Non-separable Wavelet Transform Using Learnable Convolutional Lifting Steps

Joao O. Parracho[⋆†], Eduardo A. B. da Silva[†], Lucas A. Thomaz[⋆‡], Luis M. N. Tavora[⋆‡], and Sergio M. M. Faria[⋆‡]

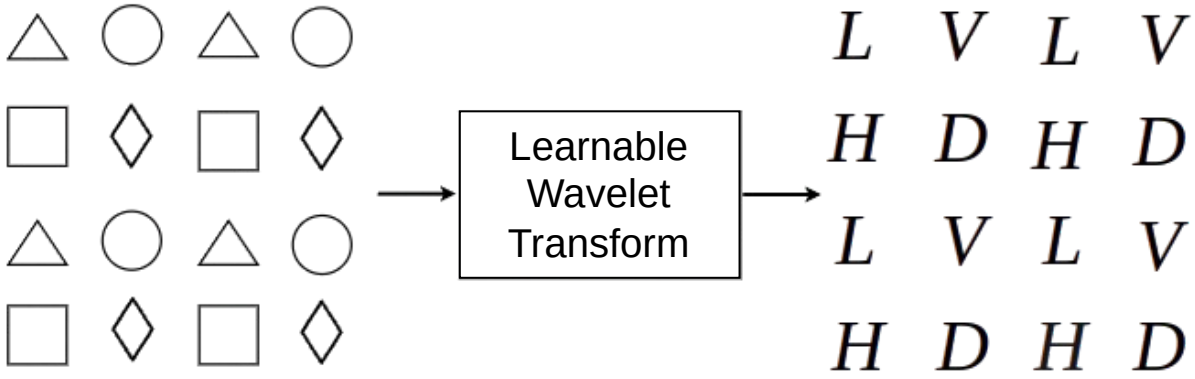
[⋆] Instituto de Telecomunicações, Leiria, Portugal
[‡]ESTG, Polytechnic of Leiria, Leiria, Portugal
[†]Electrical Engineering Program, COPPE — Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Summary

Objectives

- 2D non-separable learnable Wavelet
  - CNN-based lifting scheme
- End-to-end optimised codec

Learnable Wavelet transf

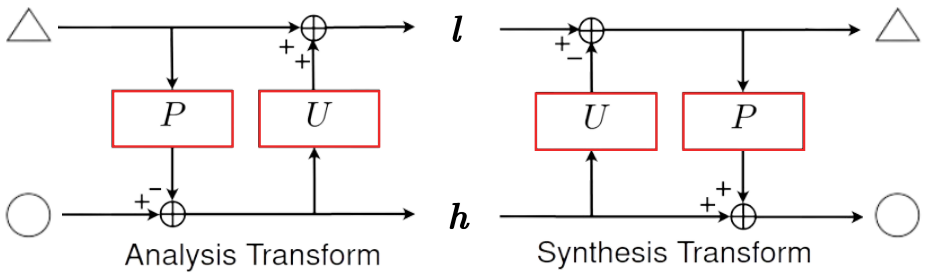


- It generates **four** distinct subbands:
  - Low-pass (L)
  - Horizontal (H)
  - Vertical (V)
  - Diagonal (D)

Learnable Wavelet

HOW ?

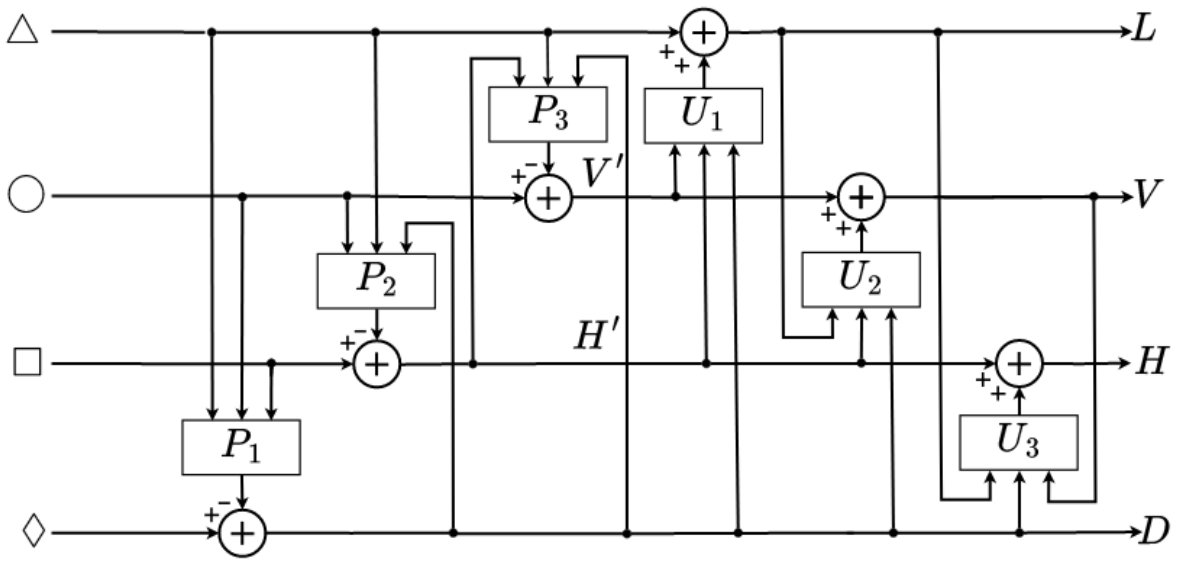
1D Lifting scheme



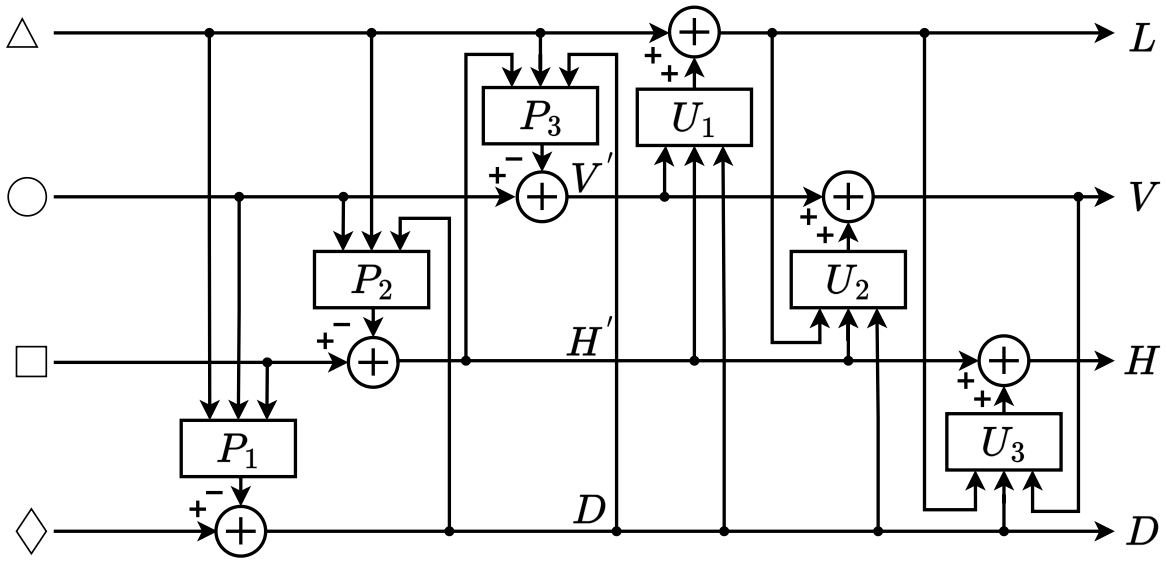
- **Predict operator** $P(\cdot)$: Predict a polyphase component based on the other one
- **Update operator** $U(\cdot)$: Refine coarse details

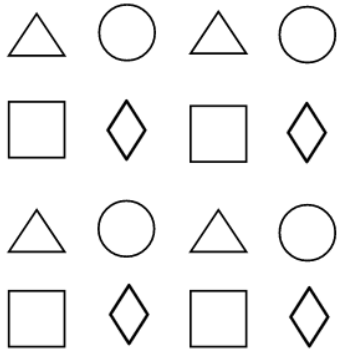
Generic Non-Separable 2D Lifting Scheme

- It comprises **three predict** (P_1, P_2, P_3) and **three update** (U_1, U_2, U_3) steps, with three inputs each
- The lifting operators can use information across all directions

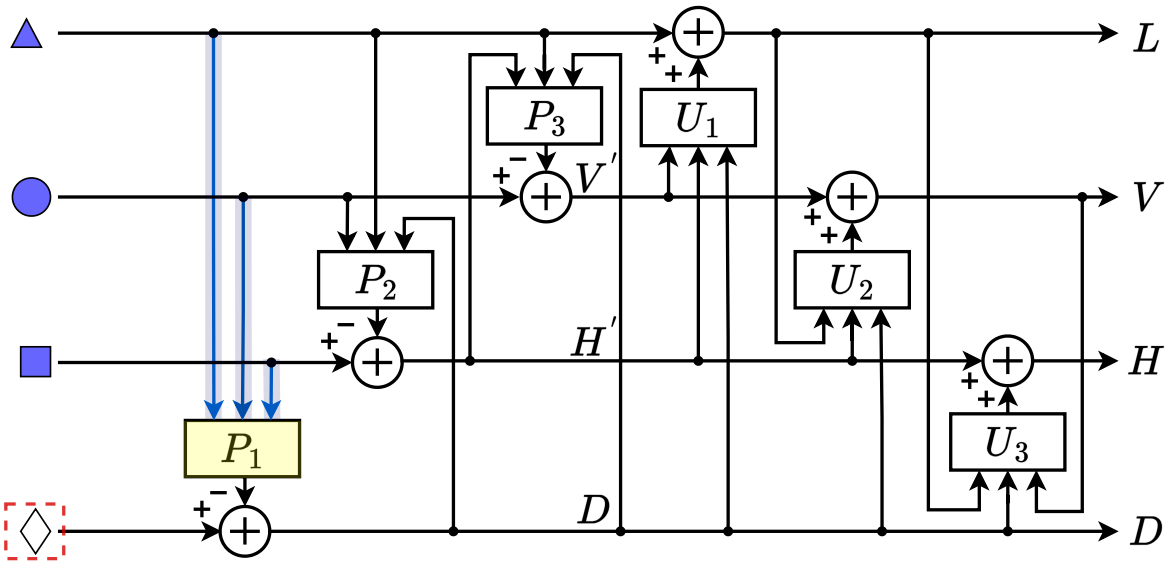


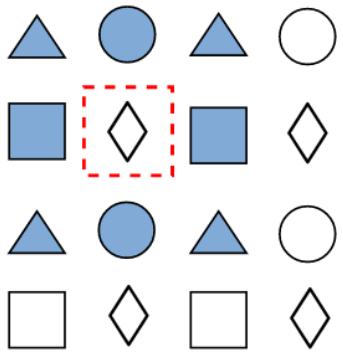
Proposed Non-Separable 2D Lifting





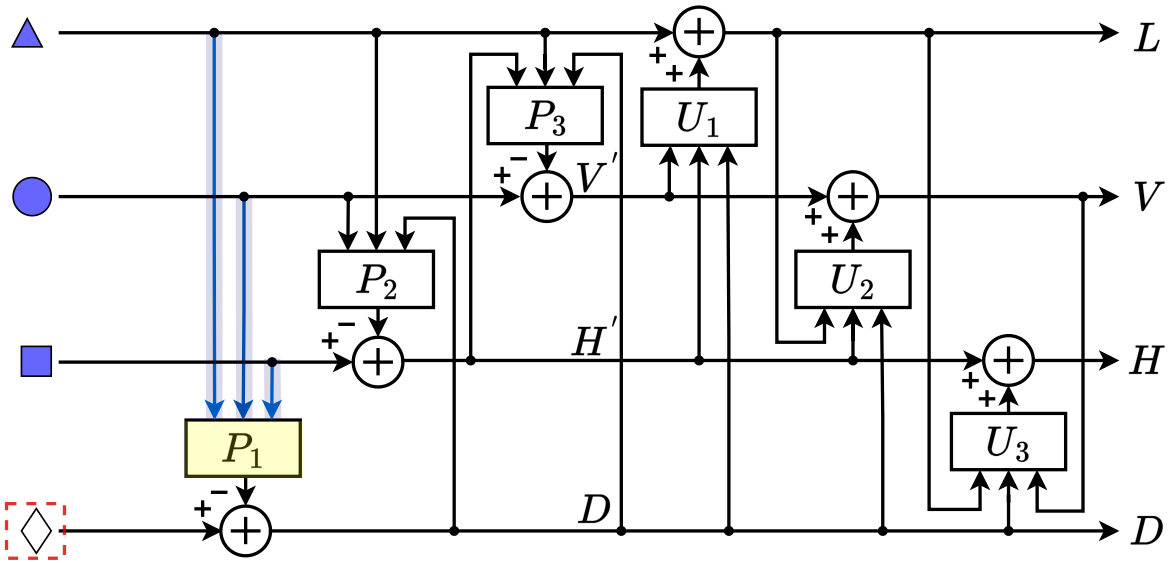
Proposed Non-Separable 2D Liftings: P_1

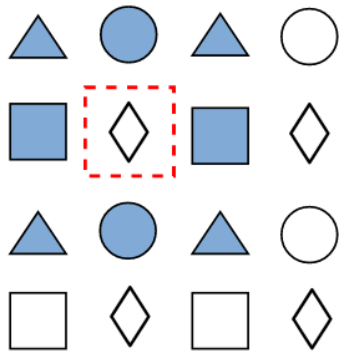




$$\begin{bmatrix} a_{01} & a_{02} & a_{03} & 0 \\ a_{10} & 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

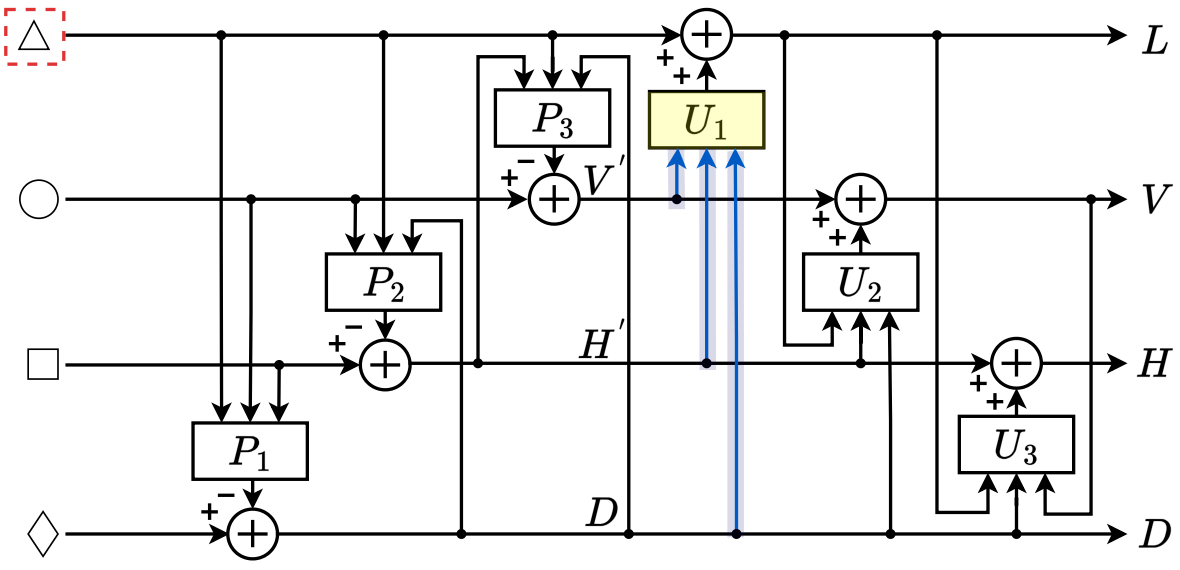
Proposed Non-Separable 2D Liftings: P_1





Stride $= 2$

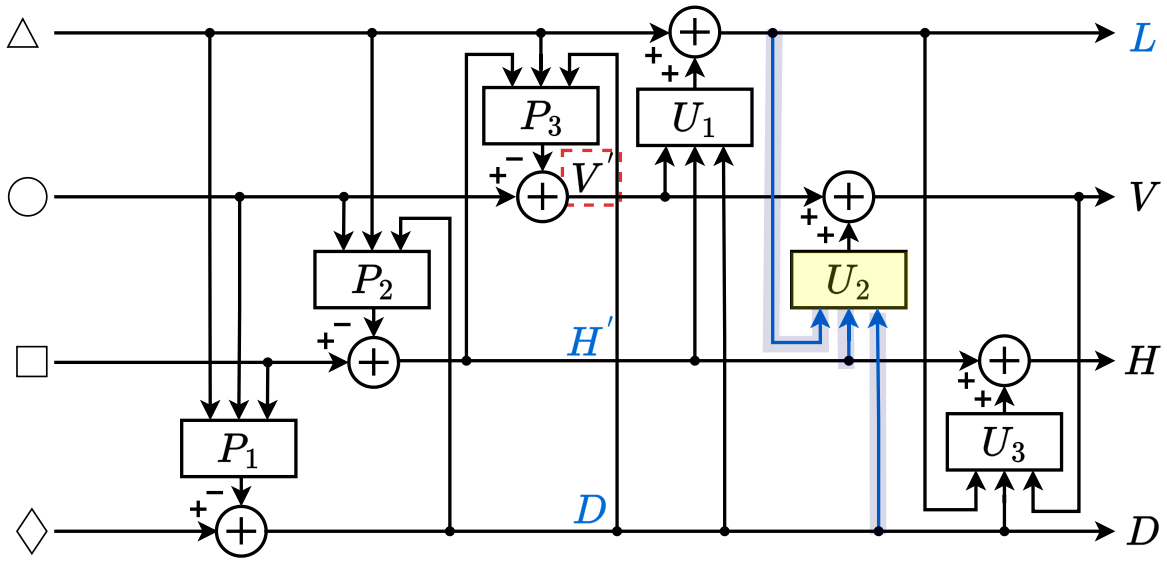
Proposed Non-Separable 2D Lifting: U_1



$$\begin{array}{cccc} \triangle & V' & \triangle & V' \\ H' & D & H' & D \\ \triangle & V' & \triangle & V' \\ H' & D & H' & D \end{array}$$

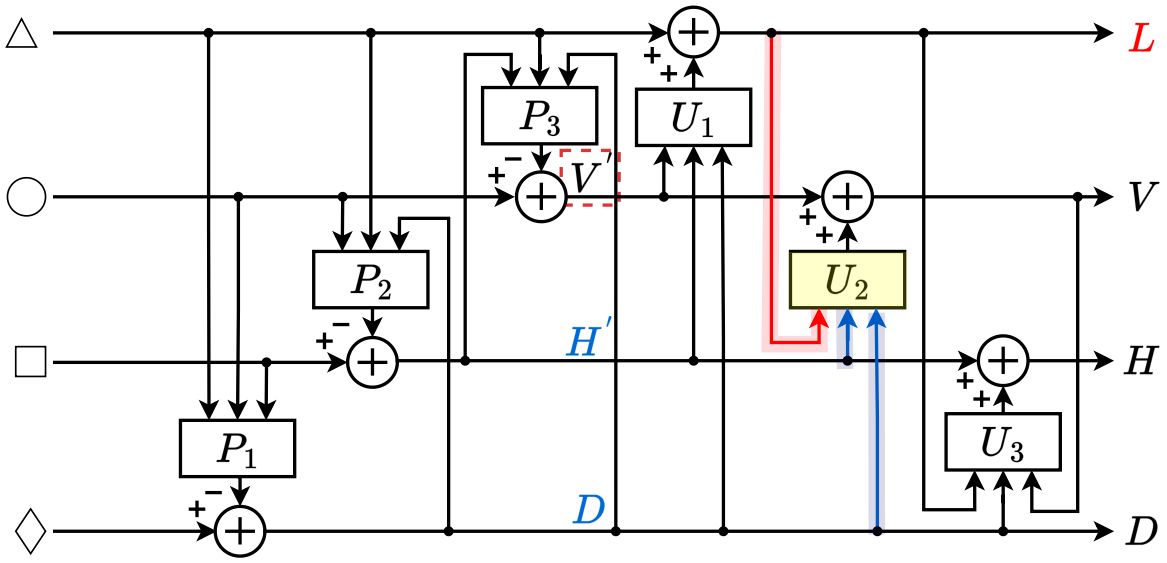
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{11} & a_{12} & a_{13} \\ 0 & a_{21} & 0 & a_{23} \\ 0 & a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Proposed Non-Separable 2D Lifting: U_2



$$\begin{array}{cccc} L & V' & L & V' \\ H' & D & H' & D \\ L & V' & L & V' \\ H' & D & H' & D \end{array}$$

Proposed Non-Separable 2D Lifting: U_2



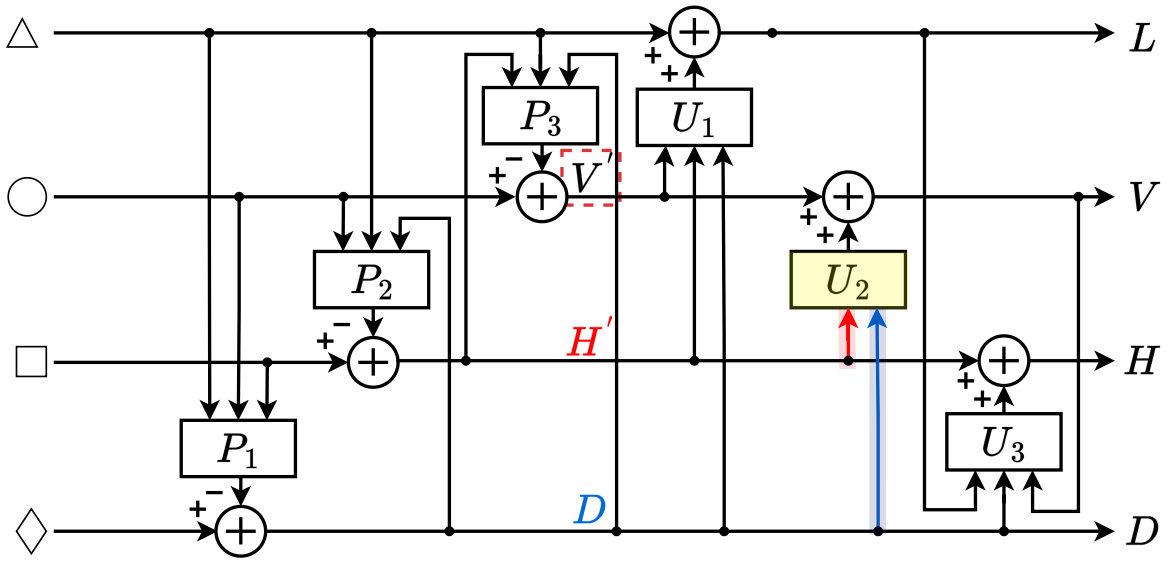
L V' L V'

H' D H' D

L V' L V'

H' D H' D

Proposed Non-Separable 2D Lifting: U_2



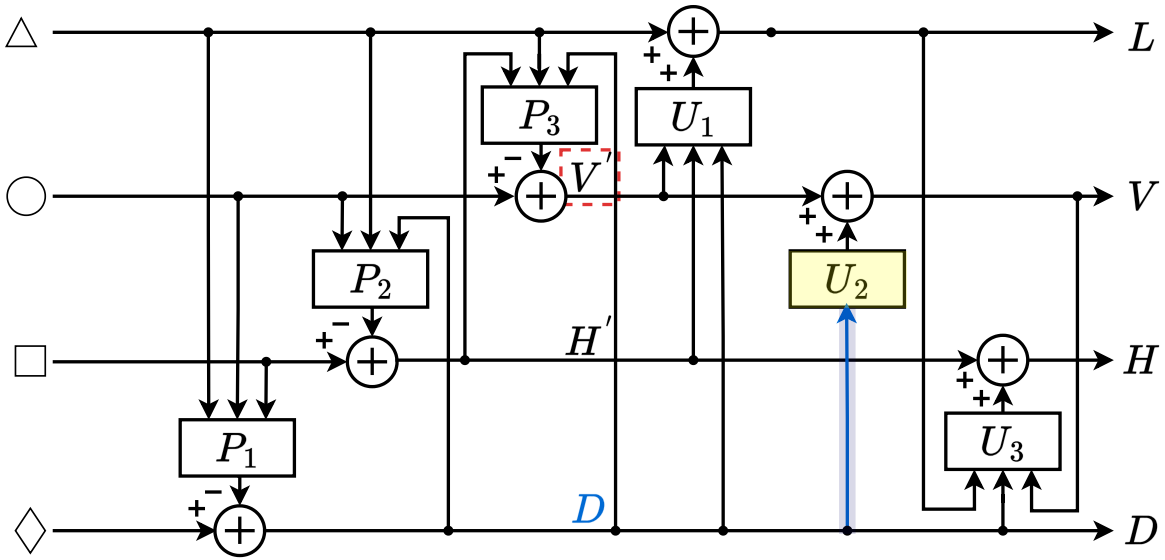
L V' L V'

H' D H' D

L V' L V'

H' D H' D

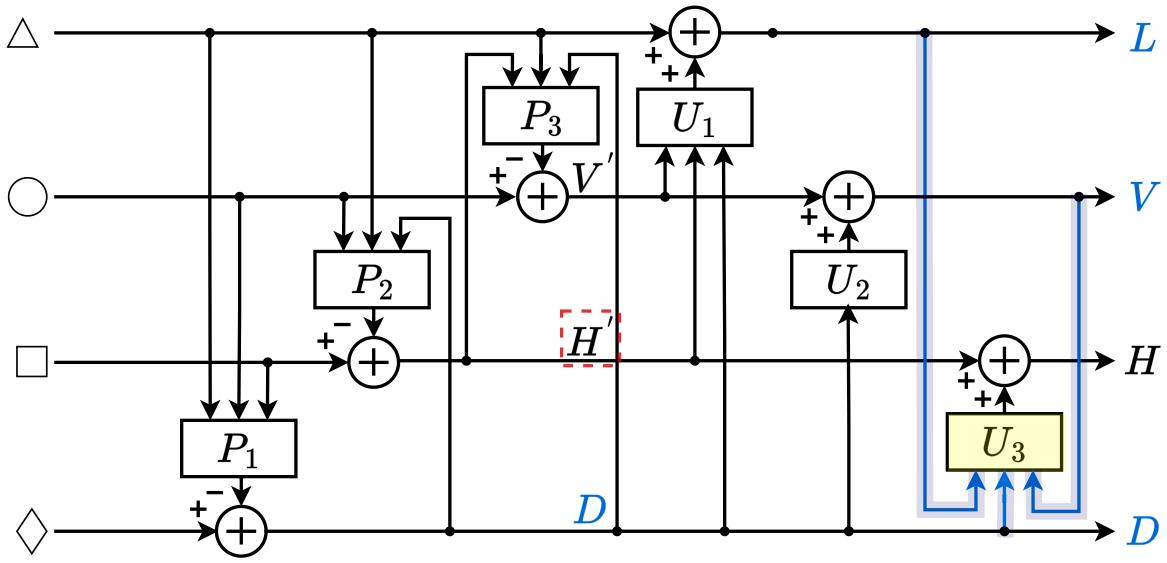
Proposed Non-Separable 2D Lifting: U_2



$$\begin{matrix} L & V' & L & V' \\ H' & D & H' & D \\ L & V' & L & V' \\ H' & D & H' & D \end{matrix}$$

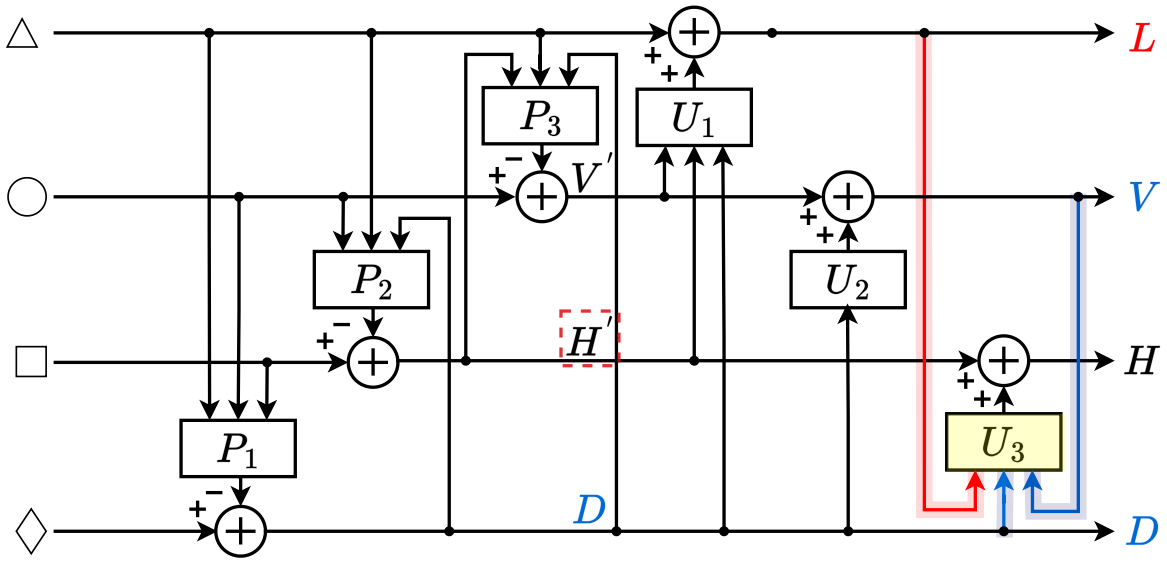
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & a_{31} & 0 & 0 \end{bmatrix}$$

Proposed Non-Separable 2D Lifting: U_3



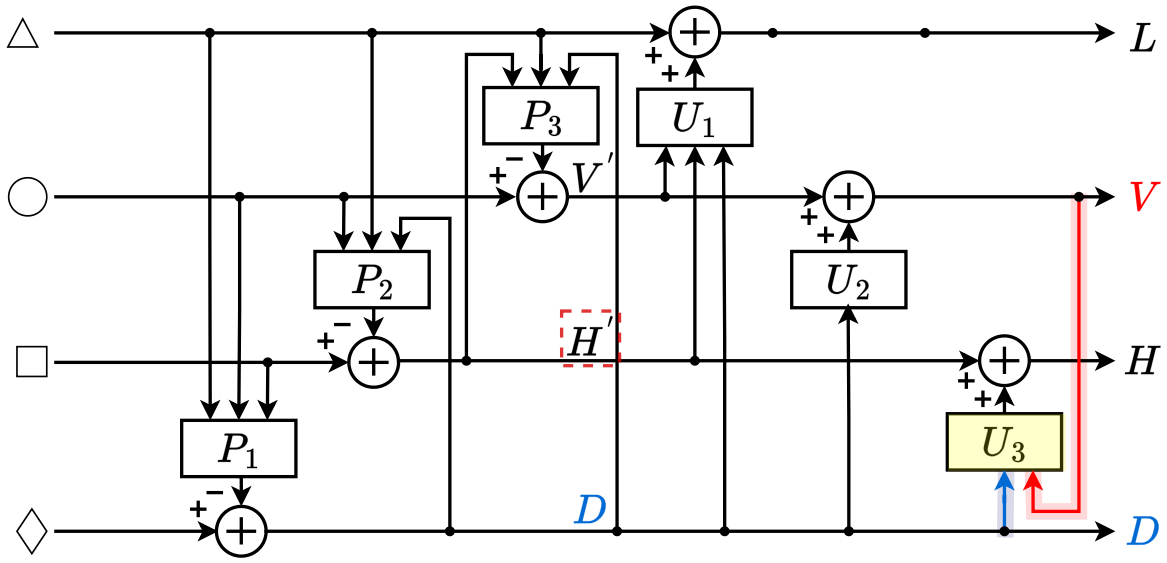
$$\begin{matrix} L & V & L & V \\ H' & D & H' & D \\ L & V & L & V \\ H' & D & H' & D \end{matrix}$$

Proposed Non-Separable 2D Lifting: U_3



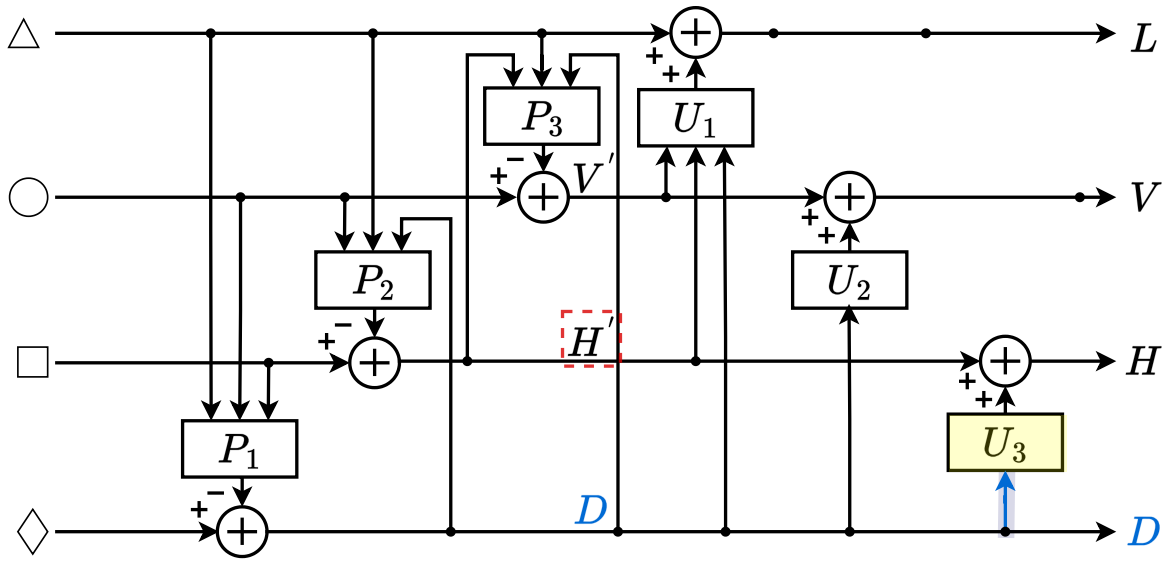
$$\begin{matrix} L & V & L & V \\ H' & D & H' & D \\ L & V & L & V \\ H' & D & H' & D \end{matrix}$$

Proposed Non-Separable 2D Lifting: U_3



$$\begin{matrix} L & V & L & V \\ H' & D & H' & D \\ L & V & L & V \\ H' & D & H' & D \end{matrix}$$

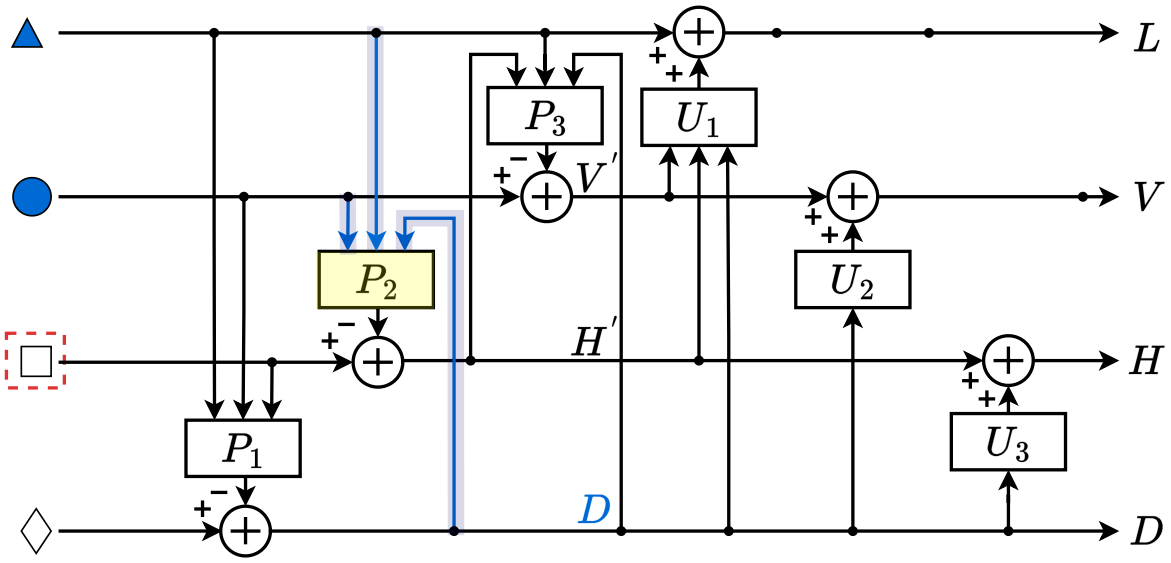
Proposed Non-Separable 2D Lifting: U_3

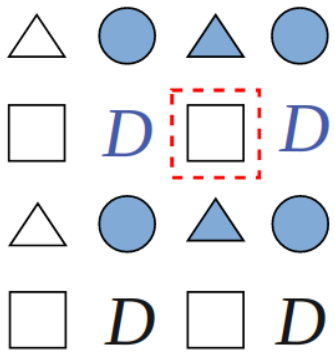


$$\begin{matrix} L & V & L & V \\ H' & D & H' & D \\ L & V & L & V \\ H' & D & H' & D \end{matrix}$$

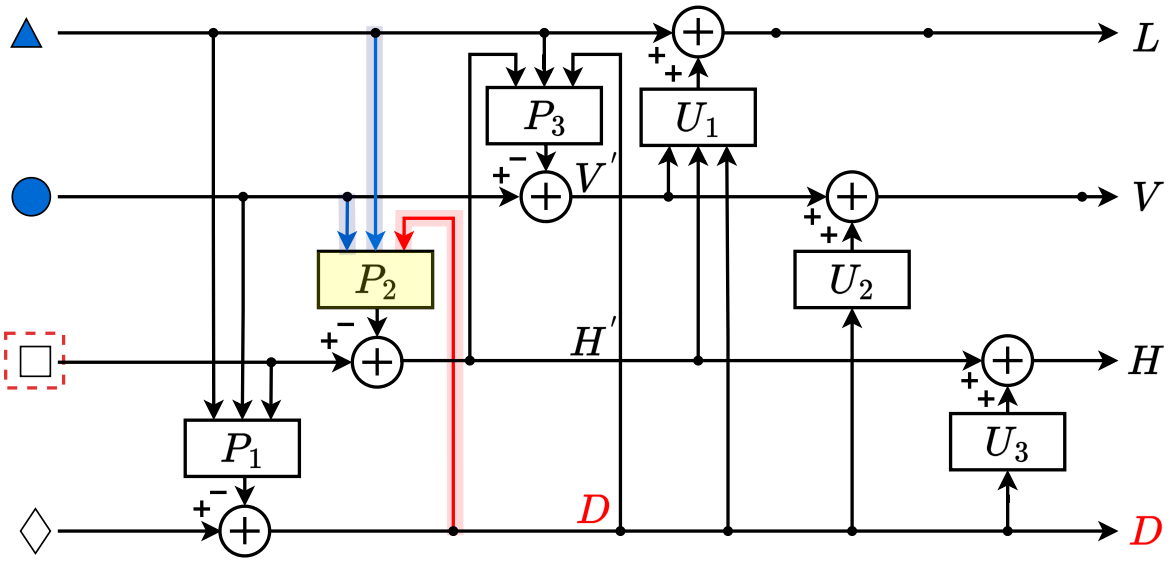
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{11} & 0 & a_{13} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

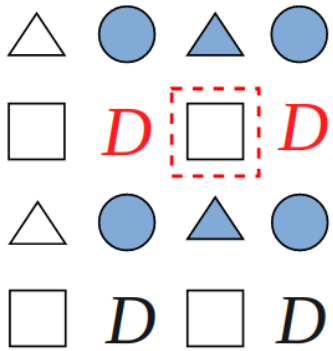
Proposed Non-Separable 2D Lifting: P_2



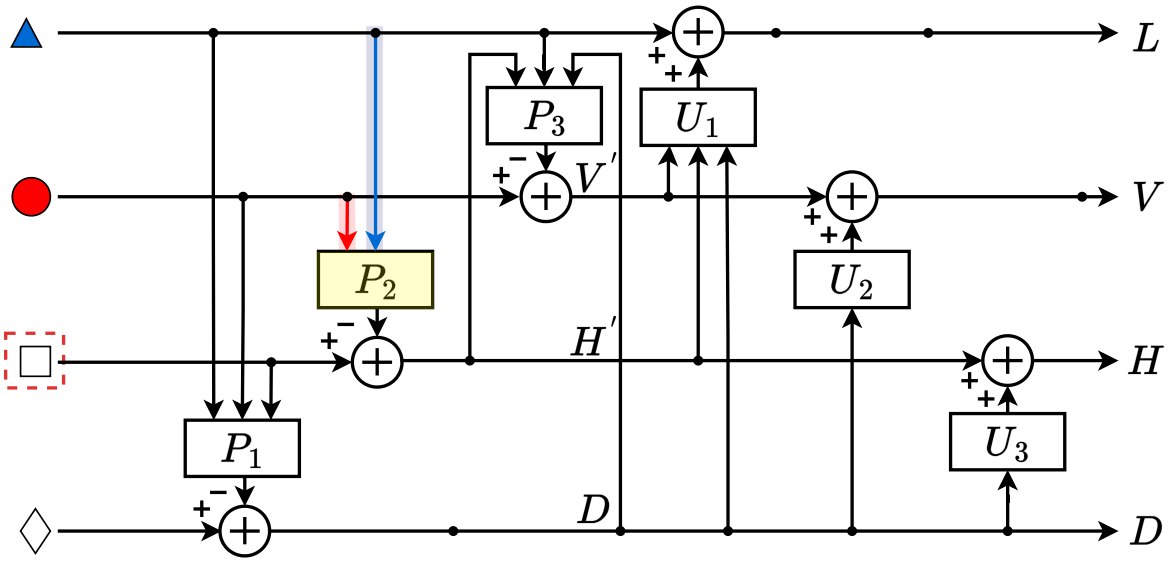


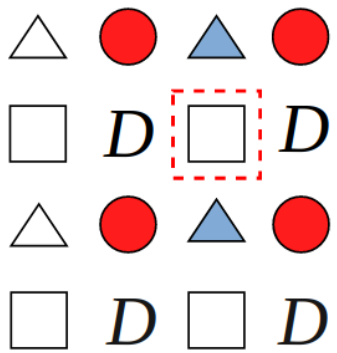
Proposed Non-Separable 2D Lifting: P_2



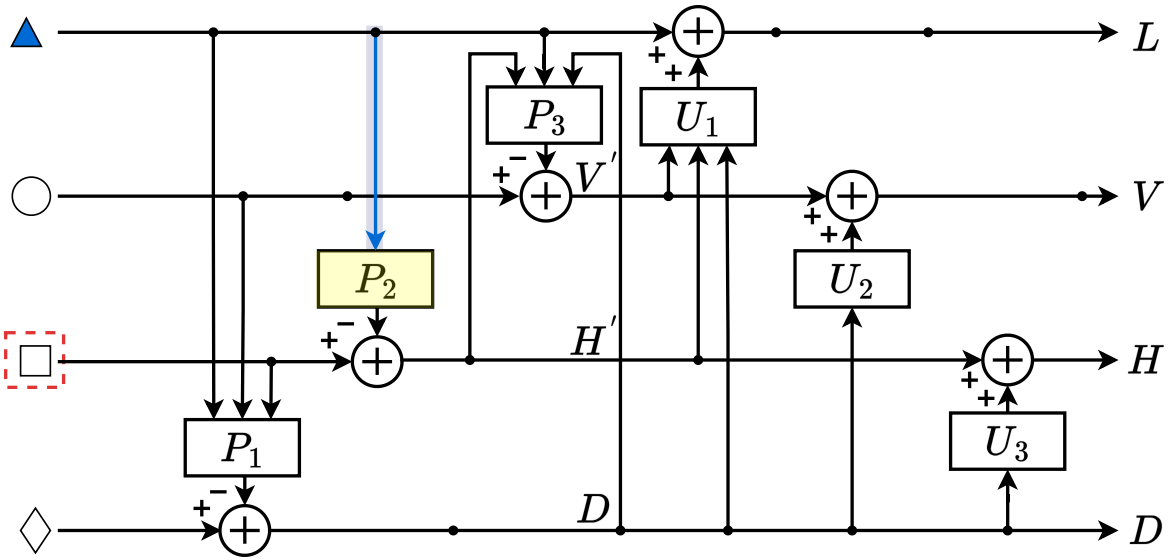


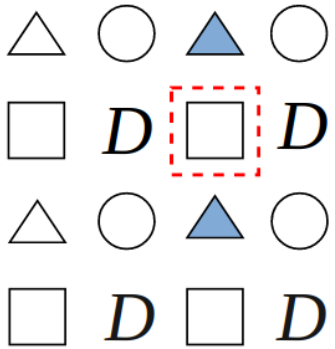
Proposed Non-Separable 2D Lifting: P_2





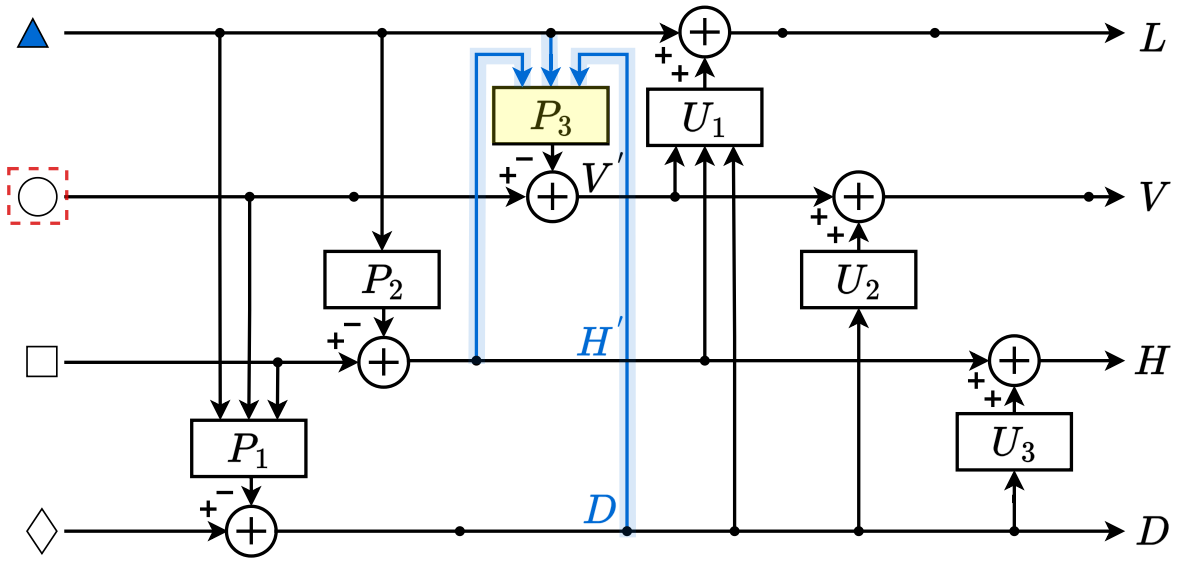
Proposed Non-Separable 2D Lifting: P_2

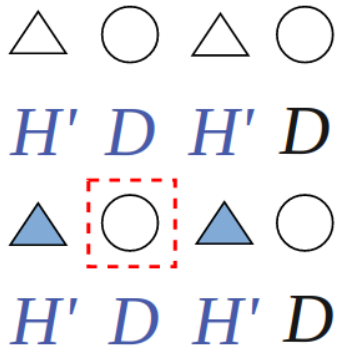




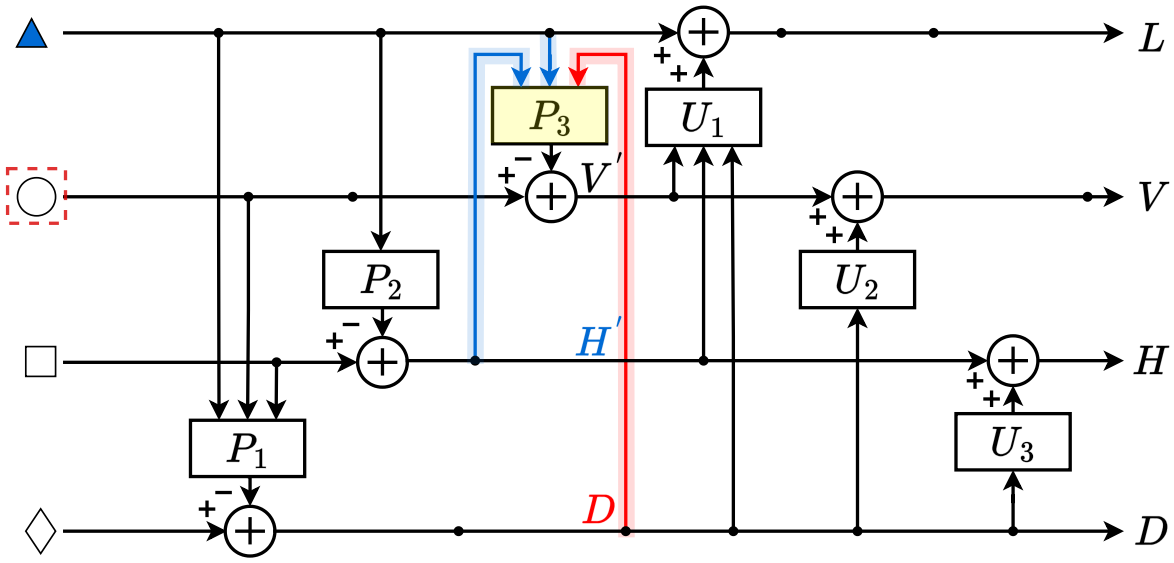
$$\begin{bmatrix} 0 & 0 & a_{03} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_{23} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

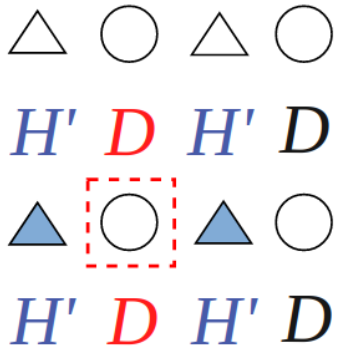
Proposed Non-Separable 2D Lifting: P_3



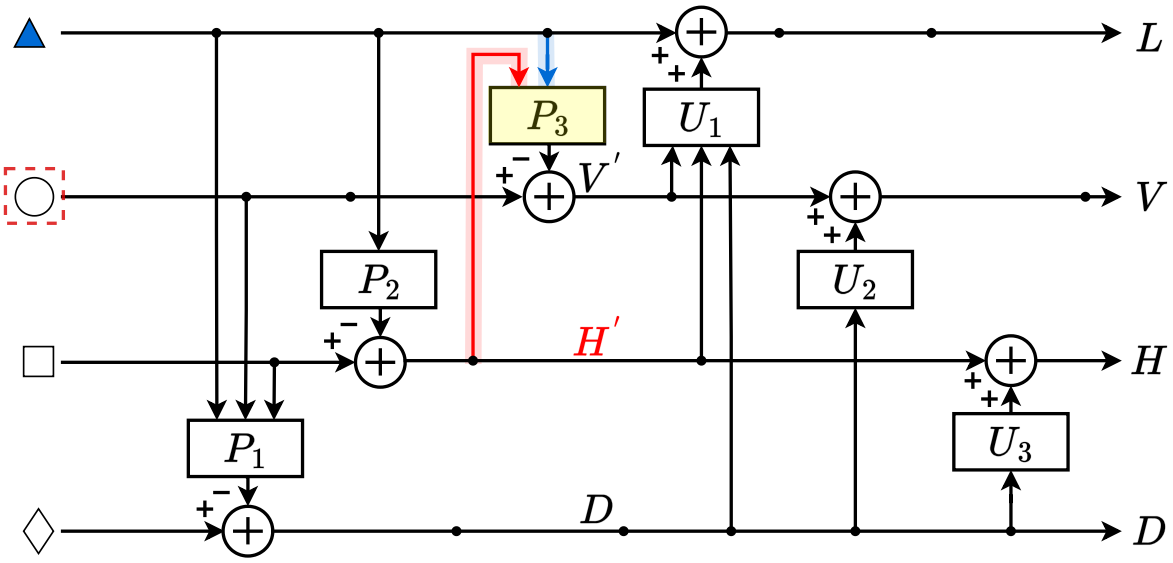


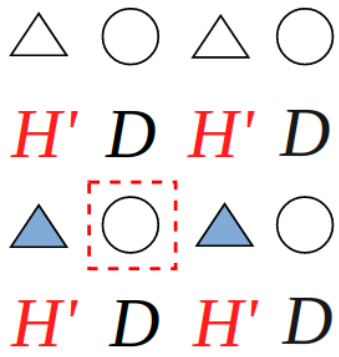
Proposed Non-Separable 2D Lifting: P_3



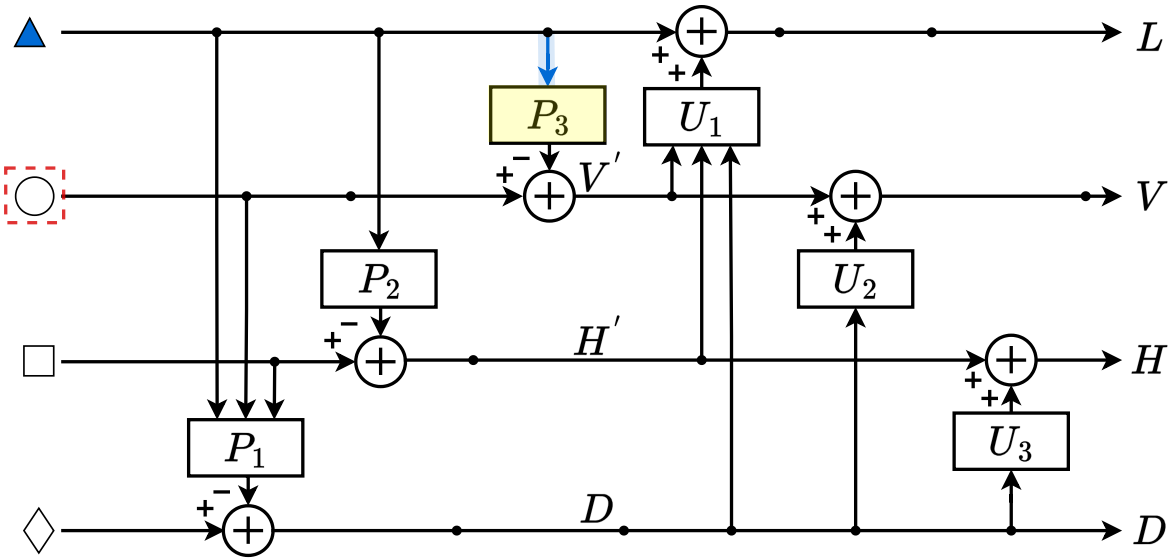


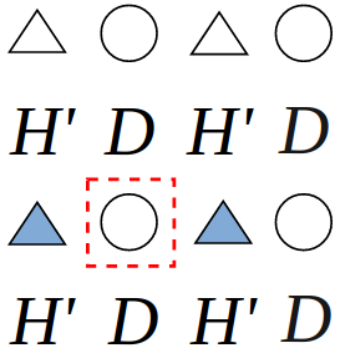
Proposed Non-Separable 2D Lifting: P_3





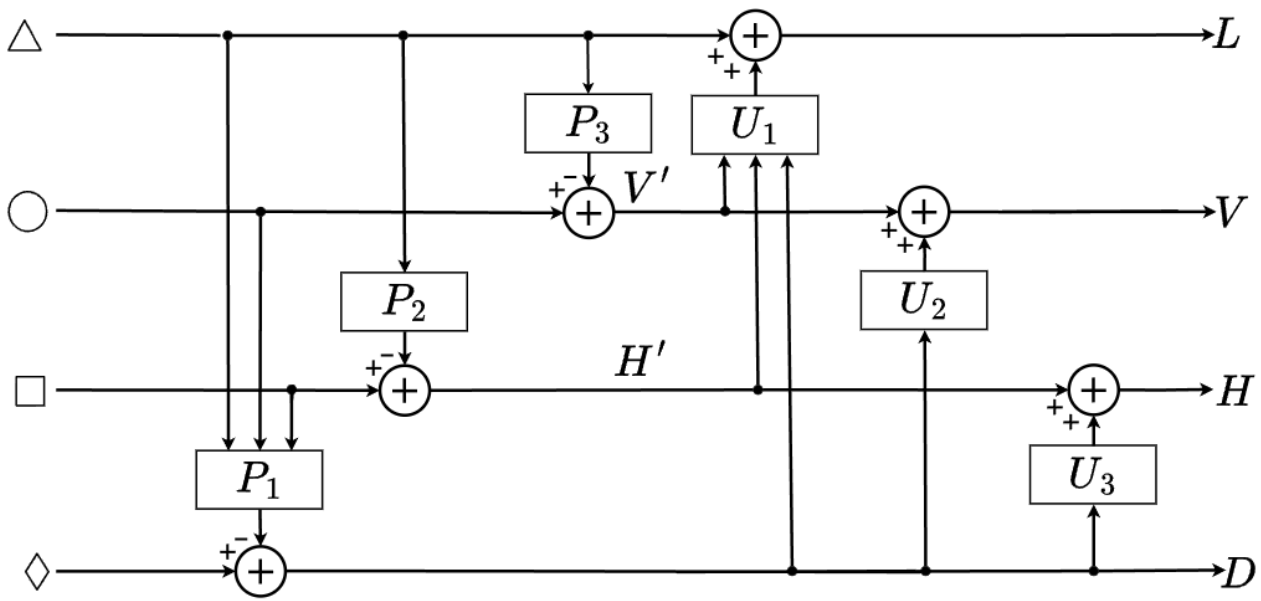
Proposed Non-Separable 2D Lifting: P_3





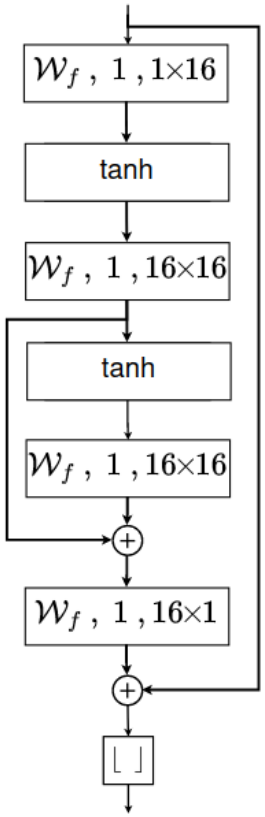
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{21} & 0 & a_{23} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Proposed Non-Separable 2D Lifting



$$\begin{matrix} L & V & L & V \\ H & D & H & D \\ L & V & L & V \\ H & D & H & D \end{matrix}$$

How can the lifting operators be implemented through CNNs?

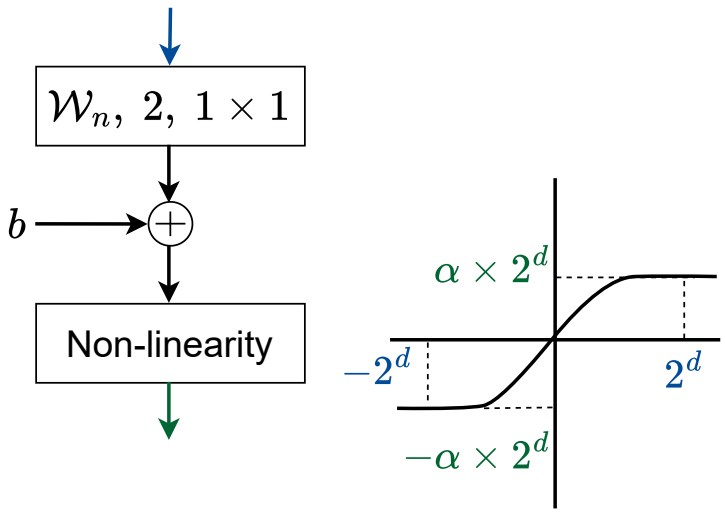


- Architecture proposed in the seminal paper iWave++ [1]
- **ResNet** architecture with **multi-channel** convolutional layers
- Separable wavelet transform

[1] H. Ma, D. Liu, N. Yan, H. Li, and F. Wu, "End-to-end optimized versatile image compression with wavelet-like transform," IEEE Tran. on Pattern Analysis and Machine Intelligence, vol. 44, no. 3, pp. 1247–1263, Sep. 2022.

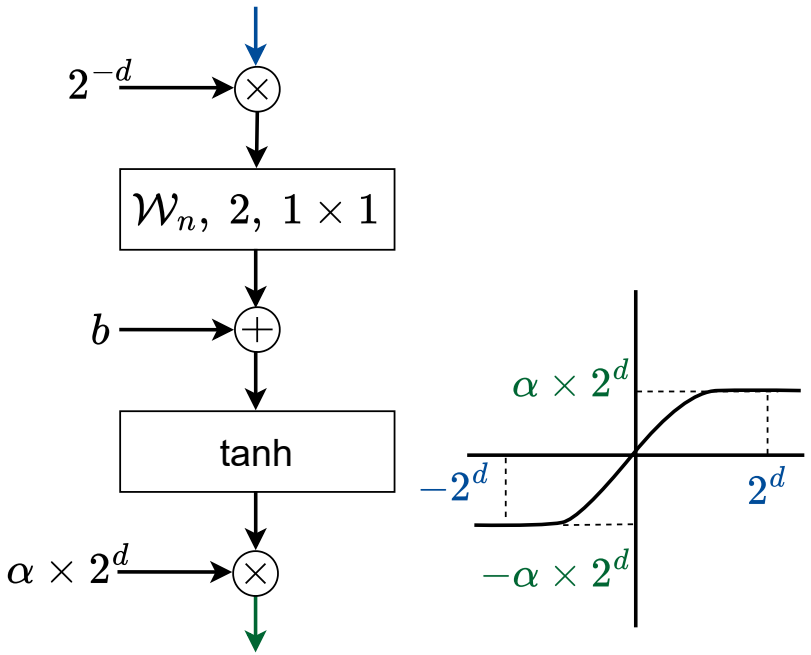
Does it need to have so many coefficients
Are that many channels needed

Proposed learnable lifting steps



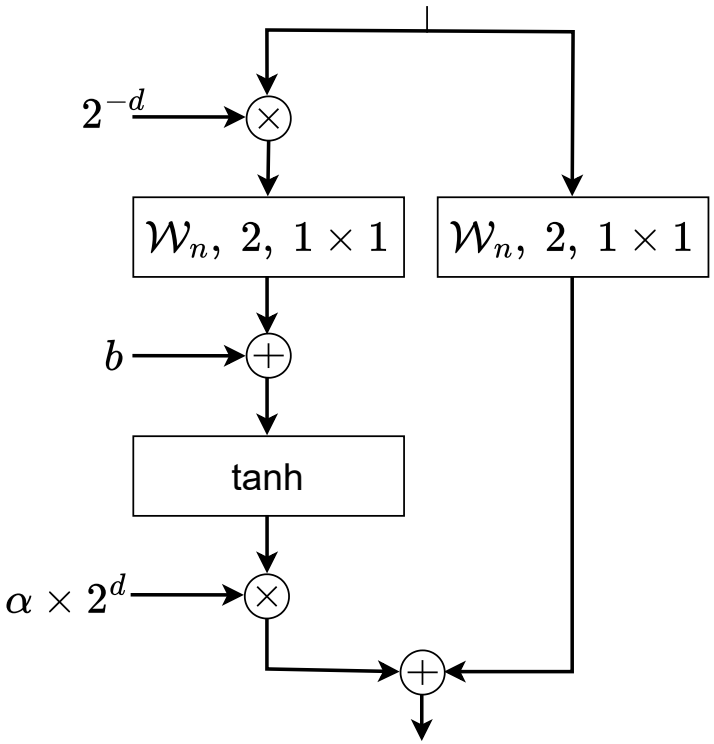
- d: Bit-depth of the input image
- α: Learnable parameter

Proposed learnable lifting steps



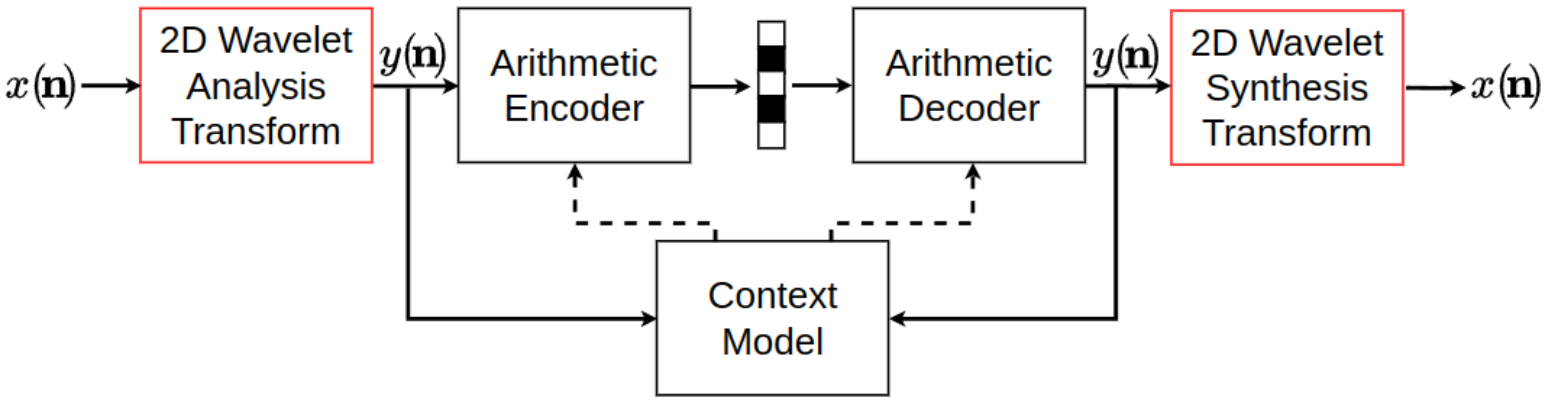
- d: bit-depth of the input image
- α: learnable parameter

Proposed learnable lifting steps



- d: bit-depth of the input image
- α: learnable parameter
- Initialised with the corresponding predict or update operator of the separable 5-3 wavelet

End-to-end coder

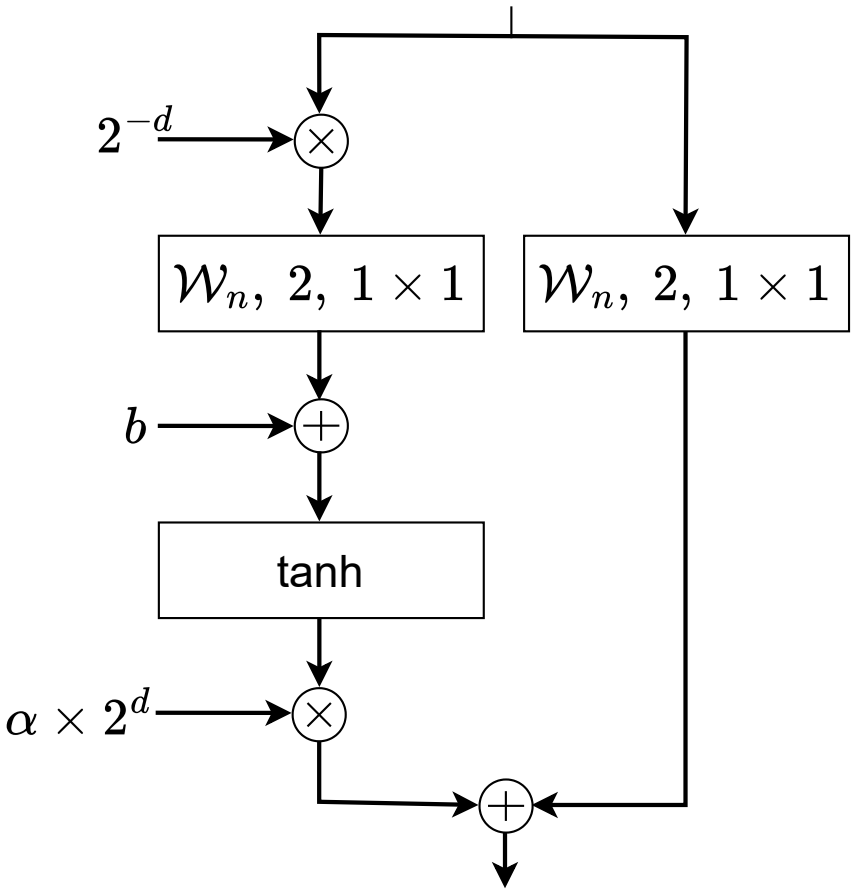


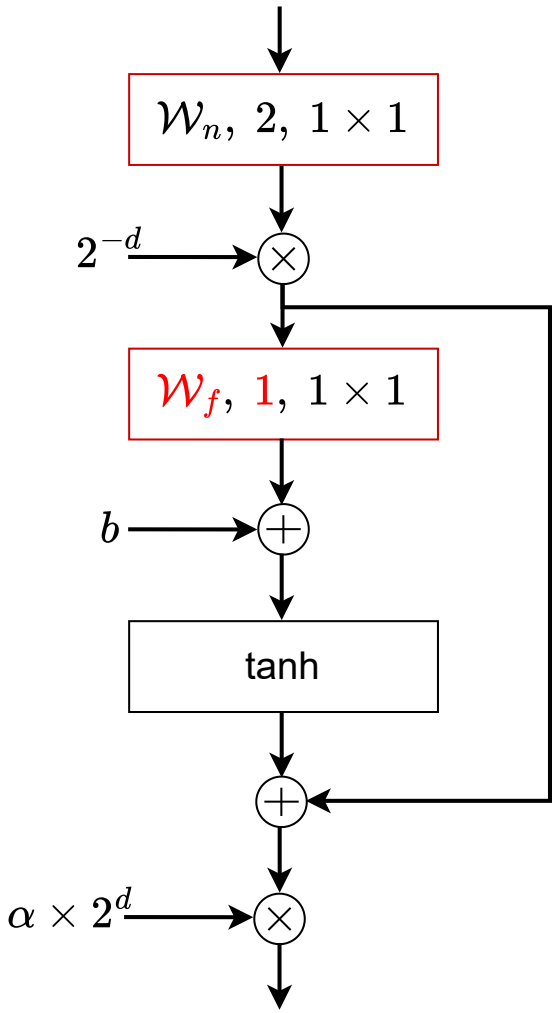
- **Reference**: Lossless version of iWave++ codec
- Replaced with the **proposed 2D non-separable wavelet transform**

2D NSWT-LCLS - Kodak

2D Wavelet transform	**Predict/Update modules**	**No. of learnable coefficients**	**Rate (bpp)**
iWave++	Multi-channel ResNet	39560	8.97
2D NSWT-LCLS	**Proposed**	**264**	**8.70**

Experiment 1: Variant A

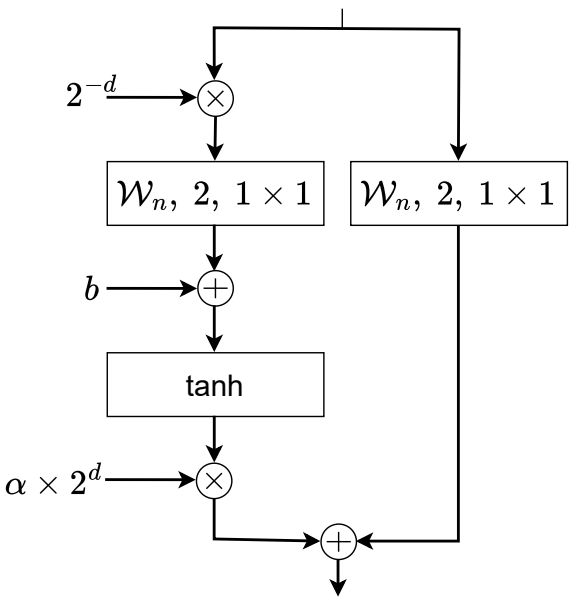


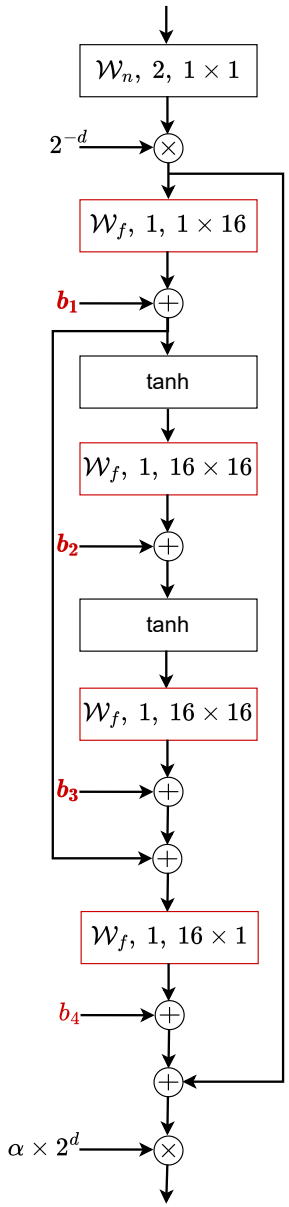


Experiment 1: Comparison of the proposed learnable lifting operator with variant A

Predict/Update Modules	**No. of learnable coefficients**	**Rate (bpp)**		
		CLIC.pro	**DIV2K**	**Kodak**
Proposed	**264**	**7.70**	**7.89**	**8.70**
Variant A	384	7.73	8.03	8.71

Experiment 2: Variant B (iWave++)

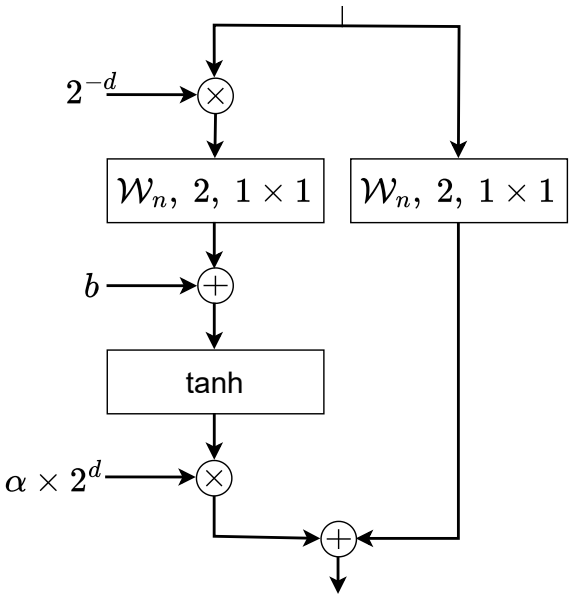


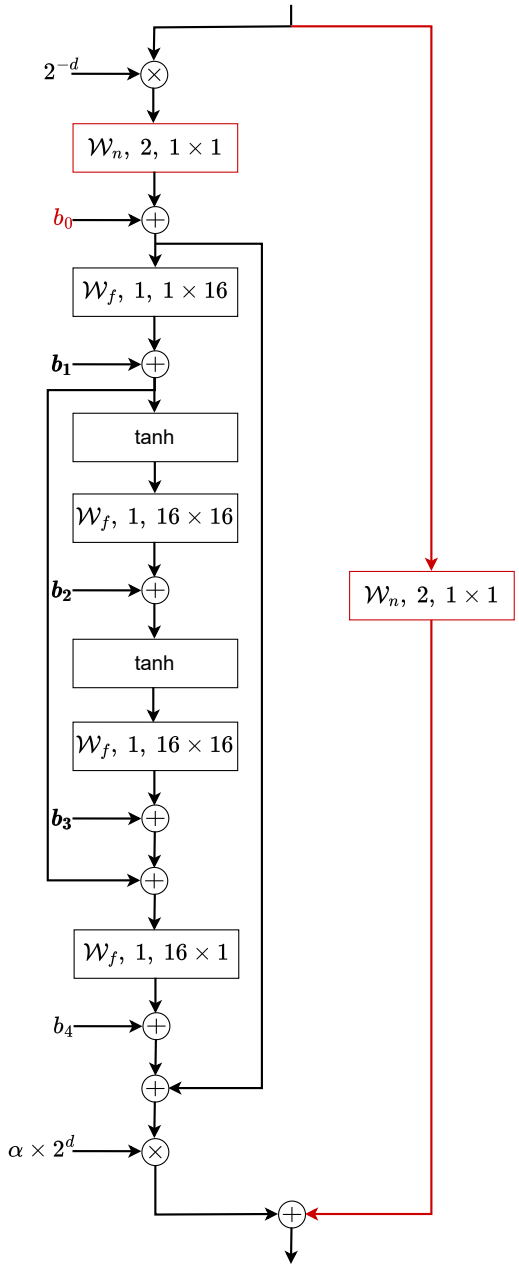


Experiment 2: Comparison of the proposed learnable lifting operator with variant B

Predict/Update Modules	**No. of learnable coefficients**	**Rate (bpp)**		
		CLIC.pro	**DIV2K**	**Kodak**
Proposed	**264**	**7.70**	**7.89**	**8.70**
Variant B	118824	7.79	8.07	8.80

Experiment 3: Variant C

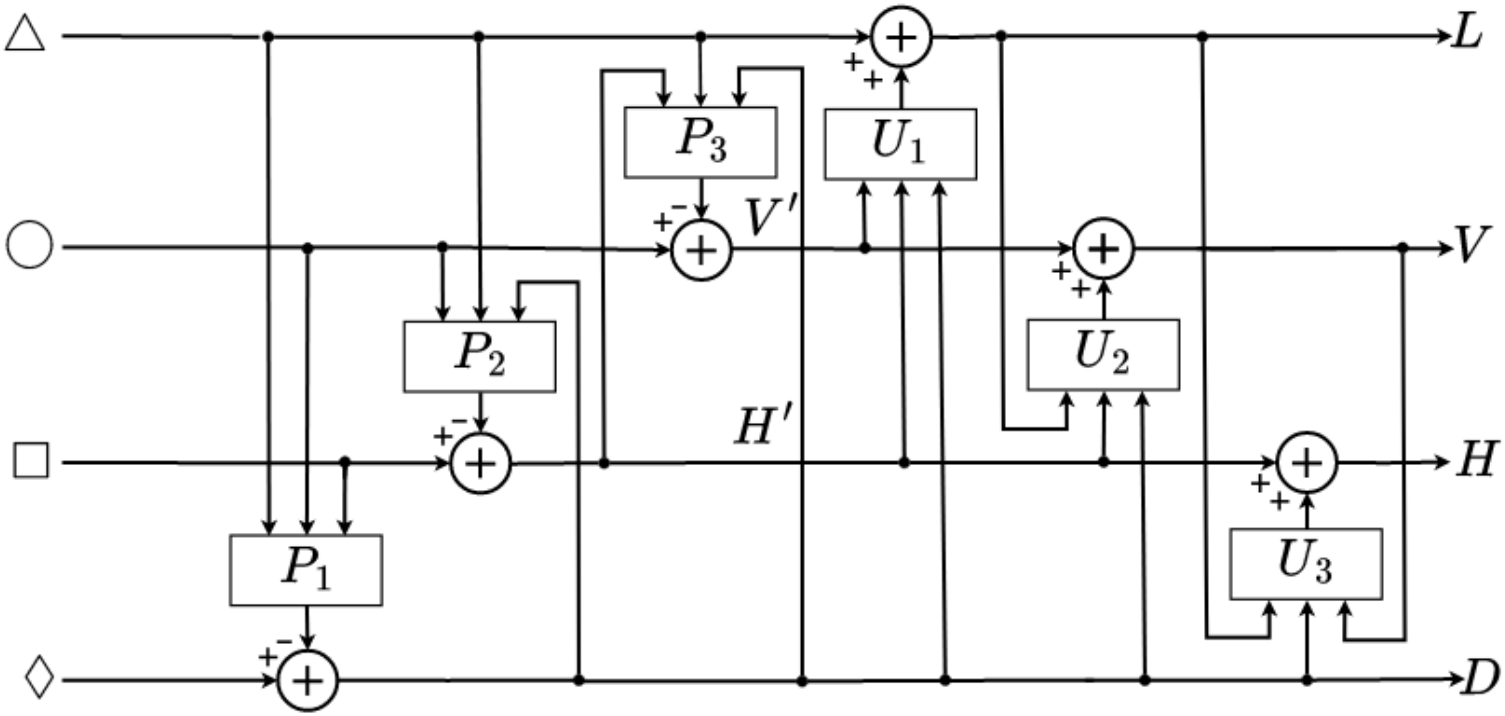




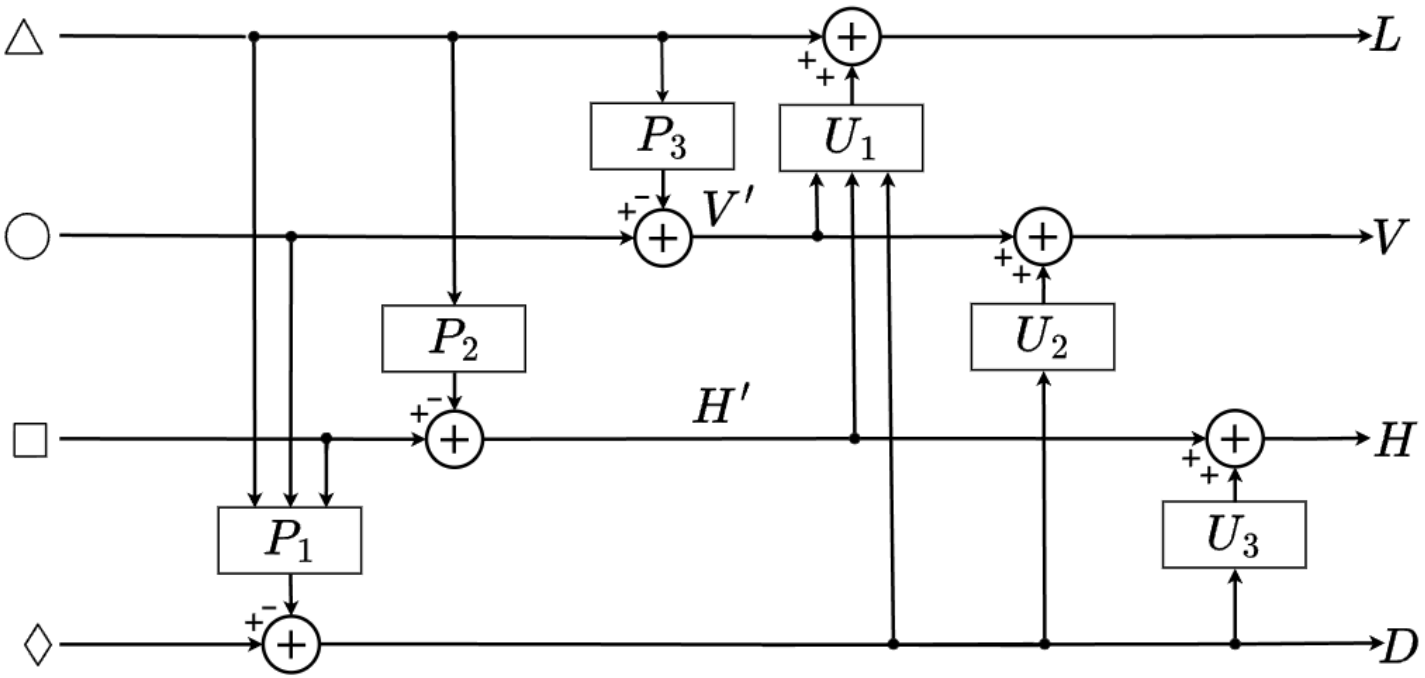
Experiment 3: Comparison of the proposed learnable lifting operator with variant C

Predict/Update Modules	**No. of learnable coefficients**	**Rate (bpp)**		
		CLIC.pro	**DIV2K**	**Kodak**
Proposed	**264**	**7.70**	**7.89**	**8.70**
Variant C	118944	7.88	8.15	8.83

Generic Non-Separable 2D Lifting



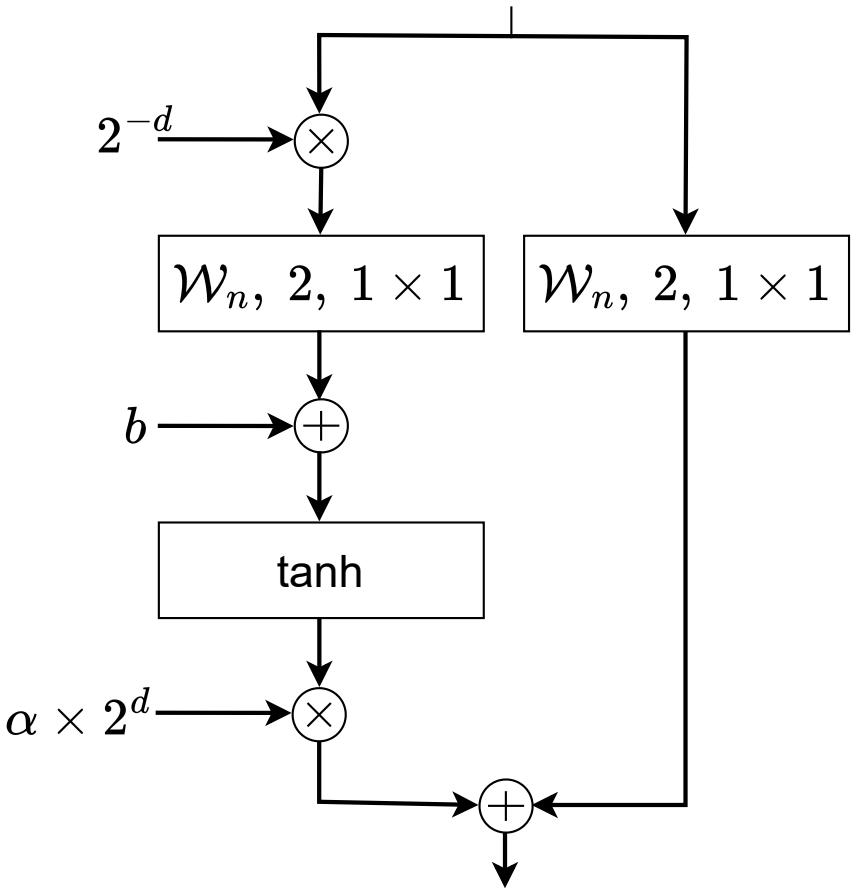
Proposed Non-Separable 2D Lifting

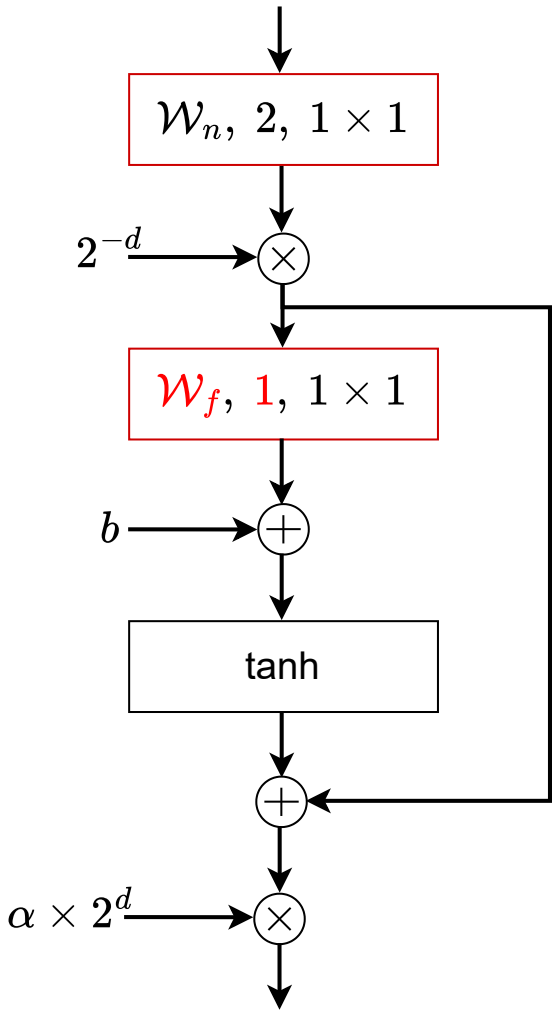


Comparison between the generic and proposed 2D non-separable lifting architecture

2D Non-Separable Lifting Architecture	**Predict/Update Modules**	**No. of learnable coefficients**	**Rate (bpp)**		
			CLIC.pro	**DIV2K**	**Kodak**
Proposed	**Proposed**	**264**	**7.70**	**7.89**	**8.70**
Generic	Proposed	456	7.82	8.10	8.82

Experiment 5: Variant A

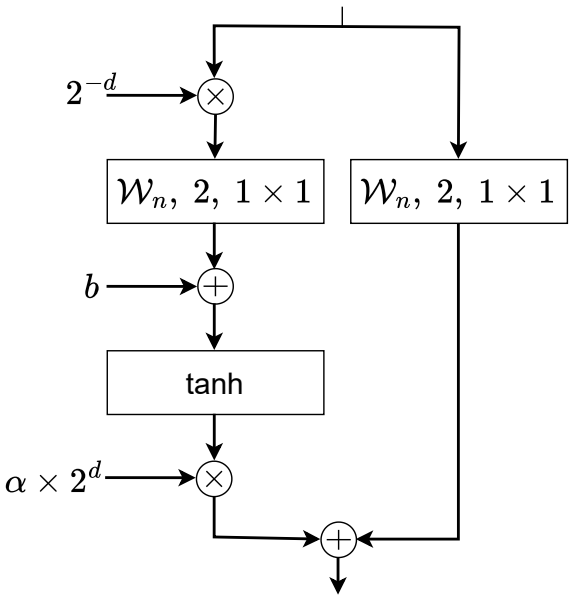


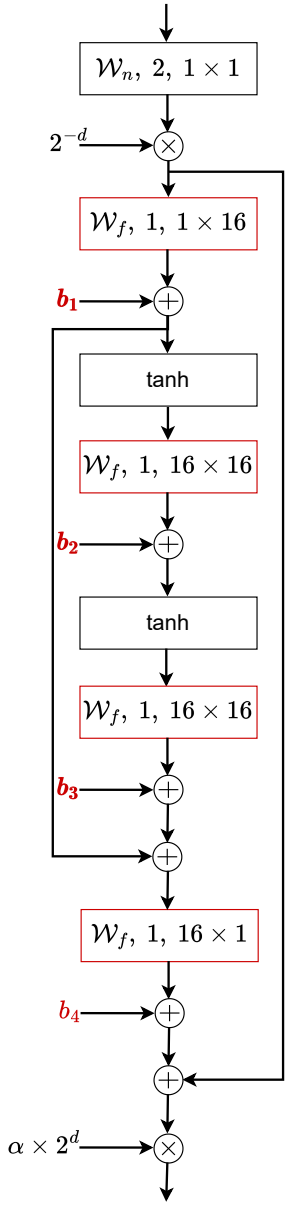


Experiment 5: Variant A

2D Non-Separable Lifting Architecture	**Predict/Update Modules**	**No. of learnable coefficients**	**Rate (bpp)**		
			CLIC.pro	**DIV2K**	**Kodak**
Proposed	**Proposed**	**264**	**7.70**	**7.89**	**8.70**
Generic	Variant A	480	7.87	8.13	8.84

Experiment 6: Variant B (iWave++)

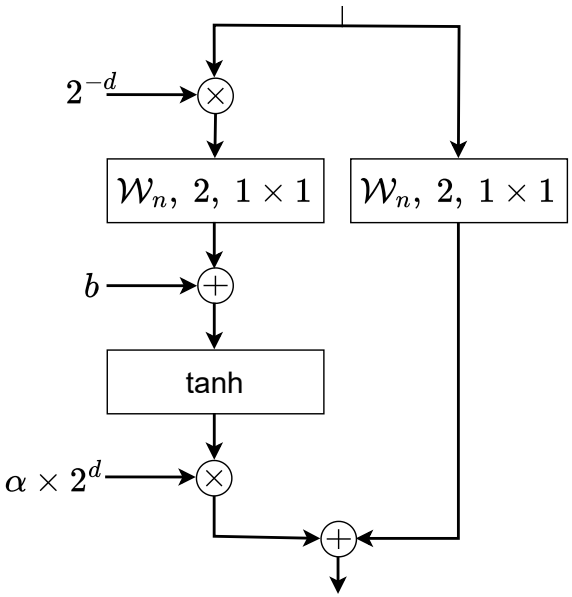


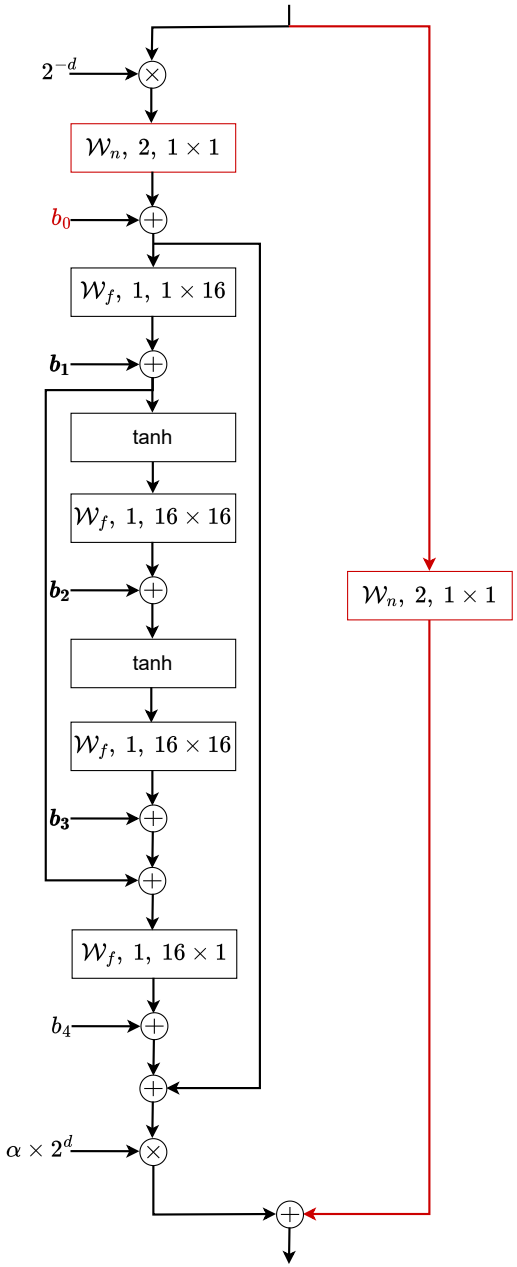


Experiment 6: Variant B (iWave++)

2D Non-Separable Lifting Architecture	**Predict/Update Modules**	**No. of learnable coefficients**	**Rate (bpp)**		
			CLIC.pro	**DIV2K**	**Kodak**
Proposed	**Proposed**	**264**	**7.70**	**7.89**	**8.70**
Generic	Variant B	118920	7.82	8.10	8.82

Experiment 3: Variant C





Experiment 3: Variant C

2D Non-Separable Lifting Architecture	**Predict/Update Modules**	**No. of learnable coefficients**	**Rate (bpp)**		
			CLIC.pro	**DIV2K**	**Kodak**
Proposed	**Proposed**	**264**	**7.70**	**7.89**	**8.70**
Generic	Variant C	119136	9.23	9.18	9.47

Conclusions of developed work

- Results highlight **superior compression performance** compared to state-of-the-art end-to-end wavelet-based frameworks
- **150-fold reduction (39560 to 264) in learnable coefficients**

Conclusions of developed work

- Results highlight **superior compression performance** compared to state-of-the-art end-to-end wavelet-based frameworks
- **150-fold reduction (39560 to 264) in learnable coefficients**

The deeper understanding and application of signal processing principles in the design of learning-based systems can lead to simpler and more effective solutions

Thank you!

joao.parracho@smt.ufrj.br