2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

These Magic Moments: Differentiable Uncertainty Quantification of Radiance Field Models

Parker Ewen* Hao Chen* Seth Isaacson Joey Wilson Katherine A. Skinner Ram Vasudevan

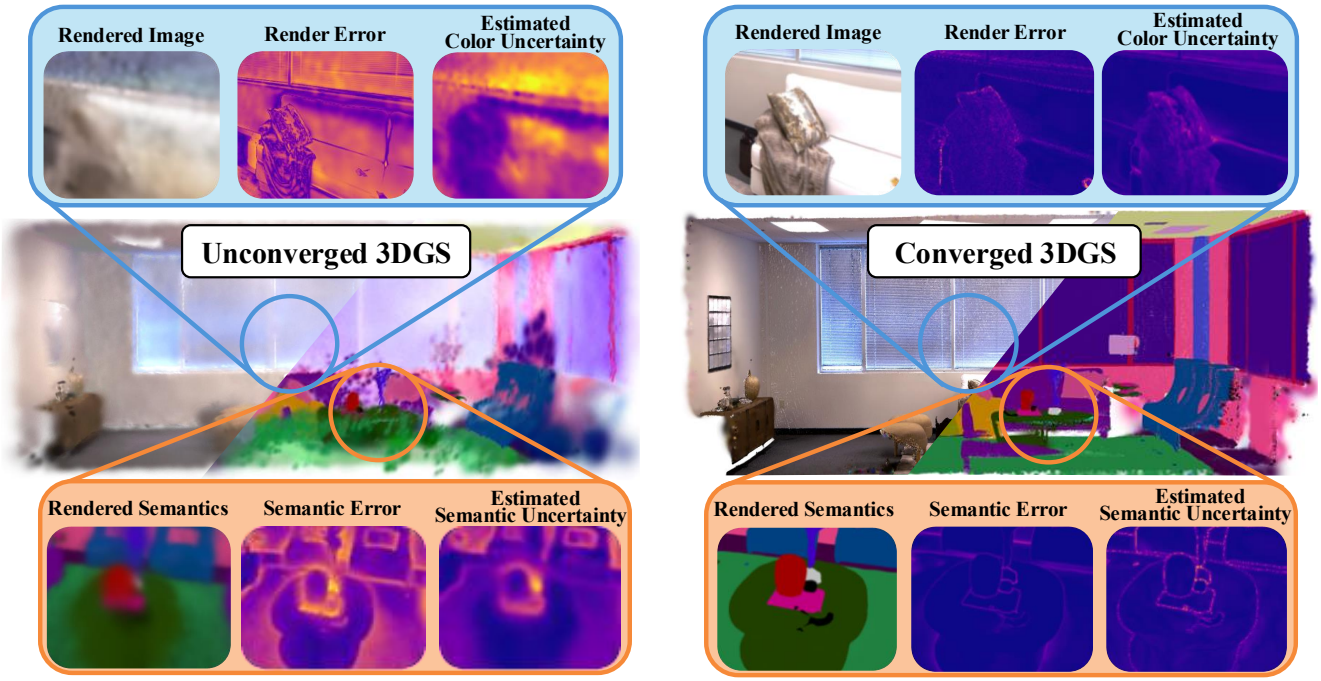

Fig. 1: Our proposed method quantifies uncertainty over rendered scenes generated through radiance field models by considering the differentiable higher-order moments of the rendering equation. The proposed approach requires no additional training, generalizes for color, depth, and semantics, and may be used for both NeRF and 3DGS methods. Shown is the variance, or second central moment, which is highly correlated to the rendering error for both color and semantics.

***Abstract*— Uncertainty quantification is crucial for autonomous systems, enabling safe and robust decision making in tasks ranging from active perception to robotic planning. This paper introduces a novel approach to quantify uncertainty for radiance fields by deriving pixel-wise moment expressions from the rendering equation. While radiance fields offer powerful scene representations, their high dimensionality and complexity have historically made uncertainty quantification computationally prohibitive for real-time applications. This paper demonstrates that the probabilistic nature of the rendering process enables efficient and differentiable computation of higher-order moments for radiance field outputs, including color, depth, and semantic predictions. The proposed method outperforms existing radiance field uncertainty estimation techniques while offering a more direct, computationally efficient, and differentiable formulation without the need for post-processing. Beyond uncertainty quantification, this paper also illustrates the utility of the proposed approach in downstream applications such as next-best-view (NBV) selection and active ray sampling for neural radiance field training. Extensive experiments on both synthetic and real-world scenes demonstrate state-of-the-art performance, confirming that principled uncertainty quantification can be seamlessly integrated into radiance field pipelines without sacrificing efficiency or accuracy.**

I. Introduction

Radiance field models have become fundamental tools for novel view synthesis, enabling the generation of photo-realistic images from different perspectives based on a set of input views [1]. These models excel at learning continuous volumetric scene representations, facilitating high-quality, view-consistent rendering beneficial for tasks including dense mapping, scene understanding, planning, and navigation [2]–[4]. Despite their widespread adoption, quantifying uncertainty in these models remains challenging. This limitation makes it difficult to determine when these models yield accurate scene representations, undermining their robustness and reliability, particularly for robotic applications that depend on trustworthy scene representation for safety-critical decision-making.

Uncertainty quantification in radiance field models provides valuable insight into regions of the environment that the model represents poorly and areas where additional data

*Denotes equal contribution.

Parker Ewen, Hao Chen, Seth Isaacson, Joey Wilson, Katherine A. Skinner and Ram Vasudevan are with the Department of Robotics, University of Michigan, Ann Arbor, MI 48109. {pewen, haochern, sethgi, wilsoniv, kskin, ramv}@umich.edu.

could improve model accuracy. This uncertainty can be leveraged for downstream applications such as active perception [5], for safety-critical systems [3], or for localization and mapping [6].

Existing uncertainty quantification approaches for radiance field models typically attempt to model uncertainty in the model parameters directly, and approximate model output uncertainty in the form of rendered uncertainty [7], [8]. However, the nonlinearity of the rendering function makes calculating pixel-wise probability densities from model parameter uncertainty mathematically intractable. Existing approaches therefore rely on approximations or heuristics to transform model parameter uncertainty into rendering uncertainty, including training neural networks to estimate uncertainty directly [5], [9], approximating uncertainty over a 3D deformation field in a post-hoc manner [8], or relying on simplifying approximations [10]. Furthermore, the complexity of these approximations often results in significantly increased computational costs.

To address these challenges, this paper presents a formulation for quantifying uncertainty of radiance field model outputs directly. We demonstrate that the probabilistic interpretation of the rendering equation from the underlying optical model enables efficient, differentiable computation of higher-order moments. This approach shows equivalent or superior performance on downstream tasks compared to existing uncertainty quantification techniques, which also proving to be more expensive to compute. The efficacy of this approach is demonstrated by first showing that the pixel-wise variance rendered using the proposed method has a higher correlation with rendering error than existing radiance field uncertainty quantification methods (Fig 1). We further demonstrate that our uncertainty measure is a useful signal for two downstream tasks: next-best-view selection and active ray selection for neural radiance field model training. In both tasks, the proposed approach outperforms baselines across all metrics.

In summary, the contributions of this paper are:

- Derivation of pixel-wise higher-order moments for radiance field model rendering;
- Demonstration that the proposed pixel-wise variance strongly correlates to rendering error for color, depth, and semantic rendering;
- Application of pixel-wise uncertainty as a signal for next-best-view selection that outperforms existing state-of-the-art approaches;
- Application of pixel-wise uncertainty for active ray selection that results in higher quality neural radiance field models.

The remainder of the paper is organized as follows: Section II presents an overview of uncertainty quantification in radiance field methods from the literature. Section III describes the probabilistic rendering process and the derivation of higher-order moments. Section IV demonstrates the efficacy of the proposed approach using three case studies: correlation with rendering error, next-best-view selection, and active ray sampling. Finally, Section V provides a discussion and concluding remarks.

II. Related Works

Uncertainty quantification plays a critical role in robot perception for enhancing model reliability and estimating prediction error [11]. Ideally, uncertainty quantification for radiance field models should correlate with rendering error and be applicable to color, depth, and semantics. Furthermore, computing the uncertainty should also be fast such that it can be used in downstream applications.

A. Uncertainty In NeRF Models

Neural radiance fields (NeRFs) use a neural network to approximate the 5-dimensional radiance field function [12] and are able to leverage existing uncertainty quantification methods for neural networks, including network dropout [13] and Bayesian neural networks (BNNs) [14], [15], two well-established techniques used to quantify uncertainty using Monte Carlo and functional approximations, respectively. While dropout has proven effective and offers a balance between computation cost and accuracy [16], BNNs can be computationally intensive to train, may not generalize under domain shift, and may have trouble converging [17]. Ensemble methods also approximate uncertainty, however, training a sufficient number of models is computationally intensive and inference time scales linearly with the number of models used [7], making ensemble methods unsuitable for many downstream tasks that require real-time operation.

Other prominent NeRF-based uncertainty quantification methods include S-NeRF [18] and CF-NeRF [9], which model uncertainty over the set of possible radiance field models via variational inference. Evidential regression [19] has also shown promise for uncertainty quantification by training a network to output an additional variance parameter [5], [9], although fundamental concerns about the accuracy of evidential learning underpinning these approaches have been raised [20], [21]. Lastly, post-hoc approaches that approximate the volumetric uncertainty of a NeRF model without modifying the training process have also been proposed [8]. It was found that approximating uncertainty of the NeRF network parameters directly was infeasible due to high correlation between the network layers, and a spatial distortion field was found to accurately approximate geometric uncertainty.

Depth-supervised NeRFs treat measured and rendered depth as probability distributions and leverage variance during training to improve convergence [22]–[24]. However, these methods do not assess model quality and cannot generalize to color or semantics. More generally, existing NeRF-based uncertainty quantification often requires training separate networks for color and semantics, estimating uncertainty independently per feature, and relies on slow, per-ray Monte Carlo approximations.

B. Uncertainty In 3DGS Models

Recently, explicit models using Gaussian basis functions to approximate the radiance field have gained popularity via

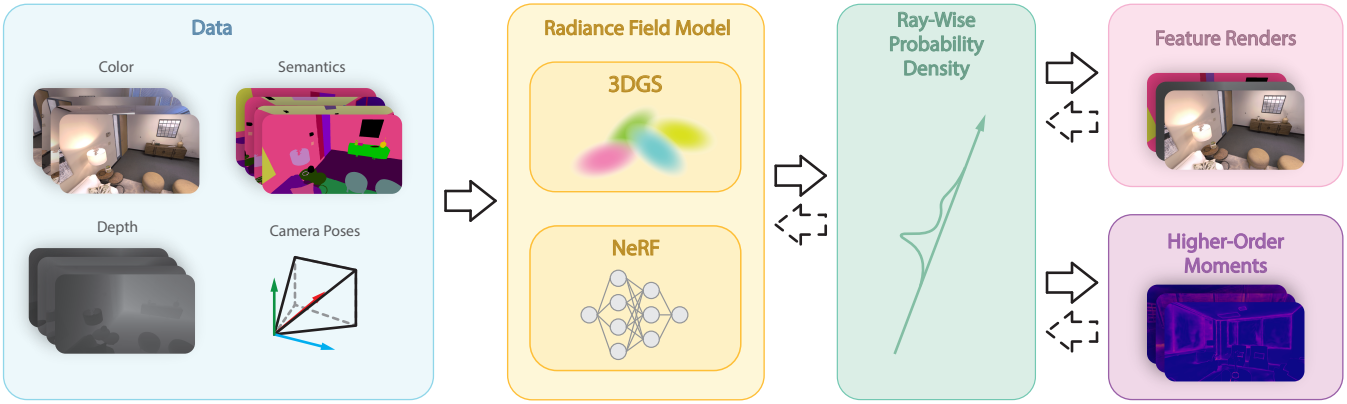

Fig. 2: Images containing color, depth, semantics, or other features of interest, along with camera poses, are used to train a radiance field model depicted. The rendered features (i.e. the expected feature of each pixel) are computed in a differentiable manner. Our proposed approach demonstrates that the same probabilistic process used to compute these renders can be used to differentiably compute the higher-order moments. Solid arrows indicate operation flow and dotted arrows indicate gradient flow.

3D Gaussian Splatting (3DGS) [25]. They have shown state-of-the-art performance in novel view synthesis and scene representation [26] and uncertainty quantification methods for 3DGS models have recently emerged. Similar to NeRF-based approaches, variational inference has also been used for uncertainty quantification in 3DGS models [27]. These variational inference methods again do not generalize to color, depth, and semantics, and require estimating uncertainty of each feature independently of the others. Semantic uncertainty quantification for closed-set semantics has also been demonstrated for 3DGS models [28]. This method uses the conjugate pair theorem [29] to track uncertainty over semantic estimates.

FisherRF [10], a recent uncertainty quantification method for next-best-view selection, takes an information-theoretic approach to estimating uncertainty in the Gaussians' parameters, then renders a heuristic approximation of this uncertainty. Unlike NeRF-based approaches, it was found that the parameters in the 3DGS model have minimal correlation [10], and second-order Laplace estimates are used to capture this parametric uncertainty. Notably, this information-theoretic uncertainty quantification using Fisher information requires a linear rendering model. As the Gaussian rasterization process is nonlinear, a Gauss-Newton approximation is made, which often over-approximates uncertainty when the number of Gaussian bases is not regularized. Other methods use heuristics for uncertainty and leverage these heuristics for training and regularization [30], [31]. While FisherRF is useful for color and depth uncertainty, it shows low correlation with rendering error in unbounded scenes and lacks semantic uncertainty quantification.

In contrast to these approaches, this paper presents a general, unified approach for estimating the uncertainty in radiance field model outputs by considering the rendering equation as a probabilistic process. The proposed approach works with both NeRF and 3DGS models and is able to render color, depth, and semantic uncertainty without any additional training. Furthermore, we demonstrate that the uncertainty correlates strongly with the rendering error and outperforms state-of-the-art uncertainty quantification methods.

III. Methodology

This section presents an overview of the proposed uncertainty quantification approach for radiance field models. First, a review of radiance fields and the associated rendering equation are presented in Section III-A. Then, the rendering equation is re-examined through the lens of a probabilistic process using the local illumination optical model. Finally, this probabilistic process is used to derive a means of rendering higher-order moments and variance using the NeRF and 3DGS radiance field models.

A. Radiance Fields

A radiance field models the image formation process by approximating a five-dimensional function that maps from a point and viewing direction to the density $\sigma : \mathbb{R}^3 \to \mathbb{R}$ and the color $c : \mathbb{R}^3 \times S^2 \to \mathbb{R}^3$ along the viewing ray. Such radiance field models have also been extended to represent depth [32] and semantics [2]. Under a radiance field model, an image is rendered by considering separately the ray that intersects the camera origin and each pixel. Consider such a ray $r : \mathbb{R} \to \mathbb{R}^3$. We define an additional quantity, the *transmittance*, which defines the unnormalized probability that a particle traveling along ray r from location $s = a$ to $s = b$ will experience a collision. This transmittance, denoted as $\mathcal{T}_a^b[r]$, is computed by

$$\mathcal{T}_a^b[r] = \exp\left(-\int_a^b \sigma(r(s))ds\right) \tag{1}$$

From the transmittance, the quantity of interest $\rho : \mathbb{R}^3 \to \mathbb{R}$, representing either color, depth, or semantics, of the pixel corresponding to ray r is computed as

$$I[\rho] = \int_{z_n}^{z_f} \rho(r(s))\sigma(r(s))\mathcal{T}_{z_n}^{s}[r]ds \tag{2}$$

where z_n, z_f are the near and far clipping planes.

B. Uncertainty Quantification of Radiance Fields

Early work in optical models considered volume rendering as a probabilistic process [33], [34]. The theory underpinning current radiance field approaches uses the so-called local illumination model, where a scene is comprised of groups of small, non-reflecting, opaque, glowing particles. When rendering using this model, a ray is cast through this sea of particles and, when a ray collides with a particle at location s along the ray, the color of the rendered pixel is the color emitted by the particle at location s. An equivalent process may be defined for rendering depth and semantics.

These optical models treat the locations of these particles as uncertain. The mass density of these particles is represented as σ, termed the density in the radiance field literature. Ray collision is then, consequently, a random event, meaning the rendered quantity is a random variable [34].

Under this probabilistic definition, $\sigma(r(s))$ represents the probability that a ray collides with a particle within the infinitesimal length $[s, ds]$ along the ray. The transparency, or transmittance, $\mathcal{T}_0^s[r]$, then represents that probability that traveling from point 0 to s along the ray does not collide with a particle. As these are independent events, the probability that a ray does not hit a particle between 0 and s, and then hits one between distances s and ds, is the multiplication of both probabilities, $\mathcal{T}_0^s[r]\sigma(r(s))$ [35]. Finally, $\mathcal{T}_0^s[r]\sigma(r(s))$ defines the probability density of particle collision along the ray taking the limit as $ds \to 0$. Under these considerations, the rendering equation (2) represents the expected value of the random variable of interest, ρ, be it color, depth, or semantics [34].

Extending this further, we propose computing the j-th moment for each pixel-wise random variable:

$$\mathbb{E}[\rho^j] = \int_{z_n}^{z_f} \rho(r(s))^j \sigma(r(s)) T_{z_n}^s[r] ds. \tag{3}$$

The variance of each pixel is then computed using (2) and (3):

$$\mathrm{Var}[\rho] = \mathbb{E}[\rho^2] - \mathbb{E}[\rho]^2. \tag{4}$$

In computer vision, the terms "variance" and "uncertainty" are often taken as equivalent statements [19], [28], [36]. We will refer to the variance as the quantified uncertainty for the remainder of this paper.

Similar formulas exist for other higher-order central moments if another another metric for uncertainty is desired. Notably, the moment and central moment computations make no assumptions on the form of the random variable. It is possible to use these moments to fit a distribution to each pixel-wise random variable via moment fitting [37].

C. Uncertainty Computation

Two prominent models have emerged for radiance field estimation. The first is neural radiance fields, which use a feed-forward network to approximate features, ρ, and density, σ, of the radiance field function. The second is 3D Gaussian Splatting models, which use a collection of Gaussian basis functions with differentiable components to represent the radiance field quantities. Below we describe how to compute the moments for each model. Figure 2 shows the algorithm for uncertainty quantification using a 3DGS model.

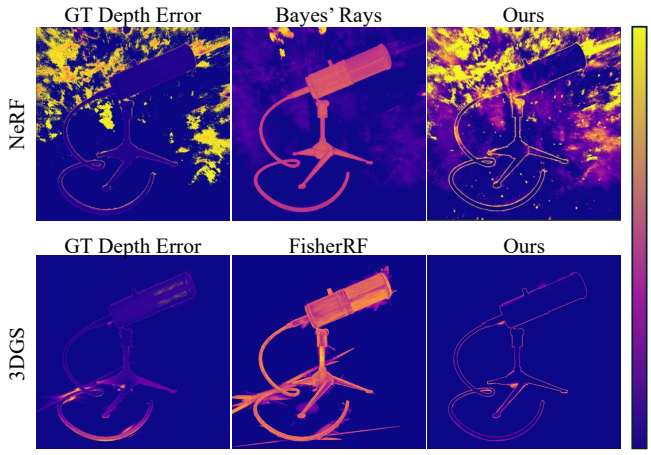

Fig. 3: Uncertainty quantification for depth estimates of NeRF and 3DGS models. We compare our proposed variance estimate to Bayes' Rays [8] for a trained NeRF model, and to FisherRF [10] for a trained 3DGS model. Unlike existing methods, our approach works with both NeRF and 3DGS models. CF-NeRF [38] and 3DGS Ensemble [7] are omitted, as they require specialized models.

1) Neural Radiance Fields: Neural radiance fields jointly estimate the features, ρ, and density, σ, of the radiance field at discrete points along a ray cast from the camera center and through a pixel. For volumetric rendering (2), the ray is first divided into N evenly spaced bins and one sample, t_i, is drawn uniformly from within each bin. These samples are then used to approximate (2) via Monte Carlo integration such that, for each feature channel:

$$\mathbb{E}[\rho] = \sum_{i=0}^{N} T_i c_i (1 - \exp(-\sigma_i \delta_i)), \tag{5}$$

where $T_i = \exp(-\sum_{k=1}^{i-1} \sigma_k \delta_k)$, and $\delta_i = t_{i+1} - t_i$ is the distance between two adjacent samples. To compute the moments of feature ρ, the same Monte Carlo Integration technique is applied to (3) such that:

$$\mathbb{E}[\rho^j] = \sum_{i=0}^{N} T_i c_i^j (1 - \exp(-\sigma_i \delta_i)). \tag{6}$$

2) 3D Gaussian Splatting: Recently, 3D Gaussian Splatting (3DGS) models have shown fast and efficient performance in representing radiance fields using 3D differentiable Gaussian basis functions. Each 3D Gaussian consists of an opacity, α, mean and covariance of the Gaussian, μ and Σ, as well as a color or semantics, here represented as the learned feature ρ [25, Section 2.3]. We combine all parameters together such that $\boldsymbol{w} = \{\alpha_i, \mu_i, \Sigma_i, \rho_i\}_{i=1}^M$, where M is the number of Gaussians in the splatting model.

Unlike neural radiance fields, Monte Carlo Integration along rays is not required for Gaussian Splatting as the 3D Gaussians are projected onto the image plane and subsequent pixel colors, depths, and semantics are computed. Given a camera pose $(\boldsymbol{x}, \boldsymbol{d}) \in SE(3)$ consisting of position $\boldsymbol{x} \in \mathbb{R}^3$

TABLE I: Correlation coefficients for color variance and rendered color error. The CF-NeRF baseline did not converge on the TUM dataset and therefore no correlation coefficients are given. Presented is the time to render uncertainty for a single image. Cells are highlighted with red, orange, and yellow for for the first, second, and third best performance, respectively. Higher is better for the correlation coefficients.

Method	Blender				Mip360				TUM Dataset			
	τ_p	τ_s	τ_k	Time [ms]	τ_p	τ_s	τ_k	Time [ms]	τ_p	τ_s	τ_k	Time [ms]
FisherRF [10]	0.580	0.215	0.137	13679.3	0.527	-0.007	-0.005	156071.0	0.482	-0.071	-0.046	46769.2
CF-NeRF [38]	0.307	0.612	0.547	180718.0	0.113	0.187	0.130	5329000.0	-	-	-	-
3DGS Ensemble [7]	0.663	0.816	0.712	29.0	0.398	0.401	0.276	79.1	0.434	0.533	0.374	237.4
Ours	0.716	0.838	0.716	2.0	0.885	0.885	0.695	6.8	0.781	0.778	0.577	20.4

TABLE II: Correlation coefficients for depth variance and rendered depth error. The CF-NeRF baseline did not converge on the TUM dataset and therefore no correlation coefficients are given. Presented is the time to render uncertainty for a single image. Cells are highlighted with red, orange, and yellow for for the first, second, and third best performance, respectively. Higher is better for the correlation coefficients.

Method	Blender				TUM Dataset			
	τ_p	τ_s	τ_k	Time [ms]	τ_p	τ_s	τ_k	Time [ms]
FisherRF [10]	0.572	0.328	0.238	13679.3	0.546	0.575	0.405	46769.2
Bayes' Rays [8]	0.758	0.473	0.364	45220.6	0.596	0.638	0.468	46399.2
CF-NeRF [38]	-0.086	-0.305	-0.346	180718.0	-	-	-	-
3DGS Ensemble [7]	0.362	0.502	0.366	29.0	0.204	0.213	0.150	237.4
Ours	0.718	0.792	0.700	2.0	0.671	0.685	0.519	20.4

and viewing direction $\boldsymbol{d} \in SO(3)$, the rendering function (2) for a given feature channel of a pixel, ρ, in the image plane is

$$\mathbb{E}[\rho] = \sum_{i \in \mathcal{I}} \rho_i \tilde{\alpha}_i(\boldsymbol{x}, \boldsymbol{d}) \prod_{k<i} (1 - \tilde{\alpha}_k(\boldsymbol{x}, \boldsymbol{d})) \tag{7}$$

where ρ_i is either the color, depth, or semantic embedding of the i-th ordered Gaussian and

$$\tilde{\alpha}_i(\boldsymbol{x}, \boldsymbol{d}) = \alpha_i \cdot \exp\left(-\frac{1}{2} \left\|\boldsymbol{x} - \boldsymbol{\mu}_i(\boldsymbol{d})\right\|^2_{\Sigma_i^{-1}(\boldsymbol{d})}\right) \tag{8}$$

is the pixel-wise opacity of the i-th ordered Gaussian in set $\mathcal{I}$. Here, the set $\mathcal{I}$ is constructed by sorting the projected Gaussians by depth with respect to the image plane, and $\boldsymbol{\mu}_i$ and Σ_i are the marginalized 2D mean and variance of the i-th splatted Gaussian. As discussed in [25], the α-blending in (7) saturates once $\sum_{i \in \mathcal{I}} \alpha_i = 1$ and all remaining Gaussians in the ordered set $\mathcal{I}$ are not considered. Notably, the ordering of the Gaussian splats is approximated because the rotation of each 3D Gaussian is neglected [25].

This splatting process may also be interpreted as marginalizing the Gaussian basis functions onto the ray creating a Gaussian Mixture Model. The weights of each mode in the GMM correspond to the opacities computed in the alpha compositing process (Fig. 2). Using the 3DGS model, the moments for feature ρ are computed from (3) as:

$$\mathbb{E}[\rho^j] = \sum_{i \in \mathcal{I}} \rho_i^j \tilde{\alpha}_i(\boldsymbol{x}, \boldsymbol{d}) \prod_{k<i} (1 - \tilde{\alpha}_k(\boldsymbol{x}, \boldsymbol{d})). \tag{9}$$

IV. Experiments

This section demonstrates the utility of the proposed uncertainty quantification scheme across two case studies and validates the variance as a reasonable choice to quantify the uncertainty. First, we demonstrate that the variance renders strongly correlate with the rendering error. Next, we show that variance renders using our proposed approach provide a better Next-Best-View (NBV) selection criteria than existing state-of-the-art methods. Additional implementation details and qualitative comparisons are presented in an online appendix on the project webpage.

A. Error Correlation

We determine the relationship between our proposed variance approach and the rendering error for color, depth, and semantics. A correlation between variance and error indicates that an approach correctly models uncertainty and highlights regions poorly represented by the radiance field.

In this case study, we compute the Pearson correlation coefficient τ_p [40], Spearman rank-order coefficient τ_s [41], and Kendall tau coefficient τ_k [42]. These coefficients each model the strength and direction of correlation using different metrics, with the Pearson coefficient being the most robust to outliers. A coefficient of 1 indicates a strong correlation, meaning regions of high rendered variance correspond to regions of high rendering error. A coefficient of -1 indicates a strong inverse correlation, and a coefficient of 0 indicates no correlation.

We compare our proposed variance rendering approach with Bayes' Rays [8], FisherRF [10], CF-NeRF [38], and 3DGS Ensembles [7], four state-of-the-art methods that approximate model output uncertainty. The 3DGS Ensemble approach is the 3DGS equivalent of NeRF ensembles [7] and can be interpreted as a Monte Carlo Sampling-based approach for computing the integral in (4).

Color: The correlation coefficients for the Blender [12], Mip360 [43], and TUM [44] datasets for all methods are presented in Table I. As shown in Table I, our method outperforms the FisherRF and CF-NeRF baselines in terms of correlation and runtime for every dataset and has comparable performance to the 3DGS Ensemble in terms of correlation but requires an order-of-magnitude less time to run.

Depth: The depth correlation coefficients are presented in Table II. Here, we see the proposed method outperforms all

TABLE III: NBV selection on the Blender dataset with a total of 12 views and the TUM dataset with a total of 300 views. Times correspond to choosing the first NBV in the experiment. Times for the random baseline are excluded since it does not consider uncertainty and relies solely on a stochastic process. Cells are highlighted with red, orange, and yellow for for the first, second, and third best performance, respectively.

Method	Blender				TUM Dataset			
	PSNR (↑)	SSIM (↑)	LPIPS (↓)	Time [s]	PSNR (↑)	SSIM (↑)	LPIPS (↓)	Time [s]
Random	23.072	0.998	0.162	-	19.175	0.703	0.382	-
ActiveNeRF [39]	14.115	0.806	0.280	1014.	10.834	0.428	0.613	1272.8
FisherRF [10]	23.658	0.998	0.148	13.8	18.978	0.700	0.388	47.4
Ours (Depth Var.)	22.867	0.997	0.185	0.6	18.946	0.697	0.388	0.9
Ours (Color Var.)	23.944	0.998	0.128	0.6	19.191	0.703	0.382	0.9

baselines across both datasets in terms of both correlation and runtime. Figure 3 shows a qualitative example of the ground truth depth error associated with a NeRF and 3DGS model and the associated uncertainty quantifications for the Bayes' Rays and FisherRF baselines and our proposed approach. Bayes' Rays and our proposed approach estimate uncertainty using the same trained NeRF model. Likewise, FisherRF and our proposed approach estimate uncertainty using the same trained 3DGS model. Notably, Bayes' Rays and FisherRF overestimate uncertainty for the object while our approach estimates uncertainty over the entire scene and at the object boundaries where the depth error is highest.

Semantics: We evaluate semantic correlation results of the proposed approach using the Mip360 datasets, obtaining correlation coefficients of $\tau_p = 0.180$, $\tau_s = 0.200$, and $\tau_k = 0.135$. The ground truth visual-language embeddings are computed by passing the training RGB images through the CLIP-LSeg network [45]. Intuitively, conflicting visual-language embeddings will result in high variance. We note that the rendered language features are compressed via an encoding network in [46], and so to compare the rendered feature vectors with the LSeg outputs, principle component analysis (PCA) is run to extract the first 128 principle components. As there are no baselines for computing variance for language embeddings, only the results for the proposed approach are shown. Furthermore, the ensemble method used for color and depth uncertainty quantification is infeasible for semantic uncertainty quantification due to the prohibitively expensive memory resources required to store the semantic feature vectors. The correlation between semantic error and variance is noticeably worse than for depth and color. We postulate that this is due to the noise inherent in the CLIP-LSeg model, where visual-language features vary for the same location depending on the viewpoint.

Latency: Lastly, we highlight the significantly reduced runtime of our proposed approach in comparison to the baselines as seen in Tables I and II. Our method computes variance directly from a radiance field model, in contrast to Cf-NeRF [38] and 3DGS Ensembles [7], and requires no additional post-processing, in contrast to FisherRF [10] and Bayes' Rays [8], without any additional memory cost.

B. Next-Best-View Selection

The Next-Best-View (NBV) problem (Fig. 5) refers to the task of determining the optimal subsequent viewpoint that maximizes the information gained, often in the context of

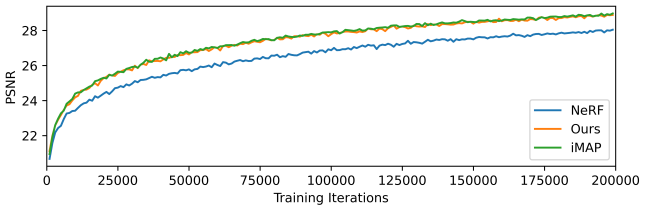

Fig. 4: Average rendering quality on the Blender dataset using active ray selection. For the NeRF baseline, rays are selected uniformly over the entire image. iMAP [47] and the proposed approach sample for high error and high variance regions of the image, respectfully. The proposed approach matches the performance of iMAP without needing to reference the ground-truth image.

3D scene reconstruction, object recognition, or inspection tasks. Existing methods for NBV estimate the uncertainty or entropy of the model parameters directly [8], [10], and then approximate model output uncertainty as a function of parameter uncertainty. The viewpoint with the highest model parameter uncertainty is picked and added to the training set, thus improving scene completeness or reconstruction accuracy by providing more information about these parameters.

We demonstrate that computing model output uncertainty directly is a stronger selection criteria for NBV selection by comparing this approach to state-of-the-art NBV methods for radiance fields. Following [10], the viewpoint with the highest uncertainty is selected as the NBV.

We compare our proposed variance rendering approach with ActiveNeRF [39] and FisherRF [10], two state-of-the-art methods that approximate model output uncertainty, and a random baseline. Notably, FisherRF [10] requires iterating over the entire training set before uncertainty is estimated for a novel viewpoint. The Blender and TUM datasets are used for this experiment. For the Blender dataset, the 3DGS model is initialized by uniformly sampling Gaussians within a cube with randomized RGB values and 60% default opacities. The model is initially trained from 9 adjacent views, and 3 new viewpoints are chosen at 1000, 3000, and 5000 iterations, respectively. Model performance is evaluated after 10000 training iterations.

For the more complex TUM dataset, the model is initially trained on 150 images. A new image is then added every 100 iterations until a total of 300 images are selected. Model performance is evaluated after 20000 training iterations. The results for both NBV experiments are presented in Table III and assess the PSNR, SSIM, and LPIPS metrics alongside

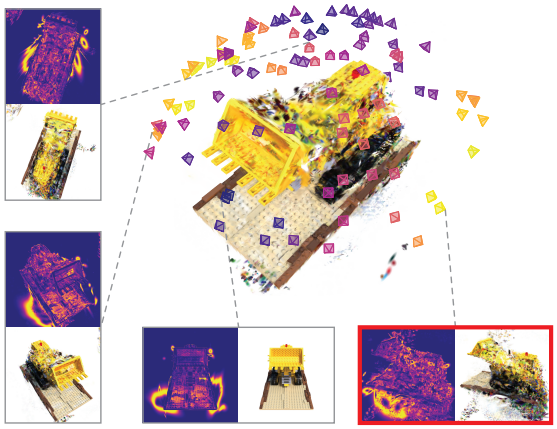

Fig. 5: Given a radiance field model trained from few views, the goal is to select the next best view (NBV) that maximally improves the model. Rendered variance for four views is shown, with the highlighted viewpoint chosen as NBV because it has the highest pixel-wise variance.

the computation time to pick an NBV.

From these results, selecting NBV with the proposed color variance outperforms the baselines on both datasets. For the Blender dataset, the intricate geometric detail translates to high entropy using FisherRF [10], resulting in a diverse set of NBV selections. In contrast, the TUM dataset contains regions of homogeneity, such as desks and walls, which translate into low entropy using the method presented, FisherRF. As a result, the viewpoints selected by this method are therefore less informative for the TUM dataset as they prioritize close-up views of small objects with intricate geometry within the room that have high entropy. Notably, the random baseline performs surprisingly well across all experiments, with exceptional performance in the TUM dataset.

The proposed variance rendering method estimates uncertainty of the model output directly, choosing viewpoints with high variance directly without estimating model parameter uncertainty. We propose that the reason the color variance is a better selection criteria than depth variance for the NBV task is that the evaluation criteria, PSNR, SSIM, and LPIPS, are representative of the quality of the color render. We highlight that our approach renders variance images faster than FisherRF [10] which requires the cumulative entropy over all training images to be evaluated.

C. Uncertainty-Aware Ray Sampling

In NeRF training, rendering and optimizing over all pixels at each iteration is prohibitively expensive. Active sampling has emerged as an effective strategy to concentrate resources on the most informative regions, accelerating training and improving final reconstruction quality.

We propose a novel active sampling approach using the proposed variance rendering scheme in (6). Similar to iMAP [47] which is a state-of-the-art approach for active ray sampling, we compute the variance for each cell of a grid spanning the image and compute the average ray variance within each cell. In contrast, iMAP uses the average rendering error for each cell. More training rays are then sampled from cells with higher variance. For additional implementation details, refer to the project webpage.

We compare the proposed approach against the original NeRF and iMAP implementations on the Blender dataset. As seen in Figure 4, our proposed approach matches the state-of-the-art performance of iMAP while only requiring variance estimates along each ray rather than the rendering error. Both methods significantly improve the final rendering quality when compared to the uniform ray sampling in the baseline NeRF implementation. This further solidifies the correlation between the rendering error, which is used for ray selection in iMAP, and variance, used in the proposed method, and demonstrates the generality of our approach for both 3DGS and NeRF models.

V. Conclusion

This paper presents a method for rendering pixel-wise uncertainty and entropy by interpreting the rendering equation as a probabilistic process. This enables the computation of higher-order moments and entropy for each pixel in the rendered image with minimal computational overhead. The proposed approach has a stronger correlation with rendering error than existing uncertainty quantification methods, indicating improved identification of poorly represented regions. Furthermore, the rendered variance is a better selection criterion than existing NBV approaches, while still operating in real time. These characteristics make our method particularly beneficial for uncertainty-sensitive robotic tasks demanding real-time response. Computing the true joint probability, and resultant uncertainty, is an area of future work.

References

[1] A. Rabby and C. Zhang, "Beyondpixels: A comprehensive review of the evolution of neural radiance fields," *arXiv preprint arXiv:2306.03000*, 2023.

[2] M. Qin, W. Li, J. Zhou, H. Wang, and H. Pfister, "Langsplat: 3d language gaussian splatting," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 20 051–20 060.

[3] J. Michaux, S. Isaacson, C. E. Adu, A. Li, R. K. Swayampakula, P. Ewen, S. Rice, K. A. Skinner, and R. Vasudevan, "Let's make a splan: Risk-aware trajectory optimization in a normalized gaussian splat," *arXiv preprint arXiv:2409.16915*, 2024.

[4] T. Chen, P. Culbertson, and M. Schwager, "Catnips: Collision avoidance through neural implicit probabilistic scenes," *IEEE Transactions on Robotics*, 2024.

[5] L. Jin, X. Chen, J. Rückin, and M. Popović, "Neu-nbv: Next best view planning using uncertainty estimation in image-based neural rendering," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 11 305–11 312.

[6] H. Matsuki, R. Murai, P. H. Kelly, and A. J. Davison, "Gaussian splatting slam," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 18 039–18 048.

[7] N. Sünderhauf, J. Abou-Chakra, and D. Miller, "Density-aware NeRF ensembles: Quantifying predictive uncertainty in neural radiance fields," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 9370–9376.

[8] L. Goli, C. Reading, S. Sellán, A. Jacobson, and A. Tagliasacchi, "Bayes' Rays: Uncertainty quantification for neural radiance fields," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 20 061–20 070.

[9] J. Shen, A. Agudo, F. Moreno-Noguer, and A. Ruiz, "Conditional-Flow NeRF: Accurate 3D modelling with reliable uncertainty quantification," in *European Conference on Computer Vision*. Springer, 2022, pp. 540–557.

[10] W. Jiang, B. Lei, and K. Daniilidis, "FisherRF: Active view selection and uncertainty quantification for radiance fields using fisher information," *arXiv preprint arXiv:2311.17874*, 2023.

[11] R. Szeliski, "Bayesian modeling of uncertainty in low-level vision," *International Journal of Computer Vision*, vol. 5, no. 3, pp. 271–301, 1990.

[12] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[13] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*. PMLR, 2016, pp. 1050–1059.

[14] I. Kononenko, "Bayesian neural networks," *Biological Cybernetics*, vol. 61, no. 5, pp. 361–370, 1989.

[15] V. Mullachery, A. Khera, and A. Husain, "Bayesian neural networks," *arXiv preprint arXiv:1801.07710*, 2018.

[16] N. Srivastava, "Improving neural networks with dropout," *University of Toronto*, vol. 182, no. 566, p. 7, 2013.

[17] P. Izmailov, S. Vikram, M. D. Hoffman, and A. G. G. Wilson, "What are Bayesian neural network posteriors really like?" in *International conference on machine learning*. PMLR, 2021, pp. 4629–4640.

[18] J. Shen, A. Ruiz, A. Agudo, and F. Moreno-Noguer, "Stochastic Neural Radiance Fields: Quantifying uncertainty in implicit 3D representations," in *2021 International Conference on 3D Vision (3DV)*. IEEE, 2021, pp. 972–981.

[19] A. Amini, W. Schwarting, A. Soleimany, and D. Rus, "Deep evidential regression," *Advances in neural information processing systems*, vol. 33, pp. 14 927–14 937, 2020.

[20] Y. Wu, B. Shi, B. Dong, Q. Zheng, and H. Wei, "The evidence contraction issue in deep evidential regression: Discussion and solution," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 19, 2024, pp. 21 726–21 734.

[21] N. Meinert, J. Gawlikowski, and A. Lavin, "The unreasonable effectiveness of deep evidential regression," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 8, 2023, pp. 9134–9142.

[22] K. Deng, A. Liu, J.-Y. Zhu, and D. Ramanan, "Depth-supervised NeRF: Fewer views and faster training for free," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022.

[23] K. Rematas, A. Liu, P. P. Srinivasan, J. T. Barron, A. Tagliasacchi, T. Funkhouser, and V. Ferrari, "Urban radiance fields," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[24] S. Isaacson, P.-C. Kung, M. Ramanagopal, R. Vasudevan, and K. A. Skinner, "Loner: Lidar only neural representations for real-time slam," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8042–8049, 2023.

[25] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian Splatting for real-time radiance field rendering." *ACM Trans. Graph.*, vol. 42, no. 4, pp. 139–1, 2023.

[26] G. Chen and W. Wang, "A survey on 3D Gaussian Splatting," *arXiv preprint arXiv:2401.03890*, 2024.

[27] L. Savant, D. Valsesia, and E. Magli, "Modeling uncertainty for Gaussian Splatting," *arXiv preprint arXiv:2403.18476*, 2024.

[28] J. Wilson, M. Almeida, M. Sun, S. Mahajan, M. Ghaffari, P. Ewen, O. Ghasemalizadeh, C.-H. Kuo, and A. Sen, "Modeling uncertainty in 3d gaussian splatting through continuous semantic splatting," *arXiv preprint arXiv:2411.02547*, 2024.

[29] P. Ewen, A. Li, Y. Chen, S. Hong, and R. Vasudevan, "These maps are made for walking: Real-time terrain property estimation for mobile robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7083–7090, 2022.

[30] S. Kheradmand, D. Rebain, G. Sharma, W. Sun, J. Tseng, H. Isack, A. Kar, A. Tagliasacchi, and K. M. Yi, "3d Gaussian Splatting as Markov Chain Monte Carlo," *arXiv preprint arXiv:2404.09591*, 2024.

[31] J. Hu, X. Chen, B. Feng, G. Li, L. Yang, H. Bao, G. Zhang, and Z. Cui, "CG-SLAM: Efficient dense RGB-D SLAM in a consistent uncertainty-aware 3d Gaussian field," *arXiv preprint arXiv:2403.16095*, 2024.

[32] J. Chung, J. Oh, and K. M. Lee, "Depth-regularized optimization for 3d gaussian splatting in few-shot images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 811–820.

[33] N. Max, "Optical models for direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 1, no. 2, pp. 99–108, 1995.

[34] N. Max and M. Chen, "Local and global illumination in the volume rendering integral," Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), Tech. Rep., 2005.

[35] A. Tagliasacchi and B. Mildenhall, "Volume rendering digest (for nerf)," *arXiv preprint arXiv:2209.02417*, 2022.

[36] J. Wilson, Y. Fu, J. Friesen, P. Ewen, A. Capodieci, P. Jayakumar, K. Barton, and M. Ghaffari, "Convbki: Real-time probabilistic semantic mapping network with quantifiable uncertainty," *IEEE Transactions on Robotics*, 2024.

[37] P. Billingsley, *Convergence of probability measures*. John Wiley & Sons, 2013.

[38] Q. Yan, Q. Wang, K. Zhao, J. Chen, B. Li, X. Chu, and F. Deng, "CF-NeRF: Camera parameter free neural radiance fields with incremental learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 6, 2024, pp. 6440–6448.

[39] X. Pan, Z. Lai, S. Song, and G. Huang, "ActiveNeRF: Learning where to see with uncertainty estimation," in *European Conference on Computer Vision*. Springer, 2022, pp. 230–246.

[40] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," *Noise reduction in speech processing*, pp. 1–4, 2009.

[41] P. Sedgwick, "Spearman's rank correlation coefficient," *Bmj*, vol. 349, 2014.

[42] P. K. Sen, "Estimates of the regression coefficient based on kendall's tau," *Journal of the American statistical association*, vol. 63, no. 324, pp. 1379–1389, 1968.

[43] J. T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, and P. P. Srinivasan, "Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 5855–5864.

[44] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *Proc. of the International Conference on Intelligent Robot Systems (IROS)*, Oct. 2012.

[45] B. Li, K. Q. Weinberger, S. Belongie, V. Koltun, and R. Ranftl, "Language-driven semantic segmentation," *arXiv preprint arXiv:2201.03546*, 2022.

[46] S. Zhou, H. Chang, S. Jiang, Z. Fan, Z. Zhu, D. Xu, P. Chari, S. You, Z. Wang, and A. Kadambi, "Feature 3dgs: Supercharging 3d gaussian splatting to enable distilled feature fields," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 676–21 685.

[47] E. Sucar, S. Liu, J. Ortiz, and A. J. Davison, "imap: Implicit mapping and positioning in real-time," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 6229–6238.