A. Proof of the main statement

In this section we provide the proof that the expected value of the variance of the scores is equivalent to the expected value of the second derivative of the noising distribution $q_t(\mathbf{X}_t) = \int p(\mathbf{X}) q_t(\mathbf{X}_t|\mathbf{X}) dX$

$$\mathbb{E} \quad \left[\left(\frac{\partial}{\partial \mathbf{X}_t} \log q(\mathbf{X}_t) \right) \left(\frac{\partial}{\partial \mathbf{X}_t} \log q(\mathbf{X}_t) \right)^\top \right] = (1)$$

$$= \quad -\mathbb{E} \left[\frac{\partial^2}{\partial \mathbf{X}_t \partial \mathbf{X}_t^\top} \log q(\mathbf{X}_t) \right]. \tag{2}$$

In the main text we use this result to gain insight about our uncertainty estimates, which approximate the expected value of the variance of the scores with a Monte Carlo estimate i.e.

$$\mathbf{U}_t \quad = \quad \text{diag} \left(\left(E_t - \bar{E}_t \right)^T \left(E_t - \bar{E}_t \right) \right) \tag{3}$$

$$\approx \quad \mathbb{E} \left[\left(\frac{\partial}{\partial \mathbf{X}_t} \log q(\mathbf{X}_t) \right) \left(\frac{\partial}{\partial \mathbf{X}_t} \log q(\mathbf{X}_t) \right)^\top \right] (4)$$

where "diag" is the diagonal operator, E_t is the matrix obtained by stacking the estimated scores $\left\{ \boldsymbol{\epsilon}_{\mathbf{X}_t} \left(\hat{\mathbf{X}}_t^i, t \right) : i = 1 \dots M \right\}$ and $\bar{E}_t$ the average of E_t.

Now we provide the proof of Eq. 1. For the sake of simplicity, we demonstrate our statement for a scalar x

Theorem Suppose that response x is real-valued, and the noising distribution $q(x)$ satisfies the following regularity conditions:

$$q(x) \in C^2 \tag{5}$$

i.e. q(x) is twice continuously differentiable and

$$\int_{-\infty}^{\infty} \left| \frac{\partial^2 \log q(x)}{\partial x^2} \right| q(x) dx < \infty \tag{6}$$

Then we have the main result:

$$\mathbb{E} \left[\left(\frac{\partial}{\partial x} \log q(x) \right)^2 \right] = \\ = -\mathbb{E} \left[\frac{\partial^2}{\partial x^2} \log q(x) \right]. \tag{7}$$

Proof To prove that LHS = RHS, we can start with the right-hand side and show that it equals the left-hand side.

1. First, we expand the RHS:

$$-\mathbb{E} \left[\frac{\partial^2}{\partial x^2} \log q(x) \right] = -\int q(x) \frac{\partial^2}{\partial x^2} \log q(x) dx \tag{8}$$

2. Using the chain rule:

$$\frac{\partial^2}{\partial x^2} \log q(x) = \frac{\partial}{\partial x} \left(\frac{1}{q(x)} \frac{\partial q(x)}{\partial x} \right) \tag{9}$$

Then by applying the product rule for differentiation, which states that $(u \cdot v)' = u \cdot v' + v \cdot u'$ we have that

$$= -\frac{1}{q(x)^2} \left(\frac{\partial q(x)}{\partial x} \right)^2 + \frac{1}{q(x)} \frac{\partial^2 q(x)}{\partial x^2} \tag{10}$$

3. Substituting this back into the integral:

$$-\int q(x) \left(-\frac{1}{q(x)^2} \left(\frac{\partial q(x)}{\partial x} \right)^2 + \frac{1}{q(x)} \frac{\partial^2 q(x)}{\partial x^2} \right) dx$$

$$= \int \frac{1}{q(x)} \left(\frac{\partial q(x)}{\partial x} \right)^2 dx - \int \frac{\partial^2 q(x)}{\partial x^2} dx$$

4. The second term becomes zero due to the property in Eq. 5 as:

$$\int \frac{\partial^2 q(x)}{\partial x^2} dx = \frac{\partial q(x)}{\partial x} |_{-\infty}^{\infty} \tag{11}$$

Finally, considering that q(x) is a probability distribution, its derivative $\frac{\partial q(x)}{\partial x}$ is 0 when diverging to $\pm\infty$, hence

$$\frac{\partial q(x)}{\partial x} |_{-\infty}^{\infty} = 0 \tag{12}$$

Now, going back to the first term

$$\int \frac{1}{q(x)} \left(\frac{\partial q(x)}{\partial x} \right)^2 dx \tag{13}$$

5. We can multiply and divide the integrand by $q(x)$ without changing the value of the integral:

$$\int \frac{q(x)}{q(x)} \left(\frac{\partial q(x)}{\partial x} \right)^2 \frac{1}{q(x)} dx \tag{14}$$

6. This can be rewritten as:

$$\int q(x) \left(\frac{1}{q(x)} \frac{\partial q(x)}{\partial x} \right)^2 dx \tag{15}$$

7. Now, we can use the following identity:

$$\frac{1}{q(x)} \frac{\partial q(x)}{\partial x} = \frac{\partial \log q(x)}{\partial x} \tag{16}$$

8. Substituting this identity into the previous expression, we get:

$$\int q(x) \left(\frac{\partial \log q(x)}{\partial x} \right)^2 dx \tag{17}$$

9. This is exactly the definition of the left-hand side of the original equation:

$$\mathbb{E} \left[\left(\frac{\partial}{\partial x} \log q(x) \right)^2 \right] \tag{18}$$

Therefore, we have shown that the right-hand side equals the left-hand side, proving the identity.

B. Additional figures

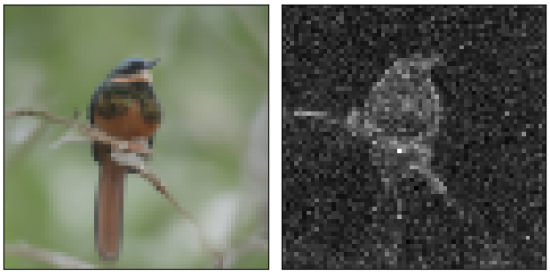

Figure 1. Left: generated image from DDPM trained on Imagenet64 with 50 steps and DDIM sampler. Right: uncertainty map of the generated image. The uncertainty map is obtained by summing the step-wise uncertainty of the sampling process. We observe that most of the uncertainty is concentrated in the foreground elements of the image.

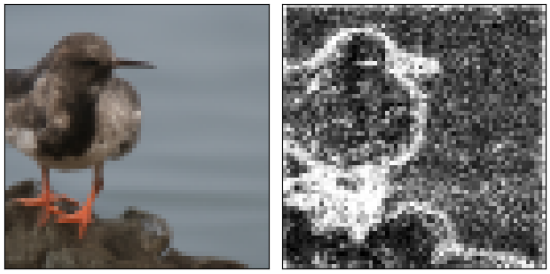

Figure 2. Left: generated image from DDPM trained on Imagenet64 with 50 steps and DDIM sampler. Right: uncertainty map from MC-Dropout of the generated image. The uncertainty map is obtained by summing the step-wise uncertainty of the sampling process. We observe that most of the uncertainty is concentrated in the edges of the foreground elements of the image.

C. Additional tables

Figure 3. Additional visual results of uncertainty guidance applied to Stable Diffusion. For each pair of images, the *top row* shows the generated image *without* uncertainty guidance while the **bottom row** shows the same image generated **with** uncertainty guidance.

Figure 4. Additional visual results of uncertainty guidance applied to Stable Diffusion. For each pair of images, the *top row* shows the generated image *without* uncertainty guidance while the **bottom row** shows the same image generated **with** uncertainty guidance.

Table 1. Comparison of the Precision and Recall between 60 000 generated images with and without the uncertainty guidance, except for Imagenet512 for memory reasons.

Model	Dataset	Precision ↑		Recall ↑	
		Random	Ours	Random	Ours
ADM	ImageNet 64	0.999	0.999	0.004	0.005
ADM	ImageNet 128	0.951	0.951	0.371	**0.380**
ADM w/2-DPM	ImageNet 128	**0.874**	0.872	0.524	**0.540**
U-ViT	ImageNet 256	0.325	**0.339**	0.762	**0.856**
U-ViT	ImageNet 512	0.791	**0.793**	0.431	**0.451**
DDPM	CIFAR-10	0.685	0.685	0.00	0.00

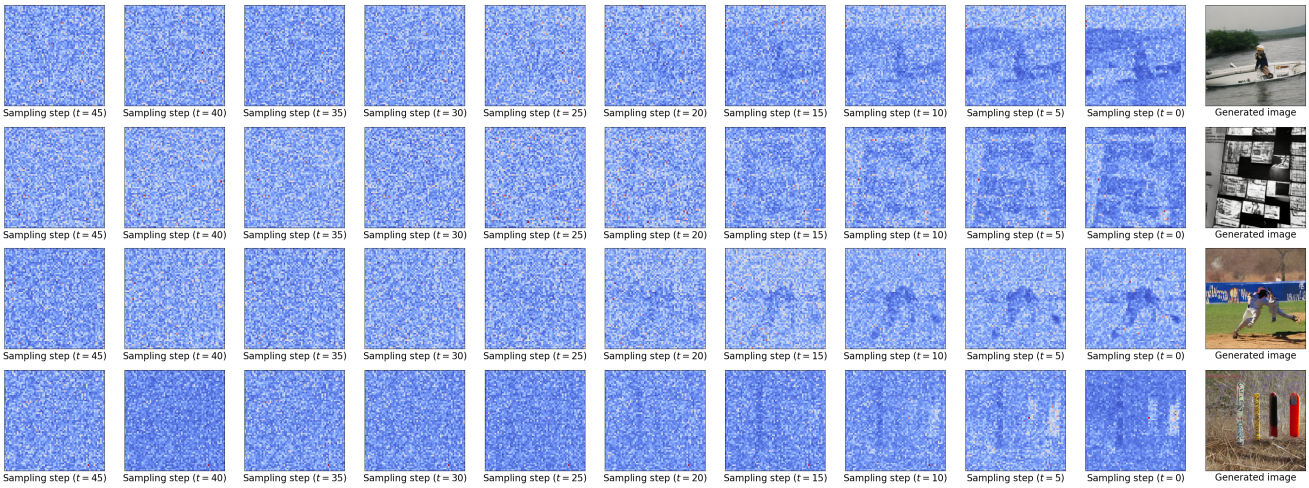

Figure 5. Uncertainty maps obtained from our proposed method. Coherently to our findings in the main article (Figure 3 in the main article), we observe high uncertainty in the first phases of the sampling process with very little differences between different samples, while most of the uncertainty related to the elements in the final image are in the last steps of the denoising process.

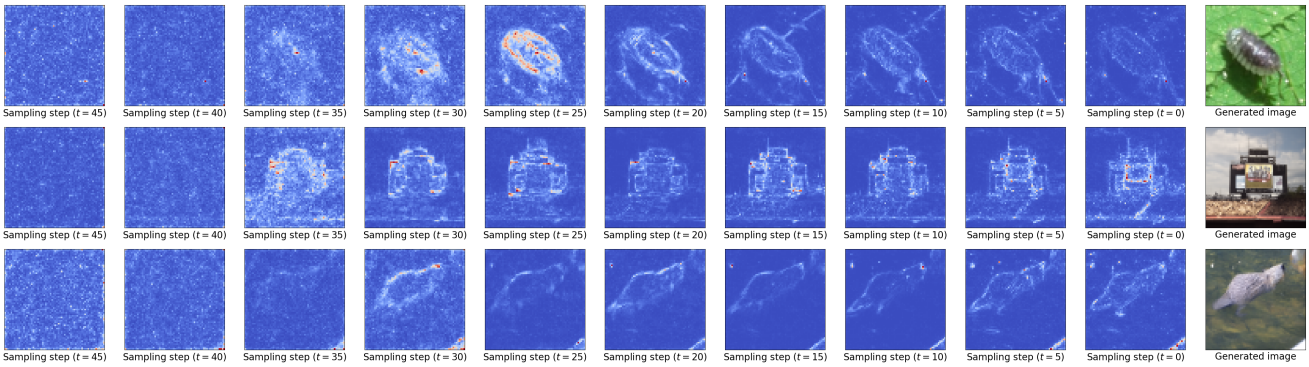

Figure 6. Uncertainty maps obtained from MC Dropout. While our method has high uncertainty on foreground objects, we observe that MC-Dropout has high uncertainty only on the edges of foreground objects

Table 2. Comparison of generation time with and without uncertainty estimation in seconds of 128 samples, using the same setup described in Section 4.1 of the main article, i.e. using M=5, 50 generation steps and compute the uncertainty between step 45 and 48.

Model	Dataset	M=5	
		Without uncertainty estimation	With uncertainty estimation
ADM	ImageNet 64	40.753	52.387
ADM	ImageNet 128	86.805	112.777
ADM w/2-DPM	ImageNet 128	86.712	112.765
U-ViT	ImageNet 256	26.272	37.058
U-ViT	ImageNet 512	32.859	47.531
DDPM	CIFAR-10	2.661	3.671

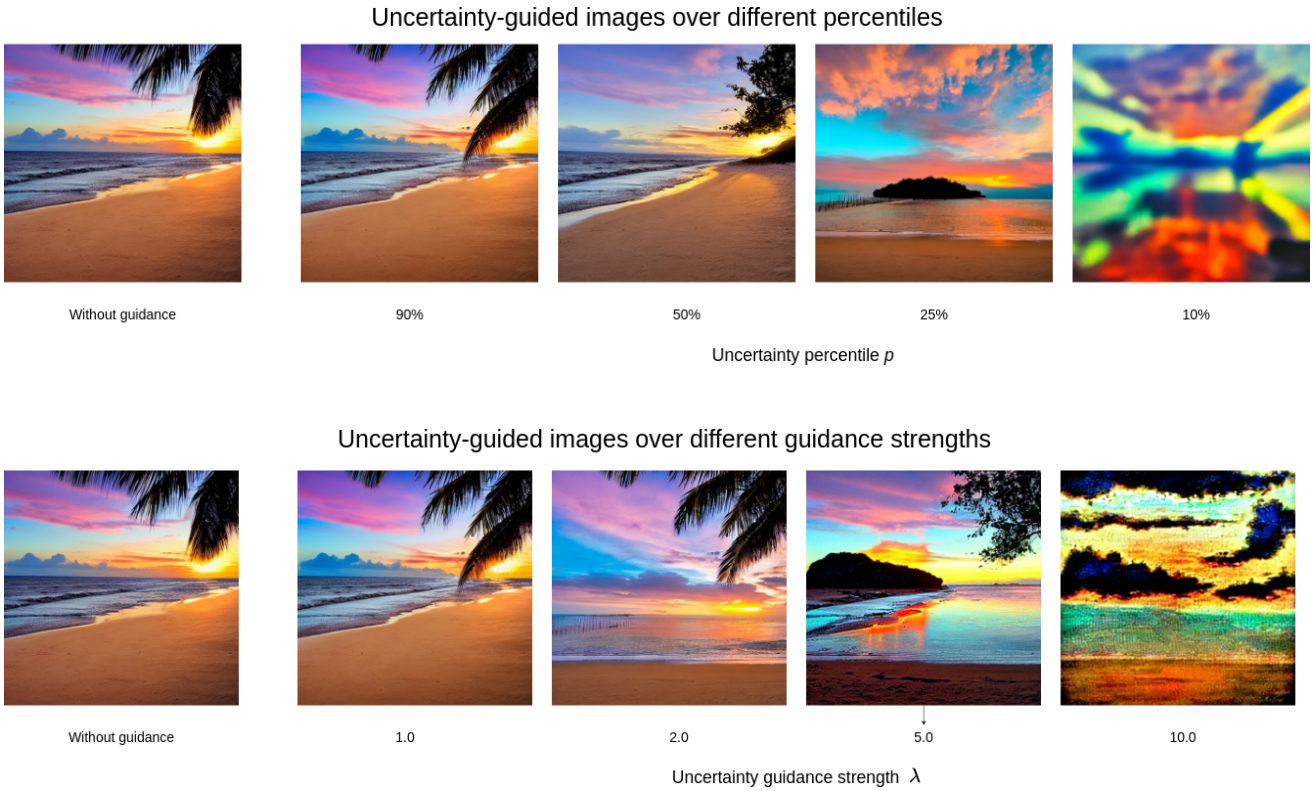

Figure 7. Hyperparameter sweep for uncertainty-guided sampling on Stable Diffusion 1.5. The top row shows the effect of varying the uncertainty percentile threshold, while the bottom row demonstrates the impact of adjusting the uncertainty strength. In the first row, by lowering the uncertainty percentile p we change important scene details as the sun. In the second row, by increasing the uncertainty guidance strength λ, we are fundamentally changing the scene structure.

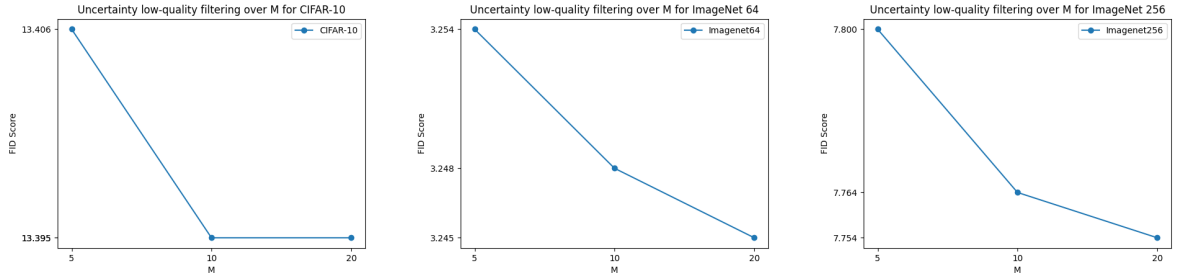

Figure 8. Uncertainty low quality filtering as in Table 1 of the main article, but using different number of perturbated samples for uncertainty estimation (M). We observed slight improvements with higher, but at the cost of higher prediction times as highlighted by Table 2 and 3

Table 3. Comparison of generation time with and without uncertainty estimation in seconds of 128 samples, using the same setup described in Section 4.1 of the main article, i.e. except for M=20, 50 generation steps and compute the uncertainty between step 45 and 48.

Model	Dataset	M=20	
		Without uncertainty estimation	With uncertainty estimation
ADM	ImageNet 64	41.013	89.316
ADM	ImageNet 128	86.768	190.939
ADM w/2-DPM	ImageNet 128	86.750	190.871
U-ViT	ImageNet 256	43.987	60.550
U-ViT	ImageNet 512	53.189	74.420
DDPM	CIFAR-10	2.726	6.302

Figure 9. Additional visual results of uncertainty guidance applied to Stable Diffusion. For each pair of images, the *top row* shows the generated image *without* uncertainty guidance while the **bottom row** shows the same image generated **with** uncertainty guidance.

Figure 10. Additional visual results of uncertainty guidance applied to Stable Diffusion. For each pair of images, the *top row* shows the generated image *without* uncertainty guidance while the **bottom row** shows the same image generated **with** uncertainty guidance. In the second column we observe the failure of the uncertainty guidance with human hands, as generating coherent hands is a very challenging task for Stable Diffusion. In the third column we observe very small changes with the uncertainty guidance, as the generated image is already of high quality. However, with hyper-parameter tuning, we can observe further improvements as demonstrated in Figure 7