



Learnable Discrete Wavelet Transforms for 3D Gaussian Splatting Frequency Modulation

Hung Nguyen, An Le, Truong Nguyen
Video Processing Lab, UC San Diego
{hun004, d0le, tqn001}@ucsd.edu



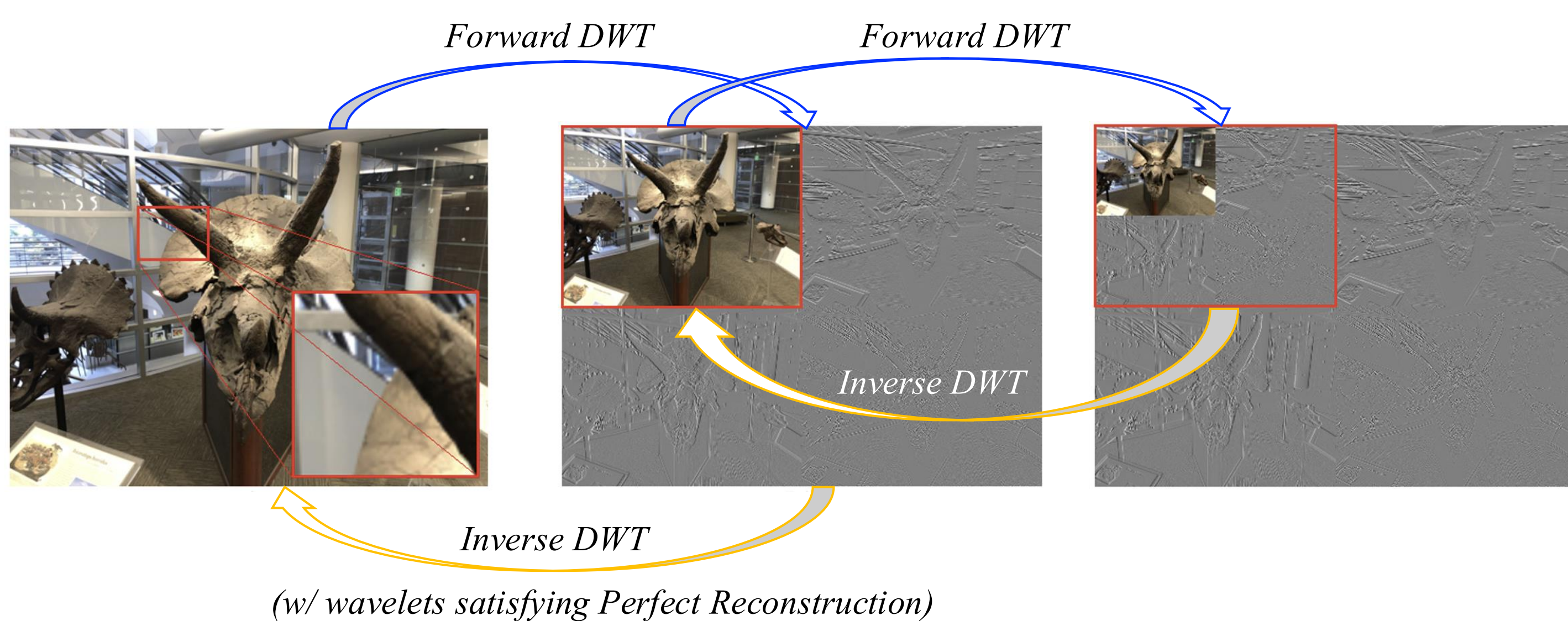
1. Motivation

- 3D Gaussian Splatting (3DGS) represents 3D scenes with optimizable 3D Gaussians, reconstructed from a collection of multi-view 2D images.
- However, Gaussian counts often grow uncontrollably during training. Therefore, reducing them while preserving rendering quality is crucial to lower GPU memory and storage costs.
- We introduce a framework for automatically optimizing 3DGS Gaussian counts by modulating the frequencies of the input images with the multi-level learnable Discrete Wavelet Transforms (DWTs).

2. Discrete Wavelet Transform (DWT)

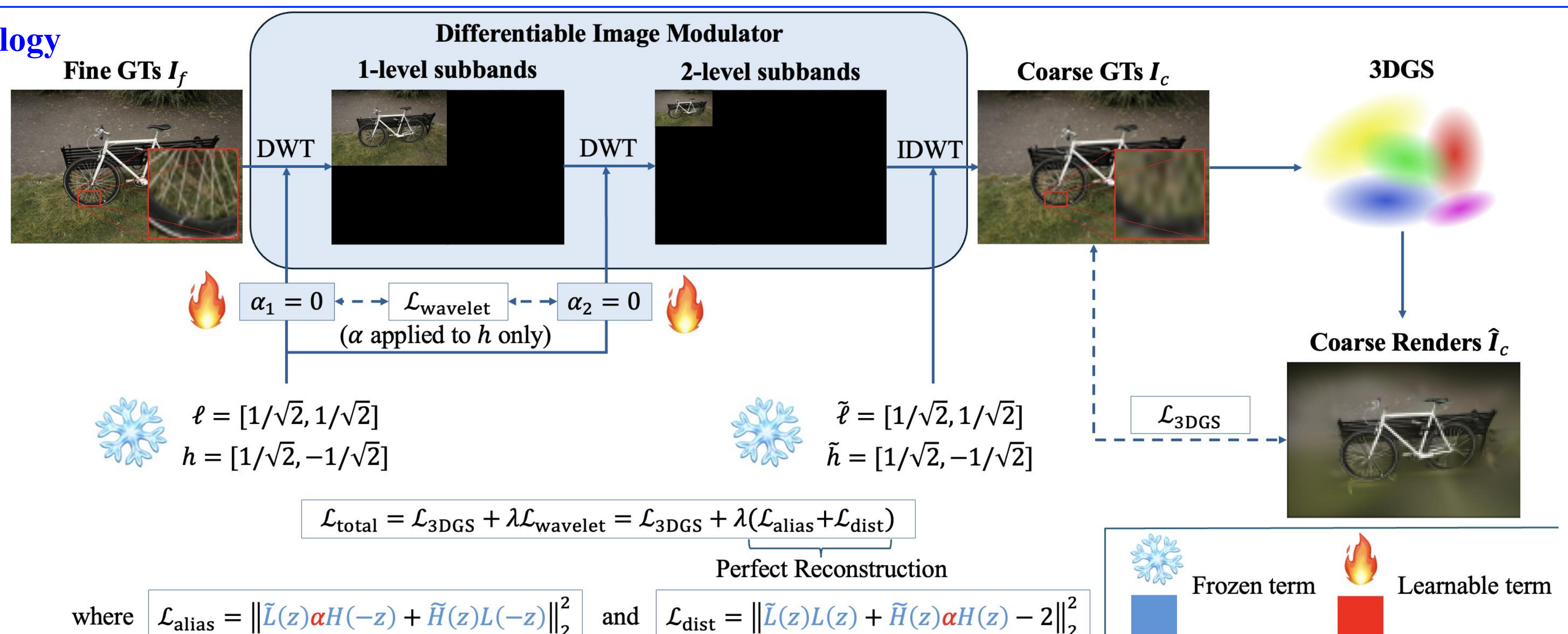
- The Forward DWT decomposes input image $\mathbf{X}$ into “subbands” of varying frequency levels:
 $\mathbf{X}_{LL} = \mathbf{LXL}^T, \mathbf{X}_{LH} = \mathbf{HXL}^T, \mathbf{X}_{HL} = \mathbf{LXH}^T, \mathbf{X}_{HH} = \mathbf{HXH}^T$
where $\mathbf{L}$ and $\mathbf{H}$ are Toeplitz matrices constructed from 1D low-pass and high-pass analysis filters ℓ and h , respectively.
- The Inverse DWT reconstructs the original image using synthesis filters $\tilde{\ell}$ and $\tilde{h}$. For the Haar wavelet:

$$\ell = \tilde{\ell} = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right], \quad h = \tilde{h} = \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right]$$



- The above diagram depicts the multi-level DWT, where the Forward DWT is recursively applied to the LL subband, and the Inverse DWT reconstructs the original image through successive synthesis stages.

3. Methodology



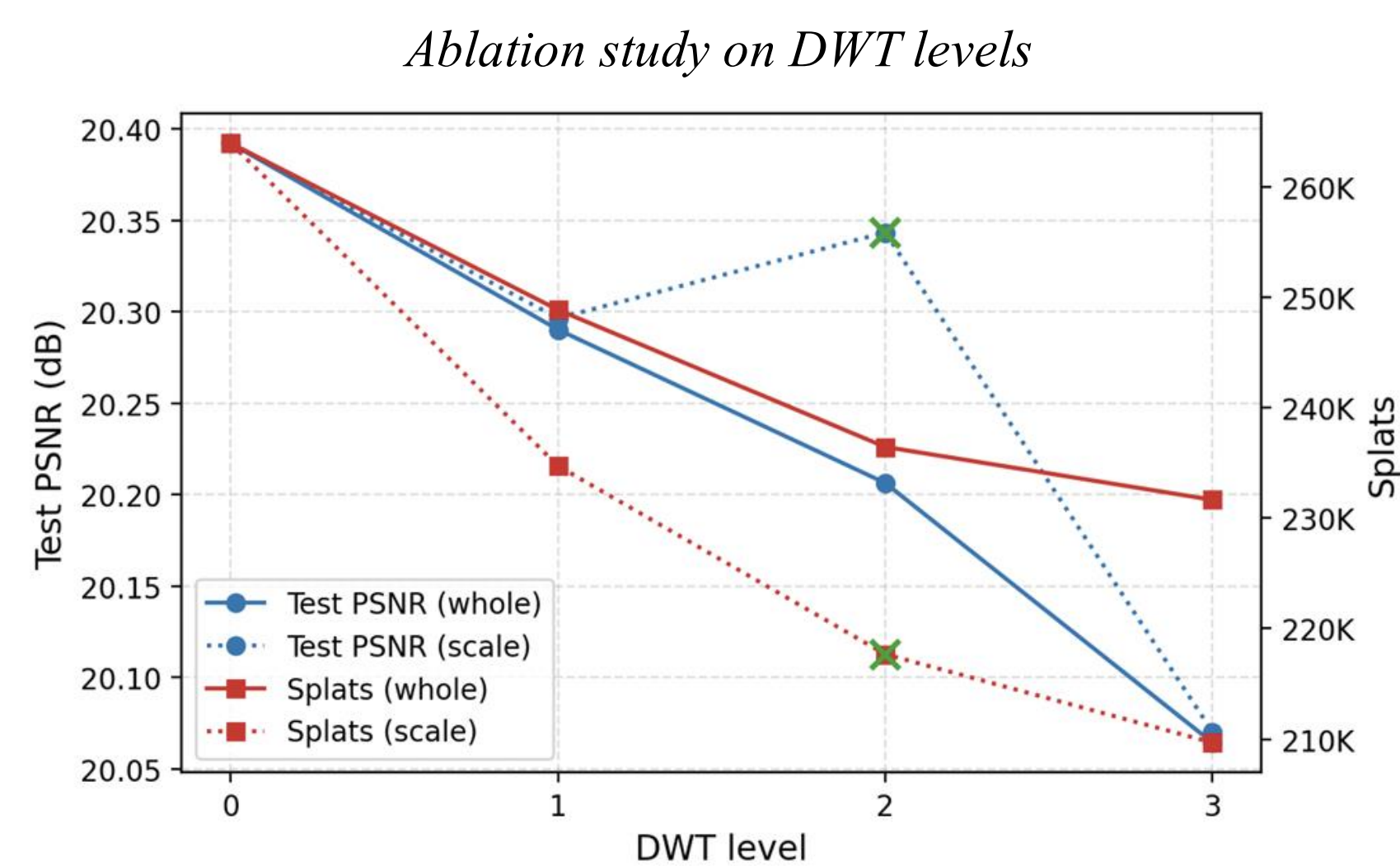
- Our **frequency modulation** framework progressively introduces high-frequency (HF) details into input images as training progresses. This reduces unnecessary Gaussian proliferation by reconstructing HF details only when needed.

6. References

- [1] Farooq et al. (2025). Optimized 3D Gaussian Splatting using Coarse-to-Fine Image Frequency Modulation. CVMP.
- [2] Nguyen et al. (2025). From Coarse to Fine: Learnable Discrete Wavelet Transforms for Efficient 3D Gaussian Splatting. ICCV workshop.
- [3] Mildenhall et al. (2019). Local Light Field Fusion: Practical View Synthesis with Prescriptive Sampling Guidelines. TOG.
- [4] Barron et al. (2022). Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. CVPR.

5. Quantitative Results

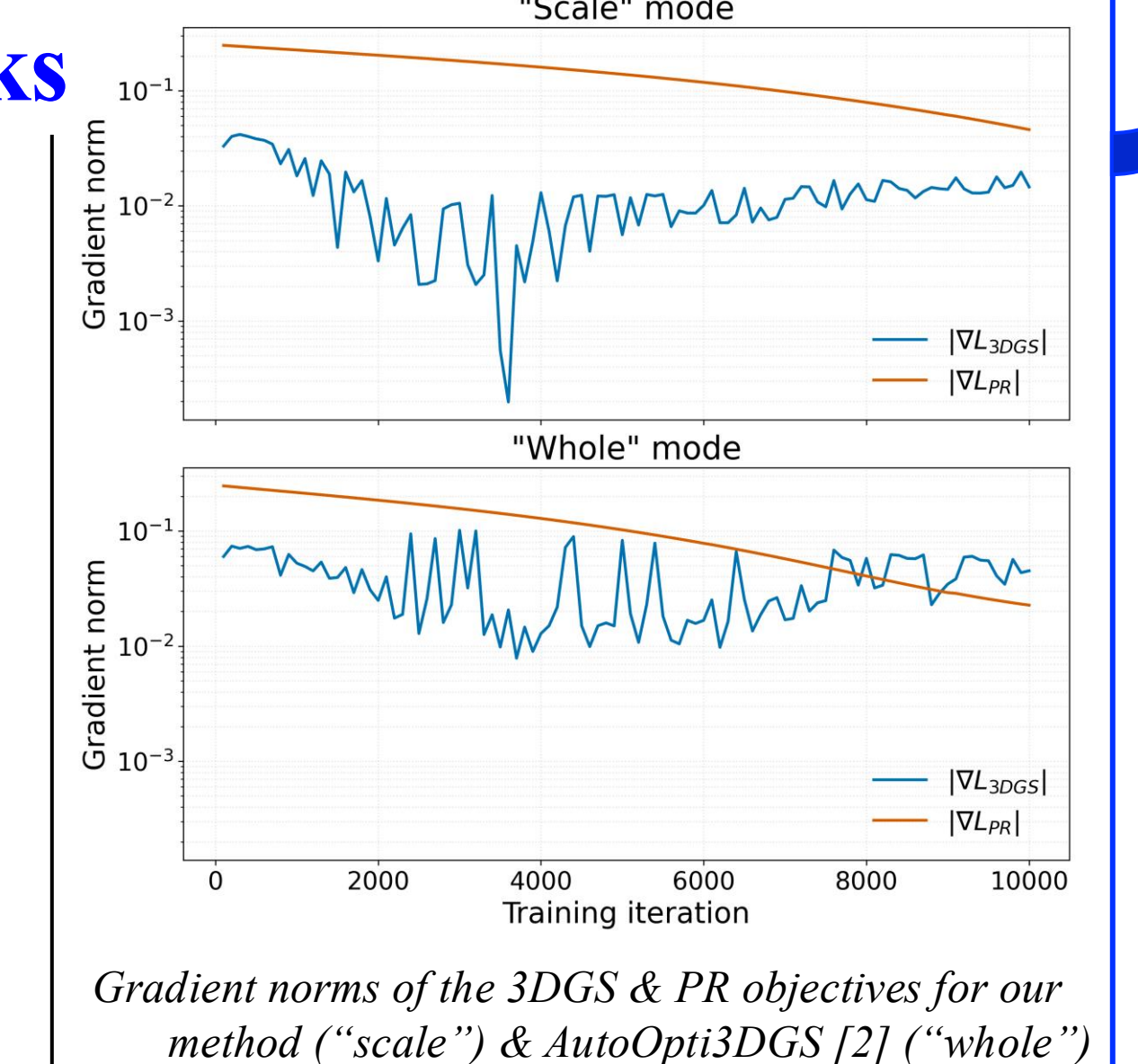
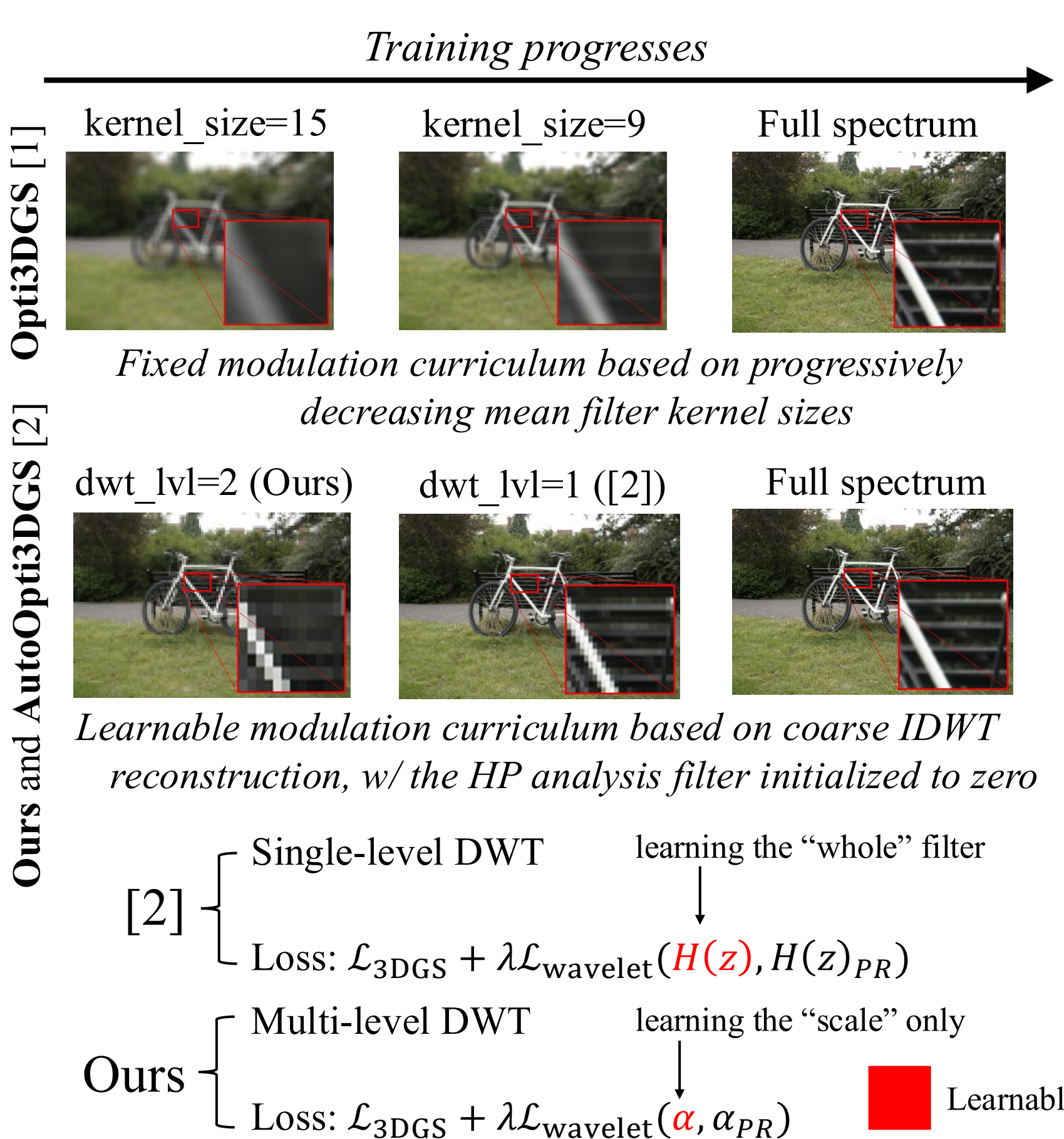
	3-view LLFF dataset [3]			12-view Mip-NeRF 360 dataset [4]		
	PSNR	SSIM	#G	PSNR	SSIM	#G
Opti3DGS (fixed mean filter)	19.59	0.660	247K	19.19	0.552	636K
AutoOpti3DGS (learnable single-level DWT)	20.29	0.703	249K	19.24	0.541	615K
Ours (learnable two-level DWT)	20.34	0.687	218K	19.29	0.560	589K
Vanilla 3DGS	20.40	0.706	272K	19.30	0.564	701K



Takeaways:

- Gaussian counts decrease with higher DWT levels.
- Learning only the scaling parameter produces fewer Gaussians than learning the whole high-pass filter.
- With the 2-level DWT, our method balances Gaussian optimization and image quality.

4. Comparison with Previous Works



Gradient norms of the 3DGS & PR objectives for our method ("scale") & AutoOpti3DGS [2] ("whole")

Takeaways:

- Our method constrains wavelet optimization to a single gradient direction, reducing conflict between the wavelet and reconstruction objectives and thereby reducing unnecessary Gaussian creation.

- The initial images are obtained through the Inverse DWT, with the analysis high-pass (HP) filter initialized as zero.
- A wavelet objective encourages the HP filter to satisfy the Perfect Reconstruction property, thus recovering HF details.