Initial State Reconstruction on Graphs

Vo Anh Khoa, Mai Thanh Nhat Truong, Imhotep Hogan, and Roselyn Williams

Abstract. The presence of noise is an intrinsic problem in acquisition processes for digital images. One way to enhance images is to combine the forward and backward diffusion equations. However, the latter problem is well known to be exponentially unstable with respect to any small perturbations on the final data. In this scenario, the final data can be regarded as a blurred image obtained from the forward process, and that image can be pixelated as a network. Therefore, we study in this work a regularization framework for the backward diffusion equation on graphs. Our aim is to construct a spectral graph-based solution based upon a cut-off projection. Stability and convergence results are provided together with some numerical experiments.

1. Statement of the initial state reconstruction on graphs

Regularization of the restoration problem is a denoising technique used in the hope that it can retain the image crucial signal features in the presence of noise (motion artifacts, blurring, distortion) during procurement course. It is worth mentioning that even modern cameras which are able to acquire high resolution images still, nowadays, produce noisy outputs. Noise in those images usually constitutes high frequencies; therefore, one way to denoise the images is to smooth “away” the noise. It is well known that one of the standard filtering processes is Gaussian smoothing. In particular, let $g_0$ be a gray scale image as a real-valued mapping; see Figure 1 for an overview of the surface representation of a gray scale image. Let $g_0^\varepsilon$ be its noisy image, where $\varepsilon \in (0, 1)$ represents the noise level. Suppose that these $g_0, g_0^\varepsilon$ belong to $L^2(\mathbb{R}^2)$. Then, the Gaussian smoothing constructs a smoothed version of $g_0^\varepsilon$ by the convolution $G_\sigma \ast g_0^\varepsilon \in C^\infty(\mathbb{R}^2)$,

$$(G_\sigma \ast g_0^\varepsilon)(x) = \int_{\mathbb{R}^2} \frac{1}{2\pi\sigma^2} e^{-|x-y|^2/(2\sigma^2)} g_0^\varepsilon(y) \, dy, \quad \sigma > 0.$$  

In terms of Partial Differential Equations (PDEs), the Gaussian convolution is obtained by solving a sequence of the parabolic PDEs forward in time. Indeed, if
we consider
\[
\begin{align*}
\frac{\partial}{\partial t} u - \Delta u &= 0 \quad \text{in } \mathbb{R}^2 \times (0, T), \\
u (x, 0) &= g_0^\varepsilon (x) \quad \text{in } \mathbb{R}^2,
\end{align*}
\]
then its solution can be expressed as \( u(x, t) = (G_{\sqrt{Tt}} \ast g_0^\varepsilon)(x) \) for \( x \in \mathbb{R}^2, t > 0 \).

We remark that due to the smoothness property of the parabolic equation, the smoothing process will destroy image characteristics such as lines and edges; see Figure 2. Stemming from the idea in [5, 16], we combine the forward and backward diffusion equations to enhance images. Henceforth, it brings us back to the initial state reconstruction problem in which the noisy final condition as our blurred image is \( g_T^\varepsilon = G_{\sqrt{Tt}} \ast g_0^\varepsilon(x) \). Unique continuity of this problem is studied in [4].

We would like to mention that image can be pixelated as a network; see Figure 3. Therefore, one can consider the linear diffusion PDE (1.1) on graphs. In this work, we concentrate on the simple graph \( G = (V, E) \). It is an ordered pair of sets, where

- \( E \subset \{ \{x_i, x_j\} \mid x_i, x_j \in V, i \neq j\} \) is a set of undirected edges;
- \( V \) is a set of vertices \( \{x_i\}_{1 \leq i \leq n} \) with \( n = |V| < \infty \) being the number of vertices.
Consider \( u_i (t) = u(x_i, t) \) as the pixel value at time \( t \) and vertex \( x_i \). Note here that the number \( n \) should be fixed because of the fixed resolution of the image. The Graph Laplacian is defined as \( L = D - A \). Here, \( A \in \mathbb{R}^{n \times n} \) is the adjacency matrix, whose the entries \( A_{ij} \) are given by

\[
A_{ij} = \begin{cases} 
1 & \text{if } \{x_i, x_j\} \in E, \\
0 & \text{otherwise}.
\end{cases}
\]

From knowledge of the adjacency matrix \( A \), we define the degree \( D_i = D(x_i) > 0 \) of a vertex \( x_i \) in the simple graph \( G \) is the number of vertices in \( G \) that are adjacent to \( x_i \). As a consequence, it holds true that

\[
\sum_{j=1}^{n} A_{ij} = D(x_i) =: D_i \quad \text{for } 1 \leq i \leq n.
\]

By this way, \( D \in \mathbb{R}^{n \times n} \) is the degree matrix which is diagonal with \( D_i \) being the degree of the \( i \)-th vertex. Henceforth, \( D \in \mathbb{R}^{n \times n} \) is of the following form:

\[
D = \begin{bmatrix}
D_1 & 0 & \cdots & 0 \\
0 & D_2 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & D_{n-1} & 0 \\
0 & 0 & \cdots & D_n
\end{bmatrix}.
\]

In this setting, our Graph Laplacian matrix \( L \in \mathbb{R}^{n \times n} \) is symmetric and its entries are given by

\[
L_{ij} = \begin{cases} 
-1 & \text{if } i \neq j \text{ and } \{x_i, x_j\} \in E, \\
D_i & \text{if } i = j, \\
0 & \text{otherwise}.
\end{cases}
\]

Let \( U = U(t) \) now be the \( n \)-dimensional vector of \( u_i (t) \) for \( 1 \leq i \leq n \), i.e.

\[
U = \begin{pmatrix}
u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n
\end{pmatrix}^T.
\]
Suppose that $U_T^g$ is the blurred image $g_T$ pixelated on the simple graph under study. In this scenario, our initial state reconstruction reads as

\[
\begin{cases}
\frac{dU}{dt}(t) + LU(t) = 0 & \text{for } t \in (0, T), \\
U(T) = U_T^g \in \mathbb{R}^n.
\end{cases}
\]

(1.2)

2. Fourier instability of the initial state reconstruction

In this section, we show the natural instability of the initial state reconstruction problem (1.2). When doing so, we need to introduce the standard inner product

\[
\langle A, B \rangle = A^T B = \sum_{i=1}^{n} a_i b_i \quad \text{for } A, B \in \mathbb{R}^n.
\]

Hereby, the corresponding $l^2$-norm $\|A\| = \sqrt{\langle A, A \rangle}$ is given by

\[
\|A\|^2 = \sum_{i=1}^{n} a_i^2.
\]

Next, we introduce in Definition 1 regarding an eigenvector of a matrix. Then, we state the standard spectral theorem without proof in Theorem 1. These are essential in our way to show the Fourier instability of problem (1.2).

Definition 1. An eigenvector of a matrix $A \in \mathbb{R}^{n \times n}$ is a (non-zero) vector $\phi$ such that $A\phi = \lambda \phi$ for some scalar $\lambda$. The value $\lambda$ is the corresponding eigenvalue, which is a root of the characteristic polynomial of $A$, $p_A = \det (\lambda I - A)$ where $I \in \mathbb{R}^{n \times n}$ denotes the identity matrix.

Theorem 1. Let $M \in \mathbb{R}^{n \times n}$ be a real and symmetric matrix. Then it has $n$ orthogonal eigenvectors and the eigenvalues are real and non-negative.

Since the Graph Laplacian matrix $L$ is real and symmetric, we know that there is an orthonormal basis $\phi_1, \ldots, \phi_n$ of $\mathbb{R}^n$ such that each $\phi_j$ is an eigenvector of $L$. Let $\lambda_j$ be the real eigenvalue corresponding to $\phi_j$, i.e. $L\phi_j = \lambda_j \phi_j$ for $1 \leq j \leq n$. Moreover, we have $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, and the sum of these eigenvalues is twice the number of edges of the simple graph under consideration (cf. (1.2)). We find $U(t)$ as a linear combination of the eigenvectors:

\[
U(t) = \sum_{i=1}^{n} B_i(t) \phi_i.
\]

(2.1)

It is clear that in this form, $U(t) \in \mathbb{R}^n$ for every $t$. Now, we multiply both sides of (1.2) (in the sense of the standard product) by $\phi_j \in \mathbb{R}^n$ and thus, arrive at

\[
\frac{d}{dt} \langle U(t), \phi_j \rangle + \langle LU(t), \phi_j \rangle = 0,
\]

equivalently,

\[
\frac{d}{dt} \langle U(t), \phi_j \rangle + \langle U(t), L\phi_j \rangle = 0.
\]

Since $L\phi_j = \lambda_j \phi_j$, we obtain the following differential equation:

\[
\frac{d}{dt} \langle U(t), \phi_j \rangle + \lambda_j \langle U(t), \phi_j \rangle = 0,
\]

(2.2)

associated with the following final condition (cf. (1.2)),

\[
\langle U(T), \phi_j \rangle = \langle U_T^g, \phi_j \rangle.
\]

(2.3)
Solving (2.2)-(2.3) as an initial-value differential problem, we obtain
\begin{equation}
\langle U(t), \phi_j \rangle = e^{\lambda_j (T-t)} \langle U_T^\varepsilon, \phi_j \rangle ,
\end{equation}
Then plugging the spectral form of $U(t)$ (cf. (2.1)) in (2.4) and using the orthogonality of $\phi_i$ yield
\begin{equation}
e^{\lambda_j (T-t)} \langle U_T^\varepsilon, \phi_j \rangle = \langle U(t), \phi_j \rangle = \sum_{i=1}^{n} B^\varepsilon_i (t) \langle \phi_i, \phi_j \rangle = B^\varepsilon_j (t).
\end{equation}
Henceforth, our solution $U(t)$ to the initial state reconstruction problem (1.2) can be computed as
\begin{equation}
U(t) = \sum_{i=1}^{n} e^{\lambda_i (T-t)} \langle U_T^\varepsilon, \phi_i \rangle \phi_i.
\end{equation}
Hence, the initial information of an pixelated noisy image, denoted by $U^\varepsilon_0 = U^\varepsilon(0)$ is formulated by
\begin{equation}
U^\varepsilon_0 = \sum_{i=1}^{n} e^{\lambda_i T} \langle U_T^\varepsilon, \phi_i \rangle \phi_i.
\end{equation}
Observe in (2.5) that the exponential kernel $e^{\lambda_i (T-t)}$ is the main natural factor of the Fourier instability. In this sense, the solution is exponentially unstable when the number of vertices, which is $n$, becomes larger and larger for a better resolution of image. If we keep refining the mesh of discretization or, in other words, increasing the number of vertices, the reconstruction process will be worse regardless of any numerical method applied.

Our aim in this work is to study a stable approximate solution of the discrete model on graphs (1.2). This new research will formulate a new numerical approach in image denoising among well known approaches such as Perona–Malik [13], p-Laplacian [12], Mumford–Shah [6] and deep learning [14].

3. A spectral graph-based approximate solution

For each noise level $\varepsilon$, we consider the set of admissible eigenvalues $\Theta (\varepsilon) = \{ i \in \mathbb{N}^*: \lambda_i \leq M_\varepsilon \}$. Here, $M_\varepsilon = M(\varepsilon) > 0$ is the so-called regularization parameter which will be chosen later. It should be that $|\Theta (\varepsilon)| \leq |V|$ as the resolution of image is fixed. If there exists $\varepsilon_0$ such that $|\Theta (\varepsilon_0)| > |V|$, for any $\varepsilon \geq \varepsilon_0$ we set $M_\varepsilon = \lambda_n$, where we recall that $\lambda_n$ is the largest eigenvalue of the Graph Laplacian matrix. We thereby construct the following cut-off projection in which the number of vertices $n$ is selected appropriately in terms of $\varepsilon$.
\begin{equation}
P^\varepsilon U(t) = \sum_{i \in \Theta (\varepsilon)} e^{\lambda_i (T-t)} \langle U(T), \phi_i \rangle \phi_i.
\end{equation}

Remark 3.1. The underlying cut-off projection is one of the conventional Fourier-based approaches that is “computable” in the regularization theory; cf. [9, 15]. We remark that the existing literature on regularization of terminal-value problems is huge. For brevity, we mention here some PDE-based approaches (e.g. quasi-reversibility method [8, 10], quasi-boundary value method [2], and references cited therein) and some minimization-based approaches (see in [3] regarding methods of total variation, sparse representation and others, and convexification in [7]).
Theorem 2. Consider \( U(T) \in \mathbb{R}^n \). Then, it holds true that
\[
\| \mathbf{P}^\varepsilon U(t) \| \leq e^{M_\varepsilon(T-t)} \| U(T) \| ,
\]

Proof. The proof can be done directly by taking into account the formulation of the cut-off projection in (3.1). Indeed, by the orthogonality of the eigenvectors, we have
\[
\| \mathbf{P}^\varepsilon U(t) \|^2 = \sum_{i \in \Theta(\varepsilon)} e^{2\lambda_i(T-t)} \| (U(T), \phi_i) \| (\phi_i, \phi_i)
\]
Thus, it yields that
\[
\| \mathbf{P}^\varepsilon U(t) \|^2 \leq e^{2M_\varepsilon(T-t)} \| U(T) \|^2 ,
\]
which completes the proof of the theorem. \( \square \)

Theorem 2 shows that the cut-off projection solution is bounded at each noise level. This result also proves the stability of the approximate solution when the input, \( U(T) \), is perturbed by noise. Let \( U_0 \in \mathbb{R}^n \) be the pixel values of the ideal image \( g_0 \), and let \( U_\varepsilon^0 \in \mathbb{R}^n \) be the pixel values of the noisy image \( g_\varepsilon^0 \). We assume that
\[
\| U_0 - U_\varepsilon^0 \| \leq \varepsilon .
\]
Henceforth, we obtain
\[
\| U_T - U_\varepsilon^T \| \leq \| U_0 - U_\varepsilon^0 \| \leq \varepsilon ,
\]
where \( U_T \) and \( U_\varepsilon^T \) are, respectively, the terminal pixel values of the ideal and noisy images after the forward solver (1.1).

Theorem 3. Let \( U_T, U_\varepsilon^T \in \mathbb{R}^n \) be the terminal data of the ideal solution \( U(t) \) and the noisy solution \( U_\varepsilon(t) \), respectively. These data satisfy (3.2). Then, the following stability estimate holds true:
\[
\| \mathbf{P}^\varepsilon U(t) - \mathbf{P}^\varepsilon U_\varepsilon(t) \| \leq e^{M_\varepsilon(T-t)} \varepsilon .
\]

Proof. The proof is straightforward because
\[
\| \mathbf{P}^\varepsilon U(t) - \mathbf{P}^\varepsilon U_\varepsilon(t) \|^2 = \sum_{i \in \Theta(\varepsilon)} e^{2\lambda_i(T-t)} \| (U_T - U_\varepsilon^T, \phi_i) \|^2 
\leq e^{2M_\varepsilon(T-t)} \varepsilon^2 .
\]

Now, we are in a position to state the following theorem for the convergence rate of the cut-off projection.

Theorem 4. Let \( \mathbf{P}^\varepsilon U_\varepsilon(t) \in \mathbb{R}^n \) be the stable approximate solution of \( U(t) \) associated with the terminal data \( U_\varepsilon^T \). Then, the following uniform-in-time estimate holds:
\[
\| U(t) - \mathbf{P}^\varepsilon U_\varepsilon(t) \| \leq M_\varepsilon^{-1} \left\| \frac{d}{dt} U(t) \right\| + e^{M_\varepsilon(T-t)} \varepsilon .
\]
Furthermore, by choosing
\[ M_\varepsilon = \frac{1}{T} \ln (\varepsilon^{-\gamma}), \quad \gamma \in (0, 1), \]
we obtain
\[ \| U(t) - P^\varepsilon U^\varepsilon(t) \| \leq \frac{T}{\ln (\varepsilon^{-\gamma})} \left\| \frac{d}{dt} U(t) \right\| + \varepsilon^{1-\gamma (1-\frac{T}{T})}. \]  

\[ (3.5) \]

**Proof.** In view of the facts that
\[ |\Theta(\varepsilon)| \leq |V| \] and
\[ U(t) = \sum_{i=1}^{n} e^{\lambda_i (T-t)} \langle U_T, \phi_i \rangle \phi_i, \] we have
\[ U(t) - P^\varepsilon U(t) = \sum_{i \notin \Theta(\varepsilon)} e^{\lambda_i (T-t)} \langle U_T, \phi_i \rangle \phi_i. \]

It follows from (3.6) that
\[ \frac{d}{dt} U(t) = \sum_{i=1}^{n} -\lambda_i e^{\lambda_i (T-t)} \langle U_T, \phi_i \rangle \phi_i. \]
Combining this with (3.7), we estimate that
\[ \| U(t) - P^\varepsilon U(t) \|^2 = \sum_{i \notin \Theta(\varepsilon)} \lambda_i^{-2} \lambda_i^2 e^{2\lambda_i (T-t)} |\langle U_T, \phi_i \rangle|^2. \]
\[ \leq M_\varepsilon^{-2} \left\| \frac{d}{dt} U(t) \right\|^2. \]

It is equivalent to
\[ (3.8) \]
\[ \| U(t) - P^\varepsilon U(t) \| \leq M_\varepsilon^{-1} \left\| \frac{d}{dt} U(t) \right\|. \]

Henceforth, by combining (3.8), (3.3), and using the triangle inequality, we obtain
\[ \| U(t) - P^\varepsilon U^\varepsilon(t) \| \leq \| U(t) - P^\varepsilon U(t) \| + \| P^\varepsilon U(t) - P^\varepsilon U^\varepsilon(t) \| \]
\[ \leq M_\varepsilon^{-1} \left\| \frac{d}{dt} U(t) \right\| + e^{M_\varepsilon (T-t) \varepsilon}. \]

This completes the proof of (3.4). Proof of (3.5) also follows. \( \Box \)

**Remark 3.2.** In the proof of Theorem 4, we obtain a logarithmic rate of convergence under the assumption that \( |\Theta(\varepsilon)| \leq |V| \). When \( \varepsilon \) is so small such that \( |\Theta(\varepsilon)| > |V| \), as mentioned above, we set \( M_\varepsilon = \lambda_n \). This means that \( U(t) \) coincides with \( P^\varepsilon U(t) \), i.e. \( U(t) - P^\varepsilon U(t) = 0 \). In this scenario, we obtain a Lipschitz rate of convergence. Last but not least, taking \( t = 0 \) in (3.5) gives us that
\[ \| U_0 - P^\varepsilon U_0^\varepsilon \| \leq \frac{T}{\ln (\varepsilon^{-\gamma})} \left\| \frac{d}{dt} U_0 \right\| + \varepsilon^{1-\gamma} \]
in which we assume the existence of \( \frac{d}{dt} U_0 \in \mathbb{R}^n \).
4. Numerical examples

In this section, we provide a series of numerical examples to demonstrate the proposed reconstruction algorithm in denoising gray images corrupted by the commonly used additive white Gaussian noise (AWGN). The algorithm is implemented using MATLAB 2016 and executed using a computer equipped with Intel® Core™ i5-10210U @ 1.60GHz and 8 GB of RAM. In our numerical experiments, we use Set12—a widely-used test dataset for benchmarking denoising algorithms; cf. [11]. The dataset consists of 12 gray images of sizes 256 × 256 or 512 × 512; however, we resize the images to the size of 128 × 128 due to the limitation of computer hardware in creating larger graphs. We apply the AWGN with variance $\sigma_{\text{noise}} = 20$ to this resized dataset and thereby, obtain noisy images.

The blurred images $U^\varepsilon_T$ are obtained via the forward process in (1.1). For simplicity, we apply in this process the standard forward Euler method with the Courant number of 0.03 for the conditional stability and with $T = 0.5$ that corresponds to the deviation $\sigma = 1$ in the Gaussian blurring process. Since we work with 128 × 128 images, the number of vertices is $n = 16384$ in our numerical results. Moreover, we fix $\varepsilon = 0.1$ and $\gamma = 0.5$ that gives $M_\varepsilon \approx 2.30$. For objective assessment, we use peak signal-to-noise ratio (PSNR) as an image quality metric in evaluating denoising performance, which is represented in the logarithmic decibel (dB) scale. In principle, it is computed as

$$\text{PSNR}(U_0, PU_0^\varepsilon) = 10 \log_{10} \left( \frac{255^2}{\frac{1}{n} \sum_{i=1}^{n} (u_{0i} - P_u_{0i})^2} \right).$$

To estimate the PSNR of our images, we compare those to the original ones that are resized from the Set12 dataset. This comparison is operated by the built-in MATLAB function \texttt{psnr}.

Figures 4 and 5 show the denoising results of the cut-off solution $PU_0^\varepsilon$ in (3.1) for all images in Set12. The denoised results exhibit better visual quality than that of noisy images. Especially in regions with rich textures, the noise is reduced while edge features are preserved without significant degradation. Table 1 shows the subjective evaluations for denoising results of all images in Set12 from approximation solution (2.6) and cut-off solution $PU_0^\varepsilon$ in (3.1) in terms of PSNR. As can be seen from the table, the cut-off solution $PU_0^\varepsilon$ outperforms the approximation solution (2.6), as the PSNR scores of cut-off solutions $PU_0^\varepsilon$ are higher than those of approximation solutions from 1.66dB (for Monarch image) to 2.64dB (for Couple image). We detail that 1.66dB for the Monarch image is calculated directly from the PSNR difference between the cut-off (18.86) and approximation (17.20) solutions; see Table 1. The average of PSNR increase is 2.17dB, i.e. 20.31dB of cut-off solutions $PU_0^\varepsilon$ compared with 18.14dB of approximation solutions (2.6), which is a significant improvement.

Last but not least, we provide Table 2 to show the time efficiency of the cut-off projection under study. As can readily be expected from the fact that $|\Theta(\varepsilon)| < |V|$, in average, the running time for the cut-off solution $PU_0^\varepsilon$ is 4.5 times less than that of the approximation (2.6). Therefore, it is potential to adapt the cut-off projection to more complex scenarios (e.g. color images with three-dimensional matrices) in upcoming research projects.

Acknowledgment. Vo Anh Khoa would like to thank Professor Paul Sacks (Iowa State University, USA) for great support of his research career.
Figure 4. Visual comparison of denoising results on standard test images in the Set12 dataset. Left column: original images; middle column: noisy images; right column: denoised images.
Figure 5. Visual comparison of denoising results on standard test images in the *Set12* dataset. Left column: original images; middle column: noisy images; right column: denoised images.
Table 1. Quantitative comparison in terms of PSNR for denoising performance between approximate solutions (2.6) and cut-off solutions (3.1).

| Image name | Approximation (2.6) | Cut-off (3.1) |
|------------|---------------------|---------------|
| Camera man | 17.89               | 20.15         |
| House      | 19.40               | 22.00         |
| Peppers    | 18.10               | 20.06         |
| Starfish   | 18.05               | 20.08         |
| Monarch    | 17.20               | 18.86         |
| Airplane   | 17.76               | 19.83         |
| Parrot     | 16.66               | 18.39         |
| Lena       | 18.42               | 20.66         |
| Barbara    | 18.34               | 20.37         |
| Boat       | 18.50               | 20.94         |
| Man        | 18.75               | 21.09         |
| Couple     | 18.64               | 21.28         |
| Avg.       | 18.14               | 20.31         |

Table 2. Comparison of running time in seconds between approximate solutions (2.6) and cut-off solutions (3.1).

| Image name | Approximation (2.6) | Cut-off (3.1) |
|------------|---------------------|---------------|
| Camera man | 1.23                | 0.28          |
| House      | 1.15                | 0.27          |
| Peppers    | 1.15                | 0.24          |
| Starfish   | 1.13                | 0.26          |
| Monarch    | 1.14                | 0.25          |
| Airplane   | 1.14                | 0.25          |
| Parrot     | 1.13                | 0.26          |
| Lena       | 1.13                | 0.26          |
| Barbara    | 1.16                | 0.27          |
| Boat       | 1.16                | 0.26          |
| Man        | 1.26                | 0.28          |
| Couple     | 1.20                | 0.27          |
| Avg.       | 1.17                | 0.26          |

References

[1] A. E. Brouwer and W. H. Haemers. *Spectra of Graphs*. Springer New York, December 2011.
[2] M. Denche and K. Bessila. A modified quasi-boundary value method for ill-posed problems. *Journal of Mathematical Analysis and Applications*, 301(2):419–426, 2005.
[3] L. Fan, F. Zhang, H. Fan, and C. Zhang. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1), 2019.
[4] A. Friedman. *Partial Differential Equations of Parabolic Type*. Dover Pubn Inc, April 2008.
[5] G. Gilboa, N. Sochen, and Y. Y. Zeevi. Forward-and-backward diffusion processes for adaptive image enhancement and denoising. *IEEE Transactions on Image Processing*, 11(7):689–703, 2002.
[6] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese. Nonlocal Mumford-Shah regularizers for color image restoration. *IEEE Transactions on Image Processing*, 20(6):1583–1598, 2011.
[7] M. V. Klibanov and A. G. Yagola. Convergent numerical methods for parabolic equations with reversed time via a new Carleman estimate. *Inverse Problems*, 35(11):115012, 2019.

[8] R. Lattès and J. L. Lions. *Méthode de Quasi-réversibilité et Applications*. Dunod, Paris, 1967.

[9] P. T. Nam, D. D. Trong, and N. H. Tuan. The truncation method for a two-dimensional nonhomogeneous backward heat problem. *Applied Mathematics and Computation*, 216:3423–3432, 2010.

[10] H. T. Nguyen, V. A. Khoa, and V. A. Vo. Analysis of a quasi-reversibility method for a terminal value quasi-linear parabolic problem with measurements. *SIAM Journal on Mathematical Analysis*, 51(1):60–85, 2019.

[11] S. Roth and M. J. Black. Fields of experts: a framework for learning image priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 860–867, 2005.

[12] K. Shi. Coupling local and nonlocal diffusion equations for image denoising. *Nonlinear Analysis: Real World Applications*, 62:103362, 2021.

[13] K. Shi. Image denoising by nonlinear nonlocal diffusion equations. *Journal of Computational and Applied Mathematics*, 395:113605, 2021.

[14] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin. Deep learning on image denoising: An overview. *Neural Networks*, 131:251–275, 2020.

[15] N. H. Tuan, V. A. Khoa, M. T. N. Truong, T. T. Hung, and M. N. Minh. Application of the cut-off projection to solve a backward heat conduction problem in a two-slab composite system. *Inverse Problems in Science and Engineering*, 27(4):460–483, 2018.

[16] M. Welk, G. Gilboa, and J. Weickert. Theoretical Foundations for Discrete Forward-and-Backward Diffusion Filtering. pages 527–538. Springer Berlin Heidelberg, 2009.

**Department of Mathematics, Florida A&M University, Tallahassee, FL 32307, USA**  
*E-mail address*: anhkhoa.vo@famu.edu

**Department of Multimedia Engineering, Dongguk University, South Korea**  
*E-mail address*: mtntruong@toantin.org

**Department of Mathematics, Florida A&M University, Tallahassee, FL 32307, USA**  
*E-mail address*: imhotep1.hogan@famu.edu

**Department of Mathematics, Florida A&M University, Tallahassee, FL 32307, USA**  
*E-mail address*: roselyn.williams@famu.edu