NONLINEAR ACOUSTIC ECHO CANCELLATION USING ELITIST RESAMPLING PARTICLE FILTER

Mhd Modar Halimeh, Christian Huemmer and Walter Kellermann

University of Erlangen-Nuremberg, Multimedia Communications and Signal Processing
Cauerstr. 7, D-91058 Erlangen, Germany
{mhd.m.halimeh, christian.huechr.huemmer, walter.kellermann}@fau.de

ABSTRACT

This paper considers an effective method for nonlinear acoustic echo cancellation (NL-AEC). More specifically, we model the nonlinear echo path by a latent state vector capturing the coefficients of a memoryless processor and a linear finite impulse response filter. To estimate the posterior probability distribution of the state vector, an elitist particle filter based on evolutionary strategies (EPFES) has been proposed, which evaluates realizations of the latent state vector based on long-term fitness measures. This method includes a manually-tuned recursive calculation of the probabilities that the observation has been produced by the state-vector realizations. For avoiding this manual tuning, we introduce a new approach denoted as Elitist Resampling Particle Filtering (ERPF) which can also be shown to combine the advantages of the Sequential Importance Sampling Particle Filter (SIS-PF) and the Sequential Importance Sampling/Resampling Particle Filter (SIR-PF). This new approach allows universal use and leads to superior system identification performance compared to both the original EPFES as well as the SIR-PF, as verified for a simulated scenario and a real smartphone recording.

Index Terms— Nonlinear System Identification, Particle Filters, Bayesian Estimation, Nonlinear Acoustic Echo Cancellation.

1. INTRODUCTION

Since decades, acoustic echo cancellation (AEC) has been an active research area. Starting from the early linear acoustic echo cancelers [1], the complexity of algorithms and models have been growing. This translated into a rich literature of algorithms which are able to estimate complex linear acoustic echo paths both efficiently and accurately [2–5]. Driven by the distortions produced by miniaturized loudspeakers in portable devices, sophisticated and more complex nonlinear AEC (NL-AEC) algorithms and models have emerged based on Volterra filters [6, 7], artificial neural networks [8], power filters [9], Legendre polynomials [10] and functional link adaptive filters [11]. While the models vary widely, their parameters are almost always determined by solving linear problems, using e.g., the Least Mean Square (LMS) algorithm or its extensions.

On the other hand, particle filters have received increasing attention in various fields to identify nonlinear systems, e.g., in finance [12], source tracking [13] and navigation [14]. Particle filters approximate the posterior density of a latent state vector by a set of weighted particles, where each particle corresponds to a possible realization of the latent state vector [15]. Unlike the extended Kalman filter, which can be used for the task of nonlinear system identification, a particle filter imposes no strong constraints on the system model, rendering it a suitable approach for a wide range of applications. This generality motivated researchers to develop many different particle filtering schemes [16–20] which vary in complexity and performance. In this paper, we propose the elitist resampling particle filter (ERPF) to fuse the advantages of the Sequential Importance Sampling Particle Filter (SIS-PF) and the Sequential Importance Sampling/Resampling Particle Filter (SIR-PF) [21] by exploiting an evolutionary selection of elitist particles introduced in [22]. The comparison between the proposed ERPF and the EPFES provides additional theoretical insights and highlights that despite requiring no tuning step, the proposed approach outperforms the original one.

This paper is organized as follow: In Section 2, the NL-AEC scenario is introduced and the concept of Significance Aware (SA) adaptive filtering [23] is reviewed. In Section 3, a brief overview of two generic particle filtering algorithms is given while in Section 4, the proposed ERPF is introduced. The relationship to previous work is discussed in Section 5 and is followed by the experimental evaluation for a simulated AEC scenario and a real smartphone recording in Section 6. Finally, conclusions are drawn in Section 7.

2. THE HAMMERSTEIN MODEL

In this section, we introduce the NL-AEC scheme characterized by a memoryless preprocessor followed by a linear finite impulse response (FIR) filter. As shown in Figure 1, the memoryless preprocessor $f(\boldsymbol{s}_n, \hat{\boldsymbol{a}}_n)$ models the nonlinearity introduced by the loudspeaker while the FIR filters $\hat{\mathbf{h}}_{\text{direct},n}$ and $\hat{\mathbf{h}}_{\text{comp},n}$ jointly model the linear acoustic echo path from the loudspeaker to the microphone [24] as discussed below.

2.1. The memoryless preprocessor

As illustrated in Figure 1, at time instant n, the speech signal segment $\boldsymbol{s}_n = [s_n, s_{n-1}, ..., s_{n-M+1}]^{\text{T}}$ is first distorted by the loudspeaker nonlinearities. This distortion is modeled by

$$\hat{\boldsymbol{d}}_n = [\hat{d}_n, \hat{d}_{n-1}, ..., \hat{d}_{n-M+1}]^{\text{T}} = \boldsymbol{f}(\boldsymbol{s}_n, \hat{\boldsymbol{a}}_n), \quad (1)$$

where $\hat{\boldsymbol{d}}_n$ is the distorted speech signal vector. The nonlinear function $\boldsymbol{f}$ is parameterized by the estimated vector $\hat{\boldsymbol{a}}_n = [\hat{a}_{0,n}, \hat{a}_{1,n}, ..., \hat{a}_{P-1,n}]^{\text{T}}$. In the following, we model the nonlinearity as a P-th order polynomial

$$\boldsymbol{f}(\boldsymbol{s}_n, \hat{\boldsymbol{a}}_n) = \sum_{\upsilon=0}^{P-1} \hat{a}_{\upsilon,n} \mathcal{L}_{2\upsilon+1}(\boldsymbol{s}_n) \quad (2)$$

The authors would like to thank the Deutsche Forschungsgemeinschaft (DFG) for supporting this work (contract number KE 890/4-2).

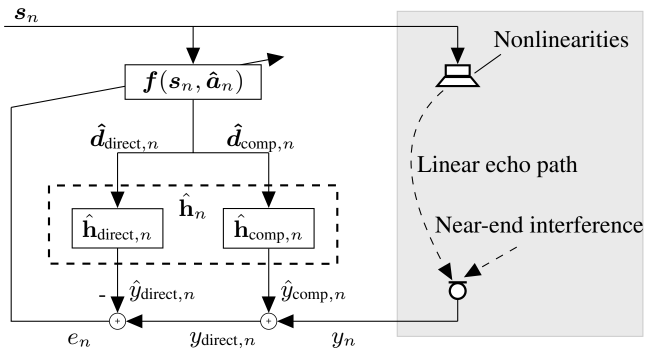

Fig. 1. NL-AEC scenario with SA-based system identification.

which is based on the odd-order Legendre polynomials of the first kind $\mathcal{L}_{2v+1}$ [10,23,24]. It is worth mentioning that the use of particle filters in the NL-AEC context is not limited to this model, so that other nonlinear functions can be used just as well.

2.2. The linear acoustic echo path

The linear acoustic path between the loudspeaker and the microphone at time n is estimated by means of a linear FIR filter of length M $\hat{\mathbf{h}}_n = [\hat{h}_{0,n}, \hat{h}_{1,n}, ..., \hat{h}_{M-1,n}]^{\text{T}}$. Thus, the microphone signal estimate can be written as

$$\hat{y}_n = \hat{\mathbf{h}}_n^{\text{T}} \boldsymbol{f}(\boldsymbol{s}_n, \hat{\boldsymbol{a}}_n). \tag{3}$$

2.3. The Significance-Aware Approach

Two sets of parameters are needed to describe the entire acoustic echo path: the memoryless preprocessor vector $\hat{\boldsymbol{a}}_n$ and the linear FIR filter $\hat{\mathbf{h}}_n$. However, since $\hat{\mathbf{h}}_n$ is typically a long filter, using a computationally expensive approach, e.g., particle filters, to estimate all coefficients is of limited practical interest. In order to overcome this problem, the SA-principle is used [23, 24], which intends to reduce the computational load by focusing on the direct-path of the estimated linear acoustic echo path $\hat{\mathbf{h}}_n$ to estimate the parameters of the nonlinear preprocessor. Thereby, we start by estimating the entire linear FIR filter $\hat{\mathbf{h}}_n$ using the Normalized Least Mean Square (NLMS) algorithm

$$\hat{\mathbf{h}}_{n+1} = \hat{\mathbf{h}}_n + \frac{\mu}{\hat{\boldsymbol{d}}_n^{\text{T}} \hat{\boldsymbol{d}}_n + \epsilon} \hat{\boldsymbol{d}}_n e_n, \tag{4}$$

where μ is a scalar step size, ϵ is a positive constant which prevents division by zero and the error signal $e_n = y_n - \hat{y}_n$ relates the current observation to its estimate. The FIR filter $\hat{\mathbf{h}}_n$ is then split into two parts: a direct acoustic path from the loudspeaker to the microphone centered around the peak at I $(0 < I < M)$, with a length $2R + 1$

$$\hat{\mathbf{h}}_{\text{direct},n} = [\hat{h}_{I-R,n}, \hat{h}_{I-R+1,n}, ..., \hat{h}_{I+R,n}]^{\text{T}} \tag{5}$$

which is expected to contain the most significant part of $\hat{\mathbf{h}}_n$, and a complementary part

$$\hat{\mathbf{h}}_{\text{comp},n} = [\hat{h}_{0,n}, ..., \hat{h}_{I-R-1,n}, \hat{h}_{I+R+1,n}, ..., \hat{h}_{M-1,n}]^{\text{T}} \tag{6}$$

which contains the rest of $\hat{\mathbf{h}}_n$ coefficients. To compensate the estimation errors in $\hat{\mathbf{h}}_n$ during the estimation of the preprocessor parameters $\hat{\boldsymbol{a}}_n$, we define the state vector according to [24]

$$\hat{\boldsymbol{x}}_n = [\hat{\boldsymbol{a}}_n, \hat{\mathbf{h}}_{\text{direct},n}]^{\text{T}} \tag{7}$$

which is estimated using particle filters. Since the system now describes the nonlinearity and the direct acoustic path only, we further decompose the nonlinearly distorted signal $\hat{\boldsymbol{d}}_n$ in a similar manner

$$\begin{aligned} \hat{\boldsymbol{d}}_{\text{direct},n} &= [d_{I-R,n}, d_{I-R+1,n}, ..., d_{I+R,n}]^{\text{T}} \\ \hat{\boldsymbol{d}}_{\text{comp},n} &= [d_{0,n}, ..., d_{I-R-1,n}, d_{I+R+1,n}, ..., d_{M-1,n}]^{\text{T}}. \end{aligned} \tag{8}$$

As illustrated in Figure 1, the direct-path and complementary components for the received microphone signal y_n are also decomposed

$$y_{\text{direct},n} = y_n - \hat{y}_{\text{comp},n} = y_n - \hat{\mathbf{h}}_{\text{comp},n}^{\text{T}} \hat{\boldsymbol{d}}_{\text{comp},n}. \tag{9}$$

The direct path components $\hat{\boldsymbol{d}}_{\text{direct},n}$ and $y_{\text{direct},n}$ are then used to estimate the latent state vector $\hat{\boldsymbol{x}}_n$ defined in (7).

3. PARAMETERS ESTIMATION USING PARTICLE FILTERS

3.1. Sequential Importance Sampling Particle Filter

The Bayesian Minimum Mean Square Error (MMSE) estimate of the state vector $\hat{\boldsymbol{x}}_n$ is given by

$$\hat{\boldsymbol{x}}_n = E\{\boldsymbol{x}_n | \boldsymbol{y}_{1:n}\}, \tag{10}$$

where $\boldsymbol{y}_{1:n} = [y_1, y_2, ..., y_n]^{\text{T}}$. However, due to the nonlinear mathematical relationship between the state vector $\boldsymbol{x}_n$ and the observation vector y_n, a closed-form solution of the Bayesian estimator is not available. A particle filter evaluates the expectation, or equivalently the integral, by approximating the posterior distribution by

$$p(\boldsymbol{x}_{0:n} | \boldsymbol{y}_{1:n}) \approx \sum_{i=1}^{N_p} w_n^{(i)} \delta(\boldsymbol{x}_{0:n} - \boldsymbol{x}_{0:n}^{(i)}), \tag{11}$$

where $\{\boldsymbol{x}_{0:n}^{(i)}, w_n^{(i)}\}_{i=1}^{N_p}$ are the N_p particles $\boldsymbol{x}_{0:n}^{(i)}$ and their associated weights $w_n^{(i)}$, which are normalized to fulfill $\sum_{i=1}^{N_p} w_n^{(i)} = 1$. The weights are calculated according to the importance sampling principle [16]: assume that the only access available to $p(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})$ is by evaluating a proportional quantity $\pi(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})$. Then, in order to approximate $p(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})$, one can draw samples from another easily accessible density denoted $q(\boldsymbol{x}_n|\boldsymbol{y}_{1:n}, \boldsymbol{x}_{0:n})$ which has a strict support over $p(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})$ and compensates for the difference through weighting the samples by

$$w_n^{(i)} = \frac{p(\boldsymbol{x}_{0:n}^{(i)} | \boldsymbol{y}_{1:n})}{q(\boldsymbol{x}_n^{(i)} | \boldsymbol{y}_{1:n}, \boldsymbol{x}_{0:n})} \propto \frac{\pi(\boldsymbol{x}_{0:n}^{(i)} | \boldsymbol{y}_{1:n})}{q(\boldsymbol{x}_n^{(i)} | \boldsymbol{y}_{1:n}, \boldsymbol{x}_{0:n})}. \tag{12}$$

By exploiting the recursive nature of the Bayesian filter [18]

$$p(\boldsymbol{x}_{0:n} | \boldsymbol{y}_{1:n}) = \frac{p(\boldsymbol{x}_n | \boldsymbol{x}_{n-1}) p(y_n | \boldsymbol{x}_n)}{p(y_n | \boldsymbol{y}_{1:n-1})} p(\boldsymbol{x}_{0:n-1} | \boldsymbol{y}_{1:n-1}), \tag{13}$$

and through an appropriate choice of the importance density $q(\boldsymbol{x}_n^{(i)}|\boldsymbol{y}_{1:n}, \boldsymbol{x}_{0:n})$, one arrives at the weight calculation [16–18]

$$w_n^{(i)} \propto w_{n-1}^{(i)} \frac{p(\boldsymbol{x}_n^{(i)} | \boldsymbol{x}_{n-1}^{(i)}) p(y_n | \boldsymbol{x}_n^{(i)})}{q(\boldsymbol{x}_n^{(i)} | \boldsymbol{x}_{n-1}^{(i)}, y_n)}. \tag{14}$$

Furthermore, if one samples from the state transition density $p(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)})$, the weight calculation in (14) reduces to [18]

$$w_n^{(i)} \propto w_{n-1}^{(i)} p(y_n | \boldsymbol{x}_n^{(i)}). \tag{15}$$

While being attractive for its simplicity, the SIS-PF exhibits an unavoidable drawback: After few iterations, all weights but one will be equal to zero, which is known as the degeneracy problem [25, 26].

On the other hand, by examining (15), the SIS-PF maintains a memory for each particle through the weight propagation. This will mitigate the impact of an observed outlier by considering the previous observations as well in addition to the current one.

3.2. Sequential Importance Sampling/Resampling Particle Filter

SIR-PFs, also known as Bootstrap filters [21], address the degeneracy problem by resampling after each iteration. In [27], resampling was motivated by the fact that a sample with a low weight will have a low impact on the estimation. Thereby, samples with small weight are discarded while samples with sufficiently large weight are replicated such that all particles will have the same weight $1/N_p$. As a result, the new weights after propagation are obtained directly via

$$w_n^{(i)} \propto p(y_n|\boldsymbol{x}_n^{(i)}). \tag{16}$$

While solving the degeneracy problem, the SIR-PF introduces another, namely sample impoverishment, i.e, after few iterations, few particles with significant weights will dominate the estimation [16].

Many adaptive resampling schemes have been developed [16–18, 21, 28]. In this paper, adaptive resampling is not discussed and, for clarity, only a basic SIR-PF which resamples after each iteration is considered. As an efficient alternative to adaptive resampling, the posterior Probability Density Function (PDF) $p(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})$ can by approximated by means of a continuous PDF. For instance, this has been realized for the GPF in [29]. However, by resampling the entire set of particles, the SIR-PF, and as a consequence the GPF, re-initiate all particles' memory, resulting in an increased focus on instantaneous likelihoods.

4. THE ELITIST RESAMPLING PARTICLE FILTER

The SIS-PF and SIR-PF can be seen as two extremes. While the SIS-PF is the optimum approach in the sense of not discarding any path, a realistic limitation of the number of particles will result in weight degradation. On the other hand, the SIR-PF tackles this issue by favoring a local solution at each iteration while eliminating many possible paths due to their low local probabilities or likelihoods. This motivates the introduction of a hybrid resampling scheme, offering a compromise between the two extremes. To this end, we divide the particle population into two sets, an elitist set $\boldsymbol{\Theta}_{\mathbf{E},n}$ and a non-elitist set $\boldsymbol{\Theta}_{\mathbf{NE},n}$. Non-elitist particles are discarded and replaced by new particles while elitist particles are not discarded by resampling at the end of each iteration, preserving the elitist particles' memory [22].

It is intuitive to use a weight threshold w_{th} in deciding which particles are elitist and which are discarded. Moreover, we state that the goal is to discard a maximum number of low-weighted particles without discarding '*good*' particles. This is accomplished by specifying *good* particles as the ones being directly drawn from the posterior distribution. From the weights equation (12), such particles will have normalized weights equal to $\frac{1}{N_{\text{p}}}$. Thus, we set the threshold as $w_{\text{th}} = \frac{1}{N_{\text{p}}}$, and assign the particles as follows

$$\boldsymbol{x}_n^{(i)} \in \begin{cases} \boldsymbol{\Theta}_{\mathbf{E},n} & \text{if } w_n^{(i)} \geq w_{\text{th}} \\ \boldsymbol{\Theta}_{\mathbf{NE},n} & \text{if } w_n^{(i)} < w_{\text{th}}. \end{cases} \tag{17}$$

After discarding the non-elitist particles, substitution particles need to be drawn. To this end, elitist particles are used to construct a continuous density $\hat{p}(\boldsymbol{x}_{0:n}^{(i)}|\boldsymbol{y}_{1:n})$, which new particles are drawn from. Afterwards, all particles are advanced through $p(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n)$, resulting in the newly spawned particles being distributed according to

$$\boldsymbol{x}_{n+1,\text{new}}^{(i)} \sim \hat{p}(\boldsymbol{x}_{0:n}|\boldsymbol{y}_{1:n})p(\boldsymbol{x}_{n+1}|\boldsymbol{x}_n). \tag{18}$$

Introducing the above density into the weight equation of (12) gives

$$w_{n+1,\text{new}}^{(i)} \propto p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)}). \tag{19}$$

One should notice that this update rule is identical to (16).
On the other hand, the elitist particles are not resampled or changed and by following the same derivation as for the SIS-PF, we arrive at

$$w_{n+1,\text{elitist}}^{(i)} \propto w_n^{(i)} p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)}). \tag{20}$$

Thus, the overall weight update rule is

$$w_{n+1}^{(i)} \propto \begin{cases} w_n^{(i)} p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)}) & \text{if } \boldsymbol{x}_n^{(i)} \in \boldsymbol{\Theta}_{\mathbf{E},n} \\ p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)}) & \text{if } \boldsymbol{x}_n^{(i)} \in \boldsymbol{\Theta}_{\mathbf{NE},n}. \end{cases} \tag{21}$$

The advantage of such an approach is that introducing the extra memory for the elitist particles increases the robustness against outliers compared to the SIR-PF. This is reflected in the weights propagation in (21), where the weights of the elitist particles depend on previous observations as well as current ones. Similar to the GPF, drawing the substitution particles from a continuous density, the particle population will have more diversity compared to the SIR-PF, thus mitigating the diversity decay effect.

5. RELATION TO PRIOR WORK

As mentioned earlier, this work is based on the EPFES introduced in [22] aiming at avoiding any heuristic parameter choices to guarantee robustness and thus allow universal use. In contrast, the approach in [22] was heuristically motivated and incorporated a manually-tuned recursive calculation of the particle weights:

$$w_{n+1}^{(i)} \propto \begin{cases} {w_n^{(i)}}^{\lambda} p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)})^{1-\lambda} & \text{if } \boldsymbol{x}_n^{(i)} \in \boldsymbol{\Theta}_{\mathbf{E},n} \\ p(y_{n+1}|\boldsymbol{x}_{n+1}^{(i)}) & \text{if } \boldsymbol{x}_n^{(i)} \in \boldsymbol{\Theta}_{\mathbf{NE},n} \end{cases} \tag{22}$$

where λ has to be optimized for the application at hand.
In this paper, we formulated the new ERPF algorithm as a hybrid combination of SIS-PF and SIR-PF in Section 4. This provides new theoretical insights and leads to a version of [22], where the weight calculation in (22) is replaced by the update of (21) thereby avoiding any tuning parameter and allowing universal use.

6. EXPERIMENTS RESULTS

The ERPF is compared to an SIR-PF [21] which uses a systematic resampling scheme, the GPF [29] and the EPFES [24] (with an optimized λ in (22)). In [24] a comparison of the EPFES to a well-known nonlinear acoustic echo canceller based on a Hammerstein Group Model (HGM) shows that the EPFES outperforms the HGM-based approach.
In both experiments the NLMS was first initialized for 0.1s to obtain an initial Room Impulse Response (RIR) estimate $\hat{\mathbf{h}}_n$ of length $L_h = 256$ to allow for an initial split of $\hat{\mathbf{h}}_n$ according to the SA-principle into two parts with length $L_{h_d} = 11$ for the direct path. The nonlinear function $\boldsymbol{f}(\boldsymbol{s}_n, \hat{\boldsymbol{a}}_n)$ in Figure 1 is approximated using the first three odd-order Legendre polynomials of the first kind, $\upsilon = \{1, 2, 3\}$ in (2). Afterwards, the different filters (ERPF, EPFES, GPF, SIR-PF) are realized with $N_p = 100$ particles to estimate the

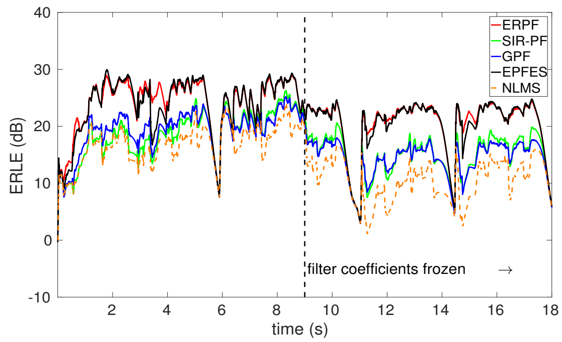

Fig. 2. ERLE of NLMS, SIR-PF, GPF, EPFES and ERPF for a synthesized NL-AEC scenario.

	Synthesized		Recorded	
Filter	$\overline{\text{ERLE}}_{\text{on}}$	$\overline{\text{ERLE}}_{\text{off}}$	$\overline{\text{ERLE}}_{\text{on}}$	$\overline{\text{ERLE}}_{\text{off}}$
ERPF	21.4	21.5	13.8	15.7
EPFES	19.7	21.2	12.9	14.6
SIR-PF	15.4	13.6	11.1	12.9
GPF	16.1	13.3	11.5	13.1
NLMS	14.0	7.4	9.8	9.3

Table 1. Temporal average ERLE for NL-AEC in dB.

state vector $\hat{\boldsymbol{x}}_n = [\hat{\mathbf{h}}_{\text{direct},n}, \hat{\boldsymbol{a}}_n]$ of length 14. Each filter produces online estimates during an adaptation period of 9s, which is half of the speech signal length. In the next 9s, the estimated state vector $\hat{\boldsymbol{x}}_n$ is kept fixed and used in evaluating how well each filter's estimate generalizes. Note that this is an important indicator for the performance during double-talk situations. As an evaluation measure we exploit the Echo Return Loss Enhancement (ERLE)

$$\text{ERLE}_n = 10\log_{10}\frac{E\{y_n^2\}}{E\{e_n^2\}}, \quad \text{where} \quad e_n = y_n - \hat{y}_n \qquad (23)$$

where $\hat{y}_n$ is the estimated microphone signal obtained using the estimated coefficients $[\hat{\boldsymbol{a}}_n, \hat{\mathbf{h}}_{\text{direct},n}, \hat{\mathbf{h}}_{\text{comp},n}]$.
Each experiment is repeated 10 times, and the average results are used and presented. In order to provide a reference for comparison, the ERLE of a linear AEC approach, i.e., the NLMS algorithm with an FIR filter of length 256 is also evaluated for each scenario.

6.1. Simulated nonlinearity

In this experiment, a speech signal $\boldsymbol{s}$ of a female speaker with a total length of 18 seconds was used. The nonlinearity was generated using

$$d_n = \frac{1}{4}\tanh(4s_n), \qquad (24)$$

as it has already been employed in [24]. Afterwards, the distorted loudspeaker signal d_n was convolved with a recorded RIR and superimposed by additive white noise at an SNR level of 30dB.

The ERLE values of each filter are shown in Figure 2. To provide a quantitative comparison, the temporal averages of the ERLE during adaptation and after freezing the parameters are provided in Table 1.

As the results show, during adaptation (up to t=9s), both the ERPF and the EPFES outperform SIR-PF and GPF by a very significant margin and despite the absence of the tuning parameters, the ERPF outperforms the EPFES. More importantly, by looking at the ERLE for frozen coefficients (t=9s to t=18s), the ERPF and the EPFES keep outperforming the GPF and SIR-PF. This shows that the ERPF and EPFES do not outperform the other filters due to an increased adaptation speed or over-adaptation to the actual signal, but rather because of a more accurate description of the unknown system. Moreover, the SIR-PF and GPF performed similarly, which is consistent with the findings in [29]. Finally, the initial convergence speed of the ERPF and the EPFES reflects how fast the filters adapt to a new system and as a consequence, it seems reasonable to expect an according tracking behavior in case of a time-varying echo path.

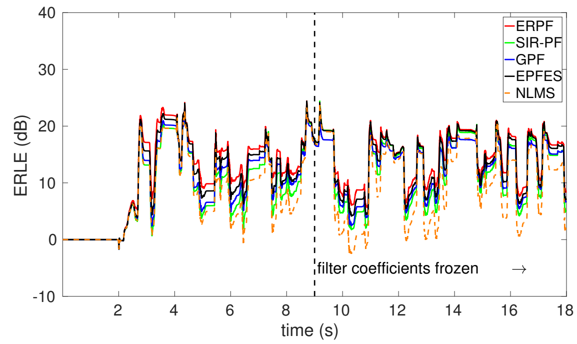

Fig. 3. Performance comparison of NLMS, SIR-PF, GPF, EPFES and ERPF for NL-AEC with real smartphone recordings.

6.2. Real smartphone recording

In this experiment, a male speech signal was emitted by a smartphone loudspeaker, which introduces an unknown nonlinearity, and was recorded by a microphone placed in a close proximity with an SNR level of 20dB. Again, the ERLE and the temporal average ERLE values are shown in Figure 3 and Table 1, respectively.

It is clear from the results, that ERPF is an effective algorithm when used with real smartphone recordings. The performance difference between the ERPF/EPFES and the other filters (GPF, SIR-PF) is relatively high, both for online adaptation (up to t=9s) and frozen filter coefficients (t=9 to t=18s). Furthermore, it is visible from Table 1 that the ERPF generalizes better compared to the (manually tuned) EPFES for the real smartphone recording.

Finally, it should be mentioned that the computational cost of the EPFES is discussed in [24] and is estimated to be not more than 1.5 times the multiplications per sample (MPS) of a stand-alone NLMS algorithm, and 0.37 times the MPS of a 4-branch HGM. Due to the similarities between the EPFES and the ERPF, these estimates also hold for the ERPF.

7. CONCLUSION

In this paper, we have introduced the ERPF as a new algorithm for NL-AEC based on the EPFES approach introduced in [22], but avoiding all tuning parameters. The ERPF is formulated as a hybrid of two generic and widely-known particle filters, namely, the SIS-PF and the SIR-PF, and there the algorithm's performance was evaluated for two speech signals in two different scenarios, namely, for a synthesized nonlinearity and for a commercial smartphone. The results have shown that despite having no tuning parameters, the ERPF achieves a better system identification performance compared to the (manually optimized) EPFES, the SIR-PF and the GPF for both scenarios so that ERPF can be viewed as a universally applicable approach to NL-AEC with superior peformance.

8. REFERENCES

[1] M. Sondhi, "An adaptive echo canceller," *Bell Syst. Tech. J.*, vol. 46, no. 3, pp. 497–511, 1967.

[2] C. Breining *et al.*, "Acoustic echo control. an application of very-high-order adaptive filters," *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 42–69, July 1999.

[3] E. Ferrara, "Fast implementations of LMS adaptive filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 4, pp. 474–475, Aug. 1980.

[4] J. Soo and K. Pang, "Multidelay block frequency domain adaptive filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 2, pp. 373–376, Feb. 1990.

[5] J. Benesty and S. Gay, "An improved PNLMS algorithm," in *IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP)*, May 2002, pp. 1881–1884.

[6] A. Stenger *et al.*, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in *IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP)*, 1999, pp. 15–19.

[7] L. Gudupudi *et al.*, "A comparison of different loudspeaker models to empirically estimated non-linearities," in *4th Joint Workshop on Hands-free Speech Commun. and Microphone Arrays (HSCMA)*, May 2014, pp. 112–116.

[8] A. Birkett and R. Goubran, "Acoustic echo cancellation using NLMS-neural network structures," in *IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP)*, May 1995, vol. 5, pp. 3035–3038.

[9] F. Kuech *et al.*, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," in *IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP)*, Mar. 2005, vol. 3, pp. 105–108.

[10] A. Carini *et al.*, "Introducing Legendre nonlinear filters," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, May 2014, pp. 7939–7943.

[11] D. Comminiello *et al.*, "Functional link adaptive filters for nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 7, pp. 1502–1512, July 2013.

[12] D. Egloff and M. Min-Oo, "Convergence of Monte Carlo algorithms for pricing american options," Dept. of Math. and Stat., McMaster University, Hamilton, Ontario, Canada, 2002.

[13] M. Arulampalam *et al.*, "Bearings-only tracking of maneuvering targets using particle filters," *EURASIP J. on Applied Signal Process.*, pp. 2351–2365, 2004.

[14] F. Gustafsson *et al.*, "Particle filters for positioning and navigation and tracking," *IEEE Trans. Signal Process.*, pp. 425–437, Feb. 2002.

[15] J. Carpenter *et al.*, "Building robust simulation-based filters for evolving data sets," Tech. Rep., Dept. of Stat. Univ. Oxford, London, U.K., 1999.

[16] M. Arulampalam *et al.*, "A Tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 5, pp. 174–188, 2002.

[17] A. Doucet and A. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," in *The Oxford Handbook of Nonlinear Filtering*, London, U.K.: Oxford Univ. Press., 2008.

[18] Z. Chen, "Bayesian filtering from Kalman filters to particle filters and beyond," Adaptive Systems Lab., McMaster Univ., Hamilton, ON, Canada, Mar. 2002.

[19] R. Van Der Merwe *et al.*, "The unscented particle filter," in *Conf. on Neural Inform. Process. Syst. (NIPS)*, Feb. 2001, vol. 13.

[20] M. Klaas *et al.*, "Toward practical N2 Monte Carlo: the marginal particle filter," in *Proc. of the Twenty-First Confer. on Uncertainty in Artificial Intell.*, Arlington, Virginia, USA., 2005, pp. 308–315.

[21] N. Gordon *et al.*, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEEE Proc. F- Radar and Signal Process.*, Apr. 1993, vol. 140, pp. 107–113.

[22] C. Huemmer *et al.*, "The elitist particle filter based on evolutionary strategies as novel approach for nonlinear acoustic echo cancellation," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, May 2014, pp. 1315–1319.

[23] C. Hofmann *et al.*, "Significance-aware Hammerstein group models for nonlinear acoustic echo cancellation," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, May 2014, pp. 5934–5938.

[24] C. Huemmer *et al.*, "The significance-aware EPFES to estimate a memoryless preprocessor for nonlinear acoustic echo cancellation," in *IEEE Global Conf. on Signal and Inform. Process. (GlobalSIP)*, Dec. 2014, pp. 557–561.

[25] A. Doucet *et al.*, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. and Computing*, vol. 10, pp. 197–208, July 2000.

[26] A. Doucet *et al.*, "Particle filters for state estimation of jump Markov and linear systems," *IEEE Trans. Signal Process.*, pp. 613–624, Mar. 2001.

[27] J. Liu and R. Chen, "Blind deconvolution via sequential imputations," *J. of the Amer. Statistical Assoc.*, vol. 90, no. 430, pp. 567–576, 1995.

[28] J. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Publishing Company, Incorporated, 2008.

[29] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2592–2601, Oct. 2003.