IS GAUSSIAN ENOUGH? ON THE LOCAL SPECIALIZATION OF KERNELS IN 3D SPLATTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent 3D splatting methods have extended Gaussian primitives toward increasingly expressive and adaptive kernel formulations. However, it remains unclear whether greater within-family adaptability makes kernel families effectively interchangeable or whether they retain distinct local representation biases. Existing studies primarily rely on scene-level reconstruction metrics, potentially obscuring heterogeneous behavior within a scene. We systematically characterize local cross-family specialization across Gaussian, generalized exponential, and deformable radial kernels using real-scene analysis, structure-conditioned characterization, controlled synthetic manipulations, and robustness controls. We find substantial local complementarity: different families achieve lower reconstruction error in different regions, with relative performance varying systematically with local scene structure. Controlled structural manipulations produce reproducible changes in cross-family performance, while specialization persists across spatial scales and under a matched primitive budget. These results show that within-family adaptability does not necessarily eliminate cross-family representation differences, motivating representations that account for compatibility between kernel families and local scene structure.

1 INTRODUCTION

The introduction of 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) marked a significant shift in neural scene representation, providing an explicit alternative to predominantly implicit neural radiance fields (Mildenhall et al., 2021). By representing scenes with anisotropic Gaussian primitives and combining differentiable rasterization with adaptive densification, 3DGS enables high-fidelity reconstruction, fast optimization, and real-time rendering. However, its prescribed Gaussian functional form imposes a characteristic spatial profile that can limit the representation of complex geometry and high-frequency structure. Recent work has therefore explored more expressive kernels. Hamdi et al. (2024) introduced Generalized Exponential Splatting (GES), which learns a shape parameter that adapts kernel falloff, while Huang et al. (2025) proposed Deformable Radial Kernel (DRK) Splatting, which employs direction-dependent radial deformation to model irregular local structure. These developments expand splatting from a predominantly Gaussian formulation toward increasingly expressive and adaptive kernel families, alongside broader work on alternative kernel formulations and structure-aware splatting (Chen et al., 2024; Held et al., 2025; Pramuditha et al., 2026; Guédon & Lepetit, 2024; Shao et al., 2025). A broader discussion of point-based rendering, alternative kernels, and structure-aware representations is provided in Appendix A.

Despite this progress, kernel families are predominantly evaluated as scene-wide representations using aggregate metrics such as peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and learned perceptual image patch similarity (LPIPS) (Chen & Wang, 2026; Liu et al., 2025). Such comparisons identify which representation performs better on average but can obscure spatially heterogeneous reconstruction behavior across smooth surfaces, sharp boundaries, fine textures, thin structures, and irregular geometry (Lavoué et al., 2015; Zhang et al., 2024; Kong et al., 2025). Because kernel families impose distinct spatial profiles and parameterizations (Liu et al., 2025), global superiority need not imply uniform local superiority. More fundamentally, it remains unclear whether increasing within-family adaptability overcomes these representation biases or whether different kernel families retain systematic advantages under different local structural conditions. Ad-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

dressing this question requires moving beyond global rankings toward characterizing *where* and *under what structural conditions* different kernel families succeed or fail.

We investigate this problem through four research questions:

- **RQ1:** Do adaptive kernel families exhibit local specialization within the same real-world scene?
- **RQ2:** Is local kernel-family preference systematically associated with scene structure?
- **RQ3:** Do controlled structural manipulations systematically alter relative cross-family performance?
- **RQ4:** Does specialization persist under matched primitive budgets and different local-analysis settings?

We study 3DGS (Kerbl et al., 2023), GES (Hamdi et al., 2024), and DRK Splatting (Huang et al., 2025) using local real-scene analysis, structure-conditioned characterization, controlled synthetic experiments, and robustness controls.

Our contributions are threefold. **First**, we systematically characterize local cross-family specialization, showing substantial complementarity between independently optimized adaptive kernel families within the same scenes that is obscured by scene-level metrics. **Second**, we show that relative cross-family performance varies systematically with local scene structure and that controlled structural manipulations produce reproducible changes in these differences. **Third**, we show that specialization persists across spatial analysis scales, tie criteria, and a matched primitive budget, indicating that unequal primitive counts alone are insufficient to explain the observed complementarity.

2 Characterizing Kernel Specialization

2.1 Adaptive Splatting Kernel Families

We consider splatting representations as collections of spatial primitives $\mathcal{P} = \{P_i\}_{i=1}^{N}$, where N is the number of primitives and each primitive P_i has center $\boldsymbol{\mu}_i$, opacity o_i, appearance $\mathbf{c}_i$, and family-specific spatial parameters $\boldsymbol{\theta}_i$. Its spatial contribution is determined by a kernel $K_i(\mathbf{x}; \boldsymbol{\mu}_i, \boldsymbol{\theta}_i)$ evaluated at spatial location $\mathbf{x}$. The representations studied here share this principle but differ in the functional form and adaptability of K_i. Figure 2 illustrates the characteristic spatial profiles and shape adaptability of the three kernel families.

Gaussian and generalized exponential kernels. 3DGS (Kerbl et al., 2023) uses an anisotropic Gaussian,

$$K_i^{\mathrm{G}}(\mathbf{x}) = \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right], \tag{1}$$

where $\boldsymbol{\Sigma}_i$ is the covariance matrix controlling the orientation and anisotropic extent of primitive i. GES (Hamdi et al., 2024) additionally learns a shape parameter β_i controlling kernel decay. For scalar position x and scalar center μ_i, its one-dimensional generalized exponential form is

$$f_{\beta_i}(x) = A_i \exp\left[-\left(\frac{|x-\mu_i|}{a_i}\right)^{\beta_i}\right], \tag{2}$$

where A_i is the amplitude, $a_i > 0$ controls spatial scale, and $\beta_i > 0$ controls the decay shape. The Gaussian is recovered for $\beta_i = 2$ under the corresponding parameterization (Hamdi et al., 2024). Thus, GES expands the admissible decay profiles while retaining Gaussian behavior as a special case.

Deformable radial kernel. DRK Splatting (Huang et al., 2025) replaces the elliptical Gaussian footprint with a learnable radial representation whose direction-dependent radial bases deform the primitive support and control local boundary shape. The formulation contains the 2D Gaussian as a special configuration of its radial bases. Its full radial parameterization is given in Appendix B.1.

We use within-family adaptation to denote optimization of primitive parameters while retaining the underlying kernel family. Under this definition, 3DGS adapts anisotropic Gaussian parameters,

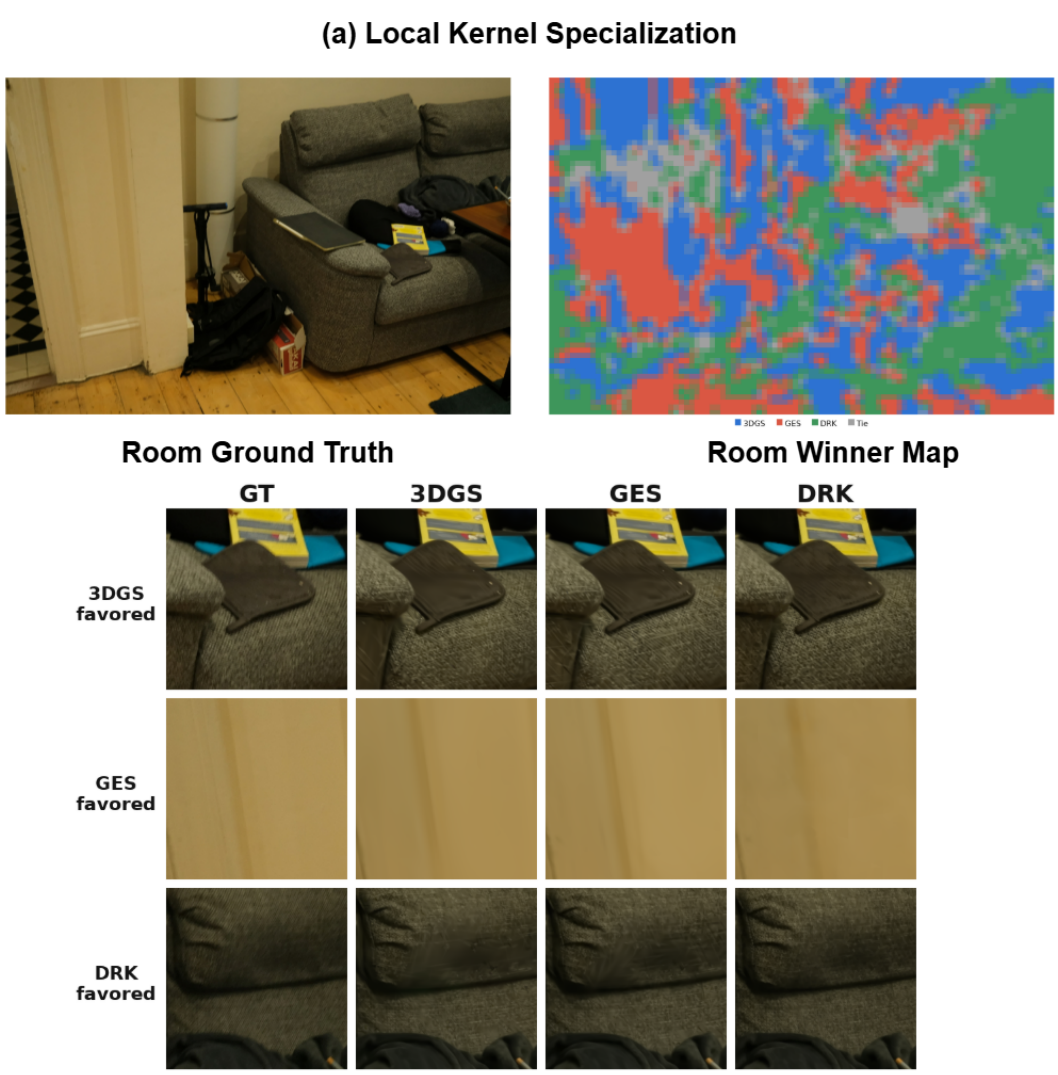


Figure 1: **Local cross-family kernel specialization on the Room scene.** Top: ground-truth view and corresponding local winner map. Bottom: representative regions locally favored by 3DGS, GES, and DRK, shown alongside the ground truth and reconstructions from all three kernel families. The examples illustrate that different kernel families attain lower reconstruction error in different regions of the same scene.

GES additionally adapts kernel decay, and DRK learns deformable radial support. We use *cross-family variation* for differences between these independently optimized representations. Because the methods also differ in parameterization and optimization, we treat them as complete adaptive kernel-family representations rather than attributing observed differences exclusively to kernel functional form.

Why cross-family specialization can persist. Consider a target signal f_p associated with local region p and a kernel family k with representable set $\mathcal{F}_k(B)$ where B denotes a fixed representation budget. Its minimum local approximation error is

$$\epsilon_k(p; B) = \inf_{\hat{f} \in \mathcal{F}_k(B)} \|f_p - \hat{f}\|_2^2. \tag{3}$$

Within-family adaptation seeks to minimize this error over the parameters available to that family, but does not remove constraints imposed by its kernel form. Gaussian kernels impose quadratic exponential decay, generalized exponential kernels additionally adapt the decay exponent, and DRKs

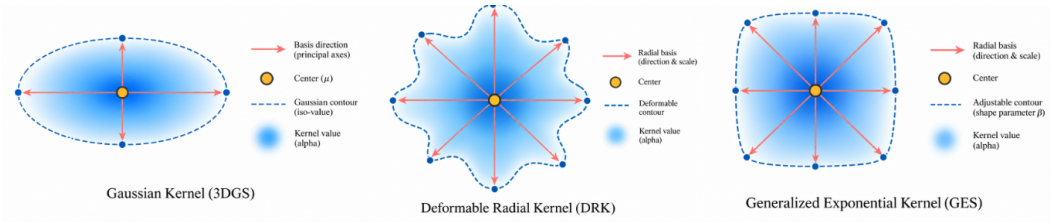


Figure 2: Comparison of the three kernel families considered in this study: Gaussian (3DGS), deformable radial kernel (DRK), and generalized exponential (GES).

permit direction-dependent support deformation. These differences induce distinct spatial and frequency responses and therefore different approximation biases for structures such as smooth regions, high-frequency detail, sharp boundaries, and curved geometry. Hence, for two families a and b, there can exist local regions p_1 and p_2 such that $\epsilon_a(p_1; B) < \epsilon_b(p_1; B)$ while $\epsilon_a(p_2; B) > \epsilon_b(p_2; B)$. Let $\mathcal{K} = \{\mathrm{G}, \mathrm{GE}, \mathrm{DRK}\}$ denote the evaluated kernel-family representations. The preferred family $k^*(p; B) = \arg\min_{k \in \mathcal{K}} \epsilon_k(p; B)$ can therefore vary with local structure, even when every family is internally optimized. Our experiments test this representation-bias hypothesis and whether such differences occur systematically in practice.

2.2 Local Cross-Family Specialization

We characterize cross-family differences at a local spatial scale rather than through scene-level rankings. For a held-out view v, let I_v denote the ground-truth image and $\hat{I}_v^{(k)}$ the rendering produced by family $k \in \mathcal{K}$. For a local region p containing $|p|$ pixels and an image with C color channels, we define the local mean-squared reconstruction error as

$$E_k(p) = \frac{1}{C|p|} \sum_{\mathbf{u} \in p} \left\| I_v(\mathbf{u}) - \hat{I}_v^{(k)}(\mathbf{u}) \right\|_2^2. \tag{4}$$

Unlike scene-level metrics, the collection of local errors preserves spatial variation in reconstruction performance within the same scene.

Pairwise relative error. For two families $a, b \in \mathcal{K}$, we define their local relative error as

$$\Delta E_{a,b}(p) = E_a(p) - E_b(p), \tag{5}$$

where negative values favor a and positive values favor b. This quantity directly measures the local performance gap between representations and serves as the response variable in our structure-conditioned analysis.

Local family preference. Let $E_{(1)}(p)$ and $E_{(2)}(p)$ denote the smallest and second-smallest local errors, respectively. To avoid interpreting negligible numerical differences as specialization, we introduce a nonnegative tolerance $\tau \geq 0$ and assign

$$S(p) = \begin{cases} \arg\min\limits_{k \in \mathcal{K}} E_k(p), & E_{(2)}(p) - E_{(1)}(p) > \tau, \\ \text{tie}, & E_{(2)}(p) - E_{(1)}(p) \leq \tau. \end{cases} \tag{6}$$

The resulting $S(p)$ forms the *winner map*; aggregating these assignments gives family winner and tie fractions. Exact definitions are provided in Appendix B.2. Sensitivity to τ and patch scale is evaluated separately in our robustness analysis.

Oracle complementarity. Winner fractions measure how often each family is locally preferred but not the magnitude of its advantage. We therefore define a post-hoc oracle with $E_{\mathrm{oracle}}(p) = \min_{k \in \mathcal{K}} E_k(p)$ and quantify its relative headroom over family k as

$$H_k = \frac{\bar{E}_k - \bar{E}_{\mathrm{oracle}}}{\bar{E}_k}, \tag{7}$$

where $\bar{E}_k$ and $\bar{E}_{\text{oracle}}$ denote the mean patch-wise errors of family k and the oracle, respectively, over the evaluated patches. The oracle is diagnostic only: it uses ground-truth reconstruction error and does not constitute a realizable inference-time selection mechanism. Detailed definitions are provided in Appendix B.2.

We define *local cross-family specialization* as a regime in which different independently optimized kernel-family representations attain the lowest reconstruction error over different regions of the same scene. We characterize it jointly through local winner variation, pairwise relative error, and oracle headroom. This definition is descriptive rather than causal: it identifies differences between complete kernel-family representations and does not attribute them exclusively to kernel functional form. Whether these local differences vary systematically with scene structure is examined next.

2.3 Structure-Conditioned Specialization

Local specialization alone does not establish that cross-family differences are systematically related to scene content. We therefore characterize each region by measurable structural properties and examine how relative cross-family reconstruction error varies with local structure.

Local structural descriptors. For each region p, structural measurements are computed exclusively from the corresponding ground-truth patch, independently of all kernel-family renderings. We use six descriptors, $\mathbf{s}(p) = [s_{\text{grad}}(p), s_{\text{edge}}(p), s_{\text{lap}}(p), s_{\text{var}}(p), s_{\text{hf}}(p), s_{\text{ent}}(p)]^\top \in \mathbb{R}^6$, measuring mean gradient magnitude, edge strength, Laplacian energy, local variance, high-frequency energy, and intensity entropy, respectively. These provide complementary measurements of local image structure rather than a single notion of complexity. Exact definitions are given in Appendix B.3.

Structure-conditioned relative error. For descriptor index $d \in \{1, \dots, 6\}$ and family pair $a, b \in \mathcal{K}$, we analyze the pairwise difference $\Delta E_{a,b}(p)$ from Eq. 5. Our primary quantity of interest is the conditional expectation $\mathbb{E}[\Delta E_{a,b}(p) \mid s_d(p)]$, which characterizes how relative cross-family performance varies with local structure. We quantify monotonic association using the Spearman rank correlation $\rho_d^{a,b} = \rho_{\text{S}}(s_d(p), \Delta E_{a,b}(p))$. For trend visualization, patches are additionally divided into ten approximately equal-sized quantiles according to s_d. For quantile $\mathcal{Q}_q$, $q \in \{1, \dots, 10\}$, we compute

$$\widetilde{\Delta E}_{a,b}^{(q)} = \text{median}\left\{\Delta E_{a,b}(p) : p \in \mathcal{Q}_q\right\}, \tag{8}$$

and fit a linear slope across normalized quantile positions as a descriptive summary of trend direction, without assuming an intrinsically linear patch-level relationship.

Uncertainty and cross-scene consistency. Because patches from the same held-out view are dependent, uncertainty is estimated using a view-level cluster bootstrap rather than treating patches as independent replications. We report 95% confidence intervals for binned trends and evaluate trend directions independently across scenes to assess cross-scene reproducibility. Multiple family–descriptor comparisons are controlled using the Benjamini–Hochberg procedure. Full statistical details and complementary winner-conditioned analyses are provided in Appendix B.4. These real-scene analyses establish association rather than causation because naturally occurring structural properties are correlated. Controlled synthetic experiments therefore separately test whether deliberate structural manipulations produce reproducible changes in relative cross-family performance.

3 Experimental Framework

3.1 Kernel-Family Implementations and Training Protocol

We evaluate independently optimized 3DGS (Kerbl et al., 2023), GES (Hamdi et al., 2024), and DRK (Huang et al., 2025) using their official implementations and native optimization procedures, retaining their respective 30k, 40k, and 35k training schedules. For real scenes, all methods initialize from the same COLMAP reconstruction and follow their native adaptive primitive-growth procedures, with primitive count unconstrained unless otherwise stated. Each held-out viewpoint is

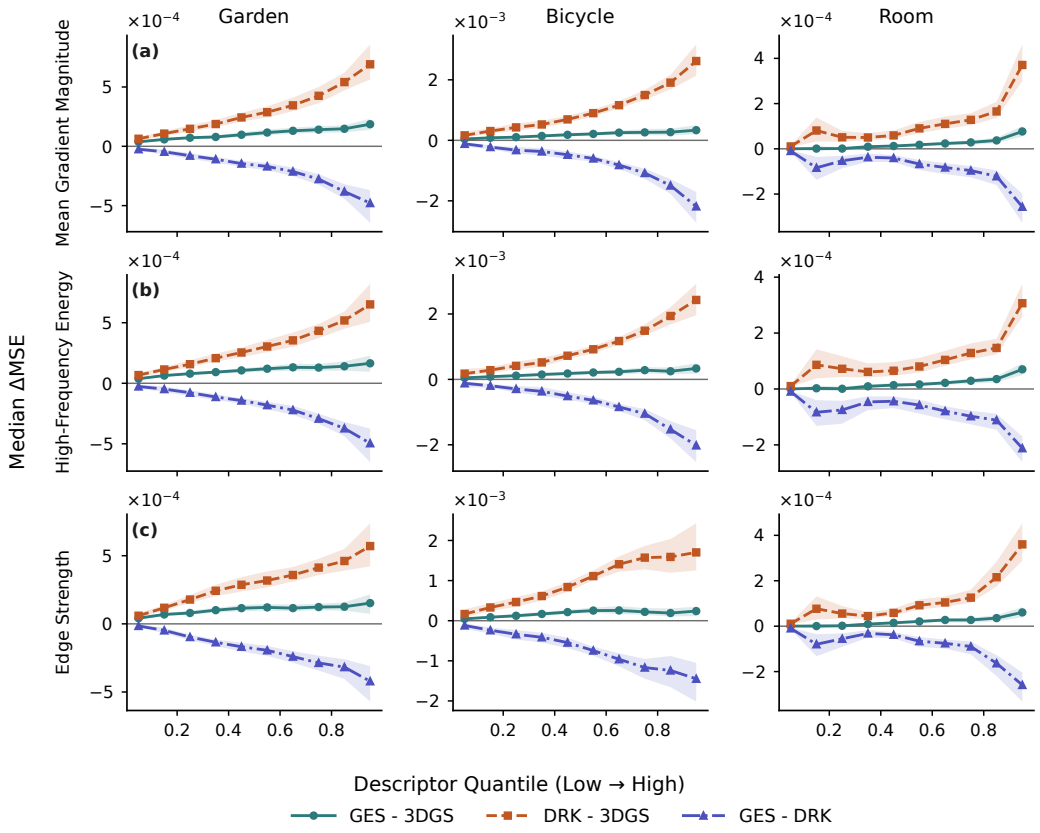


Figure 3: **Structure-conditioned cross-family reconstruction differences.** Median pairwise local MSE differences across quantiles of three representative structural descriptors for *Garden*, *Bicycle*, and *Room*. Shading denotes 95% view-level cluster-bootstrap confidence intervals; positive (negative) values favor the second (first) family.

evaluated against the same ground-truth image, with camera-view correspondence verified before local analysis. We compare complete independently optimized kernel-family representations rather than causally isolate kernel functional form; the methods retain differences in parameterization, optimization, and resulting primitive count, motivating the matched primitive-budget control in Section 3.4. Exact densification settings, and additional training details are provided in Appendix C.1.

3.2 Real-Scene Evaluation

We evaluate local cross-family specialization on three Mip-NeRF 360 scenes: *Garden*, *Bicycle*, and *Room*, spanning diverse indoor and outdoor structure. Following Section 3.1, each representation is evaluated on 24, 25, and 39 held-out views under the standard Mip-NeRF 360 evaluation split (Barron et al., 2022), respectively, with all renderings aligned to the same ground-truth views (Appendix C.2). Our primary analysis uses 32×32 patches with stride 16 and MSE tie tolerance $\tau = 10^{-5}$. Robustness to patch scale and tie tolerance is evaluated in Section 3.4. For each patch, we compute local reconstruction errors $E_k(p)$ (Eq. 4), six ground-truth structural descriptors (Section 2.3), and local family assignments (Eq. 6). Specialization is characterized through winner fractions, pairwise relative errors $\Delta E_{a,b}(p)$, post-hoc oracle headroom, and structure-conditioned trends. Because overlapping patches within a view are dependent, uncertainty is estimated by bootstrapping held-out views as clusters rather than individual patches. Cross-scene conclusions require reproducibility across all three scenes, with patch-scale and tie-tolerance sensitivity evaluated in Section 3.4.

Table 1: **Structure-conditioned cross-family specialization.** Patch-level Spearman correlations ρ between structural descriptors and pairwise local reconstruction differences $\Delta E_{a,b} = E_a - E_b$. All 18 descriptor–comparison relationships exhibit the same correlation direction across *Garden*, *Bicycle*, and *Room*.

Structural descriptors									
Descriptor	Comparison	Garden	Bicycle	Room	Descriptor	Comparison	Garden	Bicycle	Room
Mean gradient	GES–3DGS	0.193	0.168	0.142	Local variance	GES–3DGS	0.077	0.069	0.097
	DRK–3DGS	0.410	0.453	0.224		DRK–3DGS	0.272	0.289	0.229
	GES–DRK	-0.331	-0.385	-0.157		GES–DRK	-0.255	-0.270	-0.182
Edge strength	GES–3DGS	0.119	0.122	0.132	HF energy	GES–3DGS	0.149	0.159	0.136
	DRK–3DGS	0.331	0.383	0.229		DRK–3DGS	0.358	0.437	0.199
	GES–DRK	-0.304	-0.339	-0.168		GES–DRK	-0.323	-0.374	-0.137
Laplacian energy	GES–3DGS	0.161	0.176	0.138	Entropy	GES–3DGS	0.174	0.134	0.119
	DRK–3DGS	0.359	0.452	0.195		DRK–3DGS	0.383	0.385	0.239
	GES–DRK	-0.316	-0.379	-0.133		GES–DRK	-0.290	-0.330	-0.180

3.3 Controlled Structural Experiments

Real-scene associations cannot isolate individual structural properties because naturally occurring scene characteristics are correlated. We therefore construct a controlled synthetic benchmark varying three factors at low, medium, and high levels while matching other scene-generation conditions. *Edge sharpness* is controlled by the transition width w of $c(x) = \frac{1}{2} + \frac{1}{2}\tanh(x/w)$, with $w \in \{0.32, 0.12, 0.035\}$; *spatial frequency* by planar sinusoidal textures with $f \in \{2, 6, 14\}$ cycles across the scene width; and *surface curvature* by $z = \kappa(x^2 + y^2)$ with $\kappa \in \{0.02, 0.18, 0.30\}$. These factors manipulate boundary sharpness, appearance frequency, and geometric curvature, respectively. For each factor and level, we generate five instances using seeds $\{0, \ldots, 4\}$, yielding 45 synthetic scenes. Within each factor and seed, nuisance variables are fixed across levels to form matched observations. Each scene uses 24 training and 8 held-out views following the standard novel-view evaluation paradigm (Mildenhall et al., 2021), and is independently reconstructed with 3DGS, GES, and DRK, yielding 135 method–scene runs under the same family-specific training budgets as the real-scene experiments. Complete scene-generation, camera, initialization, and nuisance-variable specifications are provided in Appendix C.3.

Evaluation and statistical analysis. Novel-view reconstruction is evaluated using PSNR as the primary metric, with SSIM and LPIPS as secondary metrics. For each family pair, we analyze scene-level differences $\Delta M_{a,b} = M_a - M_b$ across structural levels, with seed as the independent replication unit. A Friedman test (Friedman, 1937) evaluates differences across the three matched conditions, with Benjamini–Hochberg (Benjamini & Hochberg, 1995) correction for multiple comparisons and Kendall's W as an effect-size measure (Kendall & Smith, 1939). Pairwise tests and additional statistical details are provided in Appendix C.4. This design complements the real-scene analysis by testing reproducible responses to controlled structural changes.

3.4 Capacity and Analysis Controls

We evaluate two potential confounds: primitive-count differences and sensitivity to the definition of local specialization. To test whether unequal primitive counts explain the observed complementarity, we perform an additional *Room* experiment in which 3DGS, GES, and DRK are each constrained to 250,000 primitives, a common budget close to the smallest naturally optimized representation and evaluated using the same held-out protocol. This control matches primitive count rather than all aspects of representation capacity, since the families retain distinct parameterizations and optimization procedures. We additionally repeat the local analysis over patch sizes $P \in \{32, 64, 128\}$ and tie thresholds $\tau \in \{0, 10^{-5}, 5 \times 10^{-5}\}$ on *Garden*, *Bicycle*, and *Room*, recomputing winner and tie fractions and oracle headroom across the resulting 3×3 sensitivity grid. Complete implementation and sensitivity details are provided in Appendix C.5.

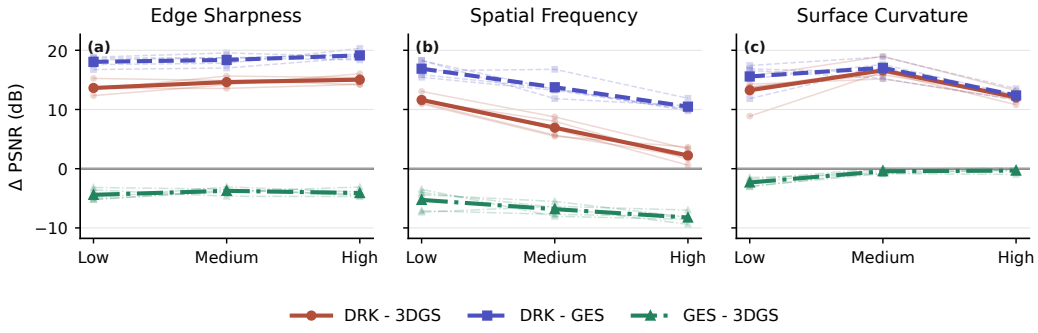


Figure 4: **Controlled structural effects on cross-family reconstruction.** Pairwise PSNR differences across low, medium, and high levels of (a) edge sharpness, (b) spatial frequency, and (c) surface curvature. Thin trajectories show five matched seeds and thick curves their mean; positive values favor the first family.

4 Results

4.1 RQ1: Local Specialization Across Real Scenes

Figure 1 shows local family assignments under the primary 32×32 patch configuration ($\tau = 10^{-5}$). Distinct 3DGS-, GES-, and DRK-preferred regions occur within all three real scenes. The winner fractions for 3DGS/GES/DRK are $64.74/19.56/10.49\%$ in *Garden*, $61.68/26.51/8.96\%$ in *Bicycle*, and $41.72/27.25/13.15\%$ in *Room*, with tie fractions of 5.21%, 2.85%, and 17.89%, respectively. Thus, regions favoring a non-3DGS family account for 30.05%, 35.47%, and 40.40% of the three scenes. The post-hoc oracle further indicates that these local differences are non-negligible. Relative to 3DGS patch MSE, oracle selection reduces error by 9.16% on *Garden*, 7.20% on *Bicycle*, and 31.16% on *Room*. Together, the spatial winner maps and oracle headroom show that scene-level preference for 3DGS does not imply uniform local superiority: alternative adaptive kernel families attain lower reconstruction error over substantial subsets of each evaluated scene. Additional qualitative examples and absolute reconstruction statistics are provided in Appendix Figures 5–6 and Table 3.

4.2 RQ2: Specialization Varies with Scene Structure

We next examine whether local cross-family differences vary systematically with ground-truth scene structure. Across all six descriptors, a consistent cross-scene pattern emerges: $\Delta E_{\mathrm{GES,3DGS}}$ and $\Delta E_{\mathrm{DRK,3DGS}}$ exhibit positive binned trends in *Garden*, *Bicycle*, and *Room*, whereas $\Delta E_{\mathrm{GES,DRK}}$ exhibits negative trends. Thus, all $6 \times 3 = 18$ descriptor–family relationships retain the same trend direction across all three scenes. Figure 3 visualizes representative trends for mean gradient magnitude, high-frequency energy, and edge strength, while Table 1 reports patch-level Spearman associations across all six descriptors. DRK–3DGS correlations are positive throughout ($\rho \approx 0.19$–0.45), GES–DRK correlations are consistently negative, and GES–3DGS associations are weaker but remain positive. Complete binned slopes and correlation statistics are reported in Appendix Table 4. Together, these results show that local cross-family differences are not merely spatially heterogeneous but vary systematically with measurable properties of the represented scene content.

4.3 RQ3: Controlled Structural Manipulation

Figure 4 shows how controlled changes in edge sharpness, spatial frequency, and surface curvature affect pairwise PSNR differences. The strongest ordered effect occurs for spatial frequency: the mean DRK–3DGS advantage decreases from 11.62 to 6.93 to 2.24 dB, with the same ordering across all five seeds ($q = 0.0403$, $W = 1.00$). DRK–GES shows a similar decrease ($16.89 \rightarrow 13.76 \rightarrow 10.48$ dB; $q = 0.0403$, $W = 0.84$), whereas GES–3DGS does not remain significant after correction ($q = 0.111$). Curvature produces significant but non-monotonic variation ($q = 0.0403$, $W = 0.84$), while edge sharpness shows comparatively weak evidence of modulation ($q = 0.505$, $W = 0.16$).

Table 2: **Robustness of local cross-family specialization.** Specialization persists under a matched 250k primitive budget and across increasingly coarse spatial analysis scales.

Matched budget: Room ($p = 32$)			**Spatial-scale sensitivity** ($\tau = 10^{-5}$)						
	Non-3DGS winners (%)	Oracle gain (%)		Non-3DGS winners (%)			Oracle gain vs. 3DGS (%)		
Setting			Scene	$p = 32$	$p = 64$	$p = 128$	$p = 32$	$p = 64$	$p = 128$
Natural	40.40	31.16	Garden	30.05	25.79	21.63	9.16	6.47	4.26
250k each	53.92	29.69	Bicycle	35.47	29.95	23.74	7.20	4.94	3.39
			Room	40.40	38.64	36.16	31.16	26.86	22.44

Thus, controlled structural changes reproducibly alter cross-family performance differences, with the form and strength of modulation depending on the manipulated factor.

4.4 RQ4: Robustness to Capacity and Analysis Choices

Table 2 shows that specialization persists under both controls. With a matched 250k primitive budget on *Room*, non-3DGS winners increase from 40.40% to 53.92%, while oracle gain remains comparable ($31.16\% \rightarrow 29.69\%$). Increasing patch size from 32 to 128 reduces both non-3DGS winner fractions and oracle headroom, but both remain nonzero across all three scenes. Thus, specialization weakens with spatial aggregation but persists across the evaluated analysis scales. The full patch-size and tie-threshold grid is reported in Appendix C.5.

5 Discussion

Implications of Cross-Family Specialization. Our results indicate that within-family adaptability does not necessarily make kernel families representationally interchangeable. Relative cross-family performance varies systematically with local structure, responds reproducibly to controlled structural changes, and remains heterogeneous under robustness controls. Thus, scene-level metrics can obscure local advantages retained by different adaptive kernel families. Rather than seeking only a universally superior kernel, representation design may therefore benefit from considering which kernel characteristics suit different local structures, motivating heterogeneous primitives or local selection among kernel families.

Limitations and Future Directions. Our conclusions should be interpreted within the scope of the evaluated systems. Because the families differ in kernel function, parameterization, and optimization, our results establish cross-family representation differences rather than causally isolating kernel functional form. Although the matched primitive-budget experiment controls one important source of capacity variation, it does not equalize all representational or optimization differences across families. We consider a finite set of families and scenes, and image-space descriptors only partially capture underlying 3D structure. Moreover, controlled manipulations alter relative performance gaps without necessarily changing the globally best family, so they demonstrate structure-dependent modulation rather than universal family switching. The oracle is likewise diagnostic rather than realizable. These limitations motivate broader evaluations and representations that jointly optimize heterogeneous kernel families and select them according to local structural requirements.

6 Conclusion

We examined whether increasingly adaptive splatting kernels become representationally interchangeable or retain distinct local biases. Across real-scene analysis, structure-conditioned characterization, controlled synthetic experiments, and robustness controls, we find persistent cross-family specialization: different kernel families achieve lower reconstruction error in different regions, with relative performance varying systematically with local scene structure. Controlled manipulations reproducibly alter these differences, while specialization persists across spatial scales and under a matched primitive budget. These results indicate that within-family adaptability does not necessarily eliminate cross-family representation differences, motivating representations that account for compatibility between kernel characteristics and local scene structure.

ETHICS STATEMENT

This work studies representation behavior in 3D scene reconstruction and does not involve human subjects, personally identifiable information, or sensitive data. Our experiments use publicly available benchmark data and synthetically generated scenes. We do not identify significant direct ethical risks arising from the methodology presented in this work. The authors are responsible for the integrity, reproducibility, and reporting of all experiments and results.

REPRODUCIBILITY STATEMENT

We provide detailed information to support reproduction of all experiments. Section 3.1 describes the evaluated kernel families and training protocol, while Sections 3.2–3.4 specify the real-scene, controlled synthetic, and robustness experiments. The appendix provides exact implementation and training details, structural descriptor definitions, synthetic scene-generation parameters, statistical procedures, and complete robustness analyses. Code, evaluation scripts, configuration files, and synthetic-scene generation utilities will be released with the supplementary material. All reported results were generated using documented experimental configurations and fixed random seeds where applicable.

AI USE STATEMENT

Generative AI tools were used to assist with grammatical and stylistic polishing of the manuscript and, in limited instances, to support refinement of experimental design and implementation. The research questions, core hypotheses, methodology, and scientific contributions originated from the authors. All AI-assisted suggestions, text, code, and experimental procedures were reviewed and verified by the authors. The authors take full responsibility for the accuracy, originality, claims, results, and final content of this work.

REFERENCES

Yanqi Bao, Tianyu Ding, Jing Huo, Yaoli Liu, Yuxin Li, Wenbin Li, Yang Gao, and Jiebo Luo. 3d gaussian splatting: Survey, technologies, challenges, and opportunities. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(7):6832–6852, 2025.

Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5460–5469. IEEE, 2022.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.

Kunal Chelani, Assia Benbihi, Torsten Sattler, and Fredrik Kahl. Edgegaussians-3d edge mapping via gaussian splatting. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 3268–3279. IEEE, 2025.

Guikun Chen and Wenguan Wang. A survey on 3d gaussian splatting. *ACM Computing Surveys*, 58 (12):1–39, 2026.

Haodong Chen, Runnan Chen, Qiang Qu, Zhaoqing Wang, Tongliang Liu, Xiaoming Chen, and Yuk Ying Chung. Beyond gaussians: Fast and high-fidelity 3d splatting with linear kernels. *arXiv preprint arXiv:2411.12440*, 2024.

Ben Fei, Jingyi Xu, Rui Zhang, Qingyuan Zhou, Weidong Yang, and Ying He. 3d gaussian splatting as a new era: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 31(8): 4429–4449, 2024.

Milton Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, 32(200):675–701, 1937.

Xiaowen Fu, Yang Zhang, Yuhan Tang, Huazhong Zhang, Tianxing Zhao, Yuhang Guo, Yu Huang, and Jinbao Wang. Scene experts: Specializing in 3d gaussian splatting with adaptive decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 4040–4048, 2026.

Yuanhao Gong. Eggs: Edge guided gaussian splatting for radiance fields. In *Proceedings of the 29th International ACM Conference on 3D Web Technology*, pp. 1–5, 2024.

Antoine Guédon and Vincent Lepetit. Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5354–5363. IEEE, 2024.

Abdullah Hamdi, Luke Melas-Kyriazi, Jinjie Mai, Guocheng Qian, Ruoshi Liu, Carl Vondrick, Bernard Ghanem, and Andrea Vedaldi. Ges: Generalized exponential splatting for efficient radiance field rendering. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19812–19822. IEEE, 2024.

Jan Held, Renaud Vandeghen, Abdullah Hamdi, Adrien Deliege, Anthony Cioppa, Silvio Giancola, Andrea Vedaldi, Bernard Ghanem, and Marc Van Droogenbroeck. 3d convex splatting: Radiance field rendering with 3d smooth convexes. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24360–21369. IEEE, 2025.

Yi-Hua Huang, Ming-Xian Lin, Yang-Tian Sun, Ziyi Yang, Xiaoyang Lyu, Yan-Pei Cao, and Xiaojuan Qi. Deformable radial kernel splatting. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21513–21523. IEEE, 2025.

Kaiwen Jiang, Venkataram Sivaram, Cheng Peng, and Ravi Ramamoorthi. Geometry field splatting with gaussian surfels. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5752–5762. IEEE, 2025.

Maurice G Kendall and B Babington Smith. The problem of m rankings. *The annals of mathematical statistics*, 10(3):275–287, 1939.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, George Drettakis, et al. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1, 2023.

Leif Kobbelt and Mario Botsch. A survey of point-based techniques in computer graphics. *Computers & Graphics*, 28(6):801–814, 2004.

Hanyang Kong, Xingyi Yang, and Xinchao Wang. Generative sparse-view gaussian splatting. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 26745–26755. IEEE, 2025.

Guillaume Lavoué, Mohamed Chaker Larabi, and Libor Váša. On the efficiency of image metrics for evaluating the visual quality of 3d models. *IEEE transactions on visualization and computer graphics*, 22(8):1987–1999, 2015.

Rong Liu, Dylan Sun, Meida Chen, Yue Wang, and Andrew Feng. Deformable beta splatting. In *Proceedings of the special interest group on computer graphics and interactive techniques conference conference papers*, pp. 1–11, 2025.

Daniel MacSwayne, Ales Leonardis, and Jianbo Jiao. 3d superquadric splatting. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 5154–5163, 2026.

Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

Joerg H Mueller, Martin Winter, and Markus Steinberger. From ex (p) to poly: Gaussian splatting with polynomial kernels. *arXiv preprint arXiv:2603.18707*, 2026.

Hashiru Pramuditha, Vinasirajan Viruthshaan, Vishagar Arunan, Saeedha Nazar, Sameera Ramasinghe, Simon Lucey, and Ranga Rodrigo. Darb-splatting: Generalizing splatting with decaying anisotropic radial basis functions. In *2026 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 4504–4514. IEEE, 2026.

Darius Rückert, Linus Franke, and Marc Stamminger. Adop: Approximate differentiable one-pixel point rendering. *ACM Transactions on Graphics (ToG)*, 41(4):1–14, 2022.

Mingwen Shao, Yuanjian Qiao, Kai Zhang, and Lingzhuang Meng. Frequency-aware uncertainty gaussian splatting for dynamic scene reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 31(5):3558–3568, 2025.

Qiangeng Xu, Zexiang Xu, Julien Philip, Sai Bi, Zhixin Shu, Kalyan Sunkavalli, and Ulrich Neumann. Point-nerf: Point-based neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5428–5438. IEEE, 2022.

Jiahui Zhang, Fangneng Zhan, Muyu Xu, Shijian Lu, and Eric Xing. Fregs: 3d gaussian splatting with progressive frequency regularization. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21424–21433. IEEE, 2024.

Qiang Zhang, Seung-Hwan Baek, Szymon Rusinkiewicz, and Felix Heide. Differentiable point-based radiance fields for efficient view synthesis. In *SIGGRAPH Asia 2022 Conference Papers*, pp. 1–12, 2022.

M. Zwicker, H. Pfister, J. van Baar, and M. Gross. Ewa splatting. *IEEE Transactions on Visualization and Computer Graphics*, 8(3):223–238, 2002. doi: 10.1109/TVCG.2002.1021576.

Matthias Zwicker, Hanspeter Pfister, Jeroen Van Baar, and Markus Gross. Surface splatting. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 371–378, 2001.

Appendix

A Related Work

B Additional Characterization Details

B.1 Additional Kernel-Family Formulations

B.2 Additional Local Specialization Metrics

B.3 Structural Descriptor Definitions

B.4 Structure-Conditioned Statistical Analysis

C Additional Experimental Details

C.1 Implementation and Training Details

C.2 Real-Scene Evaluation Details

C.3 Controlled Synthetic Benchmark Details

C.4 Controlled-Experiment Statistical Analysis

C.5 Capacity-Control and Sensitivity Details

A Related Work

Point-Based Rendering and Learned Splatting Representations. Point-based rendering has a long history in computer graphics (Kobbelt & Botsch, 2004). Surface splatting and elliptical weighted average (EWA) splatting established reconstruction kernels and anisotropic filtering as central components of point-based rendering (Zwicker et al., 2001; 2002). These principles later entered learned novel-view synthesis through explicit differentiable representations. Point-NeRF (Xu et al., 2022) represents radiance fields using neural 3D points, while Differentiable Point-Based Radiance Fields (Zhang et al., 2022) and ADOP (Rückert et al., 2022) demonstrated differentiable point-based formulations for high-quality novel-view synthesis. 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) subsequently established anisotropic Gaussian primitives as an efficient explicit representation for radiance-field rendering, motivating a rapidly expanding family of Gaussian-based methods (Fei et al., 2024; Bao et al., 2025).

Beyond Gaussian Reconstruction Kernels. Recent work has increasingly reconsidered the Gaussian kernel itself. GES (Hamdi et al., 2024) replaces the fixed Gaussian decay profile with a generalized exponential formulation whose shape parameter is learned during optimization. DRK (Huang et al., 2025) introduces deformable radial kernels with greater flexibility in support and shape, while Deformable Beta Splatting (Liu et al., 2025) further expands the space of adaptive splatting primitives beyond conventional Gaussian support. Other formulations explore linear kernels (Chen et al., 2024), smooth convex primitives (Held et al., 2025), superquadric representations (MacSwayne et al., 2026), decaying anisotropic radial-basis functions (Pramuditha et al., 2026), and polynomial kernels (Mueller et al., 2026). Together, these methods demonstrate a progression from a prescribed Gaussian primitive toward increasingly diverse and internally adaptive reconstruction families. Existing comparisons, however, predominantly evaluate each family as a scene-wide representation. Our work instead asks whether increasing within-family adaptability eliminates systematic *local* representation differences between independently optimized kernel families.

Structure-Aware Representation and Reconstruction. A complementary line of work adapts splatting representations or optimization strategies according to scene structure. SuGaR (Guédon & Lepetit, 2024) and Geometry Field Splatting (Jiang et al., 2025) incorporate surface and geometric structure into Gaussian-based representations, while EGGS (Gong, 2024) and EdgeGaussians (Chelani et al., 2025) target boundaries and discontinuities. Frequency-aware approaches explicitly address variation in visual detail (Shao et al., 2025), while Scene Experts (Fu et al., 2026) specializes Gaussian representations across spatial regions through adaptive scene decomposition. These studies establish that geometry, boundaries, frequency content, and spatial variation can motivate structure-dependent representation or optimization strategies. They primarily, however, adapt or specialize a particular representation according to scene structure. We study a distinct question: whether *independently optimized kernel families evaluated under the same local structural conditions* retain systematic differences in reconstruction behavior after each family has undergone its own internal adaptation.

B Additional Characterization Details

B.1 Additional Kernel-Family Formulations

For completeness, we provide the detailed deformable radial parameterization used by DRK Splatting Huang et al. (2025). For local tangent-plane coordinates (u, v), let $r_2 = \sqrt{u^2 + v^2}$ and let θ denote the corresponding polar angle. Given adjacent radial bases $(s_{i,k}, \theta_{i,k})$ and $(s_{i,k+1}, \theta_{i,k+1})$, DRK defines

$$K_i^{\mathrm{DRK}}(u, v) = \exp\left[-\frac{r_2^2}{2}\left(\frac{1 + \cos\Delta\theta_{i,k}}{2s_{i,k}^2} + \frac{1 - \cos\Delta\theta_{i,k}}{2s_{i,k+1}^2}\right)\right], \tag{9}$$

where

$$\Delta\theta_{i,k} = \frac{(\theta - \theta_{i,k})\pi}{\theta_{i,k+1} - \theta_{i,k}}, \qquad \theta_{i,k} < \theta \le \theta_{i,k+1}. \tag{10}$$

The learnable radial lengths and angles deform the primitive footprint, while additional DRK parameters control boundary curvature and sharpness. The formulation contains the 2D Gaussian as a special configuration of its radial bases Huang et al. (2025).

B.2 Additional Local Specialization Metrics

Winner and tie fractions. For the complete collection of analyzed regions $\mathcal{P} = \bigcup_v \mathcal{P}_v$, the winner fraction of family $k \in \mathcal{K}$ is

$$W_k = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} \mathbb{1}[S(p) = k], \tag{11}$$

and the corresponding tie fraction is

$$W_{\text{tie}} = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} \mathbb{1}[S(p) = \text{tie}], \tag{12}$$

such that $\sum_{k \in \mathcal{K}} W_k + W_{\text{tie}} = 1$. Here, $\mathbb{1}[\cdot]$ denotes the indicator function. These quantities summarize the spatial prevalence of local family preference while preserving ties as a separate outcome.

Oracle aggregation. For each region, the post-hoc oracle selects the smallest reconstruction error,

$$E_{\text{oracle}}(p) = \min_{k \in \mathcal{K}} E_k(p). \tag{13}$$

The aggregate patch-wise errors for family k and the oracle are

$$\bar{E}_k = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} E_k(p), \tag{14}$$

$$\bar{E}_{\text{oracle}} = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} E_{\text{oracle}}(p). \tag{15}$$

The oracle headroom reported in the main paper is then $H_k = (\bar{E}_k - \bar{E}_{\text{oracle}})/\bar{E}_k$. By construction, $H_k \geq 0$.

B.3 Structural Descriptor Definitions

All structural descriptors are computed exclusively from the ground-truth image patches and are independent of the renderings produced by the evaluated kernel families. Given an RGB patch, luminance is computed as

$$Y(\mathbf{x}) = 0.2126R(\mathbf{x}) + 0.7152G(\mathbf{x}) + 0.0722B(\mathbf{x}).$$

The mean gradient magnitude is

$$s_{\text{grad}}(p) = \frac{1}{|p|} \sum_{\mathbf{x} \in p} \|\nabla Y(\mathbf{x})\|_2 . \tag{16}$$

Edge strength is defined as the 90th percentile of the gradient-magnitude distribution,

$$s_{\text{edge}}(p) = Q_{0.9}\left(\{\|\nabla Y(\mathbf{x})\|_2 : \mathbf{x} \in p\}\right). \tag{17}$$

Laplacian energy measures rapid second-order intensity variation,

$$s_{\text{lap}}(p) = \frac{1}{|p|} \sum_{\mathbf{x} \in p} \left(\Delta Y(\mathbf{x})\right)^2, \tag{18}$$

where Δ denotes the discrete four-neighbor Laplacian.

Local intensity variance is

$$s_{\text{var}}(p) = \frac{1}{|p|} \sum_{\mathbf{x} \in p} \left(Y(\mathbf{x}) - \bar{Y}_p\right)^2, \qquad \bar{Y}_p = \frac{1}{|p|} \sum_{\mathbf{x} \in p} Y(\mathbf{x}). \tag{19}$$

High-frequency residual energy is computed as

$$s_{\mathrm{hf}}(p) = \frac{1}{|p|} \sum_{\mathbf{x} \in p} \left[Y(\mathbf{x}) - (B_{3\times 3} * Y)(\mathbf{x}) \right]^2 , \tag{20}$$

where $B_{3\times 3}$ is a normalized box kernel and reflected boundary padding is used.

Finally, intensity entropy is computed from a 32-bin luminance histogram over $[0, 1]$,

$$s_{\mathrm{ent}}(p) = - \sum_{j : q_j > 0} q_j \log_2 q_j, \tag{21}$$

where q_j denotes the empirical probability of histogram bin j.

B.4 Structure-Conditioned Statistical Analysis

Quantile trends. For each descriptor s_d, patches are sorted by descriptor magnitude and partitioned into $B = 10$ approximately equal-sized quantile groups $\{\mathcal{Q}_q\}_{q=1}^{B}$. For each family pair (a, b), the median pairwise error difference within each quantile is computed as defined in Eq. 8. A linear slope fitted across normalized quantile positions is used as a descriptive summary of trend direction rather than as an assumption of a linear patch-level relationship.

Winner-conditioned characterization. We complement the continuous pairwise analysis with the local assignment $S(p)$ from Eq. 6. For each descriptor, we estimate

$$P(S(p) = k \mid s_d(p) \in \mathcal{Q}_q) , \qquad k \in \mathcal{K}, \tag{22}$$

across descriptor quantiles. Because this discrete assignment discards the magnitude of cross-family error differences, it is treated as complementary to the continuous $\Delta E_{a,b}(p)$ analysis.

Clustered uncertainty. Patches extracted from the same rendered view are not statistically independent. Confidence intervals for binned error summaries are therefore obtained using a view-level cluster bootstrap. Held-out views are resampled with replacement, all patches belonging to a sampled view are retained as a cluster, and the statistic is recomputed for each bootstrap replicate. We use $2{,}000$ replicates and report percentile-based 95% confidence intervals. Multiple family–descriptor comparisons are controlled using the Benjamini–Hochberg false-discovery-rate procedure.

Cross-scene consistency. We distinguish within-scene association from cross-scene reproducibility. For each descriptor and family pair, the sign of the binned median-difference trend is evaluated independently across scenes. Consistent trend directions across scenes provide evidence that an observed structure–performance relationship is not specific to a single scene. Accordingly, patches constitute the local observations, views define the dependence structure for uncertainty estimation, and scenes provide the unit for cross-scene interpretation.

C Additional Experimental Details

C.1 Implementation and Training Details

Experiments use the official implementations released by the respective authors. For 3DGS Kerbl et al. (2023) and GES Hamdi et al. (2024). For DRK Huang et al. (2025), we use the official released implementation; the exact repository revision was not recorded in the experimental repository.

For the real-scene experiments, 3DGS, GES, and DRK are trained for 30k, 40k, and 35k iterations, respectively, and initialize their spatial primitives from the corresponding input COLMAP point cloud. All three representations perform adaptive primitive growth during optimization. Densification begins at iteration 500 and terminates at iteration 15k, with densification performed every 100 iterations and opacity reset every 3k iterations. The densification gradient thresholds are 2×10^{-4}, 3×10^{-4}, and 5×10^{-4} for 3DGS, GES, and DRK, respectively, following their method-specific configurations. GES additionally optimizes its kernel-shape parameters, while DRK uses its dense deformable-radial-kernel configuration.

Table 3: **Absolute reconstruction and local specialization results.** Global reconstruction metrics, final primitive counts, and local winner fractions for the three kernel families on the evaluated real scenes. Winner fractions use the primary 32×32 patch analysis with stride 16 and tie threshold $\tau = 10^{-5}$. Best global reconstruction metric within each scene is shown in bold.

Scene	Method	PSNR ↑	SSIM ↑	LPIPS ↓	Primitives	Local winner (%)
Garden	3DGS	**27.4781**	0.8681	**0.1061**	4,146,866	64.74
	GES	27.1017	0.8525	0.1347	2,709,992	19.56
	DRK	25.4346	**0.9193**	0.1880	614,345	10.49
Bicycle	3DGS	**25.2347**	0.7669	**0.2076**	4,925,950	61.68
	GES	24.8160	0.7324	0.2629	2,773,438	26.51
	DRK	23.3625	**0.7992**	0.3645	1,578,713	8.96
Room	3DGS	**31.4472**	0.9203	**0.2179**	1,314,498	41.72
	GES	31.3557	0.9139	0.2322	777,374	27.24
	DRK	27.9280	**0.9224**	0.3345	259,045	13.15

Unless explicitly stated otherwise, no common primitive-count constraint is imposed, allowing each representation to follow its native densification and pruning dynamics. The matched-budget experiment described in Section 3.4 separately controls final primitive count.

C.2 Real-Scene Evaluation Details

Cross-family view alignment. Local comparison requires reconstruction errors from different representations to correspond to identical image content. For each scene, we establish correspondence between the held-out renderings produced by 3DGS, GES, and DRK and a common ground-truth view. Ground-truth correspondence across the method-specific rendering pipelines is verified before computing local statistics, and comparisons are performed only between aligned renderings of the same camera view. This prevents differences in camera indexing or evaluation-view selection from being interpreted as cross-family reconstruction differences.

Patch extraction. The primary real-scene analysis uses square 32×32 image patches extracted with a stride of 16 pixels. The overlapping sampling provides spatially localized measurements while retaining sufficient image support for estimating reconstruction error and structural descriptors. Local family assignments use an MSE tolerance of $\tau = 10^{-5}$, with differences within this tolerance treated as ties.

For each scene, we compute winner fractions, pairwise relative errors, post-hoc oracle headroom, and structure-conditioned statistics. Because overlapping patches within a held-out view are dependent, uncertainty is estimated using view-level cluster resampling as detailed in Appendix B.4.

Absolute real-scene results. Table 3 reports the global reconstruction metrics, final primitive counts, and local winner fractions for all evaluated real-scene representations. These results provide the scene-level performance context for the local cross-family specialization analysis in the main paper.

C.3 Controlled Synthetic Benchmark Details

Structural manipulations. Edge sharpness is generated on a bounded planar surface using

$$c(x) = \frac{1}{2} + \frac{1}{2}\tanh\Big(\frac{x}{w}\Big), \tag{23}$$

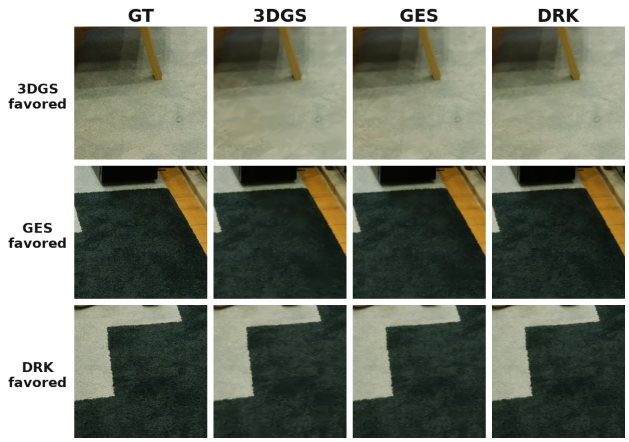
with $w \in \{0.32, 0.12, 0.035\}$ for the low-, medium-, and high-sharpness conditions, respectively. Decreasing w produces a sharper appearance discontinuity while preserving the underlying planar geometry.

Spatial frequency is manipulated on the same bounded planar geometry using an analytic two-dimensional sinusoidal texture,

$$c(x, y) \propto \sin(\phi_x(x))\sin(\phi_y(y)), \tag{24}$$



(a) Full ground-truth view

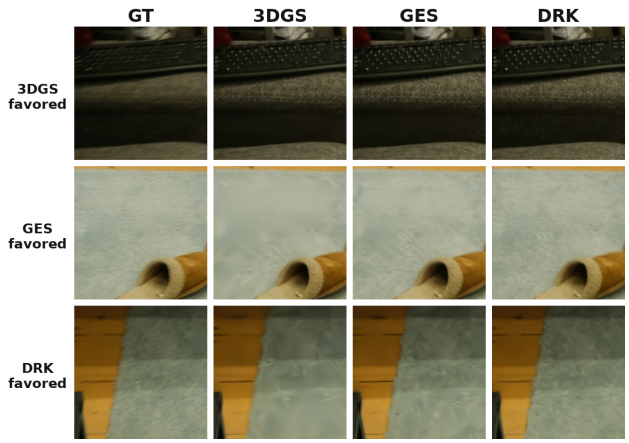


(b) Local reconstruction comparisons

Figure 5: **Additional qualitative example for Room.** The full ground-truth view is shown alongside representative regions locally favored by 3DGS, GES, and DRK. Each selected region is compared across the ground truth and all three independently optimized kernel-family representations, providing an additional example of local cross-family specialization.



(a) Full ground-truth view



(b) Local reconstruction comparisons

Figure 6: **Additional qualitative example for Room.** The full ground-truth view is shown alongside representative regions locally favored by 3DGS, GES, and DRK. Each selected region is compared across the ground truth and all three independently optimized kernel-family representations, further illustrating that local family preference occurs across different regions and viewpoints of the same scene.

with $f \in \{2, 6, 14\}$ cycles across the scene width for the low-, medium-, and high-frequency conditions, respectively.

Surface curvature is manipulated geometrically using

$$z = \kappa(x^2 + y^2), \tag{25}$$

where $\kappa \in \{0.02, 0.18, 0.30\}$ defines the three curvature levels.

Matched nuisance variation. Five independent instances are generated for every factor and level using seeds $\{0, 1, 2, 3, 4\}$. Edge scenes vary transition orientation and a small positional offset; frequency scenes vary texture orientation and sinusoidal phase; and curvature scenes vary the phase

and orientation of a low-contrast albedo pattern. For a given factor and seed, these nuisance variables are held fixed across all three structural levels, producing matched low-, medium-, and high-level observations.

Rendering and initialization. All scenes are rendered at 256×256 resolution using 24 uniformly spaced orbit views for training and 8 interleaved orbit views for testing. The camera radius is 2.8, elevation is $35°$, horizontal field of view is $50°$, scene extent is 1.15, and the background is white.

Each scene is initialized using $100{,}000$ points deterministically sampled in $[-1.25, 1.25]^3$, with the initialization seed matched to the corresponding scene instance. Synthetic scenes are consumed directly through their training and test camera transforms; no COLMAP reconstruction is used.

Cross-family reconstruction. Each of the 45 scenes is independently reconstructed using all three kernel families, producing 135 method–scene runs. We use 30k optimization iterations for 3DGS, 40k for GES, and 35k for DRK, matching the family-specific budgets used in the real-scene experiments.

C.4 Controlled-Experiment Statistical Analysis

For a reconstruction metric M and families a and b, relative performance is defined at the scene level as

$$\Delta M_{a,b} = M_a - M_b. \tag{26}$$

For PSNR and SSIM, positive values favor family a; for LPIPS, lower values are preferable and the interpretation is therefore reversed.

The random seed is treated as the independent replication unit. For each structural factor, family comparison, and evaluation metric, observations at the low, medium, and high structural levels are paired by seed. A Friedman test evaluates differences across the three ordered conditions. Pairwise contrasts are subsequently evaluated using exact two-sided Wilcoxon signed-rank tests for low versus medium, medium versus high, and low versus high. Benjamini–Hochberg false-discovery-rate correction is applied within each test family and evaluation metric.

We report Kendall's W for the Friedman analysis and the paired standardized effect size

$$d_z = \frac{\overline{\Delta D}}{s_{\Delta D}}, \tag{27}$$

where ΔD denotes the within-seed change in pairwise family-performance difference between two structural levels.

C.5 Capacity-Control and Sensitivity Details

Matched primitive-budget protocol. Under their native densification procedures, the evaluated representations can converge to substantially different primitive counts. We therefore perform an additional *Room* experiment with a common maximum budget of $250{,}000$ primitives. The same budget constraint is incorporated into all three training pipelines. Whenever densification or primitive growth causes family $k \in \mathcal{K}$ to exceed the budget, primitives with the lowest opacity are removed using the pruning mechanism of the corresponding implementation until $N_k \leq 250{,}000$.

The constraint is applied only to the budget-controlled experiment; models used in the primary analysis retain their native primitive counts. All three controlled models terminate with exactly $250{,}000$ primitives and are evaluated using the same held-out views, patch extraction, and local specialization protocol as the primary *Room* experiment. We recompute local family assignments, winner fractions, reconstruction errors, and oracle headroom.

This experiment controls final primitive count but does not equate all notions of representation capacity because the three families retain different primitive parameterizations and optimization procedures.

Patch-scale and tie-threshold sensitivity. We evaluate the full Cartesian product of patch sizes $P \in \{32, 64, 128\}$ and tie thresholds $\tau \in \{0, 10^{-5}, 5 \times 10^{-5}\}$ independently on *Garden*, *Bicycle*, and *Room*. The primary analysis uses $P = 32$ and $\tau = 10^{-5}$.

Table 4: **Complete structure-conditioned cross-family trends across real scenes.** Each scene entry reports the binned median-ΔMSE slope followed by the patch-level Spearman correlation ρ in parentheses. Pairwise differences follow $\Delta E_{a,b} = E_a - E_b$; therefore, negative values indicate lower reconstruction error for the first family in the comparison. All 18 descriptor–comparison relationships exhibit the same slope direction across *Garden*, *Bicycle*, and *Room*.

Descriptor	**Comparison**	**Garden**	**Bicycle** *slope (ρ)*	**Room**
Mean gradient magnitude	GES–3DGS	$+1.50 \times 10^{-4}$ (0.193)	$+3.07 \times 10^{-4}$ (0.168)	$+6.82 \times 10^{-5}$ (0.142)
	DRK–3DGS	$+6.42 \times 10^{-4}$ (0.410)	$+2.47 \times 10^{-3}$ (0.453)	$+2.68 \times 10^{-4}$ (0.224)
	GES–DRK	-4.70×10^{-4} (-0.331)	-1.99×10^{-3} (-0.385)	-1.74×10^{-4} (-0.157)
Edge strength	GES–3DGS	$+1.01 \times 10^{-4}$ (0.119)	$+1.97 \times 10^{-4}$ (0.122)	$+5.92 \times 10^{-5}$ (0.132)
	DRK–3DGS	$+5.18 \times 10^{-4}$ (0.331)	$+1.87 \times 10^{-3}$ (0.383)	$+2.84 \times 10^{-4}$ (0.229)
	GES–DRK	-4.14×10^{-4} (-0.304)	-1.51×10^{-3} (-0.339)	-1.91×10^{-4} (-0.168)
Laplacian energy	GES–3DGS	$+1.31 \times 10^{-4}$ (0.161)	$+3.38 \times 10^{-4}$ (0.176)	$+6.41 \times 10^{-5}$ (0.138)
	DRK–3DGS	$+6.10 \times 10^{-4}$ (0.359)	$+2.47 \times 10^{-3}$ (0.452)	$+2.03 \times 10^{-4}$ (0.195)
	GES–DRK	-4.81×10^{-4} (-0.316)	-1.97×10^{-3} (-0.379)	-1.29×10^{-4} (-0.133)
Local variance	GES–3DGS	$+6.32 \times 10^{-5}$ (0.077)	$+8.28 \times 10^{-5}$ (0.069)	$+3.40 \times 10^{-5}$ (0.097)
	DRK–3DGS	$+3.93 \times 10^{-4}$ (0.272)	$+1.15 \times 10^{-3}$ (0.289)	$+2.42 \times 10^{-4}$ (0.229)
	GES–DRK	-3.06×10^{-4} (-0.255)	-9.98×10^{-4} (-0.270)	-1.78×10^{-4} (-0.182)
High-frequency energy	GES–3DGS	$+1.25 \times 10^{-4}$ (0.149)	$+2.95 \times 10^{-4}$ (0.159)	$+6.31 \times 10^{-5}$ (0.136)
	DRK–3DGS	$+6.03 \times 10^{-4}$ (0.358)	$+2.39 \times 10^{-3}$ (0.437)	$+2.13 \times 10^{-4}$ (0.199)
	GES–DRK	-4.80×10^{-4} (-0.323)	-1.92×10^{-3} (-0.374)	-1.36×10^{-4} (-0.137)
Entropy	GES–3DGS	$+1.28 \times 10^{-4}$ (0.174)	$+2.11 \times 10^{-4}$ (0.134)	$+4.35 \times 10^{-5}$ (0.119)
	DRK–3DGS	$+5.52 \times 10^{-4}$ (0.383)	$+1.93 \times 10^{-3}$ (0.385)	$+2.84 \times 10^{-4}$ (0.239)
	GES–DRK	-3.82×10^{-4} (-0.290)	-1.57×10^{-3} (-0.330)	-2.09×10^{-4} (-0.180)

For every condition, we recompute the numbers of decisive and tied regions, family winner fractions, family-specific local reconstruction errors, oracle error, and oracle headroom. The structure-conditioned descriptor analysis is not rerun in this sweep because the purpose of the control is specifically to test whether the observed local complementarity depends on patch scale or the winner-assignment tolerance.

Table 5: **Complete structure-conditioned trend statistics.** Spearman ρ measures the patch-level monotonic association between each ground-truth structural descriptor and pairwise reconstruction difference $\Delta E_{a,b} = E_a - E_b$. Binned slopes summarize the corresponding median ΔMSE trend across descriptor quantiles. Because spatial patches overlap, interpretation emphasizes association magnitude and direction together with cross-scene consistency; view-clustered uncertainty is reported in the main analysis.

Scene	Descriptor	Comparison	Spearman ρ	Binned slope
Garden	Mean gradient	GES–3DGS	0.193	1.50×10^{-4}
		DRK–3DGS	0.410	6.42×10^{-4}
		GES–DRK	-0.331	-4.70×10^{-4}
	Edge strength	GES–3DGS	0.119	1.01×10^{-4}
		DRK–3DGS	0.331	5.18×10^{-4}
		GES–DRK	-0.304	-4.14×10^{-4}
	Laplacian energy	GES–3DGS	0.161	1.31×10^{-4}
		DRK–3DGS	0.359	6.10×10^{-4}
		GES–DRK	-0.316	-4.81×10^{-4}
	Local variance	GES–3DGS	0.077	6.33×10^{-5}
		DRK–3DGS	0.272	3.93×10^{-4}
		GES–DRK	-0.255	-3.06×10^{-4}
	HF energy	GES–3DGS	0.149	1.25×10^{-4}
		DRK–3DGS	0.358	6.03×10^{-4}
		GES–DRK	-0.323	-4.80×10^{-4}
	Entropy	GES–3DGS	0.174	1.28×10^{-4}
		DRK–3DGS	0.383	5.52×10^{-4}
		GES–DRK	-0.290	-3.82×10^{-4}
Bicycle	Mean gradient	GES–3DGS	0.168	3.08×10^{-4}
		DRK–3DGS	0.453	2.47×10^{-3}
		GES–DRK	-0.385	-1.99×10^{-3}
	Edge strength	GES–3DGS	0.122	1.97×10^{-4}
		DRK–3DGS	0.383	1.87×10^{-3}
		GES–DRK	-0.339	-1.51×10^{-3}
	Laplacian energy	GES–3DGS	0.176	3.38×10^{-4}
		DRK–3DGS	0.452	2.47×10^{-3}
		GES–DRK	-0.379	-1.97×10^{-3}
	Local variance	GES–3DGS	0.069	8.28×10^{-5}
		DRK–3DGS	0.289	1.15×10^{-3}
		GES–DRK	-0.270	-9.98×10^{-4}
	HF energy	GES–3DGS	0.159	2.95×10^{-4}
		DRK–3DGS	0.437	2.39×10^{-3}
		GES–DRK	-0.374	-1.92×10^{-3}
	Entropy	GES–3DGS	0.134	2.11×10^{-4}
		DRK–3DGS	0.385	1.93×10^{-3}
		GES–DRK	-0.330	-1.57×10^{-3}
Room	Mean gradient	GES–3DGS	0.142	6.82×10^{-5}
		DRK–3DGS	0.224	2.69×10^{-4}
		GES–DRK	-0.157	-1.74×10^{-4}
	Edge strength	GES–3DGS	0.132	5.92×10^{-5}
		DRK–3DGS	0.229	2.84×10^{-4}
		GES–DRK	-0.168	-1.91×10^{-4}
	Laplacian energy	GES–3DGS	0.138	6.41×10^{-5}
		DRK–3DGS	0.195	2.03×10^{-4}
		GES–DRK	-0.133	-1.29×10^{-4}
	Local variance	GES–3DGS	0.097	3.40×10^{-5}
		DRK–3DGS	0.229	2.42×10^{-4}
		GES–DRK	-0.182	-1.78×10^{-4}
	HF energy	GES–3DGS	0.136	6.31×10^{-5}
		DRK–3DGS	0.199	2.13×10^{-4}
		GES–DRK	-0.137	-1.36×10^{-4}
	Entropy	GES–3DGS	0.119	4.35×10^{-5}
		DRK–3DGS	0.239	2.84×10^{-4}
		GES–DRK	-0.180	-2.09×10^{-4}