



Subtractive Modulative Network (SMN) with Learnable Periodic Activations

Tiou Wang*, Zhuoqian Yang[†], Markus Flierl*, Mathieu Salzmann[†], Sabine Süsstrunk[†]
* KTH † EPFL

TL;DR A neural field arch inspired by subtractive synthesis in music synthesizers.

Presenter: Zhuoqian Yang
PhD Candidate, EPFL
May 6, 2026

Subtractive Synthesis: building a brass sound

Start loud and bright, then carve away. Watch every step on the live oscilloscope.

1 Oscillator sawtooth

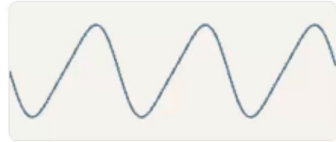
A basic waveform that is harmonic-rich. The raw material we will sculpt.



Play Stop

2 Filter low pass

Removes high frequencies.

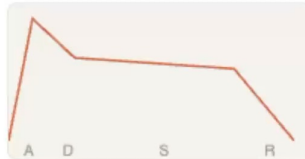


Play Stop

cutoff 1500 Hz

3 Envelope ADSR

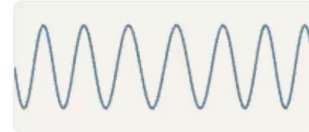
Shapes volume (amplitude) over time.



Play Stop

4 Modulation LFO

A sine wave that oscillates the pitch. Simulates the natural breathing of a player.



Play Stop

5 Unison detune

Two sawtooths, one tuned 0.5% sharp. They beat together for a thick, chorused body.



Play Stop

OSCILLOSCOPE

idle



press a Play button to see the wave

SPECTRUM

log freq · dB

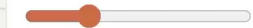


TIME SCALE



6x

VERTICAL



1.4x

TRIGGER

Auto Freeze

Brass voice

step 5 played at any pitch · keys A S D F G H J K

A

S

D

F

G

H

J

K

C

D

E

F

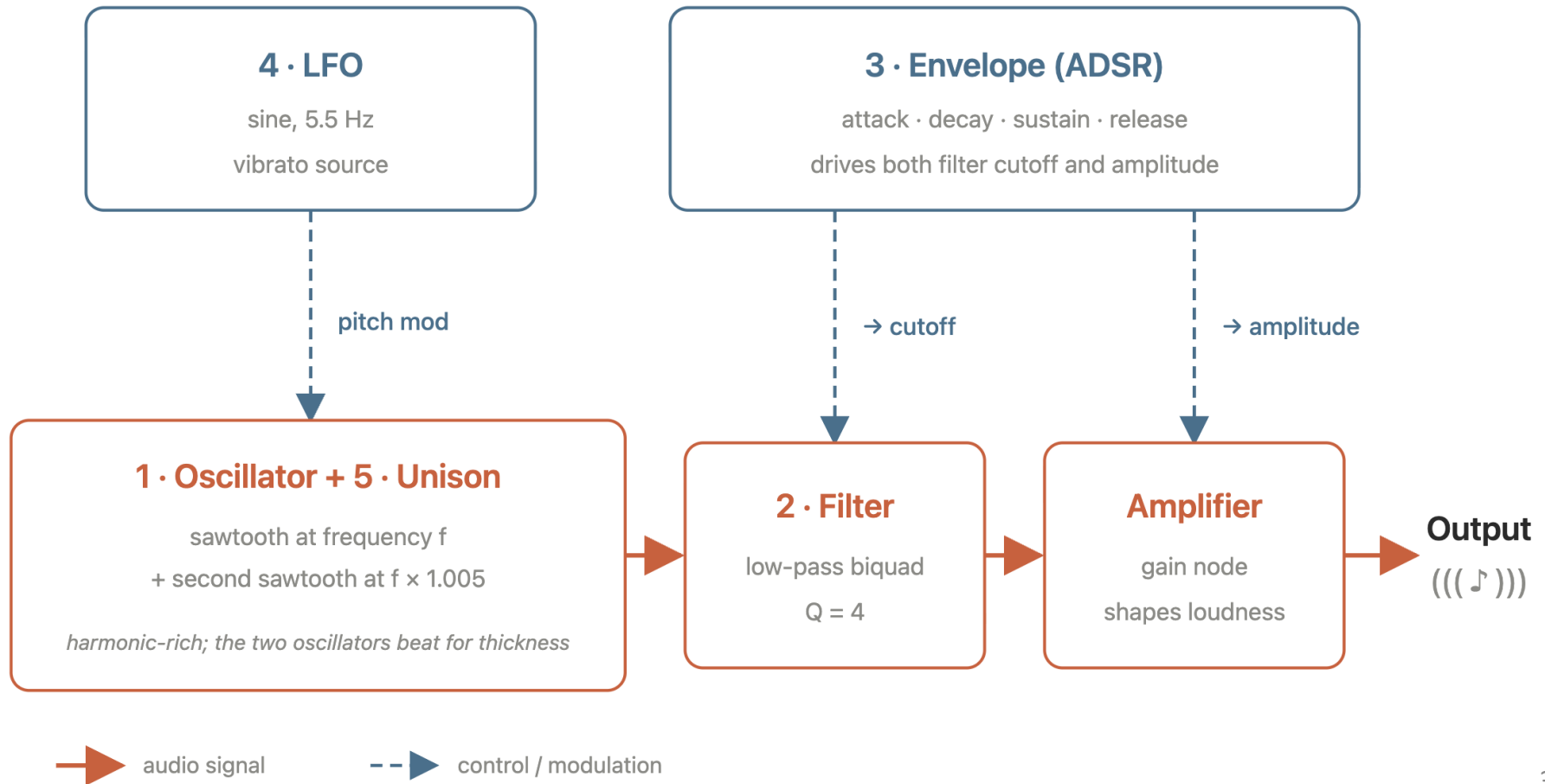
G

A

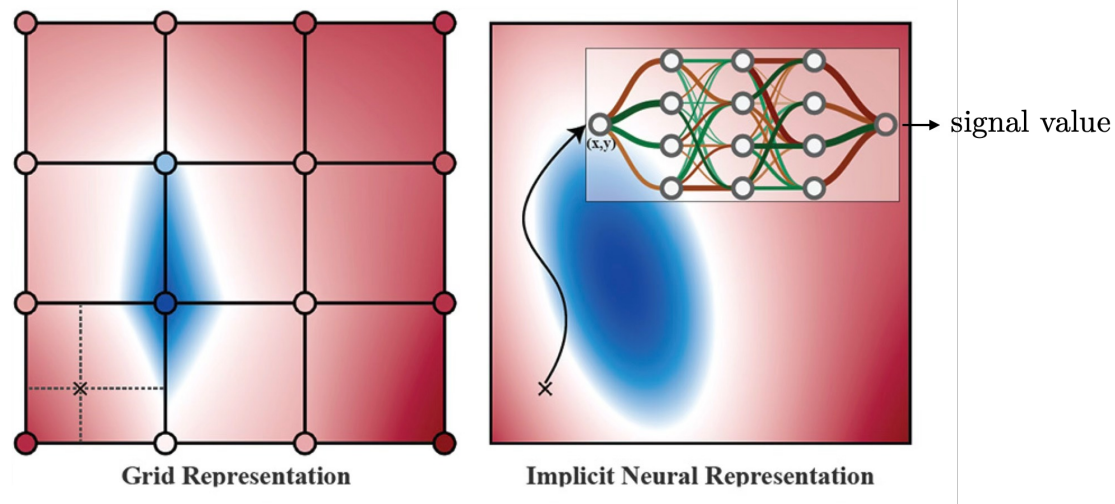
B

C

Background – Synthesizers (musical instrument)



Background - Neural Fields



- A neural field represent a signal as a **continuous function** $\mathbf{y} = f_{\theta}(\mathbf{x}) \in \mathbb{R}^m$
- \mathbf{x} : input coordinate.
- \mathbf{y} : signal value at location \mathbf{x} .

Background - Neural Fields

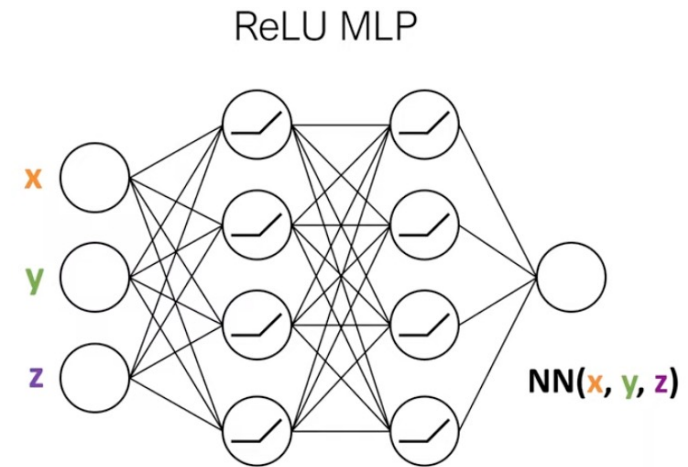
Images



Shapes

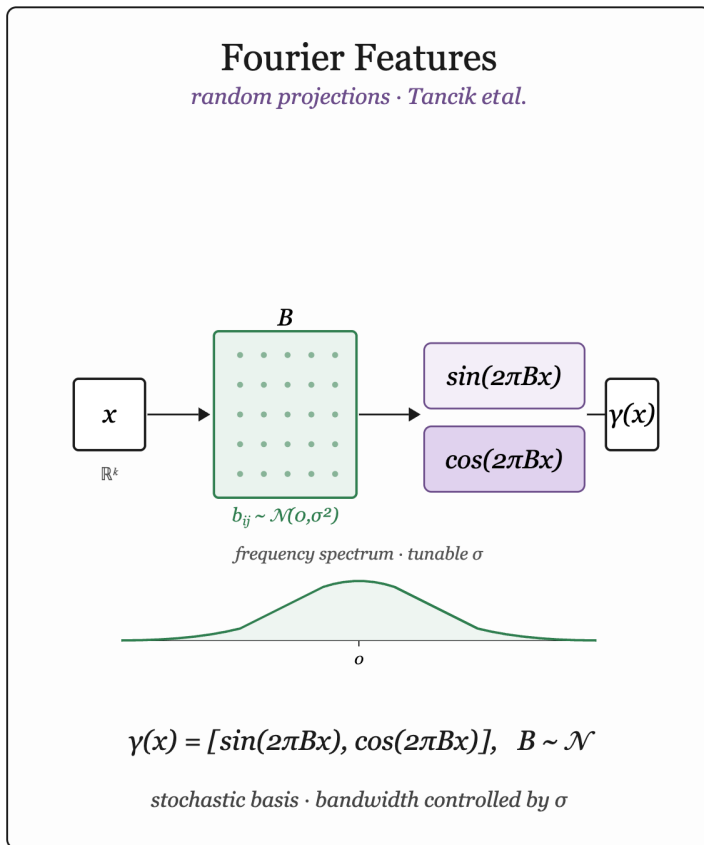


- ReLU MLPs fail to reconstruct high-frequency details

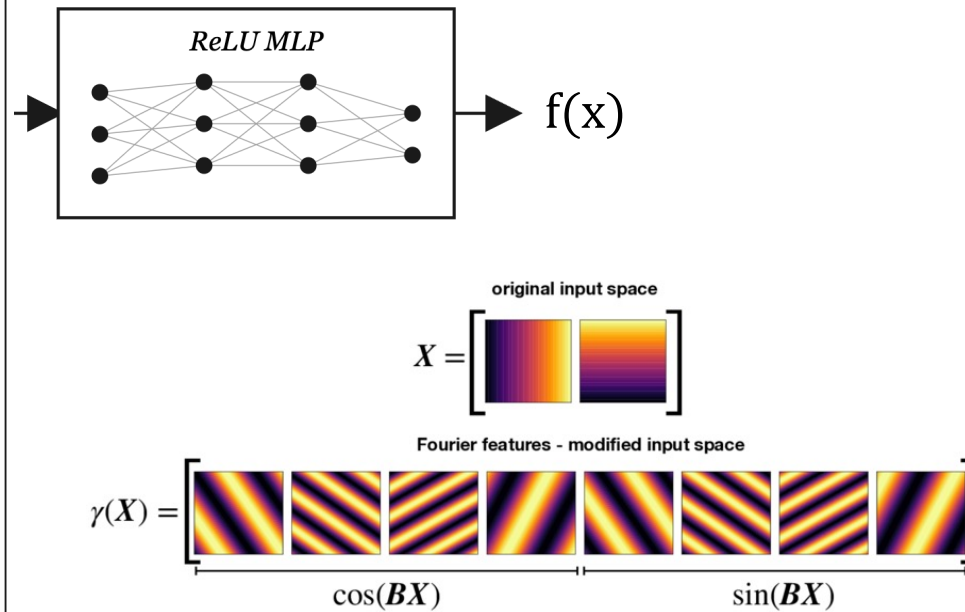


Sitzmann, Vincent, et al. "Implicit neural representations with periodic activation functions." Advances in neural information processing systems 33 (2020): 7462-7473.

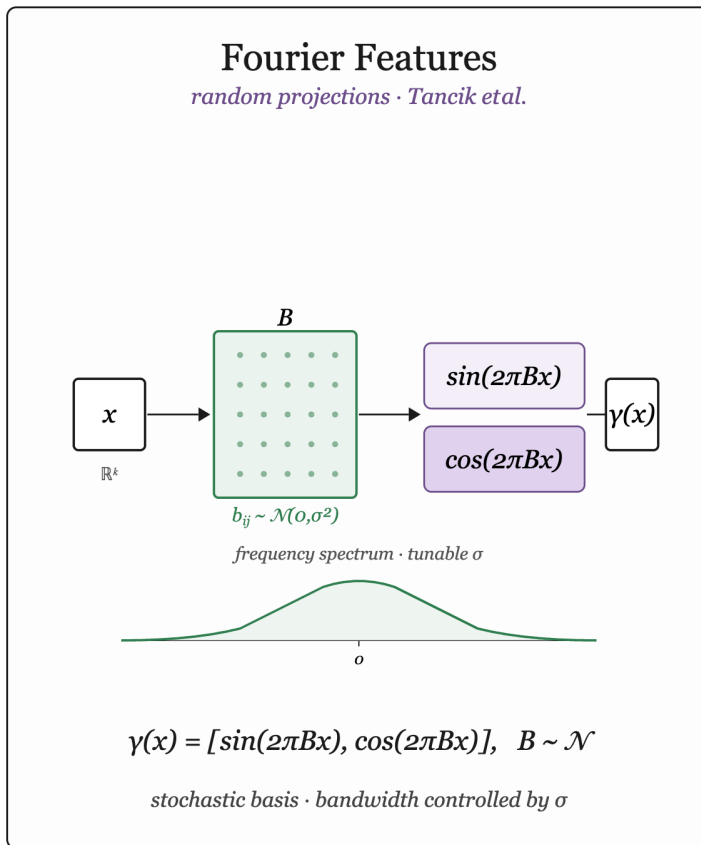
Background - Neural Fields



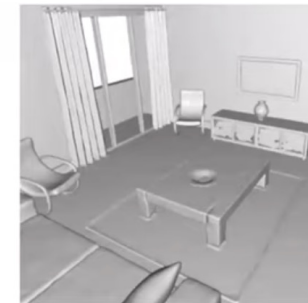
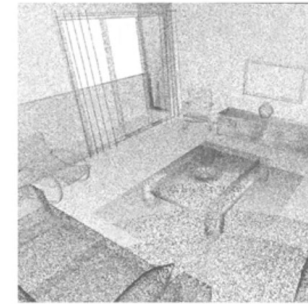
- Solution
 - Fourier Feature Network (FFN)



Background - Neural Fields



- Solution
 - Fourier Feature Network (FFN)



Mildenhall, Ben, et al. "Nerf: Representing scenes as neural radiance fields for view synthesis." Communications of the ACM 65.1 (2021): 99-106.

Sitzmann, Vincent, et al. "Implicit neural representations with periodic activation functions." Advances in neural information processing systems 33 (2020): 7462-7473.

Tancik, Matthew, et al. "Fourier features let networks learn high frequency functions in low dimensional domains." Advances in neural information processing systems 33 (2020): 7537-7547.

Motivation

- Neural fields (oversimplified):
Feed-forward MLPs with special first layers γ

Motivation

- Neural fields (oversimplified):
Feed-forward MLPs with special first layers γ
- The **expressive power** of neural fields is **restricted to** functions that can be expressed as a linear combination of certain **harmonics** of the feature mapping γ [Yüce et al.]

$$f_{\theta}(\mathbf{r}) = \sum_{\omega' \in \mathcal{H}(\Omega)} c_{\omega'} \sin(\langle \omega', \mathbf{r} \rangle + \phi_{\omega'}),$$

where

$$\mathcal{H}(\Omega) \subseteq \left\{ \sum_{t=0}^{T-1} s_t \Omega_t \mid s_t \in \mathbb{Z} \wedge \sum_{t=0}^{T-1} |s_t| \leq K^{L-1} \right\}.$$

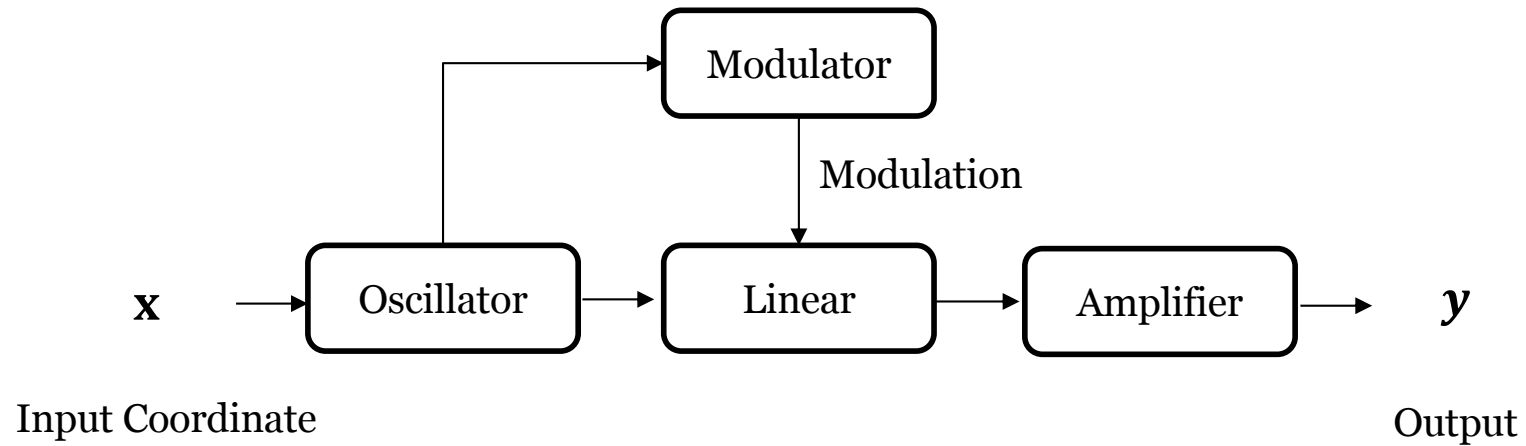
Motivation

- Neural fields (oversimplified):
Feed-forward MLPs with special first layers γ
- γ defines a set of **base frequencies** – like the *Oscillator*
- The rest of the layers expand spectral support into a collection of higher order **harmonics**

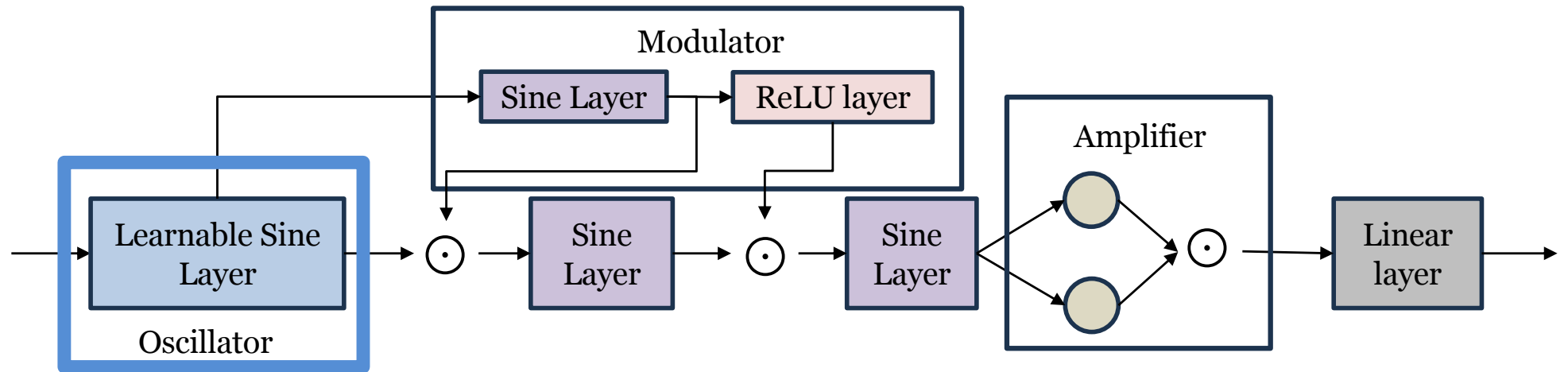
Motivation

- Neural fields (oversimplified):
Feed-forward MLPs with special first layers γ
- γ defines a set of **base frequencies** – like the *Oscillator*
- The rest of the layers expand spectral support into a collection of higher order **harmonics**
- **How to strengthen the expressive power?**
- **What inspiration can we draw from synthesizers / traditional signal processing pipelines?**

Synthesizer-Inspired Framework



Learnable Sine Layer (Oscillator)

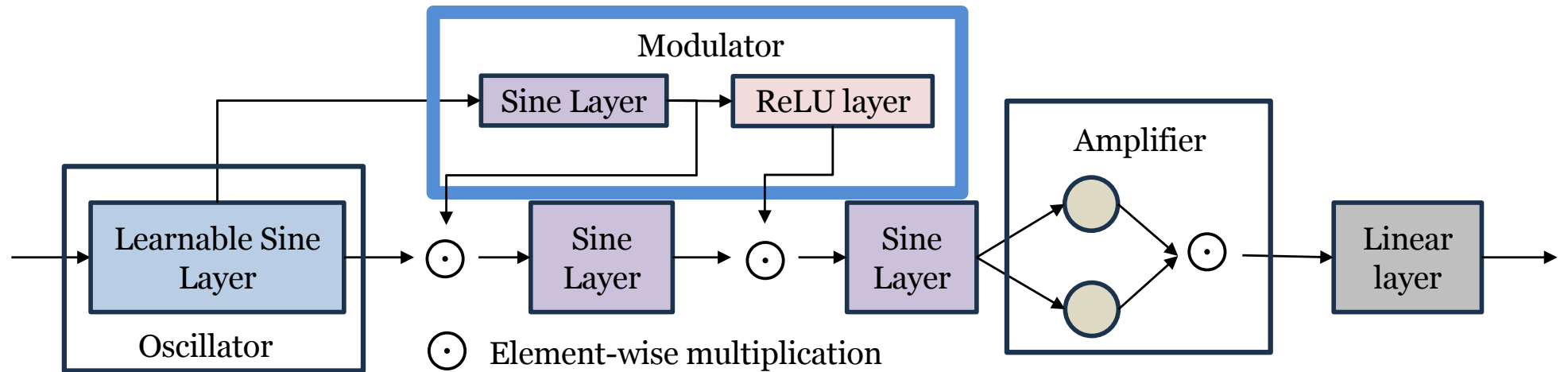


- A Linear layer with a **learnable sinusoidal activation** function

$$\gamma(\mathbf{x}) = \sum_{i=0}^k a_i \sin(\omega_i \mathbf{W} \mathbf{x}), \mathbf{W} \sim \mathcal{N}(0, \mathbf{I})$$

- \mathbf{W} – linear layer weights
- a_i, ω_i - learnable parameters
- **Provide diverse base frequencies**

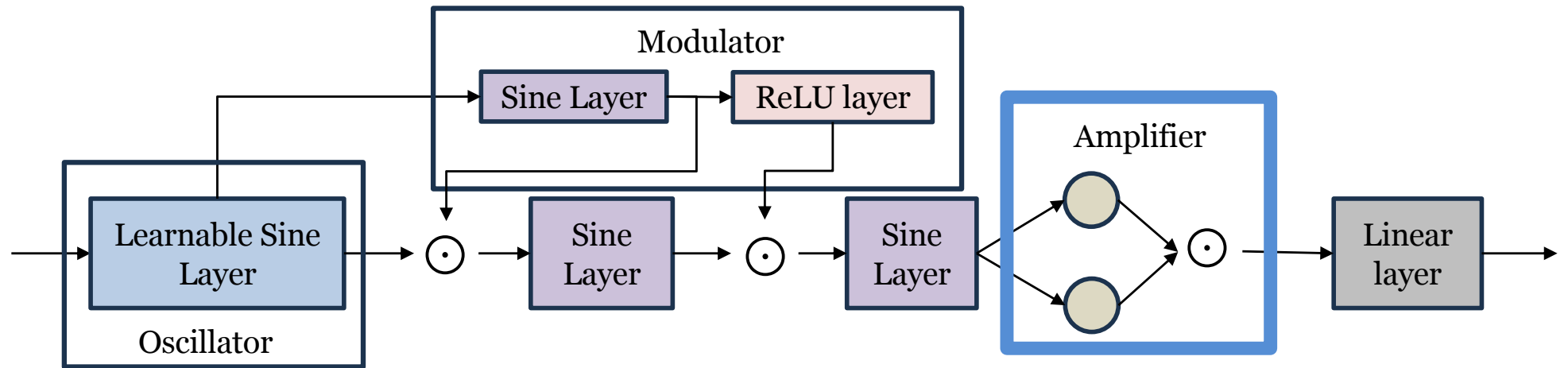
Modulation



- Learned multiplicative modulation
- **Create sum and difference frequencies** beyond integer harmonics

$$\alpha \sin(v_1 x) \cdot \beta \sin(v_2 x) = \frac{\alpha\beta}{2} [\cos((v_1 - v_2)x) - \cos((v_1 + v_2)x)]$$

Self-Multiplication (Amplifier)



- Amplify strong signals
- Create second harmonics, sum and difference frequencies with **no parameter overhead**

Evaluation

Kodak Image Dataset: 24 high-quality RGB images with a resolution of 768×512 pixels.

DIV2K Validation Set (LR-mild): 24 high-quality RGB images with a resolution of 510×339 pixels.

NeRF synthesis dataset: 8 objects, 100 pictures for each object to be trained and another 25 for testing. All the images are 400×400 .



Results on 2D (Quantitative)

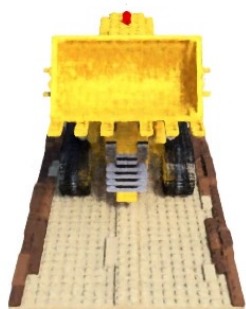
	Parameters	Average PSNR on Kodak	Average PSNR on DIV2K
Gaussian ³	272703	37.90	38.34
SIREN ²	272703	33.65	33.73
WIRE ⁴	265523	40.24	38.90
RINR ¹	289716	32.96	34.03
Ours(SMN)	264216	41.40	42.53



Results on NeRF Novel View Synthesis

	Parameters	Average PSNR
PE+Gaussian ²	287749	32.00
PE+Siren ²	287749	29.06
PE+Wire ⁴	291307	25.14
PE+RINR ¹	314 703	26.84
PE+MLP ⁵	290 370	26.66
PE+SMN	287749	32.98

Results on NeRF Novel View Synthesis



Ablation – Learnable Sine Layer

1 : $\sin(\omega_0 x) + \sin(\omega_1 x) + \sin(\omega_3 x)$

2 : $a_0 \sin(\omega_0 x)$

3 : $a_0 \sin(\omega_0 x) + a_1 \sin(\omega_1 x)$

4 : $a_0 \sin(\omega_0 x) + a_1 \sin(\omega_1 x) + a_2 \sin(\omega_3 x)$

	Average PSNR
1	35.08
2	42.87
3	42.97
4	43.68

Ablation – Additive vs. Multiplicative Modulation

Multiplication $f(\mathbf{x}) = \sin(\boldsymbol{\alpha} \odot \mathbf{x})$

scales the input, effectively changing the frequency response

Addition $f(\mathbf{x}) = \sin(\mathbf{x} + \boldsymbol{\beta})$

shifts the phase but does not affect the frequency content

	Average PSNR
Mul	41.40
Add	40.25

of Parameters: 264216



Paper, code, supplementary

Thanks!

p.s. Check out our new work that generates neural fields in weight space
Weight Space Representation Learning via Neural Field Adaptation