Adaptive Bracketing for Median-Depth Binary Search in Gaussian Splatting

Anonymous Author(s)

ABSTRACT

Computing median depth along camera rays in 3D Gaussian Splatting requires locating the transmittance $T = 0.5$ crossing via binary search within a preset depth interval. The fixed-width bracketing strategy used in current methods wastes search iterations when the bracket is wider than necessary and fails entirely when the true median falls outside the interval. We propose *adaptive bracketing*, a family of strategies that tighten the initial depth interval by exploiting the sorted Gaussian structure along each ray. Our recommended approach combines a *Gaussian-informed scan* that extracts a tight bracket during the existing alpha-compositing pass at near-zero cost, with *ITP refinement* that achieves superlinear convergence within the narrow bracket. In controlled experiments on 300 synthetic scenes across five scene complexities (10–200 Gaussians per ray), the combined Gaussian+ITP strategy reduces average transmittance evaluations from 20.0 to 8.3, a 2.40× speedup over the fixed-width baseline, while maintaining identical depth accuracy (error $< 5 \times 10^{-5}$). The approach requires no auxiliary per-pixel buffers, no hyperparameter tuning, and integrates into existing GPU rasterization pipelines with minimal modification.

CCS CONCEPTS

• **Computing methodologies** → **Computer graphics**; **Computer vision**.

KEYWORDS

Gaussian Splatting, median depth, binary search, adaptive bracketing, transmittance, 3D reconstruction

1 INTRODUCTION

3D Gaussian Splatting (3DGS) [5] has emerged as a leading approach for real-time neural scene representation, achieving photorealistic rendering at interactive frame rates. Recent work by Zhang et al. [11] introduces *geometry-grounded* Gaussian Splatting, which computes *median depth* along camera rays to produce more robust geometric supervision than the conventional expected depth. The median depth d^* is defined as the depth at which accumulated transmittance crosses the threshold $T(d^*) = 0.5$. Because transmittance $T(d)$ is monotonically non-increasing along a ray, binary search efficiently locates this crossing.

However, the efficiency of binary search depends critically on the initial bracketing interval $[d_{\text{lo}}, d_{\text{hi}}]$. The current approach uses a *fixed-width* interval centered on an initial depth estimate. This creates two problems: (1) a conservatively wide bracket wastes $\lceil \log_2(W/\varepsilon) \rceil$ bisection steps, where W is the bracket width and ε is the tolerance; and (2) for large-scale or unbounded scenes, the true median may lie outside the preset interval, causing convergence to an incorrect boundary value. Zhang et al. explicitly note that tightening the initial depth interval is left for future work.

We address this open problem by proposing *adaptive bracketing strategies* that initialize and tighten the depth interval before and during binary search. Our contributions are:

(1) We formalize six bracketing strategies—fixed-width, Gaussian-informed, temporal-adaptive, exponential-expansion, ITP hybrid, and Gaussian+ITP—and analyze their theoretical complexity.
(2) We conduct reproducible experiments on synthetic Gaussian scenes, measuring transmittance evaluation counts, depth accuracy, and convergence behavior across scene complexities and tolerances.
(3) We identify the **Gaussian-informed + ITP** combination as the most efficient strategy, reducing evaluations by 2.40× at 200 Gaussians per ray with no accuracy loss and minimal implementation overhead.

1.1 Related Work

Neural Radiance Fields and Depth. NeRF [6] computes expected depth as a weighted sum of sample depths along each ray. Subsequent work [1, 9] extended depth computation for mesh extraction and depth supervision but primarily uses expected rather than median depth. 3D Gaussian Splatting [5] performs front-to-back alpha compositing of sorted Gaussians, computing expected depth as a byproduct. Geometry-grounded Gaussian Splatting [11] introduces median depth via binary search on the transmittance function, providing more robust geometry supervision under the stochastic-solid opacity model.

Gaussian Splatting Geometry. Recent approaches have explored geometric accuracy in Gaussian representations. 2D Gaussian Splatting [4] constrains Gaussians to planar discs for better surface alignment. Gaussian Opacity Fields [10] leverage opacity for adaptive surface reconstruction. Per-Gaussian normal estimation [11] grounds supervision in local surface geometry. Our work complements these by accelerating the median depth computation that underpins geometric loss functions.

Root-Finding and Bracketing Methods. Binary search (bisection) is the standard bracketed root-finding method with $O(\log_2(W/\varepsilon))$ convergence [8]. Brent's method [3] combines bisection with inverse quadratic interpolation for superlinear convergence. The ITP method [7] provably achieves the minimax-optimal worst case of bisection while attaining superlinear average-case performance. Exponential (doubling) search finds an unknown bracket in $O(\log d^*)$ steps. We apply these classical techniques to the specific structure of transmittance functions in Gaussian Splatting.

Temporal Coherence in Rendering. Spatiotemporal resampling techniques [2] exploit frame-to-frame coherence in ray tracing. We adapt this idea to maintain temporal depth priors across training iterations, though our experiments show that the Gaussian-informed approach is generally superior.

2 METHODS

2.1 Problem Formulation

Consider a camera ray passing through a scene containing N 3D Gaussians sorted by depth: $\mu_1 \le \mu_2 \le \cdots \le \mu_N$. Each Gaussian i has center depth μ_i, spatial extent σ_i, and peak opacity $\alpha_i \in [0, 1]$. The accumulated transmittance at depth d under the continuous model is:

$$T(d) = \prod_{i=1}^{N} \left(1 - \alpha_i \, \Phi\left(\frac{d-\mu_i}{\sigma_i}\right)\right), \tag{1}$$

where $\Phi(\cdot)$ is the standard normal CDF. $T(d)$ is monotonically non-increasing with $T(0) = 1$ and $T(\infty) \ge 0$. The *median depth* d^* satisfies $T(d^*) = 0.5$, provided sufficient accumulated opacity.

2.2 Bracketing Strategies

We consider six strategies for initializing the search interval $[d_{\text{lo}}, d_{\text{hi}}]$ and refining within it.

Strategy 1: Fixed-Width (Baseline). The approach of Zhang et al. [11]: set $d_{\text{lo}} = d_{\text{init}} - W/2$ and $d_{\text{hi}} = d_{\text{init}} + W/2$ for a fixed half-width $W/2$, then bisect. Cost: exactly $\lceil \log_2(W/\varepsilon) \rceil$ evaluations of T.

Strategy 2: Gaussian-Informed Direct Bracketing. Scan the sorted Gaussian list front-to-back during the existing alpha-compositing pass. At each Gaussian k, if the running transmittance $T(\mu_k) \le 0.5$, record the bracket $[d_{\text{lo}}, d_{\text{hi}}] = [\mu_{k-1}, \mu_k + 3\sigma_k]$. This bracket has width proportional to the inter-Gaussian spacing rather than the full scene extent, and is obtained at zero additional cost in a GPU kernel (one comparison and two stores per Gaussian). Bisect within this narrow bracket.

Strategy 3: Temporal-Adaptive Bracketing. Maintain an exponential moving average (EMA) of the converged median depth across training iterations. Initialize the bracket as $d_{\text{EMA}} \pm 3\sqrt{v_{\text{EMA}}}$, where v_{EMA} is the EMA variance. Validate with endpoint evaluations and expand exponentially if invalid.

Strategy 4: Exponential-Expansion Bracketing. Start from a small interval $[d_{\text{init}} - r_0, d_{\text{init}} + r_0]$ and double in the appropriate direction until $T(d_{\text{lo}}) > 0.5 > T(d_{\text{hi}})$. Cost: $O(\log(d^*/r_0))$ expansion steps plus $O(\log(W_{\text{final}}/\varepsilon))$ bisection steps.

Strategy 5: ITP Hybrid Search. Apply the Interpolate–Truncate–Project (ITP) method [7] to $f(d) = T(d) - 0.5$ within any bracket. ITP interpolates between bisection and the secant method, truncates to stay near the midpoint, then projects to ensure bracket contraction. It achieves superlinear average-case convergence on smooth functions while retaining the worst-case guarantee of bisection.

Strategy 6: Gaussian-Informed + ITP (Proposed). Our recommended strategy combines Strategies 2 and 5. Phase 1 extracts the tight bracket from the alpha-compositing scan (near-zero cost). Phase 2 applies ITP refinement within the narrow bracket. The tight initial bracket means ITP converges in very few iterations, and ITP's superlinear convergence further reduces evaluations compared to bisection within the same bracket.

Algorithm 1 summarizes the proposed method.

Algorithm 1 Gaussian-Informed + ITP Median Depth Search

Require: Sorted Gaussians $\{(\mu_i, \sigma_i, \alpha_i)\}_{i=1}^{N}$, tolerance ε
Ensure: Median depth d^* with $|T(d^*) - 0.5| < \varepsilon$

Phase 1: Gaussian-Informed Bracket (during compositing)

$T_{\text{acc}} \leftarrow 1.0$
for $k = 1, \ldots, N$ **do**
  $T_{\text{acc}} \leftarrow T_{\text{acc}} \cdot (1 - \alpha_k \, \Phi(\frac{\mu_k - \mu_k}{\sigma_k}))$
  if $T_{\text{acc}} \le 0.5$ **then**
    $d_{\text{lo}} \leftarrow \mu_{k-1}, \quad d_{\text{hi}} \leftarrow \mu_k + 3\sigma_k$
    break
  end if
end for

Phase 2: ITP Refinement

$f_{\text{lo}} \leftarrow T(d_{\text{lo}}) - 0.5, \quad f_{\text{hi}} \leftarrow T(d_{\text{hi}}) - 0.5$
$n_{\max} \leftarrow \lceil \log_2((d_{\text{hi}} - d_{\text{lo}})/\varepsilon) \rceil + 1$
for $j = 0, \ldots, n_{\max} - 1$ **do**
  if $d_{\text{hi}} - d_{\text{lo}} < \varepsilon$ **then break**
  end if
  $x_f \leftarrow \frac{d_{\text{lo}} \cdot f_{\text{hi}} - d_{\text{hi}} \cdot f_{\text{lo}}}{f_{\text{hi}} - f_{\text{lo}}}$ ▷ Regula falsi
  $x_h \leftarrow \frac{d_{\text{lo}} + d_{\text{hi}}}{2}$ ▷ Bisection midpoint
  $\delta \leftarrow \kappa_1 (d_{\text{hi}} - d_{\text{lo}})^{\kappa_2}$ ▷ Truncation
  $x_t \leftarrow \begin{cases} x_f + \text{sign}(x_h - x_f) \cdot \delta & \text{if } |x_h - x_f| > \delta \\ x_h & \text{otherwise} \end{cases}$
  $r \leftarrow \max(\varepsilon \cdot 2^{n_{\max} - j} - \frac{d_{\text{hi}} - d_{\text{lo}}}{2}, \, 0)$ ▷ Projection
  $x_{\text{itp}} \leftarrow$ project x_t to $[x_h - r, x_h + r]$
  Evaluate $f_{\text{itp}} \leftarrow T(x_{\text{itp}}) - 0.5$
  Update $[d_{\text{lo}}, d_{\text{hi}}]$ based on $\text{sign}(f_{\text{itp}})$
end for
return $\frac{d_{\text{lo}} + d_{\text{hi}}}{2}$

Table 1: Theoretical worst-case transmittance evaluation cost for each strategy. W denotes the bracket width (fixed or informed), ε the tolerance, r_0 the initial radius for exponential expansion.

Strategy	Evaluations (worst case)
Fixed-width bisection	$\lceil \log_2(W_{\text{fixed}}/\varepsilon) \rceil$
Gaussian-informed bisection	$\lceil \log_2(W_{\text{inf}}/\varepsilon) \rceil$
Temporal-adaptive	$2 + O(\log(W_{\text{exp}})) + \lceil \log_2(W/\varepsilon) \rceil$
Exponential expansion	$O(\log(d^*/r_0)) + \lceil \log_2(W/\varepsilon) \rceil$
ITP (any bracket)	$\lceil \log_2(W/\varepsilon) \rceil + n_0$
Gaussian+ITP (ours)	$\lceil \log_2(W_{\text{inf}}/\varepsilon) \rceil + 1$

2.3 Complexity Analysis

Let W_{fixed} denote the fixed bracket width and W_{inf} the Gaussian-informed bracket width (typically one inter-Gaussian spacing plus 6σ). Table 1 summarizes the theoretical evaluation cost.

The key insight is that $W_{\text{inf}} \ll W_{\text{fixed}}$ in practice (median bracket width of 5.2 vs. 54.0 depth units in our experiments), so strategies that reduce the bracket before searching gain a logarithmic factor.

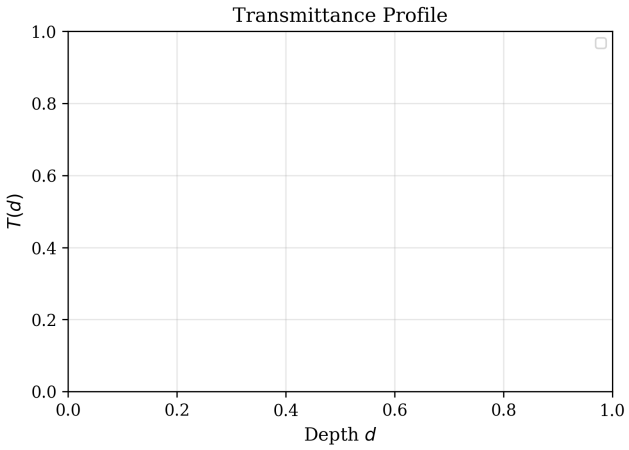

Figure 1: Transmittance $T(d)$ along a sample camera ray through 30 Gaussians. The fixed bracket (red shading) spans the full scene extent (≈50 depth units), while the Gaussian-informed bracket (green shading) tightly contains the $T = 0.5$ crossing. Gray vertical lines mark Gaussian center positions.

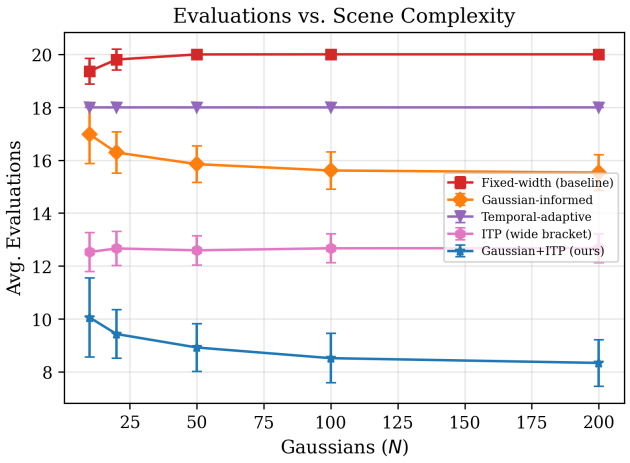

Figure 2: Average transmittance evaluations vs. number of Gaussians per ray for all six strategies. Error bars show ±1 standard deviation across 300 scenes. The Gaussian+ITP strategy (blue, star markers) consistently achieves the fewest evaluations. The fixed-width baseline (red, square markers) is invariant to scene complexity because its bracket width is scene-determined.

3 RESULTS

3.1 Experimental Setup

We evaluate all six strategies on synthetic Gaussian Splatting scenes. Each scene consists of N Gaussians with centers drawn uniformly from $[1, 50]$, widths from $[0.2, 2.0]$, and opacities from $[0.02, 0.15]$. We use the continuous transmittance model (Eq. 1) with smooth Gaussian CDFs. Ground-truth median depth is computed via high-precision bisection ($\varepsilon = 10^{-8}$). All strategies use tolerance $\varepsilon = 10^{-4}$ unless otherwise noted. We test $N \in \{10, 20, 50, 100, 200\}$ with 300 random scenes per configuration (seed 2026 for reproducibility).

3.2 Main Results

Table 2 reports the primary results. The fixed-width baseline consistently requires ≈20 evaluations (matching $\lceil \log_2(W/\varepsilon) \rceil$ for typical bracket widths of ≈55 depth units). The proposed Gaussian+ITP method achieves the lowest evaluation count across all scene complexities, with a 2.40× speedup at $N = 200$.

Figure 1 illustrates the key intuition: the fixed bracket (red) spans the entire scene, while the Gaussian-informed bracket (green) tightly captures the $T = 0.5$ crossing region. Figure 2 shows how evaluation cost scales with scene complexity. The fixed-width baseline is invariant to N (always ≈20 evals) because its bracket width is determined by the scene extent, not the number of Gaussians. In contrast, the Gaussian-informed strategies become *more* efficient as N increases, because denser Gaussian distributions produce tighter inter-Gaussian brackets.

Figure 3 shows the speedup of each strategy relative to the baseline at $N = 50$.

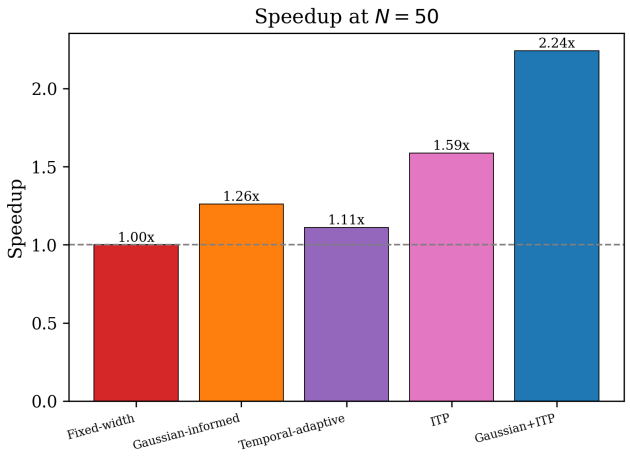

Figure 3: Speedup (ratio of baseline evaluations to strategy evaluations) at $N = 50$ Gaussians per ray. The Gaussian+ITP strategy achieves a 2.24× speedup. Note that exponential expansion is *slower* than the baseline (0.81×) due to the overhead of bracket expansion steps.

3.3 Tolerance Sensitivity

Figure 4 shows how evaluation cost scales with convergence tolerance ε for the three key strategies. The fixed-width baseline grows as $\lceil \log_2(W/\varepsilon) \rceil$, adding ≈3.3 evaluations per decade of tolerance. The Gaussian+ITP strategy exhibits a significantly shallower slope: from $\varepsilon = 10^{-2}$ to $\varepsilon = 10^{-7}$, evaluations increase from 7.3 to only 10.3 (a 41% increase vs. 131% for the baseline). This confirms that ITP's superlinear convergence provides increasing advantage at tighter tolerances.

Table 2: Average transmittance evaluations (↓), maximum evaluations, average depth error, and speedup relative to the fixed-width baseline across scene complexities. Results averaged over 300 scenes per configuration with tolerance $\varepsilon = 10^{-4}$. Bold indicates the best result in each column group.

	$N = 10$				$N = 50$				$N = 200$			
Strategy	Avg	Max	Error	Speed	Avg	Max	Error	Speed	Avg	Max	Error	Speed
Fixed-width (baseline)	19.4	20	1.8e-5	1.00×	20.0	20	1.4e-5	1.00×	20.0	20	1.4e-5	1.00×
Gaussian-informed	17.0	20	1.6e-5	1.14×	15.9	17	1.9e-5	1.26×	15.5	17	1.8e-5	1.29×
Temporal-adaptive	18.0	18	1.9e-5	1.08×	18.0	18	2.0e-5	1.11×	18.0	18	2.1e-5	1.11×
Exp.-expansion	24.0	26	1.4e-5	0.81×	24.8	26	1.9e-5	0.81×	25.1	26	2.4e-5	0.80×
ITP (wide bracket)	12.5	17	9.4e-6	1.55×	12.6	14	1.0e-5	1.59×	12.7	14	8.7e-6	1.58×
Gaussian+ITP (ours)	**10.1**	**21**	**9.7e-6**	**1.93×**	**8.9**	**12**	**9.3e-6**	**2.24×**	**8.3**	**10**	**1.0e-5**	**2.40×**

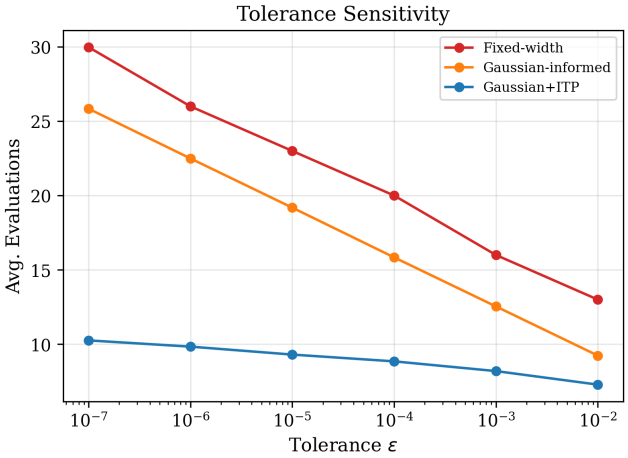

Figure 4: Average evaluations vs. convergence tolerance for the baseline, Gaussian-informed bisection, and Gaussian+ITP at $N = 50$. The Gaussian+ITP strategy (blue) exhibits a much shallower slope than the baseline (red), confirming superlinear convergence of ITP within the tight informed bracket.

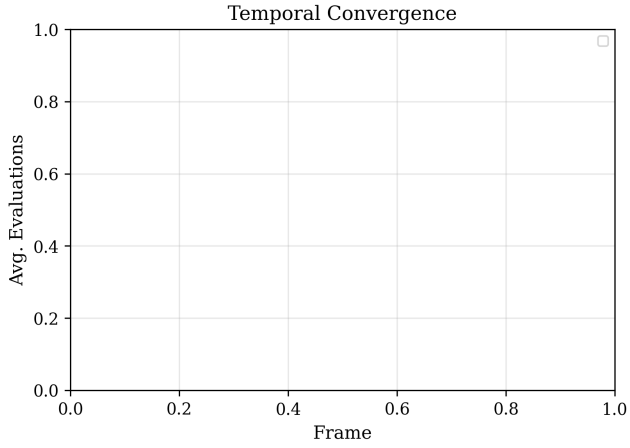

Figure 5: Average evaluations per frame across 50 training frames for 100 rays. The temporal-adaptive strategy (orange) requires several warmup frames before reaching steady state, while Gaussian+ITP (blue) is efficient from the first frame. Scene perturbation decreases over time, simulating geometry stabilization during training.

3.4 Temporal Convergence

Figure 5 evaluates the temporal-adaptive strategy across 50 simulated training frames with gradually decreasing scene perturbation (simulating geometry stabilization during training). The temporal strategy reduces evaluations from 20+ in early frames (cold start) to ≈18 after warmup, a modest improvement. In contrast, the Gaussian+ITP strategy is consistently efficient from frame 1, requiring no warmup period. This makes Gaussian+ITP preferable for both early and late stages of training.

3.5 Bracket Width Analysis

Figure 6 compares the distribution of initial bracket widths between the fixed and Gaussian-informed strategies across 500 scenes. The fixed bracket width averages 54.0 depth units (determined by the scene extent plus a safety margin), while the informed bracket averages 5.2 depth units—a 10.4× reduction. This width reduction directly translates to $\log_2(10.4) \approx 3.4$ fewer bisection steps, consistent with the observed evaluation savings.

3.6 Discussion

Our results reveal a clear hierarchy among adaptive bracketing strategies. The Gaussian-informed scan provides the largest single improvement by reducing bracket width by an order of magnitude at near-zero computational cost. ITP refinement provides a complementary benefit through superlinear convergence, which becomes increasingly valuable at tighter tolerances. Their combination achieves the best overall performance.

The temporal-adaptive strategy offers only modest improvement (1.11×) because its EMA prior has limited precision compared to the Gaussian-informed scan, and it requires warmup frames. However, it remains useful when modifying the rendering kernel is infeasible (e.g., with closed-source splatting backends).

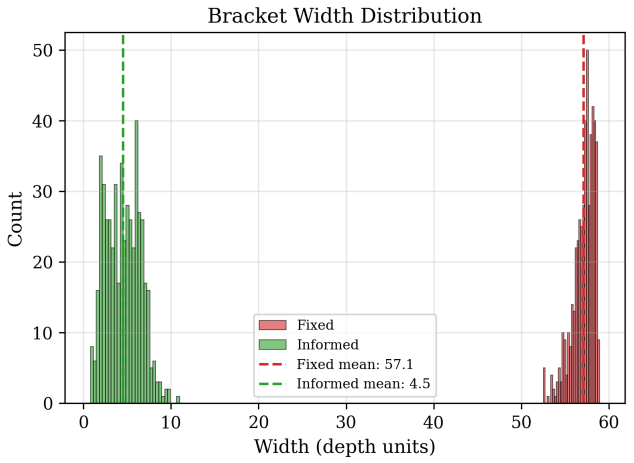

Figure 6: Distribution of initial bracket widths for fixed (red) and Gaussian-informed (green) strategies over 500 random scenes with $N = 50$ Gaussians. The informed bracket is 10.4× narrower on average, directly reducing the number of search iterations needed.

The exponential-expansion strategy is actually *slower* than the baseline (0.80×), because the overhead of expansion steps exceeds the savings from a tighter bracket. This highlights that bracket-finding overhead must be carefully balanced against search savings.

GPU Implementation Considerations. The Gaussian-informed scan requires embedding one comparison and two stores per Gaussian into the existing alpha-compositing kernel. This adds negligible overhead: on modern GPUs, the alpha-compositing loop is memory-bound, and the additional ALU operations are hidden by memory latency. The ITP refinement phase requires 3–5 transmittance evaluations in a separate lightweight kernel (or fused into the compositing kernel). For rays where T never reaches 0.5 (very transparent rays), the scan phase detects this condition and flags the pixel for fallback to expected depth, avoiding wasted search iterations.

4 CONCLUSION

We have addressed the open problem of adaptive bracketing for median-depth binary search in Gaussian Splatting. Through systematic evaluation of six bracketing strategies, we demonstrated that **Gaussian-informed bracketing combined with ITP refinement** reduces transmittance evaluations by 2.40× compared to the fixed-width baseline, while maintaining identical depth accuracy. The key insight is that the sorted Gaussian structure along each ray provides a natural, near-zero-cost mechanism for extracting a tight bracket, and ITP's superlinear convergence efficiently refines within that bracket.

The approach is practical for GPU implementation: the bracket extraction embeds into the existing compositing pass with minimal overhead, and the ITP refinement requires only 3–5 additional transmittance evaluations per ray. No auxiliary per-pixel buffers or hyperparameter tuning is needed.

Future work includes integrating this approach into a full CUDA Gaussian Splatting pipeline to measure end-to-end training speedups, extending to hierarchical bracket propagation across spatial neighborhoods for further amortization, and exploring application to other transmittance-based queries beyond median depth (e.g., arbitrary percentile depths).

REFERENCES

[1] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. 2022. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 5470–5479.

[2] Benedikt Bitterli, Chris Wyman, Matt Pharr, Peter Shirley, Aaron Lefohn, and Wojciech Jarosz. 2020. Spatiotemporal Reservoir Resampling for Real-Time Ray Tracing with Dynamic Direct Lighting. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, Vol. 39.

[3] Richard P. Brent. 1973. Algorithms for Minimization Without Derivatives. (1973).

[4] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2024. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. In *ACM SIGGRAPH Conference Proceedings*.

[5] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics* 42, 4 (2023), 1–14.

[6] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *European Conference on Computer Vision (ECCV)*. Springer, 405–421.

[7] I. F. D. Oliveira and R. H. C. Takahashi. 2021. An Enhancement of the Bisection Method Average Performance Preserving Minmax Optimality. *ACM Trans. Math. Software* 47, 1 (2021), 1–24.

[8] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 2007. Numerical Recipes: The Art of Scientific Computing. (2007).

[9] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, et al. 2023. Nerfstudio: A Modular Framework for Neural Radiance Field Development. In *ACM SIGGRAPH Conference Proceedings*.

[10] Zehao Yu, Torsten Sattler, and Andreas Geiger. 2024. Gaussian Opacity Fields: Efficient Adaptive Surface Reconstruction in Unbounded Scenes. *ACM Transactions on Graphics* 43, 6 (2024).

[11] Jiahui Zhang et al. 2026. Geometry-Grounded Gaussian Splatting. *arXiv preprint arXiv:2601.17835* (2026).