

Optimised Anisotropic Poisson Denoising

Georg Radow¹(✉), Michael Breuß¹, Laurent Hoeltgen¹, and Thomas Fischer²

¹ Chair for Applied Mathematics, Brandenburg Technical University,
Platz der Deutschen Einheit 1, 03046 Cottbus, Germany
`{radow,breuss,hoeltgen}@b-tu.de`

² Central Analytical Laboratory, Brandenburg Technical University,
Konrad-Wachsmann-Allee 6, 03046 Cottbus, Germany
`thomas.fischer@b-tu.de`

Abstract. The aim of this paper is to deal with Poisson noise in images arising in electron microscopy. We consider here especially images featuring sharp edges and many relatively large smooth regions together with smaller strongly anisotropic structures. To deal with the denoising task, we propose a variational method combining a data fidelity term that takes into account the Poisson noise model with an anisotropic regulariser in the spirit of anisotropic diffusion. In order to explore the flexibility of the variational approach also an extension using an additional total variation regulariser is studied. The arising optimisation problems can be tackled by efficient recent algorithms. Our experimental results confirm the high quality obtained by our approach.

Keywords: Poisson denoising · Variational methods · Electron microscopy

1 Introduction

The aim of this paper is to deal with noise in certain types of images arising in electron microscopy. A typical example for some of our applications of interest is depicted in Fig. 1. The displayed test image shows gold sputtered Polytetrafluoroethylene (PTFE) particles that were recorded using the secondary electron detector of a ZEISS DSM 962 scanning electron microscope.

Let us briefly discuss the properties of images we deal with here as exemplified via Fig. 1. Beginning with image content, we observe that there are many relatively large smooth (and round) regions with strong edges at particle outlines and particle overlaps. However, there are also relatively thin, elongated structures. Turning to the image disturbances, the image is characterised by random electron discharges appearing as bright horizontally striped artefacts. To circumvent sample destruction by electron impact at the image acquisition process, we had to choose a high scanning speed, which further resulted in a grainy texture.

Our Modelling Approach. Variational methods have been very successful for image denoising and restoration tasks [4]. Their main conceptual advantages are

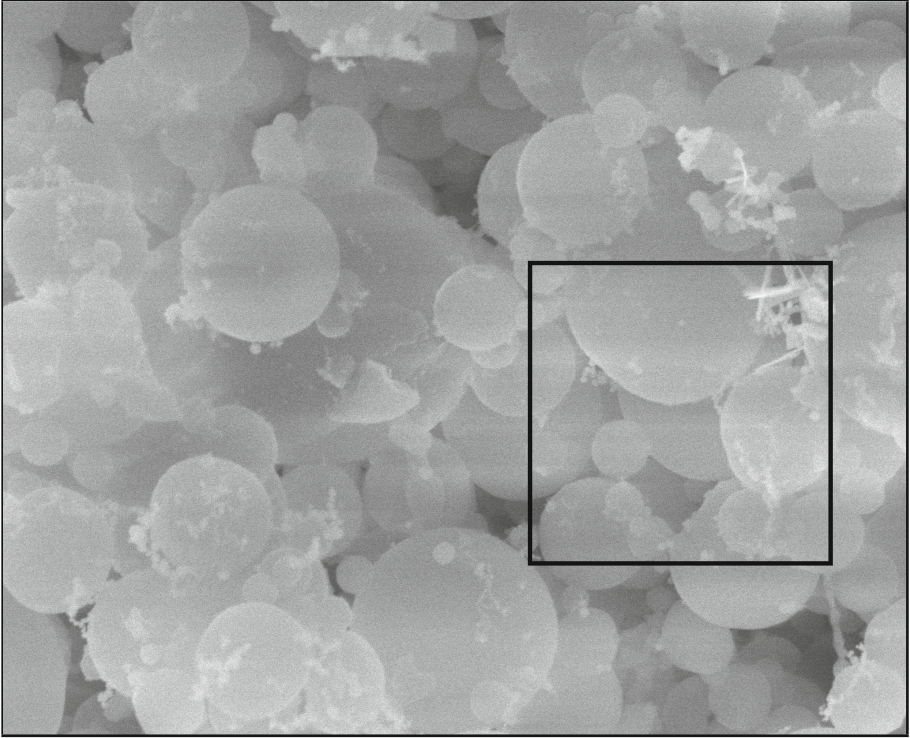


Fig. 1. Noisy image from an electron microscope as an example for a typical application in our lab, image size is $M \times N = 2048 \times 1664$. For image acquisition, we used an accelerating voltage of 20 kV, a working distance of 6 mm and a magnification of 5000 with a ZEISS DSM 962 scanning electron microscope. The marked image region is later on employed for demonstrating filtering results.

their flexibility with respect to the possible model components and the inherent robustness due to the use of suitable regularisers. Moreover, the components of a variational method can often be easily interpreted and fine tuned. As the images in our applications may show various types of structures of potential interest, we opt to employ a variational approach because of its flexibility. As a further benefit we may devise a model for enhancing specific image structures.

Let us now turn to the variational model. The task is to consider the terms in a corresponding energy functional $E(u)$ that shall be minimised for finding the optimal solution u^* . The usual design incorporates terms for data fidelity and regularisation, respectively, so that on a coarse level of abstraction the variational model reads as:

$$E(u) := R(u) + \tau S(u; u^p). \quad (1)$$

Thereby, u^p denotes the given degraded input image, and $\tau > 0$ is a parameter defining the balance between data term $S(u; u^p)$ and regulariser $R(u)$.

Let us first consider the data term $S(u; u^p)$. We may safely assume that an adequate model for noise in electron microscopy images u^p is Poisson noise [18]. For the regulariser $R(u)$ we need to take into account the types of prominent image structures. We choose here to employ an anisotropic regulariser that is able to preserve the strong edges and the thin elongated structures that are prominent in our input images. Furthermore, exploring the flexibility of the variational approach, we also add a total variation (TV) regulariser in order to enhance edges and achieve visually more plateau-like structures corresponding to individual particles in our application.

Related Work. As the subject of this paper touches various aspects of image processing, we can only attempt to cover here some of the most important works that influenced the considerations for the construction of our approach.

Several methods and approaches for Poisson denoising (PD) have been proposed in previous literature. Often a variance stabilizing transformation (VST) is applied to transform Poisson noisy data into data with Gaussian noise. One such transformation was proposed by Anscombe in [1]. Methods to optimise the denoising with this transformation were presented by Mäkitalo and Foi in [13] and by Azzari and Foi in [2]. An extension of the Anscombe transformation and its application to PD was proposed by Zhang *et al.* [22]. Also, some methods from the field of machine learning have been considered, see *e.g.* Giryes and Elad [9] and Salmon *et al.* [17]. Poisson denoising methods based on multiscale models were studied by Lefkimiatis *et al.* [11].

There also have been some variational methods for PD. Rudin *et al.* proposed a model for TV based image denoising in [16]. The basic denoising model was customised by Le *et al.* in [10] for PD. As this model is technically relevant for our approach, let us give some more details. As the data term naturally includes the formulation for Poisson noise, the data term in [10] resembles the one we will formulate, as is also the case in [7, 8]. However, we will introduce a slight technical modification enabling better algorithmical treatment. In contrast to our work, Le *et al.* consider only TV regularisation. As another major difference we make use of a modern, non-trivial numerical optimisation approach for the numerical realisation, whereas Le *et al.* consider a simple finite difference method solving the Euler-Lagrange equation for their functional. Figueiredo *et al.* employ for the regulariser besides TV also frame-based models. The arising optimisation problems are tackled with the alternating direction method of multipliers. The recent variational model of Feng *et al.* follows the field of experts methodology for devising the regulariser [7].

In summary, we introduce several alternatives in modeling and numerics compared to the works [7, 8, 10]. Especially, we consider a novel anisotropic model for the problem and make use of efficient numerical tools.

Turning to the use of anisotropic diffusion, Weickert presented an extensive framework in [20]. The anisotropic regulariser we employ here has been proposed before in the context of diffusion [20]. In [21] accurate numerics, such as a family of specific finite difference schemes is given. We also take into account these recent, dedicated approximations within our numerical framework.

2 Poisson Noise

Suppose we have a Poisson noisy image $u^p \in \mathbb{N}_0^{MN}$ with $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, that is the result of a random experiment with the probability distribution

$$\mathbf{P}(U = u^p | u^0) := \prod_{(i,j) \in I} P(U_{i,j} = u_{i,j}^p | u_{i,j}^0), \quad (2)$$

where for every $(i, j) \in I := \{1, \dots, M\} \times \{1, \dots, N\}$ the parameter $u_{i,j}^0 \geq 0$ is supplied by the original image u^0 and P is the discrete Poisson distribution

$$P(X = x | y) := \begin{cases} \frac{(y)^x}{x!} \exp(-y), & y > 0, \\ \delta_{\{0\}}(x), & y = 0. \end{cases} \quad (3)$$

Here, $\delta_{\mathcal{S}}(z)$ is the Kronecker delta function that equals 1 if $z \in \mathcal{S}$ and 0 otherwise.

For any $u \in \mathbb{R}^{MN}$ we define sets of indices $I^+(u) := \{(i, j) \in I : u_{i,j} > 0\}$, $I^0(u) := \{(i, j) \in I : u_{i,j} = 0\}$ and $I^-(u) := \{(i, j) \in I : u_{i,j} < 0\}$. While \mathbf{P} in (2) is defined only for nonnegative u^0 , the likelihood of any $u \in \mathbb{R}^{MN}$ to be the image that led to u^p can be computed by

$$\mathcal{P}(u | u^p) := \delta_{\{\emptyset\}}(I^-(u)) \prod_{(i,j) \in I \setminus I^-(u)} P(U_{i,j} = u_{i,j}^p | u_{i,j}). \quad (4)$$

With the probability distribution (3) we immediately see that $\mathcal{P}(u | u^p) > 0$ if and only if $I^-(u) = \emptyset$ and $u_{i,j} > 0$ for all $i \in I^+(u^p)$. Because of $\lim_{x \rightarrow 0} x \log x = 0$, the log-likelihood is given by

$$\begin{aligned} \mathcal{L}(u | u^p) := \log(\mathcal{P}(u | u^p)) &= \sum_{(i,j) \in I^+(u)} (u_{i,j}^p \log u_{i,j} - \log u_{i,j}^p! - u_{i,j}) \\ &\quad - \gamma_{\{\emptyset\}}(I^-(u)) - \sum_{(i,j) \in I^0(u)} \gamma_{\{0\}}(u_{i,j}^p), \end{aligned} \quad (5)$$

where $\gamma_{\mathcal{S}}(z)$ is the indicator function that equals 0 if $z \in \mathcal{S}$ and ∞ otherwise. For a given u^p , maximizing $\mathcal{L}(u | u^p)$ is equivalent to minimizing

$$\sum_{(i,j) \in I^+(u)} (u_{i,j} - u_{i,j}^p \log u_{i,j}) + \gamma_{\{\emptyset\}}(I^-(u)) + \sum_{(i,j) \in I^0(u)} \gamma_{\{0\}}(u_{i,j}^p). \quad (6)$$

Minimizing a data fidelity in form of (6) is an approach for PD, which was used *e.g.* in [7, 8, 10].

For the noise model sometimes also a mixed Poisson distribution is used, *e.g.* in [12]. We assume that the image u^p has only Poisson noise, for a detailed overview of noise sources in scanning electron microscopy see [18].

3 Anisotropic Diffusion

As we rely in this paper on the variational formulation of anisotropic diffusion in the sense of Weickert [20] let us briefly recall some details.

The anisotropic diffusion flow within the image plane during a given time can be described by the partial differential equation (PDE) $\partial_t u = \text{div}(D\nabla u)$. Thereby D denotes the diffusion tensor with $D = v_1 v_1^\top + \lambda_2 v_2 v_2^\top$, where the vectors $v_1 \parallel \nabla u^\sigma$ and $v_2 \perp \nabla u^\sigma$ are of unit length and $\lambda_2 = g(\|\nabla u^\sigma\|^2)$, see also [20]. The function $g(\cdot)$ denotes the diffusivity function. By u^σ we denote the convolution of u with a Gaussian of standard deviation σ . For the spatial derivatives in ∇u^σ we use here central differences.

In [21] a discretisation of the term $\text{div}(D\nabla u)$ is obtained by minimizing

$$R(u) := \frac{1}{2} \int_{\Omega} \nabla^\top u D \nabla u dx dy, \tag{7}$$

where D is in this notation the time-invariant diffusion tensor

$$D = \begin{pmatrix} a(x, y) & b(x, y) \\ b(x, y) & c(x, y) \end{pmatrix}. \tag{8}$$

Following [21] we remark that (7) can be discretised and written as

$$R(u) := -\frac{1}{2} u^\top A u \quad \text{and} \quad \nabla R(u) = -A u \tag{9}$$

with some negative semidefinite matrix A . An explicit time discretisation can be computed by iterating $u^{(l+1)} = u^{(l)} + \alpha A(u^{(l)})u^{(l)}$, where α is a suitable step size and $A(u^{(l)})$ is the matrix A , which stems from the diffusion tensor using the iterate $u^{(l)}$, see [21] for more details. Similar to that, we employ a *lagged diffusivity* approach and evaluate A always at the given iterate in our optimisation method.

4 Our Variational Approach

As indicated we aim to find a local minimiser $u^* \in \mathbb{R}^{MN}$ of

$$E(u) := R(u) + \tau S(u; u^p), \tag{10}$$

see (1). By E we want to combine anisotropic diffusion with data fidelity $S(\cdot; u^p)$ customised for a certain noise model. Among other models that we also test here, we especially consider the log-likelihood data term as developed before in (5). In the total, we used for $S := S(u; u^p)$ the following functions:

$$S_2 := \frac{1}{2} \|u - u^p\|_2^2, \quad S_1 := \|u - u^p\|_1, \tag{11}$$

$$S_\ell := \sum_{(i,j) \in I^+(u)} (u_{i,j} - u_{i,j}^p \log u_{i,j}) + \gamma_{\{\emptyset\}}(I^-(u)) + \sum_{(i,j) \in I^0(u)} \gamma_{\{\emptyset\}}(u_j^p). \tag{12}$$

For the diffusivity $g := g(s^2)$, that contributes to R through the eigenvalue λ_2 , some possibilities that we consider are

$$g_{\text{PM}} := \frac{1}{1 + s^2/\lambda^2}, \quad g_{\text{ePM}} := \exp(-s^2/(2\lambda^2)), \quad (13)$$

$$g_{\text{Ch}} := \frac{1}{\sqrt{1 + s^2/\lambda^2}}, \quad g_{\text{W}} := \begin{cases} 1, & s^2 = 0, \\ 1 - \exp\left(\frac{-3.31488}{(s/\lambda)^8}\right), & s^2 > 0. \end{cases} \quad (14)$$

For the Perona-Malik diffusivity g_{PM} and the exponential Perona-Malik diffusivity g_{ePM} see [15], for the Charbonnier diffusivity g_{Ch} see [5] and for the Weickert diffusivity g_{W} see [20]. Each of the diffusivity functions depend on the contrast parameter λ . While the impact of λ on the diffusion process is different in each diffusivity, edges with a contrast below λ are smoothed out more than those with a contrast above λ , generally speaking.

5 A Numerical Solution Strategy

Our variational model $E(u) := R(u) + \tau S(u; u^p)$ requires the minimisation of a non-convex and non-smooth cost function. A common choice to handle the non-convexity is to embed our method into a lagged diffusivity fixed-point iteration, thereby fixing the diffusivities at the previous iteration and thus, considering

$$R(u; \hat{u}^{(l)}) := -\frac{1}{2} u^\top A(\hat{u}^{(l)}) u. \quad (15)$$

Now, the minimisation of $R(\cdot) + \tau S(\cdot; u^p)$ comes down to a series of convex optimisation tasks. $R(\cdot; \hat{u}^{(l)})$ is smooth and convex, although not necessarily strictly, nor strongly convex. On the other hand, our smoothness term $S_2(\cdot; u^p)$ is convex and smooth, whereas $S_1(\cdot; u^p)$ and $S_\ell(\cdot; u^p)$ are convex but continuous at the very best. Our convex energies $R(\cdot; \hat{u}^{(l)}) + \tau S(\cdot; u^p)$ also suggest a natural splitting into a sum of two terms, here $R(\cdot; \hat{u}^{(l)})$ and $\tau S(\cdot; u^p)$. One of our goals is to present a single concise but modular framework that is flexible enough to handle all presented setups efficiently. To this end we propose to use the inertial proximal algorithm for nonconvex optimization (iPiano) by Ochs et al. [14]. In its most generic form it considers the minimisation of cost functions $f + g$, where g is a convex (possibly non-smooth) and f a smooth (possibly non-convex) function. Thus, it suits our requirements well. The algorithm, inspired by the Heavy-ball method of Polyak, combines a forward-backward splitting scheme with an additional inertial force to improve the convergence properties. In its generic form it iterates

$$x^{(k+1)} = \text{prox}_{\alpha_k g} \left(x^{(k)} - \alpha_k \nabla f \left(x^{(k)} \right) + \beta_k \left(x^{(k)} - x^{(k-1)} \right) \right), \quad (16)$$

where α_k and β_k are parameters that stem from the numerical scheme. Here, the function $\text{prox}_{\alpha g}$ denotes the proximal operator given by

$$\text{prox}_{\alpha g}(y) := \arg \min_x \left(\frac{1}{2} \|x - y\|^2 + \alpha g(x) \right). \quad (17)$$

Algorithm 1. Poisson Image denoising via iPiano

Choose $\sigma = \frac{3}{4}$, a diffusivity $g(\cdot)$, a diffusivity parameter $\lambda > 0$, a data fidelity $S(\cdot; u^p)$, $\tau > 0$, $c = \frac{1}{1000}$, $\beta = \frac{4}{5}$

Initialise $\hat{u}^{(1)} = u^p$

repeat

 Compute D for $\hat{u}^{(l)}$ and the matrix $A(\hat{u}^{(l)}) = A_l$ that realises (9)

 Estimate Lipschitz constant L_l by Geršgorin circle theorem

 Compute $\alpha_l = \frac{2(1-\beta)}{L_l+c}$

 Initialise $u^{(0)} = u^{(1)} = \hat{u}^{(l)}$

repeat

$$u^{(k+1)} = \text{prox}_{\alpha_l \tau S(\cdot; u^p)} \left(u^{(k)} + \alpha_l A_l u^{(k)} + \beta_k \left(u^{(k)} - u^{(k-1)} \right) \right)$$

until convergence of the iPiano scheme towards $u^{(\infty)}$

 Set $\hat{u}^{(l+1)} = u^{(\infty)}$

until convergence of the lagged diffusivity scheme

Applying the iPiano formalism to our setup in (10) yields the following iterative strategy:

$$u^{(k+1)} = \text{prox}_{\alpha_k \tau S(\cdot; u^p)} \left(u^{(k)} - \alpha_k \nabla R \left(u^{(k)}; \hat{u}^{(l)} \right) + \beta_k \left(u^{(k)} - u^{(k-1)} \right) \right) \quad (18)$$

with initial values $u^{(0)} = u^{(1)}$ and certain parameters α_k and β_k .

Whether the iPiano algorithm is fast hinges on an effective evaluation of the proximal mapping. For our choices of S_1 , S_2 , and S_ℓ the corresponding proximal mappings are well known and can be expressed in closed form. The proximal mapping of S_1 is given by the soft shrinkage operation, while it corresponds to a simple averaging for S_2 . Finally, the proximal mapping for S_ℓ is easily found by setting the first derivative to 0. We also remark in this context that it suffices in each case to consider the 1D formulation. The computation of the proximal mapping decouples for all our choices of S .

Besides the computation of the proximal mapping we must also specify an update strategy for the free parameters α_k and β_k in (18) (resp. (17)). Here, we follow the recommendations found in [14] and set $\beta_k = \frac{4}{5}$ and $\alpha_k = \frac{2(1-\beta_k)}{L+c}$, where $c = \frac{1}{1000}$ and L being an upper bound of the spectral norm of $A(\hat{u}^{(l)})$, that is, the Lipschitz constant of $\nabla R(\cdot; \hat{u}^{(l)})$. For performance reasons we opt to estimate L by means of the Geršgorin circle theorem. A detailed listing of our approach is given in Algorithm 1.

We also use the following stopping criteria in our implementation. The outer iteration stops when $\hat{u}^{(l+1)}$ fulfils either

$$\frac{|E_{l+1}(\hat{u}^{(l+1)}) - E_l(\hat{u}^{(l)})|}{|E_{l+1}(\hat{u}^{(l+1)}) + E_l(\hat{u}^{(l)})|} < 5 \cdot 10^{-7} \quad (19)$$

or

$$\frac{1}{MN} \left\| \hat{u}^{(l+1)} - \hat{u}^{(l)} \right\|_1 < 10^{-6} \|u^p\|_\infty, \quad (20)$$

where E_l denotes the function E with the diffusion tensor fixed with the iterate $\hat{u}^{(l)}$. For the iPiano algorithm we use nearly the same break criteria, we substitute $\hat{u}^{(l+1)}$ with $u^{(k+1)}$, $\hat{u}^{(l)}$ with $u^{(k)}$ and E_{l+1} with E_l .

6 Additional Model Improvements

Even though our approach already delivers convincing results, we investigate further means to improve the quality of our findings. Our real world data is assumed to have hard edges with almost constant regions in between in the ideal case. Such results can, for example, be obtained by adding a second regulariser in form of a TV term. Thus, we must minimise the following energy.

$$F(u) := R(u) + \tau S(u ; u^p) + \kappa \text{TV}(u) \tag{21}$$

Algorithm 2. Poisson Image denoising via three-operator splitting

Choose $\sigma = \frac{3}{4}$, a diffusivity $g(\cdot)$, a diffusivity parameter $\lambda > 0$, a data fidelity $S(\cdot ; u^p)$, $\tau > 0$, $c = \frac{1}{1000}$, $\kappa > 0$

Initialise $\hat{u}^{(1)} = u^p$

repeat

 Compute D for $\hat{u}^{(l)}$ and the matrix $A(\hat{u}^{(l)}) = A_l$ that realises (9)

 Estimate Lipschitz constant L_l by Geršgorin circle theorem

 Compute $\alpha_l = \frac{2}{L_l + c}$

 Initialise $u^{(1)} = \hat{u}^{(l)}$

repeat

$$x^{(k)} = \text{prox}_{\alpha_l \kappa T} \left(u^{(k)} \right)$$

$$y^{(k)} = \text{prox}_{\alpha_l \tau S(\cdot ; u^p)} \left(2x^{(k)} - u^{(k)} + \alpha_l A_l x^{(k)} \right)$$

$$u^{(k+1)} = u^{(k)} + y^{(k)} - x^{(k)}$$

until convergence of the three-operator splitting scheme towards $y^{(\infty)}$

 Set $\hat{u}^{(l+1)} = y^{(\infty)}$

until convergence of the lagged diffusivity scheme

Here, $\text{TV}(u)$ defines the discretised form of the total variation of a smooth function.

$$\text{TV}(u) := \sum_{(i,j) \in I} \left\| \begin{pmatrix} d_x u_{i,j} \\ d_y u_{i,j} \end{pmatrix} \right\|_2, \tag{22}$$

where d_x and d_y denote standard forward differences. To find a minimiser u^* we proceed similarly as before. We use a lagged diffusivity fixed-point iteration to overcome the nonlinearities in the regulariser R . Similarly as before, we obtain a series of convex problems. The difference lies in the fact that we must minimise a sum of three terms now. Even though one could apply iPiano, we opt for a

better adapted algorithmic approach. We use the three-operator splitting scheme presented in [6] to handle the additional TV term. It requires us to evaluate the proximal mapping of the TV term

$$\text{prox}_{\alpha\text{TV}}(\tilde{u}) = \arg \min_u \left(\frac{\|u - \tilde{u}\|_2^2}{2} + \alpha \sum_{(i,j) \in I} \left\| \begin{pmatrix} d_x u_{i,j} \\ d_y u_{i,j} \end{pmatrix} \right\|_2 \right). \quad (23)$$

In this work, we use the algorithm of Chambolle and Pock [3]. It is appealing due to its simple form and decent efficiency. The complete algorithm, including the numerical strategy with the three-operator splitting is listed in Algorithm 2.

Here, we use the same stopping criteria as for Algorithm 1, see (19) and (20). In addition, we abort the iterative algorithm for computing $\text{prox}_{\alpha\text{TV}}(\tilde{u})$ when the following condition is met.

$$\frac{1}{MN} \|u^{(m+1)} - u^{(m)}\|_1 < 10^{-6} \left(\frac{9}{10}\right)^l \|u^p\|_\infty \quad (24)$$

Here, l corresponds to the loop counter for the lagged diffusivity scheme. By tightening the convergence requirements while progressing through the outer iteration of Algorithm 2, we aim to shorten computing time at the start and get a better final result. It seems plausible that the first few iterates can be rather rough, whereas accurate estimates are much more important towards the end.

7 Numerical Results

We use three different synthetic test images for a qualitative analysis of our methods. These images are chosen in such a way that they contain features commonly occurring in microscopic images. The considered samples are depicted in Fig. 2. For our evaluation we corrupt each image three times with Poisson noise in three different strengths, yielding a total of nine noisy images. One of them is shown

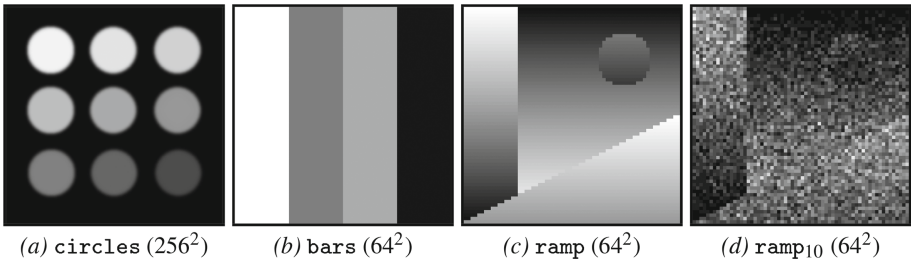


Fig. 2. (a–c) Noise-free original versions of our considered test images. The numbers in brackets indicate the size of the square images. (d) Exemplary noisy version of **ramp**. The lowercase index states the peak value, that the original image was scaled to before computing the Poisson noisy image. The noisy image has a range going from 0 to 19.

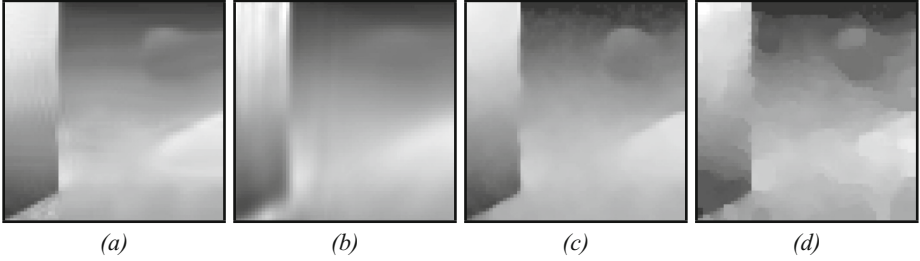


Fig. 3. (a) denoised image obtained by using the method from [2] for \mathbf{ramp}_{10} (SSIM = 0.7845) (b) denoised image obtained by using the method from [9] for \mathbf{ramp}_{10} (SSIM = 0.7437) (c) result of Algorithm 1 optimised for maximal SSIM with S_ℓ , g_{ePM} , $\lambda \approx 0.2446$, $\tau \approx 0.2443$ for \mathbf{ramp}_{10} (SSIM = 0.7817) (d) result of Algorithm 2 optimised for maximal SSIM with S_1 , g_W , $\lambda \approx 10.39$, $\tau \approx 96.74$, $\kappa \approx 226.2$ for \mathbf{ramp}_{10} (SSIM = 0.6552)

Table 1. MSE and SSIM for all tested images. Numbers in bold show the best result of all considered methods. The index after the image name denotes the peak value to which the image was scaled for the computation of the noisy image. Our approach outperforms the reference algorithms in terms of MSE for **bars** and **ramp**. Also, we obtain a higher SSIM for **bars** and **circles**.

Image	MSE			SSIM		
	[2]	[9]	Algorithm 1	[2]	[9]	Algorithm 1
bars ₂₅₅	1.3295	21.0788	1.6132	0.5381	0.4535	0.8264
bars ₁₀	23.9553	25.3805	17.0658	0.4448	0.4156	0.5685
bars ₁	428.8934	296.7009	292.8159	0.2744	0.3033	0.3122
ramp ₂₅₅	9.4805	160.4250	18.4692	0.9618	0.8086	0.9375
ramp ₁₀	114.0484	186.8883	111.2820	0.7845	0.7437	0.7817
ramp ₁	544.2952	702.1416	459.8652	0.4758	0.4836	0.5196
circles ₂₃₉	1.6322	0.8585	0.8934	0.8284	0.8176	0.9394
circles ₁₀	20.9328	9.7073	9.4121	0.6912	0.6563	0.8226
circles ₁	103.5318	56.7450	85.3593	0.2714	0.4334	0.6910

in Fig. 2, in an exemplary manner. In order to simulate different noise levels we rescale the images such that their peak value equals 255, 10, and 1 respectively before computing the sample of the probability distribution (2). In the last case the strength of noise is fairly high and the features of the original image are barely visible. We also use the mean squared error (MSE) and the structural similarity (SSIM) as a means to measure the quality of our reconstructions (see [19] for a reference on the SSIM). In addition, we consider the methods proposed in [2] and [9] for comparison. The authors of these works provide reference implementations of their algorithms, allowing us to do an objective evaluation.

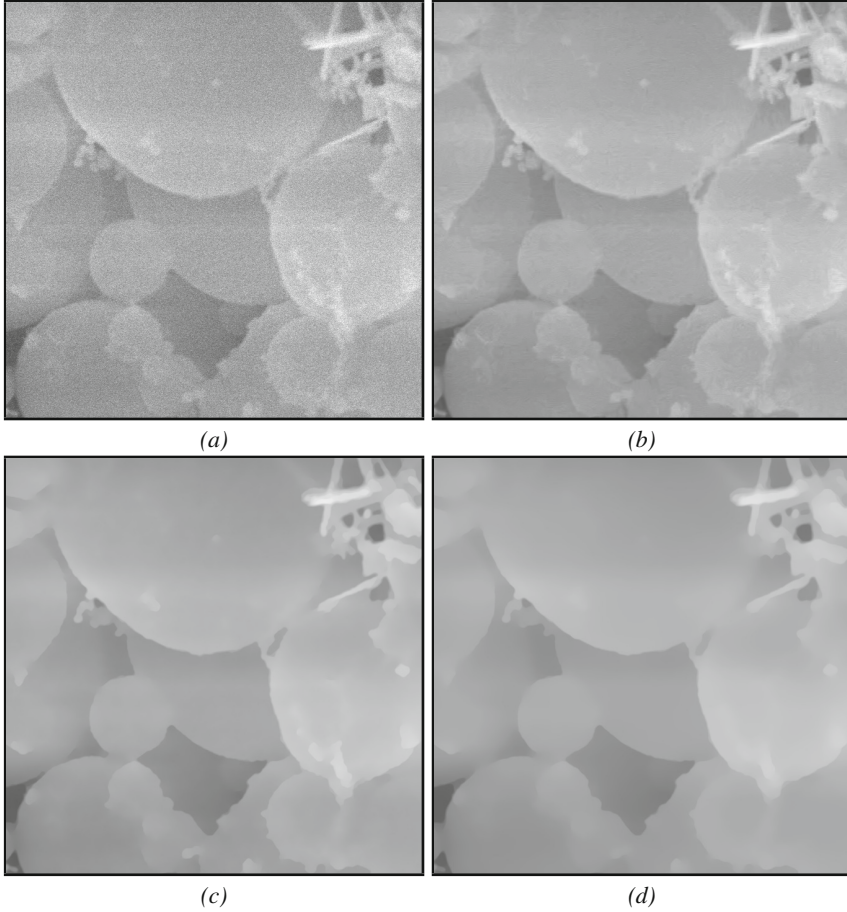


Fig. 4. (a) Original noisy image, cut out from Fig. 1 (b) result of VST from [2] (c) result of Algorithm 1 with S_ℓ , g_w $\lambda = 1.25$ and $\tau = 2.5$ (d) Result of Algorithm 2 with S_ℓ , g_w $\lambda = 1.25$ and $\tau = 2.5$ and $\kappa = 0.5$

Table 1 presents our findings for all considered images. Figure 3 shows one of the denoised images for every considered algorithm. The results of Algorithms 1 and 2 were achieved by thoroughly tuning all the model parameters to minimise the MSE respectively maximise the SSIM. The findings from Table 1 indicate that we achieve competitive results in most cases. If we focus solely on the SSIM, then our results are notably better (0.8226 vs. 0.6563 resp. 0.6912 for circles₁₀).

Finally, Fig. 4 presents the results of our algorithms on a real world example. The images depict gold sputtered PTFE particles recorded by a secondary electron detector of a ZEISS DSM 962 scanning electron microscope at an accelerating voltage of 20 kV. We have used a working distance of 6 mm and a magni-

fication factor of 5000. Since the ground truth is unknown, the parameters have been tuned to obtain the most visually appealing result. Clearly, we are able to remove noise and other artefacts while preserving the contrast and sharp edges.

8 Conclusion

We have shown that modern optimisation methods and discretisation schemes can be applied with benefit in electron microscopy. Although our basic anisotropic model is relatively simple, it already yields visually convincing results in our test application. The flexible variational framework enables many possible extensions that can be used for specific applications. For future work we may especially include a thorough study of possible optimisation methods for optimal balancing of data term and regularisers.

References

1. Anscombe, F.J.: The transformation of Poisson, binomial and negative-binomial data. *Biometrika* **35**(3/4), 246–254 (1948)
2. Azzari, L., Foi, A.: Variance stabilization for noisy+estimate combination in iterative Poisson denoising. *IEEE Signal Process. Lett.* **23**(8), 1086–1090 (2016)
3. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2010)
4. Chan, T.F., Shen, J.: *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. Society for Industrial and Applied Mathematics, Philadelphia (2005)
5. Charbonnier, P., Blanc-Feraud, L., Aubert, G., Barlaud, M.: Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.* **6**(2), 298–311 (1997)
6. Davis, D., Yin, W.: A three-operator splitting scheme and its optimization applications (2015). <https://arxiv.org/abs/1504.01032>
7. Feng, W., Qiao, H., Chen, Y.: Poisson noise reduction with higher-order natural image prior model. *SIAM J. Imaging Sci.* **9**(3), 1502–1524 (2016)
8. Figueiredo, M.A.T., Bioucas-Dias, J.M.: Restoration of Poissonian images using alternating direction optimization. *IEEE Trans. Image Process.* **19**(12), 3133–3145 (2010)
9. Giryes, R., Elad, M.: Sparsity-based Poisson denoising with dictionary learning. *IEEE Trans. Image Process.* **23**(12), 5057–5069 (2014)
10. Le, T., Chartrand, R., Asaki, T.J.: A variational approach to reconstructing images corrupted by Poisson noise. *J. Math. Imaging Vis.* **27**(3), 257–263 (2007)
11. Lefkimiatis, S., Maragos, P., Papandreou, G.: Bayesian inference on multiscale models for Poisson intensity estimation: applications to photon-limited image denoising. *IEEE Trans. Image Process.* **18**(8), 1724–1741 (2009)
12. Luisier, F., Blu, T., Unser, M.: Image denoising in mixed Poisson-Gaussian noise. *IEEE Trans. Image Process.* **20**(3), 696–708 (2011)
13. Mäkitalo, M., Foi, A.: Optimal inversion of the anscombe transformation in low-count poisson image denoising. *IEEE Trans. Image Process.* **20**(1), 99–109 (2011)
14. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: Inertial proximal algorithm for nonconvex optimization. *SIAM J. Imaging Sci.* **7**(2), 1388–1419 (2014)

15. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990)
16. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D* **60**(1–4), 259–268 (1992)
17. Salmon, J., Harmany, Z., Deledalle, C.A., Willett, R.: Poisson noise reduction with non-local PCA. *J. Math. Imaging Vis.* **48**(2), 279–294 (2013)
18. Timischl, F., Date, M., Nemoto, S.: A statistical model of signal-noise in scanning electron microscopy. *Scanning* **34**(3), 137–144 (2011)
19. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
20. Weickert, J.: *Anisotropic diffusion in image processing*. Teubner, Stuttgart (1998)
21. Weickert, J., Welk, M., Wickert, M.: L^2 -stable nonstandard finite differences for anisotropic diffusion. In: Kuijper, A., Bredies, K., Pock, T., Bischof, H. (eds.) *SSVM 2013. LNCS*, vol. 7893, pp. 380–391. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-38267-3_32](https://doi.org/10.1007/978-3-642-38267-3_32)
22. Zhang, B., Fadili, J., Starck, J.: Wavelets, ridgelets, and curvelets for Poisson noise removal. *IEEE Trans. Image Process.* **17**(7), 1093–1108 (2008)