KERNEL-BASED IDENTIFICATION OF HAMMERSTEIN SYSTEMS FOR NONLINEAR ACOUSTIC ECHO-CANCELLATION

*Steven Van Vaerenbergh**

Dept. of Communications Engineering
University of Cantabria, Spain
steven@gtas.dicom.unican.es

Luis A. Azpicueta-Ruiz†

Dept. of Sig. Theory and Communications
Universidad Carlos III de Madrid, Spain
azpicueta@tsc.uc3m.es

ABSTRACT

Traditional acoustic echo cancelers use a linear model to represent the echo path. Nevertheless, many consumer devices include loudspeakers and audio power amplifiers that may generate significant nonlinear distortions, creating the need for acoustic echo cancelers to produce a nonlinear filter response. To address this issue, we propose a nonlinear acoustic echo cancellation algorithm based on the framework of kernel methods. We model the echo path as a Hammerstein system, and we propose a resource-efficient strategy to identify the nonlinear and linear parts. While the basic algorithm is presented as an iterative batch method, we show that a simple extension allows it to be used in online scenarios as well. Results for both types of scenarios show that the algorithm produces good results on a system with a clipping nonlinearity and a realistic room impulse response.

***Index Terms*—** acoustic echo cancellation, nonlinear distortions, kernel methods, Hammerstein systems

1. INTRODUCTION AND PROBLEM STATEMENT

The widespread utilization of smartphones, laptops and other devices has modified the conception of today's communication systems, considering hands-free operation and mobility a must. In order to avoid acoustic echoes that may turn the communication uncomfortable, teleconferencing and videoconferencing applications used on these devices include an acoustic echo canceler (AEC) [1].

Fig. 1 describes a general setup of the acoustic echo cancellation problem. We have access to the speech or audio signal $x(n)$, which is generated by the far-end user and fed into the loudspeaker. The microphone outputs a signal $d(n) = y_h(n) + e_0(n)$ where $y_h(n) = x(n) * h(n)$ is the echo signal to be canceled, $h(n)$ represents the impulse response that describes the echo path, $*$ denotes convolution and $e_0(n)$ is the background noise at the microphone position. The goal of the AEC is to estimate a copy of such echo signal, denoted as $y(n)$, seeking to minimize the power of the cancellation error $e(n) = d(n) - y(n) = y_h(n) - y(n) + e_0(n)$ to avoid that the far-end user receives a delayed and modified version of his/her voice $x(n)$. Since the echo path $h(n)$ depends highly on different factors that affect the acoustic propagation in the room and that can vary with time (such as the position of the microphone, loudspeaker, objects in the room and temperature), AECs include an adaptive filter as a key component to adaptively cancel the echo.

*The work of S. Van Vaerenbergh was partly supported by MINECO grants TEC2010-19545-C04-03 (COSIMA) and CONSOLIDER-INGENIO 2010 CSD2008-00010 (COMONSENS).

†The work of L.A. Azpicueta-Ruiz was partly supported by MINECO projects TEC2011-22480 and PRI-PIBIN-2011-1266.

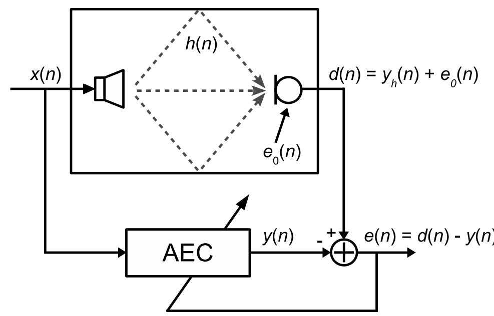

Fig. 1. General setup of the acoustic echo cancellation problem.

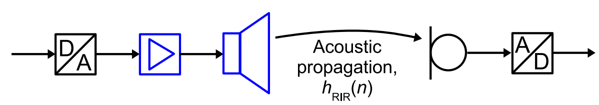

Fig. 2. Block diagram of the systems involved in the echo path, where the main sources of nonlinear distortion (the audio power amplifier and the loudspeaker) have been blue-highlighted.

However, the aforementioned devices usually include low-cost loudspeakers with amplifiers driven at high power levels, which cause non-negligible and easily perceived nonlinear distortions. Traditional AECs are based on linear adaptive filters, which rely on the assumption of a purely linear echo path. In the presence of nonlinear distortions these linear AECs provide insufficient cancellation, and they should be extended to include nonlinear models that enable them to eliminate the nonlinear distortions [2–4].

Fig. 2 represents a block diagram of an echo path including the typical electroacoustic chain: Digital-to-analog converter, power amplifier and loudspeaker at the emitter side; acoustic propagation in the room with room impulse response (RIR) $h_{\mathrm{RIR}}(n)$; and microphone and analog-to-digital converter at the receptor side. Though all of these elements are potential sources of nonlinearities (excepting the acoustic propagation denoted as $h_{\mathrm{RIR}}(n)$ that is linear in nature), the most important nonlinear distortions are caused by the action of the loudspeaker and the power amplifier.

Many techniques have been proposed in the literature on nonlinear acoustic echo cancelation (NAEC). One of the most popular approaches in the last years is based on adaptive Volterra filters (VFs) [5–7]. However, the very high computational cost associated with the operation of VFs limits its use only to cases where the order of the nonlinearity is quadratic or cubic at most [8, 9]. Some recent improvements that lower the computational complexity of the basic VF scheme can be found in [3, 10].

In addition to VFs, several different structures have been proposed to model nonlinear distortions, such as even mirror Fourier nonlinear filters [11], block-based Wiener-Hammerstein models [12, 13], functional link adaptive filters [14], and nonlinear transformations based on different curves such as polynomial saturation curves [15], sigmoid functions [16], and spline functions [17].

More recently, kernel methods [18] have had considerable success in nonlinear signal processing. Nevertheless, they have not yet been applied widely to real-time applications such as NAEC, mainly due to the challenges present in designing online and adaptive kernel algorithms. In particular, kernel methods typically have computational requirements that exceed the budget of real-time implementations [19], and *online* kernel methods require to include specific mechanisms to avoid that their computational requirements grow unboundedly during operation [20]. An online kernel-based NAEC was recently proposed in [21]. It modeled the acoustic echo path as a black-box input-output system, for which it used a specifically tuned kernel with a linear and a nonlinear part, and to maintain its computational complexity fixed it used a sliding-window strategy.

In this paper we follow a different kernel-based approach. We model the echo path as a Hammerstein system, which is motivated by Fig. 2, and we design a kernel-based identification algorithm for this model. By using a Hammerstein model we take away some of the flexibility that exists in black-box approaches, which in turn allows us to learn the input-output relationship with far fewer data and therefore less computational resources.

The rest of the paper is organized as follows: Section 2 includes a detailed explanation of the proposed Hammerstein system identification algorithm based on kernel methods, which will be experimentally assessed in Section 3 in a typical nonlinear acoustic echo cancellation scenario. Section 4 contains the conclusions of this work as well as an advance of the future research lines.

2. KERNEL-BASED IDENTIFICATION OF HAMMERSTEIN SYSTEMS

We will model both parts of the Hammerstein system separately: For the nonlinear part we use a regression technique based on kernel methods, and for the linear part we employ standard filtering theory. The proposed algorithm aims to learn both representations jointly.

2.1. Kernel methods and ridge regression

Kernel methods are based on a nonlinear transformation of the input data x, scalars in our case, into a high-dimensional *feature space*, in which it is more likely that the transformed data $\Phi(x)$ are linearly separable. Due to its high dimensionality, however, it is impractical to perform explicit calculations in feature space. Nevertheless, inner products in this space can be calculated by using a positive definite kernel function satisfying Mercer's condition [22]: $\kappa(x,x') = \langle\Phi(x),\Phi(x')\rangle$. This idea, known as the "kernel trick", allows to perform inner-product based algorithms implicitly in feature space by replacing all inner products with kernels in the input space. A commonly used kernel function is the Gaussian kernel

$$\kappa(x,x') = \exp(-|x-x'|^2/2\sigma^2), \tag{1}$$

where σ is the *kernel width*.

Given a set of N input-output data pairs $\{x(n), d(n)\}_{n=1}^N$, the technique of kernel ridge regression (KRR) [23] aims to obtain the projection $\mathbf{w}$ that minimizes the cost function

$$J_{\text{KRR}} = \sum_{n=1}^{N} |d(n) - \Phi(x(n))^\top \mathbf{w}|^2 + c\mathbf{w}^\top\mathbf{w}, \tag{2}$$

where c is a Tikhonov regularization constant. Since $\mathbf{w}$ lies in the high-dimensional feature space, it is impractical to calculate explicitly. Fortunately, the Representer Theorem [18] states that $\mathbf{w}$ can be expressed as a linear combination of the transformed training data $\Phi(x(n))$, which, in turn, allows to employ the kernel trick.

In order to reduce the complexity of the method we will use only a subset of all training data to represent $\mathbf{w}$, specifically

$$\mathbf{w} = \sum_{m=1}^{M} \alpha_m \Phi(x_m^s), \tag{3}$$

and we will refer to the chosen points x_m^s as the *support points*. If all training data are used for the support, i.e., $M = N$, the obtained solution is optimal. For $M < N$ an approximation to the optimal $\mathbf{w}$ is obtained. Based on Eq. (3) and the kernel trick, the output of the estimated nonlinear function for a test input x' becomes

$$f(x') = \Phi(x')^\top\mathbf{w} = \sum_{m=1}^{M} \kappa(x', x_m^s)\alpha_m, \tag{4}$$

By substituting (3) in (2) and adopting a matrix notation, the cost function (2) reads

$$J_{\text{KRR}} = \|\mathbf{d} - \mathbf{K}\boldsymbol{\alpha}\|^2 + c\boldsymbol{\alpha}^\top\mathbf{K}_s\boldsymbol{\alpha}, \tag{5}$$

where $\mathbf{K}$ is the kernel matrix with elements $\mathbf{K}_{nm} = \kappa(x(n), x_m^s)$, $\mathbf{d} = [d(1), d(2), \ldots, d(N)]^\top$ contains the stacked output data, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]^\top$, and $\mathbf{K}_s$ is the $M \times M$ kernel matrix calculated only for the support points. The solution is obtained as

$$\boldsymbol{\alpha} = (\mathbf{K}^\top\mathbf{K} + c\mathbf{K}_s)^{-1}\mathbf{K}^\top\mathbf{d}. \tag{6}$$

2.2. Batch identification algorithm

As motivated in Section 1, we model the observed echo path as a Hammerstein system, i.e., a cascade of a nonlinearity and a linear filter. The output of the microphone then becomes

$$d(n) = h(n) * f(x(n)) + e_0(n), \tag{7}$$

where $h(n)$ is the linear impulse response, including the impulse response of the echo path $h_{\text{RIR}}(n)$ and other linear parts, and $f(\cdot)$ represents the nonlinear distortion. In order to identify both parts of the system, we propose to minimize the following KRR-based cost function:

$$J = \|\mathbf{d} - \mathbf{h} * \mathbf{K}\boldsymbol{\alpha}\|^2 + c_\alpha\boldsymbol{\alpha}^\top\mathbf{K}_s\boldsymbol{\alpha} + c_h\mathbf{h}^\top\mathbf{h}, \tag{8}$$

with respect to $\boldsymbol{\alpha}$ and $\mathbf{h}$. Here, $\mathbf{d} = [d(1), \ldots, d(N)]^\top$, $\mathbf{h} = [h(1), \ldots, h(L)]^\top$, and the operator $*$ refers to the convolution of $\mathbf{h}$ on the columns of $\mathbf{K}$. In contrast to the previously introduced formulations, this cost function does not have a closed form solution. Nevertheless, if an estimate of the linear coefficients, $\hat{\mathbf{h}}$, were available, it would be possible to obtain the corresponding coefficients $\hat{\boldsymbol{\alpha}}$ in closed form. In this case, the last term can be discarded as it does not affect the minimization, and the cost function reduces to

$$J_{\boldsymbol{\alpha}} = \|\mathbf{d} - \mathbf{K}_h\hat{\boldsymbol{\alpha}}\|^2 + c_\alpha\hat{\boldsymbol{\alpha}}^\top\mathbf{K}_s\hat{\boldsymbol{\alpha}}, \tag{9}$$

Algorithm 1 Kernel-based identification of Hammerstein systems (KIHAM)

Initialize $\hat{\mathbf{h}}$ through Eq. (14)
while J not converged **do**
  Update $\mathbf{K}_h$, and update $\hat{\boldsymbol{\alpha}}$ through Eq. (10)
  Update $\mathbf{K}_\alpha$ and $\hat{\mathbf{h}}$ through Eq. (12) and Eq. (13)
end while
Output: $\hat{\boldsymbol{\alpha}}$ and $\hat{\mathbf{h}}$

where $\mathbf{K_h} = \hat{\mathbf{h}} * \mathbf{K}$. The solution is given by

$$\hat{\boldsymbol{\alpha}} = (\mathbf{K}_h^\top \mathbf{K}_h + c_\alpha \mathbf{K}_s)^{-1} \mathbf{K}_h^\top \mathbf{d}. \quad (10)$$

In turn, if an estimate $\hat{\boldsymbol{\alpha}}$ were available, then $\hat{\mathbf{h}}$ could be obtained in a similar fashion from Eq. (8). The cost function now reduces to

$$J_\mathbf{h} = \|\mathbf{d} - \mathbf{K}_\alpha \hat{\mathbf{h}}\|^2 + c_h \|\hat{\mathbf{h}}\|^2, \quad (11)$$

where $\mathbf{K}_\alpha$ is an $N \times L$ matrix that contains the elements of the vector $\mathbf{k}_\alpha = \mathbf{K}\hat{\boldsymbol{\alpha}}$ on its columns,

$$\mathbf{K}_\alpha = \begin{bmatrix} k_\alpha(1) & 0 & \dots & 0 \\ k_\alpha(2) & k_\alpha(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ k_\alpha(N) & k_\alpha(N-1) & \dots & k_\alpha(N-L+1) \end{bmatrix}. \quad (12)$$

The solution is obtained as

$$\hat{\mathbf{h}} = (\mathbf{K}_\alpha^\top \mathbf{K}_\alpha + c_h \mathbf{I})^{-1} \mathbf{K}_\alpha^\top \mathbf{d}. \quad (13)$$

This suggests an iterative algorithm that alternates between updating estimates of the linear channels $\mathbf{h}$ and the coefficients $\boldsymbol{\alpha}$ of the nonlinearity. Convergence is guaranteed because each update may either decrease or maintain the cost function Eq. (8). The algorithm is summarized in Alg. 1. Initialization may be performed by setting the coefficients $\mathbf{h}$ to their linear solution,

$$\hat{\mathbf{h}} = (\mathbf{X}^\top \mathbf{X} + c_h \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{d}, \quad (14)$$

where $\mathbf{X}$ contains time-embedded vectors of the data $x(n)$, stacked as rows. A Matlab implementation of this algorithm is available at `http://gtas.unican.es/people/steven`.

2.3. Online strategy

We propose a simple extension of the batch algorithm for application in online and adaptive scenarios. The main source of nonstationarity in the echo path is the acoustic propagation described by the RIR, which is comprised in the linear part of the Hammerstein system model. The nonlinear distortions show only very small and very slow changes that may typically be discarded. Therefore, once the nonlinearity has been estimated satisfactorily using the batch algorithm, it is reasonable to perform tracking only on the linear part. The outline of the online strategy is as follows:

1. Start performing AEC by a standard linear adaptive filter on $x(n)$ and $d(n)$, obtaining the linear filter $\hat{\mathbf{h}}(n)$.
2. At the same time, fill a buffer with $x(n)$ and $d(n)$.
3. When the buffer is filled with N_b input and output data, apply the batch algorithm to obtain $\hat{\boldsymbol{\alpha}}(n)$ and a new estimate of the linear filter, $\hat{\mathbf{h}}(n)$, which replaces the current estimate.
4. Continue adapting the linear filter, now using input $f(x(n))$, calculated through Eq. (4), and output $d(n)$.

The nonlinearity can be re-estimated periodically if significant variations in its response are expected.

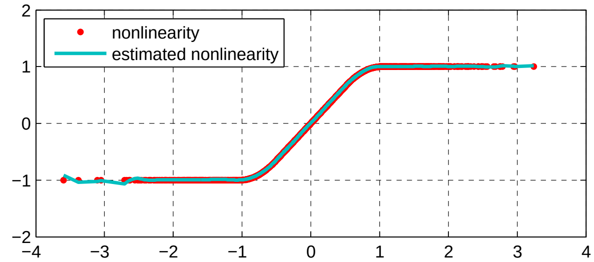

Fig. 3. Input vs. output of the nonlinear distortion, and input vs. estimated output, in the offline experiment. Note that the true values of the output are not directly observed.

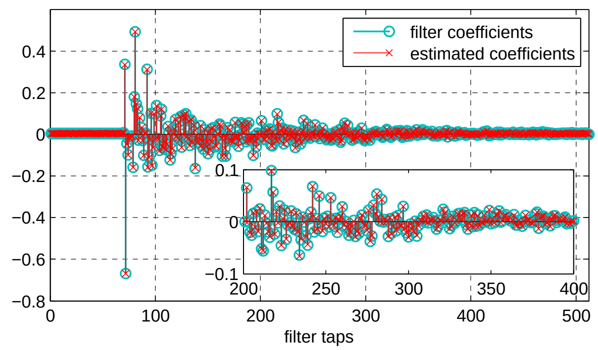

Fig. 4. True and estimated impulse response coefficients of the RIR, in the offline experiment. A zoom from tap 200 to 400 is included.

3. EXPERIMENTAL EVALUATION

We now assess the capabilities of the proposed scheme in a typical nonlinear echo cancellation scenario. To this end, we consider a setup where the main source of nonlinearity is due to the clipping of the power amplifier. The used clipping function is shown in Fig. 3. The signal $x(n)$ that is fed into the amplifier is generated as USASI noise with a speech-like spectrum [24]. The room impulse response that describes the acoustic propagation $\mathbf{h}_{\text{RIR}}(n)$ has been measured in a typical office and has been limited to 512 taps (considering a sampling frequency of 8 kHz). Background noise $e_0(n)$ is generated as zero-mean white noise, uncorrelated with $x(n)$, and its variance is set to obtain a signal to noise ratio of 30 dB at microphone location.

The figure of merit employed to measure the behavior of the algorithms is the echo return loss enhancement, expressed in decibels and defined as

$$\text{ERLE}(n) := 10 \cdot \log_{10} \frac{\text{E}\{d^2(n)\}}{\text{E}\{e^2(n)\}}. \quad (15)$$

We evaluate the performance of the proposed algorithm considering an offline and an online scenario.

3.1. Offline cancellation

In the offline scenario we consider that $N = 2048$ data pairs $\{x(n), d(n)\}_{n=1}^N$ are available. To these data we apply the proposed KIHAM algorithm, as described in Alg. 1, with a Gaussian kernel. We choose the number of support points to be $M = 50$, and we place them equally spaced over the range of the input data, resulting

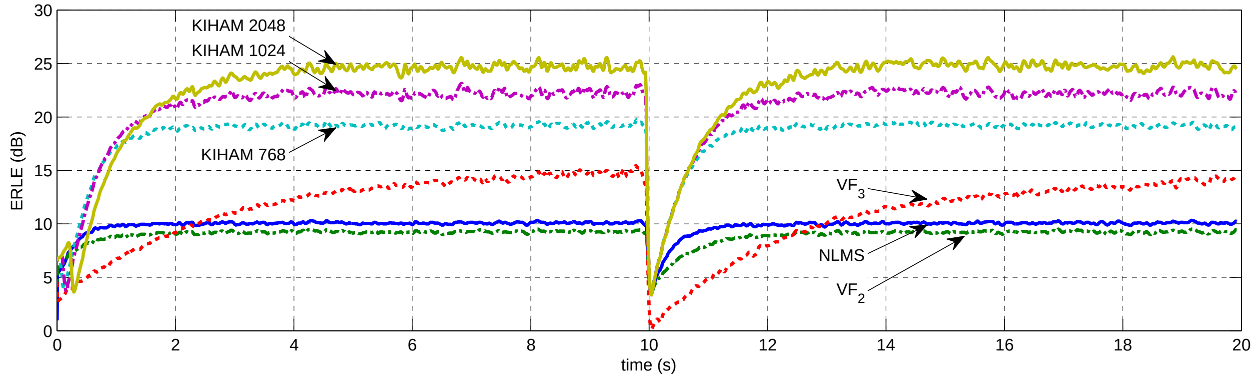

Fig. 5. Online ERLE performance of the proposed method ("KIHAM", considering buffer lengths of 756, 1024 and 2048), compared to a linear NLMS filter, a quadratic VF ("VF_2") and an ad-hoc designed cubic VF ("VF_3").

in a separation of 0.14. The kernel width is chosen as $\sigma = 0.25$, in line with rules of thumb from kernel density estimation [25]. The regularization constants are set to $c_\alpha = 10^{-2}$ and $c_h = 1$.

The offline algorithm converges after 6 iterations, in this scenario. Figs. 3 and 4 display the estimated nonlinearity and linear channel, which match the true values very well.

3.2. Online cancellation

We now apply the proposed algorithm to an online setup during 20 seconds, at a sampling frequency of 8 kHz. An abrupt change in the RIR is triggered after 10 seconds, in order to assess the algorithm's ability to reconverge. The parameters of the algorithm are kept the same as in the previous experiment, except for the buffer length, which is set to 768, 1024 and 2048 in three separate tests. For the adaptive part of the algorithm, we use a normalized least-mean squares (NLMS) filter with step size $\mu = 1$.

In addition, we include the results obtained by a purely linear NLMS filter with $\mu = 1$, and by a VF with linear and quadratic kernels that adapt with $\mu_L = 1$ and $\mu_{NL} = 0.3$, respectively, following the Separate NLMS (SNLMS) algorithm [26]. For a fair comparison, only coefficients on the main diagonal of the quadratic kernel are updated, since we assume it is known that the nonlinearity produced by the clipping has no memory. We have also applied the scheme of [21] to this scenario, though our preliminary results are not significantly better than the linear filter, and therefore we have not included them in the comparison. This may be due to the nonlinearity of the distortion in our scenario, which seems particularly hard to identify.

Fig. 5 compares the results obtained by all algorithms, averaged out over 50 independent realizations of the experiment. The online configuration of the proposed KIHAM algorithm obtains a proper performance, reaching higher ERLE levels when the buffer length increases. The linear NLMS filter produces fairly bad results for obvious reasons. The result obtained by the VF filter ("VF2") could seem surprising (it behaves slightly worse than the linear filter), although it coincides with that shown in [27]. The reason of this behavior is because the clipping operation in audio amplifiers only produces odd harmonics [28] and the VF is limited to identify nonlinearities of second order. This can be demonstrated by evaluating the performance of a VF that has been designed *ad-hoc* considering only linear and cubic kernels (again with only coefficients in the main diagonal for the case of the third-order kernel). As shown in Fig. 5 this ad-hoc VF ("VF3") reaches a higher ERLE than that of both the quadratic VF and the NLMS filter, but its convergence is slower than the other algorithms.

We have also carried out additional experiments considering harder clippings that give rise to larger levels of nonlinear distortion. First results indicate that the proposed method maintains its performance reasonably well, while the behavior of the other evaluated schemes degrades quickly.

4. CONCLUSIONS AND FUTURE WORK

We have presented a novel kernel-based nonlinear acoustic echo canceler that models the echo path as a Hammerstein system. This model assumes that the nonlinear distortion is caused mainly by the clipping operation of the audio power amplifier and behavior of the loudspeaker. The proposed algorithm jointly learns a kernel-based representation to identify the nonlinearity of the Hammerstein system and a linear filter to model the linear channel.

We described two different implementations of the algorithm: a batch scheme and an online configuration. The batch algorithm iterates between updating the nonlinearity and the linear filter until convergence is reached. The online strategy extends this scheme with an adaptive linear filter to update the time-varying linear channel. Experiments show that the batch algorithm is capable of satisfactorily identifying both the nonlinearity and the linear channel of the Hammerstein echo path, and the online version allows to adapt the model correctly in time-varying scenarios. The quality of the solution improves as more data is used for the estimation of the nonlinearity.

Future research will focus on the implementation of a fully online version that updates the nonlinearity in an online manner, including while it is filling its buffer. In addition, we will evaluate the performance of our proposal considering other scenarios, such as nonlinear distortions with memory.

5. REFERENCES

[1] C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic echo control. An application of very-high-order adaptive filters," *IEEE Signal Processing Magazine*, vol. 16, no. 4, pp. 42–69, 1999.

[2] M. Mossi, N. Evans, and C. Beaugeant, "An assessment of linear adaptive filter performance with nonlinear distortions,"

in *Proc. of Acoustics, Speech, and Signal Processing*, 2010, pp. 313–316.

[3] L. Azpicueta-Ruiz, M. Zeller, A. Figueiras-Vidal, J. Arenas-García, and W. Kellermann, "Adaptive combination of Volterra kernels and its application to nonlinear acoustic echo cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 97–110, 2011.

[4] F. Küch, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," in *Proc. of Acoustics, Speech, and Signal Processing*, 2005, vol. 3, pp. 105–108.

[5] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, Wiley, New York, NY, USA, 2000.

[6] F. Küch and W. Kellermann, *Nonlinear Acoustic Echo Cancellation*, Springer, Heidelberg, Germany, 2006.

[7] A. Guerin, G. Faucon, and R. Le Bouquin-Jeannès, "Nonlinear acoustic echo cancellation based on Volterra filters," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 672–683, 2003.

[8] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in *Proc. of Acoustics, Speech, and Signal Processing*, 1999, pp. 877–880 vol.2.

[9] F. Küch and W. Kellermann, "Partitioned block frequency-domain adaptive second-order Volterra filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 564–575, 2005.

[10] T. Burton, R. Goubran, and F. Beaucoup, "Nonlinear system identification using a subband adaptive Volterra filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 5, pp. 1389–1397, 2009.

[11] A. Carini and G. Sicuranza, "Fourier nonlinear filters," *Signal Processing*, vol. 94, pp. 183–194, 2014.

[12] J.-P. Costa, A. Lagrange, and A. Arliaud, "Acoustic echo cancellation using nonlinear cascade filters," in *Proc. of Acoustics, Speech, and Signal Processing*, 2003, vol. 5, pp. 389–392.

[13] J. Vörös, "An iterative method for Hammerstein-Wiener systems parameter identification," *Journal of Electrical Engineering*, vol. 55, pp. 328–331, 2004.

[14] D. Comminiello, M. Scarpiniti, L. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, "Functional link adaptive filters for nonlinear acoustic echo cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 7, pp. 1502–1512, 2013.

[15] T. Mäkelä and R. Niemistö, "Effects of harmonic components generated by polynomial preprocessor in acoustic echo control," in *Proc. of Intl. Workshop Acoustic Echo Noise Control (IWAENC)*, 2003, pp. 139–142 vol.2.

[16] J. Fu and W.-P. Zhu, "A nonlinear acoustic echo canceller using sigmoid transform in conjunction with RLS algorithm," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 10, pp. 1056–1060, 2008.

[17] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Nonlinear spline adaptive filtering," *Signal Processing*, vol. 93, no. 4, pp. 772 – 783, 2013.

[18] B. Schölkopf and A. J. Smola, *Learning with Kernels*, The MIT Press, Cambridge, MA, USA, 2002.

[19] J. Ohmer, F. Maire, and R. Brown, "Implementation of kernel methods on the GPU," in *Proc. of the Digital Image Computing: Techniques and Applications (DICTA 2005)*, 2005.

[20] S. Van Vaerenbergh, *Kernel methods for nonlinear identification, equalization and separation of signals*, Ph.D. thesis, University of Cantabria, 2010.

[21] J. Gil-Cacho, M. Signoretto, T. van Waterschoot, M. Moonen, and S. Jensen, "Nonlinear acoustic echo cancellation based on a sliding-window leaky kernel affine projection algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 9, pp. 1867–1878, 2013.

[22] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag New York, Inc., 1995.

[23] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis*, Cambridge University Press, 2004.

[24] "NRSC AM preemphasis/deemphasis and broadcast audio transmission bandwidth specifications," *Standard ANSI/EIA-54988*, 1998.

[25] B. W. Silverman, *Density estimation for Statistics and Data Analysis*, Chapman & Hall/CRC, London, UK, April 1986.

[26] M. Zeller and W. Kellermann, "Coefficient pruning for higher-order diagonals of Volterra filters representing Wiener-Hammerstein models," in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2008.

[27] D. Comminiello, L. Azpicueta-Ruiz, M. Scarpiniti, A. Uncini, and J. Arenas-García, "Functional link based architectures for nonlinear acoustic echo cancellation," in *Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, 2011, pp. 180–184.

[28] M. Hoyerby and M. A. E. Andersen, "Carrier distortion in hysteretic self-oscillating class-D audio power amplifiers: Analysis and optimization," *IEEE Transactions on Power Electronics*, vol. 24, pp. 714–729, 2009.

Corrigendum

In the published version of this paper, a factor $\frac{1}{2}$ was present in the cost functions (2), (5) ,(8), (9), and (11), though it was not taken into account for further calculations. We have left this factor out.