Fast deconvolution using a combination of Richardson-Lucy iterations and diffusion regularization

Thibaut Modrzyk, Ane Etxebeste, Elie Bretin, Voichita Maxim

Fast deconvolution using a combination of Richardson-Lucy iterations and diffusion regularization

Thibaut Modrzyk[1], Ane Etxebeste[1], Elie Bretin[2], Voichita Maxim[1]
[1] INSA-Lyon, CREATIS UMR 5220 U1294, Villeurbanne, F-69621, France
[2] INSA-Lyon, CNRS UMR 5208, Institut Camille Jordan, Villeurbanne, F-69621, France.

***Abstract*—This paper presents a novel hybrid approach to address the non-blind deconvolution inverse problem, from the viewpoint of a future application to medical imaging. Motivated by the efficiency of diffusion models applied to inverse problems, we propose to integrate Diffusion Posterior Sampling (DPS) with variational methods through Richardson-Lucy iterations. The benefit of our method is to reduce by a factor of 10 the number of sampling steps currently required by DPS. Several numerical experiments demonstrate that our framework achieves comparable performance to existing deconvolution strategies, while seemingly reducing the computing time and the hallucination effect, both particularly important for medical applications.**

***Index Terms*—Inverse Problems, Non-blind Deconvolution, Diffusion models, Richardson-Lucy**

I. Introduction

In many situations it is necessary to recover an image x from its blurry and noisy version y. A relatively easy situation is when the blur is produced by convolution with a known filter k, and the problem is modeled as

$$y = k * x + n \tag{1}$$

with $*$ the convolution operation and n some noise, often considered as independent and identically distributed (*i.i.d.*) following a zero-mean Gaussian distribution, although non-additive Poisson noise is also frequent. Recovering x from y is a particular case of inverse problem, called deconvolution. The kernel k may be a smooth function, as the point spread function (PSF) of some imaging device, often modelled as a Gaussian function, or some sparse motion filter. For smooth filters, it is easy to see by considering the Fourier transform that the high frequencies of the signal x are strongly attenuated. A Gaussian noise however will equally affect all frequencies. To prevent noise from high frequencies to be enhanced through deconvolution and to hide the signal, prior information has to be included in the process. The deconvolution problem may be written in a variational form as finding $\widehat{x}$ solution of the minimization problem

$$\min_x d(k * x, y) + \lambda R(x) \tag{2}$$

where $d(k * x, y)$ is the data fidelity term, often considered to be the square of the ℓ_2 norm or the Kullback-Leibler distance, $R(x)$ is a penalty expressing prior on data, and λ is a constant balancing the two terms.

Both Wiener filtering [1] and the Richardson-Lucy (RL) algorithm [2], aim to recover the low frequencies of the signal and as much high frequencies as possible, while still controlling the signal to noise ratio (SNR). For the latter, this is possible either by early stopping or by adding some regularization term. Total variation (TV) regularization coupled with RL iterations allows to enhance edges and thus contributes to recovery of high frequencies [3]. However, small details are still removed and texture is modified. Other works use dictionaries and redundancies in images [4].

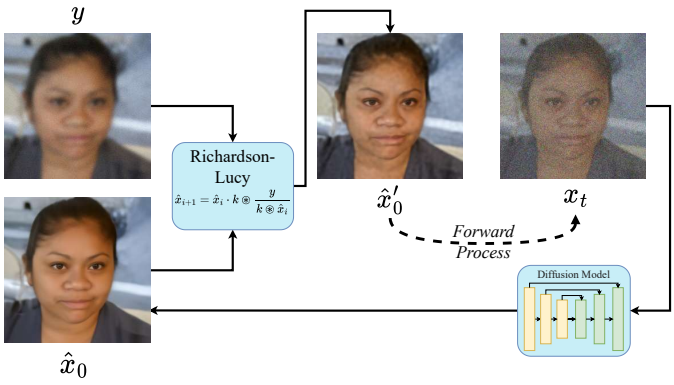

Fig. 1. Diagram illustrating our proposed Richardson-Lucy diffusion algorithm. Firstly, we obtain an estimate of the deconvolved image using several RL iterations. Next, we add noise to the image according to the forward diffusion process and then apply the diffusion model as a score-matching regularizer. This process can be repeated to further deconvolve the image.

Recently the success of deep learning in various image processing tasks has inspired efforts to develop new methods for solving inverse problems. Deep learning methods allow to learn either the entire deconvolution process in an end-to-end network [5] [6], or to combine the convolution model from Equation (1) with some learned processing/regularization [7] [8]. The former consist in learning a filter that could be successfully applied to any image of interest. This means that in principle a training should be done for each convolution filter k that may occur. The latter consists to exploit the learned prior distribution $p(x)$ as a regularizer for the inverse problem. In both cases a large data base is required for training, in order to accurately sample $p(x)$. The learning process is however

This work was partially funded by the PRIMES LABEX (ANR-11-LABX-0063) at the University of Lyon, within the "Investissements d'Avenir" program (ANR-11-IDEX-0007).

independent on the convolution kernel and potentially more robust to distribution shifts, *i.e.*, out-of-distribution samples.

In this work we focus on priors expressed as trained diffusion models [9]. Diffusion models can sample from the prior images with impressive detail. Data fidelity optimization aims to keep the solution in agreement with the data and prevent network hallucinations. This idea was previously explored for instance in [10], where sampling from the posterior $p(x|y)$ is done by taking a gradient descent step on the data fidelity at each diffusion time step. However, this algorithm can be slow, as diffusion models typically require training with 1000 time steps. Additionally, the low effectiveness of the gradient descent in deconvolution problems makes that the diffusion prior takes over the data-fidelity constraint, resulting in hallucinated details. These are not rare, as diffusion models produce samples from a distribution and these samples may be outliers.

We propose a formulation based on the Richardson-Lucy algorithm, known to be faster for operating deconvolution, and diffusion models, leading to an algorithm in the flavour of plug-and-play methods.

II. Related works

A. Classical approaches

Traditionally, computing the ground-truth image x from the measurement y is done by minimizing an objective function comprising a data-fidelity term and regularization term as expressed in equation (2). Standard resolution methods include Wiener filtering and iterative methods such as the RL algorithm which can be written, for a symmetric kernel, as:

$$x^{(i+1)} = x^{(i)} \cdot k * \frac{y}{k * x^{(i)}}, \tag{3}$$

with $x^{(0)}$ set to strictly positive values. Here $\cdot$ is the matrix Hadamard product which represents an element-wise multiplication. The division is also performed element-wise.

The RL algorithm is slower compared to Wiener filtering due to its iterative nature but usually performs better. It is still sensitive to noise and can produce visible artefacts. A solution to this issue is early stopping of iterations, at the cost of an insufficient recovery of high frequencies. An other solution is to incorporate regularization into the reconstruction process. Although there are numerous regularization techniques employed in image processing inverse problems, they frequently result in the smoothed edges. TV regularization, which was initially introduced by Rudin *et al.* [11] for optical images, has the advantage of preserving sharp edges and has been widely used for various inverse problems, including deconvolution [3].

B. Diffusion models

Diffusion models are a class of generative models which aim to approximate the *reverse* of the noising process known as the *diffusion process*. Diffusion models are also called *score-matching models* and were introduced in [9] and [12]. The diffusion process, or *forward process*, is a noising process that increasingly adds noise to the input data x following the variance schedule β_t. It can formally be described as a stochastic differential equation:

$$dx = -\frac{1}{2}\beta_t x dt + \sqrt{\beta_t} dw \tag{4}$$

where $t \in \{0, \dots, T\}$ with T often chosen to be 1000, and dw is the standard Wiener process. This process is designed such that the distribution x_T converges to a normal Gaussian distribution, i.e. $x_T \sim \mathcal{N}(0, 1)$. The reverse process is usually defined as:

$$dx = \left[-\frac{1}{2}\beta_t x - \beta_t \nabla_{x_t} \log p(x_t)\right] dt + \sqrt{\beta_t} d\bar{w} \tag{5}$$

where $\nabla_{x_t} \log p(x_t)$ is the score function, and $d\bar{w}$ is the reverse standard Wiener process. The score function is estimated by denoising score matching:

$$\arg\min_{\theta} \mathbb{E}_{t,x_t,x_0}[||s_\theta(x_t, t) - \nabla_{x_t} \log p(x_t|x_0)||_2^2] \tag{6}$$

The trained model s_{θ^*} is then used as a surrogate for the score function.

C. Diffusion models for inverse problems

While diffusion models gained popularity with their success in generative tasks, they have also been adapted to solve inverse problems by sampling from the posterior $p(x|y)$. Using Bayes' formula, the corresponding score of the posterior can be written as:

$$\nabla_{x_t} \log p(x_t|y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y|x_t), \tag{7}$$

where the prior term can easily be obtained using the trained model s_{θ^*}, while the likelihood term is intractable in general. To be able to sample from the posterior, a previous work by Chung *et al.* [10] uses the approximation $p(y|x_t) \approx p(y|\hat{x}_0)$ where $\hat{x}_0$ is the denoised estimate of the image at step t. This estimate can be computed using Tweedie's formula as follows:

$$\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1 - \bar{\alpha}_t) s_{\theta^*}(x_t, t)) \tag{8}$$

More precisely they also choose this expression for the gradient of the log-likelihood:

$$\nabla_{x_t} \log p(y|x_t) = -\nabla_{x_t} ||y - k * \hat{x}_0||_2^2 \tag{9}$$

Finally, the reverse process reads as:

$$dx = -\frac{1}{2}\beta_t x dt - \beta_t \nabla_{x_t} \log\, p(x_t|y) + \sqrt{\beta_t} d\bar{w}. \tag{10}$$

Referred as Diffusion Posterior Sampling (DPS), the algorithm by Chung *et al.* is detailed in Algorithm 1 under the "DPS refinement" section.

Song *et al.* [13] also developed a method inspired both by diffusion models and variational approaches. They propose a loss where the diffusion model is a regularizer as in the RED framework, with a weighting mechanism that balance the effects of regularization and data-fidelity.

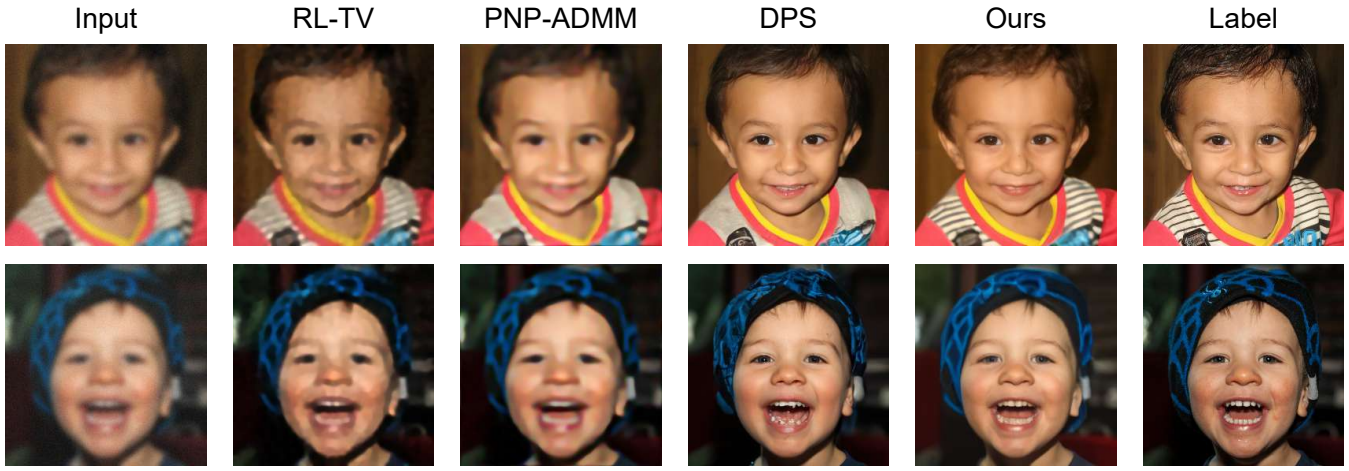

Fig. 2. Examples of deconvolutions using several deconvolution methods, including diffusion based methods (DPS). Note the improved consistency of our method compared to DPS, for instance on the shirt in the first row or on the hat on the second row.

D. Limitations of diffusion approaches

While these methods give impressive results on several inverse problems, such as impainting, super-resolution or Gaussian deconvolution, the sampling time could be a limitation for many applications. In fact, sampling through 1000 steps using DPS takes about 90 seconds to reconstruct a single image on our system (see section IV-A for details).

Although posterior sampling can produce high-resolution reconstructions on datasets such as FFHQ or ImageNet, it often generates details that are not present in the original image. Those details are coherent with the produced image, making them difficult to detect and are commonly referred to as 'hallucinations'. In the field of medical imaging, hallucinations are considered as a major drawback of deep learning methods. Iterative algorithms are considered safer because they utilize the physics of the inverse problem, and any artifacts they produce are easier to identify.

III. Proposed method

The main inspiration for our method is plug-and-play, where the iterative algorithm and the deep learning regularizer are decoupled. We combine the RL algorithm, which is well-adapted for data-fidelity minimization, with a diffusion model that provides a strong information on the prior distribution.

A natural mathematical framework is that of the Expectation Maximization (EM) algorithm [14], which allows to numerically compute a solution x'_0 of the maximum a posteriori (MAP) problem:

$$\arg\max_x \log p(x|y) = \arg\max_x \left(\log p(y|x) + \log p(x)\right), \quad (11)$$

where $p(x)$ is the prior distribution learned by the diffusion model and $p(y|x)$ is the likelihood of the observations y. Note that, contrary to the likelihood $p(y|x_t)$ from equation (7), this likelihood is tractable and related to the forward model.

This solution x'_0 is computed iteratively. Each iteration is splitted in two steps: an expectation (E) step which is in practice a step from the RL algorithm, and a maximization (M) step. The objective of the second one is to find $\hat{x}_0$ that is more likely than x'_0 in the sense of the prior $p(x)$. This can be obtained by solving:

$$\arg\max_x \left(\log p(x'_0|x) + \log p(x)\right) = \arg\max_x \; p(x|x'_0). \quad (12)$$

Instead of computing the maximum in Equation (12), we perform one sampling step of the diffusion model. We also run several RL iterations before each diffusion sampling. This strategy leads us to the first half of Algorithm 1.

Algorithm 1 Our algorithm

Require: $k, y, N, M, L, \{\zeta_i\}_{t=1}^N, \{\tilde{\sigma}_i\}_{i=1}^N$

$\hat{x}_0 = y$
for $t = M + N - 1$ to M **do** ▷ Our contribution
  for $j = 1$ to L **do**
    $x'_0 = \hat{x}_0 \cdot k * \frac{y}{k * \hat{x}_0}$
  end for
  $z \sim \mathcal{N}(0, I)$
  $x_t \leftarrow \sqrt{\bar{\alpha}_t} x'_0 + \sqrt{1 - \bar{\alpha}_t} z$
  $\hat{x}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}} \left(x_i + (1 - \bar{\alpha}_i) s_{\theta^*}(x_t, t)\right)$
end for
for $t = M - 1$ to 0 **do** ▷ DPS Refinement
  $\hat{x}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}} (x_t + (1 - \bar{\alpha}_t) s_{\theta^*}(x_t, t))$
  $z \sim \mathcal{N}(0, I)$
  $x'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}} \beta_t}{1 - \bar{\alpha}_t} \hat{x}_0 + \tilde{\sigma}_t z$
  $x_{t-1} \leftarrow x'_{t-1} - \zeta_t \nabla_{x_t} \|y - k * \hat{x}_0\|_2^2$
end for
return $\mathbf{x}_0$

We start the algorithm from the observation y rather than Gaussian noise. This idea comes from the simple observation that most of the sampling process is spent forming basic shapes and semantic objects, while only the last steps are responsible for the formation of high frequency details, as it was already observed in [15]. We thus begin the diffusion

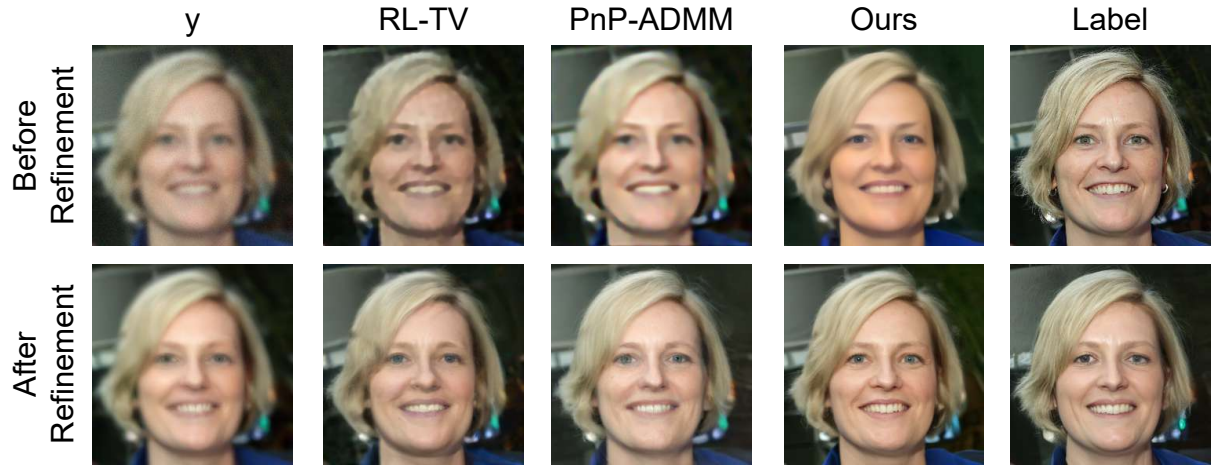

Fig. 3. Ablation study: comparison of reconstructions using different deconvolution algorithms before and after the refinement step using DPS. On the bottom right is the reconstruction using DPS for 1000 steps for reference.

sampling at timestep $t = 100$ instead of $T = 1000$, enabling us to bypass most of the sampling process. The noise schedule of the diffusion model is adjusted to maintain a balance between preserving the information in intermediate reconstructions and masking the RL artefacts with noise, allowing the prior of the diffusion model to eliminate them. We use the parameter $\beta_T = 0.002$ instead of the usual value $\beta_T = 0.02$.

Although our method enforces strengthened data-consistency compared with DPS, we have observed that it struggles to reconstruct high-frequency details. For this reason we refine our solution with a small number of DPS steps in order to recover high frequencies. It may seem that applying the refining step to the observation y or another reconstruction would yield similar results. We have verified that this is not the case by conducting an ablation study.

IV. Experiments

A. Experimental setup

Dataset We tested our algorithm's capabilities on the FFHQ 256×256 dataset [16] using the corresponding pre-trained diffusion model from [10]. The model was trained by Chung *et al.* on 49k images of FFHQ, leaving 1k images for validation.

Blurry images were produced by convolution with a kernel of size 61×61 pixels and a standard deviation of 3.0 pixels. Gaussian noise with a $\sigma = 0.05$ was then added to the blurry images. The convolution was applied using a padding of 30 pixels replicating the values on the border of the image.

Metrics We report the average peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) to measure the fidelity of the reconstruction to the original image. We also measure the perceptual quality of the output images using the Learned Perceptual Image Patch Similarity (LPIPS) and the Fréchet Inception Distance (FID). This ensures the quality of our reconstruction by accounting for aspects of the reconstruction that may be overlooked by pixel-wise metrics. For instance blurry reconstruction are known to have an advantage in terms of PSNR and SSIM while being perceptually worse. All experiments were run on an RTX 6000 Ada GPU with 48GB RAM.

B. Methods for comparison

We compared our method with three other approaches:

Diffusion Posterior Sampling [10] For DPS we adopted the same configuration as Chung *et al.* with $\beta_1 = 0.0001$ and $\beta_T = 0.02$ with $T = 1000$.

Richardson-Lucy with TV regularization [17] For RL-TV we tried to optimize the number of steps and the regularization parameter λ_{TV} on a few images that were not part of the validation set, based on the PSNR of the reconstruction. We found that 25 steps with a regularization parameter $\lambda_{TV} = 0.005$ provided the best results.

Plug-and-play alternating direction method of multipliers [18] We used the implementation provided in the scico library for Python (Balke et al., 2022). We employed the pre-trained DnCNN [19] model to replace the proximal operator. We used the default parameters provided in the library, which are $\rho = 0.2$ with ρ the ADMM penalty parameter, and the maximum number of iterations was set to 12.

To implement our algorithm, we applied $N = 20$ steps of the first reconstruction stage with $L = 20$ RL steps at each iteration and $M = 80$ refinement steps using DPS. These parameters were chosen empirically to achieve satisfactory performance, but we did not perform a thorough grid search to determine the best hyper-parameters.

C. Results

The objective of the first experiment is to compare our method with other conventional reconstruction methods. Regarding DPS, it was found that reconstructed images contained high-frequency details that may appear impressive at first glance, but a significant part of these details were actually hallucinated and not present in the original image. Although the reconstruction remains believable, its faithfulness to the label is limited as can be seen in Figure 2. This can be

observed in Table I through distortion metrics, where DPS performs worse than the other tested methods. It should also be noted that variational methods often yield results that are unsatisfactory from a perceptual standpoint, despite achieving good distortion metrics.

Again the benefit of our method is that it runs significantly faster than DPS to perform the deconvolution. Indeed, it takes around 90 s for DPS to reconstruct the image while for our method it takes only 30 s. The main reason for this improvement is that we perform only 100 Neural Function evaluations compared to the 1000 evaluations required by DPS.

TABLE I
COMPARISON OF SEVERAL DECONVOLUTION METHODS ON THE TEST DATASET

Method	PSNR (dB) ↑	SSIM ↑	LPIPS ↓	FID ↓
RL-TV	23.15	0.72	0.21	95.90
PNP-ADMM	**25.50**	**0.73**	0.21	79.80
DPS	24.33	0.68	**0.14**	**32.42**
Ours	<u>24.40</u>	**0.73**	**0.14**	<u>44.23</u>

Ablation study The aim of our second experiment is to determine whether the refinement step is responsible for our method's good performances. As the refinement step can be applied to any intermediate reconstruction, we replaced our method by (i) the measurement y, (ii) the result of the TV-regularized RL algorithm, (iii) the result of the PnP-ADMM deconvolution. We performed 80 refinement steps with DPS on the output of each method.

Under visual inspection, our method offers the best compromise between pixel-wise precision and perceptual qualities. It can be seen in Figure 3 that applied to other approximates, for instance the observation y, the refinement results in blurry reconstructions. In table II, it can be seen that our method strikes the best balance between all considered metrics. It is worth noting that RL-TV, close to one multiple (E) iteration of our method, provides comparable performances.

TABLE II
ABLATION OF THE MAIN ALGORITHM WITH THE SAME REFINEMENT TECHNIQUE

Input	PSNR (dB) ↑	SSIM ↑	LPIPS ↓	FID ↓
y + DPS	24.11	0.69	0.22	64.74
PNP-ADMM + DPS	23.09	0.66	<u>0.17</u>	**40.85**
RL-TV + DPS	**24.97**	<u>0.72</u>	<u>0.17</u>	53.89
Ours + DPS	<u>24.40</u>	**0.73**	**0.14**	<u>44.23</u>

V. CONCLUSION AND DISCUSSIONS

We propose a method for fast non-blind deconvolution using Richardson-Lucy iterations combined with diffusion regularization. The method strikes a balance between diffusion prior and consistent variational methods. The method involves alternating RL with a diffusion step in the first phase, followed by a refinement method using pure DPS with limited iterations in the second phase. Numerical experiments showed that our method reduces the number of steps required by the sampling process while remaining more efficient than variational methods. Our method produces results comparable to the state of the art and it could be a first step towards utilizing diffusion models in a plug-and-play fashion. Although there is currently no mathematical formalism for our algorithm, we believe it is possible to justify its consistency with the EM framework [14].

Although our method enforces strengthened data-consistency compared with DPS, we have observed that it struggles to reconstruct high-frequency details. We are uncertain about the cause of this behaviour, but an efficient mechanism to balance between diffusion prior and deconvolution could be necessary to improve the performance of our method. Instead, a more straightforward approach is used, which involves refining the reconstruction by applying DPS for a small number of steps. It may seem that applying the refining step to the observation y or another reconstruction would yield similar results, but our experiments demonstrate otherwise. We have verified that our method is crucial to the quality of the reconstructions by conducting an ablation study, as shown in Table II.

REFERENCES

[1] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press, 1964.

[2] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration*," *JOSA*, vol. 62, no. 1, pp. 55–59, Jan. 1972.

[3] N. Dey *et al.*, "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," *Microscopy Research and Technique*, vol. 69, no. 4, pp. 260–266, Apr. 2006.

[4] F. Soulez, "A "learn 2D, apply 3D" method for 3D deconvolution microscopy," in *11th IEEE ISBI*, Apr. 2014, pp. 1075–1078.

[5] L. Xu *et al.*, "Deep Convolutional Neural Network for Image Deconvolution," in *NeurIPS*, vol. 27. Curran Associates, Inc., 2014.

[6] J. D. Rego *et al.*, "Robust Lensless Image Reconstruction via PSF Estimation," in *IEEE WACV*, Jan. 2021, pp. 403–412.

[7] Z. Wang *et al.*, "Image Deconvolution with Deep Image and Kernel Priors," in *IEEE/CVF ICCVW*, Oct. 2019, pp. 980–989.

[8] J. Dong *et al.*, "Deep wiener deconvolution: Wiener meets deep learning for image deblurring," *NeurIPS*, vol. 33, pp. 1048–1059, 2020.

[9] Y. Song *et al.*, "Score-based generative modeling through stochastic differential equations," in *ICLR*, 2021.

[10] H. Chung *et al.*, "Diffusion posterior sampling for general noisy inverse problems," in *ICLR*, 2023.

[11] L. I. Rudin *et al.*, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, Nov. 1992.

[12] J. Ho *et al.*, "Denoising diffusion probabilistic models," in *NeurIPS*, ser. NIPS'20. Red Hook, NY, USA: Curran Associates Inc., 2020.

[13] J. Song *et al.*, "Pseudoinverse-guided diffusion models for inverse problems," in *ICLR*, 2023.

[14] A. P. Dempster *et al.*, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[15] J. Yu *et al.*, "Freedom: Training-free energy-guided conditional diffusion model," *IEEE/CVF ICCV*, 2023.

[16] T. Karras *et al.*, "A style-based generator architecture for generative adversarial networks," *IEEE Transactions on Pattern Analysis; Machine Intelligence*, vol. 43, no. 12, pp. 4217–4228, dec 2021.

[17] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, May 2011.

[18] O. Elgendy, "Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications," *IEEE Transactions on Computational Imaging*, Jan. 2017.

[19] K. Zhang *et al.*, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.