mathematics

MDPI

Article

Variational Inference and Error Estimation for Inverse Heat Dissipation Models with Quasi-Reversibility Regularization

Zhiqi Tong [1], Zheng Han [1], Qinghua Hu [2] and Jing Zheng [2,*]

[1] School of Mathematics, Hangzhou Normal University, Hangzhou 311121, China
[2] Department of Statistics, Hangzhou Dianzi University, Hangzhou 310018, China
* Correspondence: zhengjing@hdu.edu.cn

Abstract

While the inverse heat dissipation model has shown success in image generation, it is intrinsically ill-posed. To address this issue, we introduce a quasi-reversibility regularization technique and propose a quasi-reversibility-regularized inverse heat dissipation model. We derive the variational formulation of the objective function for the proposed model and establish an upper bound on the error between the regularized solution obtained from noisy observations and the exact solution in the relevant function space. These findings provide a theoretical basis for improving inverse heat dissipation models. The experimental results demonstrate that compared with the original inverse heat dissipation model, the new model achieves an 34.6% reduction in Fréchet Inception Distance on the MNIST dataset.

Keywords: inverse heat dissipation models; variational inference; quasi-reversibility regularization

MSC: 62F15

1. Introduction

Diffusion models have become one of the most active areas of generative modeling, achieving remarkable advances in image processing [1,2]. Inspired by nonequilibrium thermodynamics, Sohl-Dickstein et al. [3] first proposed diffusion models in 2015. These models describe a process in which data are gradually corrupted by noise and subsequently recovered through a reverse denoising process, thereby laying a theoretical foundation for subsequent research. In 2020, Ho et al. [4] introduced the Denoising Diffusion Probabilistic Model (DDPM), which made model training more stable by simplifying the training objective. Building on this work, researchers incorporated conditional variables such as class labels and textual descriptions, leading to the development of conditional diffusion models, including classifier guidance [5] and classifier-free approaches [6]. The former requires an additionally trained classifier, whereas the latter learns conditional and unconditional models simultaneously, albeit at a higher training cost. Subsequently, the Stable Diffusion model [7] compressed images into a latent space for processing and then reconstructed them, substantially reducing the computational burden.

All the aforementioned diffusion models initialize the process with random Gaussian noise and continuously inject stochastic perturbations throughout the reverse path, resulting in different outputs from the same initial condition. Although this generative diversity is advantageous for data augmentation and artistic creation, it poses challenges for inverse imaging problems—such as medical image restoration—which require recovery of a unique ground-truth image.

Check for updates

Academic Editor: Sabri Arik

Received: 21 August 2026
Revised: 19 October 2026
Accepted: 26 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

To mitigate these limitations, Rissanen et al. [8] proposed the Inverse Heat Dissipation Model (IHDM). Whereas conventional diffusion models destroy structure through random noise, IHDM simulates heat diffusion to blur an image and then recovers it according to physical laws, thereby sharpening the image. This approach better captures local structural consistency and preserves more details. However, the reverse denoising process of IHDM is severely ill-posed [9]: mathematically, even slight errors or noise in the input data are amplified exponentially during the reverse process. Thus, a tiny deviation at the beginning of the reverse process can be magnified drastically during recovery, ultimately compromising the quality of the generated images.

Regularization is a primary approach to addressing ill-posed problems. Classical methods include quasi-reversibility regularization [10–13], Tikhonov regularization [14,15], and spectral truncation [16–21]. In this paper, we apply the method proposed in [11] to systematically reformulate IHDM. Specifically, we incorporate a quasi-reversibility regularization term into the original equation, thereby transforming the ill-posed inverse denoising equation into a well-posed one and preventing the amplification of high-frequency noise at its source. We first derive the objective function for model training via variational inference, an aspect not addressed in [13]. We then establish an error estimate for the regularized solution in the image domain and derive an upper bound on the error between the regularized solution under noisy observations and the ground-truth solution. Unlike [13], which estimates this upper bound using the Fourier transform, we use the method of separation of variables. Furthermore, to the best of our knowledge, this is the first application of these methods to generative models.

The contributions of this paper can be summarized as follows:

- A quasi-reversibility regularization module is integrated into IHDM to form a new model that alleviates the ill-posedness of the original IHDM;
- The training objective function of the proposed model is derived using variational inference;
- An upper bound is established on the error between the regularized solution obtained from noisy observations and the exact solution in the relevant function space.

2. IHDM with Quasi-Reversibility Regularization

2.1. Original IHDM

In the original IHDM, the forward process is a gradual heat dissipation process termed "information melting." Let $u : \mathrm{D} \times R_+ \to R$ be the function of a color channel of an image on the two-dimensional plane, where $D = [0, W] \times [0, H]$ is a domain in R^2. The forward process of IHDM is governed by the heat equation with Neumann boundary conditions:

$$\begin{cases} \dfrac{\partial u(x,y,t)}{\partial t} = \Delta u(x,y,t), & (x,y) \in D, t \geq 0 \\ \dfrac{\partial u}{\partial n} = 0, & (x,y) \in \partial D \\ u(x,y,0) = u_0(x,y) \end{cases} \tag{1}$$

where $\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator, ∂D is the boundary of D and n is the out normal vector of ∂D. Since this process is subject to Neumann boundary conditions, which imply that as $t \to \infty$, each color channel of the image eventually averages to the global mean of the original pixel intensities.

To solve Equation (1), we need the eigenvalues and eigenfunctions of the negative Laplacian operator with Neumann boundary conditions:

$$\begin{cases} -\Delta\phi(x,y) = \lambda\phi(x,y), & (x,y) \in D \\ \dfrac{\partial\phi(x,y)}{\partial n} = 0, & (x,y) \in \partial D \end{cases}$$

It is well known that the eigenvalues are

$$\lambda = \lambda_{n,m} = \pi^2\left(\frac{n^2}{W^2} + \frac{m^2}{H^2}\right), n, m \in \mathbb{Z},$$

and the corresponding eigenfunctions are

$$\phi(x,y) = \phi_{n,m}(x,y) = \frac{4}{WH}\cos(\frac{n\pi x}{W})\cos(\frac{m\pi x}{H}).$$

These eigenfunctions form a basis for the solution of Equation (1), and we can write

$$u(x,y,t) = \sum_{n,m} \hat{u}(n,m,t)\phi_{n,m}(x,y).$$

Thus, the heat equation is

$$\sum_{n,m} \frac{d}{dt}\hat{u}(n,m,t)\phi_{n,m}(x,y) = -\sum_{n,m} \lambda_{n,m}\hat{u}(n,m,t)\phi_{n,m}(x,y).$$

Therefore, we get an ordinary equation

$$\frac{d}{dt}\hat{u}(n,m,t) = -\lambda_{n,m}\hat{u}(n,m,t)$$

with initial condition

$$\hat{u}(n,m,0) = \hat{u}_0(n,m) = \int_D u_0(x,y)\phi_{n,m}(x,y)dxdy.$$

It is easy to obtain

$$\hat{u}(n,m,t) = \hat{u}_0(n,m)e^{-\lambda_{n,m}t}.$$

Suppose the frequency-domain data acquired at time T is $\hat{g}(n,m)$, we can directly recover the initial image as

$$\hat{u}_0(n,m) = \hat{g}(n,m)e^{\lambda_{n,m}T}. \tag{2}$$

This expression reveals the "catastrophic" consequence of the reverse denoising process of IHDM, which is an ill-posed problem [12], the factor $e^{\lambda_{n,m}T}$ amplifies noise. Any perturbation at microscopic scales in the observed data, after evolution through the inverse heat equation, is magnified within a very short time into macroscopic, meaningless violent oscillations that completely overwhelm the true image signal.

2.2. The Inverse Heat Diffusion Equation with Quasi-Reversibility Regularization

Quasi-reversibility regularization was first proposed by Lattès and Lions [10]. Its core idea lies in using a well-posed system with a regularization parameter α to address ill-posed problems in differential equations. Based on the theory, we introduce a quasi-reversibility regularization term that depends on the mixed spatial-temporal derivative, $\alpha \cdot \Delta u_t(x,y,t)$, into Equation (1), and construct the following improved forward diffusion equation:

$$\begin{cases} \dfrac{\partial u(x,y,t)}{\partial t} = \Delta u(x,y,t) + \alpha\Delta u_t(x,y,t), & (x,y) \in D, t \geq 0 \\ \dfrac{\partial u}{\partial n} = 0, & (x,y) \in \partial D \\ u(x,y,0) = f(x,y) & (x,y) \in D \end{cases} \tag{3}$$

where the parameter α ($\alpha > 0$) is referred to as the regularization parameter. When α is sufficiently small, the solution of Equation (3) approximates that of Equation (1). This construction carries profound physical implications. The term $\alpha \cdot \Delta u_t$ is a higher-order operator that couples the second-order spatial derivative with the first-order time derivative.

In the forward diffusion process, it acts as an additional damping mechanism, altering the rate of energy dissipation of the original equation. More importantly, in the reverse generative process, this higher-order term effectively modifies the dispersion relation of the equation in the extremely high-frequency regime, thereby suppressing the high-frequency oscillations induced by ill-posedness.

To illustrate how Equation (3) mitigates the amplification of high-frequency noise, we solve the above partial differential equation using a method similar to that used for Equation (1).

Write $u(x, y, t)$ as

$$u(x, y, t) = \sum_{n,m} \hat{u}(n, m, t)\phi_{n,m}(x, y),$$

and the partial differential Equation (3) as

$$\sum_{n,m} \frac{d}{dt}\hat{u}(n, m, t)\phi_{n,m}(x, y) = -\sum_{n,m} \lambda_{n,m}(\hat{u}(n, m, t) + \alpha\frac{d}{dt}\hat{u}(n, m, t))\phi_{n,m}(x, y),$$

and we obtain the following first-order ordinary differential equation in time t:

$$(1 + \alpha\lambda_{n,m})\frac{d}{dt}\hat{u}(n, m, t) = -\lambda_{n,m}\hat{u}(n, m, t)$$

with initial condition

$$\hat{u}(n, m, 0) = \hat{f}(n, m).$$

We have its exact analytical solution of the above differential equation is

$$\hat{u}(n, m, t) = \hat{f}(n, m)e^{-\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}t}. \tag{4}$$

Define the frequency-domain damping function as

$$D_\alpha(n, m) = \frac{\lambda_{n,m}}{1 + \alpha\lambda_{n,m}}$$

and consider the asymptotic behavior of this function across different frequency bands.

1. Low-frequency region ($\lambda_{n,m}$ is small)

In an image, low-frequency components correspond to large-scale contours, background tones and other global structural priors. The low frequency means that $1 + \alpha\lambda_{n,m} \approx 1$, and $D_\alpha(n, m) \approx \lambda_{n,m}$. This indicates that in the low-frequency regime, the diffusion behavior of the quasi-reversibility-regularized heat equation is almost identical to that of the standard heat equation, preserving the multi-scale information erasure and fusion capabilities required by the generative model.

2. High-frequency region ($\lambda_{n,m} \to \infty$)

High-frequency components correspond to edges, textural details, and potential numerical noise in an image. As the frequency tends to infinity, L'Hôpital's rule gives:

$$\lim_{n^2+m^2\to\infty} D_\alpha(n, m) = \lim_{n^2+m^2\to\infty} \frac{\lambda_{n,m}}{1 + \alpha\lambda_{n,m}} = \frac{1}{\alpha}.$$

This shows that quasi-reversibility regularization confines the high-frequency error amplification factor to a finite constant, resolving the divergence issue caused by the ill-posedness of the inverse heat equation. In practice, the damping function $D_\alpha(n, m)$ avoids these problems such as gradient explosion during network training, and the appearance of random high-frequency noise or texture collapse in generated images, ensuring the numerical stability of the entire reverse generation process. By appropriately setting the parameter α, the model can achieve an optimal balance between preserving the physical characteristics of image degradation and controlling the numerical stability of reverse generation.

2.3. Framework Design of the Quasi-Reversibility-Regularized Inverse Heat Diffusion Model

In Section 2.2, we derived the quasi-reversibility-regularized heat equation in the finite domain from the partial differential equation, and proved that this equation can suppress high-frequency oscillations during the reverse evolution. However, in image generation tasks of computer vision, it must be discretized and integrated into the probabilistic framework of deep generative models. In this section, we elaborate on how to construct the forward process and the reverse generative process of the new model with the quasi-reversibility regularization theory.

2.3.1. Quasi-Reversibility-Regularized Forward Blurring Process

First, we rewrite the partial differential equation in the continuous spatial domain as a discrete operator acting on the initial image. Let the input image $\mathrm{u}_0 = \mathrm{f}(\mathrm{x}, \mathrm{y})$. According to the frequency-domain analytical solution derived in Section 2.2, the solution of the regularized heat equation can be expressed as a linear evolution operator acting on the input:

$$\mathrm{u}(\mathrm{t}) = \mathcal{F}_t \mathrm{u}_0 = \sum_{\mathrm{n,m}=1}^{\infty} \hat{\mathrm{f}}(\mathrm{n}, \mathrm{m}) e^{-\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}} t} \phi_{n,m}.$$

In image processing, this evolution operator $\mathcal{F}_t$ can be implemented via the discrete cosine transform (DCT). It is equivalent to applying an effective attenuation function $e^{-\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}} t}$ to the image in the frequency domain. We denote this discrete-time evolution operator as the transition matrix $\mathcal{F}_{t_k} \in R^{N\times N}$.

Next, we discretize the time interval $[0, T]$ into K discrete time steps $t_1,\ t_2, \ldots, t_K$. These time steps correspond to a sequence of latent variables $u_1,\ u_2, \ldots, u_K$. Meanwhile, we superimpose an isotropic Gaussian observation noise with variance σ^2 on the deterministic evolution results. Consequently, the forward process of our model is defined as a non-Markovian joint distribution:

$$\mathrm{q}(\mathrm{u}_{1:\mathrm{K}}|\mathrm{u}_0) \triangleq \prod_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{q}(\mathrm{u}_\mathrm{k}|\mathrm{u}_0) = \prod_{\mathrm{k}=1}^{\mathrm{K}} \mathcal{N}\left(\mathrm{u}_\mathrm{k}\middle|\mathcal{F}_{t_\mathrm{k}} \mathrm{u}_0, \sigma^2 \mathrm{I}\right).$$

The above expression, we write the blurred latent variable u_k at the k time step as a random variable with mean equal to the value solution at t_k, and a noise standard deviation σ. In fact, we make such a choice not only because it is required as a probability model, but also because the solution of the partial differential equation $\mathcal{F}_{t_k} u_0$ is an infinite series, and in computations only finite many terms can be selected, with the high-frequency part being discarded; so the noise, in a certain sense, can be regarded as a supplement to the high-frequency part.

2.3.2. Parameterized Reverse Generative Process

Having defined the forward blurring process, the core task of the generative model is to learn a parameterized reverse process, so as to reconstruct clear images from the extremely blurred prior distribution. At the final time step K of the forward process, the image has been deeply blurred by the regularized heat equation, resulting in the loss of most spatial structural information; and the latent variable u_K collapses to a neighborhood of a low-dimensional mean manifold. We define the marginal distribution of this state as the prior distribution $p(u_K)$ of the generative process. The reverse process is a Markov chain that starts from u_K, progressively generates $u_{K-1},\ u_{K-2}, \ldots$, and eventually recovers u_0. We define the joint probability distribution of the reverse trajectory as follows:

$$\mathrm{p}_\theta(\mathrm{u}_{0:\mathrm{K}}) = \mathrm{p}(\mathrm{u}_\mathrm{K}) \prod_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{p}_\theta(\mathrm{u}_{\mathrm{k}-1}|\mathrm{u}_\mathrm{k}), \tag{5}$$

where $p_\theta(u_{k-1}|u_k) = \mathcal{N}\big(u_{k-1}|\mu_\theta(u_k,k),\delta^2 I\big)$.

Through the design of the forward and reverse frameworks described above, we have transformed the original severely ill-posed inverse heat-conduction partial differential equation into a probabilistic generative Markov chain governed by the quasi-reversibility operator and parameterized and fitted by neural networks. This improved formulation not only provides a rigorous well-posedness guarantee at the mathematical level but also establishes a complete probabilistic foundation for the subsequent application of variational inference to derive the loss function.

3. Variational Inference of the Objective Function

Similarly to standard diffusion probabilistic models, we adopt the variational inference technique and optimize the neural network parameters θ by maximizing the variational lower bound of the marginal likelihood in this section.

Our optimization objective is to minimize the negative log-likelihood $-logp_\theta(u_0)$. According to the basic principle of variational inference, we introduce the forward inference distribution $q(u_{1:K}|u_0)$ as an approximation to the posterior distribution $p_\theta(u_{1:K}|u_0)$. By Jensen's inequality, we obtain an upper bound on the negative log-likelihood:

$$-\log \mathrm{p}_\theta(\mathrm{u}_0) \le \mathrm{E}_{\mathrm{q}(\mathrm{u}_{1:\mathrm{K}}|\mathrm{u}_0)}\left[-\log \frac{\mathrm{p}_\theta(\mathrm{u}_{0:\mathrm{K}})}{\mathrm{q}(\mathrm{u}_{1:\mathrm{K}}|\mathrm{u}_0)}\right].$$

Substituting Equation (5) from Section 2.3.2 into the above equation and using the properties of logarithms to convert the product into a sum, we obtain:

$$-\log \mathrm{p}_\theta(\mathrm{u}_0) \le \mathbb{E}_\mathrm{q}\left[-\log\frac{\mathrm{p}_\theta(\mathrm{u}_{0:\mathrm{K}})}{\mathrm{q}(\mathrm{u}_{1:\mathrm{K}}|\mathrm{u}_0)}\right] = \mathbb{E}_\mathrm{q}\left[-\log\frac{\mathrm{p}_\theta(\mathrm{u}_\mathrm{K})}{\mathrm{q}(\mathrm{u}_\mathrm{K}|\mathrm{u}_0)} - \sum_{\mathrm{k}=2}^{\mathrm{K}}\log\frac{\mathrm{p}_\theta(\mathrm{u}_{\mathrm{k}-1}|\mathrm{u}_\mathrm{k})}{\mathrm{q}(\mathrm{u}_{\mathrm{k}-1}|\mathrm{u}_0)} - \mathrm{logp}_\theta(\mathrm{u}_0|\mathrm{u}_1)\right].$$

We can rewrite the expectation terms above as a combination of multiple Kullback–Leibler (KL) divergences. After marginalizing the redundant integral terms, the objective function can be rigorously decomposed into the following three independent parts:

$$\mathbb{E}_q\left[\underbrace{D_{KL}[\,q(u_K|u_0)\|p(u_K)\,]}_{L_K} + \sum_{k=2}^{K}\underbrace{D_{KL}\big[q(u_{K-1}|u_0)\|p_\theta(u_{k-1}|u_k)\big]}_{L_{K-1}} - \underbrace{logp_\theta(u_0|u_1)}_{L_0}\right],$$

these three parts correspond respectively to the prior matching loss L_K, the reverse evolution transition loss L_{K-1}, and the final reconstruction loss L_0.

1. The prior matching loss L_K

The term measures the discrepancy between the final-state distribution $q(u_K|u_0)$ of the forward process and the assumed prior distribution $p(u_K)$ of the model. Since it contains no neural network parameters θ to be optimized, it does not need to be incorporated into the objective function for gradient-based optimization during the neural network reverse generation process.

2. The reverse evolution transition loss L_{K-1} $(k > 1)$

This is the most crucial part of the variational objective function. It measures the match between the reverse one-step state distribution $p_\theta(u_{k-1}|u_k)$ predicted by the network, conditioned on the current blurred state u_K, and its corresponding forward distribution $q(u_{k-1}|u_0)$.

It is known that both of these distributions are isotropic multivariate Gaussian distributions:

$$q(u_{k-1}|u_0) = \mathcal{N}\big(u_{k-1}\big|\mathcal{F}_{t_{k-1}}u_0,\sigma^2 I\big),$$

$$p_\theta(u_{k-1}|u_k) = \mathcal{N}\big(u_{k-1}\big|\mu_\theta(u_k,k),\delta^2 I\big),$$

According to the KL divergence formula for two multivariate Gaussian distributions:

$$D_{KL}\left[\mathcal{N}\left(\mu_1, \sum_1\right)\middle\|\mathcal{N}\left(\mu_2, \sum_2\right)\right] = \frac{1}{2}\left[\mathrm{tr}\left(\sum_2^{-1}\sum_1\right) - N + (\mu_2-\mu_1)^T\sum_2^{-1}(\mu_2-\mu_1) + \ln\frac{|\sum_2|}{|\sum_1|}\right],$$

we have

$$\mathbb{E}_q[L_{k-1}] = \frac{1}{2}\left(\frac{\sigma^2}{\delta^2}N - N + \frac{1}{\delta^2}\mathbb{E}_{q(u_k|u_0)}\left[\left||\mu_\theta(u_k,k) - \mathcal{F}_{t_{k-1}}u_0\right|\right|_2^2\right] + 2Nlog\frac{\delta}{\sigma}\right).$$

In the above formula, N denotes the total number of pixels in the image, σ is the standard deviation of the observation noise injected in the forward process, and δ is the standard deviation of the sampling noise in the reverse process. A closer inspection reveals that the first, second, and fourth terms are all constant terms independent of the network parameters θ. Therefore, the gradient-based optimization task with respect to θ is entirely concentrated on the third term.

3. The final reconstruction loss L_0

This term characterizes the model's ability to fully recover the extremely clear original image u_0 from the slightly blurred state u_1 at time step t_1. Since $p_\theta(u_0|u_1) = \mathcal{N}(u_0|\mu_\theta(u_1,1),\delta^2 I)$, we have

$$E_q[L_0] = \frac{1}{2\delta^2}E_{q(u_1|u_0)}\left[||\mu_\theta(u_1,1) - u_0||_2^2\right] + Nlog\left(\delta\sqrt{2\pi}\right).$$

In summary, variational inference decomposes the complex reverse generative process into a sequence of continuous deblurring regression tasks, each guided by mean-squared error.

4. Error Estimation

Assume that the ideal clear image function $f(x,y)$ and all solutions of equations belong to the space $L^2(D)$. This is a Hilbert space with inner product

$$\langle f, g\rangle = \int_D f(x,y)g(x,y)dxdy$$

and the corresponding norm of this space is defined as:

$$\|f\| = \left(\iint_D |f(x,y)|^2 dxdy\right)^{\frac{1}{2}}.$$

Since

$$\left\langle \phi_{n,m}, \phi_{n',m'}\right\rangle = \delta_{n,n'}\delta_{m,m'},$$

it is easy to get

$$\langle f, g\rangle = \sum_{n,m}\hat{f}(n,m)\hat{g}(n,m),$$

where

$$f(x,y) = \sum_{n,m}\hat{f}(n,m)\phi_{n,m}(x,y),\quad g(x,y) = \sum_{n,m}\hat{g}(n,m)\phi_{n,m}(x,y),$$

where $\delta_{n,m}$ is Kronecker delta, that means $\delta_{n,m} = 1$ when $n = m$, $\delta_{n,m=0}$ and $n \neq m$. So

$$\|f\| = (\sum_{n,m}\hat{f}(n,m)^2)^{\frac{1}{2}},$$

which facilitates our computation of the error bound in the frequency domain. In real computational environments, the ideal initial data $f(x,y)$ cannot be obtained with absolute

precision. Whether acquired by sensors or computed through floating-point arithmetic on a computer, the image $f(x,y)$ is inevitably subject to a certain degree of perturbation, due to the signal-to-noise ratio limitation of physical devices and the rounding errors inherent in low-level computational logic. We denote the actual measured data with observation noise as $f^{\varepsilon}(x,y)$. In the theoretical framework of this paper, we assume that the measurement error is bounded in the L^2 norm and satisfies the following condition:

$$\|f - f^{\varepsilon}\| \leq \varepsilon. \tag{6}$$

Given the observation error ε in the data, how large is the deviation between the approximate solution obtained by the quasi-reversibility regularization method (with parameter α) and the ideal exact solution of Equation (1)? The following theorem addresses this question.

Theorem 1. *Error estimation of the quasi-reversibility-regularized solution. Let $u(\cdot,\cdot,t)$ be the solution of Equation (1) driven by the exact initial data $u(x,y,T) = f(x,y)$. Let $u_{\alpha,\varepsilon}(\cdot,\cdot,t)$ be the quasi-reversibility-regularized solution of Equation (2) driven by the noisy data $u_{\alpha,\varepsilon}(x,y,T) = f^{\varepsilon}(x,y)$ with regularization parameter α. Under the conditions (6) and $\|u(x,y,0)\| \leq E$ and $t < T$, the following estimation holds:*

$$\left\|u_{\alpha,\varepsilon}(\cdot,\cdot,t) - u(\cdot,\cdot,t)\right\| \leq \varepsilon e^{\frac{T-t}{\alpha}} + E \cdot \frac{4\alpha(T-t)e^{-2}}{t}.$$

Proof of Theorem 1. Letting $u_{\alpha}(\cdot,\cdot,t)$ be the solution of Equation (3) driven by the exact initial data $f(x,y)$, we have

$$\begin{aligned}\left\|u_{\alpha,\varepsilon}(\cdot,\cdot,t) - u(\cdot,\cdot,t)\right\| &= \left\|u_{\alpha,\varepsilon}(\cdot,\cdot,t) - u_{\alpha}(\cdot,\cdot,t) + u_{\alpha}(\cdot,\cdot,t) - u(\cdot,\cdot,t)\right\| \\ &\leq \left\|u_{\alpha,\varepsilon}(\cdot,\cdot,t) - u_{\alpha}(\cdot,\cdot,t)\right\| + \left\|u_{\alpha}(\cdot,\cdot,t) - u(\cdot,\cdot,t)\right\| \\ &= A_1 + A_2\end{aligned}$$

The data propagation error A_1 measures the deviation in the solution caused by the initial noise $f^{\varepsilon} - f$ under the same regularized evolution operator. By the method we have used above of the regularized solution derived in Section 2.2 (Equation (4)), it is easy to have:

$$\begin{aligned}A_1 &= \left(\sum_{n,m}\left(e^{\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)}\hat{f}^{\varepsilon}(n,m) - e^{\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)}\hat{f}(n,m)\right)^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{n,m} e^{\frac{2\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)}\left(\hat{f}^{\varepsilon}(n,m) - \hat{f}(n,m)\right)^2\right)^{\frac{1}{2}} \\ &\leq \sup_{\lambda_{n,m}} e^{\frac{\lambda_{n,m}(T-t)}{1+\alpha\lambda_{n,m}}}\left(\sum_{n,m}\left(\hat{f}^{\varepsilon}(n,m) - \hat{f}(n,m)\right)^2\right)^{\frac{1}{2}} \leq \varepsilon e^{\frac{T-t}{\alpha}}\end{aligned} \tag{7}$$

For the last step, we use the result that $e^{\frac{\lambda_{n,m}(T-t)}{1+\alpha\lambda_{n,m}}}$ is increasing about $\lambda_{n,m} \geq 0$, so

$$\sup_{\lambda_{n,m}} e^{\frac{\lambda_{n,m}(T-t)}{1+\alpha\lambda_{n,m}}} = \lim_{\lambda_{n,m}\to\infty} e^{\frac{\lambda_{n,m}(T-t)}{1+\alpha\lambda_{n,m}}} = e^{\frac{T-t}{\alpha}}$$

The term A_2 measures the structural inherent deviation between the analytical solution of the regularized partial differential Equation (2) and that of the standard heat Equation (1) under ideal noise-free data. Expanding in the frequency domain, we have:

$$\begin{aligned} A_2 &= \left(\sum_{n,m} \left(e^{\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)} \hat{f}(n,m) - e^{\lambda_{n,m}(T-t)} \hat{f}(n,m) \right)^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{n,m} \left(e^{\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)} - e^{\lambda_{n,m}(T-t)} \right)^2 \hat{f}^2(n,m) \right)^{\frac{1}{2}} \\ &= \left(\sum_{n,m} \left(e^{-\lambda_{n,m}t} - e^{\frac{\lambda_{n,m}}{1+\alpha\lambda_{n,m}}(T-t)-\lambda_{n,m}T} \right)^2 e^{2\lambda_{n,m}T} \hat{f}^2(n,m) \right)^{\frac{1}{2}} \\ &= \left(\sum_{n,m} e^{-2\lambda_{n,m}t} \left(1 - e^{\frac{-\alpha\lambda_{n,m}^2}{1+\alpha\lambda_{n,m}}(T-t)} \right)^2 e^{2\lambda_{n,m}T} \hat{u}^2(n,m,0) \right)^{\frac{1}{2}} \end{aligned} \tag{8}$$

Here we have used equation $\hat{u}(n,m,0) = e^{\lambda_{n,m}T}\hat{f}(n,m)$ in the last step of the above equation. Write the value of $\lambda_{n,m} = \pi^2\left(\frac{n^2}{W^2} + \frac{m^2}{H^2}\right)$ as $\mathrm{k} = \xi^2+\eta^2 (\mathrm{k} \geq 0)$ in (8) and we obtain:

$$\begin{aligned} A_2 &\leq E \cdot \sup_{\xi,\eta\in R^2} \left| e^{-(\xi^2+\eta^2)t} \left(1 - e^{-\frac{\alpha(\xi^2+\eta^2)^2(T-t)}{1+\alpha(\xi^2+\eta^2)}} \right) \right| \\ &\leq E \cdot \sup_{\xi,\eta\in R^2} \left(e^{-(\xi^2+\eta^2)t} \cdot \frac{\alpha(\xi^2+\eta^2)^2(T-t)}{1+\alpha(\xi^2+\eta^2)} \right) \\ &\leq E \cdot \sup_{\xi,\eta\in R^2} \left(\alpha(T-t)(\xi^2+\eta^2)^2 e^{-(\xi^2+\eta^2)t} \right). \end{aligned} \tag{9}$$

To obtain the above supremum, we consider the maximum of the function $G(\mathrm{k}) = \alpha(T-t)\mathrm{k}^2 e^{-\mathrm{kt}}$. Compute the first derivative of $G(\mathrm{k})$ with respect to the variable k and set it to zero in order to find the critical points:

$$\mathrm{G}'(\mathrm{k}) = \alpha(\mathrm{T}-\mathrm{t})\left(2\mathrm{k}e^{-\mathrm{kt}} - \mathrm{k}^2\mathrm{t}e^{-\mathrm{kt}}\right) = \alpha(\mathrm{T}-\mathrm{t})\mathrm{k}e^{-\mathrm{kt}}(2-\mathrm{kt}) = 0,$$

solving the above equation yields the critical point at which the function attains its maximum as $k = \frac{2}{t}$. Substituting this critical value into (9), we have:

$$A_2 \leq E \cdot \frac{4\alpha(T-t)}{t} e^{-2}. \tag{10}$$

Combining (7) and (10), we finally obtain the overall error estimate of the quasi-reversibility-regularized solution with respect to the exact solution:

$$\left\| u_{\alpha,\varepsilon}(\cdot,\cdot,t) - u(\cdot,\cdot,t) \right\| \leq \varepsilon e^{\frac{\mathrm{T}-\mathrm{t}}{\alpha}} + \mathrm{E} \cdot \frac{4\alpha(T-t)e^{-2}}{t}. \tag{11}$$

The proof of Theorem 1 is complete. □

https://doi.org/10.3390/math14213965

From the estimation results, the first term A_1 is bounded by a constant multiple of ε, whereas the second term A_2 is more problematic: A_2 tends to infinity as $t \to 0$. In fact, for the values of T and α we tested, most of the error estimates fall between 10^9 and 10^{27}, which makes the idea of constructing a forward process using the heat equation with dynamic variance difficult to realize in practice. In the inverse problem, this remains a dilemma, and additional information is needed to improve the results. Therefore, in our design, the forward process is IHDM with quasi-reversibility regularization.

5. Experimental Results

In this section, we evaluate our model in the MNIST dataset. MNIST contains 70,000 grayscale hand-written digit images of 28 × 28 pixels, among which 60,000 are used for training and 10,000 for testing. Despite its low resolution and lack of color information, it possesses clear edges and high-contrast structures, making it well-suited as a benchmark for verifying the model's reconstruction capability for basic geometric structures.

The hyper-parameters related to our model are the same as [8] on MNIST. For training, given the medium-to-low resolution of the datasets, a relatively high learning rate of 1×10^{-4} is used to accelerate convergence. Adam is adopted with default momentum parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and a numerical stability term $\varepsilon_0 = 1 \times 10^{-8}$. Gradient norm clipping (threshold 1.0) and a linear learning-rate warm-up over the first 5000 iterations are applied to prevent early gradient explosion. For generation, the pseudo-inverse regularization parameter $\alpha = 0.5$. The total number of evolution steps is $K = 100$, and the constant noise variance ($\sigma = 0.01$). The corresponding reverse-sampling noise still obeys the empirical amplification ratio $\delta = 1.25\sigma$.

We consider the original IHDM as baselines. Both models are trained under identical computational environments, dataset splits, and the classical U-Net neural network architecture as the backbone for 50,000 training iterations. Dataset is availiable at https://ossci-datasets.s3.amazonaws.com/mnist/ through torchvision at 16 October 2026.

We evaluate the FID scores with five independent random seeds to eliminate incidental errors. The results are presented in Table 1. The mean FID score of the proposed method achieves a substantial reduction of 34.57% compared with the baseline model. A more important finding is that the FID scores of the proposed model remain quite stable, which indicates that the proposed method substantially improves the stability of the model.

Table 1. FID scores of different diffusion models ($\alpha = 0.5$).

Random Seed	Original IHDM	Ours
12	22.688	16.117
22	22.345	15.776
30	22.119	15.444
42	23.359	15.533
50	30.036	16.003
Mean (SD)	24.109 (3.346)	15.774 (0.290)

We investigate the relationship between the parameter α and their robustness to different choices. We run models and calculated FID scores with $\alpha \in \{0.1, 0.5, 1, 2, 5\}$, which are shown in Figure 1. Note that the proposed model is quite robust to α, as long as it is not too close to zero, and the optimal value is about $\alpha = 1$.

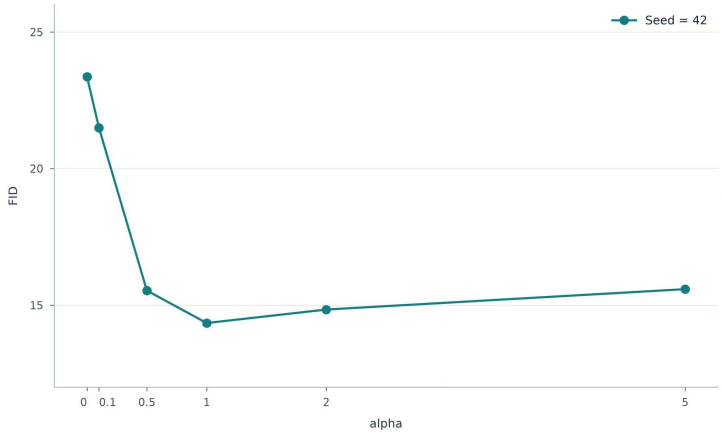


Figure 1. FID scores for MNIST trained with different α values.

6. Conclusions

This paper presents a quasi-reversibility-regularized inverse heat diffusion model for image generation. The proposed framework bridges deterministic PDE evolution and probabilistic generative modeling. Theoretically, we prove that the regularized equation yields a well-posed inverse solution. We further derive the variational objective function and establish a rigorous error bound, which guides the design of our dynamic noise-scheduling strategy.

Despite these promising results achieved in the work, several directions remain open for future exploration. A valuable research direction is to treat the regularization parameter α and the noise floor ε as learnable variables within the end-to-end training framework, allowing the model to adaptively calibrate its diffusion dynamics according to the spectral properties of input data. Furthermore, extending the proposed framework to higher-resolution image generation and video generation tasks constitutes a promising direction for further research.

Author Contributions: Conceptualization, Z.H. and J.Z.; methodology, Z.H. and Z.T.; software, Q.H.; validation, Z.T. and Q.H.; formal analysis, Z.T.; writing—original draft preparation, Z.T.; writing—review and editing, Z.H.; visualization, Q.H.; supervision, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Social Science Fund of China, grant number 25BTJ023.

Data Availability Statement: The data presented in this study are openly available in [mnist] at [https://ossci-datasets.s3.amazonaws.com/mnist/, (accessed on 16 October 2025)]. The model validation results generated during this study are available from the corresponding author upon reasonable request.

Acknowledgments: We would like to thank the editors and the anonymous reviewers for furnishing us with helpful reviewed comments, which have been valuable for revising and improving our manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. He, C.; Shen, Y.; Fang, C.; Xiao, F.; Tang, L.; Zhang, Y.; Zuo, W.; Guo, Z.; Li, X. Diffusion Models in Low-Level Vision: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 4630–4651. [CrossRef] [PubMed]
2. Li, X.; Ren, Y.; Jin, X.; Lan, C.; Wang, X.; Zeng, W.; Wang, X.; Chen, Z. Diffusion Models for Image Restoration and Enhancement: A Comprehensive Survey. *Int. J. Comput. Vis.* **2025**, *133*, 8078–8108. [CrossRef]
3. Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. *Int. Conf. Mach. Learn.* **2015**, *37*, 2256–2265.

4. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.
5. Dhariwal, P.; Nichol, A. Diffusion models beat gans on image synthesis. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 8780–8794.
6. Ho, J.; Salimans, T. Classifier-free diffusion guidance. *arXiv* **2022**, arXiv:2207.12598.
7. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. *arXiv* **2021**, arXiv:2112.10752. [CrossRef]
8. Rissanen, S.; Heinonen, M.; Solin, A. Generative modelling with inverse heat dissipation. *arXiv* **2022**, arXiv:2206.13397.
9. Hadamard, J. Sur les problèmes aux dérivées partielles et leur signification physique. *Princet. Univ. Bull.* **1902**, *13*, 49–52.
10. Lattès, R.; Lions, J.-L. *The Method of Quasi-Reversibility: Applications to Partial Differential Equations*; Elsevier: New York, NY, USA, 1969.
11. Gajewski, H.; Zacharias, K. Zur Regularisierung einer Klasse nichtkorrekter Probleme bei Evolutionsgleichungen. *J. Math. Anal. Appl.* **1972**, *38*, 784–789. [CrossRef]
12. Koba, H.; Matsuoka, H. Generalized Quasi-Reversibility Method for a Backward Heat Equation with a Fractional Laplacian. *Analysis* **2015**, *35*, 47–57. [CrossRef]
13. Hou, J.; Xiong, X. The Regularization Methods for Solving the Time-Inverse Heat Conduction Problems in Two Dimensions and Error Estimation. *J. Sichuan Norm. Univ. Nat. Sci.* **2023**, *46*, 497–502.
14. Tikhonov, A.N.; Goncharsky, A.V.; Stepanov, V.V.; Yagola, A.G. *Numerical Methods for the Solution of Ill-Posed Problems*; Kluwer: Dordrecht, The Netherlands, 1995.
15. Tikhonov, A.N.; Arsenin, V.Y. *Solutions of Ill-Posed Problems*; V. H. Winston & Sons: Washington, DC, USA, 1977.
16. Nair, M.T. On Truncated Spectral Regularization for an Ill-Posed Evolution Equation. *arXiv* **2019**, arXiv:1907.11076.
17. Trong, D.D.; Tuan, N.H. Regularization and Error Estimates for Nonhomogeneous Backward Heat Problems. *Electron. J. Differ. Equ.* **2006**, *4*, 1–10.
18. Singhal, M.; Goyal, K.; Singla, R.K. A Review of Regularization Strategies and Solution Techniques for Ill-Posed Inverse Problems, with Application to Inverse Heat Transfer Problems. *Rev. Math. Phys.* **2024**, *36*, 2330007. [CrossRef]
19. Brociek, R.; Pleszczyński, M.; Miarka, J.; Goik, M. Enhanced Hybrid Algorithms for Inverse Problem Solutions in Computed Tomography. *Appl. Syst. Innov.* **2025**, *8*, 31. [CrossRef]
20. Grzymkowski, R.; Pleszczyński, M.; Słota, D. Application of the Adomian Decomposition Method for Solving the Heat Equation in the Cast-Mould Heterogeneous Domain. *Arch. Foundry Eng.* **2009**, *9*, 57–62.
21. Woodbury, K.A.; Najafi, H.; de Monte, F. *Inverse Heat Conduction: Ill-Posed Problems*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2023. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.