EFFICIENT ADAPTIVE FILTERING IN COMPRESSIVE DOMAINS FOR SPARSE SYSTEMS AND RELATION TO TRANSFORM-DOMAIN ADAPTIVE FILTERING

Herbert Buchner[1], *Karim Helwani*[2], *Bashar I. Ahmad*[1], *and Simon Godsill*[1]

[1] Information Engineering Division, University of Cambridge, Cambridge, UK
[2] Huawei European Research Center, Munich, Germany

ABSTRACT

In this paper we introduce a novel class of efficient multichannel adaptive filtering algorithms for sparse FIR systems. By suitably integrating ideas from compressed sensing and adaptive filter theory, this class of algorithms allows to significantly reduce the actual number of adaptive coefficients in an efficient way. These algorithms, termed compressive-domain adaptive filters, can be interpreted as a novel type of transform-domain techniques. They can also be seen as adaptive approach in an efficiently self-learning manifold based on the prior knowledge of sparseness of the system. An important property of this concept is that it does not place additional restrictions on the input signal characteristics. Based on the well-known RLS algorithm as a reference, the simulation results confirm that the proposed algorithm converges at acceptable rates, even for strongly colored signals such as speech and audio.

***Index Terms*—** Compressive domains, Adaptive filtering, System identification.

1. INTRODUCTION

Linear adaptive filters have found applications in diverse fields including communications, control, robotics, sonar, radar, seismics and biomedical engineering, to name a few [1]. The main classes of tackled problems (inverse modelling, prediction, linear prediction, and system identification) share a structure in which a cost function is minimized iteratively. In this paper, we mainly focus on the system identification problem, although the results will also carry over to the other classes of problems. Adaptive system identification has several application areas, such as acoustic echo cancellation (AEC), layered earth modelling (LEM), propagation channel estimation and others [1].

The complexity of the utilised adaptive filter typically depends on the length -dimensions of the target system impulse response, which can be excessively high. Nevertheless, in several application areas, such as network AEC, LEM and time domain reflectometry [2], the system impulse response can be sparse (i.e., only a small percentage of its coefficients has significant magnitudes) or compressible (magnitudes of the ordered coefficients are fast decaying). In this paper, we propose a novel multichannel adaptive filtering approach, which achieves a substantial reduction of the computational complexity of the system identification and improves convergence by effectively leveraging the underlying sparse nature of the sought system impulse response. Whilst the locations of the significant or non-zero coefficients of the impulse response are unknown *a priori*, the introduced method delivers the savings on complexity by transforming the adaptation problem into a lower dimensional system-dependent manifold.

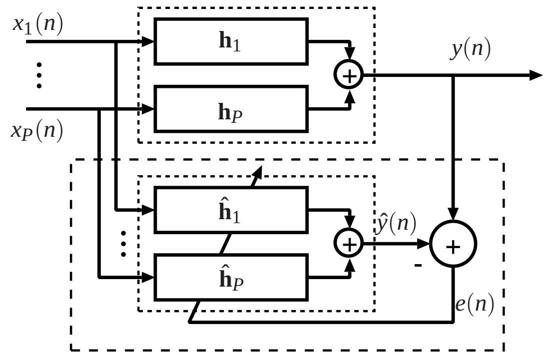

Fig. 1. Basic multichannel adaptive system identification setup.

1.1. Problem Formulation

We consider the multichannel problem depicted in Figure 1. Let P be the number of input channels over which the unknown system is excited; $x_p(n)$ is the input signal in the p^{th} channel at the discrete-time instant t_n. The sought system impulse response of the overall length $N = PL$, is denoted by $\mathbf{h} = \left[\mathbf{h}_1^{\text{T}}, \mathbf{h}_2^{\text{T}}, ..., \mathbf{h}_P^{\text{T}}\right]^{\text{T}}$ such that $\mathbf{h}_p = [h_{p,0}, h_{p,1}, ..., , h_{p,L-1}]^{\text{T}}$ pertains to the p^{th} subsystem in the multiple-input single-output (MISO) overall system model. Subsequently, the output at t_n is given by $y(n) = \mathbf{x}^{\text{T}}(n)\mathbf{h}$ with $\mathbf{x}(n) = [\mathbf{x}_1^{\text{T}}(n), \mathbf{x}_2^{\text{T}}(n), \cdots, \mathbf{x}_P^{\text{T}}(n)]^{\text{T}}$ and $\mathbf{x}_p(n) = [x_p(n), x_p(n-1), ..., x_p(n-L+1)]^{\text{T}}$. The objective in this paper is to adaptively estimate the system impulse response $\mathbf{h}$, which is assumed to be sparse/compressible, i.e. $|\text{supp}(\mathbf{h})| = \|\mathbf{h}\|_0 \leqslant K$ and $K \ll PL$.

1.2. Related Work and Outline

Most of the popular adaptive filtering algorithms are based on the least-squares error minimization, where the cost function is commonly defined by

$$J\left(\widehat{\mathbf{h}}(n)\right) := \hat{\mathcal{E}}\{e^2(n)\} = \hat{\mathcal{E}}\left\{\left(y(n) - \mathbf{x}^{\text{T}}(n)\widehat{\mathbf{h}}(n)\right)^2\right\}, \qquad (1)$$

such that $\hat{\mathcal{E}}\{.\}$ is an estimation of the expectation (usually a weighted sum over time) [1]. The impulse response estimate at t_n is denoted by $\widehat{\mathbf{h}}(n)$. It comprises PL MISO coefficient vector composed of the P subfilters $\widehat{\mathbf{h}}_p(n) = [\hat{h}_{p,0}(n), \hat{h}_{p,1}(n), \cdots, \hat{h}_{p,L-1}(n)]^{\text{T}}$. Unlike minimizing (1), in this paper we capitalise on the premise that the system impulse response $\mathbf{h}$ (not the input signal) is sparse/compressible to develop an efficient adaptive filtering technique.

Several studies consider regularized versions of (1) by incorporating known signal priors, perhaps the simplest and most popular example being the energy-based Tikhonov regularization [3]. Another very important class of priors are sparsity-promoting priors, e.g., [3, 4, 5, 6, 7]. For multichannel adaptive signal processing, structured hybrid norm regularizers have also been shown to be very effective [6]. In general, such approaches (regardless of the particular type of regularization) perform adaptation at each time step for all the components of $\widehat{\mathbf{h}}(n)$, including the zero components of the sparse vector. This can lead to a high computational cost and unnecessarily slow convergence. Conversely, the formulation introduced here addresses this issue by aiming to adapt only the non-zero coefficients of the impulse response vector (whose locations are not known).

On the other hand, the Compressed Sensing (CS) paradigm enables the concurrent sensing and compression of signals, which are sparse in an appropriate transform domain [8, 9, 10]. In CS, we have $\mathbf{y} = \mathbf{\Phi}\boldsymbol{\theta}$, where $\mathbf{y} \in \mathbb{C}^M$ are linear measurements of a target vector $\boldsymbol{\theta} \in \mathbb{C}^N$, $\|\boldsymbol{\theta}\|_0 \leqslant K$ and $K < N$. The sensing matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$ is typically dictated by the physical constraints of collecting the samples in $\mathbf{y}$. This matrix has to satisfy certain condition(s), e.g., Restricted Isometry Property (RIP), to guarantee the accurate recovery of $\boldsymbol{\theta}$ from $M < N$ observations via computationally tractable algorithms, e.g., convex relaxation and greedy methods [9, 10]. In some previous works, compressed sensing results were directly applied to the sparse channel estimation problem in communication systems by concatenating the measured system output (i.e. received signal) of the exciting signal (e.g. transmitted pilots) at multiple time instants as per $\mathbf{y} = \tilde{\mathbf{X}}\mathbf{h}$ where $\mathbf{y} = [y(n_0), y(n_0+1), \ldots, y(n_0+M)]^{\mathrm{T}}$ and $\tilde{\mathbf{X}} = [\mathbf{x}(n_0), \mathbf{x}(n_0+1), \ldots, \mathbf{x}(n_0+M)]^{\mathrm{T}}$ is an $M \times N$ Toeplitz matrix [11, 10, 12]. Hence, the propagation channel(s) impulse response $\mathbf{h}$ can be estimated by employing one of the CS recovery techniques. However, this is subject to $\tilde{\mathbf{X}}$, which is composed of the exciting signal, satisfying certain conditions such as RIP. Whilst in communication systems a wide range of signals can be transmitted during a training period (e.g. spread spectrum or OFDM signalling), in the majority of other applications (e.g. AEC or LEM) guaranteeing that the input signal meets stringent requirements, such as the RIP condition, can be impractical, overly restrictive and costly.

In this paper, we propose a novel compressive-domain adaptive filtering approach where the compression matrix can be chosen offline, without imposing any constraints on the nature or characteristics of $\mathbf{\Phi}$. Thus, the utilized compression matrix $\mathbf{\Phi}$ can be chosen such that it satisfies particular conditions related to performance guarantees, e.g., a RIP condition. Most importantly, this choice is independent of the input signal statistics. We recall that the objective here is to adaptively estimate the system impulse response $\mathbf{h}$, rather than recovering a sparse signal vector from its compressed measurements. Building upon our previous work [13, 14], we will also introduce in this paper a computationally efficient algorithm exploiting efficiently the lower dimensional manifold of the (initially unknown) sparse system, i.e., we compress the system impulse response to reduce its dimensionality based on the sparsity prior. In [14] it was demonstrated that this class of algorithms can indeed be considered as an adaptive manifold learning algorithm, which simultaneously identifies the sparse system and the corresponding lower-dimensional manifold. Additionally, the algorithm introduced below can be interpreted as a novel transform-domain adaptive algorithm in which the transformation is learned from the data. Due to the novel efficient formulation of this type of manifold learning/recovery algorithm here, the resulting realization has a substantially lower computational complexity compared with other benchmark adaptive techniques.

2. COMPRESSIVE DOMAINS FOR SPARSE SYSTEMS

Recent studies on compressive sensing state that a sparse signal, e.g., $\widehat{\mathbf{h}}$, can be perfectly reconstructed from its undersampled version

$$\underline{\widehat{\mathbf{h}}} := \mathbf{\Phi}\widehat{\mathbf{h}} \tag{2}$$

with $\mathbf{\Phi}$ a random observation matrix [15]. The originality of the theory of compressed sensing bases on its implicit statement that a subspace spanned by

$$M = O(K \log_2(PL/K)) \tag{3}$$

uncorrelated white vectors is dense in the space of K-sparse signals of length PL [16]. This motivates exploring the possibility of formulating an adaptive filtering solution for sparse systems in compressive domains with random compression matrices and without an explicit knowledge about the relevant support of the sparse systems. Since the compression matrix is not given by an explicit eigenspace of the sparse system, we will formulate the adaptive algorithm via the reconstruction approach from compressed sensing. An optimal reconstruction by a transformation matrix can be obtained from a typical compressed sensing cost function which is based on exploiting the sparsity of the system, given by

$$J'\left(\widehat{\mathbf{h}}(n)\right) = \lambda \left\|\widehat{\mathbf{h}}(n)\right\|_1 + \left\|\underline{\widehat{\mathbf{h}}}(n) - \mathbf{\Phi}\widehat{\mathbf{h}}(n)\right\|_2^2, \tag{4}$$

where λ denotes the Lagrange-multiplier. A minimum of the cost function can be found by setting its gradient w.r.t $\widehat{\mathbf{h}}$ to zero.

With

$$\left\|\widehat{\mathbf{h}}\right\|_1 = \mathrm{sgn}\left\{\widehat{\mathbf{h}}\right\}^{\mathrm{T}}\widehat{\mathbf{h}},$$

hereby, $\mathrm{sgn}\{\cdot\}$ stands for the sign function. Hence, the gradient reads

$$\nabla_{\widehat{\mathbf{h}}} J' = 2\lambda \mathrm{sgn}\left\{\widehat{\mathbf{h}}(n)\right\} - 2\mathbf{\Phi}^{\mathrm{T}}\left[\underline{\widehat{\mathbf{h}}}(n) - \mathbf{\Phi}\widehat{\mathbf{h}}(n)\right] = \mathbf{0},$$

i.e., the reconstruction of the coefficient vector $\hat{\mathbf{h}}$ is given as the solution of the (nonlinear) system of equations

$$\lambda \mathrm{sgn}\left\{\widehat{\mathbf{h}}(n)\right\} + \mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\widehat{\mathbf{h}}(n) = \mathbf{\Phi}^{\mathrm{T}}\underline{\widehat{\mathbf{h}}}(n). \tag{5}$$

3. COMPRESSIVE-DOMAIN ADAPTIVE FILTERING AS SELF-LEARNING TRANSFORM-DOMAIN ADAPTIVE FILTERING

The sign function in (5) can be approximated by

$$\mathrm{sgn}\left\{\widehat{\mathbf{h}}\right\} = \mathbf{E}^{-1}\widehat{\mathbf{h}}, \quad \text{with} \quad \mathbf{E} := \mathrm{diag}\left\{\left|\widehat{\mathbf{h}}\right| + \varepsilon\right\}, \tag{6}$$

where ε is a parameter that prevents a division by zero. By substituting this approximation into (5), we obtain

$$\widehat{\mathbf{h}}(n) = \left(\lambda \mathbf{E}^{-1}(n) + \mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^{\mathrm{T}}\underline{\widehat{\mathbf{h}}}. \tag{7}$$

Hence, we can now write the reconstruction process by a multiplication of $\underline{\widehat{\mathbf{h}}}$ with a reconstruction matrix that we define as

$$\mathbf{\Phi}^{+}(n) := \left(\lambda \mathbf{E}^{-1}(n) + \mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^{\mathrm{T}}. \tag{8}$$

Since $\widehat{\mathbf{h}}$ is *a priori* unknown, an iterative computation for the regularization matrix $\mathbf{E}^{-1}$ is performed, i.e.,

$$\mathbf{E}(n) = \mathrm{diag}\left\{\left|\widehat{\mathbf{h}}(n-1)\right| + \varepsilon\right\}, \tag{9}$$

where $\mathbf{E}(0)$ is set to the unity matrix. Thereby, the reconstruction matrix can be understood as an adaptive backtransformation matrix from the domain where the system has a compressed dense representation to the domain where it is sparsely represented. This backtransformation matrix is adaptively adjusted to the sparse structure of the system.

Having defined the reconstruction matrix $\mathbf{\Phi}^{+}(n)$ according to (8), i.e., $\widehat{\mathbf{h}}(n) = \mathbf{\Phi}^{+}(n)\underline{\widehat{\mathbf{h}}}(n)$, we can now express the (uncompressed) output signal of the adaptive filter as follows:

$$\hat{y}(n) = \widehat{\mathbf{h}}^{\mathrm{T}}(n)\mathbf{x}(n) = \left[\mathbf{\Phi}^{+}(n)\underline{\widehat{\mathbf{h}}}(n)\right]^{\mathrm{T}}\mathbf{x}(n) = \underline{\widehat{\mathbf{h}}}^{\mathrm{T}}(n)\mathbf{\Phi}^{+\mathrm{T}}(n)\mathbf{x}(n).$$

By introducing the transformed input vector

$$\underline{\mathbf{x}}(n) := \mathbf{\Phi}^{+\mathrm{T}}(n)\mathbf{x}(n), \tag{10}$$

we can finally express the output signal $\hat{y}(n)$ as

$$\hat{y}(n) = \underline{\widehat{\mathbf{h}}}^{\mathrm{T}}(n)\underline{\mathbf{x}}(n). \tag{11}$$

The actual adaptive filter optimization can then be expressed/performed completely in the corresponding transform domain. We can write the original cost function (1) of the adaptive filter equivalently as

$$J\left(\underline{\widehat{\mathbf{h}}}(n)\right) = \hat{\mathcal{E}}\left\{(y(n) - \hat{y}(n))^2\right\}, \tag{12}$$

with the new definition $\hat{y}(n) = \underline{\widehat{\mathbf{h}}}^{\mathrm{T}}(n)\underline{\mathbf{x}}(n)$. Note, however, that in each adaptation step, the reconstruction matrix $\mathbf{\Phi}^{+}(n)$ depends on the previous coefficient vector $\widehat{\mathbf{h}}(n)$ via $\mathbf{E}(n)$. In other words, the optimization is performed in each step n for a given $\mathbf{\Phi}^{+}(n)$ in a local Euclidean space. This mechanism is precisely in line with the manifold learning framework [17], as illustrated in [14].

The least-squares solution in the compressed domain, i.e., in the local Euclidean space, is given as

$$\underline{\widehat{\mathbf{h}}}_{\mathrm{opt}}(n) = \mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}^{-1}(n)\mathbf{r}_{\underline{\mathbf{x}}y}(n). \tag{13}$$

Typically, the correlation matrix is estimated iteratively using the formula

$$\mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}(n) = \alpha\,\mathbf{R}_{\underline{\mathbf{x}}\underline{\mathbf{x}}}(n-1) + \underline{\mathbf{x}}(n)\underline{\mathbf{x}}^{\mathrm{T}}(n), \tag{14}$$

where α denotes a forgetting factor. This leads to the well-known recursive least-squares (RLS) algorithm [1], applied here in the lower-dimensional compressed domain. The uncompressed estimated filter coefficient vector at step n is then given by

$$\widehat{\mathbf{h}}(n) = \mathbf{\Phi}^{+}(n)\underline{\widehat{\mathbf{h}}}(n). \tag{15}$$

It should be mentioned that instead of the RLS algorithm, we can essentially apply any adaptive filtering algorithm in the compressive domain. This is due to the fact that through $\hat{y}(n) = \underline{\widehat{\mathbf{h}}}^{\mathrm{T}}(n)\underline{\mathbf{x}}(n)$ the cost function J can always be expressed exclusively by the compressed parameter vector $\underline{\widehat{\mathbf{h}}}(n)$. In general, the zeros of $\nabla_{\underline{\widehat{\mathbf{h}}}}J$ can be determined iteratively with the Newton algorithm. The main advantage of Newton-type adaptation algorithms is its quadratic convergence rate compared to the linear convergence rate of the gradient-based algorithms [1].

4. EXACT REDUCED-COMPLEXITY ALGORITHM FOR COMPRESSIVE-DOMAIN ADAPTIVE FILTERING

Using the matrix inversion lemma in the form

$$\left(\mathbf{A} + \mathbf{B}\mathbf{B}^{\mathrm{T}}\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{I} + \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}, \tag{16}$$

we can reformulate (8) as

$$\mathbf{\Phi}^{+} = \frac{1}{\lambda}\mathbf{E}\mathbf{\Phi}^{\mathrm{T}} - \frac{1}{\lambda}\mathbf{E}\mathbf{\Phi}^{\mathrm{T}}\left(\mathbf{I} + \frac{1}{\lambda}\mathbf{\Phi}\mathbf{E}\mathbf{\Phi}^{\mathrm{T}}\right)^{-1}\mathbf{\Phi}\frac{1}{\lambda}\mathbf{E}\mathbf{\Phi}^{\mathrm{T}}. \tag{17}$$

Note that instead of the inversion of a $PL \times PL$ matrix in (8), this equation only requires an inversion of a $K \times K$ matrix, where $K \ll PL$ in sparse systems. By introducing the two intermediate quantities

$$\mathbf{D} := \frac{1}{\lambda}\mathbf{E}\mathbf{\Phi}^{\mathrm{T}} = \frac{1}{\lambda}\mathrm{diag}\left\{\left|\hat{\mathbf{h}}(n-1)\right| + \boldsymbol{\varepsilon}\right\}\mathbf{\Phi}^{\mathrm{T}} \tag{18}$$

and

$$\mathbf{C} := \mathbf{\Phi}\mathbf{D}, \tag{19}$$

we obtain

$$\mathbf{\Phi}^{+} = \mathbf{D} - \mathbf{D}(\mathbf{I} + \mathbf{C})^{-1}\mathbf{C}. \tag{20}$$

The resulting algorithm for adaptive system identification in compressive domains is summarized in Table 1. Figure 2 illustrates this scheme. In the next section, it is demonstrated that the proposed class of algorithms can significantly reduce the complexity of the coefficient estimation compared with other benchmark techniques, namely the RLS.

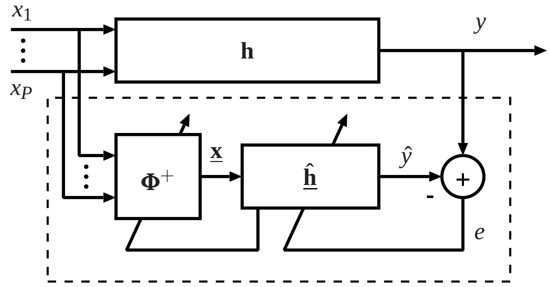

Fig. 2. Multichannel adaptive system identification in compressive domain.

5. EXPERIMENTS

As a proof of concept for the novel compressive domain adaptive filtering method we simulate a 2-channel MISO system identification scenario. Here a notably sparse system environment was simulated by only $K = 8$ randomly weighted pulses, i.e., 4 pulses in each channel, and the overall filter length was set to $L = 1000$ for each channel. The pulses were randomly distributed over the entire filter length L. The channels were excited by an audio signal at a sampling rate 44.1kHz (as an example for colored system excitation). Additive white noise was added to the captured signal to obtain an SNR of 60dB in $y(n)$.

Table 1: Novel Compressive-Domain Adaptive Filter Algorithm

Initialization:

$$\begin{aligned} \underline{\mathbf{h}} &= \mathbf{0} \\ \mathbf{\Phi} &= \mathrm{randn}(M, PL), \text{ where } M \ll PL \end{aligned}$$

Reconstruction matrix and input compression:

$$\begin{aligned} \mathbf{D}(n) &= \begin{cases} \frac{1}{\lambda}\mathbf{\Phi}^{\mathrm{T}} & \text{for } n = 0 \\ \mathrm{diag}\left\{\left|\hat{\mathbf{h}}(n-1)\right| + \boldsymbol{\varepsilon}\right\}\mathbf{D}(0) & \text{for } n = 1, 2, \ldots \end{cases} \\ \mathbf{C}(n) &= \mathbf{\Phi}\mathbf{D}(n) \\ \mathbf{\Phi}^{+}(n) &= \mathbf{D}(n) - \mathbf{D}(n)\left(\mathbf{I} + \mathbf{C}(n)\right)^{-1}\mathbf{C}(n) \\ \underline{\mathbf{x}}(n) &= \mathbf{\Phi}^{+\mathrm{T}}(n)\mathbf{x}(n) \end{aligned}$$

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

$$\begin{aligned} e(n) &= y(n) - \hat{\underline{\mathbf{h}}}^{\mathrm{T}}(n-1)\underline{\mathbf{x}}(n) \\ \mathbf{R}_{\underline{\mathbf{xx}}}(n) &= \alpha\mathbf{R}_{\underline{\mathbf{xx}}}(n-1) + \underline{\mathbf{x}}(n)\underline{\mathbf{x}}^{\mathrm{T}}(n) \\ \hat{\underline{\mathbf{h}}}(n) &= \hat{\underline{\mathbf{h}}}(n-1) + \mathbf{R}_{\underline{\mathbf{xx}}}^{-1}(n)\underline{\mathbf{x}}(n)e(n) \end{aligned}$$

Reconstruction of sparse coefficient vector:

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

Figure 3 shows a comparison between coefficient misalignment convergence curves, i.e., the normalized ℓ_2 norm of the coefficient error in dB. As a reference, the blue (dashed) curve shows the performance achieved by the original two-channel RLS algorithm with $L = 1000$ coefficients for each channel (i.e., non-compressive adaptation). The green (solid) curve shows the convergence of the adaptation of the RLS algorithm in the compressed domain according to Table 1. In contrast to the noncompressed case, the number of adaptive filter parameters was reduced significantly from $PL = 2000$ to $M = 50$, i.e, in the compressive domain we used only 25 coefficients for each channel. Despite this reduction of the coeficients (and the associated complexity reduction from $O(P^2L^2)$ to $O(M^2)$ in this case), we were able to maintain the low final misalignment level of approximately -60dB. Note that in order to take into account the possible variability of the results by the different initializations of $\mathbf{\Phi}$, the green curve was averaged over the results of 20 simulation runs. The width of the 95% confidence interval was ± 4.7dB over the simulated time span. The number $M = 50$ of required coefficients for our scenario was found experimentally, and it lies roughly at the order of the predicted value $M = \lceil K \log_2(PL/K) \rceil = 64$ according to (3).

6. CONCLUSION

In this paper we introduced a novel class of efficient multichannel adaptive filtering algorithms in compressive domains. This approach can be interpreted as a novel type of transform-domain algorithms or, more generally, as an adaptive technique in an efficiently self-learning manifold. For sparse systems, it was demonstrated that the number of adaptive coefficients can be reduced significantly in this way, while the original convergence characteristics can be maintained (or in some cases even be improved due to the reduced number of coefficients to be learned). The presented simulation results show that the concept of compressive domain adaptive filtering is efficiently realizable and at acceptable convergence rates even for colored signals such as speech and audio.

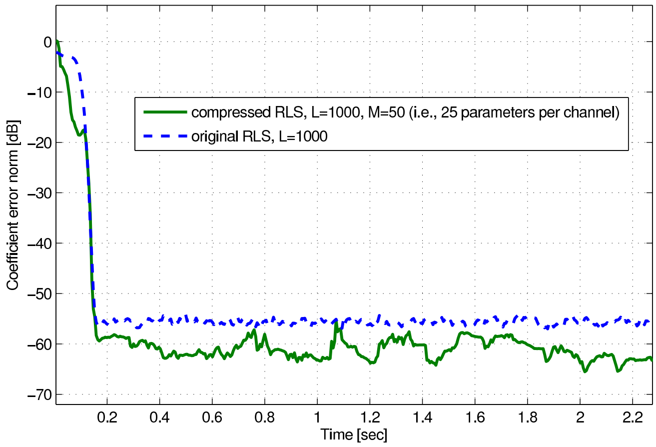

Fig. 3. Comparison of identification performance for the proposed algorithm and the RLS method.

7. REFERENCES

[1] S. Haykin, *Adaptive filter theory*, Prentice Hall, Inc., 1991.

[2] S.B. Jones, J.M. Wraith, and D. Or, "Time domain reflectometry measurement principles and applications," *Hydrological processes*, vol. 16, no. 1, pp. 141–153, 2002.

[3] C.M. Bishop, *Pattern recognition and machine learning*, springer, 2006.

[4] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[5] J. Benesty and S.L. Gay, "An improved pnlms algorithm," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2002, vol. 2, pp. II–1881.

[6] K. Helwani, H. Buchner, and S. Spors, "Multichannel adaptive filtering with sparseness constraints," in *International Workshop on Acoustic Signal Enhancement (IWAENC)*. VDE, 2012.

[7] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for \ell_1-minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.

[8] D.L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory,*, vol. 52, no. 4, pp. 1289–1306, 2006.

[9] J. Tropp and S.J. et al Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[10] M.F. Duarte and Y.C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[11] W.U. Bajwa, J. Haupt, A.M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse

multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[12] X. Rao and V.K.N. Lau, "Distributed compressive csit estimation and feedback for fdd multi-user massive mimo systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, 2014.

[13] K. Helwani and H. Buchner, "Multichannel adaptive filtering in compressive domains," in *Conf. Rec. Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014.

[14] H. Buchner and K. Helwani, "Adaptive dynamical systems in compressive domains as a manifold learning framework," in *Proc. Int. Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, 2015.

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

[16] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[17] H. Buchner, "A systematic approach to incorporate deterministic prior knowledge in broadband adaptive MIMO systems," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, 2010.