PARAMETER ESTIMATION VIA DIFFERENTIABLE MODAL PLATE SYNTHESIS

Filippo Garofalo *, *Alessandro Antonio Lillo* *, *Alessandro Ilic Mezza* †, *Riccardo Giampiccolo, and Alberto Bernardini*

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
Milan, Italy
[filippo.garofalo, alessandroantonio.lillo]@mail.polimi.it
alessandroilic.mezza@polimi.it

ABSTRACT

We present our submission to Task A of the 1st DAFx Parameter Estimation Challenge, which concerns the estimation of the physical parameters of a vibrating plate from a synthetic impulse response. Our approach introduces a differentiable modal plate synthesizer and estimates the plate parameters through inference-time gradient-based optimization of the synthesizer parameters. The six target parameters are recovered by minimizing a multi-scale spectral loss via backpropagation through the differentiable plate model. To handle the non-convexity of the loss landscape, we adopt a two-phase training strategy consisting of multiple short-term probe optimizations, followed by full-scale refinement initialized from the best candidate. We evaluate the approach on eight impulse responses synthesized with the official challenge dataset generator. Compared with a constant-value predictor and the particle swarm optimization baseline provided by the challenge, the proposed method reduces the prediction error by approximately one order of magnitude.

1. INTRODUCTION

Plate reverberation is a canonical example of a weakly damped acoustic system whose digital emulation has a long history in audio effects processing [1–3]. While the forward problem, i.e., synthesizing a plate impulse response from physical parameters, is well understood [4, 5], the inverse problem of recovering those parameters from an impulse response remains challenging. This task can be formulated as a classic system identification problem in which, given a measured or target response, the goal is to estimate the underlying physical parameters that would reproduce it under the forward model. For plate reverberation, however, dense modal overlap, frequency-dependent damping, and strong parameter coupling make the inverse problem highly non-convex. In recent years, several approaches have thus turned to machine learning, e.g., using deep neural networks to estimate the parameters of physical models directly from data [6–8].

An alternative approach is to rely on search-based methods such as Particle Swarm Optimization (PSO) [9], which directly explores the physical parameter space by iteratively updating a population of randomly initialized candidate solutions based on each particle's own best-known position and the best solution found by the entire swarm. On the downside, however, gradient-free methods are known to struggle to scale effectively in high-dimensional, non-convex settings with many local minima.

Recent advances in differentiable digital signal processing [10] have shown that embedding physically motivated models directly into the forward pass of a learning algorithm enables effective parameter estimation via backpropagation. This paradigm was systematically explored by Mezza et al. [11], who demonstrated near-perfect parameter recovery for several lumped-element models by treating the physical simulator as part of a differentiable computational graph and optimizing it with standard gradient-based methods. Crucially, this white-box formulation preserves full physical interpretability: unlike neural network approaches, every trainable variable corresponds directly to a measurable physical quantity. This approach has recently gained traction as an effective strategy for optimizing the parameters of differentiable artificial reverberation models [12–17], as well as for white-box virtual analog models of audio circuits [18,19]. Notably, differentiable modal synthesis has also proven effective for the inverse modeling of vibrating strings, membranes, and plates [20].

In this work, we apply this framework to Task A of the 1st DAFx Parameter Estimation Challenge, which requires estimating a subset of plate parameters from a synthetic impulse response [21]. We implement a fully differentiable modal plate synthesizer in PyTorch [22] and optimize its parameters end-to-end using a Multi-Scale Spectral (MSS) loss. To mitigate the non-convexity of the problem, we adopt a two-phase strategy: in the first phase, multiple candidate initializations are drawn via Latin Hypercube Sampling (LHS) [23] and each is refined through a short gradient descent run; the best candidate is then used to initialize the full-scale optimization in the second phase. We evaluate the proposed method on eight impulse responses synthesized using the official dataset generator of the 1st DAFx Parameter Estimation Challenge, comparing it against a constant-value predictor and the PSO baseline provided by the challenge organizers [21]. The proposed method, which requires an average of 13.42 minutes per impulse response, reduces the prediction error by approximately one order of magnitude with respect to both baselines, matching the ground truth to at least three significant digits in five of the eight test cases.

The remainder of this paper is organized as follows. Section 2 reviews the physical model of the vibrating plate. Section 3 describes our differentiable synthesizer and the two-phase optimization strategy. Section 4 presents experimental results on eight synthetic impulse responses. Finally, Section 5 concludes this work.

2. PHYSICAL MODEL

We briefly review the plate model underlying the challenge [21]. The plate is described using modal synthesis, in which the vibrat-

* Equal contribution. † Corresponding author.

ing structure is represented as a superposition of independently excited resonant modes [24]. A thin rectangular plate of dimensions $L_x \times L_y$ and thickness h undergoes transverse displacement $u(x, y, t)$ governed by the damped Kirchhoff–Love equation [1]:

$$\rho h \frac{\partial^2 u}{\partial t^2} = -\rho h \mathcal{K} u - 2\rho h(\eta_0 + \eta_1 \mathcal{K}) \frac{\partial u}{\partial t} + \delta(\mathbf{x} - \mathbf{x}_i) f_i(t), \quad (1)$$

where $\delta(\mathbf{x} - \mathbf{x}_i) := \delta(x - x_i)\delta(y - y_i)$ is a two-dimensional Dirac delta distribution centered at the input position $\mathbf{x}_i := (x_i, y_i)$, and $f_i(t)$ is an external input, set to $f_i(t) = \delta(t)$ for impulse response measurements. Furthermore,

$$\mathcal{K} = -(T_0/\rho h)\Delta + (D/\rho h)\Delta^2 \quad (2)$$

is the Kirchhoff–Love operator, ρ is the volumetric density, T_0 is the in-plane tension per unit length, $D = Eh^3/12(1 - \nu^2)$ is the flexural rigidity, where E is Young's modulus, ν is Poisson's ratio, and η_0, η_1 are frequency-independent and frequency-dependent loss coefficients, respectively.

Assuming simply-supported boundary conditions, the eigenproblem admits the analytical solution

$$\Phi_m(x, y) = \frac{2}{\sqrt{L_x L_y}} \sin\left(\frac{m_x \pi x}{L_x}\right) \sin\left(\frac{m_y \pi y}{L_y}\right), \quad (3)$$

with modal frequencies

$$\Omega_m = \frac{1}{\sqrt{\rho h}} \gamma_m \left(T_0 + D\gamma_m^2\right)^{1/2}, \quad (4)$$

$$\gamma_m^2 = \frac{m_x^2 \pi^2}{L_x^2} + \frac{m_y^2 \pi^2}{L_y^2}, \quad (5)$$

where $m_x, m_y \in \mathbb{Z}^+$ are the modal indices. The general solution is a superposition of modal contributions

$$u = \sum_m \Phi_m(\mathbf{x}) q_m(t), \quad (6)$$

where each modal coordinate $q_m(t)$ satisfies a damped harmonic oscillator equation with loss coefficient $\sigma_m = \eta_0 + \eta_1 \Omega_m^2$.

The discrete-time update follows the exact all-pole biquad discretization

$$q_m[k+1] = 2r_m \cos(\Omega_m T_\mathrm{s}) q_m[k] - r_m^2 q_m[k-1] + b_m f_i[k], \quad (7)$$

with $r_m = e^{-\sigma_m T_\mathrm{s}}$, sampling interval T_s, and modal gain

$$b_m = 4T_\mathrm{s}^2 \Phi_m(\mathbf{x}_i)\Phi_m(\mathbf{x}_o) r_m / (\rho h L_x L_y). \quad (8)$$

Given the set of plate parameters

$$\mathcal{P} = \{\rho, E, \nu, L_x, L_y, h, T_0, \eta_0, \eta_1, x_i, y_i, x_o, y_o\}, \quad (9)$$

Task A of the challenge requires estimating the parameter subset

$$\mathcal{S} = \{\mu, D/\mu, T_0/\mu, L_y, x_o, y_o\}, \quad (10)$$

where $\mu = \rho h$, from a single impulse response, with L_x, ν, τ_0, τ_1, x_i, y_i fixed and known.

3. PROPOSED METHOD

3.1. Differentiable Modal Synthesizer

The core of our approach is a differentiable implementation of the modal plate synthesizer in PyTorch, realized as a `nn.Module` whose forward pass computes the plate impulse response from the six learnable parameters in $\mathcal{S}$. A loss function is then evaluated between the synthesized and reference impulse responses, and the parameters are optimized via gradient-based updates using derivatives obtained through reverse-mode automatic differentiation [25].

Rather than computing the modal grid bounds dynamically from the current parameter estimates as in [21], we fix $N_x = 110$ and $N_y = 439$, where $N_x = \max(m_x)$ and $N_y = \max(m_y)$. These values are determined analytically as the worst-case upper bounds over the full parameter ranges defined in Table I of [21]. Specifically, they correspond to the maximum number of modes below $\Omega_{\max}$ achievable by any valid plate configuration, obtained by evaluating (5) at the extremes of the parameter space. This choice proved more stable than dynamically recomputing N_x and N_y at every training iteration based on the current parameter estimates, as modes entering or leaving the valid set were observed to result in inconsistent convergence behavior. By contrast, fixed bounds ensure a constant computational graph topology throughout training, at the cost of evaluating a number of out-of-range modes, i.e., modes with $\Omega_m > \Omega_{\max}$ for the current parameter estimates. These are subsequently discarded via boolean indexing: a mask $\Omega_m \leq \Omega_{\max}$ is applied to the modal frequency vector Ω_m immediately after computing (5), retaining only the valid modes before they enter the modal sum in (12).

Rather than implementing the recursive update of (7), we derive the closed-form solution of the damped harmonic oscillator excited by a unit pulse at $k = 0$, with $k = 0, \ldots, K - 1$ the discrete time index:

$$q_m[k] = \frac{b_m}{\sin(\Omega_m T_\mathrm{s})} r_m^{k-1} \sin(\Omega_m T_\mathrm{s} k). \quad (11)$$

The output signal is thus obtained by summing all modal contributions:

$$y[k] = \sum_{m=1}^{M} q_m[k], \quad (12)$$

where $M \leq N_x N_y$ is the number of modes satisfying $\Omega_m \leq \Omega_{\max}$. This formulation avoids sequential state updates, enabling fully parallelized computation across both modes and time steps. To manage memory during backpropagation, the modal sum is computed in chunks of 300 modes, accumulating gradients incrementally without materializing the full $M \times K$ matrix.

The model operates at double precision (`float64`) throughout, which we found essential for numerical stability given the wide dynamic range of modal amplitudes.

3.2. Parameter Mapping

The six physical parameters span several orders of magnitude and are subject to box constraints defined by the challenge specification [21]. Let $\theta \in [\theta_{\min}, \theta_{\max}] \subset \mathbb{R}$ be a learnable parameter. Optimizing θ directly in its physical domain is challenging, as gradient steps of uniform size would be inappropriate across parameters with vastly different scales. Therefore, for each parameter, we introduce an unconstrained surrogate $\tilde{\theta} \in \mathbb{R}$, mapped to its physical counterpart θ via a differentiable bijection. The full set of unconstrained surrogates is

$$\tilde{\mathcal{S}} = \{\tilde{\mu}, \widetilde{D/\mu}, \widetilde{T_0/\mu}, \tilde{L}_y, \tilde{x}_o, \tilde{y}_o\}. \tag{13}$$

For parameters with logarithmically distributed ranges, i.e., μ, D/μ, T_0/μ, we apply a log-sigmoid mapping:

$$\theta = \exp\left(\log\theta_{\min} + g(\alpha\,\tilde{\theta}) \cdot \log\frac{\theta_{\max}}{\theta_{\min}}\right), \tag{14}$$

where $g(z) = 1/(1+e^{-z})$ is the standard logistic function, and $\alpha > 0$ is a weight controlling the slope.

For parameters with linearly distributed ranges, i.e., L_y, x_o, y_o, we use a linear-sigmoid mapping:

$$\theta = \theta_{\min} + g(\alpha\,\tilde{\theta}) \cdot (\theta_{\max} - \theta_{\min}). \tag{15}$$

Both (14) and (15) are smooth and strictly monotone, guaranteeing that the physical parameters θ remain within their valid ranges for any value of the raw parameters $\tilde{\theta}$, while preserving well-defined gradients throughout. In the following, we set $\alpha = 2$ for D/μ and T_0/μ, and $\alpha = 1$ for the remaining parameters.

3.3. Loss Function

Parameter estimation is driven by a MSS loss function operating in the time-frequency domain at different resolutions [10]. Let $\mathbf{S}^{(w)} \in \mathbb{C}^{F\times N}$ denote the Short-Time Fourier Transform (STFT) of a signal, computed with FFT size w, hop length $w/4$, and Hann window, where F and N denote the number of frequency bins and time frames, respectively, and $\mathbf{S}^{(w)}_{n,f}$ denotes the entry of $\mathbf{S}^{(w)}$ at frequency bin $f \in \{1,\ldots,F\}$ and time frame $n \in \{1,\ldots,N\}$. The loss function compares the predicted and target impulse responses across multiple STFT resolutions, corresponding to the set of window sizes $\mathcal{W} = \{64, 128, 256, 512, 1024, 2048, 4096\}$. For $w \in \mathcal{W}$, we have

$$\mathcal{L}^{(w)} = \frac{\|\,|\hat{\mathbf{S}}^{(w)}| - |\mathbf{S}^{(w)}|\,\|_F}{\|\,|\mathbf{S}^{(w)}|\,\|_F} + \frac{1}{NF}\sum_{n=1}^{N}\sum_{f=1}^{F}\left|\log\frac{|\hat{\mathbf{S}}^{(w)}_{n,f}|}{|\mathbf{S}^{(w)}_{n,f}|}\right|, \tag{16}$$

where $\mathbf{S}^{(w)}$ and $\hat{\mathbf{S}}^{(w)}$ are the target and predicted complex STFTs at scale w, respectively, $|\cdot|$ denotes the magnitude, and $\|\cdot\|_F$ denotes the Frobenius norm. The first term in (16) penalizes large spectral differences, whereas the log-magnitude term emphasizes low-energy components such as late modal decay. The MSS loss is obtained by averaging over all scales, i.e.,

$$\mathcal{L}_{\text{MSS}} = \frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}}\mathcal{L}^{(w)}, \tag{17}$$

where $|\mathcal{W}|$ denotes the cardinality of $\mathcal{W}$.

3.4. Two-Phase Optimization

The loss surface of the plate parameter estimation problem is highly non-convex. Specifically, dense modal overlap and closely spaced resonant frequencies can cause distinct parameter sets to produce similar spectro-temporal features, giving rise to poor yet stable local minima. Consequently, the optimizer may become trapped in the basin of one of many suboptimal solutions rather than recovering the true underlying parameters. Because a single gradient descent run from a random initialization is unlikely to navigate such a complex landscape to the global optimum, we adopt a two-phase strategy aimed at separating exploration from exploitation.

Phase 1: Multi-start probing. We sample 500 candidate initializations using LHS over the six-dimensional physical parameter space $\mathcal{S}$, ensuring uniform coverage of the feasible region across all six target parameters. Each candidate is subjected to a short gradient descent run of 100 iterations using the Adam optimizer [26], with learning rates of 0.01 for $\{\mu, L_y, x_o, y_o\}$ and 0.05 for $\{D/\mu, T_0/\mu\}$. To reduce computational cost, the probes operate on a truncated impulse response of 0.2 s, use a smaller set of STFT scales, namely $\{256, 1024, 2048\}$, and impose a reduced frequency ceiling $\Omega^{\text{probe}}_{\max} = 2\pi \cdot 2500$ rad/s, thereby restricting early optimization to the most prominent low-frequency modes. This provides a robust macroscopic alignment before the high-frequency details and late modal decay are refined in Phase 2. Gradient norms are clipped at 1 to prevent instability during early exploration. Probes are terminated early if the loss remains above 1 after 20 iterations, and Phase 1 is stopped early if any candidate reaches a loss below 0.4. The candidate yielding the lowest final probe loss is selected as the initialization for Phase 2.

Phase 2: Full-scale optimization. Starting from the best probe, we run 1500 iterations of gradient descent using the Adam optimizer, with full frequency ceiling $\Omega_{\max} = 2\pi \cdot 10^4$ rad/s and all seven STFT scales active. A duration curriculum progressively increases the length of the impulse response from 0.05 s to 3 s over the first 1000 iterations, allowing the optimizer to first match the early transient before fitting the late decay. As in Phase 1, gradient norms are clipped at 1 throughout Phase 2 to prevent large, destabilizing parameter updates during training. Learning rates are set to 0.01 for all parameters, and a `ReduceLROnPlateau` scheduler reduces them by a factor of 10 after 500 iterations without improvement, with a minimum learning rate of 10^{-4}.

4. EVALUATION

4.1. Experimental Setup

We evaluate our method on eight synthetic impulse responses generated with the official challenge dataset generator [21], with parameters uniformly sampled from the predefined ranges specified by the challenge. The fixed parameters are set as per Task A specification: $L_x = 1.0$ m, $\nu = 0.25$, $\tau_0 = 6$ s, $\tau_1 = 2$ s, $x_i = 0.335\,L_x$, and $y_i = 0.467\,L_y$. All impulse responses are generated at $f_s = 44.1$ kHz with a duration of 5 s. The experiments are run on a single workstation equipped with an AMD Ryzen 9 9900X CPU with 32 GB of DDR5 RAM (CL36), and an NVIDIA GeForce RTX 5070 Ti GPU (PNY, 16 GB VRAM). We compare our differentiable modal plate synthesizer (DMPS) against two baselines: a naive constant predictor (Const.), which always outputs the midpoint of each parameter's admissible range, and the PSO baseline provided by the challenge organizers [21].

4.2. Numerical Results

Performance is evaluated using the Normalized Mean Squared Error (NMSE), defined as [21]

$$\text{NMSE} = \frac{1}{|\mathcal{S}|}\sum_{\theta\in\mathcal{S}}\frac{(\hat{\theta}-\theta)^2}{(\theta_{\max}-\theta_{\min})^2}, \tag{18}$$

where $\mathcal{S}$ is the parameter set defined in (10), $\hat{\theta}$ is the estimated value of parameter θ, and $[\theta_{\min}, \theta_{\max}]$ is its admissible range as

Table 1: Ground truth parameters (GT), constant predictor (Const.), PSO baseline, and proposed method (DMPS) estimates for the eight test impulse responses.

IR	Method	μ	D/μ	T_0/μ	L_y (m)	x_o (m)	y_o (m)
1	GT	15.566	11.723	8.144	3.142	0.600	0.806
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	31.444	2.788	14.411	3.078	0.650	0.823
	DMPS	15.566	11.723	8.143	3.142	0.600	0.806
2	GT	55.173	24.168	8.147	2.566	0.947	0.853
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	42.847	21.554	11.714	3.356	0.932	0.916
	DMPS	21.729	24.179	8.131	2.566	0.980	0.852
3	GT	72.830	7.002	5.038	2.562	0.842	0.840
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	51.983	10.247	17.985	3.161	0.649	0.956
	DMPS	73.068	7.001	5.032	2.562	0.841	0.840
4	GT	15.922	26.952	19.958	2.571	0.884	0.936
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	13.735	23.316	51.892	2.897	0.894	0.958
	DMPS	15.937	26.950	19.958	2.571	0.884	0.936
5	GT	40.354	19.259	20.006	3.866	0.761	0.895
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	83.348	8.128	11.214	2.846	0.860	0.786
	DMPS	40.356	19.261	20.008	3.866	0.761	0.895
6	GT	42.878	3.109	10.269	1.770	0.994	0.624
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	63.278	17.735	4.530	2.662	0.952	0.791
	DMPS	16.115	3.105	2.551	1.514	0.999	0.637
7	GT	74.826	24.545	11.673	3.686	0.520	0.605
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	91.529	22.315	0.428	3.062	0.834	0.819
	DMPS	82.948	20.133	55.443	3.508	0.512	0.605
8	GT	34.524	15.035	9.563	3.028	0.765	0.816
	Const.	35.485	15.194	23.200	2.550	0.755	0.755
	PSO	58.451	6.154	11.126	2.424	0.831	0.770
	DMPS	34.524	15.035	9.563	3.028	0.765	0.816

specified in Table I of [21]. Normalizing each squared error by the corresponding range ensures that all six parameters, despite spanning very different physical scales, contribute comparably to the overall metric.

Table 1 reports the ground truth and estimated values for all six parameters across the eight test impulse responses (IR 1–IR 8). Table 2 reports the per-IR and average NMSE for the three methods. On average, DMPS achieves an NMSE of 4.29×10^{-3}, roughly one order of magnitude lower than both PSO (3.98×10^{-2}) and the constant predictor (4.61×10^{-2}), confirming that gradient-based optimization of the differentiable synthesizer consistently outperforms gradient-free search on this task.

For five of the eight impulse responses (IR 1, IR 3, IR 4, IR 5, and IR 8), DMPS recovers all six parameters almost exactly, with NMSE ranging from 2.22×10^{-11} to 1.08×10^{-6}. In these cases the estimated values match the ground truth to at least three significant digits for every parameter.

The three remaining cases (IR 2, IR 6, IR 7) account for most of the average (Avg.) NMSE reported in the last row of Table 2. Inspection of the per-parameter estimates in Table 1 reveals that the discrepancy is concentrated in the $\{\mu, D/\mu, T_0/\mu\}$ subset. For IR 2, the surface density μ is considerably underestimated (21.73 vs. 55.17), whereas D/μ, T_0/μ, L_y, x_o, y_o are recovered accurately. For IR 7, conversely, T_0/μ converges to a value nearly five times larger than the ground truth (55.44 vs. 11.67), whereas μ, L_y, x_o, y_o are well estimated. IR 6 shows a combination of the two effects, with μ (16.12 vs. 42.88), T_0/μ (2.55 vs. 10.27), and L_y (1.51 vs. 1.77) all deviating from the ground truth.

Table 2: Per-IR and average NMSE for the constant predictor (Const.), the PSO baseline, and the proposed method (DMPS).

IR	Const.	PSO	DMPS
1	3.17×10^{-2}	6.30×10^{-3}	6.99×10^{-11}
2	3.86×10^{-2}	1.77×10^{-2}	1.81×10^{-2}
3	3.25×10^{-2}	4.91×10^{-2}	1.08×10^{-6}
4	4.08×10^{-2}	3.65×10^{-3}	1.15×10^{-8}
5	4.83×10^{-2}	6.49×10^{-2}	7.63×10^{-11}
6	6.54×10^{-2}	4.37×10^{-2}	1.26×10^{-2}
7	1.04×10^{-1}	1.12×10^{-1}	3.66×10^{-3}
8	7.37×10^{-3}	2.08×10^{-2}	2.22×10^{-11}
Avg.	4.61×10^{-2}	3.98×10^{-2}	4.29×10^{-3}

Nevertheless, Table 2 shows that the proposed method outperforms the constant predictor on all eight cases. Furthermore, when compared to PSO, DMPS achieves superior results on two of the three most challenging cases (IR 6 and IR 7), while performing comparably to the baseline method on IR 2.

The full two-phase optimization pipeline required a total of 6443.58 s (approximately 1.79 hours) to process all eight impulse responses, corresponding to an average runtime of 805.45 s (approximately 13.42 minutes) per impulse response.

5. CONCLUSIONS

We presented a gradient-based method for physical parameter estimation of plate reverberation, developed as a submission to Task A of the 1st DAFx Parameter Estimation Challenge. The proposed approach embeds a differentiable modal plate synthesizer directly into the learning pipeline, enabling end-to-end parameter identification via gradient-based optimization, with derivatives computed via reverse-mode automatic differentiation. A sigmoid-based parameter mapping enforces physical constraints, while a two-phase training strategy addresses the non-convexity of the loss landscape by employing multi-start probing followed by a full-scale gradient descent stage. Experiments on eight synthetic impulse responses generated with random parameters show that the proposed method achieves an average NMSE of 4.29×10^{-3}, outperforming both a PSO baseline (3.98×10^{-2}) and a constant predictor (4.61×10^{-2}) by roughly one order of magnitude. For five of the eight cases, parameter estimation is highly accurate, matching the ground truth to at least three significant digits for every parameter, whereas the remaining three cases exhibit larger errors concentrated in the $\{\mu$, D/μ, $T_0/\mu\}$ subset. The full pipeline required on average 13.42 minutes per impulse response. Future work could investigate the robustness of the proposed method to measurement noise, better strategies to escape local minima, and its evaluation on real measured plate impulse responses, where model mismatch may introduce additional challenges.

6. REFERENCES

[1] A. Chaigne and C. Lambourg, "Time-domain simulation of damped impacted plates. I. Theory and experiments," *The Journal of the Acoustical Society of America*, vol. 109, no. 4, pp. 1422–1432, 2001.

[2] S. Bilbao and M. Van Walstijn, "A finite difference plate model," in *Proceedings of the 31st International Computer Music Conference (ICMC)*, Barcelona, Spain, Sept. 2005, pp. 119–122.

[3] S. Bilbao, K. Arcas, and A. Chaigne, "A physical model for plate reverberation," in *Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006, pp. V–V.

[4] S. Bilbao, "A digital plate reverberation algorithm," *Journal of the Audio Engineering Society*, vol. 55, no. 3, pp. 135–144, 2007.

[5] M. Ducceschi and C. J. Webb, "Plate reverberation: Towards the development of a real-time physical model for the working musician," in *Proceedings of the 22nd International Congress on Acoustics (ICA)*, Buenos Aires, Argentina, Sept. 2016, pp. 1–11.

[6] A. T. Cemgil and C. Erkut, "Calibration of physical models using artificial neural networks with application to plucked string instruments," in *Proceedings of the Institute of Acoustics*, vol. 19, 1997, pp. 213–218.

[7] L. Gabrielli, S. Tomassetti, S. Squartini, and C. Zinato, "Introducing deep machine learning for parameter estimation in physical modelling," in *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx)*, Edinburgh, UK, Sept. 2017, pp. 11–16.

[8] L. Gabrielli, S. Tomassetti, S. Squartini, C. Zinato, and S. Guaiana, "A multi-stage algorithm for acoustic physical model parameters estimation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 8, pp. 1229–1240, 2019.

[9] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the International Conference on Neural Networks (ICNN)*, vol. 4, Perth, WA, Australia, Nov. 1995, pp. 1942–1948.

[10] J. Engel, C. Gu, and A. Roberts, "DDSP: Differentiable digital signal processing," in *Proceedings of the International Conference on Learning Representations (ICLR)*, online, Apr. 2020.

[11] A. I. Mezza, R. Giampiccolo, and A. Bernardini, "Data-driven parameter estimation of lumped-element models via automatic differentiation," *IEEE Access*, vol. 11, pp. 143 601–143 615, 2023.

[12] A. I. Mezza, R. Giampiccolo, E. De Sena, and A. Bernardini, "Data-driven room acoustic modeling via differentiable feedback delay networks with learnable delay lines," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2024, no. 1, p. 51, 2024.

[13] ——, "Differentiable scattering delay networks for artificial reverberation," in *Proceedings of the 28th International Conference on Digital Audio Effects (DAFx)*, Ancona, Italy, Sept. 2025, pp. 202–209.

[14] G. Dal Santo, K. Prawda, S. J. Schlecht, and V. Välimäki, "Differentiable feedback delay network for colorless reverberation," in *Proceedings of the 26th International Conference on Digital Audio Effects (DAFx)*, Copenhagen, Denmark, Sept. 2023, pp. 244–251.

[15] ——, "Optimizing tiny colorless feedback delay networks," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2025, no. 1, p. 13, 2025.

[16] A. I. Mezza, R. Giampiccolo, and A. Bernardini, "Modeling the frequency-dependent sound energy decay of acoustic environments with differentiable feedback delay networks," in *Proceedings of the 27th International Conference on Digital Audio Effects (DAFx)*, Ancona, Italy, Sept. 2024, pp. 238–245.

[17] R. Giampiccolo, A. I. Mezza, S. Koyama, M. Pezzoli, A. Bernardini, F. Antonacci, and A. Sarti, "On the role of trainable parameters in differentiable feedback delay networks," in *Proceedings of the 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2026, pp. 22 597–22 601.

[18] F. Esqueda, B. Kuznetsov, and J. D. Parker, "Differentiable white-box virtual analog modeling," in *Proceedings of the 24th International Conference on Digital Audio Effects (DAFx)*, Vienna, Austria, Sept. 2021, pp. 41–48.

[19] O. Massi, A. I. Mezza, R. Giampiccolo, L. Casale, and A. Bernardini, "Gradient-based optimization of mems loudspeaker equivalent circuit models via automatic differentiation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 72, no. 10, pp. 1468–1472, 2025.

[20] R. Diaz and M. Sandler, "Fast differentiable modal simulation of non-linear strings, membranes, and plates," in *Proceedings of the 28th International Conference on Digital Audio Effects (DAFx)*, Ancona, Italy, Sept. 2025, pp. 134–141.

[21] M. Ducceschi and L. Gabrielli, "The 1st DAFx parameter estimation challenge: A benchmark for plate reverb system identification," Boston, MA, USA, Sept. 2026.

[22] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, Vancouver, Canada, Dec. 2019.

[23] M. D. Mckay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.

[24] J.-M. Adrien, *The missing link: modal synthesis*. Cambridge, MA, USA: MIT Press, 1991, p. 269–298.

[25] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: a survey," *Journal of Machine Learning Research*, vol. 18, no. 153, pp. 1–43, 2018.

[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.