SYMMETRY IS ALL YOU NEED: IMAGE GENERATION USING PRE-TRAINED DEEP DIFFUSION PROBABILISTIC MODELS

Stephen D. Liang
Hewlett Packard Enterprise
6280 America Center Dr
San Jose, CA 95002

ABSTRACT

Denoising Diffusion Probabilistic Models (DDPMs) typically rely on zero-mean white Gaussian noise during both training and sample generation. In this paper, we demonstrate that image generation in pre-trained DDPMs is not restricted to Gaussian noise; any zero-mean, symmetrically distributed noise or signal can be equally effective. We present theoretical and empirical evidence that symmetry, rather than the specific distribution type, is a sufficient condition for successful image generation using a pre-trained DDPM. Our findings enable DDPMs to operate more flexibly on resource-constrained devices by utilizing alternative noise or signal types, making them suitable for applications such as semantic communication, where symmetrical noise and signals are prevalent. This work establishes new avenues for efficient and adaptable generative models in real-world environments.

1 INTRODUCTION

Denoising Diffusion Probabilistic Models (DDPMs) are a powerful category of generative models that have gained prominence for their ability to produce high-quality images. In contrast to traditional generative models like Generative Adversarial Networks (GANs) or Variational Autoencoders (VAEs), DDPMs use a diffusion process inspired by statistical physics (Sohl-Dickstein et al., 2015; Ho et al., 2020). This approach involves modeling data distributions by reversing a noise-corruption process that gradually distorts the data over time. The diffusion process consists of two phases: a forward phase, where Gaussian noise is progressively added to data, and a reverse phase, where a neural network learns to denoise the data to recover the original distribution.

A typical implementation of DDPMs employs white Gaussian noise during the forward diffusion and denoising process. However, this reliance on white Gaussian noise may limit flexibility, especially on devices with computational constraints, such as smartphones, where generating high-quality white Gaussian noise may be infeasible or unnecessary. This raises a critical question: could alternative noise distributions, such as colored Gaussian noise or non-Gaussian, symmetric noise or signal, also support effective image generation? By examining these looser conditions for the noise distribution in pre-trained DDPMs, we aim to identify configurations that are more suitable for resource-constrained environments.

In this paper, we focus on exploring how pre-trained DDPMs can operate with non-standard noise distributions to facilitate image generation on devices with limited computational power. Specifically, we investigate whether symmetric noise or signal can replace white Gaussian noise in the denoising process without compromising image quality. This line of inquiry could broaden the applicability of DDPMs, making them more practical for mobile or low-power applications where starting from scratch to train models is impractical.

The contributions of this paper are as follows:

1. We apply the Information Bottleneck (IB) principle to DDPMs to analyze and derive key design objectives for improving image generation on computationally constrained devices.

2. We establish that a symmetric noise distribution suffices for effective image generation in pre-trained DDPMs, challenging the assumption that white Gaussian noise is essential.

3. We demonstrate that random noise from symmetric distributions can successfully generate images with DDPMs and that different images can be generated by simply reshuffling noise samples.

4. We validate that symmetrical random signals are effective for generating samples in DDPMs, paving the way for optimized usage in resource-limited settings.

5. We extend our findings to semantic communication, where DDPMs generate content under Additive White Gaussian Noise (AWGN) corruption. Additionally, we show that using an interleaver is an efficient method for content generation in DDPM-based semantic communication.

2 Related Work

In DDPMs, white Gaussian noise has traditionally been used for both the forward and reverse processes (Ho et al., 2020; Choi et al., 2022). This type of noise is characterized by a flat power spectral density (PSD) across all frequencies (Kuo, 2018). In (Nachmani et al., 2021), alternative noise types, such as Gamma noise and a mixture of Gaussian noises, were applied during the diffusion process, leading to performance improvements. Further optimization of DDPMs was achieved through noise level estimation and scheduling adjustments, as explored in (San-Roman et al., 2021). The work in (Stevens et al., 2023) examined the removal of structured noise in diffusion models. Additionally, (Wyatt et al., 2022) introduced a multi-scale simplex noise diffusion method for anomaly detection, showing that simplex noise outperformed Gaussian noise in certain tasks. More recently, blue noise, which has a PSD that increases with frequency, was incorporated into diffusion models (Huang et al., 2024). This approach integrates correlated noise into model training, resulting in improved image quality compared to models using white Gaussian noise (Huang et al., 2024). In Huberman-Spiegelglas et al. (2024), an alternative latent noise space for DDPM was proposed for editing operations via simple means and an inversion method. In Wen et al. (2024), a generative change detection (CD) model called GCD-DDPM was proposed to directly generate CD maps by exploiting DDPM. Recently, Mix-DDPM was proposed to randomly select a Gaussian component and then add the chosen Gaussian noise to perturb the signals into a known distribution Wang et al. (2024).

Regarding modifications to DDPMs, several improvements have been introduced, such as enhanced noise scheduling and reduced gradient noise (Nichol & Dhariwal, 2021). A Discrete Denoising Diffusion Probabilistic Model (D3PM) was proposed in (Austin et al., 2021) for discrete data. Iterative Latent Variable Refinement (ILVR), which generates high-quality images using a reference image, was introduced in (Choi et al., 2021). In (Zhao et al., 2023), a multimodal image fusion method using DDPMs was explored, while (Ryu & Ye, 2022) introduced a pyramidal DDPM for high-resolution image generation. Recently, a multi-task denoising diffusion framework, DiffusionMTL, was proposed in (Ye & Xu, 2024) for refining task prediction. A Denoising Diffusion Implicit Model (DDIM) was introduced in (Song et al., 2020), featuring a faster reverse denoising process. Denoising Diffusion Restoration Models (DDRM) (Kawar et al., 2022) leveraged pre-trained models to solve linear inverse problems, while a Single Image Denoising Diffusion Model (SinDDM) (Kulikov et al., 2023) employed a multi-scale diffusion process to learn internal image statistics. Bilateral Denoising Diffusion Models (BDDM) (Lam et al., 2021) reduced the number of steps required for high-quality sample generation, and Latent Diffusion Models (LDMs) (Rombach et al., 2022), utilizing autoencoders' latent space, provided both quality and flexibility, forming the basis for Stable Diffusion (Bolya & Hoffman, 2023; Tang et al., 2022).

In terms of theoretical studies in DDPMs, the variational lower bound of diffusion models was simplified using the signal-to-noise ratio (Kingma et al., 2021). Sampling from diffusion probabilistic models (DPMs) was investigated in (Lu et al., 2022), where an exact solution for diffusion ODEs was proposed. The Higher-Order Denoising Diffusion Solver (GENIE) (Dockhorn et al., 2022) introduced higher-order gradients to accelerate synthesis, while (Bao et al., 2022) proposed an analytic diffusion probabilistic model for variance estimation without additional training.

3 Introduction to DDPM

The DDPM consists of two processes: the forward diffusion process and the reverse denoising process (Ho et al., 2020). The model is trained to reverse the forward diffusion process by learning how to denoise noisy samples step by step.

In the forward diffusion process, noise is progressively added to the data over T steps, transforming a data sample x_0 into pure noise x_T. Each step in this process can be described as a small perturbation of the sample, controlled by a variance schedule β_t (Ho et al., 2020):

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbf{I}) \tag{1}$$

Starting with a data sample x_0, the sample at step t is obtained as (Ho et al., 2020):

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)\mathbf{I}) \tag{2}$$

where $\alpha_t = 1 - \beta_t$, and $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$ represents the cumulative product of the variance schedule. The noise scheduler β_t is designed such that at the end of the diffusion process $\bar{\alpha}_t \to 0$, resulting in (Ho et al., 2020):

$$q(\mathbf{x}_T) \approx \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I}), \tag{3}$$

The reverse denoising process aims to learn how to denoise a sample x_t at each step by predicting x_{t-1} from x_t. The learned reverse process is parameterized as (Ho et al., 2020):

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \tag{4}$$

Typically, Σ_θ is set to a time-dependent constant, and the mean μ_θ is predicted by a neural network. The model is trained by minimizing the following variational bound (Ho et al., 2020):

$$L_{\text{vlb}} = \mathbb{E}_q \left[D_{\text{KL}}(q(x_{t-1}|x_t, x_0)||p_\theta(x_{t-1}|x_t)) \right] \tag{5}$$

Alternatively, the training objective can be simplified to predict the noise ϵ added at each step (Ho et al., 2020):

$$L_{\text{simple}} = \mathbb{E}_{t,x_0,\epsilon} \left[\|\epsilon - \epsilon_\theta(x_t, t)\|^2 \right] \tag{6}$$

where $\epsilon_\theta(x_t, t)$ is the noise predicted by the model, and ϵ is the true noise added in the forward process.

Once trained, the model generates samples by starting from Gaussian noise $x_T \sim \mathcal{N}(0, \mathbf{I})$ and iteratively applying the reverse process to generate progressively less noisy samples (Ho et al., 2020):

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z \tag{7}$$

where $z \sim \mathcal{N}(0, \mathbf{I})$ is sampled from a standard Gaussian distribution, and σ_t controls the noise variance at each step.

For ease of analysis, let us denote the output from the reverse denoising process as $\hat{\mathbf{x}}_0$. In most cases, the denoised output $\hat{\mathbf{x}}_0$ differs from the initial input $\mathbf{x}_0$.

4 Information Bottleneck Principle for DDPM

The Information Bottleneck (IB) principle is a concept from information theory that provides a framework for compressing data while preserving the most relevant information for a specific task (Tishby et al., 2000). The IB principle has been explored to gain insights into how deep learning models (Tishby & Zaslavsky, 2015), particularly deep neural networks (DNNs), generalize from training data to unseen data. In deep learning, the goal is to learn representations of the input data X that are useful for predicting the target labels Y. This learning process can be viewed through the lens of the IB principle, which seeks to balance the compression of input data with the preservation of predictive information.

The IB principle is applied to deep learning as follows (Tishby & Zaslavsky, 2015):

$$\min_{p(T|X)} \mathcal{L}(p(T|X)) = I(X;T) - \beta I(T;Y) \tag{8}$$

where T represents the hidden layer activations in the neural network; $I(X;T)$ quantifies how much information the hidden representation T retains about the input X; and $I(T;Y)$ measures how much information T provides about the target output Y.

In deep networks, each layer progressively transforms the input data into representations that compress irrelevant information (reducing $I(X;T)$) while retaining the information crucial for predicting Y (maximizing $I(T;Y)$). This can be viewed as each layer acting like a successive stage in IB optimization. Research has shown that during the training of deep neural networks, two distinct phases occur (Tishby & Zaslavsky, 2015):

- **Fitting phase:** The network fits the training data by increasing $I(T;Y)$, ensuring that the hidden representations capture meaningful patterns in the data.
- **Compression phase:** The network compresses the information in the input by reducing $I(X;T)$, discarding irrelevant details and improving generalization.

The IB principle provides a theoretical framework for understanding how deep learning models balance fitting the data and generalizing well to new inputs. It offers insights into the trade-off between model complexity and predictive accuracy, key for avoiding overfitting.

In DDPM, the initial data sample $\mathbf{x}_0$, the final diffusion outcome $\mathbf{x}_T$, and the denoised estimate $\hat{\mathbf{x}}_0$ form a Markov chain:

$$\mathbf{x}_0 \to \mathbf{x}_T \to \hat{\mathbf{x}}_0. \tag{9}$$

The input $\mathbf{x}_0$ to the diffusion process can either be a matrix or a vector. If it is a matrix, we can vectorize it (Barratt, 2018), and assume the vector length is N.

In DDPM design, two key goals should be achieved based on the IB principle:

$$\max I(\mathbf{x}_T;\hat{\mathbf{x}}_0), \tag{10}$$

$$\min I(\mathbf{x}_0;\mathbf{x}_T). \tag{11}$$

5 Symmetry Is All You Need

In this paper, we study image generation based on a trained DDPM, so we focus on (10). The mutual information $I(\mathbf{x}_T;\hat{\mathbf{x}}_0)$ measures the amount of information that the noisy image $\mathbf{x}_T$ shares with the denoised output $\hat{\mathbf{x}}_0$:

$$I(\mathbf{x}_T;\hat{\mathbf{x}}_0) = h(\hat{\mathbf{x}}_0) - h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T) \tag{12}$$

where $h(\hat{\mathbf{x}}_0)$ is the differential entropy of $\hat{\mathbf{x}}_0$; $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$ is the conditional differential entropy of $\hat{\mathbf{x}}_0$ given $\mathbf{x}_T$. Our goal is to maximize $I(\mathbf{x}_T;\hat{\mathbf{x}}_0)$ by appropriately choosing the distribution of $\mathbf{x}_T$ and designing the denoising process.

A random variable $\mathbf{x}$ has a symmetric distribution if its probability density function (PDF) satisfies:

$$f_{\mathbf{x}}(x) = f_{\mathbf{x}}(-x) \quad \text{for all } x \in \mathbb{R} \tag{13}$$

for example, Gaussian Distribution, Laplace Distribution, Uniform Distribution, etc. Such variables often have the properties: 1) Zero Mean: Symmetric around the mean (often zero). 2) Equal Tail Behavior: The probability of deviations above and below the mean is the same. 3) Moment Characteristics: Odd moments (about the mean) are zero; even moments capture the distribution's spread.

For entropy $h(\mathbf{x}_T)$, symmetric distributions can have high entropy, contributing positively to $I(\mathbf{x}_T;\hat{\mathbf{x}}_0)$. For conditional entropy $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$, a well-designed denoising function can exploit symmetry to minimize this term.

Symmetric signal and noise have some good properties: 1) Symmetry allows for more accurate predictions of $\hat{\mathbf{x}}_0$ from $\mathbf{x}_T$. 2) Symmetric noise models can be inverted effectively during denoising. 3) Symmetry ensures that information is not skewed or lost in certain directions. Symmetric distributions often have higher entropy than skewed distributions for a given variance, enhancing $h(\hat{\mathbf{x}}_0)$. So a symmetric $\mathbf{x}_T$ can lead to a large $h(\hat{\mathbf{x}}_0)$ and a small $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$, maximizing $I(\mathbf{x}_T;\hat{\mathbf{x}}_0)$. The symmetry can be leveraged to design denoising functions that reduce uncertainty in $\hat{\mathbf{x}}_0$ given $\mathbf{x}_T$, to minimize $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$

In the current image generation based on a trained DDPM, Gaussian noise is often used for $\mathbf{x}_T$. The Gaussian distribution maximizes entropy among all distributions with the same variance. However, when considering mutual information, it is the difference in entropies that matters. Non-Gaussian symmetric distributions can have entropy levels close to that of the Gaussian. The denoising process can be tailored to specific symmetric distributions to minimize $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$.

In the denoising process, we need to develop a denoising function $\hat{\mathbf{x}}_0 = g(\mathbf{x}_T)$ that effectively reconstructs $\mathbf{x}_0$ from $\mathbf{x}_T$, maximizing $I(\mathbf{x}_T; \hat{\mathbf{x}}_0)$. The appropriate denoising function reduces $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$, increasing $I(\mathbf{x}_T; \hat{\mathbf{x}}_0)$.

Any symmetric distribution of $\mathbf{x}_T$ can be used to maximize $I(\mathbf{x}_T; \hat{\mathbf{x}}_0)$ when the denoising function is appropriately designed. While Gaussian distributions have desirable properties, they are not the only distributions that can maximize mutual information in DDPMs. This allows for flexibility in designing DDPMs to handle various types of noise and data characteristics. Using different symmetric distributions can improve robustness to noise types encountered in practice. While Gaussian distributions maximize entropy, mutual information depends on both $h(\hat{\mathbf{x}}_0)$ and $h(\hat{\mathbf{x}}_0 \mid \mathbf{x}_T)$. Balancing these can be achieved with other symmetric distributions.

This analysis shows that a zero-mean Gaussian distribution for $\mathbf{x}_T$ is sufficient to satisfy the IB principle. Gaussian distribution is a symmetric distribution. If we relax the Gaussian condition and only require that $\mathbf{x}_T$ has a symmetric distribution, such as Laplacian, uniform, cosine, or logistic (Johnson et al., 2000), the DDPM still performs well in the denoising process.

Symmetric distributions (like Laplace or uniform) can approximate Gaussian behavior, especially when aggregated over many steps. According to the Central Limit Theorem (CLT) (Rosenblatt, 1956; Islam, 2018), when independent random variables from any distribution with finite mean and variance are added, their sum tends toward a Gaussian distribution, regardless of the original distribution.

Thus, symmetric noise distributions may produce behavior similar to Gaussian noise because, over multiple time steps, noise aggregation will tend toward Gaussianity due to the CLT. As long as the noise distribution is symmetric and has finite variance, it will approach Gaussian behavior as it propagates through the reverse diffusion steps. Symmetric distributions with finite variance meet this criterion, allowing the neural network to model both mean and variance effectively.

6 Experiments

We used the LSUN-Church dataset (Yu et al., 2015) with a resolution of $3 \times 256 \times 256$ for our experiments. This dataset is used in the scenario where the forward diffusion involves white Gaussian noise, and the reverse denoising process may employ different noise types, such as Laplacian noise. Our DDPM design followed the architecture described in (Ho et al., 2020). We conducted experiments to validate our theoretical findings using random noises and signals, with a 2D UNet employed for the DDPM (Ronneberger et al., 2015).

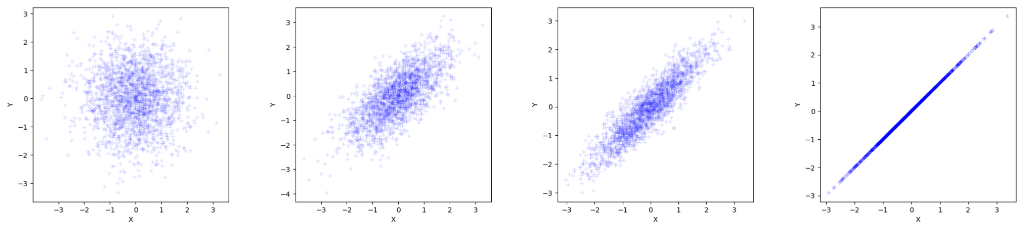
Gaussian noise can be either white or colored. In Fig. 1, we illustrate four zero-mean Gaussian noises with different covariance matrices $\mathbf{K}_\mathbf{n}$. These noises share the same average power:

$$P = \mathrm{tr}(\mathbf{K}_\mathbf{n}), \tag{14}$$

since their traces are identical. Fig. 1(a) depicts white Gaussian noise with a unitary covariance matrix, while Figs. 1(b)-(d) show examples of colored Gaussian noise. Notably, Fig. 1(d) illustrates singular noise, where $|\mathbf{K}_\mathbf{n}| = 0$.

In Fig. 1, observe that white Gaussian noise (Fig. 1a) exhibits no directional preference. In contrast, colored noises (Figs. 1b-d) have directional biases. These noises exhibit stronger power (higher eigenvalue of $\mathbf{K}_\mathbf{n}$) in the direction $\begin{bmatrix}1\\1\end{bmatrix}$ compared to the orthogonal direction $\begin{bmatrix}-1\\1\end{bmatrix}$. The singular noise (Fig. 1d) has power 2 in the stronger direction and power 0 in the weaker direction.

We performed the denoising process over 1000 iterations. Fig. 2 illustrates the generated samples by DDPM using the four different Gaussian noises across several iterations. All types of Gaussian noise performed well in generating the church images. However, the skies generated by singular noise had no clouds, reflecting the absence of detail in certain regions.



(a) (b) (c) (d)

Figure 1: Illustration of Gaussian noise. (a) White Gaussian noise with zero-mean and $\mathbf{K_n} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, (b) Colored Gaussian noise with zero-mean and $\mathbf{K_n} = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$, (c) Colored Gaussian noise with zero-mean and $\mathbf{K_n} = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$, (d) Singular Gaussian noise with zero-mean and $\mathbf{K_n} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

We also used Laplacian noise (with power 1) to generate samples, as shown in Fig. 3, row 1. Since the noise follows a Laplacian distribution, applying an interleaver to the noise will not alter its distribution. Block Interleaving is a technique used in signal processing, data transmission, and error correction to enhance the robustness of data against burst errors, where a cluster of adjacent data elements may be corrupted simultaneouslySadjadpour et al. (2001)Takeshita & Costello (2000)Yin et al. (2019). The process involves breaking data into smaller, evenly-sized blocks and rearranging them in a way that distributes sequential elements across non-adjacent positions. This redistribution of data is particularly beneficial in noisy environments, as it allows errors to be spread across multiple blocks, making them appear more random Yin et al. (2020). This, in turn, makes the errors easier to detect and correct.

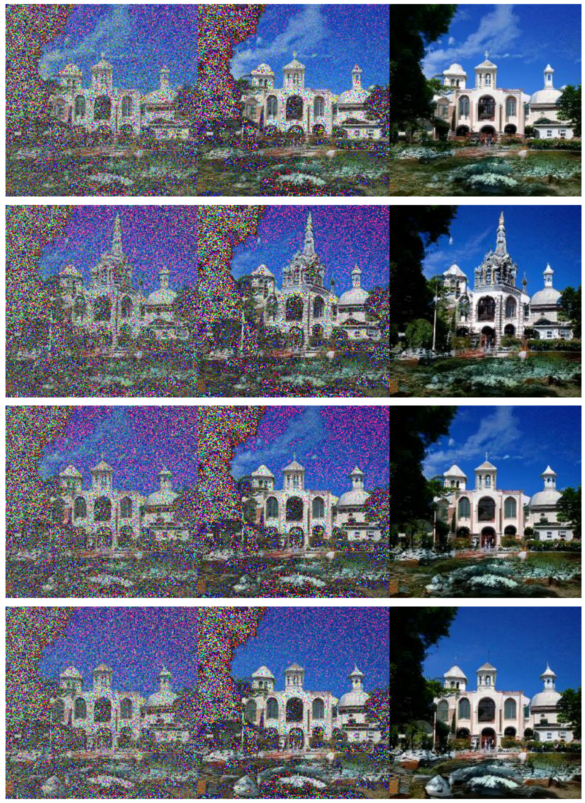
In the context of image processing and deep learning, block interleaving can be used to rearrange image data to modify the spatial distribution of noise or features. By interleaving noisy image blocks, noise patterns become more uniformly dispersed, aiding in smoother noise removal and higher quality outputs. We applied 4×8 and 16×8 block interleavers to the noise and used them to generate samples with the trained DDPM. The results, shown in rows 2 and 3 of Fig. 3, indicate that the interleaved noise produced entirely different church images. This demonstrates that same noise distribution can generates different images by simply reshuffling the noises.

In addition to Laplacian noise, we explored three other symmetric noise types—uniform, cosine, and logistic—for image generation in DDPM. In Fig. 4 (row 1), we plot the generated outputs of DDPM with uniform noise in the range $[-\sqrt{3}, \sqrt{3}]$ so that the noise has a power of 1. In Fig. 4 (row 2), we plot the generated outputs of DDPM with cosine noise, which was normalized to have a power of 1. In Fig. 4 (row 3), we plot the generated outputs of DDPM with logistic noise, which was normalized to have a power of 1. Note that all noises with symmetrical distributions successfully generated churches using DDPM.

More experimental results (ablation study, random signals for image generation, and application to semantic communications) are provided in the Appendix.

7 Conclusions and Future Work

In this paper, we applied the IB principle to DDPMs, analyzing two essential design objectives within the DDPM framework. We demonstrated that symmetric distributions are sufficient for image generation in pre-trained DDPMs, providing empirical validation that random noise with symmetric properties can effectively generate diverse images through simple reshuffling. Additionally, we showed that symmetrical random signals can be reliably utilized for image generation in DDPMs, underscoring the robustness of this approach. Further, we extended DDPMs to semantic communication, where symmetric random signals, even when corrupted by AWGN, enabled efficient content



iteration 900 iteration 950 iteration 1000

Figure 2: Generated outputs of DDPM with zero-mean white Gaussian noise (row 1), zero-mean colored Gaussian noise with covariance matrix $\begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$ (row 2), zero-mean colored Gaussian noise with covariance matrix $\begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$ (row 3), and singular Gaussian noise with covariance matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (row 4).

generation. The integration of an interleaver further improved the efficacy of DDPM-based semantic communication. These findings suggest new possibilities for optimizing DDPM architectures by simplifying noise generation, enhancing both efficiency and performance, and broadening the application scope in communication systems.

Building on the use of DDPMs for semantic communication, future work could explore more sophisticated interleaving techniques that leverage the unique properties of diffusion models. Future work could delve deeper into the design of specialized interleavers and explore error correction mechanisms that work synergistically with DDPMs. This could lead to improved robustness against channel impairments such as fading or interference in wireless communication systems. While we demonstrated the efficacy of symmetric noise distributions for DDPM-based generation, optimizing the structure and characteristics of this noise for specific communication tasks remains an open question. Tailoring noise generation methods to fit particular data types or communication channels could improve both efficiency and quality of generated content.

Another key challenge is understanding how diffusion models perform under extreme channel conditions, such as low-SNR environments or highly dynamic fading scenarios. Future research could explore whether adaptive noise scheduling within DDPMs can enhance signal recovery when communication conditions are unstable. Additionally, investigating hybrid approaches that combine DDPMs with other generative models, such as normalizing flows or VAEs, may yield more efficient

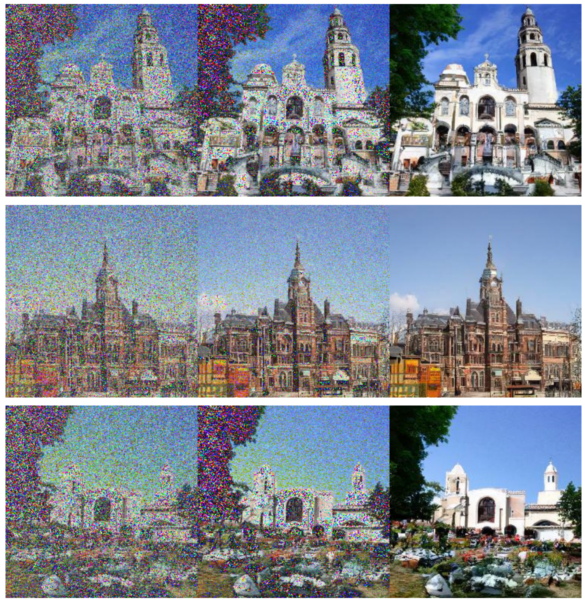


Figure 3: Generated outputs of DDPM using Laplacian noise (row 1), Laplacian noise with 4×8 block interleaver (row 2), and Laplacian noise with 16×8 block interleaver (row 3).

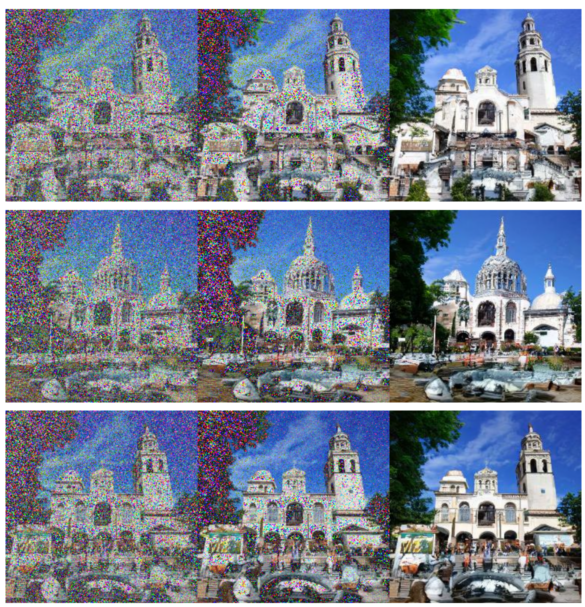


Figure 4: Generated outputs of DDPM with uniform noise in $[-\sqrt{3}, \sqrt{3}]$ (row 1). Generated outputs of DDPM with cosine noise (row 2). Generated outputs of DDPM with logistic noise (row (3).

transmission methods. The integration of reinforcement learning with diffusion-based semantic communication is another promising avenue, allowing for dynamic adaptation to varying network conditions. Finally, extending DDPMs to handle multimodal data (e.g., text, images, audio) in semantic communication frameworks presents another interesting research direction. Future work could investigate how DDPMs could be adapted to encode and transmit semantically relevant information across multiple data types within a single framework, improving efficiency and versatility.

REFERENCES

Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Advances in Neural Information Processing Systems*, 34:17981–17993, 2021.

Fan Bao, Chongxuan Li, Jun Zhu, and Bo Zhang. Analytic-dpm: an analytic estimate of the optimal reverse variance in diffusion probabilistic models. *arXiv preprint arXiv:2201.06503*, 2022.

Shane Barratt. A matrix gaussian distribution. *arXiv preprint arXiv:1804.11010*, 2018.

Daniel Bolya and Judy Hoffman. Token merging for fast stable diffusion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4599–4603, 2023.

Jooyoung Choi, Sungwon Kim, Yonghyun Jeong, Youngjune Gwon, and Sungroh Yoon. Ilvr: Conditioning method for denoising diffusion probabilistic models. *arXiv preprint arXiv:2108.02938*, 2021.

Jooyoung Choi, Jungbeom Lee, Chaehun Shin, Sungwon Kim, Hyunwoo Kim, and Sungroh Yoon. Perception prioritized training of diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11472–11481, 2022.

Daniel J Costello and G David Forney. Channel coding: The road to channel capacity. *Proceedings of the IEEE*, 95(6):1150–1177, 2007.

Tim Dockhorn, Arash Vahdat, and Karsten Kreis. Genie: Higher-order denoising diffusion solvers. *Advances in Neural Information Processing Systems*, 35:30150–30166, 2022.

Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Xingchang Huang, Corentin Salaun, Cristina Vasconcelos, Christian Theobalt, Cengiz Oztireli, and Gurprit Singh. Blue noise for diffusion models. In *ACM SIGGRAPH 2024 Conference Papers*, pp. 1–11, 2024.

Inbar Huberman-Spiegelglas, Vladimir Kulikov, and Tomer Michaeli. An edit friendly ddpm noise space: Inversion and manipulations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12469–12478, 2024.

Mohammad Rafiqul Islam. Sample size and its role in central limit theorem (clt). *Computational and Applied Mathematics Journal*, 4(1):1–7, 2018.

Richard A Johnson, Irwin Miller, and John E Freund. *Probability and statistics for engineers*, volume 2000. Pearson Education London, 2000.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. *Advances in Neural Information Processing Systems*, 35:23593–23606, 2022.

Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models. *Advances in neural information processing systems*, 34:21696–21707, 2021.

Vladimir Kulikov, Shahar Yadin, Matan Kleiner, and Tomer Michaeli. Sinddm: A single image denoising diffusion model. In *International conference on machine learning*, pp. 17920–17930. PMLR, 2023.

Hui-Hsiung Kuo. *White noise distribution theory*. CRC press, 2018.

Max WY Lam, Jun Wang, Rongjie Huang, Dan Su, and Dong Yu. Bilateral denoising diffusion models. *arXiv preprint arXiv:2108.11514*, 2021.

Stephen D. Liang. Vision language models for massive mimo semantic communication. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2025.

Jialu Liu, Deng Cai, and Xiaofei He. Gaussian mixture model with local consistency. In *Proceedings of the AAAI conference on artificial intelligence*, volume 24, pp. 512–517, 2010.

Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps. *Advances in Neural Information Processing Systems*, 35:5775–5787, 2022.

Xuewen Luo, Hsiao-Hwa Chen, and Qing Guo. Semantic communications: Overview, open issues, and future research directions. *IEEE Wireless Communications*, 29(1):210–219, 2022.

Eliya Nachmani, Robin San Roman, and Lior Wolf. Non gaussian denoising diffusion models. *arXiv preprint arXiv:2106.07582*, 2021.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International conference on machine learning*, pp. 8162–8171. PMLR, 2021.

John G Proakis and Masoud Salehi. *Digital communications*. McGraw-hill, 2008.

Theodore S Rappaport. *Wireless communications: principles and practice*. Cambridge University Press, 2024.

Carl Rasmussen. The infinite gaussian mixture model. *Advances in neural information processing systems*, 12, 1999.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241. Springer, 2015.

Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the national Academy of Sciences*, 42(1):43–47, 1956.

Dohoon Ryu and Jong Chul Ye. Pyramidal denoising diffusion probabilistic models. *arXiv preprint arXiv:2208.01864*, 2022.

Hamid R Sadjadpour, Neil JA Sloane, Masoud Salehi, and Gabriele Nebe. Interleaver design for turbo codes. *IEEE Journal on Selected Areas in Communications*, 19(5):831–837, 2001.

Robin San-Roman, Eliya Nachmani, and Lior Wolf. Noise estimation for generative diffusion models. *arXiv preprint arXiv:2104.02600*, 2021.

Guangming Shi, Yong Xiao, Yingyu Li, and Xuemei Xie. From semantic communication to semantic-aware networking: Model, architecture, and open problems. *IEEE Communications Magazine*, 59(8):44–50, 2021.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. PMLR, 2015.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.

Gokul Sridharan, Abishek Kumarasubramanian, Andrew Thangaraj, and Srikrishna Bhashyam. Optimizing burst erasure correction of ldpc codes by interleaving. In *2008 ieee international symposium on information theory*, pp. 1143–1147. IEEE, 2008.

Tristan SW Stevens, Hans van Gorp, Faik C Meral, Junseob Shin, Jason Yu, Jean-Luc Robert, and Ruud JG van Sloun. Removing structured noise with diffusion models. *arXiv preprint arXiv:2302.05290*, 2023.

Oscar Y Takeshita and Daniel J Costello. New deterministic interleaver designs for turbo codes. *IEEE Transactions on Information Theory*, 46(6):1988–2006, 2000.

Raphael Tang, Linqing Liu, Akshat Pandey, Zhiying Jiang, Gefei Yang, Karun Kumar, Pontus Stenetorp, Jimmy Lin, and Ferhan Ture. What the daam: Interpreting stable diffusion using cross attention. *arXiv preprint arXiv:2210.04885*, 2022.

Alberto Tarable, Sergio Benedetto, and Guido Montorsi. Mapping interleaving laws to parallel turbo and ldpc decoder architectures. *IEEE Transactions on Information Theory*, 50(9):2002–2009, 2004.

Naftali Tishby and Noga Zaslavsky. Deep learning and the information bottleneck principle. In *2015 ieee information theory workshop (itw)*, pp. 1–5. IEEE, 2015.

Naftali Tishby, Fernando C Pereira, and William Bialek. The information bottleneck method. *arXiv preprint physics/0004057*, 2000.

Hanzhang Wang, Deming Zhai, Xiong Zhou, Junjun Jiang, and Xianming Liu. Mix-ddpm: Enhancing diffusion models through fitting mixture noise with global stochastic offset. *ACM Transactions on Multimedia Computing, Communications and Applications*, 20(9):1–24, 2024.

Yihan Wen, Xianping Ma, Xiaokang Zhang, and Man-On Pun. Gcd-ddpm: A generative change detection model based on difference-feature guided ddpm. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

Julian Wyatt, Adam Leach, Sebastian M Schmon, and Chris G Willcocks. Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 650–656, 2022.

Jialong Xu, Bo Ai, Wei Chen, Ang Yang, Peng Sun, and Miguel Rodrigues. Wireless image transmission using deep source channel coding with attention modules. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(4):2315–2328, 2021.

Jialong Xu, Tze-Yang Tung, Bo Ai, Wei Chen, Yuxuan Sun, and Deniz Deniz Gündüz. Deep joint source-channel coding for semantic communications. *IEEE Communications Magazine*, 61(11): 42–48, 2023.

Hanrong Ye and Dan Xu. Diffusionmtl: Learning multi-task denoising diffusion model from partially annotated data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 27960–27969, 2024.

Hoover HF Yin, Ka Hei Ng, Xishi Wang, and Qi Cao. On the minimum delay of block interleaver for batched network codes. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 1957–1961. IEEE, 2019.

Hoover HF Yin, Ka Hei Ng, Xishi Wang, Qi Cao, and Lucien KL Ng. On the memory requirements of block interleaver for batched network codes. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1652–1657. IEEE, 2020.

Fisher Yu, Yinda Zhang, Shuran Song, Ari Seff, and Jianxiong Xiao. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*, 2015.

Zixiang Zhao, Haowen Bai, Yuanzhi Zhu, Jiangshe Zhang, Shuang Xu, Yulun Zhang, Kai Zhang, Deyu Meng, Radu Timofte, and Luc Van Gool. Ddfm: denoising diffusion model for multi-modality image fusion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8082–8093, 2023.

A MORE EXPERIMENTAL RESULTS

A.1 ABLATION STUDY

We studied two asymmetrically distributed noises—Gamma noise and exponential noise. In Fig. 5, we plot the generated outputs of DDPM with Gamma noise, which was normalized to have a power of 1. In Fig. 6, we plot the generated outputs of DDPM with exponential noise, which was also normalized to have a power of 1. As seen in Figs. 5 and 6, asymmetrically distributed noises failed to generate any images.

iteration 900 iteration 950 iteration 1000

Figure 5: Generated outputs of DDPM with Gamma noise.

iteration 900 iteration 950 iteration 1000

Figure 6: Generated outputs of DDPM with exponential noise.

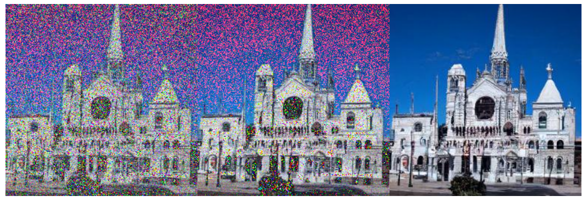
Recognizing that asymmetrical noise tends to degrade generation quality, we adapted these noise types by centering their distributions around zero, effectively making them symmetrical. This was achieved by adjusting each noise sample to have a mean of zero, aligning the noise distribution around $x = 0$. In Figs. 7 and 8, we illustrate the performance of the processed Gamma noise and exponential noise, respectively. Observe that with this adjustment, the previously asymmetrical noise distributions performed comparably to symmetrical noise distributions commonly used in DDPMs. This finding suggests that symmetry around zero plays a crucial role in enabling effective noise removal and high-quality sample generation in DDPMs.

iteration 900 iteration 950 iteration 1000

Figure 7: Generated outputs of DDPM with processed Gamma noise (zero-mean).

A.2 EXPERIMENTAL RESULTS WITH RANDOM SIGNALS

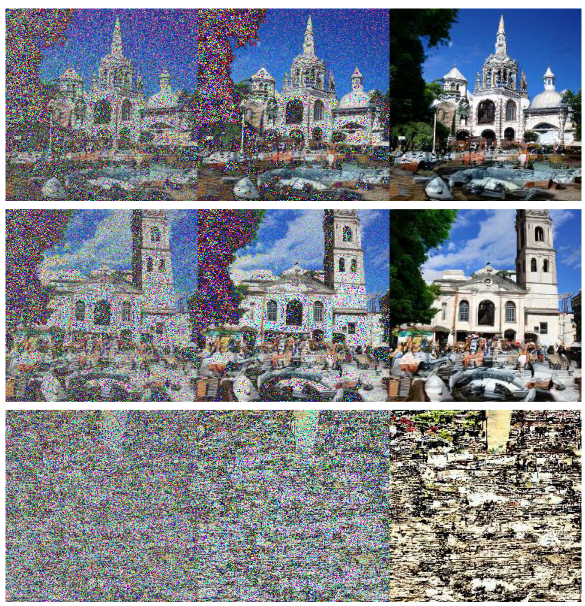
In our study, we extended traditional DDPM sample generation approaches by using structured random signals instead of purely random noise, investigating both symmetric and asymmetric equal-probability signals. Specifically, we used symmetric signals with values from $\{-1, +1\}$ and



iteration 900 iteration 950 iteration 1000

Figure 8: Generated outputs of DDPM with processed exponential noise (zero-mean).

$\{-2, -1, 0, +1, +2\}$, as well as asymmetric signals with values from $\{-12, +2, +4, +6\}$, all centered around a mean of zero. As shown in Fig. 9, symmetric signals enabled the DDPM to generate clear, detailed images, such as church structures, whereas asymmetric signals failed to produce coherent images. This outcome suggests that the symmetry of input signals around zero is critical for generating high-quality samples, as it allows the model to maintain a balanced distribution that aligns well with DDPM's training assumptions. Consequently, our findings highlight that symmetrical, zero-centered signals are effective substitutes for random noise in DDPM sample generation, potentially broadening the scope of signal types usable in diffusion models for structured image synthesis.



iteration 900 iteration 950 iteration 1000

Figure 9: Generated outputs of DDPM using symmetric random signals from $\{-1, +1\}$ (row 1), $\{-2, -1, 0, +1, +2\}$ (row 2), and asymmetric random signals from $\{-12, +2, +4, +6\}$ (row 3).

This observation is important because many real-world signals, such as communication signals (e.g., PAM, QAM), are symmetric (Proakis & Salehi, 2008; Goldsmith, 2005). The ability of symmetric signals to successfully generate samples using DDPM may have implications for semantic communications (Luo et al., 2022).

A.3 Application to Semantic Communication

In this experiment, we consider both random signals and noises. The principle that "symmetry is all you need" in DDPM sample generation can be applied to a variety of real-world applications, including semantic communication. Many communication signals are symmetric, and noise distributions often exhibit symmetry (Proakis & Salehi, 2008; Rappaport, 2024). We have demonstrated that such symmetric signals and noise can successfully generate images using DDPM. Semantic communication is an emerging concept in the communications field, where the focus shifts from transmitting raw data (bits) to transmitting meaningful information (semantics). The goal of semantic communication is to convey the intended meaning or interpretation of the message (Shi et al., 2021; Luo et al., 2022). This approach is particularly useful in applications such as human-machine interaction, natural language processing, and autonomous systems, where context and understanding are critical.

In Fig. 10, we plot the histogram of 4-PAM (Pulse Amplitude Modulation) signals corrupted by additive white Gaussian noise (AWGN) with signal-to-noise ratio (SNR) at $10dB$. The resulting signal follows a Gaussian mixture distribution Liu et al. (2010)Rasmussen (1999). A Gaussian mixture distribution is a weighted sum of multiple Gaussian distributions, each with its own mean and variance. It's clear that the resulting signals are symmetrical.

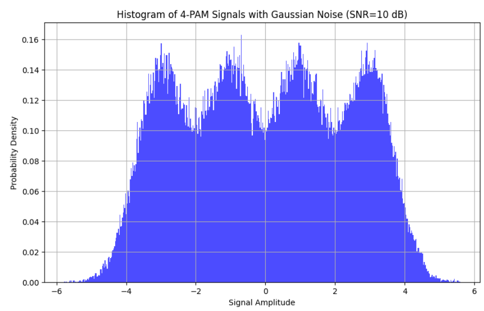
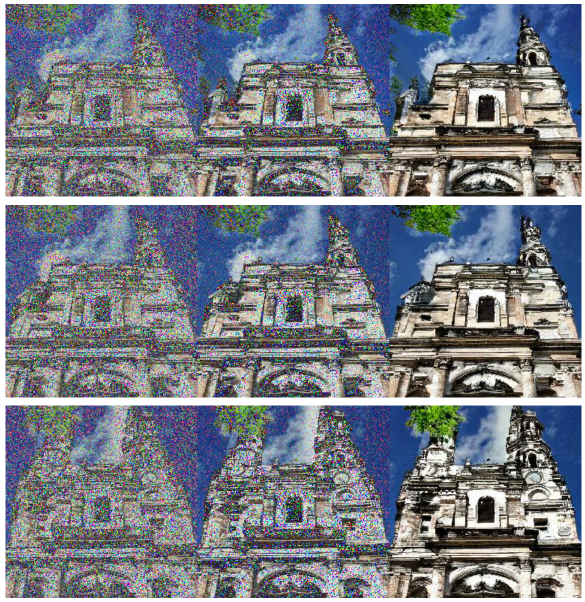


Figure 10: Histogram of 4-PAM signals corrupted by AWGN with $SNR = 10dB$.

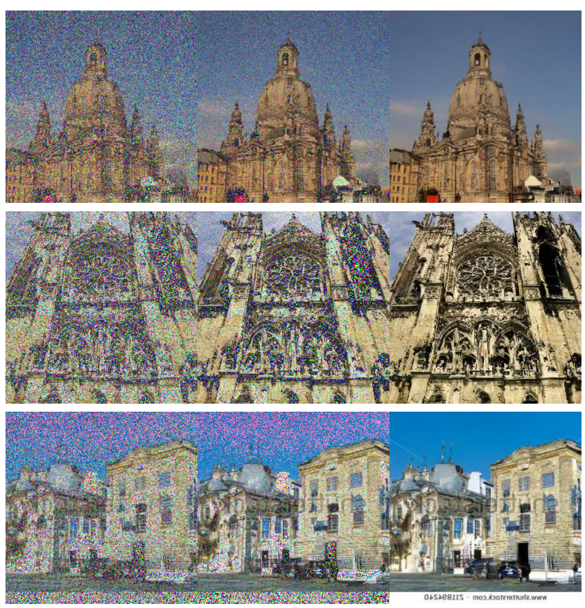
In Fig. 11, we illustrate the generated images of DDPM using 4-PAM signals corrupted by AWGN with SNR at $10dB$, $5dB$, and $0dB$. Observe that the generated image qualities are consistently good regardless the SNR, which demonstrates that DDPM has strong capability to handle noises. Traditionally communication channels with low SNR always perform worse than the channels with higher SNR because signals are corrupted by AWGN, so channel coding is always used to overcome the noise Costello & Forney (2007)Xu et al. (2021). Recently, joint souce and channel coding was applied to semantic communications Xu et al. (2023). We have demonstrated that DDPM performance is not very sensitive to the noise level in semantic communications, so channel coding may not needed for DDPM-based semantic communication.

Interleaver has been widely used in channel coding for wireless communications Tarable et al. (2004)Sridharan et al. (2008). We applied block interleavers with sizes 4×8, 8×8, and 16×8 to 4-PAM signals corrupted by AWGN at $SNR = 10dB$, and the generated images are illustrated in Fig. 12. Observe that different styles of churches are generated. This shows that DDPM-based semantic communication relies more on the space relations of noisy signals rather than on their SNR values in the channel. This observation will help effective content generation in semantic communications. In Liang (2025), Vision Language Models were successfully applied to semantic communication without using channel coding.



iteration 900 iteration 950 iteration 1000

Figure 11: Generated outputs of DDPM using 4-PAM signals+AWGN with $SNR = 10dB$ (row 1), $SNR = 5dB$ (row 2), and $SNR = 0dB$ (row 3).



iteration 900 iteration 950 iteration 1000

Figure 12: Generated outputs of DDPM using 4-PAM signals+AWGN with $SNR = 10dB$ but with different interleavers, 4×8 (row 1), 8×8 (row 2), and 16×8 (row 3).