Supplementary Material for "Noise-Modeled Diffusion Models for Low-Light Spike Image Restoration"

Ruonan Liu[1] Lin Zhu[1,*] Xijie Xiang[2] Lizhi Wang[3] Hua Huang[3]
[1]Beijing Institute of Technology [2]Peking University [3]Beijing Normal University
{liuruonan,linzhu}@bit.edu.cn

S1. Derivation

S1.1. Derivation of Eq.(10)

According to the transition distribution of Eq.(9), the relationship between x_t and x_0 can be expressed as:

$$x_t = A_t x_{t-1} + \alpha_t(\frac{L_d}{q}\Delta I + I_d) + \beta_t n + k\gamma_t \epsilon, \tag{S1}$$

where $\epsilon \sim \mathcal{N}(0, \mathbf{I})$. By recursively applying Eq. (S1), we can obtain the relationship between x_t and x_0 as follows:

$$x_t = (\prod_{i=1}^{t} A_i)x_0 + [\sum_{i=1}^{t}(\alpha_i \prod_{j=i+1}^{t} A_j)](\frac{L_d}{q}\Delta I + I_d) + [\sum_{i=1}^{t}(\beta_i \prod_{j=i+1}^{t} A_j)]n + k[\sum_{i=1}^{t}(\gamma_i \prod_{j=i+1}^{t} A_j)]\epsilon. \tag{S2}$$

S1.2. Derivation of Eq.(12) and Eq.(13)

According to the Bayes's theorem, we have:

$$q(x_{t-1}|x_t, x_0, y) \propto q(x_t|x_{t-1}, y)q(x_{t-1}|x_0, y), \tag{S3}$$

where $q(x_t|x_{t-1}, y) = \mathcal{N}(x_t; A_{t-1}x_{t-1} + \alpha_{t-1}(\frac{L_d}{q}\Delta I + I_d) + \beta_{t-1}n, k^2\gamma_{t-1}^2\mathbf{I})$ from Eq.(9), and $q(x_{t-1}|x_0, y) = \mathcal{N}(x_{t-1}; \overline{A_{t-1}}x_0 + \overline{\alpha_{t-1}}(\frac{L_d}{q}\Delta I + I_d) + \overline{\beta_{t-1}}n, k^2\overline{\gamma_{t-1}}^2\mathbf{I})$ from Eq.(10). Then, We focus on the quadratic component in the exponent of $q(x_{t-1}|x_0, y)$:

$$\begin{aligned} q(x_{t-1}|x_0, y) \propto & exp\{-\frac{[x_t - A_t x_{t-1} - \alpha_t(\frac{L_d}{q}\Delta I + I_d) - \beta_t n]^2}{2k^2\gamma_t^2} - \\ & \frac{[x_{t-1} - \overline{A_{t-1}}x_0 - \overline{\alpha_{t-1}}(\frac{L_d}{q}\Delta I + I_d) - \overline{\beta_{t-1}}n]^2}{2k^2\overline{\gamma_{t-1}}^2}\} \\ = & exp\{-\frac{1}{2}(\frac{A_t^2\overline{\gamma_{t-1}}^2 + \gamma_t^2}{k^2\gamma_t^2\overline{\gamma_{t-1}}^2})x_{t-1}^2 + [\frac{A_t x_t - A_t\beta_t n - A_t\alpha_t(\frac{L_d}{q}\Delta I + I_d)}{k^2\gamma_t^2} + \\ & \frac{\overline{A_{t-1}}x_0 + \overline{\beta_{t-1}n} + \overline{\alpha_{t-1}}(\frac{L_d}{q}\Delta I + I_d)}{k^2\overline{\gamma_{t-1}}^2}]x_{t-1} + const.\}, \end{aligned} \tag{S4}$$

where $const.$ denotes the component that is unrelated to x_{t-1}. Then, compare with the expected Gaussian distribution $exp\{\frac{(x_{t-1}-\mu_\theta(x_t,y,t))^2}{2\sigma_t^2} + const.\}$, we can derive $\mu_\theta(x_t, y, t)$ and $\sigma_\theta(x_t, y, t)$ as represents in Eq.(12) and Eq.(13), respectively.

*Corresponding author.

S1.3. Derivation of Eq.(14)

The posterior distribution $p(x_0|y)$ can be formulated as:

$$p(x_0|y) = \int p(x_T|y) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t, y) d_{x_{1:T}}, \tag{S5}$$

where $p(x_T|y) = \mathcal{N}(x_T|y, k^2\overline{\gamma_t}^2\mathbf{I})$ represents the starting point of the sampling process, $p_\theta(x_{t-1}|x_t, y)$ defines the reverse transition from the current state x_t to the previous state x_{t-1} with a learnable parameter θ:

$$p_\theta(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, y, t), \Sigma_\theta(x_t, y, t)). \tag{S6}$$

The parameter θ is optimized by minimizing the negative ELBO:

$$\mathcal{L}_\theta(x_t, y, t) = \sum_t D_{KL}[q(x_{t-1}|x_t, x_0, y)||p_\theta(x_{t-1}|x_t, y)], \tag{S7}$$

where $D_{KL}[\cdot||\cdot]$ denotes the KL divergence. Based on Eq. (S7), Eq.(14) can be derived as follows:

$$\begin{aligned}
\mathcal{L}_\theta(x_t, y, t) &= \sum_t D_{KL}[q(x_{t-1}|x_t, x_0, y)||p_\theta(x_{t-1}|x_t, y)] \\
&= \sum_t D_{KL}[\mathcal{N}(x_{t-1}; \mu_q, \Sigma_q(t))||\mathcal{N}(x_{t-1}; \mu_\theta, \Sigma_q(t))] \\
&= \frac{1}{2\Sigma_q^2(t)}||\mu_\theta - \mu_q||_2^2 \\
&= \frac{A_t^2\overline{\gamma_{t-1}}^2 + \gamma_t^2}{2k^2\gamma_t^2\overline{\gamma_{t-1}}^2} \cdot \frac{\overline{A_{t-1}}^2\gamma_t^4}{(A_t^2\overline{\gamma_{t-1}}^2 + \gamma_t^2)^2} \\
&= \frac{\overline{A_{t-1}}^2\gamma_t^2}{2k^2\overline{\gamma_{t-1}}^2(A_t^2\overline{\gamma_{t-1}}^2 + \gamma_t^2)}||\hat{x}_0^t - x_0||_2^2.
\end{aligned} \tag{S8}$$

S1.4. Derivation of Eq.(18)

Eq.(18) can be derived as follows:

$$\overline{\alpha_t} = \sum_{i=1}^{t}(\alpha_i \prod_{j=i+1}^{t} A_j) = \sum_{i=1}^{t} \alpha_i \frac{a_t}{a_i} = a_t \sum_{i=1}^{t} \frac{\alpha_i}{a_i} = \eta_t, \tag{S9}$$

$$\frac{\alpha_t}{a_t} = \frac{\eta_t}{a_t} - \frac{\eta_{t-1}}{a_{t-1}}, \tag{S10}$$

$$\alpha_t = \eta_t - \frac{a_t}{a_{t-1}}\eta_{t-1} = \eta_t - A_t\eta_{t-1}. \tag{S11}$$

S2. More Explanation

Overall novelty. Overall, our method deeply integrates spike noise modeling into the diffusion model and is the first to do so specifically for spike image restoration. The closest prior work, RSIR, employs separate network structures for each spike noise component and uses regression-based training for restoration under general illumination. In contrast, we target the more challenging low-light conditions and model spike image degradation within the diffusion process, embed unified global degradation guidance into the network design, thus exploiting the generative power of diffusion models.
End-to-end framework. Fig. S1 shows the end-to-end framework of our method.
Sampling process. Detailed sampling process can be seen in Alg. S1.

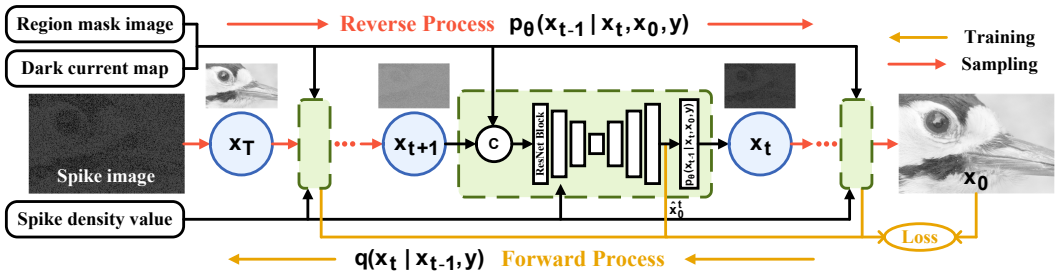

Figure S1. End-to-end framework.

Algorithm S1 Sampling

Input: Spike image y, trained model with parameter θ.
$x_T \sim \mathcal{N}(x_T; y; k^2\gamma_t^2\mathbf{I})$
for $t = T, \ldots, 1$ **do**
  $\epsilon \sim \mathcal{N}(\epsilon; 0, \mathbf{I})$ if t>1 else $\epsilon = 0$
  $x_{t-1} = \tilde{\mu_\theta}(x_{t-1}, y, t) + \sqrt{\tilde{\Sigma}_\theta(x_{t-1}, y, t)}\epsilon$
end for
return x_0

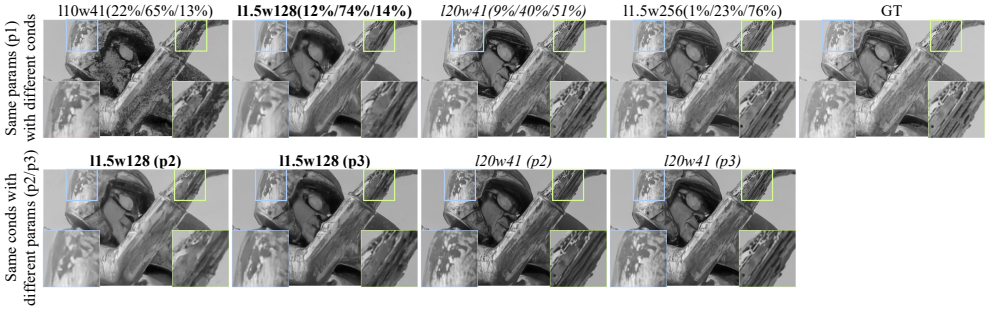

Figure S2. **Hallucination analysis.** 'l': light scale. 'w': window length. Percentages: the proportion of regions with 0, 1 and $\geq$2 spike signals in the spike stream.

S3. More Experimental Results

Hallucination analysis. Fig. S2 shows that our generated results do not exhibit unreasonable hallucinations at the selected and lighter degradation cases, except in more severe ones ($l10w41$). The results become closer to the ground truth with reduced diversity as degradation weakens ($l \uparrow$ $w \uparrow$), and show reasonable variations under different noise parameters. Our diffusion prior adapts to spike noise, parameters, and degradation prompts, enabling more faithful restoration.
More qualitative results. More comparison of qualitative results on simulated and real-world datasets are shown in Fig. S3 and Fig. S4, respectively.
More ablation study results. **(i) p-value.** More ablation studies on p-value are provided in S1. If the p-value is set too extreme, it may adversely affect performance. **(ii) Timesteps t, noise hyperparameter k, and LPIPS loss weight w.** Ablation studies on timesteps t, noise hyperparameter k, and LPIPS loss weight w are provided in S2. Our method requires only a small number of sampling steps, and the inclusion of LPIPS loss significantly enhances perceived quality.

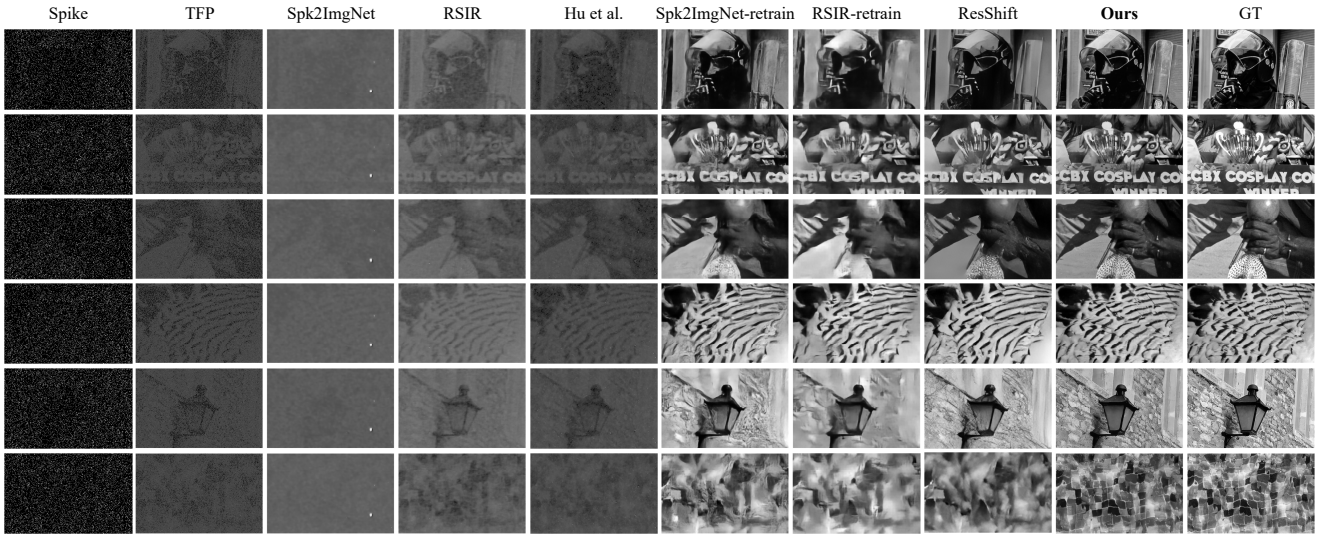

Figure S3. **More qualitative results on simulated datasets.** Cols 2–5 are gamma-corrected for clearer visualization. Zoom in for better view.

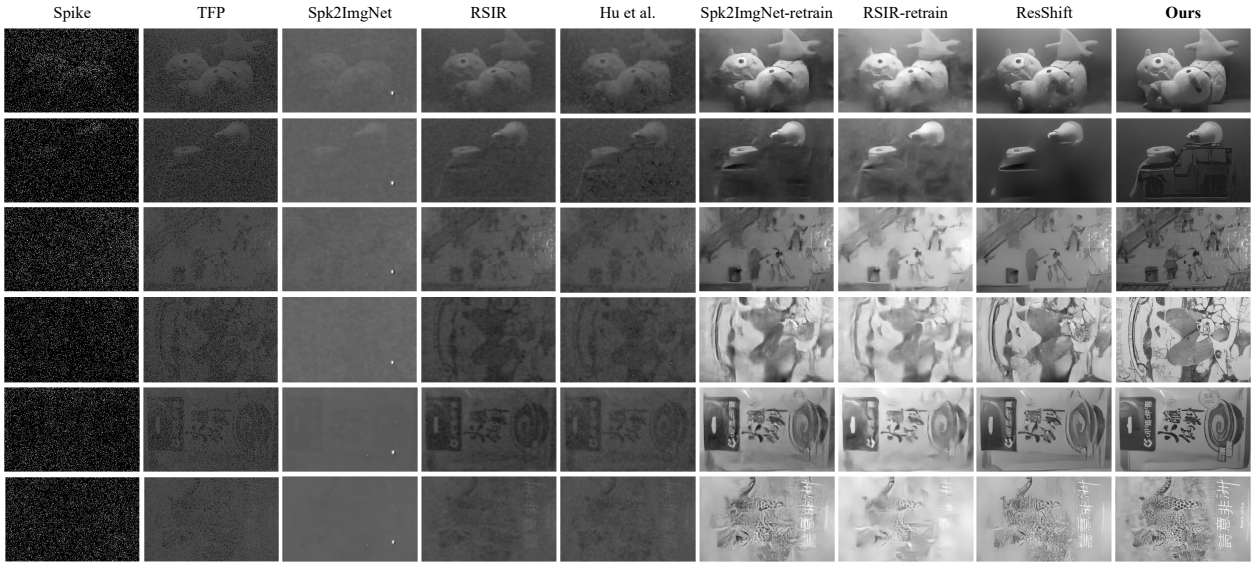

Figure S4. **More qualitative results on real-world datasets.** Cols 2–5 are gamma-corrected for clearer visualization. Zoom in for better view.

Table S1. Ablation study of schedule.

p-value			Metrics		
p_1	p_2	p_3&p_4	PSNR↑	SSIM↑	LPIPS↓
2			27.16	0.7790	0.1542
4			27.19	0.7805	0.1393
6	3	0.005	27.16	0.7807	0.1401
8			26.99	0.7786	0.1378
	1		27.05	0.7764	0.1525
	2		26.95	0.7776	0.1370
6	3	0.005	27.16	0.7807	0.1363
	6		26.68	0.7688	0.1351
		0.001	26.84	0.7767	0.1339
6	3	0.005	27.16	0.7807	0.1353
		0.01	26.83	0.7767	0.1355
		0.02	25.74	0.7685	0.1387

Table S2. Ablation study on timesteps t, noise hyper-parameter k, and LPIPS loss weight w.

hyper-parameter			Metrics		
t	k	w	PSNR↑	SSIM↑	LPIPS↓
4			26.80	0.7793	0.1491
6	0.1	2	27.16	0.7807	0.1352
8			26.43	0.7795	0.1444
	0.01		26.89	0.7783	0.1335
6	0.1	2	27.16	0.7807	0.1352
	0.5		26.53	0.7733	0.1368
		0	27.13	0.7838	0.1965
6	0.1	2	27.16	0.7807	0.1352
		4	26.92	0.7757	0.1299