Audio Engineering Society

Convention Paper 10694

Presented at the AES 156th Convention
2024 June 15–17, Madrid, Spain

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Advances in Group Delay Modeling and Optimization

Sunil Bharitkar[1]

[1] *DMS Audio Lab, Samsung Research America*

Correspondence should be addressed to Sunil Bharitkar (s.bharitkar@samsung.com)

ABSTRACT

This paper introduces a stochastic optimization algorithm to model and equalize group delay functions with all-pass filters. The method employs two techniques that are used to optimize the hyperparameters of deep neural networks and machine learning models. Various examples are used to demonstrate that the stochastic search algorithm (employing either of the two techniques) presented consistently outperforms baseline or state-of-the-art (SOTA) techniques. Moreover, the resulting all-pass filters are stable because the pole amplitudes are constrained within the unit circle during optimization. Initialization with poles obtained from closed-form expressions of second-order all-pass group delay functions further leads to an excellent approximation of the desired group delay function.

1 Introduction

Modeling and optimizing group delay functions with all-pass filters is essential in many audio applications, including (i) phase compensation to linearize phase response in the pass band of a loudspeaker [1],[2],[11] (ii) representation of musical instruments for sound synthesis as well as creating audio effects [5], (iii) decorrelation of audio signals between channels for improved spatiality perception [6], [7], (iv) feedback delay networks (FDNs) to synthesize reverberation [3], (v) reducing the peak amplitude of signals [4], and (vi) coherent summation of loudspeaker and subwoofer responses measured in the room to compensate for spectral notches in the total response [8],[9], [10], [12], [13].

All-pass filters are a computationally efficient signal processing building block with a unique property: the numerator polynomial can be uniquely derived purely from the denominator polynomial by a simple reflection of the coefficients of the denominator polynomial. The Hermitian (mirror-image) symmetry advantageously yields computationally efficient filter realizations with a minimum number of multipliers [14]. Additionally, the system transfer function is lossless (viz. $|H_{ap}(z)| = |H_{ap}(e^{j\omega})| = |H_{ap}(\omega)| = 1$); and, in the case of stable filters, the poles corresponding to the roots of the denominator polynomial are inside the unit circle. All-pass filters are cascaded to obtain a selective group delay during modeling and equalization. In group delay research, the two main topics are (i) identifying group delay perception thresholds as a function of frequency for headphones and loudspeakers, and (ii) computational algorithms to model and equalize group delay functions.

Hoshino *et al.* [18] present data on the acceptable group delay distortion values due to the low-pass filters' cutoff frequency on tone quality to identify perception thresholds. Ashley [19] presents group and phase delay requirements for loudspeakers, confirming a 2 ms requirement for group delay tolerance between the woofer and tweeter. The thresholds for group de-

lay distortion with monoaural/binaural presentation (on headphones) and on loudspeakers are given by Flanagan *et al.* [20]. Møller *et al.* studied the perceptual effects of group delay distortion with second-order all-pass filters in the range of 1-12 kHz in the context of binaural playback on headphones (where the all-pass was in either both channels or in one channel). Liski *et al.* [22] establish audibility perception limits for positive and negative group delays between 500 Hz and 4 kHz for four test signals. The results lead to a deeper understanding of the equalization of audio systems using perceptual thresholds for group delay artifacts. There is a great deal of research in the area of perception of monaural phase distortion that includes [15],[16], [17]. There are also extensive studies on the influence of interaural phase delay and group delay in binaural hearing (e.g., [23]).

In the context of computational algorithms to model and equalize group delay functions, Rämö *et al.* [24] present a graphical design with user-adjustable all-pass parameters based on [25] to model the group delay of stiff-string sounds. A parametric model of strings using cascade and nested all-pass filters is shown in [26]. A heuristic search technique is given in [27] for equalization of the group delay. Techniques for group delay modeling using all-pass filters employing minimax/gradients are shown in ([9],[28],[29], [32],[39], [33]), whereas second-order Hessian is given in [35]. Additionally, all-pass filters derived by leveraging the minimum phase nature of the stable all-pass function are given in ([40],[43],[34]), and with an eigenfilter approach is provided in ([30], [31],[41]). Appiah *et al.* [44] present an approach to precisely match the desired group delay in a few discrete frequency bins while interpolating between the exact match using Blaschke interpolation. Kanaya *et al.* [42] provide constraints on the second-order all-pass filter coefficients to ensure stability for extremal values of the corresponding group delay function. A vector space projection method [45] is presented to solve a set of non-convex constraints in an all-pass filter design, where the technique is similar in its characterization to the eigenfilter approach.

In this paper, we advance the research for modeling group delay functions using two stochastic search techniques that are currently used to tune the hyperparameters of deep neural networks and other machine learning models. We compare and contrast the two presented stochastic search methods (Bayesian optimization and Simulated Annealing) by developing an algorithm around these techniques for modeling group delays. Initializing search algorithms with locally optimal solutions (e.g., derived from baseline or SOTA techniques [34], [35]) is shown to yield significantly improved performance over SOTA. The techniques presented loosely incorporate perception thresholds ([22]) to help model group delays of a perceptual bass enhancement algorithm [46].

Section 2 summarizes the two baseline SOTA techniques ([34], [35]) that are used to compare with the presented method. Section 3 presents Bayesian optimization and Simulated annealing techniques and the algorithm that includes these techniques for modeling a desired group delay function. Closed-form expressions for computing pole amplitude from all-pass group delays are also presented in this section. These pole amplitudes also serve as initialization values for the presented stochastic search algorithm. Section 4 includes results by comparing the algorithm presented with the baseline techniques for various examples. Section 5 concludes the paper.

2 Baseline or State-of-the Art (SOTA) Group-delay Modeling Techniques

Traditional approaches for group delay analysis fall under two categories: (a) forward-modeling (Fig. 1a) to introduce a delay $\tau_{model}(\omega) \approx \tau_d(\omega) + R_1$ (where R_1 is a fixed delay) in the side chain (lower path) that models the group-delay $\tau_d(\omega)$ in the main chain (top path) to time-synchronize both chains; (b) equalizing the main chain (Fig. 1b) with a group-delay equalizer $\tau_{eq}(\omega) \approx -\tau_d(\omega) + R_2$ to obtain a constant or approximately flat group delay in the passband. In this paper, we restrict the investigation to forward modeling, since group delay equalization essentially involves modeling the negative of the group delay of the system $\tau_d(\omega)$. The two SOTA models used in this paper are [34] and [35], since the former encompasses the work done using the Hilbert transform property. The latter incorporates minimizing a least-squares criterion (in the minimax sense), which is a standardized technique in Matlab [36] as the *iirgrpdelay* function in the DSP toolbox. We do not consider the eigenfilter approach, as the results ([44], Figs. 2-4) show inadequate modeling performance. Nevertheless, the presented algorithm employing stochastic search techniques (Bayesian optimization and Simulated Annealing) can be readily adapted to work with the eigenfilter approaches.

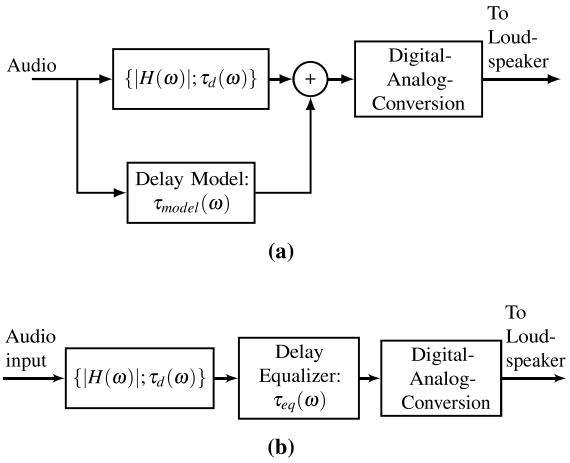


Fig. 1: Group-delay (a) forward-modeling, (b) equalization.

2.1 Autoregressive (AR) Model [34]

The process for computing the denominator all-pass polynomial coefficients $\{a_k; k = 1,2,...,P\}$ is shown in Fig. 2, where the desired group-delay to be modeled is $\tau_D(m)$ (with m being the DFT bin index),

$$\tau_D(m) = \begin{cases} \tau_D(m) & ;m = \{1,2,\ldots,N+1\} \\ \tau_D(2N+2-m) & ;m = \{N+2,\ldots,2N\} \end{cases} \quad (1)$$

The group-delay values $\tau_D(m)$ are inverse Fourier transformed to give corresponding time-domain values $x(n)$, which in turn are used to compute the cepstral coefficients $\{c(1) = 0, c(2), \ldots, c(n)\}, (n = 2,\ldots,2N)$ as,

$$2c(n) = \begin{cases} x(n)/(n-1) & ;n = \{2,\ldots,N+1\} \\ 2c(2N+2-n) & ;n = \{N+2,\ldots,2N\} \end{cases} \quad (2)$$

The cepstral coefficients are Fourier transformed to give the magnitude spectrum (i.e., pole spectrum). The resulting pole spectrum is sent through an exponential function and then inverse-transformed to obtain the autocorrelation coefficients $R(n)$. The autocorrelation coefficients $R(n)$ are normalized to give the normalized autocorrelation coefficients $r(n) = R(n)/R(0)$. Subsequently, the all-pass denominator coefficients are obtained by solving the normal equations from the normalized autocorrelation coefficients using the Levinson algorithm. The all-pass coefficients of the numerator b_k are obtained by using mirror image symmetry (viz., $b_k = a_{P-k}$), where P is the order of the all-pass filter.

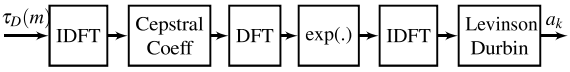


Fig. 2: Autoregressive (AR) group-delay model

2.2 Hessian-based Model [35], [36]

The p-th norm distance function with all-pass filter $A(e^{j\omega})$ and a weighting function $w(\omega)$ is expressed as,

$$\|L(A)\|_p = \{\int_0^\pi w(\omega)|\tau_d(\omega) - \tau_A(\omega)|^p d\omega\}^{1/p} \quad (3)$$

In the work by [35] (and in this paper), a cascade of second-order all-pass filters is used to minimize the influence on the error caused by any quantization effects and finite coefficient size. Additionally, the stability of a second-order cascade of all-pass filters can be readily tested, since computing the roots of a quadratic denominator polynomial (i.e., poles) requires low complexity. The solution to (3) reduces to the minimization of a non-linear function for each of the poles that contribute to the cascaded filter[1]. Accordingly [35] uses the algorithm of Fletcher and Powell [38] that incorporates a second-order gradient (viz., a positive definite Hessian matrix) to solve this nonlinear minimization problem to obtain the all-pass filter coefficients.

3 Stochastic Search Optimization

3.1 Second-order All-pass Group-delay

The transfer function of a second-order all-pass filter with complex conjugate poles $\{z_k, z_k^*\}$ can be expressed as

$$\begin{aligned} H_{ap}(z,z_k) &= \frac{z^{-1}-z_k^*}{1-z_k z^{-1}} \frac{z^{-1}-z_k}{1-z_k^* z^{-1}} \\ &= \frac{|z_k|^2 - 2z^{-1}\Re(z_k) + z^{-2}}{1-2z^{-1}\Re(z_k) + z^{-1}|z_k|^2} \quad (4) \\ H_{ap}(z) &= \prod_{k=1}^{P} H_{ap}(z,z_k) \\ H_{model}(z) &= z^{-R} H_{ap}(z) \end{aligned}$$

[1] For a single second-order all-pass filter, this would correspond to two variables associated with a single pole (viz., a pole amplitude and the corresponding pole angle), and with the complex conjugate of this pole forming the second pole with the result being a filter of order 2.

where $\Re(.)$ denotes the real part of the complex number $z_i = r_i e^{j\omega_i}$, and R is a fixed delay. The phase response of a second-order all-pass filter $\phi_{ap}(\omega)$ [37], is

$$\begin{aligned}\phi_{ap}(\omega) &= -2\omega - 2\arctan\left(\frac{r_i \sin(\theta_i^-)}{1 - r_i\cos(\theta_i^-)}\right) \\ &- 2\arctan\left(\frac{r_i \sin(\theta_i^+)}{1 - r_i\cos(\theta_i^+)}\right) \end{aligned} \quad (5)$$

with $\theta_i^{\pm} = \omega \pm \omega_i$. From (5), the group delay $\tau_{ap}(\omega)$ is

$$\begin{aligned}\tau_{ap}(\omega) &= \frac{2(1 - r_i\cos(\theta_i^-))}{1 - 2r_i\cos(\theta_i^-) + r_i^2} \\ &+ \frac{2r_i(\cos(\theta_i^+) - r_i)}{1 - 2r_i\cos(\theta_i^+) + r_i^2} \end{aligned} \quad (6)$$

At pole-frequencies $\omega = \omega_i$ and with a choice of desired group-delay values $\beta_i = \tau_d(\omega_i)$; and, by algebraically simplifying (6) the all-pass pole amplitude r_i can be obtained as the roots of the following third-order polynomial,

$$\begin{aligned}&(\beta_i + 2)r_i^3 - (2\beta_i\cos(2\omega_i) + 2\cos(2\omega_i) + \beta_i)r_i^2 + \\ &(\beta_i - 2\cos(2\omega_i) + 2\beta_i\cos(2\omega_i))r_i + (2 - \beta_i) \end{aligned} \quad (7)$$

The valid solution is the one that yields $|r_i| < 1$.

3.2 Bayesian Optimization

Bayesian optimization ([47],[48]) is a global hyperparameter optimization technique constrained by the limits of hyperparameters that is best suited for optimization with around 20 or fewer hyperparameters. The method builds a surrogate function for the objective and quantifies the uncertainty in that surrogate using Gaussian process regression. Furthermore, several parameters are used for initialization, including the type of acquisition function that guides the sampling of optimal hyperparameters [48]. Recent advances can be found in [49]. The search parameter bounds for the pole amplitudes are set on the basis of the maximum and minimum values of the group delay to be modeled. Similarly, the limits for the pole frequencies are set to be the edge frequencies over which the pole optimization is to be achieved while minimizing an error function.

3.3 Simulated Annealing (SA)

Simulated Annealing (SA) models the physical process of heating a material and slowly lowering the temperature to decrease defects, thus minimizing the energy of the system towards the optimal solution. At each iteration of the SA algorithm, a new set of hyperparameters is randomly generated within the limits specified as box constraints, from which the distance of the trial point from the current point is calculated with a probability distribution $Pr(\nabla\bar{\psi}_{SD}) = 1/(1 + \exp(\nabla\bar{\psi}_{SD}/T_{anneal,k}))$ with current temperature $T_{anneal,k}$. The algorithm accepts all new points that lower the objective function (9) but also, with a certain probability, points that increase the energy. The annealing schedule used is typically exponential in nature $T_{anneal,k} = T^0/\log(k)$ with T^0 being the initial temperature at $k = 1$.

3.4 Group Delay Model Algorithm

The technique incorporating either search method is shown in Algorithm 1. The error function for minimization with the stochastic search techniques used in this paper is

$$\begin{aligned}\sigma &= 10\log_{10}\left(\frac{1}{N}\sum_{m=1}^{N} W(m)\|\tau_d(\omega_m) - \tau_{model}(\omega_m)\|_2^2\right) \\ \tau_{model}(\omega_m) &= \sum_{k=1}^{P} \tau_{ap}^{(k)}(\omega_m) \end{aligned} \quad (8)$$

where $\tau_{ap}^{(k)}(\omega_m)$ corresponds to the k-th second order all-pass filter (6), and $W(m)$ is a weighting function (unless otherwise specified in the Results section, the weighting function is set to unity). The constraint $|r_k| < 1, \forall k$ is ensured during optimization using search techniques. For examples 1-3 (Results section), Algorithm 1 was initialized by the locally optimal poles generated by [35]. For Example 4 in the Results section, the initial pole frequencies and amplitudes were assigned using a random distribution (uniform distribution). In Example 5, the pole frequencies were not optimized but kept the same as the initialization values, whereas the pole amplitudes were initialized using (7). The successful results show that the initialization of Algorithm 1 can be done using one of the SOTA techniques, using (7), or using a random assignment of the poles. Both *bayesopt* and *simulannealbnd*

can be set to run for T seconds or for Q iterations.

Algorithm 1: All-pass pole-optimization using stochastic search (Bayesian optimization *or* Simulated Annealing)

Result: Result: (r_k^*, ω_k^*, R^*), $k = 1, \ldots, P$
minimum : σ

1 Initialize: Set the initial values of the poles $(r_k^{(q)}, \omega_k^{(q)})$ and fixed delay $R^{(q)}$ (where $q = 0$ is the iteration) randomly or in a prescribed fashion (e.g., (7) or from the optimized poles of [35], [34]) depending on the problem. Desired group delay to model $\tau_d(\omega)$
2 Set pole amplitude constraints $\{r_{k,min}, r_{k,max}\}$ (where $r_{k,min} > 0$, $r_{k,max} < 1$), pole angles $\{\omega_{k,min}, \omega_{k,max}\}$, and fixed delay limits $\{R_{min}, R_{max}\}$
3 Set number of iterations (or maximum amount of time T) for termination ;
4 **while** $t \leq T$ *seconds* $\vee$ $q \leq Q$ *iterations* **do**
5 Compute $\tau_{ap}(\omega_m)$ with $(r_k^{(t)}, \omega_k^{(t)})$ using (8);
6 Compute σ using (8);
7 Update $(r_k^{(t)}, \omega_k^{(t)})$ using *bayesopt* or *simulannealbnd* to minimize (8) ;
8 $q = q + 1$;
9 **end**

4 Results

We shall demonstrate the improvement in the group delay modeling performance over the SoTA techniques using Algorithm 1 through several synthetic examples and measurements. Examples 1-3 are recreated from [34], whereas Example 4 is from [25], and Example 5 is the side-chain group delay from a perceptual bass extension technique [46]. In this paper, we have initialized Algorithm 1 with pole amplitudes $r_k, (k = 1, 2, \ldots, K)$ derived from the locally optimal solution generated by [35] for Examples 1-3. For Example 4, the initialization of the pole amplitude and frequencies was performed with a uniform distribution, whereas for Example 5 the initialization of r_k was done using (7).

Example 1: The first example has group-delay characteristics (or a quadratic phase response) prescribed as

$$\tau_d(\omega) = 2.5\omega; \quad 0 \leq \omega \leq \pi \tag{9}$$

The bin values of the group delay are expressed as

$$\tau_D(m) = \begin{cases} \frac{5\pi(m-1)}{512} & ;m = \{1, 2, \ldots, 257\} \\ \tau_D(514 - m) & ;m = \{258, \ldots, 512\} \end{cases} \tag{10}$$

Figure 3 (blue curve) shows the true group delay, the red curve is the result obtained using the AR model [34] and the magenta shows the result obtained with [35] (viz., Matlab *iirgrpdelay(.))* using four poles[2]. As mentioned above, we initialized Algorithm 1 with the four poles obtained from [35] (magenta), with $(r_{k,min}, r_{k,max}) = (0.1, 0.9)$ and $(\omega_{k,min}, \omega_{k,max}) = (0.01, \pi - 0.01)$. The results of Algorithm 1 with Bayesian (cyan) and Simulated Annealing (black curve) show a clear visual improvement over both techniques for the same number of poles. The error function (8) is $\sigma_{AR} = 1.75$ dB, $\sigma_{minimax} = -5.26$ dB, $\sigma_{Bayes} = -7.2$ dB, and $\sigma_{SA} = -7.7$ dB.

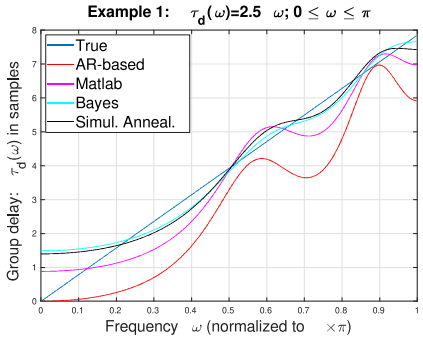


Fig. 3: Example 1 modeling performance (results using (8) shown in Table 1).

Example 2: The second example has group-delay characteristics prescribed as

$$\tau_d(\omega) = (10\omega - 3\omega^2); \quad 0 \leq \omega \leq \pi \tag{11}$$

The DFT bin-values of the group delay are expressed as

$$\tau_D(m) = \begin{cases} \frac{\pi(m-1)}{512} \times (10 - \frac{6\pi(m-1)}{512}) & ;m = \{1, \ldots, 257\} \\ \tau_D(514 - m) & ;m = \{258, .., 512\} \end{cases} \tag{12}$$

Figure 4 (blue curve) shows the desired group delay, the red curve is the result obtained using the AR model

[2] For low computational complexity applications it may be desirable to have a cascade of only a few all-pass filters

[34], and the magenta shows the result obtained with [35]. Both results were generated with four poles. For algorithm 1, we again initialized both Bayesian and Simulated Annealing algorithms with the four poles from [35] (as the approximation with this technique was better than the AR model). The results from Algorithm 1 with Bayesian (green) and Simulated Annealing (black curve) show a clear visual improvement over both techniques for the same number of poles over a wider frequency domain. The error function (8) is $\sigma_{AR} = 15.93$ dB, $\sigma_{minimax} = 3.42$ dB, $\sigma_{Bayes} = -0.4$ dB and $\sigma_{SA} = -0.45$ dB.

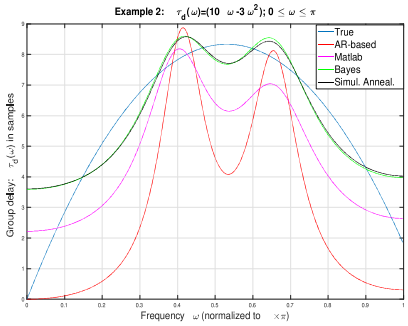


Fig. 4: Example 2 modeling performance (results using (8) shown in Table 1).

Example 3: This example has staircase like group-delay characteristics

$$\tau_D(m) = \begin{cases} 8 & ;m = \{1,2,\ldots,65\} \\ 6 & ;m = \{66,67,\ldots,129\} \\ 4 & ;m = \{130,131,\ldots,193\} \\ 2 & ;m = \{194,195,\ldots,257\} \\ \tau_D(514-m) & ;m = \{258,\ldots,512\} \end{cases} \quad (13)$$

Figure 5 (blue curve) shows the desired group delay, the black curve is the result obtained using the AR model [34], and the magenta shows the result obtained with [35]. Both results were generated with 20 poles. For algorithm 1, we initialized the Bayesian and Simulated Annealing algorithms again with the twenty poles of [35] (the approximation with this technique was similar to the approximation with the AR model with (8)). The results of Algorithm 1 with Simulated Annealing (green curve) show a clear visual improvement over both techniques for the same number of poles with a less propensity for creating high-density ripples over each frequency domain of the multistep function. The error function (8) is $\sigma_{AR} = -4.83$ dB, $\sigma_{minimax} = -4.9$ dB, $\sigma_{Bayes} = -4.1$ dB, and $\sigma_{SA} = -7$ dB.

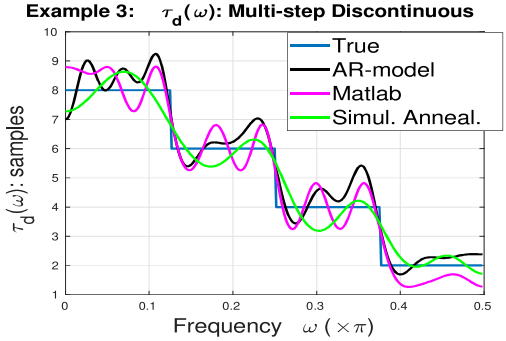


Fig. 5: Example 3 modeling performance (results using (8) shown in Table 1).

Example 4: This example is a dispersion filter (cf. [25], Fig. 6 spring reverberator element) where the group delay in milliseconds was converted to samples using a sampling frequency of 48 kHz. The frequency region of interest (DC-2.5 kHz) was kept the same for modeling. Figure 6(c) (blue curve) shows the true group delay to be modeled. The optimal number of poles for both the AR model and [35] was determined by first computing the error σ in (8) and then finding the number of poles that produced the lowest error. Figure 6(a) shows the graph that relates the number of poles with σ, where the AR model (red circles) failed to give stable poles beyond six poles because the Toeplitz matrix of normalized autocorrelation coefficients derived from normal equations is poorly conditioned. For example, the large condition number (associated with the right y-axis) is shown in Fig. 6(a) for the Toeplitz matrix based on the use of 8 poles. Similar results are obtained with [35] where Matlab (shown as a blue "+" symbol) gave an error associated with *poorly conditioned Hessian* when the number of poles was 26. The best result for [35] was obtained with 36 poles, which gave the lowest result σ. Consequently, both search techniques in Algorithm 1 were initialized with a uniform distribution of pole frequencies and amplitudes for 36 poles to compare with the baseline techniques. The resulting initial group delay curve, from random initialization, is shown in Fig. 6(b) in magenta. The comparisons of the

resulting performance plots (after using Algorithm 1) are shown in Fig. 6(c), where both stochastic search techniques from Algorithm 1 outperformed the baseline techniques. The error function (8) is $\sigma_{AR} = 36$ dB, $\sigma_{minimax} = 41$ dB, $\sigma_{Bayes} = 21.95$ dB, and $\sigma_{SA} = 22.25$ dB.

Example 5: This example is a group delay from the side chain of a perceptual bass extension technique by Bharitkar *et al.* [46]. Figure 7(a) (blue curve) shows the true group delay. The frequency domain for the approximation is DC-5 kHz, and this limit was set loosely based on the experimental detection threshold results reported by Liski *et al.* ([22], cf. Fig. 11, minimum detection threshold ≈ 1.5 ms corresponding to 72 samples at 48 kHz). The minimax technique [35] gave an ill-conditioned Hessian matrix for all orders when approximating the group delay function from DC to 5 kHz. Therefore, a band-edge technique was incorporated into the Matlab implementation *iirgrpdelay(.)*. In this technique, the center frequency of each pole ω_i was selected at the peak locations of the true group delay, and the edge of each frequency band was empirically selected to be approximately midway between two adjacent peak frequency locations. Then, each second-order all-pass filter was computed. The sum of group delay functions of each second-order section results in the magenta curve in Fig. 7(a). The modeling is reasonable except for the region covering 1.5 - 4 kHz. The presented search optimization techniques were initialized at the same pole frequencies that correspond to the peaks in the group delay function. The group delay values for initialization at these frequencies were obtained from the desired group delay (blue curve). From there, the initial pole amplitudes $(r_k^{(0)}, \omega_k^{(0)})$ for Algorithm 1 were obtained using (7). Figure 7(b) shows the results of Bayesian optimization (red curve). Overlaid on the red curve is a black dotted curve, the result obtained using a weighting function of Fig. 7(c) for improved low-frequency modeling. Figure 7(d) shows the results with Simulated Annealing as an overlaid red curve. The black curve is Bayes for illustrative comparisons. The error function (8) is $\sigma_{minimax} = 40.1$ dB, $\sigma_{Bayes} = 36.4$ dB, and $\sigma_{SA} = 35.8$ dB. A summary of the modeling error results is given in Table 1, showing improvements over the two baseline or SoTA techniques.

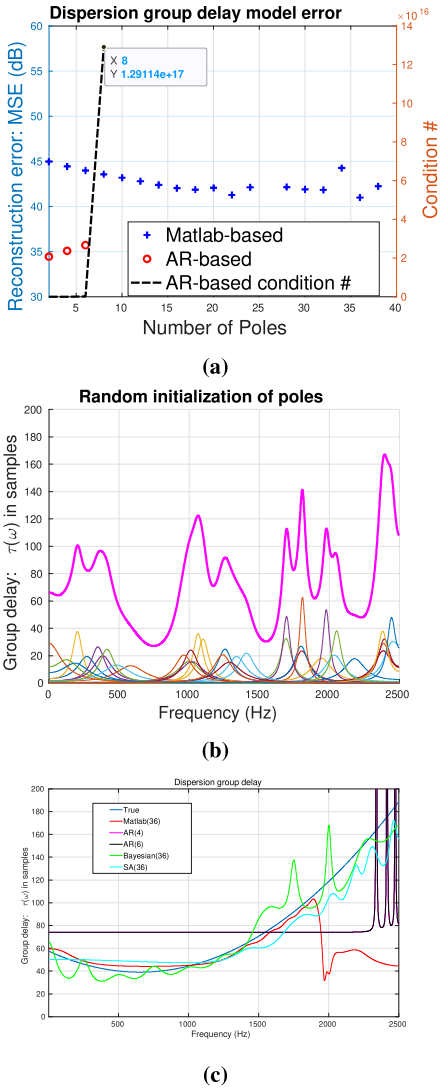


Fig. 6: Dispersion filter: (a) error, (b) random initialization of 36 poles leading to an initial group delay (solid magenta), (c) modeling perf. comparisons (results using (8) shown in Table 1)

5 Summary

This paper presents a technique for modeling group delay functions with a second-order cascade of all-pass

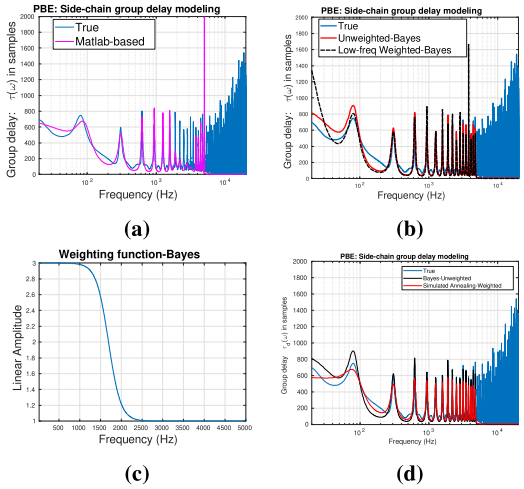


Fig. 7: Perceptual bass extension [46] side-chain modeling (results using (8) shown in Table 1): (a) Matlab [35], (b) Bayesian (weighted vs unweighted), (c) weighting function, (d) Simul. Anneal.

filters. The stochastic optimization methods presented in this paper (Bayesian optimization and Simulated Annealing) have been used in other areas of research, such as tuning hyperparameters of deep neural networks and machine learning models. The paper demonstrates through various examples that the presented Algorithm 1 (employing either of the two methods) consistently outperforms conventional techniques. The results show that the initialization of Algorithm 1 can be done by any of the baseline/SOTA techniques, or by (7), or randomly. In practice, Algorithm 1 can be run twice (once with Bayesian optimization and next with Simulated Annealing) to obtain the best modeling or equalization performance. An extension of this work would be to define the number of cascaded second-order filters as a hyper-parameter during optimization with Algorithm 1. Future research will be directed at adapting the algorithm presented to include perception criteria (detection thresholds of group delays) along with listening tests.

6 Aknowledgements

The authors express gratitude to Thaddeus Paez who helped create Figs. 1-2. Samsung Electronics and Samsung Research America supported this work. The authors would like to thank Samsung's US Audio Lab staff, who helped with all aspects of this research, offered insightful suggestions, and contributed to this work.

Table 1: Group-delay reconstruction error: σ_{dB} (8)

Problem	AR	Deczky	Bayes opt.	Simul. Anneal.
Ex. 1	1.75	-5.26	-7.2	**-7.7**
Ex. 2	15.93	3.42	-0.4	**-0.45**
Ex. 3	-4.83	-4.9	-4.1	**-7**
Disp.	36	41	**21.95**	22.25
PBE	NA	40.1	36.4	**35.8**

References

[1] S. Cecchi, L. Romoli, F. Piazza, and A. Carini, "Multipoint room response equalization with group delay compensation: Subjective listening tests," *7th Int. Symp. Image and Sig. Proc. and Analysis* (ISPA), 2011, Dubrovnik (Croatia).

[2] B. Radlovic and R. Kennedy, "Nonminimum-phase equalization and its subjective importance in room acoustics," *IEEE Trans. Speech & Audio Proc.*, 8(6), Nov. 2000, pp. 728–737.

[3] S. Schlecht, "Allpass feedback delay networks," *IEEE Trans. Sig. Proc.*, 69, Jan. 2021, pp. 1028–1038.

[4] S. Schlecht, L. Fierro, V. Välimäki, and J. Backman, "Audio peak reduction using a synced allpass filter," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), 2022, Singapore (Singapore).

[5] V. Välimäki, J. Abel, and J. Smith, "Spectral delay filters," *J. Audio Eng. Soc.*, 57 (7/8), July / August 2009, pp. 521–531.

[6] E. Canfield and J. Abel, "Signal decorrelation using perceptually informed allpass filters," *Proc. 19th Int. Conf. Digital Audio Effects* (DAFx), 2016, Brno (Czech Republic).

[7] S. Disch, "Decorrelation for immersive audio applications and sound effects," *Proc. 26th Int. Conf. Digital Audio Effects* (DAFx), 2023, Copenhagen (Denmark).

[8] S. Bharitkar, C. Kyriakakis, and T. Holman, "Time-alignment of multi-way speakers with group delay equalization-I," *Proc. 124th Audio Eng. Soc. Conv.* (AES), May 2008, Amsterdam (Netherlands).

[9] C. Lin and Y. Chen, "Phase compensation for multichannel low-frequency response using minimax approximation," *2012 Int. Conf. Audio, Lang. and Image Proc.* (ICALIP), 2012, Shanghai (China).

[10] V. Adam and S. Benz, "Correction of crossover phase distortion using reversed time all-pass IIR filter," *Proc. 122nd Audio Eng. Soc. Conv.* (AES), May 2007, Vienna (Austria).

[11] A. Mäkivirta, J. Liski, and V. Välimäki, "Modeling and delay-equalizing loudspeaker responses," *J. Audio Eng. Soc.*, 58 (7/8), July / August 2010, pp. 547–562.

[12] S. Bharitkar and C. Kyriakakis, "Comparison between time delay based and non-uniform phase based equalization for multichannel loudspeaker-room responses," *Proc. 119th Audio Eng. Soc. Conv.* (AES), Oct. 2010, New York (USA).

[13] S. Hosoi, H. Hamada, and N. Kameyama, "An improvement in sound quality of LFE by flattening group delay," *Proc. 116th Audio Eng. Soc. Conv.* (AES), May 2004, Berlin (Germany).

[14] P. Regalia, S. Mitra, and P. Vaidyanathan, "The digital all-pass filter: A versatile signal processing building block," *Proc. IEEE,*, 76(1), Jan. 1988.

[15] D. Preis, "Phase distortion and phase equalization in audio signal processing: A tutorial review," *J. Audio Eng. Soc.*, 30(11), Nov. 1982, pp. 774-–794.

[16] S. Choisel and G. Martin, "Audibility of phase response differences in a stereo playback system. Part 2: Narrow-band stimuli in headphones and loudspeakers," *Proc. 125th Audio Eng. Soc. Conv.*, Oct. 2008, San Francisco (USA).

[17] N. Thiele, "Phase considerations in loudspeaker systems," *Proc. 110th Audio Eng. Soc. 110th Conv.*, May 2001, Amsterdam (Netherlands).

[18] Y. Hoshino and T. Takegahara, "Permissible value of group delay distortion on tone quality due to low-pass filters," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), 1986, Tokyo (Japan).

[19] J. Ashley, "Group and phase delay requirements for loudspeaker systems," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), 1980, Denver (Japan).

[20] S. Flanagan, B. Moore, and M. Stone, "Discrimination of group delay in clicklike signals presented via headphone and loudspeakers," *J. Audio Eng. Soc.*, 53(7/8), July/August 2005, pp. 593–611.

[21] H. Møller, P. Minnaar, S. Olesen, F. Christensen, and J. Plogsties, "On the audibility of all-pass phase in electroacoustical transfer functions," *J. Audio Eng. Soc.*, 55(3), March 2007, pp. 115–134.

[22] J. Liski, A. Mäkivirta, V. Välimäki, "Audibility of group-Delay equalization," *IEEE/ACM Trans. Audio, Speech, Lang. Proc.*, 29, June 2021, pp. 2189–2201.

[23] J. Blauert, *Spatial Hearing: The Psychophysics of human Sound Localization*, The MIT Press, Cambridge (MA), 1996.

[24] J. Rämö and V. Välimäki, "Graphic delay equalizer," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), 2019, Brighton (UK).

[25] J. Abel and J. Smith, "Robust design of very high-order allpass dispersion filters," *Proc. 9th Int. Conf. Digital Audio Effects* (DAFx), 2006, Montreal (Canada).

[26] V. Välimäki, J. Parker, and J. Abel, "Parametric spring reverberation effect," *J. Audio Eng. Soc.*, 58 (7/8), July / August 2010, pp. 547–562.

[27] M. Quélhas, A. Petraglia, and M. Petraglia, "Efficient group delay equalization of discrete-time IIR filters," *Proc. 12th Eur. Signal Process. Conf.* (EUSIPCO), 2004, Vienna (Austria).

[28] P. Bernhardt, "Simplified design of high-order recursive group-delay filters," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 5, Oct. 1980, pp. 498–503.

[29] J. Vrbata, M. Laipert, and M. Vlček, "Synthesis of the group delay equalizers," *9th Int. Conf. Elec., Circ. and Syst.*, Sept. 2002, Dubrovnik (Croatia).

[30] J. Tapia, J. LoCicero, and G. Atkin, "The eigen-delay all-pass correction algorithm," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), April 2002, Orlando (Florida).

[31] T. Nguyen, "The design of arbitrary FIR digital filters using the eigenfilter method," *IEEE Trans. Sig. Proc.*, 41(3), March 1993, pp. 1128–1139.

[32] H-C. Chao, C-S. Lin, and B.-C Chieu, "Minimax design of FIR digital all-pass filters," *IEEE Trans. Circ. Syst.-II: Analog & Dig. Sig. Proc.*, 47(6), June 2000, pp. 576–580.

[33] X. Lai and Z. Lin, "Design and application of allpass filters with equiripple group delay errors," *IEEE Int. Symp. Circ. Syst.* (ISCAS), May 2013, Beijing (China).

[34] B. Yegnanarayana, "Design of recursive group-delay filters by autoregressive modeling," *IEEE Trans. Acoust., Speech, Sig. Proc.*, ASSP-30 (4), Aug. 1982, pp. 632–637.

[35] A. Deczky, "Synthesis of recursive digital filters using the minimum p-error criterion," *IEEE Trans. Audio & Electroacoust.*, AU-20(4), Oct. 1972, pp. 257–263.

[36] MATLAB,`https://www.mathworks.com/products/matlab.html`, R2023b, 2023.

[37] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Aug. 2009.

[38] R. Fletcher and M. Powell, "A rapidly convergent descent method for minimization," *Comput. J*, 6, 1963, pp. 163–168.

[39] C-K Chen and J-H. Lee, "Design of digital all-pass filters using a weighted least squares approach," *IEEE Trans. Circ. Syst.-II: Analog & Dig. Sig. Proc.*, 43(5), May 1994, pp. 346–350.

[40] G. Reddy and M. Swamy, "Digital all-pass filter design through discrete Hilbert transform," *Proc. IEEE Int. Symp. Circ. Syst.*, May 1990, New Orleans (Louisiana).

[41] S-C.Pei and J-J. Shyu, "Eigen-approach for designing FIR filters and all-pass phase equalizers with prescribed magnitude and phase response," *IEEE Trans. Circ. Syst.-II: Analog & Dig. Sig. Proc.*, 39(3), March 1992, pp. 137–146.

[42] N. Kanaya and S. Yuta, 'A digital all-pass transfer function with all possible coefficient combinations," *IEEE Trans. Circ. Syst.*, vol. CAS-28(12), Dec. 1981, pp. 1171–1174.

[43] H. Martinez, "A method for the design of phase equalizers," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), April 1979, Washington D.C (US).

[44] K. Appaiah and D. Pal, "All-pass filter design using Blaschke interpolation," *IEEE Sig. Proc. Lett.*, 27, Jan. 2020, pp. 226–230.

[45] K. Haddad, Y. Yang, N. Galatsanos, and H. Stark, "Allpass filter design using projection-based method under group-delay constraints," *Proc. IEEE Conf. Acoust. Speech & Sig. Proc.* (ICASSP), May 2001, Salt Lake City (Utah).

[46] S. Bharitkar, E. Blanes, G. Kubota, and A. Rawat, "Advances in perceptual bass extension for music and cinematic Content," *Proc. 154th Audio Eng. Soc. Conv.* (AES), May 2023, Helsinki (Finland).

[47] J. Snoek, H. Larochelle, and R. Adams, "Practical Bayesian optimization of machine learning algorithms," *Proc. Neural Inf. Proc. Syst. (NIPS)*, 2012.

[48] P. I. Frazier, "A tutorial on Bayesian optimization," *arXiv:1807.02811v1*[stat.ML], July, 2018.

[49] J. R. Doppa, V. Aglietti, and J. Gardner, "Advances in Bayesian Optimization," *Neural Inf. Proc. Syst. (NIPS) Wkshp.*, Dec. 2022.