\underline{\hat{\mathbf{h}}}$ and its sparse reconstruction, we apply an iterative procedure alternating between an error minimization of $e(n)$ (typically in least-squares sense) to obtain $\underline{\hat{\mathbf{h}}}$, and the ℓ_1-constrained minimization of $\lambda \left\|\widehat{\mathbf{h}}(n)\right\|_1 + \left\|\underline{\widehat{\mathbf{h}}}(n) - \mathbf{\Phi}\widehat{\mathbf{h}}(n)\right\|_2^2$, where λ is a Lagrange multiplier. Under the assumption of small iterative changes of $\hat{\mathbf{h}}(n)$, it turns out that the reconstruction step can be expressed using a matrix resembling the pseudoinverse of $\mathbf{\Phi}$ but with an additive $\hat{\mathbf{h}}(n)$-dependent nonlinear modification term, as shown in the algorithm summary in Table 1. In other words, the adaptation of $\underline{\hat{\mathbf{h}}}(n)$ is performed in a transform domain with a significantly reduced number of dimensions $K \ll L$ (see also the illustration in Fig. 1).

Table 1: Adaptive system identification in compressive domains

Reconstruction matrix and input compression:

$$\mathbf{E}(n) = \begin{cases} \mathbf{I} & \text{for } n = 0 \\ \operatorname{diag}\left\{\left|\hat{\mathbf{h}}(n-1) + \boldsymbol{\epsilon}\right|\right\} & \text{for } n = 1, 2, \ldots \end{cases}$$

$$\Phi^{+}(n) = \left(\lambda \mathbf{E}^{-1}(n) + \Phi^{\mathrm{T}}\Phi\right)^{-1}\Phi^{\mathrm{T}}$$

$$\underline{\mathbf{x}}(n) = \Phi^{+^{\mathrm{T}}}(n)\mathbf{x}(n)$$

Adaptive filtering algorithm (e.g., RLS) in the compressed-input domain (i.e., expressed in terms of $\underline{\mathbf{h}}$, $\underline{\mathbf{x}}$, y, e):

$$e(n) = y(n) - \underline{\hat{\mathbf{h}}}^{\mathrm{T}}(n-1)\underline{\mathbf{x}}(n)$$

$$\mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}(n) = \alpha \mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}(n-1) + \underline{\mathbf{x}}(n)\underline{\mathbf{x}}^{\mathrm{T}}(n)$$

$$\underline{\hat{\mathbf{h}}}(n) = \underline{\hat{\mathbf{h}}}(n-1) + \mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}^{-1}(n)\underline{\mathbf{x}}(n)e(n)$$

Reconstruction of sparse coefficient vector:

$$\hat{\mathbf{h}}(n) = \Phi^{+}(n)\underline{\hat{\mathbf{h}}}(n)$$

II. Adaptive Filtering on Arbitrary Partly Smooth Manifolds and Relation to Compressive Domains

Having expressed the algorithm in Table 1 essentially by means of filter updates in a lower-dimensional *local* subspace (depending on the previous $\hat{\mathbf{h}}(n-1)$), we now further formalise this approach and generalise it to a manifold learning framework for arbitrary supervised, blind and semi-blind broadband adaptive filtering problems with multiple inputs and multiple outputs (MIMO). This unification is facilitated by utilising TRINICON ('TRIple-N Independent component analysis for CONvolutive mixtures'), a previously introduced general framework for broadband adaptive MIMO filtering, e.g., [9], [10], [11], [12].

Most of the well known basic adaptation algorithms are expressed in Euclidean space [1]. In contrast, a manifold is a topological space that is *locally* Euclidean. For smooth (i.e., differentiable) manifolds so-called *maps* relate smoothly to each other and are described by certain mapping functions [13], [14]. Hence, we may expect that the overall scheme in Tab. 1 can accurately be described as a specific form of manifold learning. In [15] the general TRINICON framework was introduced on arbitary partly smooth manifolds and it was shown to be a very versatile tool for taking into account prior information on the unknown system by simply plugging in a suitable map $\varphi_{\mathbf{W}'}(\Delta\mathbf{T})$, where $\mathbf{W}'$ is the matrix of current MIMO filter coefficients and $\Delta\mathbf{T}$ denotes the update in the local tangent space. Figures 3 and 4 outline the basic idea of the concept. Indeed, as we can show, by plugging in the map

$$\boldsymbol{\varphi}_{\mathbf{W}(n-1)}(\Delta\mathbf{T}) = \Phi^{+}(n)\Phi\mathbf{W}(n-1) + \Phi^{+}(n)\Delta\mathbf{T},$$

we immediately obtain a generalisation of Table 1 to arbitrary MIMO setups. Hence, in conclusion, the new manifold-based approach provides a useful interpretation as an adaptation in the tangent space where the compressed sensing theory shows us how to learn a suitable manifold from the input data. Using TRINICON, it also provides a very powerful generalisation for a wide area of applications.

FIGURES

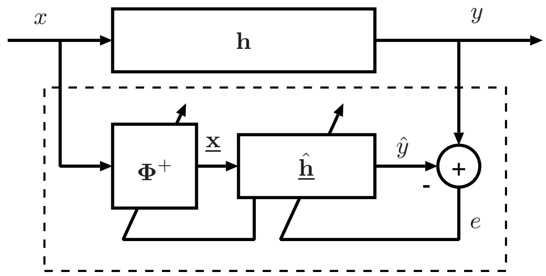

Fig. 1. Supervised single-channel adaptive system identification in compressive domain.

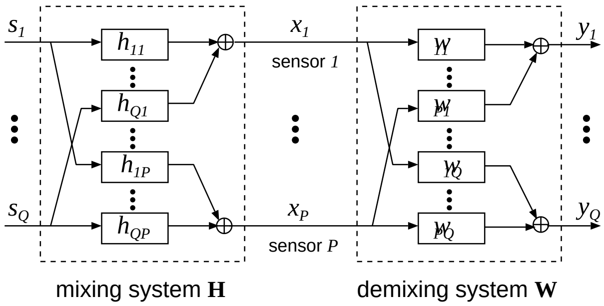

Fig. 2. General setup for MIMO signal processing.

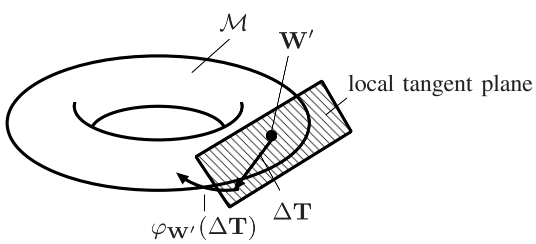

Fig. 3. Example for a two-dimensional manifold $\mathcal{M}$.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall Inc., Englewood Cliffs, NJ, 4th edition, 2002.

[2] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, May 2004, vol. 4, pp. 117–120.

[3] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, Academic Press, Boston, MA, 3rd edition, 2009.

[4] K. Helwani, H. Buchner, and S. Spors, "Multichannel adaptive filtering with sparseness constraints," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aachen, Germany, Sep. 2012.

[5] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[6] Y.C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.

[7] S. Foucart and H. Rauhut, *A mathematical Introduction to Compressive Sensing*, Springer, 2013.

[8] K. Helwani and H. Buchner, "Multichannel adaptive filtering in compressive domains," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Antibes, France, Sep. 2014.

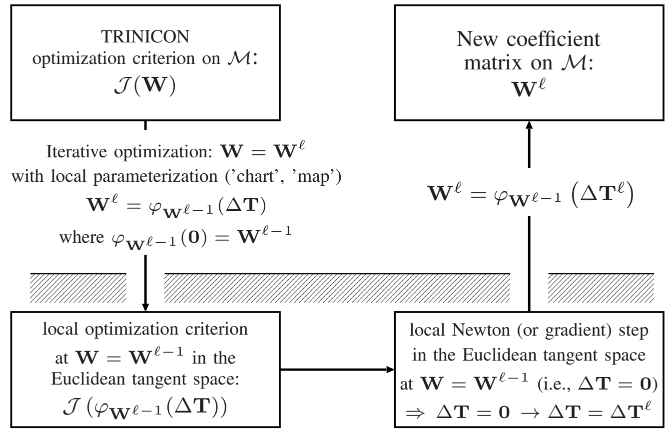

Fig. 4. Basic approach for TRINICON-based optimization on an arbitrary partly smooth manifold $\mathcal{M}$.

[9] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: A versatile framework for multichannel blind signal processing," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, May 2004, vol. 3, pp. 889–892.

[10] H. Buchner, R. Aichner, and W. Kellermann, "Blind source separation for convolutive mixtures: A unified treatment," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, J. Benesty and Y. Huang, Eds., pp. 255–293. Kluwer Academic Publishers, Boston, Apr. 2004.

[11] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON-based blind system identification with application to multiple-source localization and separation," in *Blind Speech Separation*, S. Makino, T.-W. Lee, and S. Sawada, Eds., pp. 101–147. Springer, Berlin, Sept. 2007.

[12] H. Buchner and W. Kellermann, "TRINICON for dereverberation of speech and audio signals," in *Speech Dereverberation*, P.A. Naylor and N.D. Gaubitch, Eds., pp. 311–385. Springer, London, Jul. 2010.

[13] M. Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, Harper Collins Publishers, 1965.

[14] W.M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York, 1986.

[15] H. Buchner, "A systematic approach to incorporate deterministic prior knowledge in broadband adaptive MIMO systems," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2010.