Fast Feedforward 3D Gaussian Splatting Compression

ICLR 2025

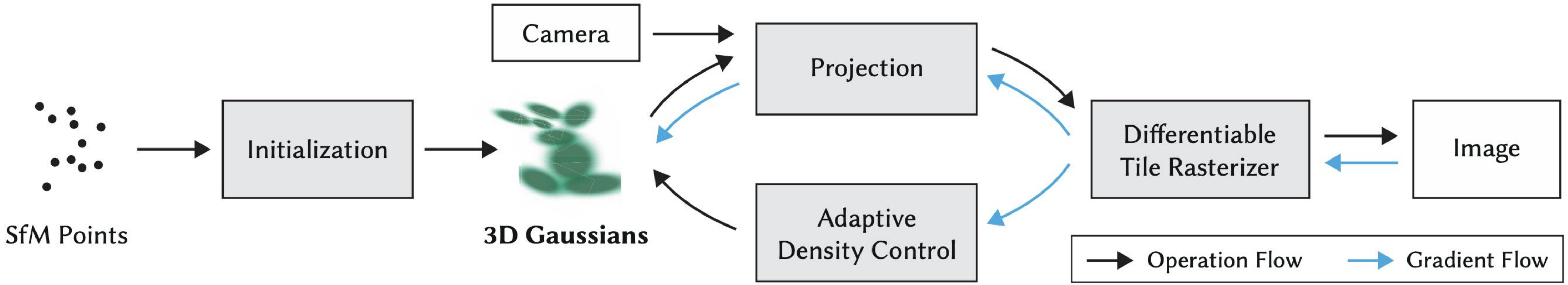
Yihang Chen[1,2], Qianyi Wu[2], Mengyao Li[3,2], Weiyao Lin[1], Mehrtash Harandi[2], Jianfei Cai[2]

[1]Shanghai Jiao Tong University, [2]Monash University, [3]Shanghai University

{yhchen.ee, wylin}@sjtu.edu.cn, {qianyi.wu, mehrtash.harandi, jianfei.cai}@monash.edu, sdlmy@shu.edu.cn

Code: https://github.com/yihangchen-ee/fcgs/
Page: https://yihangchen-ee.github.io/project_fcgs/

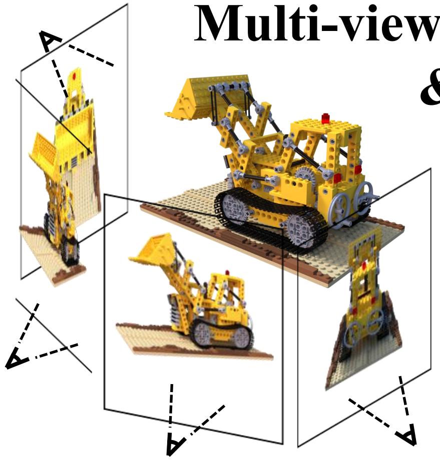
Background: 3D Gaussian Splatting (3DGS)



1. 3DGS represents scenes using **explicit** learnable attributed Gaussian points.
2. Millions of Gaussians in the scene leads to **LARGE** storage size!!

Kerbl B, Kopanas G, Leimkühler T, et al. 3D Gaussian Splatting for Real-Time Radiance Field Rendering[J]. ACM Trans. Graph., 2023, 42(4): 139:1-139:14.

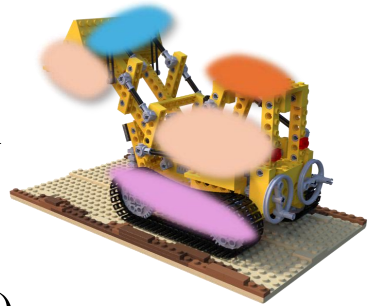
Multi-view images & Views



Generate 3DGS via:

1. Optimization (*Minutes*) (*e.g.*, vanilla 3DGS [1])
2. Feed-forward models (*Seconds*) (*e.g.*, MVSplat[2], LGM [3])

3DGS

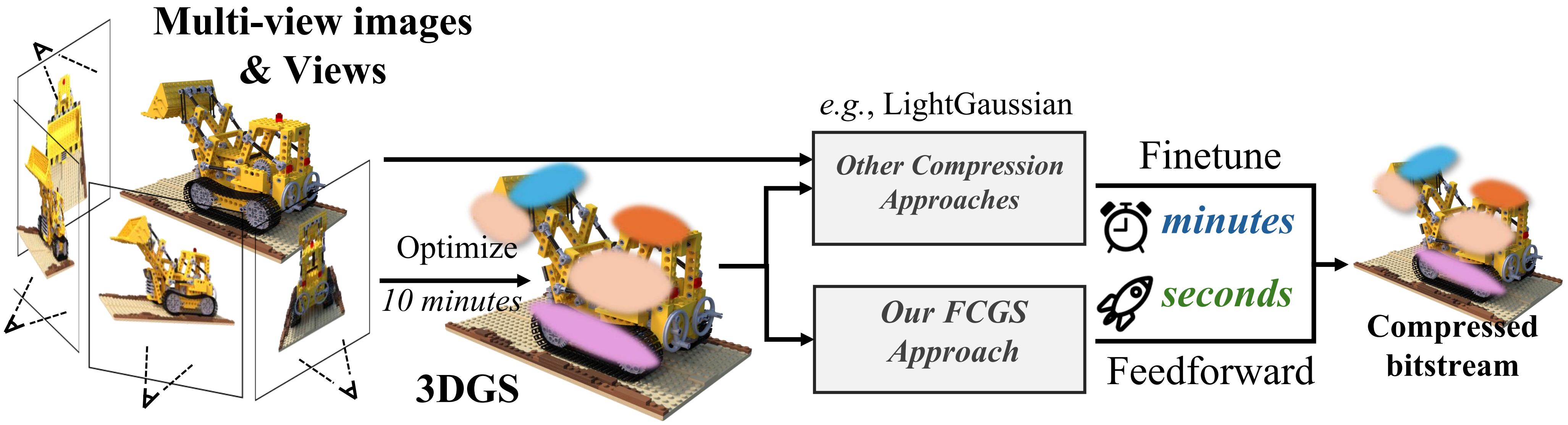


How to Compress it?

Compressed bitstreams

1. Pruning, Vector quantization …
2. Entropy constraints, Context model …

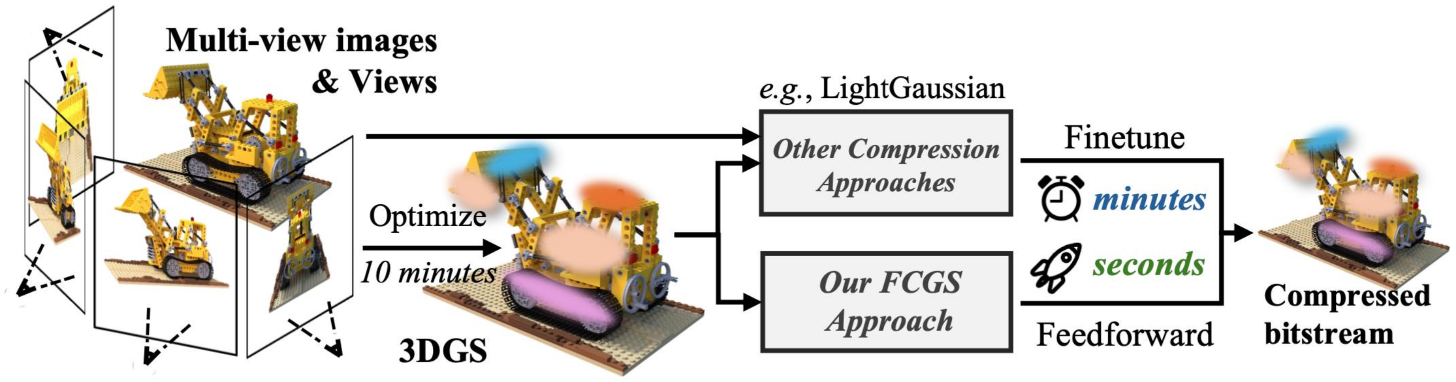
[1] Kerbl B, Kopanas G, Leimkühler T, et al. 3D Gaussian Splatting for Real-Time Radiance Field Rendering[J]. ACM Trans. Graph., 2023, 42(4): 139:1-139:14.
[2] Chen Y, Xu H, Zheng C, et al. Mvsplat: Efficient 3d gaussian splatting from sparse multi-view images[C]//ECCV 2024: 370-386.
[3] Tang J, Chen Z, Chen X, et al. Lgm: Large multi-view gaussian model for high-resolution 3d content creation[C]//ECCV 2024: 1-18.

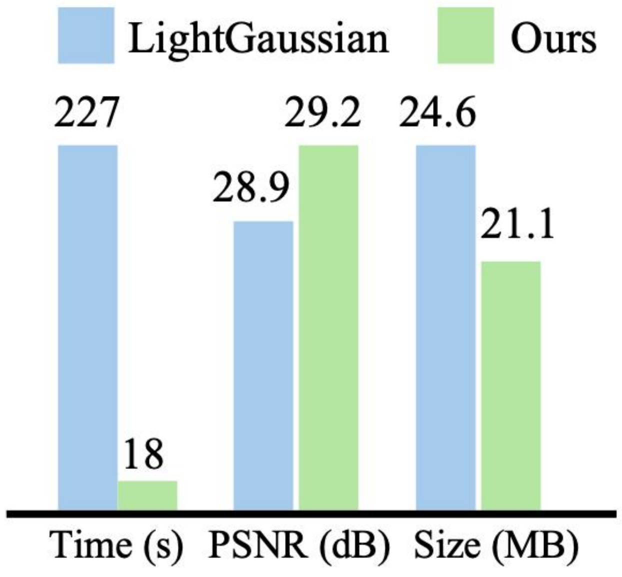


- **Per-scene optimization pipeline**.
  Pros: Per-scene adaptation for better RD performance.
  Cons: Slow, computationally expensive, need GT.

- **Generalizable optimization-free pipeline**.
  Pros: Fast, low computation cost, hassle-free.
  Cons: Limitation in RD performance.

Fan Z, Wang K, Wen K, et al. Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200+ fps[C]//NIPS, 2024





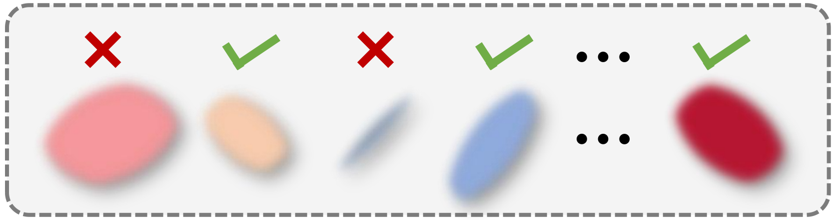
DL3DV-GS dataset

We have still achieved excellent compression performance !

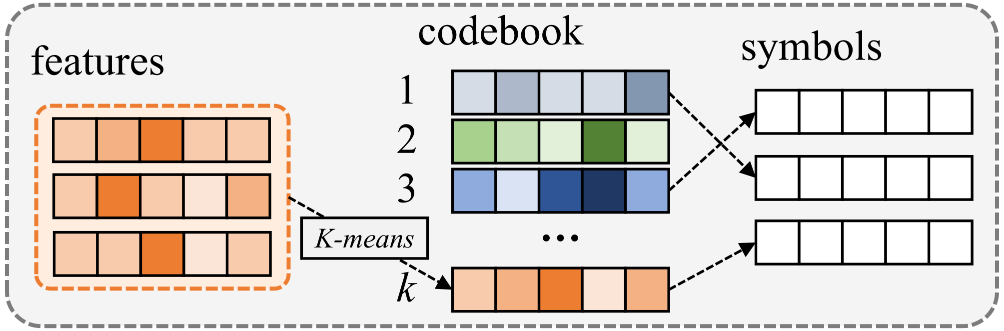
- **Per-scene optimization pipeline**.
  Pros: Per-scene adaptation for better RD performance.
  Cons: Slow, computationally expensive, need GT.
- **Generalizable optimization-free pipeline**.
  Pros: Fast, low computation cost, hassle-free.
  Cons: Limitation in RD performance.

Fan Z, Wang K, Wen K, et al. Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200+ fps[C]//NIPS, 2024

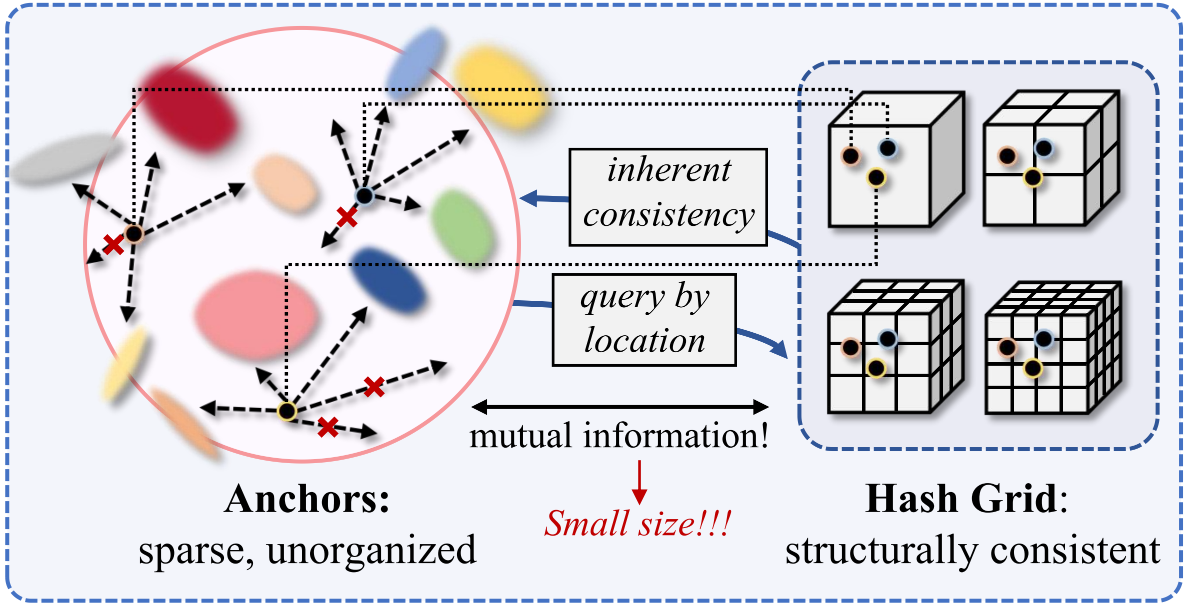
- **How to achieve compression without per-scene optimization?**



1) Pruning



2) Codebook + Vector Quantization



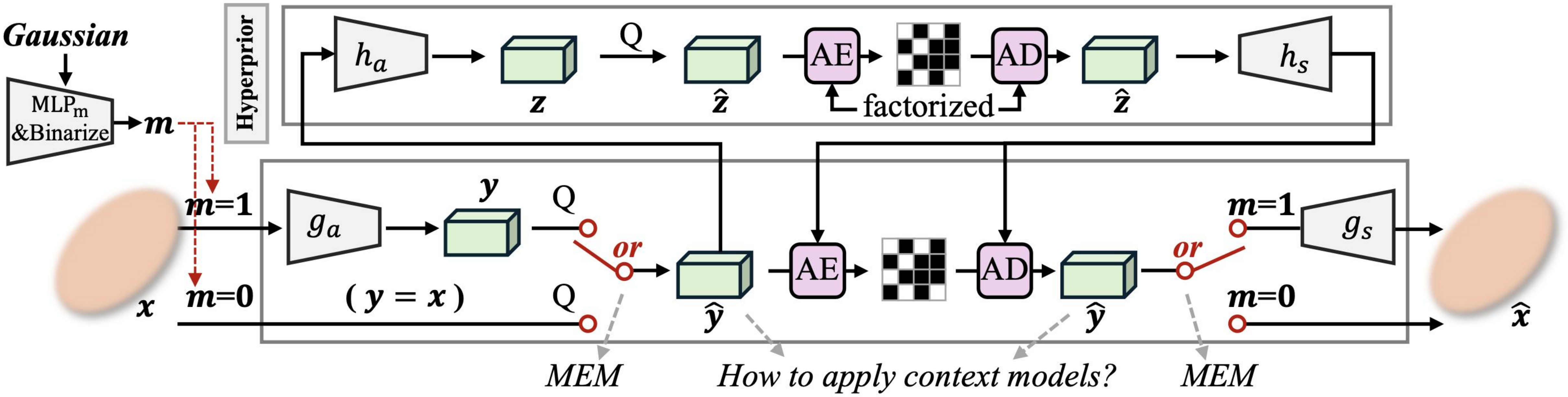
HAC-like context models

Pruning & VQ may alter the original values

The hash grids need optimization to obtain

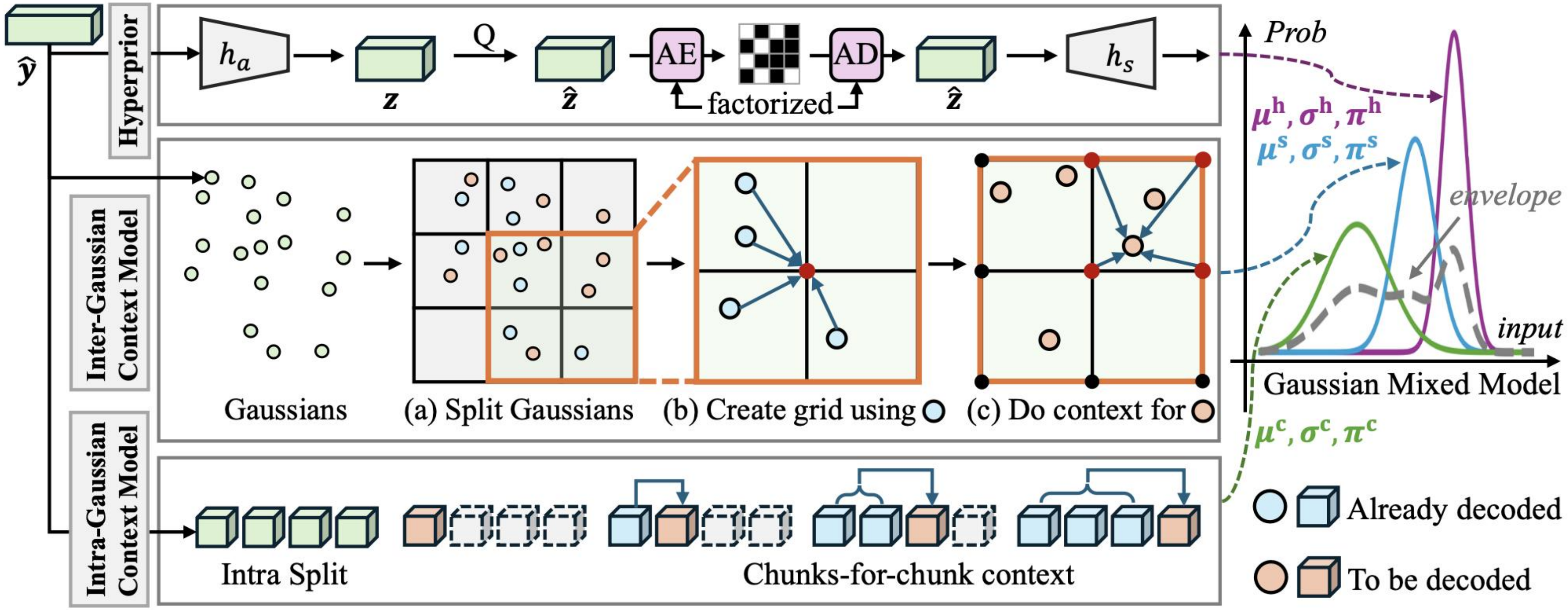
They are not suitable for the optimization-free pipeline

- **Solution: An autoencoder-based network architecture with context models**



1. A **Multi-path Entropy Module (MEM)** for adaptive balance of fidelity and size of each Gaussian.
   1) For **geometry attributes**, they are all in the m=0 path to maintain fidelity.
   2) For **color attributes**, m can be either 0 or 1 to either maintain fidelity or eliminate redundancies.

- **Solution: An autoencoder-based network architecture with context models**



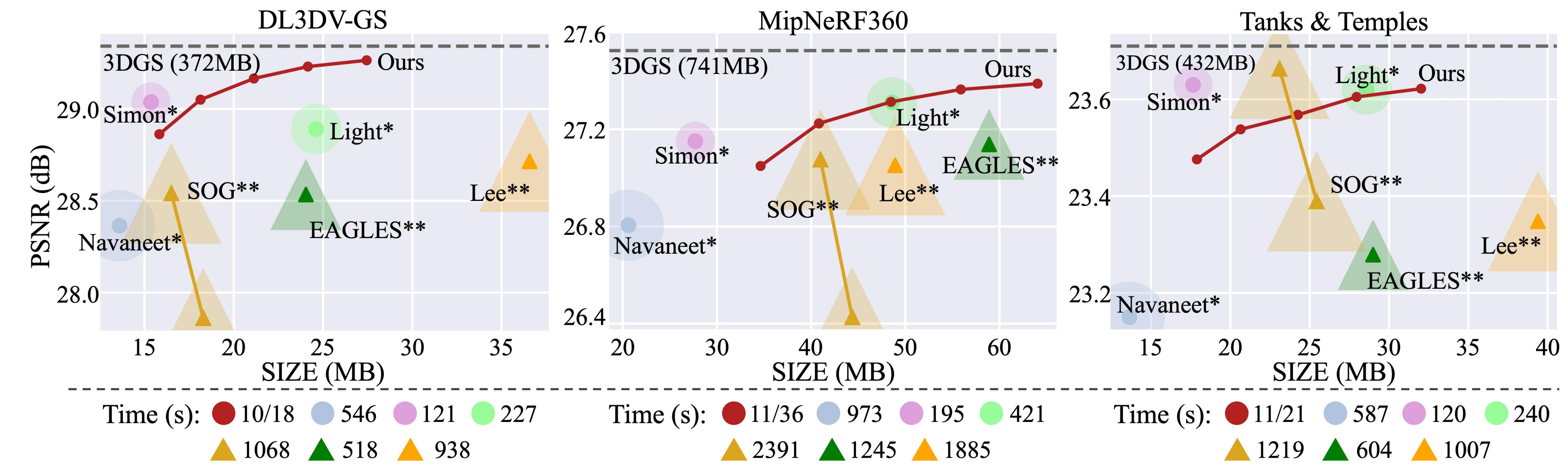
2. **Inter- and intra-Gaussian context models** to build relations among Gaussians for entropy modeling.

   1) **Inter-Gaussian context models** can *create* grids for context, instead of learning hash grids.

   2) **Intra-Gaussian context models** construct contextual relations along channels.

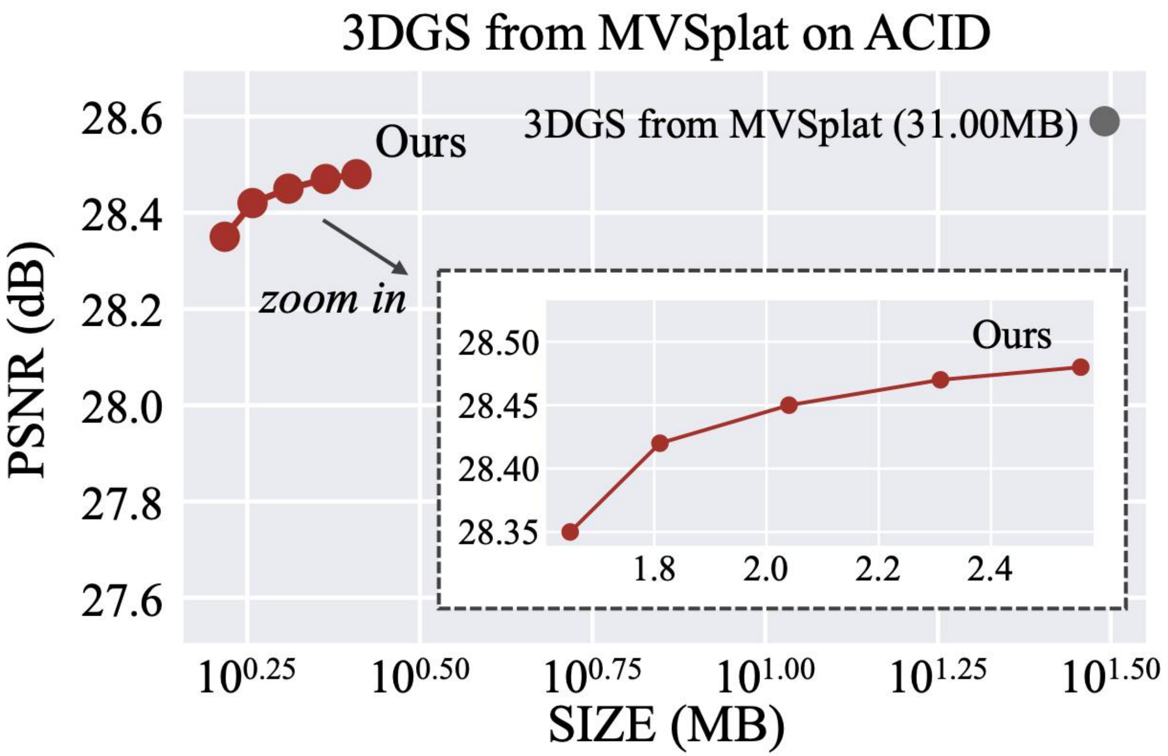
   3) **Gaussian Mixed Model (GMM)** is utilized to fuse the three sets of Gaussian probabilities.

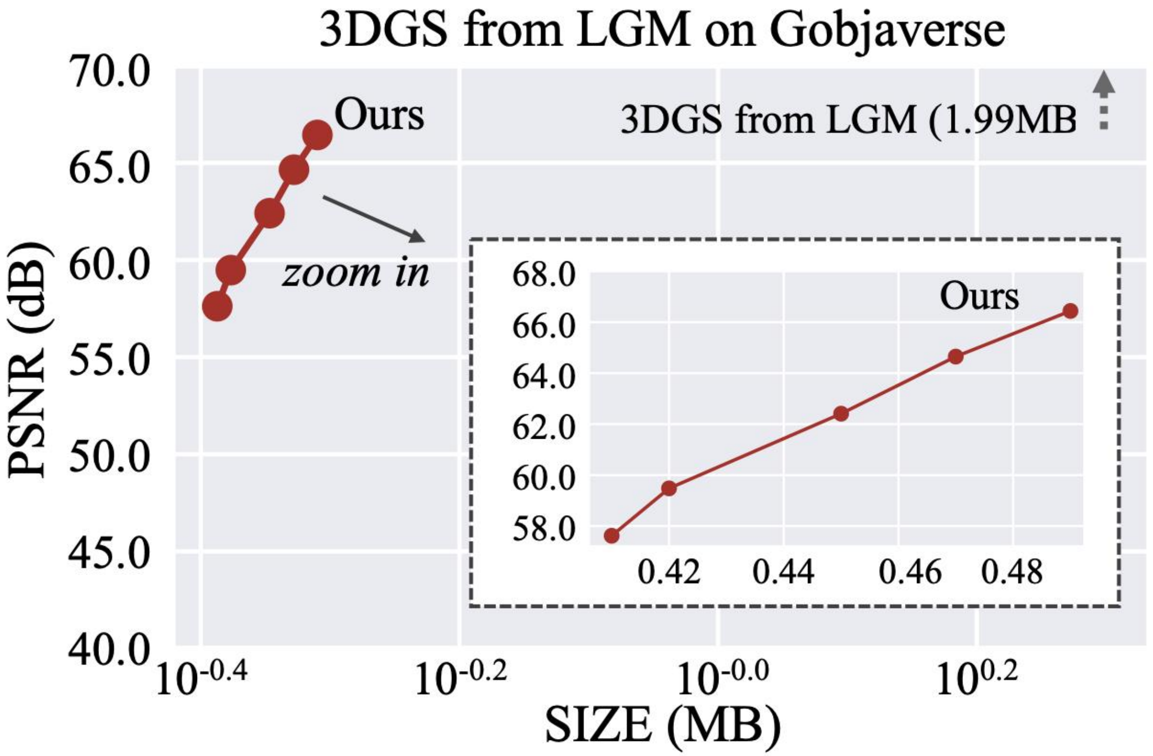
- **Performance: Compress 3DGS from optimization**



1. Although unable to do per-scene optimization, FCGS still **surpasses most optimization-based approaches**.
2. Much **faster runtime** to achieve compression, from *minutes* to *seconds*.

- **Performance: Compress 3DGS from feed-forward models**

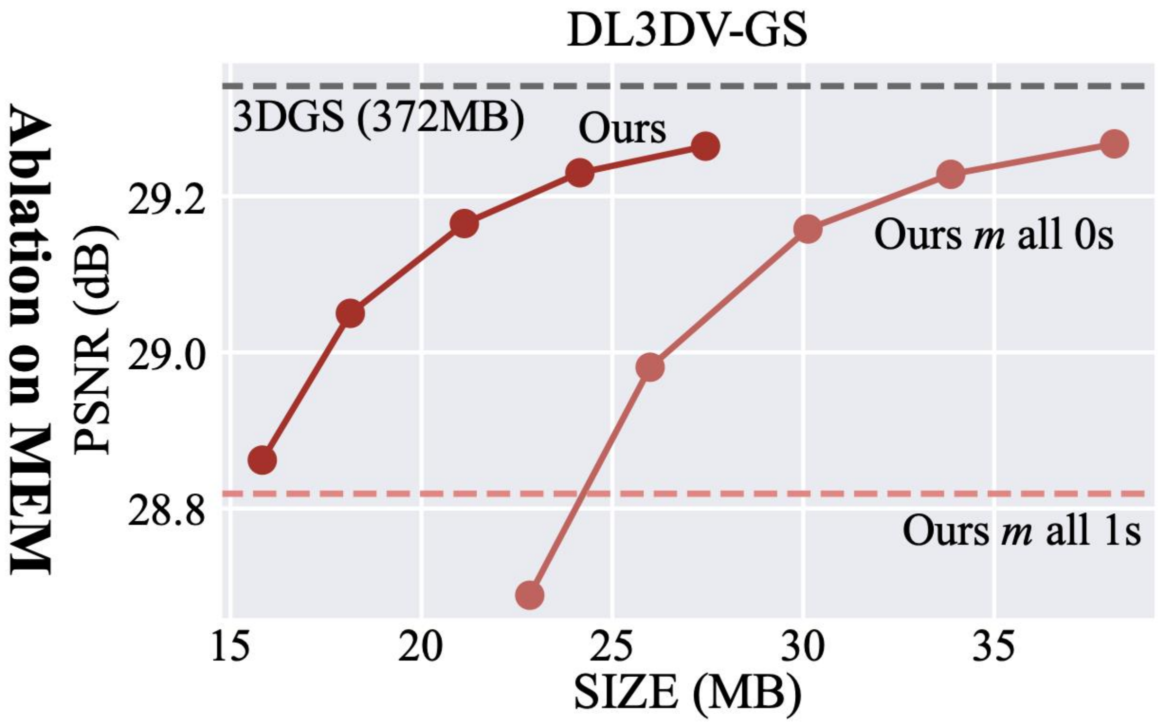


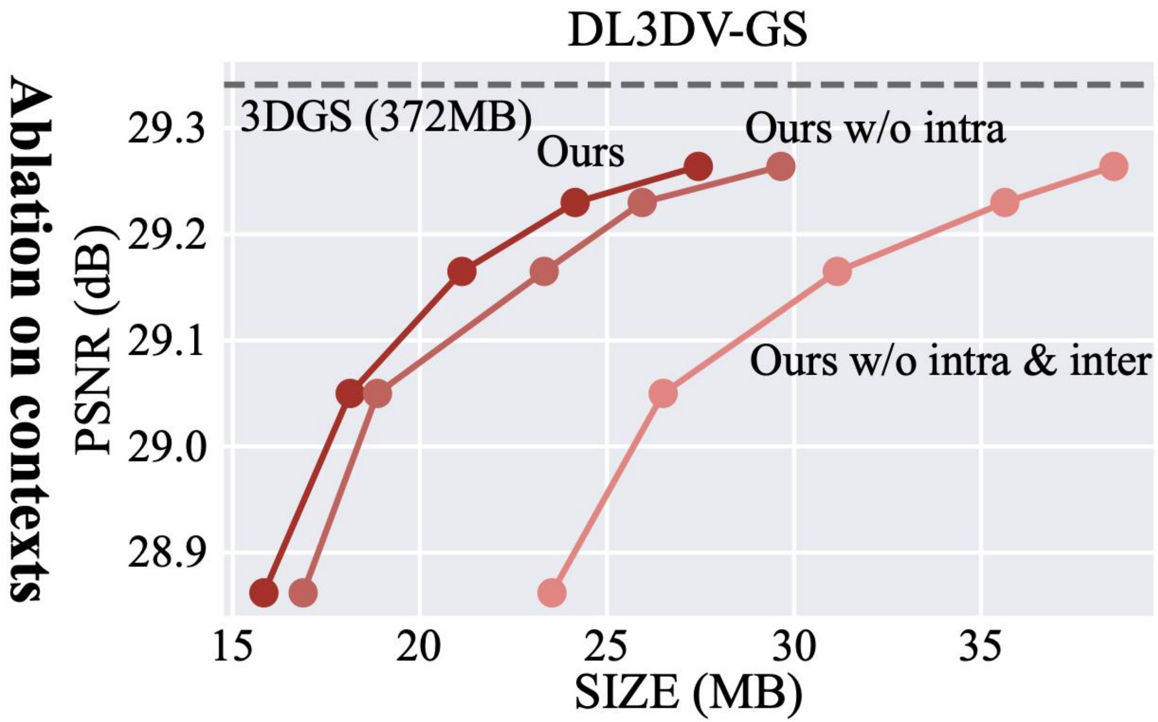


FCGS can also **compress 3DGS from feed-forward models** such as MVSplat [1] and LGM [2].

[1] Chen Y, Xu H, Zheng C, et al. Mvsplat: Efficient 3d gaussian splatting from sparse multi-view images[C]//ECCV 2024: 370-386.
[2] Tang J, Chen Z, Chen X, et al. Lgm: Large multi-view gaussian model for high-resolution 3d content creation[C]//ECCV 2024: 1-18.

- **Ablation study**





Both the **MEM** module and **context models** are demonstrated to be effective.

Fast Feedforward 3D Gaussian Splatting Compression

- **Qualitative results**

Simon* | SOG** | Ours ($\lambda = 1e-4$) | Ours ($\lambda = 16e-4$) | 3DGS | Ground Truth

Thank you for your kind attention!