Aalto University
School of Electrical
Engineering

DAFx 25
Digital Audio Effects Conference 2025

Building Robust Audio DDSP Pipelines

A Case Study on Artificial Reverb

Gloria Dal Santo, Aalto University, Finland

gloria.dalsanto@aalto.fi

Gloria Dal Santo

PhD student, **Audio Signal Processing**

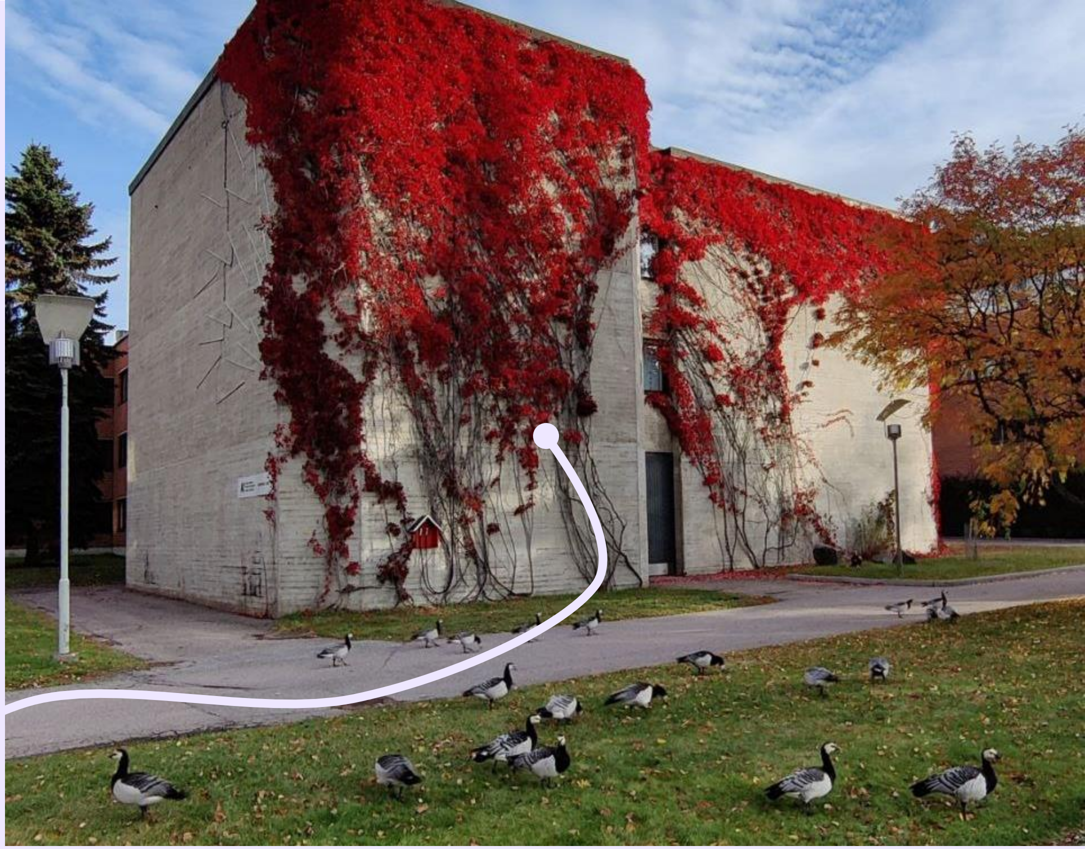
Aalto University (FI)

PhD Team:

- Prof. Vesa Välimäki
- Prof. Sebastian J. Schlecht (FAU, Germany)
- Dr. Karolina Prawda (University of York, UK)
- Many fantastic collaborators

Topic: Machine Learning-based Artificial Reverb

- Feedback Delay Network
- Differentiable DSP
- Audio similarity measures



Outline

- **Differentiable Digital Signal Processing**
  - Definition
  - A few examples
  - Typical scenario and its challenges
- **Building DDSPs with Frequency Sampling**
  - Frequency sampling
  - Advantages and challenges
  - The FLAMO library
- **Differentiable Feedback Delay Network**
  - General FDN structure
  - Differentiable FDN application scenarios
  - Challenges in FDN optimization
- **Choosing the right loss**
  - Analyzing the loss with real data
  - Loss landscape analysis

+ Demos on FLAMO

A! DAFx25

Differentiable Digital Signal Processing

Definition
A few examples
Typical scenario and its challenges

Differentiable DSP

Definition

DDSP: Concept of including digital signal processing components in machine learning models and backpropagating loss function gradients through them

- interpretability of the DSP parameters
- domain knowledge imparted into the model
- Incorporation of known signal models into neural networks

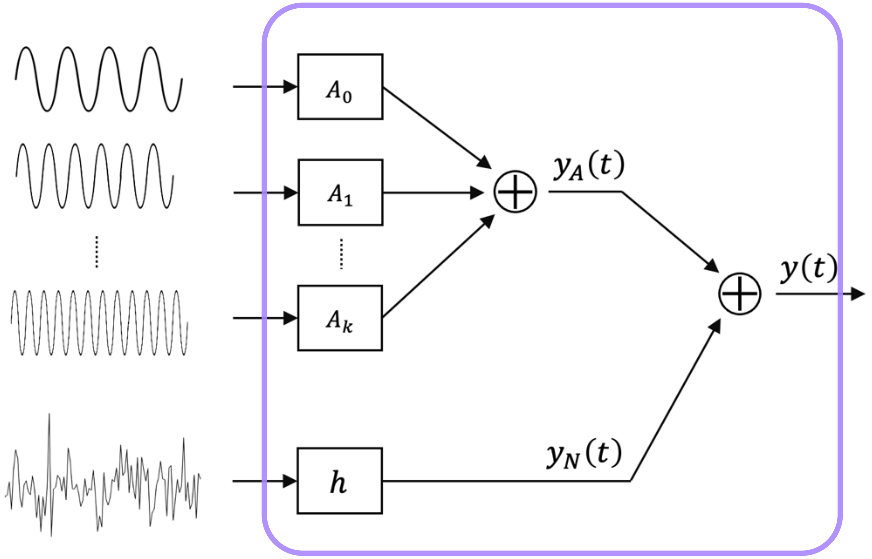
Magenta's seminal work
Jesse H. Engel et al., 2020

Harmonic + Noise model

Additive Synthesizer
Bank of weighted oscillators

Subtractive Synthesizer
Filter the noise source

Differentiable synthesizer



Differentiable DSP

Definition

DDSP: Concept of including digital signal processing components in machine learning models and backpropagating loss function gradients through them

- interpretability of the DSP parameters
- domain knowledge imparted into the model
- Incorporation of known signal models into neural networks

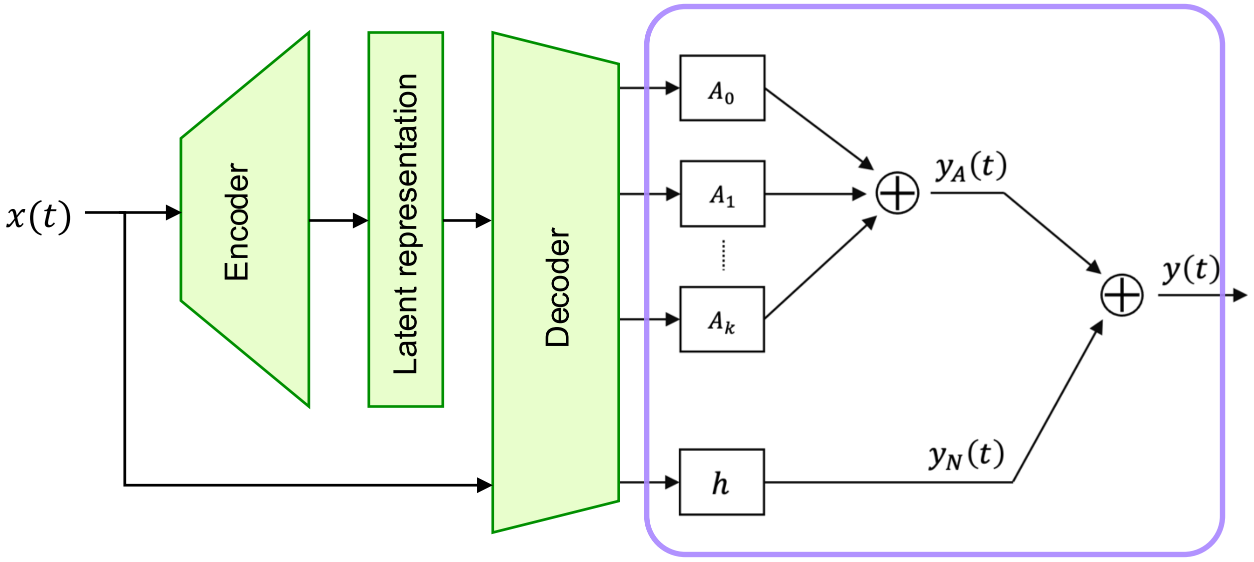
Magenta's seminal work

Jesse H. Engel et al., 2020

<u>Timbre transfer</u>

Violin → Flute

Differentiable synthesizer



Differentiable DSP
A few examples

In this tutorial, we will focus only on the DDSP, and not on the neural net used to derive its parameters

Besides the harmonic + noise model ...

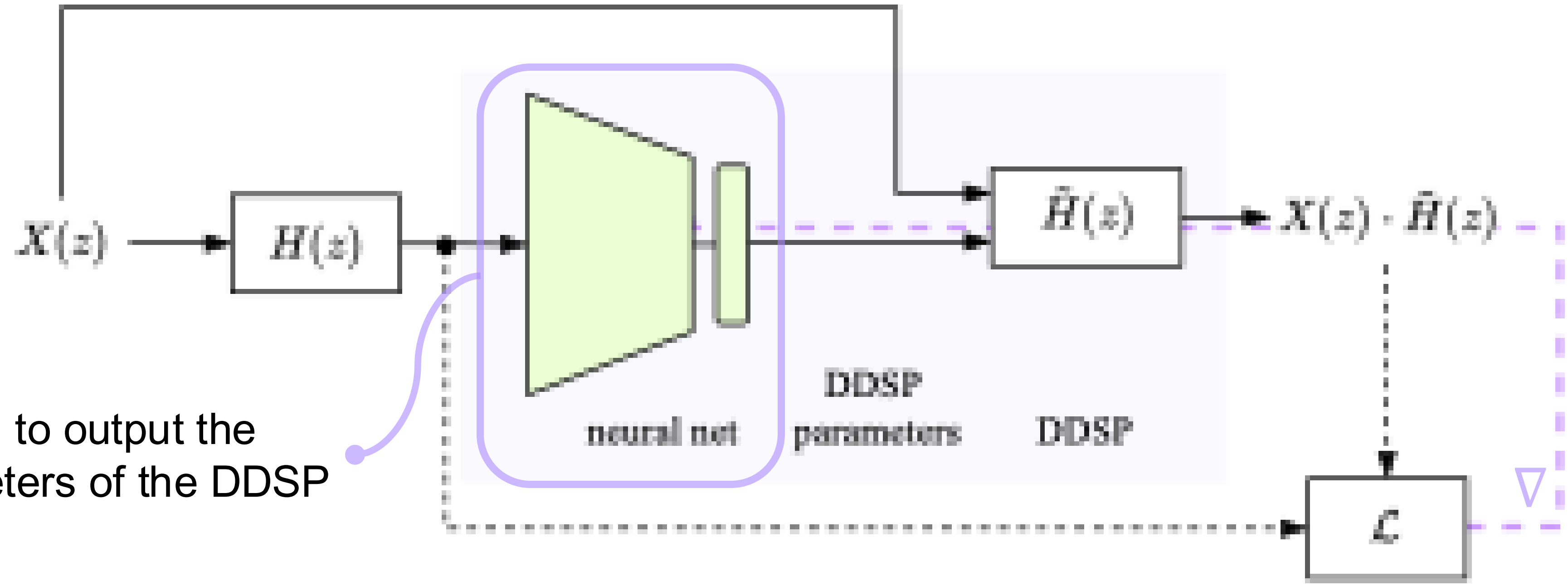
- B. Kuznetsov, J. D. Parker, and F. Esqueda, “**Differentiable IIR filters for machine learning applications**,” DAFx 2020
- S. Nercessian, “**Neural parametric equalizer matching using differentiable biquads**,” DAFx 2020
- P. Bhattacharya, P. Nowak, and U. Zölzer, ”**Optimization of cascaded parametric peak and shelving filters with backpropagation algorithm**.” DAFx 2020
- C. J. Steinmetz, K. I. Vamsi, and P. Calamia. "**Filtered noise shaping for time domain room impulse response estimation from reverberant speech**.” WASPAA, 2021
- S. Shan, L. Hantrakul, J. Chen, M. Avent, and D. Trevelyan, “**Differentiable wavetable synthesis**” ICASSP, 2022
- C. J. Steinmetz, N. J. Bryan, and J. D. Reiss, “**Style transfer of audio effects with differentiable signal processing**,” JAES, 2022
- S. Lee, H.-S. Choi, and K. Lee, “**Differentiable artificial reverberation**,” IEEE/ACM TASLP, 2022
- L. Renault, R. Mignot, and A. Roebel, “**Differentiable piano model for MIDI-to-audio performance synthesis**.” DAFx 2022
- F. Caspe, A. McPherson, and M. Sandler, “**DDX7: Differentiable FM synthesis of musical instrument sounds.**” arXiv:2208.06169, 2022
- G. Dal Santo, K. Prawda, S. Schlecht, and V. Välimäki, “**Differentiable feedback delay network for colorless reverberation**.” DAFx 2023
- A. Carson, S. King, C. Valentini Botinhao, and S. Bilbao, “**Differentiable grey-box modelling of phaser effects using frame-based spectral processing**,” DAFx, 2023
- N. Masuda, and D. Saito , “**Improving semi-supervised differentiable synthesizer sound matching for practical applications**.” IEEE/ACM TASLP, 2023.
- C. Y. Yu, C. Mitcheltree, A. Carson, S. Bilbao, J. D. Reiss, and G. Fazekas, “**Differentiable all-pole filters for time-varying audio systems,**” DAFx, 2024
- Y. Liu, C. Jin, and D. Gunawan, “**DDSP-SFX: Acoustically-guided sound effects generation with differentiable digital signal processing**.” DAFx 2024
- G. M. De Bortoli, G. Dal Santo, K. Prawda, T. Lokki, V. Välimäki, and S. Schlecht, “**Differentiable active acoustics: optimizing stability via gradient descent.**” DAFx2024

- 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, 2024.

... these are just a few examples of relevant papers!

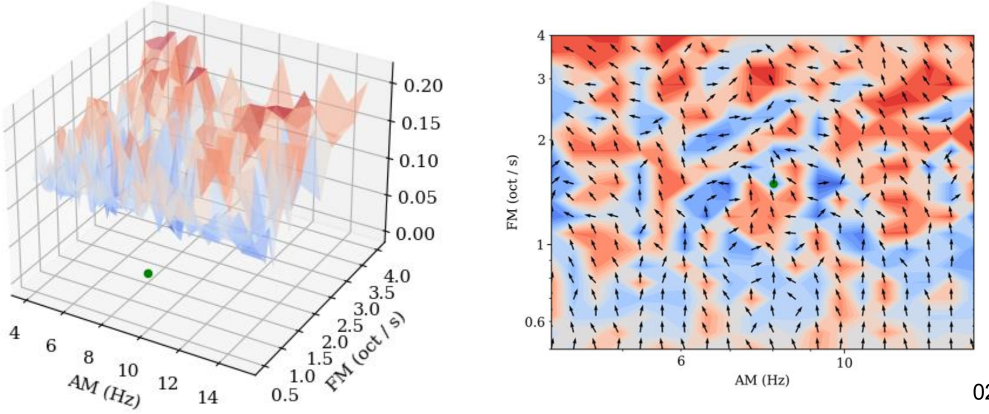
Differentiable DSP

Typical Scenario and its Challenges - Filter



Trained to output the parameters of the DDSP

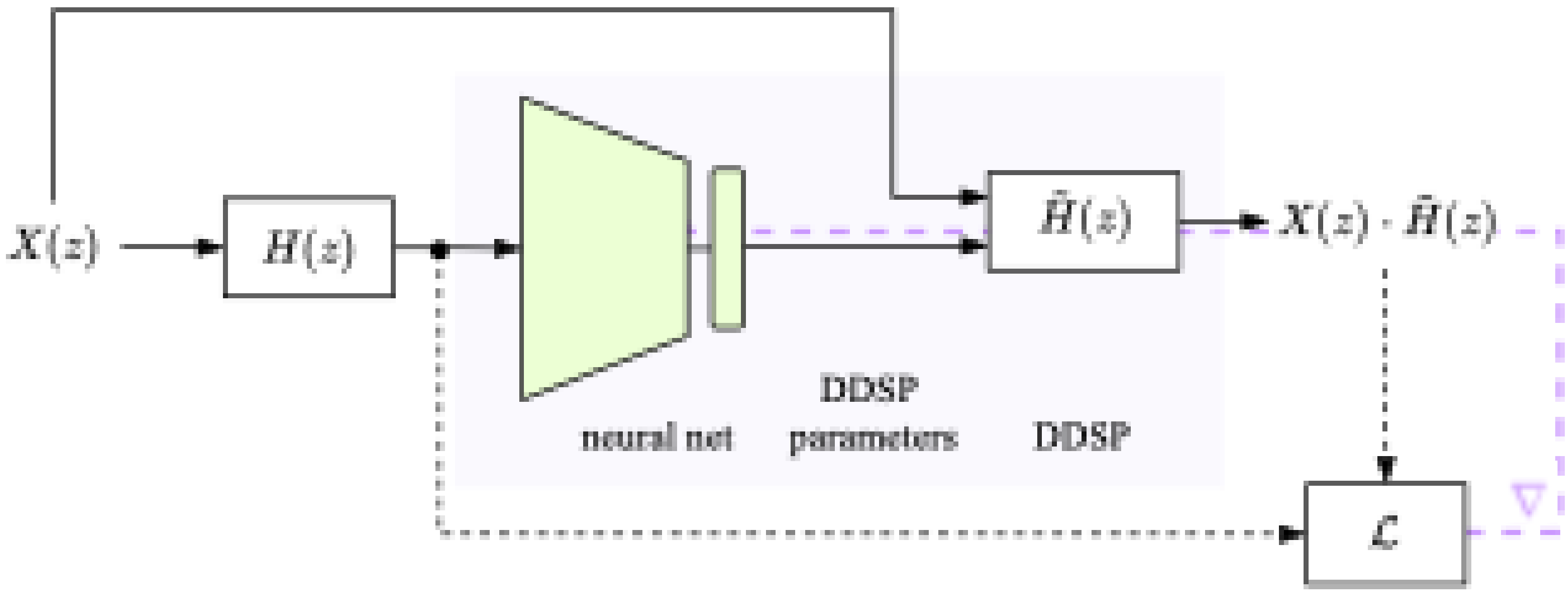
- The **loss landscape** with respect to DDSP parameters can be **highly irregular**
- **Restricts the neural network** to a low-dimensional space with physical meaning
- The desired output is achieved with only **specific DDSP parameter combinations**



Vahidi C. et al. 2023

Differentiable DSP

Typical Scenario and its Challenges - Filter



Key elements when designing DDSP pipelines:

- Differentiable implementation of the DSP
- Domain of stability and plausibility of the learnable parameters
- Smoothness of the loss landscape

No neural network
Just gradient descent optimization

Building DDSP with Frequency Sampling

Frequency Sampling
Advantages and challenges
The FLAMO library
Example: Differentiable Biquads for Parametric EQ

Building DDSP with Frequency Sampling

Differentiable Filter Formulations

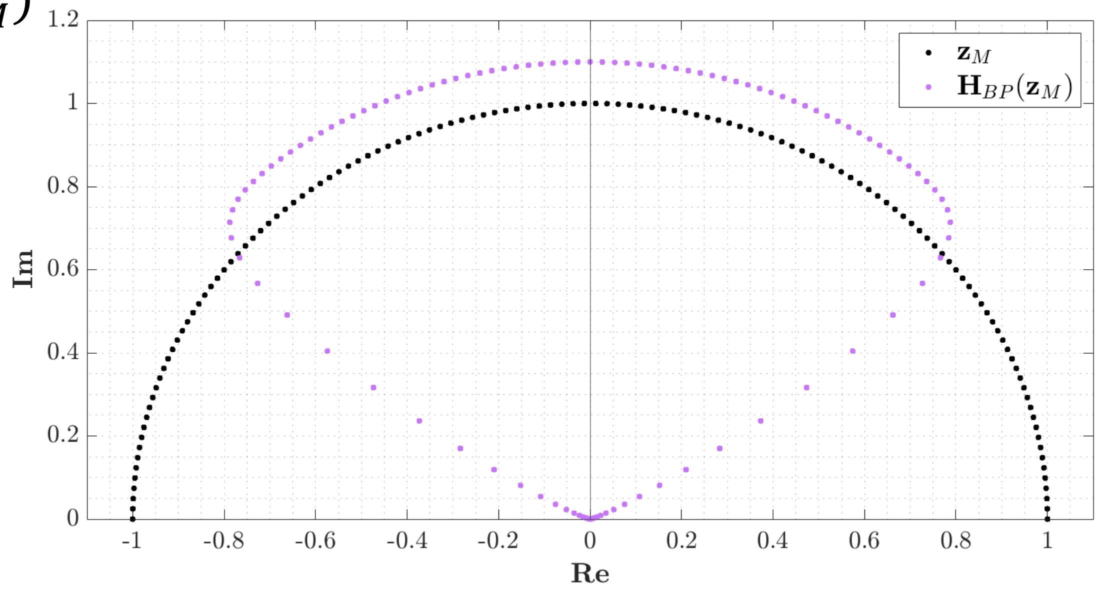
- Backpropagation through time, instantaneous backpropagation
- Convolutional layers with linear activations and zero bias
- **Frequency sampling** of (locally) time-invariant systems

Optimization of the FIR approximation

Transfer function coefficients $\mathbf{b}$ and $\mathbf{a} \rightarrow H(\boldsymbol{z}_M)$

$$\boldsymbol{z}_M = [e^{\jmath\pi\frac{0}{M}}, e^{\jmath\pi\frac{1}{M}}, \dots, e^{\jmath\pi\frac{M-1}{M}}]$$

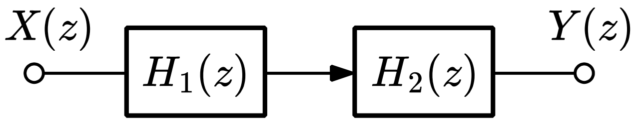
$$H(\boldsymbol{z}_M) = \frac{\mathrm{FFT}(\mathbf{b})}{\mathrm{FFT}(\mathbf{a})} = \frac{b_0 + b_1\boldsymbol{z}_M^{-1} + \cdots + b_N\boldsymbol{z}_M^{-N+1}}{a_0 + a_1\boldsymbol{z}_M^{-1} + \cdots + a_N\boldsymbol{z}_M^{-N+1}}$$

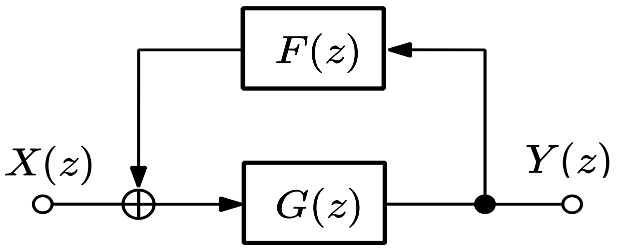


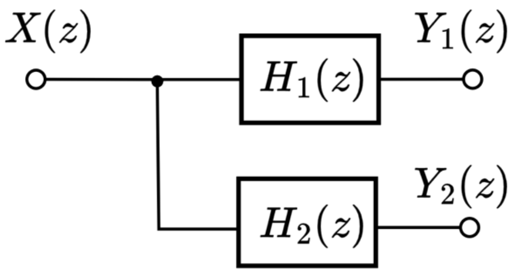
Building DDSP with Frequency Sampling

Advantages

- Each sample is independent → samples of $H(z)$ can be generated in parallel
- Filters can be chained by complex multiplication → efficiency







$$H(\mathbf{z}_M) = H_1(\mathbf{z}_M)H_2(\mathbf{z}_M)$$

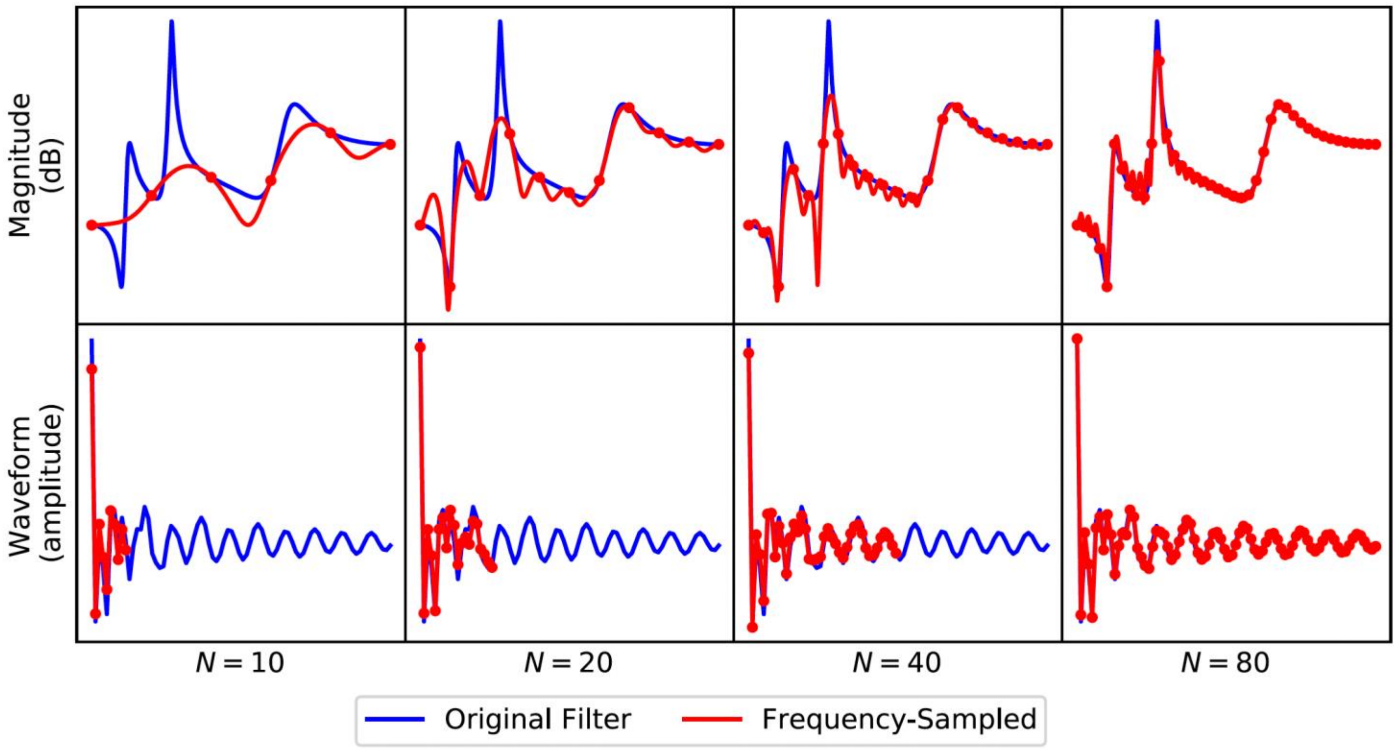
$$H(\mathbf{z}_M) = \left(I - G(\mathbf{z}_M)F(\mathbf{z}_M)\right)^{-1}G(\mathbf{z}_M)$$

$$H(\mathbf{z}_M) = \operatorname{diag}(H_1(\mathbf{z}_M), H_2(\mathbf{z}_M))$$

Building DDSP with Frequency Sampling
Disadvantages

- Sensitivity to the number of frequency samples → **time-aliasing**

  The number of frequency sampling points *N* depends on the decay rate of the modeled IR



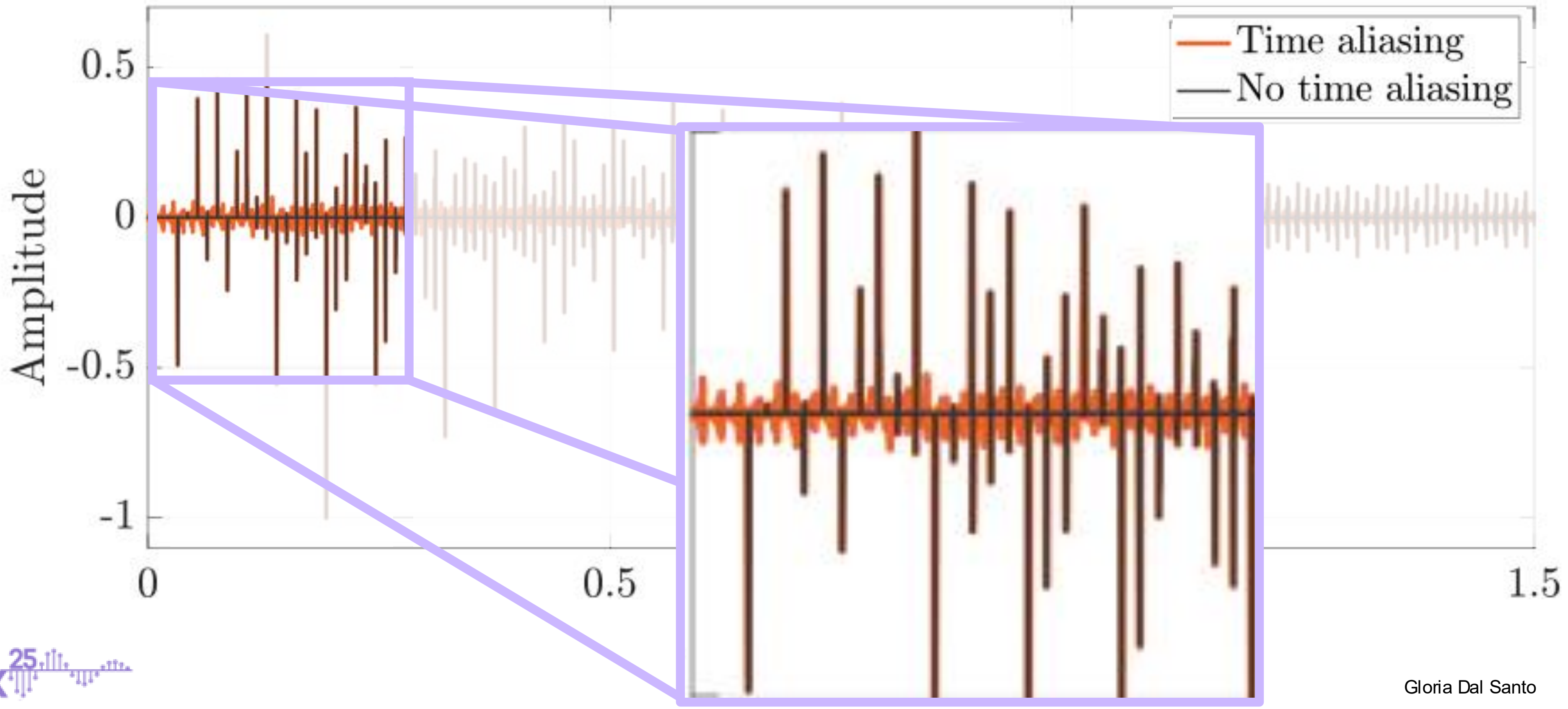
Lee S. et al. 2022

Building DDSP with Frequency Sampling

Disadvantages

- Sensitivity to the number of frequency samples → **time-aliasing**

  The number of frequency sampling points *N* depends on the decay rate of the modeled IR



Building DDSP with Frequency Sampling

Mitigating the time aliasing

- Frequency domain: sampling the frequency response outside the unit circle

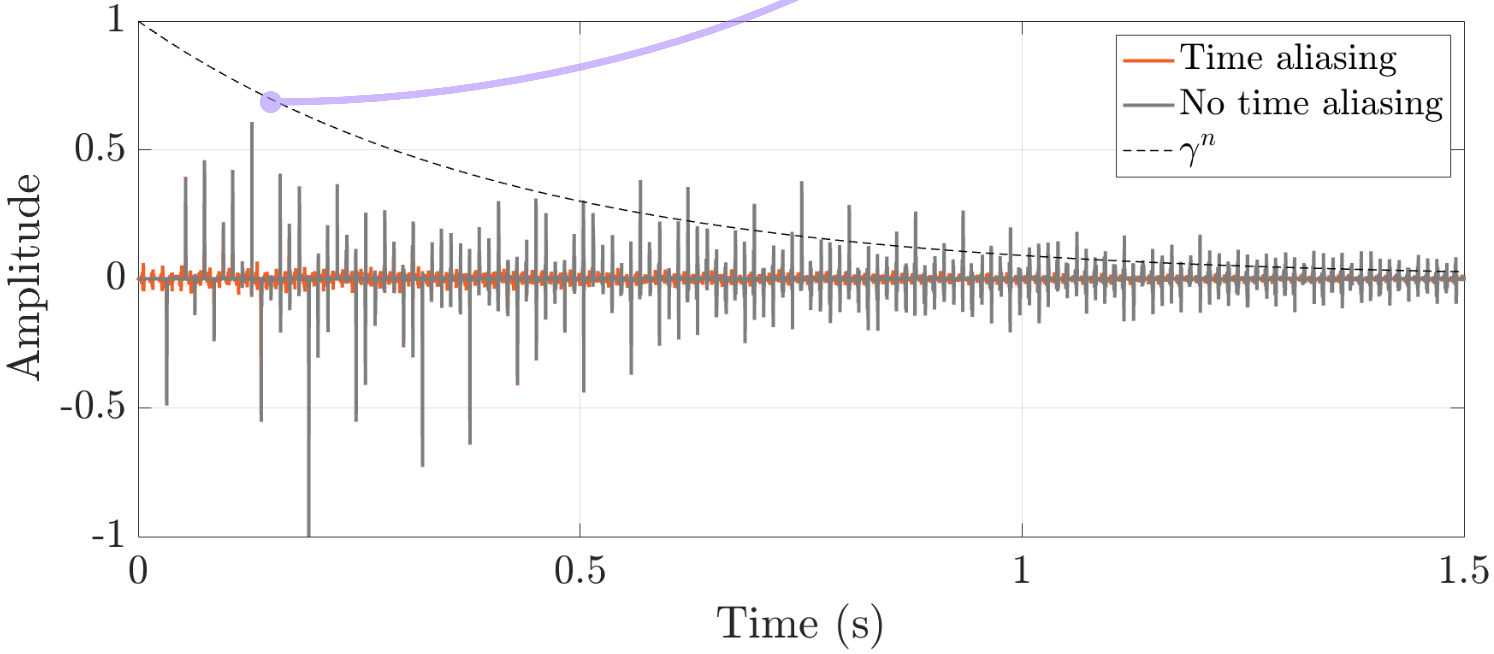
$$\hat{H}(e^{\jmath\omega}) = H(e^{\jmath\omega}/\gamma)$$

- Time domain envelope: exponentially decaying function γ^n where $0 < \gamma \leq 1$

$$\hat{h}[n] = h[n]\gamma^n$$

$$h[n] = \hat{h}[n]\gamma^{-n} = \mathrm{IDFT}(\hat{H}(e^{\jmath\omega}))\gamma^{-n}$$

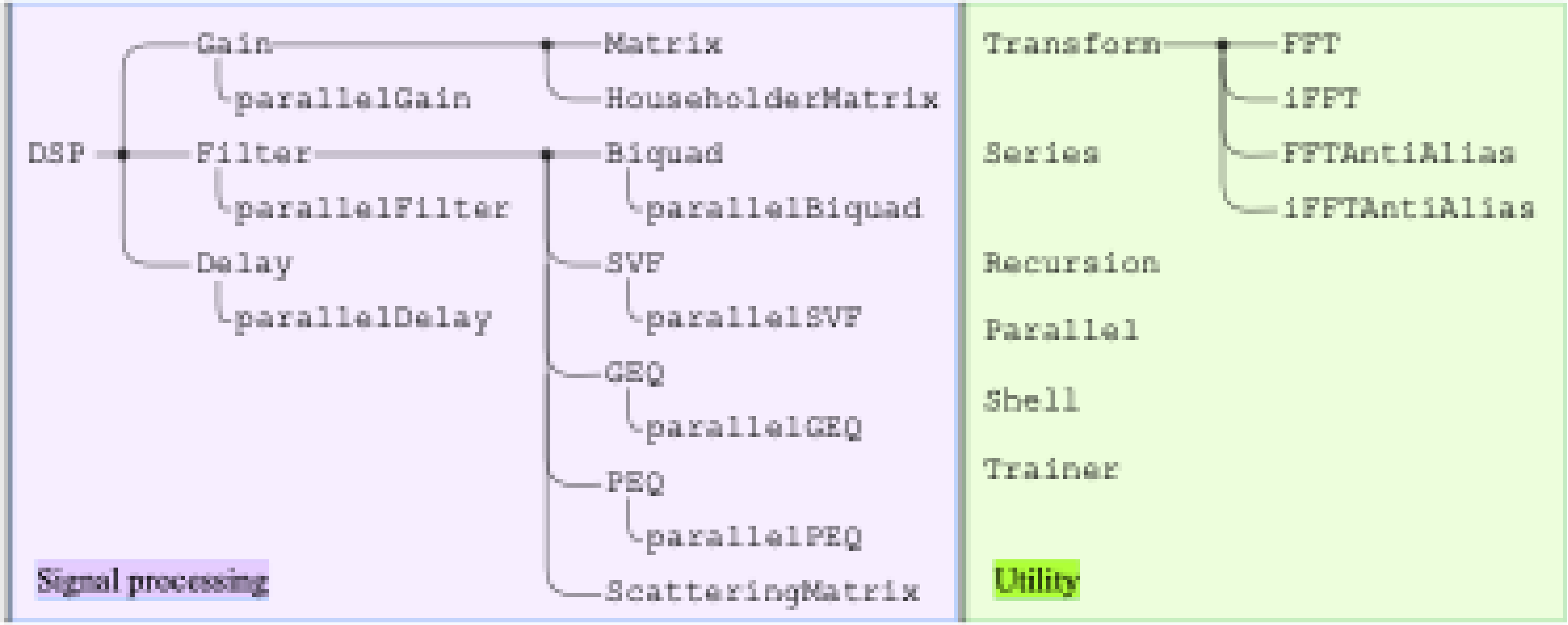
! It will amplify numerical noise



Building DDSP with Frequency Sampling

FLAMO library (ICASSP '25)

- **FLAMO**: **F**requency-sampling **L**ibrary for **A**udio-**M**odule **O**ptimization
- Open-source library for implementing and optimizing differentiable LTI audio systems
- Modules support the time-aliasing mitigation system

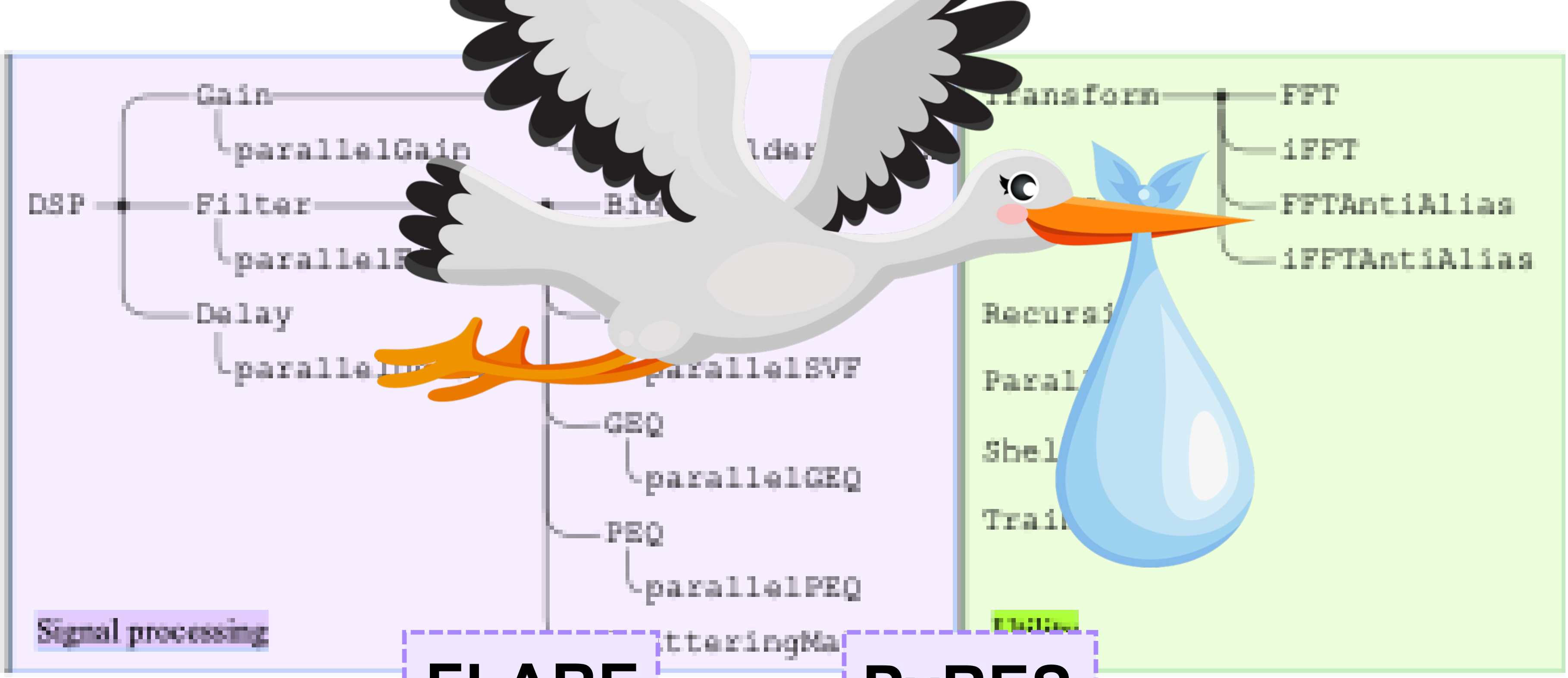


Building DDSP with Frequency Sampling

FLAMO library (ICASSP ‘25)

FLAMO

- **FLAMO**: **F**requency-sampling **L**ibrary for **A**udio-**M**odule **O**ptimization
- Open-source library for implementing and optimizing differentiable LTI audio systems
- Modules support the time-aliasing mitigation system



FLARE

PyRES

A!DAFx25

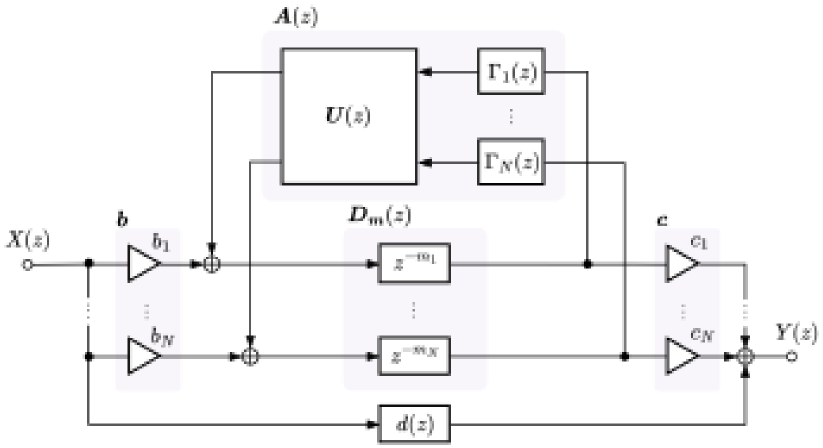
Building DDSP with Frequency Sampling

FLAMO's Children

FLARE PyRES

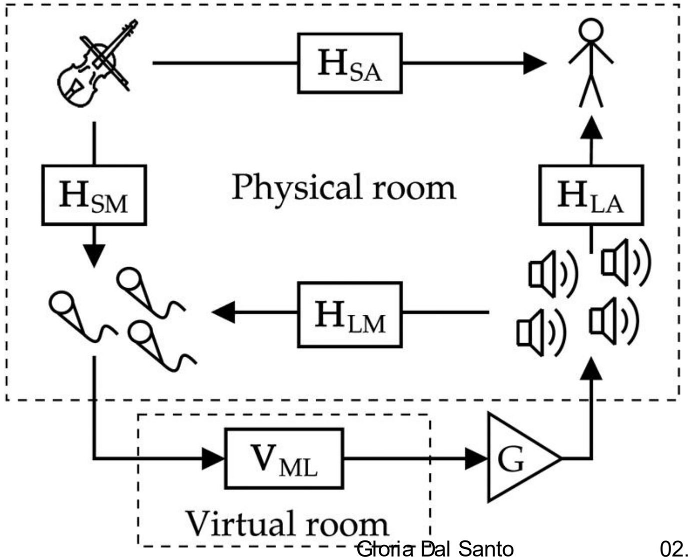
- **FLARE**
  - Library for synthesis and analysis of RIRs using Feedback Delay Networks
  - Designed as a data-augmentation tool

AES AIMLA Late Breaking Demo Paper



- **PyRES** (G. M. De Bortoli et al.)
  - Library of classes and functionality for the development, evaluation, and simulation of Reverberation Enhancement Systems (RES)
  - Utilities for the equalization and evaluation of RESs

Catch Gian Marco on Thursday at the 11 AM poster session to learn more ;)



Building DDSP with Frequency Sampling

Differentiable Biquads for PEQ (Nercessian, '20)

Parametric Equalizers (PEQ) are widely used in audio DSP, including artificial reverberation

PEQ using Biquads

cascade of 2 shelving filters and K-2 peak filters

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} \qquad H_{eq}(z) = \prod_{k=0}^{K-1} H_k(z)$$

$$\omega_0 = 2\pi \frac{f}{f_s} \qquad A = 10^{g/40}$$

Biquad coefficient formulae for different filter types.

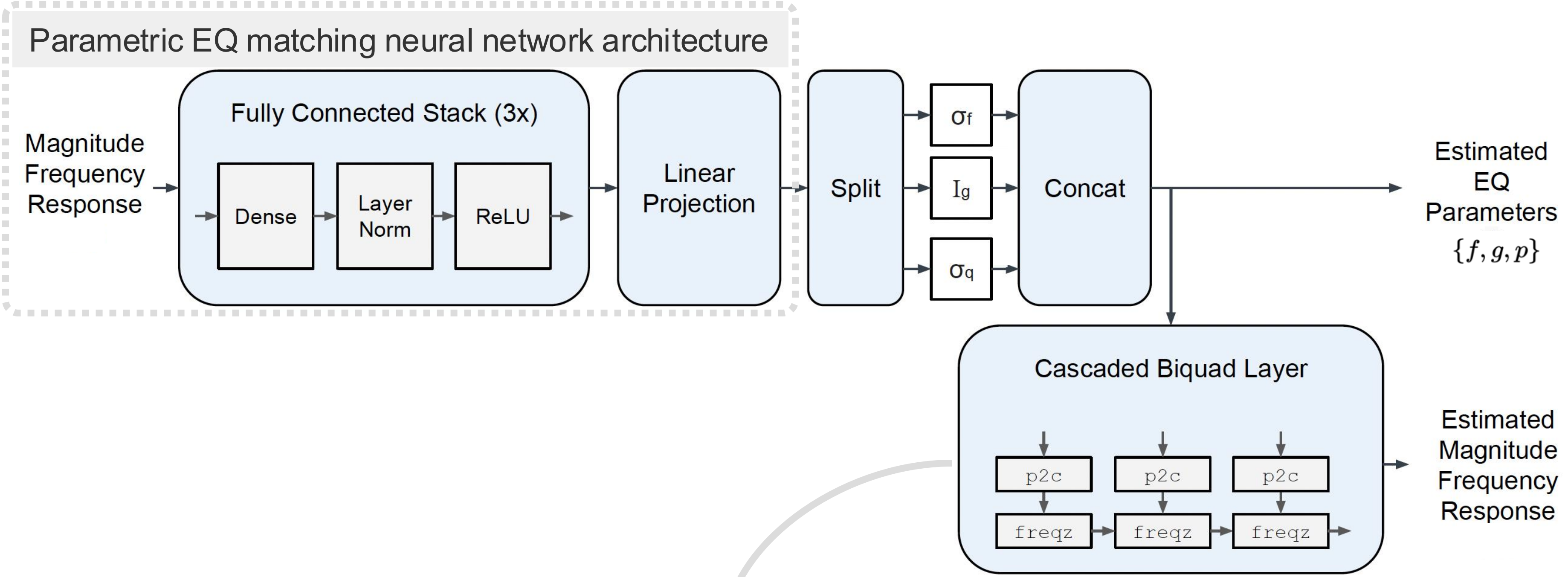
Coefficient	Low shelf	High shelf	Peak
α	$\sin(\omega_0)\sqrt{(A^2+1)(1/q-1)+2A}$	$\sin(\omega_0)\sqrt{(A^2+1)(1/q-1)+2A}$	$\frac{\sin(\omega_0)}{2q}$
b_0	$A((A+1)-(A-1)\cos(\omega_0)+\alpha)$	$A((A+1)+(A-1)\cos(\omega_0)+\alpha)$	$1+\alpha * A$
b_1	$2A((A-1)-(A+1)\cos(\omega_0))$	$-2A((A-1)+(A+1)\cos(\omega_0))$	$-2\cos(\omega_0)$
b_2	$A((A+1)-(A-1)\cos(\omega_0)-\alpha)$	$A((A+1)+(A-1)\cos(\omega_0)-\alpha)$	$1-\alpha * A$
a_0	$(A+1)+(A-1)\cos(\omega_0)+\alpha$	$(A+1)-(A-1)\cos(\omega_0)+\alpha$	$1+\alpha/A$
a_1	$-2A((A-1)+(A+1)\cos(\omega_0))$	$2A((A-1)-(A+1)\cos(\omega_0))$	$-2\cos(\omega_0)$
a_2	$(A+1)+(A-1)\cos(\omega_0)-\alpha$	$(A+1)-(A-1)\cos(\omega_0)-\alpha$	$1-\alpha/A$

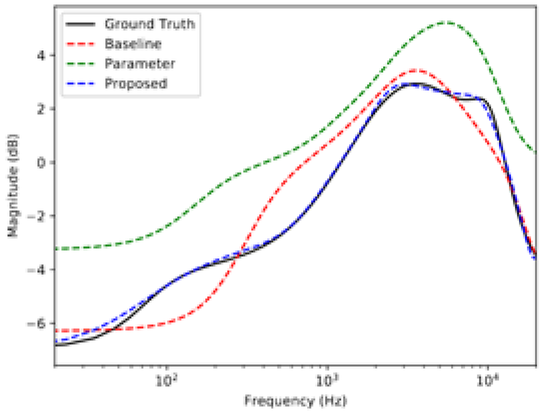
Learnable parameters for each band

$\theta = \{f, g, p\}$

Building DDSP with Frequency Sampling

Differentiable Biquads for PEQ (Nercessian, '20)





$$\{f, g, p\}_k \longrightarrow \{\mathbf{b}, \mathbf{a}\}_k \longrightarrow H_k(\mathbf{z}_M) = \frac{b_0 + b_1 \mathbf{z}_M^{-1} + b_2 \mathbf{z}_M^{-2}}{a_0 + a_1 \mathbf{z}_M^{-1} + a_2 \mathbf{z}_M^{-2}}$$

Differentiable Feedback Delay Network

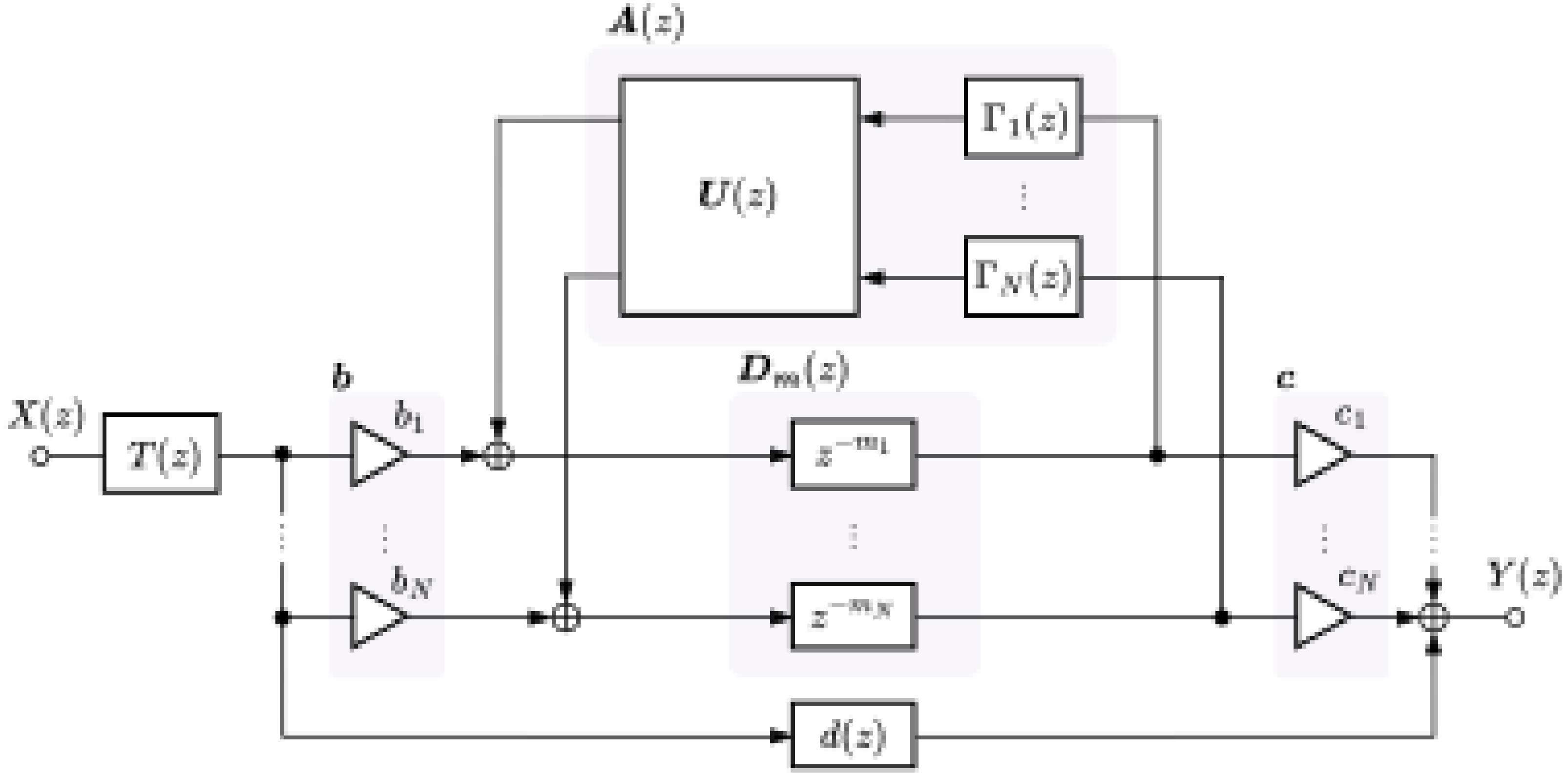
General FDN Structure
Differentiable FDN and its applications
Challenges in FDN optimization
Example: Optimization of a Feedback Delay Network

Differentiable FDN

Feedback Delay Networks for Reverb Synthesis

FDNs are recursive filters based on a generalization of the parallel comb filter structure

- cheap and highly parametrizable artificial reverberators



$\boldsymbol{A}(z)$ *mixing matrix*
$\boldsymbol{m}$ *delays*
$\boldsymbol{b}$ *input gains*
$\boldsymbol{c}$ *output gains*
$\Gamma_i(z)$ *attenuation filters*
$T(z)$ *tone corrector*
$d(z)$ *direct path*

Reference

FDN IR

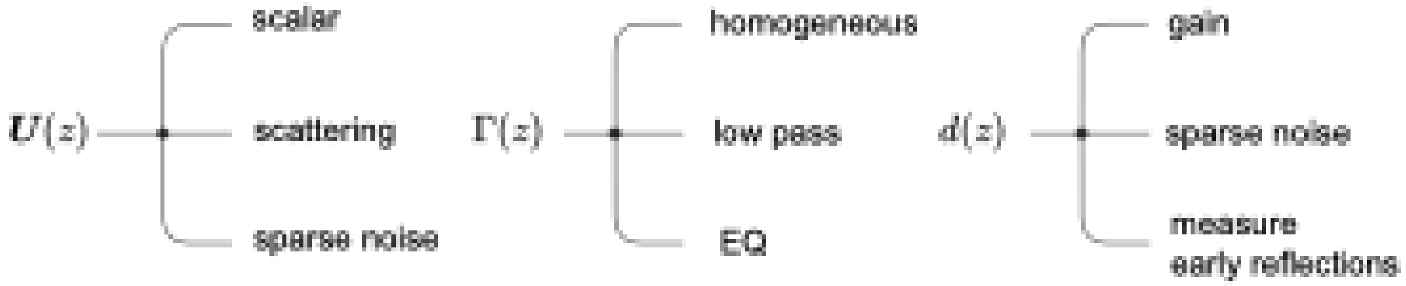
Differentiable FDN

Feedback Delay Networks for Reverb Synthesis

Parameters can be carefully designed to meet the required T_{60} accuracy and echo density

$$\Gamma_i(z) \approx \frac{-60}{f_s T_{60}(\omega)}\, m_i$$

$$T(z) \approx \varepsilon_h\,(0)$$

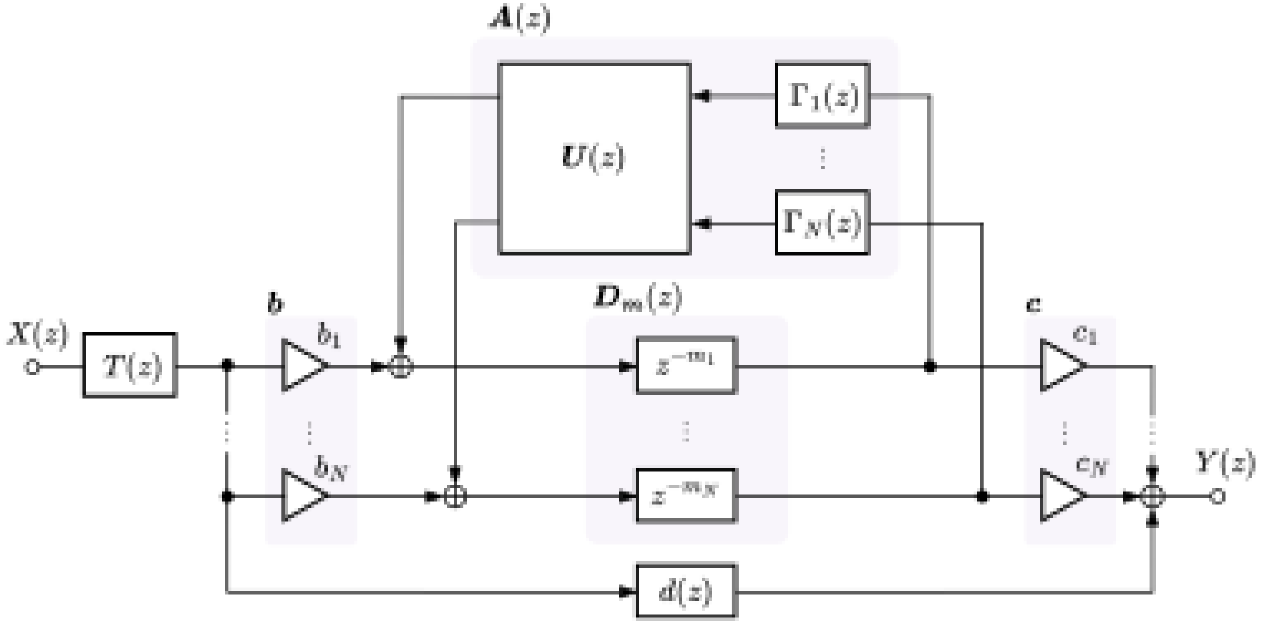


...so why would we need differentiable FDNs?

- Parameters have non-trivial relations
- Improve echo density, given computational constraints
- Estimating the filters' response given a target RIR (or Spatial RIR)
- Blind estimation from reverberant signals
- Automatic mixing
- Improve the stability of reverberation enhancement systems

Differentiable FDN

FDN Optimization



Differentiable via frequency-sampling

- FIR approximation

$$\boldsymbol{z}_M = [e^{\jmath\pi\frac{0}{M}}, e^{\jmath\pi\frac{1}{M}}, \ldots, e^{\jmath\pi\frac{M-1}{M}}]$$

$$H(\mathbf{z}_m) = T(\mathbf{z}_m)\left(\mathbf{c}^\top\left[\mathbf{D_m}(\mathbf{z}_m)^{-1} - \mathbf{A}(\mathbf{z}_m)\right]^{-1}\mathbf{b} + d\right)$$

Active field of research:

- Orchisama Das et al. – grouped FDNs for coupled spaces (arxiv pre-print)
- Alessandro Mezza, Riccardo Giampiccolo et al. – MIMO, HOM
- Ilias Ibnyahya et al. – cheaper differentiable $\Gamma(z)$
- and ofc us :D – Coloration and echo density optimization

.... and other things I keep procrastinating on putting into words -.-’

Differentiable FDN

Challenges in FDN optimization

- Common audio losses alone cannot capture common artifacts

Example: Multi-scale STFT Loss

$$\mathcal{L}_{SC}(y,\tilde{y}) = \frac{\| |\mathrm{STFT}(y)| - |\mathrm{STFT}(\tilde{y})| \|_F}{\| |\mathrm{STFT}(y)| \|_F}$$

$$\mathcal{L}_{SM}(y,\tilde{y}) = \frac{1}{N} \| \log(|\mathrm{STFT}(y)|) - \log(|\mathrm{STFT}(\tilde{y})|) \|_1$$

$$\boxed{\mathcal{L}_{\mathrm{MR}}(\hat{y},y) = \frac{1}{M}\sum_{m=1}^{M}(\mathcal{L}_{\mathrm{SC}}(\hat{y},y) + \mathcal{L}_{\mathrm{SM}}(\hat{y},y))}$$

$$\mathcal{L}(h_{\mathrm{FDN1}} h_{\mathrm{REF}}) = \mathcal{L}(h_{\mathrm{FDN2}} h_{\mathrm{REF}})$$

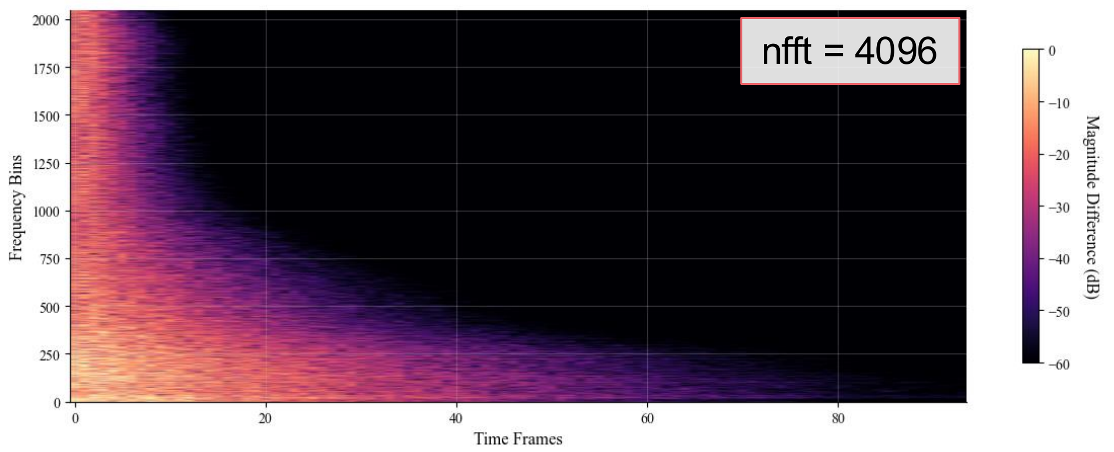
Reference

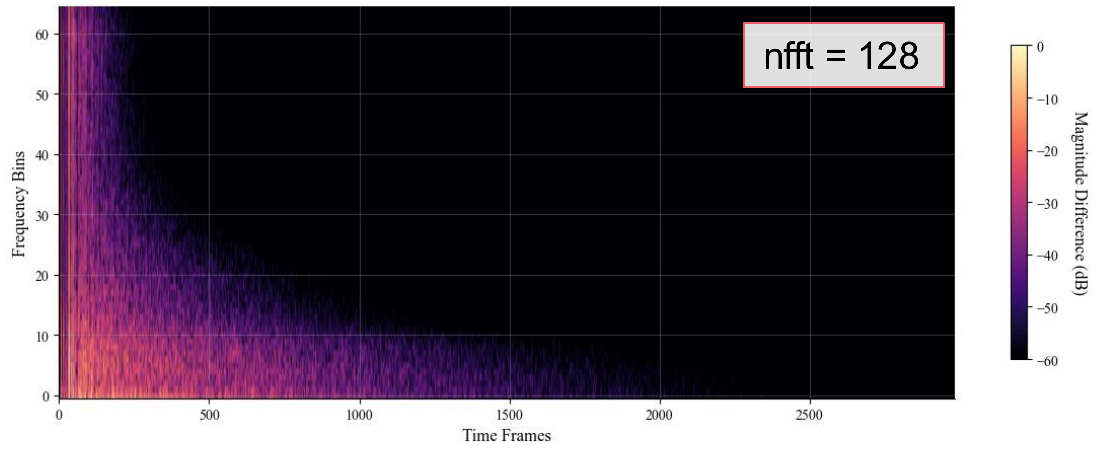
FDN 1

FDN 2

Equally likely solutions

difference between H_{FDN1} H_{FDN2}





Differentiable FDN

Challenges in FDN optimization

- Common audio losses alone are not able to capture common artifacts

- Need for **system-specific losses**
  Two examples:

Colorless optimization

Init *Optim*

$$\mathcal{L} = \mathcal{L}_{\text{spectral}}(\boldsymbol{H}(z)) + \mathcal{L}_{\text{sparsity}}(\boldsymbol{U}(z))$$

$$\mathcal{L}_{\text{spectral}}(\boldsymbol{H}(\boldsymbol{z}_\mu)) = \frac{1}{\mu} \sum_{z \in \boldsymbol{z}_\mu} (|H(z)| - 1)^2$$

$$\mathcal{L}_{\text{sparsity}}(\boldsymbol{U}) = \frac{N\sqrt{N} - \sum_{i,j}|U_{ij}|}{N(\sqrt{N}-1)}$$

RT optimization

$$\varepsilon(t; f_c) = \sum_{\tau=t}^{L} h_{f_c}^2(\tau)$$

$$\mathcal{L}_{\text{EDC}} = \frac{1}{|\mathcal{C}|} \sum_{f_c \in \mathcal{C}} \frac{\sum_{t=0}^{L} (\varepsilon_{\text{dB}}(t; f_c) - \hat{\varepsilon}_{\text{dB}}(t; f_c))^2}{\sum_{t=0}^{L} \varepsilon_{\text{dB}}^2(t; f_c)}$$

Choosing the Right Loss

Analyzing the loss with real data
Loss landscape analysis

Choosing the Right Loss

Optimizing Coloration and Attenuation

Colorlessness optimization – independent on target RIR

Attenuation optimization – <u>dependent on target RIR</u>

?

Can we design a loss function that

- Captures only the energy decay of the RIRs
- It’s smooth despite the noise-like behavior of late reverb
- Does not interfere with the colorlessness optimization

Dal Santo, G. et al. “Similarity metrics for late reverberation.” In *2024 58th Asilomar* IEEE *Conference on Signals, Systems, and Computers*

Choosing the Right Loss

Optimizing Recursive Attenuation Filters

Candidate Losses

Multi-Scale Spectral Loss

$$\mathcal{L}_{\mathrm{MSS}}(h,\hat{h}) = \frac{1}{M}\sum_{m=1}^{M}\left(\mathcal{L}_{\mathrm{SC}}(h,\hat{h}) + \mathcal{L}_{\mathrm{SM}}(h,\hat{h})\right)$$

Power Convergence Loss

$$\mathcal{L}_{\mathrm{PC}} = \frac{\left\| |H(t,f)|^2 * W - |\hat{H}(t,f)|^2 * W \right\|_{\mathrm{F}}}{\left\|(|H(t,f)|^2 * W)\right\|_{\mathrm{F}} \left\|(|\hat{H}(t,f)|^2 * W)\right\|_{\mathrm{F}}}$$

- time-frequency averaging using Hann window *W*

Energy Decay Curve Loss

$$\mathcal{L}_{\mathrm{EDC}} = \frac{1}{|\mathcal{C}|}\sum_{f_c\in\mathcal{C}} \frac{\sum_{t=0}^{L}(\varepsilon_{\mathrm{dB}}(t;f_c) - \hat{\varepsilon}_{\mathrm{dB}}(t;f_c))^2}{\sum_{t=0}^{L}\varepsilon_{\mathrm{dB}}^2(t;f_c)}$$

- one-third-octave bands 20-12.5kHz
- EDCs are normalized to 0dB prior to computing the loss

Low Anchor Loss: Error to Signal Ratio (time domain)

$$\mathcal{L}_{\mathrm{ESR}}(h,\hat{h}) = \frac{\sum_{t_{\mathrm{mix}}}^{L}|h(t) - \hat{h}(t)|^2}{\sum_{t_{\mathrm{mix}}}^{L}|h(t)|^2}$$

Choosing the Right Loss

Optimizing Recursive Attenuation Filters – Analyzing the loss with real data

Variable Acoustics Dataset (K. Prawda et al. 2022)

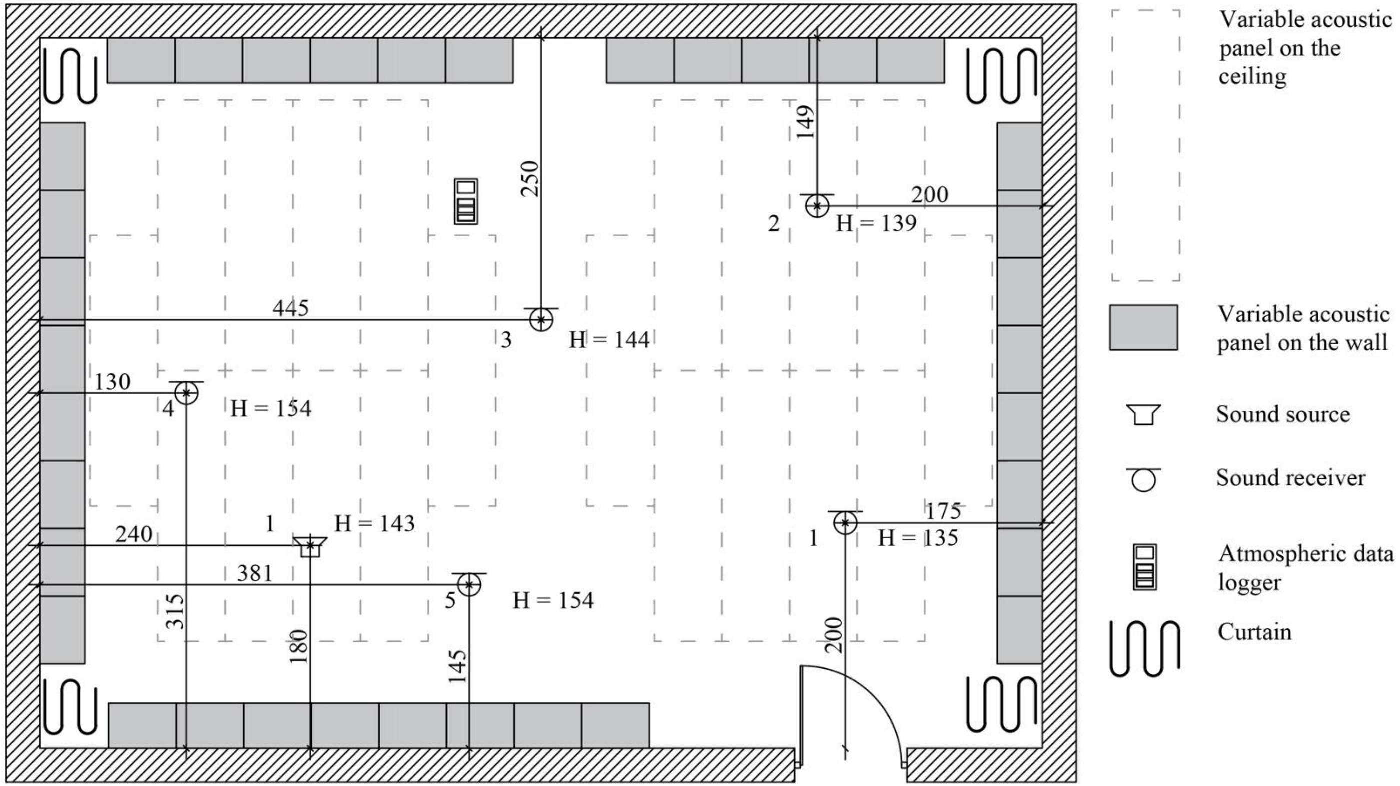
5342 recorded panel configurations

- 55 variable acoustics panels
- 5 microphone positions

Subset sampling

11 partitions {0-4, 5-9, ..., 45-49, 50-55}

- 25 RIRs per partition (5 RIRs per mic position)



Floor plan of the variable acoustics room at Aalto Acoustics Lab

Choosing the Right Loss

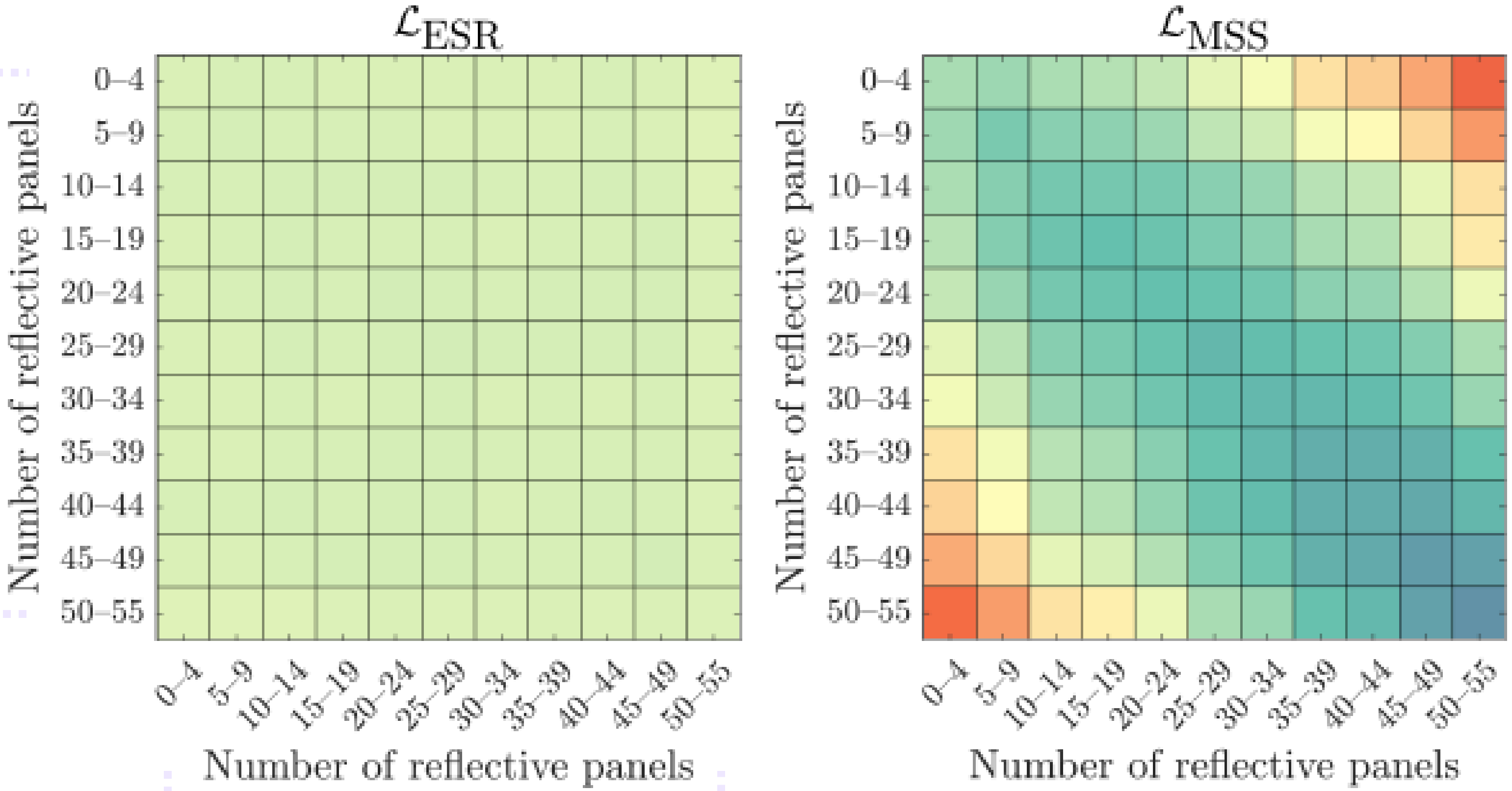
Optimizing Recursive Attenuation Filters – Analyzing the loss with real data

Reverberation condition differences

Median of loss values after normalization to $\mu = 0$ and $\sigma = 1$

reverberant

dry



reverberant

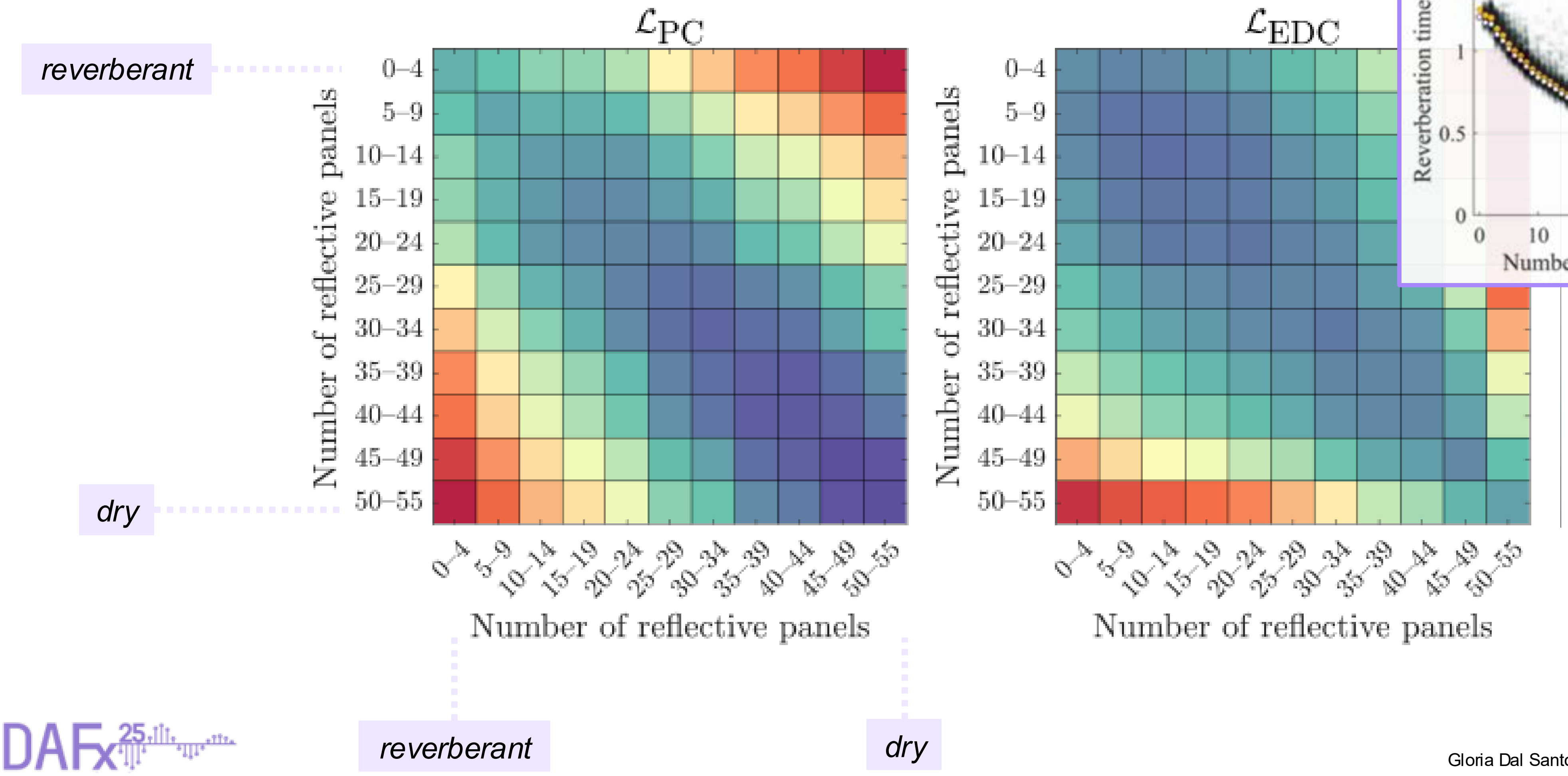
dry

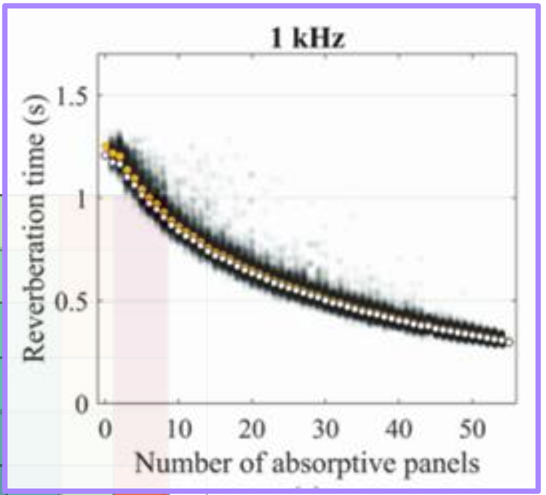
Choosing the Right Loss

Optimizing Recursive Attenuation Filters – Analyzing the loss with real data

Reverberation condition differences

Median of loss values after normalization to $\mu = 0$ and $\sigma = 1$



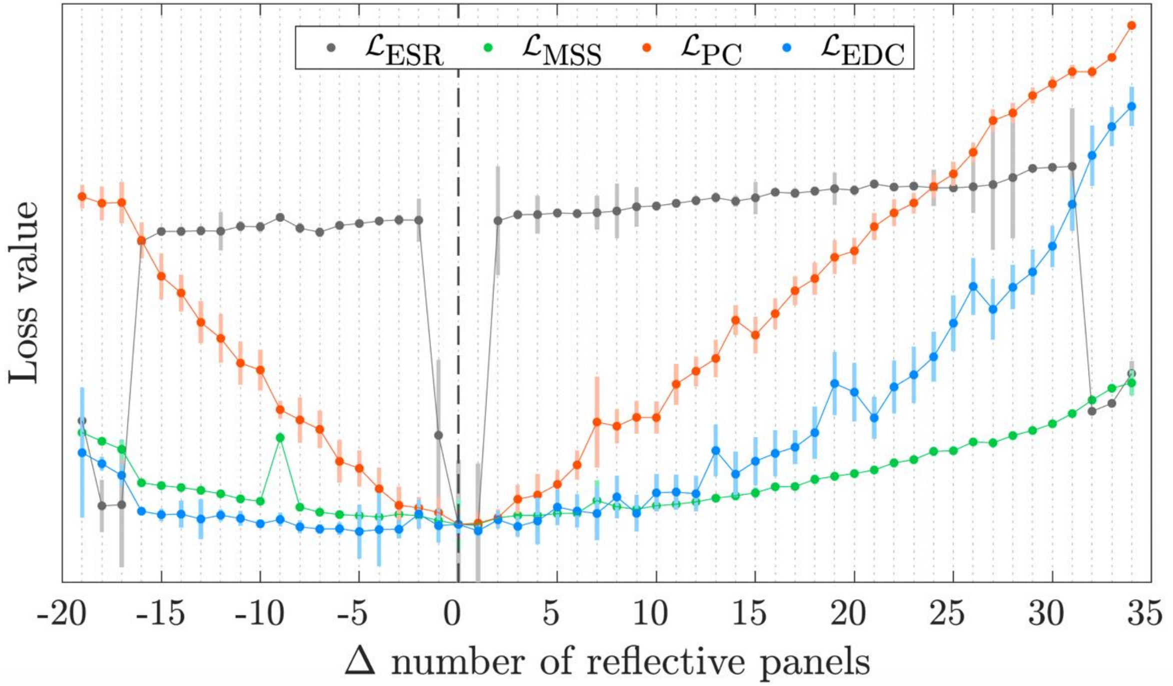


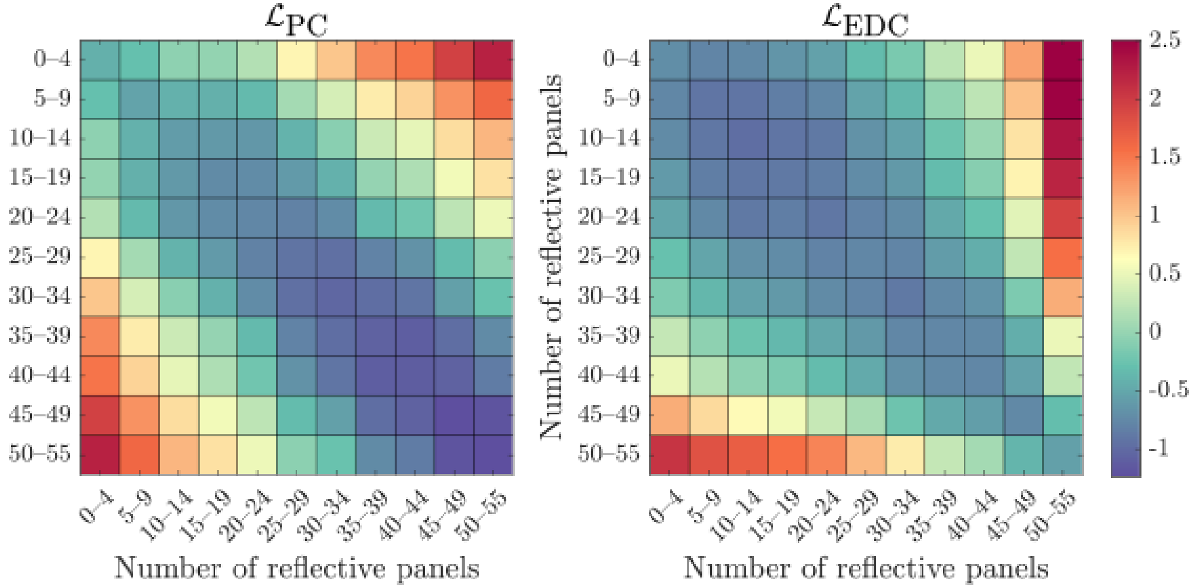
Choosing the Right Loss

Optimizing Recursive Attenuation Filters – Analyzing the loss with real data

Reverberation condition differences

Median of loss values after normalization to $\mu = 0$ and $\sigma = 1$





Choosing the Right Loss

Optimizing Recursive Attenuation Filters – Loss Landscape Analysis

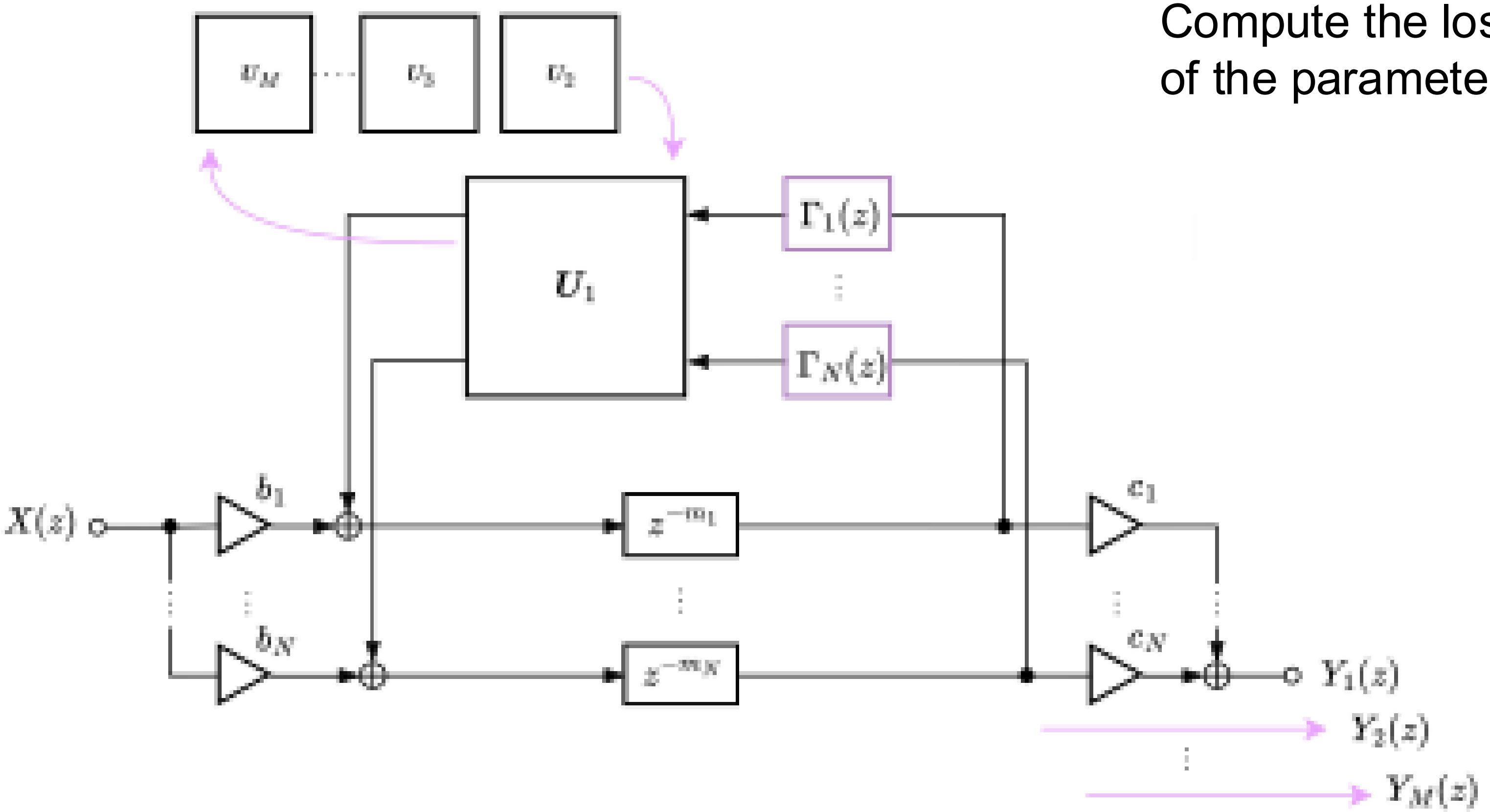
How do certain parameters interfere with the optimization of other parameters?

Example:

- Attenuation optimization
- Mixing matrix variations

How does the attenuation loss get influenced by variations of $\boldsymbol{U}$?

Compute the loss at numerous instances of the parameters under analysis

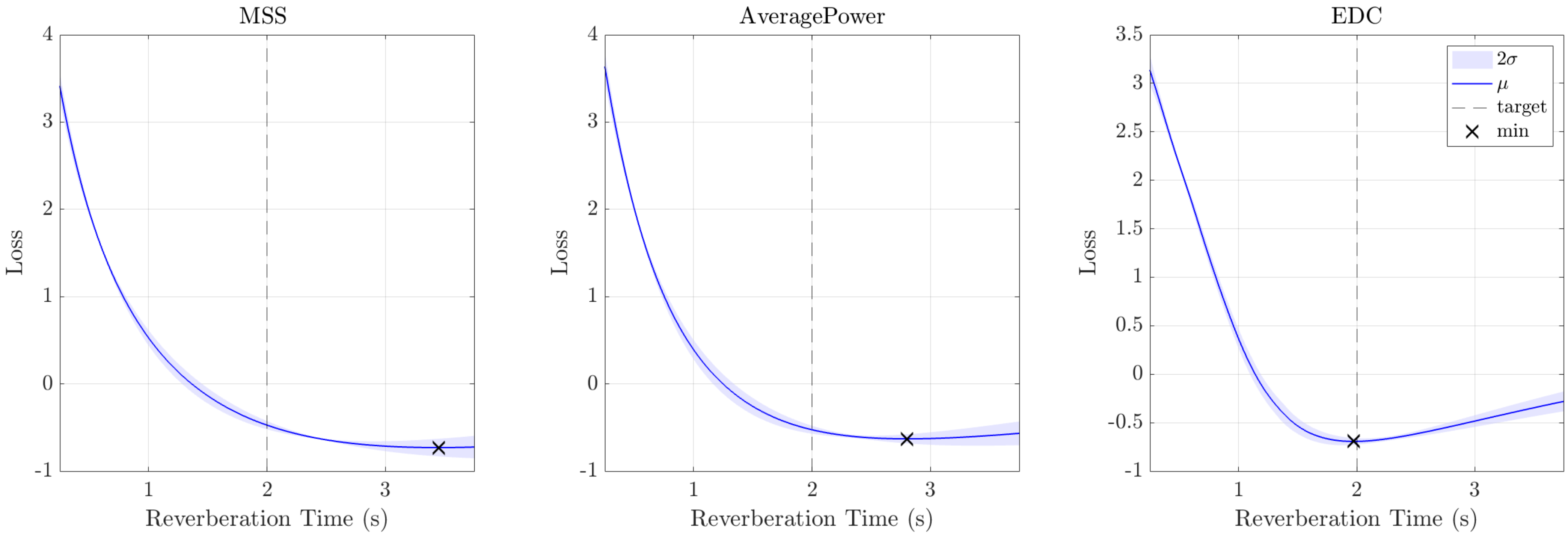


Choosing the Right Loss

Effect of feedback matrix

Target: shaped Gaussian noise
Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 200 steps
- 25 perturbations

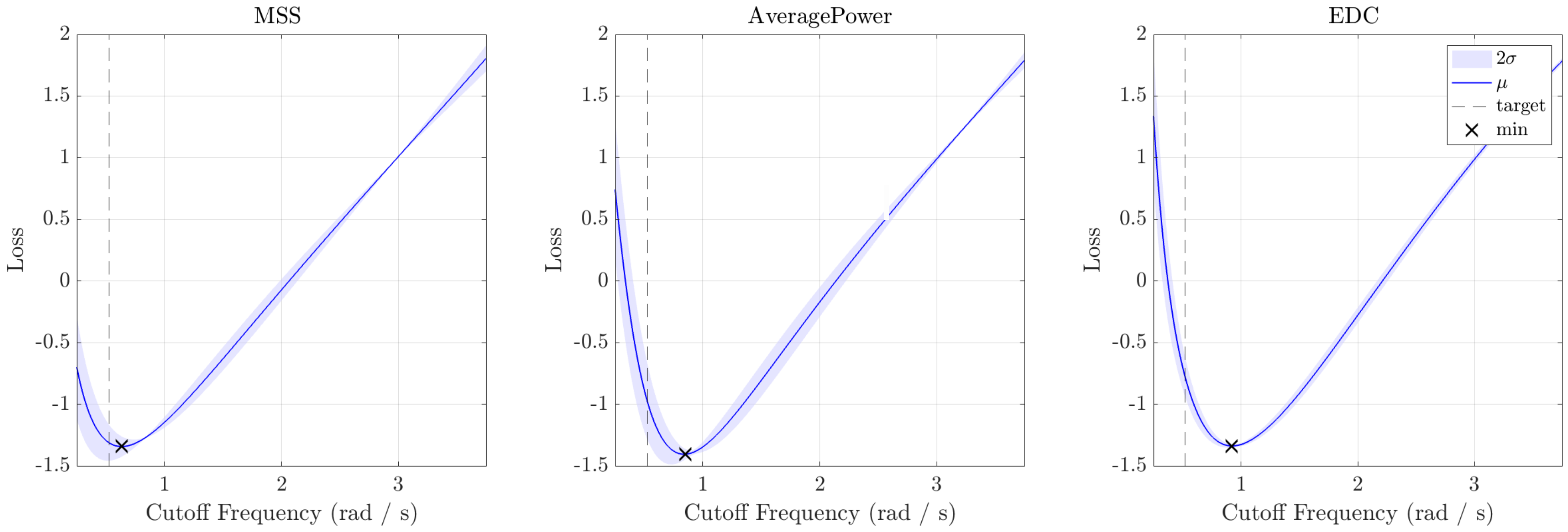


Choosing the Right Loss

Effect of feedback matrix

Target: shaped Gaussian noise
Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 200 steps
- 25 perturbations



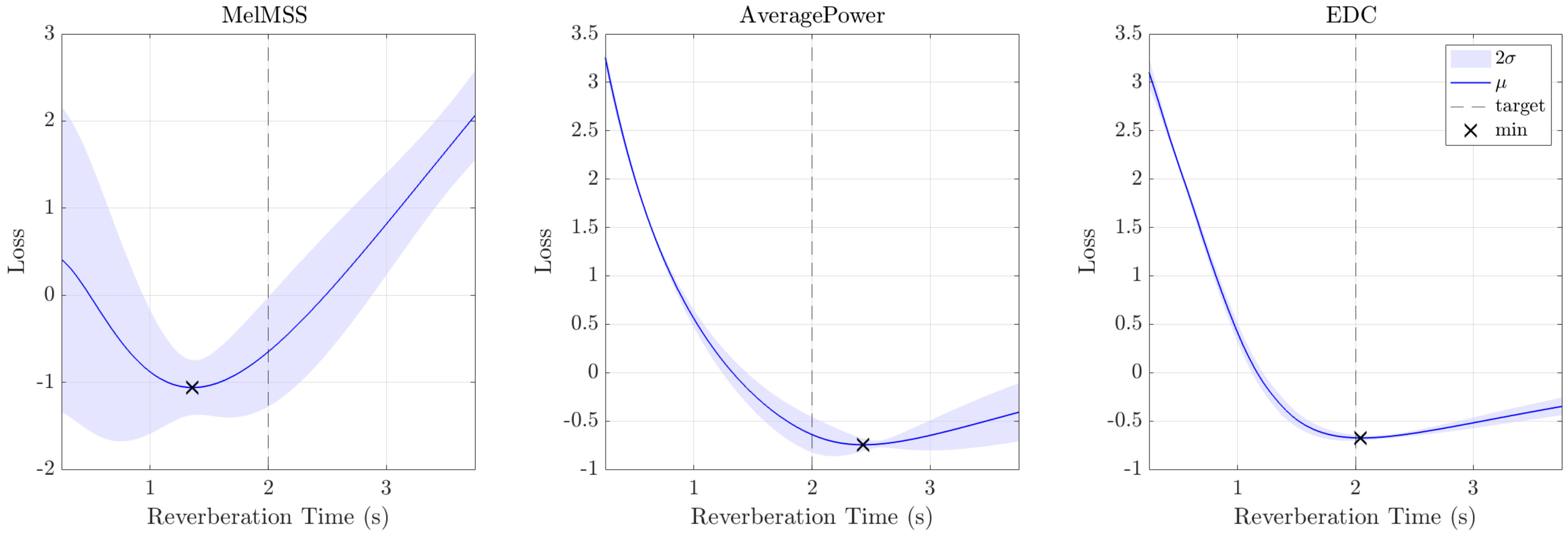
Choosing the Right Loss

Effect of feedback matrix

Target: FDN response

Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 200 steps
- 25 perturbations



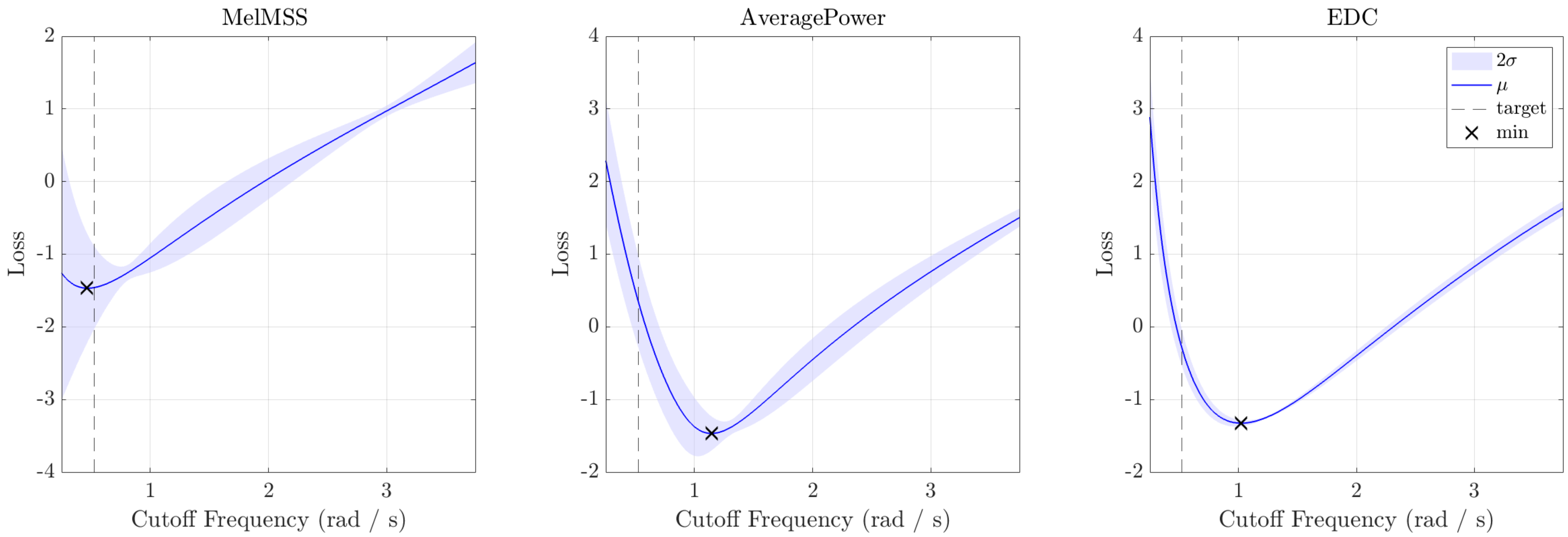
Choosing the Right Loss

Effect of feedback matrix

Target: FDN response

Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 200 steps
- 25 perturbations



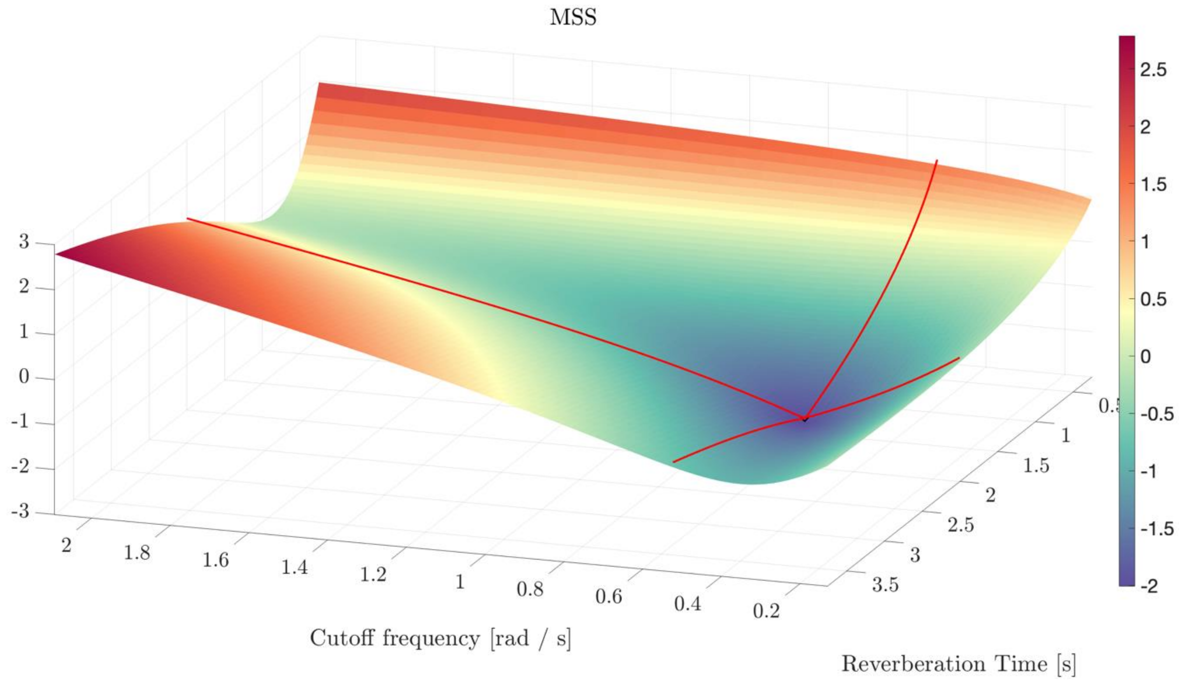
Choosing the Right Loss

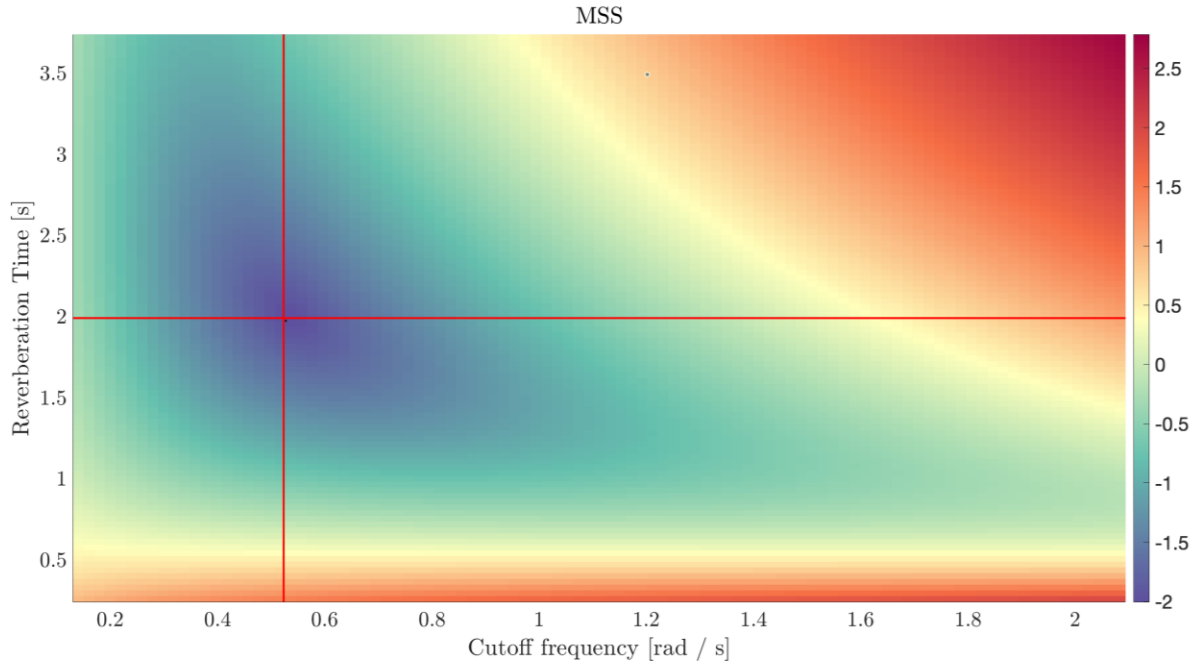
Loss Landscape

Target: shaped Gaussian noise

Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 100 steps



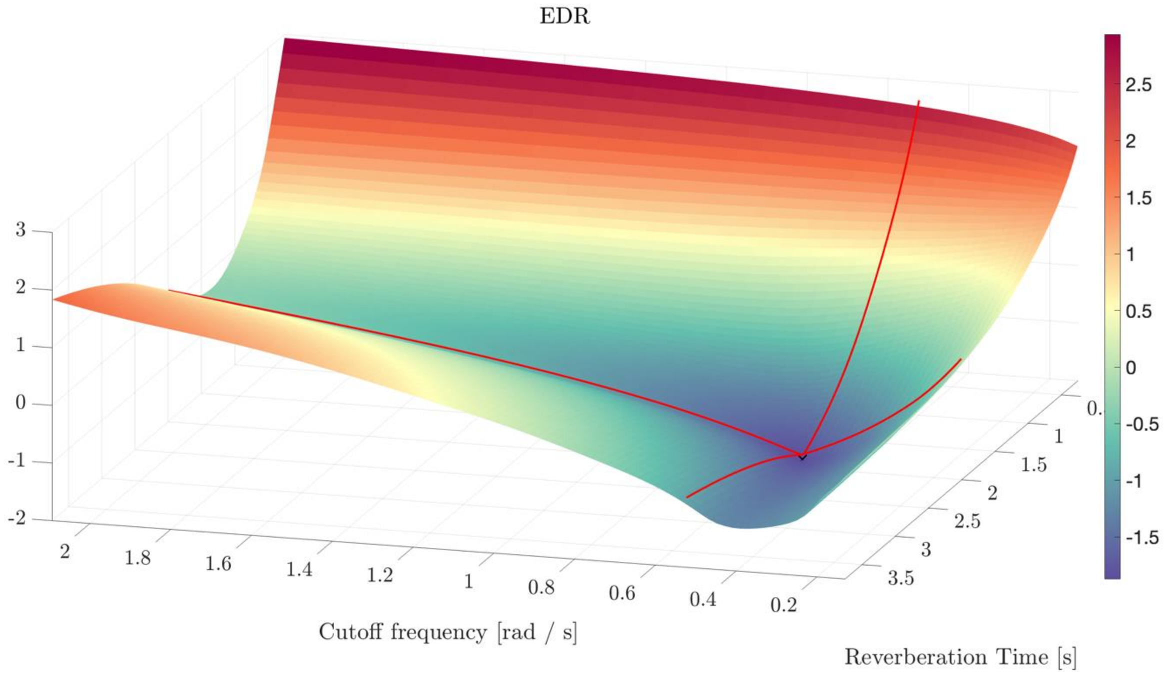


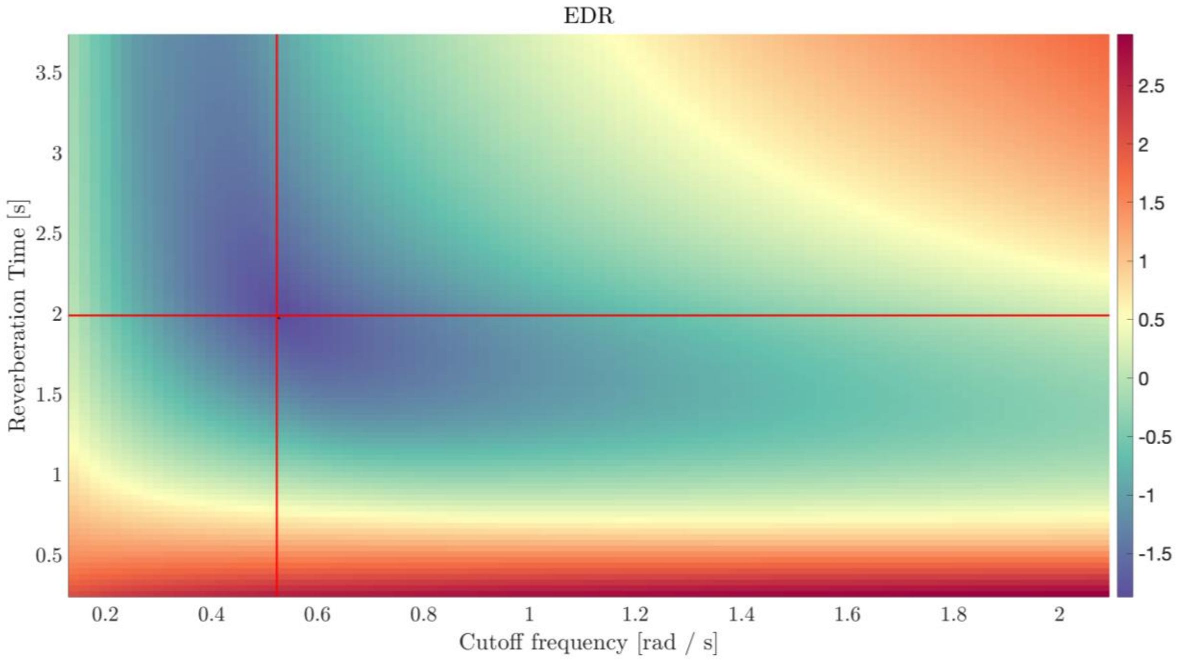
Choosing the Right Loss

Loss Landscape

Target: shaped Gaussian noise
Filter: First-order low-pass with parameters RT at DC, and cutoff frequency

- 100 steps





Choosing the Right Loss

Key Takeaways

When optimizing DDSP models, the choice of loss function is crucial.
Ask yourself:

- Does the loss function actually capture the parameter changes I want to optimize?
- How do different losses interact?
  – Does optimizing one objective interfere with another?
- What does the loss landscape look like?
  – Is it smooth enough to allow stable optimization, or too rough/noisy?

The effectiveness of DDSP pipelines depends not only on the model architecture, but equally on well-chosen, task-specific loss functions.

Choosing the Right Loss

Key Takeaways

When optimizing DDSP models, the choice of loss function is crucial.
Ask yourself:

- Does the loss function actually capture the parameter changes I want to optimize?
- How do different losses interact?
  – Does optimizing one objective interfere with another?
- What does the loss landscape look like?
  – Is it smooth enough to allow stable optimization, or too rough/noisy?

The effectiveness of DDSP pipelines depends not only on the model architecture, but equally on well-chosen, task-specific loss functions.

Summary

- **Differentiable Digital Signal Processing**
  - Brings interpretability and domain knowledge into audio deep learning models
- **Building DDSPs with Frequency Sampling**
  - Enables efficient differentiable filter design, but requires aliasing mitigation
  - FLAMO: Open-source library for DDSP modules (with FLARE, PyRES extensions)
- **Differentiable Feedback Delay Network**
  - Powerful for reverb synthesis, but optimization remains challenging
- **Choosing the right loss**
  - Must be adequate for the underlying model and learned parameters

Bibliography

- Engel, J., Hantrakul, L., Gu, C., & Roberts, A. (2020). DDSP: Differentiable digital signal processing. *arXiv preprint arXiv:2001.04643*.
- Müller, M., Liem, C., McFee, B., & Schwär, S. (2025). Learning with Music Signals: Technology Meets Education (Dagstuhl Seminar 24302). *Dagstuhl Reports*, *14*(7), 115-152.
- Hayes, B., Shier, J., Fazekas, G., McPherson, A., & Saitis, C. (2024). A review of differentiable digital signal processing for music and speech synthesis. *Frontiers in Signal Processing*, *3*, 1284100.
- Vahidi, C., Han, H., Wang, C., Lagrange, M., Fazekas, G., & Lostanlen, V. (2023). Mesostructures: Beyond spectrogram loss in differentiable time-frequency analysis. *arXiv preprint arXiv:2301.10183*.
- Nercessian, S. (2020, September). Neural parametric equalizer matching using differentiable biquads. In *Proc. Int. Conf. Digital Audio Effects (eDAFx-20)* (pp. 265-272).
- Yu, C. Y., Mitcheltree, C., Carson, A., Bilbao, S., Reiss, J. D., & Fazekas, G. (2024). Differentiable all-pole filters for time-varying audio systems. *arXiv preprint arXiv:2404.07970*.
- Lee, S., Choi, H. S., & Lee, K. (2022). Differentiable artificial reverberation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *30*, 2541-2556.
- Dal Santo, G., De Bortoli, G. M., Prawda, K., Schlecht, S. J., & Välimäki, V. (2025, April). FLAMO: An Open-Source Library for Frequency-Domain Differentiable Audio Processing. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.

- Nercessian, S. (2020, September). Neural parametric equalizer matching using differentiable biquads. In *Proc. Int. Conf. Digital Audio Effects (eDAFx-20)* (pp. 265-272).

Bibliography

- Nercessian, S. (2020, September). Neural parametric equalizer matching using differentiable biquads. In *Proc. Int. Conf. Digital Audio Effects (eDAFx-20)* (pp. 265-272).
- J.-M. Jot and A. Chaigne, “Digital delay networks for designing artificial reverberators,” in Proc. 90th Conv. Audio Eng. Soc., Paris, France, Jan. 1991, preprint no. 3030.
- Dal Santo, G., Prawda, K., Schlecht, S. J., & Välimäki, V. (2025). Optimizing tiny colorless feedback delay networks. *EURASIP Journal on Audio, Speech, and Music Processing*, *2025*(1), 13.
- Välimäki, V., Prawda, K., & Schlecht, S. J. (2024). Two-stage attenuation filter for artificial reverberation. *IEEE Signal Processing Letters*, *31*, 391-395.
- K. Prawda, S. J. Schlecht, and V. Välimäki, “Calibrating the Sabine and Eyring formulas,” J. Acoustical Society of America, 2022.
- Dal Santo, G., Prawda, K., Schlecht, S. J., & Välimäki, V. (2024, October). Similarity metrics for late reverberation. In *2024 58th Asilomar Conference on Signals, Systems, and Computers* (pp. 1409-1413). IEEE.

Q&A Time!

Building Robust Audio DDSP Pipelines

A Case Study on Artificial Reverb