AFX-Research: AN EXTENSIVE AND FLEXIBLE REPOSITORY OF RESEARCH ABOUT AUDIO EFFECTS

Marco Comunità **Joshua D. Reiss**
Centre for Digital Music
Queen Mary University of London, UK

ABSTRACT

We present *AFX-Research*, a repository[1] and associated website[2] gathering scientific literature about research on audio effects. Our database includes publications on topics like: modeling, classification and identification, estimation and regression, removal, style transfer, processing. It also includes review papers, which themselves survey the scientific literature. Furthermore, publications included in our database are not limited to specific methods or techniques (e.g. neural networks or wave digital filters) in an effort to highlight the long tradition of audio effects research and the many different approaches adopted along the decades. While our website contains a table with all details about each publication, the table itself allows to search, filter and order the publications; making it extremely easy to retrieve relevant information. At the same time the repository allows anyone to submit requests to add a new publication or update/modify an existing one. Considering the constant contributions of the audio effects research community to such topics and the fast pacing of machine learning approaches, we hope *AFX-Research* will facilitate up-to-date literature reviews and help explore novel ideas, as well as reorganizing the content in line with new research avenues.

1 Introduction

Research on audio effects has significantly expanded over the last few decades [1, 2, 3], driven by advances in digital signal processing [4], machine learning [5], and auditory perception. As the body of literature grows, so does the need for a centralized repository that not only gathers this research but also offers tools for easy access and exploration. *AFX-Research* aims to fulfill this need by providing a comprehensive database of publications related to audio effects. This repository is designed to be both extensive and flexible, accommodating a wide variety of research topics and methodologies.

The goal of *AFX-Research* is to serve as a go-to resource for researchers, educators, and practitioners in the field of audio effects. By offering a platform where publications can be easily accessed, searched, and organized, we aim to streamline the process of literature review and foster the exploration of new research directions. Additionally, the repository encourages community participation by allowing users to contribute to the database, ensuring that it remains up-to-date with the latest research developments.

2 Repository and Website

The *AFX-Research* repository is hosted on a dedicated website designed to facilitate user interaction with the database. The website features a user-friendly interface (see Fig. 1) that allows for the comprehensive exploration of the repository's contents. Key functionalities of the website include search, filter, and sort options, which enable users to efficiently navigate through the extensive collection of publications.

[1]https://github.com/mcomunita/AFX-Research
[2]https://mcomunita.github.io/AFX-Research

The website also supports community-driven updates, where users can submit new publications or suggest modifications to existing entries. This feature ensures that the repository remains current and continues to reflect the evolving landscape of audio effects research.

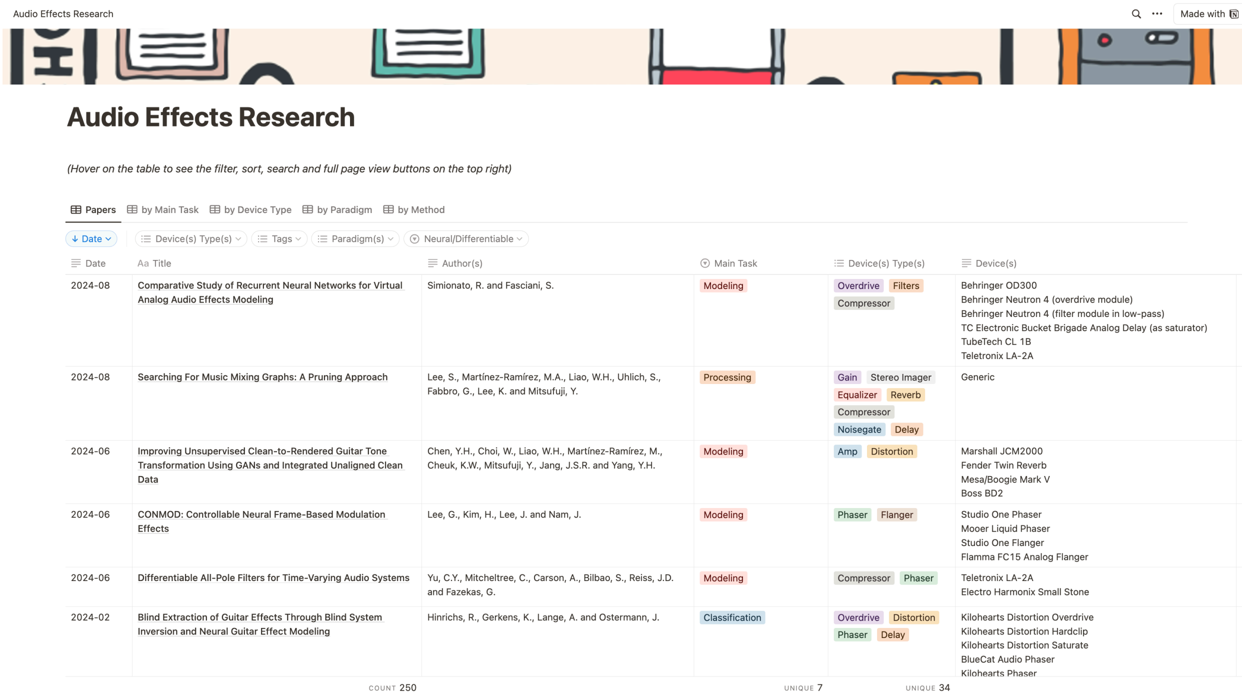
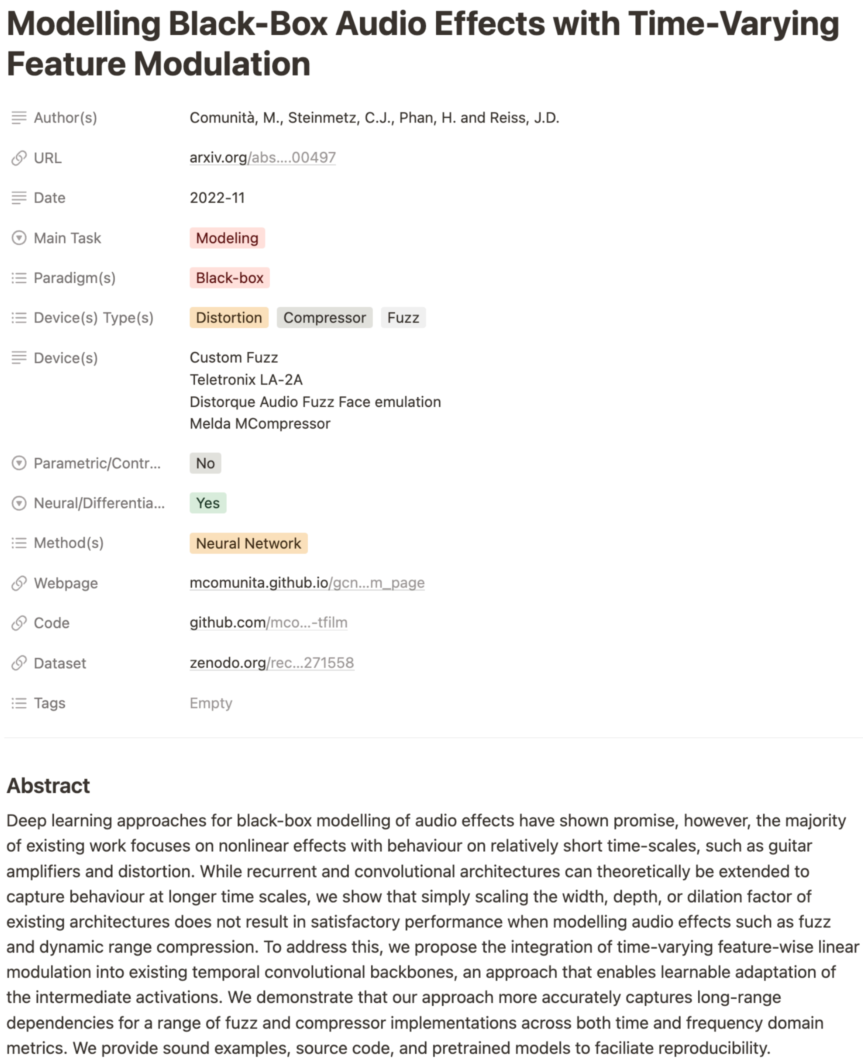


Figure 1: *AFX-Research* website UI - table of publications

3 Database

The database offers detailed metadata for each publication (see Fig. 2), allowing users to assess its relevance without needing to access the full text. At the moment of publication, the metadata includes:

- Title
- Author(s)
- URL: URL to the publication
- Date
- Main Task: classification, estimation, modeling, processing, removal, style transfer, review
- Paradigm(s): what paradigm(s) is the publication using (i.e., black-, gray-, white-box)
- Device(s) Type(s): what type of effects the publication is about (e.g., reverb, delay)
- Device(s): what specific device(s)/circuit(s) are considered (e.g., Ibanez Tube Screamer or vacuum tube stage)
- Parametric/Controllable: whether the approach includes a form of parametric control
- Neural/Differentiable: whether the publication uses approches that support gradient backpropagation
- Method(s): which method(s) or combination of methods is the publication based on (e.g., neural network, Wiener-Hammerstein or state-space)
- Webpage: URL of the page associated with the publication
- Code: URL of the repo associated with the publication
- Dataset: URL of the data associated with the publication
- Abstract



Modelling Black-Box Audio Effects with Time-Varying Feature Modulation

Author(s)	Comunità, M., Steinmetz, C.J., Phan, H. and Reiss, J.D.
URL	arxiv.org/abs....00497
Date	2022-11
Main Task	Modeling
Paradigm(s)	Black-box
Device(s) Type(s)	Distortion Compressor Fuzz
Device(s)	Custom Fuzz Teletronix LA-2A Distorque Audio Fuzz Face emulation Melda MCompressor
Parametric/Contr...	No
Neural/Differentia...	Yes
Method(s)	Neural Network
Webpage	mcomunita.github.io/gcn...m_page
Code	github.com/mco...-tfilm
Dataset	zenodo.org/rec...271558
Tags	Empty

Abstract

Deep learning approaches for black-box modelling of audio effects have shown promise, however, the majority of existing work focuses on nonlinear effects with behaviour on relatively short time-scales, such as guitar amplifiers and distortion. While recurrent and convolutional architectures can theoretically be extended to capture behaviour at longer time scales, we show that simply scaling the width, depth, or dilation factor of existing architectures does not result in satisfactory performance when modelling audio effects such as fuzz and dynamic range compression. To address this, we propose the integration of time-varying feature-wise linear modulation into existing temporal convolutional backbones, an approach that enables learnable adaptation of the intermediate activations. We demonstrate that our approach more accurately captures long-range dependencies for a range of fuzz and compressor implementations across both time and frequency domain metrics. We provide sound examples, source code, and pretrained models to faciliate reproducibility.

Figure 2: *AFX-Research* website UI - example of publication page and metadata

3.1 Tasks

The database categorizes publications based on the primary research tasks they address. These tasks include but are not limited to:

- Classification and identification: studies that classify different types of audio effects (e.g., distortion, phaser, reverb) or identify specific devices (e.g., ProCo Rat distortion, Teletronix LA2A compressor) from audio signals [6, 7, 8].
- Estimation, regression, extraction: works concerned with estimating the controls settings (e.g., gain, cutoff frequency, modulation speed) used to process a certain audio example [6, 7, 9] or related to estimating the internal coefficients of certain processing blocks (e.g., allpass filter, biquad filter, low-frequency oscillator) [10, 11, 12].
- Modeling: research focused on developing mathematical or computational models of audio effects [13, 14, 15, 16, 17, 18].
- Removal: research aimed at removing audio effects from processed signals [19, 20, 21].
- Style Transfer: studies concerned with replicating the sonic characteristics of a reference audio example when applied to an input audio example, regardless of the content or specific audio effects and effects implementations used to process the reference and the input examples [22, 23, 24, 25].

- Processing: broad category of research about processing audio signals. This includes: automatic audio effects control [26], automatic mixing [27], audio processing graph estimation [28, 29], anti-aliasing techniques [30], creative uses of audio effects or derivation of new audio effects that do not strictly model specific devices [31, 32].
- Review: overviews of a specific subtopic or task in the field of audio effects research [1, 3, 5].

3.2 Paradigms

The database includes publications that employ various modeling or emulation paradigms:

- White-box: emulation is based on complete knowledge or thorough understanding of the system (e.g., circuit schematic) and typically employs ordinary/partial differential equations to describe its behaviour and numerical methods to solve them in the continuous or discrete domain [11, 33, 34, 35, 36]. Therefore, such methods are often associated with a time consuming design process and computationally demanding and non-transferable implementations.
- Gray-box: combine a partial theoretical structure - referred to as block-oriented model - with data - typically input/output measurements - to complete the model [15, 37, 38, 39, 40, 41]. Although they reduce prior knowledge necessary to model a device, gray-box approaches still require ad hoc measurements and optimization procedures and knowledge of the underlying implementation.
- Black-box: modeling requires minimal knowledge of the system and mostly relies on input-output measurements. A major advantage is that black-box models simplify the process to collecting adequate data [18, 13, 14, 42]. However, these models often lack interpretability and might entail time-consuming optimizations.

3.3 Methods

The methods section categorizes publications based on the technical approaches and tools used in the research. Common methods include:

- Differentiable DSP: a family of techniques in which loss function gradients are backpropagated through digital signal processors, facilitating their integration into neural networks [43].
- Dynamic Convolution: techniques where the impulse response or processing kernels of a system are varied as a function of the present and/or past input amplitude to model non-linear or hysteretic behaviours of a system [44].
- Equations: techniques where modeling or emulation is based on solving or approximating the physical equations describing a system's behavior. Often based on iterative or numerical methods [45, 46].
- Neural Network: techniques where neural networks are used to solve a task by learning from data [5, 16].
- State-space: the state-space approach represents an electronic circuit as a system of first-order differential equations, describing the circuit's dynamics in terms of state variables, inputs, and outputs [47, 48, 49, 50]. This method uses matrix algebra to model the relationships between circuit components, allowing for efficient digital simulation of complex analog systems. Methods developed for the simulation of state-space systems include the K-method, NK-method and DK-method.
- Wave Digital Filters: a method for digitally modeling analog circuits based on the theory of traveling waves (scattering theory). It uses wave variables instead of standard circuit variables (voltage and current) to represent the behavior of circuit elements [51, 33, 34, 52]. Wave digital filters are designed to preserve key properties of analog circuits, such as passivity and energy conservation, making them stable and robust for digital implementation.
- Port-Hamiltonian: a method for digitally modeling analog circuits based on the principles of Hamiltonian mechanics, focusing on the conservation of energy within a system [53, 54].
- Volterra Series: mathematical tool used for modeling and simulating nonlinear systems. It extends the concept of linear systems by representing a nonlinear system as an infinite series of integral operators, analogous to a Taylor series expansion but in the context of systems and signals [55, 56, 57, 58]. In contrast to linear systems, where the output is directly proportional to the input, nonlinear systems exhibit a more complex relationship. The output in such systems depends on both the current and past inputs in a nonlinear manner. A Volterra series expresses the output of a nonlinear system as a sum of convolutions of the input signal with a series of kernels (functions).

- Waveshaping: application of a nonlinear function to an input audio signal. This function can be mathematically defined or represented as a waveshaping curve. The input signal, usually a simple waveform like a sine wave, is passed through this nonlinear function, which modifies the amplitude of the signal in a non-linear fashion, thereby altering its harmonic content [59, 60].
- Wiener-Hammerstein: a class of models used to emulate the nonlinear and dynamic behavior of audio devices and other systems. These models combine linear and nonlinear elements in a structured way, making them particularly well-suited for capturing the characteristics of audio processing equipment like amplifiers, distortion units, and other effects that exhibit both linear filtering and nonlinear distortion [15, 61, 62, 63]. A Wiener model consists of a linear block followed by a nonlinear block. A Hammerstein model consists of a nonlinear block followed by a linear block. A Wiener-Hammerstein model combines both structures, placing a linear block both before and after the nonlinear block.

4 Conclusion

AFX-Research is a valuable resource for the audio effects research community. By providing a centralized, easily accessible, and regularly updated repository, it facilitates comprehensive literature reviews and supports the exploration of new research ideas. The repository's flexible design, which accommodates a wide range of tasks, paradigms, and methods, ensures that it remains relevant to researchers with diverse interests and expertise. As the field of audio effects continues to evolve, *AFX-Research* will play a crucial role in helping researchers stay informed and inspired.

References

[1] T. Wilmering, D. Moffat, A. Milo, and M.B. Sandler. A history of audio effects. *Applied Sciences*, 10(3):791, 2020.

[2] U. Zölzer, X. Serra, M.B. Sandler, et al. Digital audio effects, 2011.

[3] V. Välimäki, J.D. Parker, L. Savioja, J.O. Smith, and J.S. Abel. Fifty years of artificial reverberation. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(5):1421–1448, 2012.

[4] J.D. Reiss and A. McPherson. *Audio effects: theory, implementation and application*. CRC Press, 2014.

[5] T. Vanhatalo, P. Legrand, M. Desainte-Catherine, P. Hanna, A. Brusco, G. Pille, and Y. Bayle. A review of neural network-based emulation of guitar amplifiers. *Applied Sciences*, 12(12):5894, 2022.

[6] M. Comunità, D. Stowell, and J.D. Reiss. Guitar effects recognition and parameter estimation with convolutional neural networks. *Journal of the Audio Engineering Society*, 69(7/8):594–604, 2021.

[7] R. Hinrichs, K. Gerkens, A. Lange, and J. Ostermann. Convolutional neural networks for the classification of guitar effects and extraction of the parameter settings of single and multi-guitar effects from instrument mixes. *EURASIP Journal on Audio, Speech, and Music Processing*, 2022(1):28, 2022.

[8] J. Guo and B. McFee. Automatic recognition of cascaded guitar effects. In *Proceedings of the International Conference on Digital Audio Effects*. DAFx Board, 2023.

[9] C. Peladeau and G. Peeters. Blind estimation of audio effects using an auto-encoder approach and differentiable digital signal processing. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 856–860. IEEE, 2024.

[10] J.T. Colonel, C.J. Steinmetz, M. Michelen, and J.D. Reiss. Direct design of biquad filter cascades with deep learning by sampling random polynomials. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3104–3108. IEEE, 2022.

[11] A.R. Bargum, S. Serafin, C. Erkut, and J.D. Parker. Differentiable allpass filters for phase response estimation and automatic signal alignment. *arXiv preprint arXiv:2306.00860*, 2023.

[12] B. Kuznetsov, J.D. Parker, and F. Esqueda. Differentiable iir filters for machine learning applications. In *Proc. Int. Conf. Digital Audio Effects (eDAFx-20)*, pages 297–303, 2020.

[13] M. Comunità, C.J. Steinmetz, H. Phan, and J.D. Reiss. Modelling black-box audio effects with time-varying feature modulation. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1–5. IEEE, 2023.

[14] C. Mitcheltree, C.J. Steinmetz, M. Comunità, and J.D. Reiss. Modulation extraction for lfo-driven audio effects. *arXiv preprint arXiv:2305.13262*, 2023.

[15] J.T. Colonel, M. Comunità, and J.D. Reiss. Reverse engineering memoryless distortion effects with differentiable waveshapers. In *Audio Engineering Society Convention 153*. Audio Engineering Society, 2022.

[16] A. Wright, E. Damskägg, L. Juvela, and V. Välimäki. Real-time guitar amplifier emulation with deep learning. *Applied Sciences*, 10(3):766, 2020.

[17] M.A. Martínez Ramírez, E. Benetos, and J.D. Reiss. Deep learning for black-box modeling of audio effects. *Applied Sciences*, 10(2):638, 2020.

[18] R. Simionato and S. Fasciani. Comparative study of recurrent neural networks for virtual analog audio effects modeling. *arXiv preprint arXiv:2405.04124*, 2024.

[19] M. Rice, C.J. Steinmetz, G. Fazekas, and J.D. Reiss. General purpose audio effect removal. In *2023 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5. IEEE, 2023.

[20] J. Imort, G. Fabbro, M.A. Martínez-Ramírez, S. Uhlich, Y. Koyama, and Y. Mitsufuji. Distortion audio effects: Learning how to recover the clean signal. *arXiv preprint arXiv:2202.01664*, 2022.

[21] T. Tanaka, K. Yatabe, M. Yasuda, and Y. Oikawa. Applade: Adjustable plug-and-play audio declipper combining dnn with sparse optimization. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1011–1015. IEEE, 2022.

[22] C.J. Steinmetz, N.J. Bryan, and J.D. Reiss. Style transfer of audio effects with differentiable signal processing. *arXiv preprint arXiv:2207.08759*, 2022.

[23] C.J. Steinmetz, S. Singh, M. Comunità, I. Ibnyahya, S. Yuan, E. Benetos, and J.D. Reiss. St-ito: Controlling audio effects for style transfer with inference-time optimization. In *International Society for Music Information Retrieval (ISMIR) Conference, 2024*, 2024.

[24] S.S. Vanka, C.J. Steinmetz, J. Rolland, J.D. Reiss, and G. Fazekas. Diff-mst: Differentiable mixing style transfer. *arXiv preprint arXiv:2407.08889*, 2024.

[25] J. Koo, M.A. Martínez-Ramírez, W. Liao, S. Uhlich, K. Lee, and Y. Mitsufuji. Music mixing style transfer: A contrastive learning approach to disentangle audio effects. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1–5. IEEE, 2023.

[26] S. Singh, G. Bromham, D. Sheng, and G. Fazekas. Intelligent control method for the dynamic range compressor: A user study. *Journal of the Audio Engineering Society*, 69(7/8):576–585, 2021.

[27] Christian J Steinmetz, Jordi Pons, Santiago Pascual, and Joan Serra. Automatic multitrack mixing with a differentiable mixing console of neural audio effects. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 71–75. IEEE, 2021.

[28] S. Lee, M.A. Martínez-Ramírez, W. Liao, S. Uhlich, G. Fabbro, K. Lee, and Y. Mitsufuji. Searching for music mixing graphs: A pruning approach. *arXiv preprint arXiv:2406.01049*, 2024.

[29] S. Lee, J. Park, S. Paik, and K. Lee. Blind estimation of audio processing graph. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1–5. IEEE, 2023.

[30] F. Esqueda, S. Bilbao, and V. Välimäki. Aliasing reduction in clipped signals. *IEEE Transactions on Signal Processing*, 64(20):5255–5267, 2016.

[31] D.J. Gillespie, S. Schachter, and N.J. Fanwood. Model bending: Teaching circuit models new tricks. In *Proc. Int. Conf. Digital Audio Effects (DAFx20in22), Vienna, Austria*, 2022.

[32] C.J. Steinmetz and J.D. Reiss. Steerable discovery of neural audio effects. *arXiv preprint arXiv:2112.02926*, 2021.

[33] R. Giampiccolo, A. Natoli, A. Bernardini, A. Sarti, et al. Parallel wave digital filter implementations of audio circuits with multiple nonlinearities. *Journal of the Audio Engineering Society*, 70(6):469–484, 2022.

[34] J. Chowdhury and C.J. Clarke. Emulating diode circuits with differentiable wave digital filters. In *Proceedings of the 19th Sound and Music Computing Conference. Zenodo, Saint-Etienne, France*, pages 2–9, 2022.

[35] F. Esqueda, B. Kuznetsov, and J.D. Parker. Differentiable white-box virtual analog modeling. In *2021 24th International Conference on Digital Audio Effects (DAFx)*, pages 41–48. IEEE, 2021.

[36] J.D. Parker, F. Esqueda, and A. Bergner. Modelling of nonlinear state-space systems using a deep neural network. In *Proceedings of the International Conference on Digital Audio Effects (DAFx), Birmingham, UK*, pages 2–6, 2019.

[37] J. Colonel and J.D. Reiss. Approximating ballistics in a differentiable dynamic range compressor. In *Audio Engineering Society Convention 153*. Audio Engineering Society, 2022.

[38] A. Carson, C. Valentini-Botinhao, S. King, and S. Bilbao. Differentiable grey-box modelling of phaser effects using frame-based spectral processing. *arXiv preprint arXiv:2306.01332*, 2023.

[39] C. Yu, C. Mitcheltree, A. Carson, S. Bilbao, J.D. Reiss, and G. Fazekas. Differentiable all-pole filters for time-varying audio systems. *arXiv preprint arXiv:2404.07970*, 2024.

[40] A. Wright and V. Välimäki. Grey-box modelling of dynamic range compression. In *Proceedings of the 25th International Conference on Digital Audio Effects (DAFx20in22)*, pages 304–311, 2022.

[41] S. Nercessian, A. Sarroff, and K.J. Werner. Lightweight and interpretable neural modeling of an audio distortion effect using hyperconditioned differentiable biquads. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 890–894. IEEE, 2021.

[42] G. Lee, H. Kim, J. Lee, and J. Nam. Conmod: Controllable neural frame-based modulation effects. *arXiv preprint arXiv:2406.13935*, 2024.

[43] B. Hayes, J. Shier, G. Fazekas, A. McPherson, and C. Saitis. A review of differentiable digital signal processing for music and speech synthesis. *Frontiers in Signal Processing*, 3:1284100, 2024.

[44] M.J. Kemp. Analysis and simulation of non-linear audio processes using finite impulse responses derived at multiple impulse amplitudes. In *Audio Engineering Society Convention 106*. Audio Engineering Society, 1999.

[45] L. Köper and M. Holters. Taming the red llama—modeling a cmos-based overdrive circuit. In *Proc. 23th Int. Conf. Digital Audio Effects (DAFx-2020), Vienna, Austria*, 2020.

[46] J. Chowdhury. Real-time physical modelling for analog tape machines. In *Proc. Int. Conf. Digital Audio Effects (DAFX)*, page 7, 2019.

[47] M. Bennett. *State-Space Modelling of the Fuzz Face Guitar Pedal*. PhD thesis, McGill University, 2022.

[48] M. Holters and U. Zölzer. A generalized method for the derivation of non-linear state-space models from circuit schematics. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 1073–1077. IEEE, 2015.

[49] K. Dempwolf, M. Holters, and U. Zölzer. Discretization of parametric analog circuits for real-time simulations. In *Proceedings of the 13th international conference on digital audio effects (DAFx'10)*, 2010.

[50] D.T. Yeh. *Digital implementation of musical distortion circuits by analysis and simulation*. Stanford University, 2009.

[51] A. Fettweis. Wave digital filters: Theory and practice. *Proceedings of the IEEE*, 74(2):270–327, 1986.

[52] A. Bernardini, A.E. Vergani, and A. Sarti. Wave digital modeling of nonlinear 3-terminal devices for virtual analog applications. *Circuits, Systems, and Signal Processing*, 39(7):3289–3319, 2020.

[53] A. Van Der Schaft. Port-hamiltonian systems: an introductory survey. In *International congress of mathematicians*, pages 1339–1365. European Mathematical Society Publishing House (EMS Ph), 2006.

[54] A. Falaize and T. Hélie. Passive guaranteed simulation of analog audio circuits: A port-hamiltonian approach. *Applied Sciences*, 6(10):273, 2016.

[55] S. Arora and C. Bennett. Fast algorithms for black box modelling of static hammerstein nonlinearities. In *Audio Engineering Society Convention 153*. Audio Engineering Society, 2022.

[56] S. Orcioni, A. Terenzi, S. Cecchi, F. Piazza, and A. Carini. Identification of volterra models of tube audio devices using multiple-variance method. *Journal of the Audio Engineering Society*, 66(10):823–838, 2018.

[57] T. Schmitz and J. Embrechts. Nonlinear guitar loudspeaker simulation. In *Audio Engineering Society 134th Convention*, 2013.

[58] A. Farina and E. Armelloni. Emulation of not-linear, time-variant devices by the convolution technique. In *Congresso Audio Engineering Society Italia, Como*, 2005.

[59] M. Le Brun. Digital waveshaping synthesis. *Journal of the Audio Engineering Society*, 27(4):250–266, 1979.

[60] C. Roads. A tutorial on non-linear distortion or waveshaping synthesis. *Computer Music Journal*, pages 29–34, 1979.

[61] F. Eichas and U. Zölzer. Gray-box modeling of guitar amplifiers. *Journal of the Audio Engineering Society*, 66(12):1006–1015, 2018.

[62] M. Schoukens and K. Tiels. Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85:272–292, 2017.

[63] A. Novak, L. Simon, F. Kadlec, and P. Lotton. Nonlinear system identification using exponential swept-sine signal. *IEEE Transactions on Instrumentation and Measurement*, 59(8):2220–2229, 2009.