Deblurring Poissonian images by split Bregman techniques

S. Setzer[a], G. Steidl[a], T. Teuber*,[a]

[a] *University of Mannheim, Dept. of Mathematics and Computer Science, A5, 68131 Mannheim, Germany*

Abstract

The restoration of blurred images corrupted by Poisson noise is an important task in various applications such as astronomical imaging, electronic microscopy, single particle emission computed tomography (SPECT) and positron emission tomography (PET). In this paper, we focus on solving this task by minimizing an energy functional consisting of the I-divergence as similarity term and the TV regularization term. Our minimizing algorithm uses alternating split Bregman techniques (alternating direction method of multipliers) which can be reinterpreted as Douglas-Rachford splitting applied to the dual problem. In contrast to recently developed iterative algorithms, our algorithm contains no inner iterations and produces nonnegative images. The high efficiency of our algorithm in comparison to other recently developed algorithms to minimize the same functional is demonstrated by artificial and real-world numerical examples.

Key words: Deblurring, Poisson noise, I-divergence, Kullback-Leibler divergence, TV, alternating split Bregman algorithm, Douglas-Rachford splitting.

1. Introduction

Blurred images corrupted by Poisson noise appear in various applications such as astronomical imaging [1], electronic microscopy [2, 3, 4], SPECT [5] and PET [6]. The MAP estimation via the negative log likelihood function resulting from Poissonian observations $f : \Omega \to \mathbb{R}$ of a convolved image, $\Omega \subset \mathbb{R}^2$ bounded and measurable, leads to the following variational problem

$$\operatorname*{argmin}_{u \geq 0} \Big\{ \int_\Omega Ku - f \log(Ku)\, dx + \lambda\, R(u) \Big\}, \quad \lambda > 0, \tag{1}$$

where K is a convolution kernel, which is supposed to be known. Since in most applications f is an image captured by some imaging system which detects, e.g., photons or electrons, the nonnegativity constraint $u \geq 0$ guarantees that no negative intensities occur in the restored image. Now consider $v = Ku$. The term

$$I(f, v) := \int_\Omega f \log \frac{f}{v} - f + v\, dx = \int_\Omega v - f \log v + f \log f - f\, dx$$

is known as Csiszár's *I-divergence* [7] or as *generalized Kullback-Leibler* divergence. If it is finite, it coincides with the *Bregman distance* [8] of the *Boltzmann-Shannon entropy*. Therefore, it shares the useful properties of a Bregman distance, in particular $I(f, v) \geq 0$. In this paper, we focus on the TV regularization term $R(u) := \sup \big\{ \int_\Omega u \operatorname{div} \psi\, dx : \psi \in C_c^1(\Omega, \mathbb{R}^2), |\psi(x)| \leq 1\ \forall x \in \Omega \big\}$, which reads for functions with (weak) derivatives in L_1 as follows:

$$R(u) := \int_\Omega |\nabla u|\, dx.$$

*Corresponding author

Email addresses: ssetzer@kiwi.math.uni-mannheim.de (S. Setzer), steidl@math.uni-mannheim.de (G. Steidl), tteuber@kiwi.math.uni-mannheim.de. T. T. gratefully acknowledges partial support by DFG Grant Ste 571/9-1. (T. Teuber)

Let BV denote the space of functions of bounded variation. It is well-known that for images g corrupted by additive white Gaussian noise, the minimizer of the Rudin-Osher-Fatemi (ROF) functional [9] is defined as

$$\operatorname*{argmin}_{u\in BV}\Big\{\int_\Omega \frac{1}{2}\,(g-u)^2\,dx+\lambda\int_\Omega |\nabla u|\,dx\Big\},\quad \lambda>0 \tag{2}$$

can be used as a denoised version of g with preserved edges. Recently, various algorithms were developed based on convex analysis to compute the minimizer of a discrete version of (2), e.g.,

- gradient descent reprojection algorithms [10], see also [11],
- Nesterov's algorithm [12],
- the (monotone) FISTA algorithm [13],
- alternating split Bregman algorithms [14].

The efficiency of the first three (first-order) algorithms for the denoising problem increases in the given order. For a comparison of first-order algorithms see also [15, 16, 17].

In the following, we will work *exclusively in a discrete setting*. To this end, we are stacking pixels of digital images columnwise into vectors in $\mathbb{R}^N$. If not stated otherwise the multiplication of vectors, their square roots, logarithms etc. are meant componentwise. We are interested in the discrete version of problem (1) given by

$$\operatorname*{argmin}_{u\in\mathbb{R}^N,u\geq 0}\left\{\langle 1, Ku-f\log Ku\rangle+\lambda\,\phi(Du)\right\}, \tag{3}$$

where $f\geq 0$, $K\in\mathbb{R}^{N,N}$ is a blur matrix, $D:=\begin{pmatrix}D_x\\D_y\end{pmatrix}\in\mathbb{R}^{2N,N}$ is the finite difference discretization of the gradient operator proposed in [11] and the functional $\phi:\mathbb{R}^{2N}\to\mathbb{R}$ is defined by

$$\phi(p):=\|\sqrt{p_1^2+p_2^2}\,\|_1,\quad p:=\begin{pmatrix}p_1\\p_2\end{pmatrix},\ p_1,p_2\in\mathbb{R}^N.$$

Of course, other consistent discretizations of the absolute value of the gradient are possible, see, e.g., [18]. For simplicity, we assume that $K^*1_N=1_N$, where 1_N denotes the vector consisting of N ones. This is usually fulfilled for blur matrices. The functional in (3) is proper, convex and lower semi-continuous (l.s.c.). If $K1_N\neq 0$, i.e., K does not annihilate constant functions as D does, then the functional is coercive. If K is invertible, the functional is strictly convex, so that there exists a unique minimizer of (3).

Recently, various algorithms were proposed to compute the minimizer of (3). In [19] Sawatzky et al. proposed an EM-TV algorithm which appears to be more efficient than other known (TV penalized Richardson-Lucy) algorithms as, e.g., those in [2, 3, 5, 20]. This algorithm combines alternatingly an EM step and a TV step as follows:

Algorithm (EM-TV)
Initialization: $u^{(0)}=f$.
For $k=1,\dots$ repeat until a stopping criterion is reached

$$u^{(k+\frac{1}{2})} = u^{(k)}K^*\big(\frac{f}{Ku^{(k)}}\big) \qquad \text{(EM step)}, \tag{4}$$

$$u^{(k+1)} = \operatorname*{argmin}_{v\in BV}\Big\{\frac{1}{2}\sum_{i=1}^N\frac{(u_i^{(k+\frac{1}{2})}-v_i)^2}{u_i^{(k)}}+\lambda\phi(Dv)\Big\} \quad \text{(TV step)}. \tag{5}$$

Indeed, the TV step requires us to minimize an ROF (-like) functional

$$\operatorname*{argmin}_w\left\{\frac{1}{2}\|w-h\|_2^2+\lambda\phi(D\Lambda w)\right\},\quad w:=\frac{v}{\sqrt{u^{(k)}}},\ h:=\frac{u^{(k+\frac{1}{2})}}{\sqrt{u^{(k)}}},\ \Lambda:=\operatorname{diag}\Big(\sqrt{u^{(k)}}\Big)_{k=1}^N.$$

This can be done by one of the above-mentioned algorithms. An analytical examination of the EM-TV algorithm was announced in [21].

Our approach is mainly based on the Poisson image deconvolution by augmented Lagrangian (PIDAL) algorithm of Figueiredo and Bioucas-Dias in [22]. Both algorithms can be deduced by alternating split Bregman techniques as described in Section 3. One major difference consists in the different operator splittings. As a consequence, our algorithm *avoids inner iteration loops* and the associated problem of choosing an appropriate stopping rule. By contrast, the PIDAL algorithm needs an inner iteration loop for computing the minimizer of a discrete ROF functional similar to the EM-TV algorithm. Moreover, it only ensures that $Ku^{(k)}$ and finally $K\hat{u}$ is nonnegative. However, even if K has only nonnegative entries, this does not imply that $\hat{u}$ is nonnegative. In our algorithm the iterates $u^{(k)}$ fulfill indeed a *nonnegativity constraint*. Numerical examples show that our new algorithm is more efficient than the EM-TV and the PIDAL algorithms. For another algorithm for removing Poisson noise which uses operator splittings we refer to [23].

This paper is structured as follows. In Section 2, we review the alternating split Bregman algorithm from the point of view we need for our settings. Then, we apply the results to obtain the known PIDAL algorithm and our novel algorithm in Section 3. In Section 4, we illustrate the good performance of our algorithm by numerical experiments. In particular, we compare our algorithm with the EM-TV algorithm and the PIDAL algorithm.

2. Alternating split Bregman algorithm

Let $\overline{\mathbb{R}}$ denote the extended real numbers $\mathbb{R} \cup \{+\infty\}$. The alternating split Bregman algorithm was introduced in [14] to efficiently solve problems of the form

$$\operatorname*{argmin}_{u \in \mathbb{R}^N} \{\Psi(u) + \Phi(Cu)\}, \tag{6}$$

where $C \in \mathbb{R}^{M,N}$ and $\Phi : \mathbb{R}^N \to \overline{\mathbb{R}}$ and $\Psi : \mathbb{R}^M \to \overline{\mathbb{R}}$ are proper, convex and l.s.c. functionals. In the following, we assume that (6) has a solution and

$$0 \in \operatorname{int}(C\operatorname{dom}\Psi - \operatorname{dom}\Phi), \qquad 0 \in \operatorname{int}(\operatorname{dom}\Psi^* + C^*\operatorname{dom}\Phi^*).$$

To derive the alternating split Bregman algorithm, we consider the following constrained version of (6), cf. [24],

$$\min_{u \in \mathbb{R}^N, w \in \mathbb{R}^M} \{\Psi(u) + \Phi(w)\} \quad \text{s.t.} \quad Cu = w. \tag{7}$$

This constrained problem can be solved via the following *split Bregman method* which is the same as the *augmented Lagrangian method* applied to (7):

$$\begin{aligned} (u^{(k+1)}, w^{(k+1)}) &= \operatorname*{argmin}_{u \in \mathbb{R}^N, w \in \mathbb{R}^M} \big\{\Psi(u) + \Phi(w) + \frac{1}{2\gamma}\|b^{(k)} + Cu - w\|_2^2\big\}, \\ b^{(k+1)} &= b^{(k)} + Cu^{(k+1)} - w^{(k+1)}. \end{aligned}$$

Minimizing alternatingly with respect to v and w yields the *alternating split Bregman method*, introduced in [14]:

$$u^{(k+1)} = \operatorname*{argmin}_{u \in \mathbb{R}^N} \big\{\Psi(u) + \frac{1}{2\gamma}\|b^{(k)} + Cu - w^{(k)}\|_2^2\big\}, \tag{8}$$

$$w^{(k+1)} = \operatorname*{argmin}_{w \in \mathbb{R}^M} \big\{\Phi(w) + \frac{1}{2\gamma}\|b^{(k)} + Cu^{(k+1)} - w\|_2^2\big\}, \tag{9}$$

$$b^{(k+1)} = b^{(k)} + Cu^{(k+1)} - w^{(k+1)}. \tag{10}$$

As Esser pointed out in [25], this algorithm can be traced back to the *alternating direction method of multipliers* proposed in [26, 27, 28]. Interestingly, there exists a connection of (8)-(10) to the *Douglas-Rachford algorithm* applied to the dual problem of (6)

$$\hat{p} = \operatorname*{argmin}_{p \in \mathbb{R}^M} \{\Psi^*(-C^*p) + \Phi^*(p)\}, \tag{11}$$

where Ψ^* denotes the conjugate function of Ψ. To derive the Douglas-Rachford algorithm for this problem, we rewrite (11) using Fermat's rule and that $0 \in \text{int}(\,\text{dom}\Psi^* + C^*\text{dom}\Phi^*)$, which gives

$$0 \in A(\hat{p}) + B(\hat{p}), \tag{12}$$

where $A := \partial(\Psi^* \circ (-C^*))$ and $B := \partial\Phi^*$. Since Ψ and Φ are proper, convex and l.s.c., the set-valued operators A and B are maximal monotone and their *resolvents* defined by $J_A := (I + A)^{-1}$ are single-valued again, cf., [29]. Then, the Douglas-Rachford splitting algorithm reads as follows:

Algorithm (Douglas-Rachford)
Initialization: $t^{(0)}, p^{(0)}$.
For $k = 1, \ldots$ repeat until a stopping criterion is reached

$$\begin{aligned} t^{(k+1)} &= J_{\eta A}(2p^{(k)} - t^{(k)}) + t^{(k)} - p^{(k)}, \\ p^{(k+1)} &= J_{\eta B}(t^{(k+1)}). \end{aligned}$$

It is well-known, see [30, 31, 32, 33], that in finite-dimensional spaces the sequences $(t^{(k)})_{k\in\mathbb{N}}$ and $(p^{(k)})_{k\in\mathbb{N}}$ converge for any initialization, where the latter converges to a solution of (12). The relation between the alternating split Bregman algorithm and the Douglas-Rachford algorithm is given as follows, see [32, 28, 34].

Proposition 2.1 *For $\eta := 1/\gamma$ and the starting values $t^{(0)} := \eta(b^{(0)} + w^{(0)})$ and $p^{(0)} := \eta b^{(0)}$, the Douglas-Rachford splitting algorithm coincides with the alternating Split Bregman algorithm* (8)-(10) *in the sense that*

$$t^{(k)} = \eta(b^{(k)} + w^{(k)}), \quad p^{(k)} = \eta b^{(k)}, \quad k \geq 0.$$

Using this relation and the known convergence properties of the Douglas-Rachford splitting algorithm we obtain with a little more effort concerning the convergence of $u^{(k)}$ the following convergence result for the alternating split Bregman algorithm, see [32, 28, 34].

Proposition 2.2 *For all starting values and any step sizes $\gamma > 0$, the sequences $(b^{(k)})_{k\in\mathbb{N}}$ and $(w^{(k)})_{k\in\mathbb{N}}$ generated by the alternating split Bregman method* (8)-(10) *converge. Denote by $\hat{b}$ and $\hat{w}$ the limits of these sequences. Then, $\frac{1}{\gamma}\hat{b}$ is a solution of the dual problem* (11)*. Furthermore, the sequence $(u^{(k)})_{k\in\mathbb{N}}$ calculated by* (8) *converges to a solution of the primal problem* (6) *if one of the following conditions are fulfilled:*
i) The primal problem has a unique solution.
ii) The problem $\operatorname{argmin}_{u\in\mathbb{R}^N} \left\{ \Psi(u) + \frac{1}{2\gamma}\|\hat{b} + Cu - \hat{w}\|_2^2 \right\}$ *has a unique solution.*

For our applications we consider instead of (6) the more general primal problem

$$\operatorname*{argmin}_{u\in\mathbb{R}^N} \sum_{i=1}^{m} f_i(C_i u), \tag{13}$$

where $C_i \in \mathbb{R}^{M_i,N}$ and $f_i : \mathbb{R}^{M_i} \to \overline{\mathbb{R}}$. Now, we can apply the splitting idea even further and write problem (13) as

$$\operatorname*{argmin}_{u\in\mathbb{R}^N, w_i\in\mathbb{R}^{M_i}} \langle 0, u\rangle + \sum_{i=1}^{m} f_i(w_i) \quad \text{s.t.} \quad w_i = C_i u. \tag{14}$$

Setting $\Psi := \langle 0, \cdot \rangle$ and $\Phi(w) := \sum_{i=1}^m f_i(w_i)$ in (8)-(10) we obtain the following alternating split Bregman algorithm

$$u^{(k+1)} = \operatorname*{argmin}_{u \in \mathbb{R}^N} \big\{ \langle 0, u \rangle + \frac{1}{2\gamma} \| \begin{pmatrix} b_1^{(k)} \\ \vdots \\ b_m^{(k)} \end{pmatrix} + \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} u - \begin{pmatrix} w_1^{(k)} \\ \vdots \\ w_m^{(k)} \end{pmatrix} \|^2 \big\}, \tag{15}$$

$$\begin{pmatrix} w_1^{(k+1)} \\ \vdots \\ w_m^{(k+1)} \end{pmatrix} = \operatorname*{argmin}_{w_i \in \mathbb{R}^{M_i}} \big\{ \sum_{i=1}^m f_i(w_i) + \frac{1}{2\gamma} \| \begin{pmatrix} b_1^{(k)} \\ \vdots \\ b_m^{(k)} \end{pmatrix} + \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} u^{(k+1)} - \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \|^2 \big\}, \tag{16}$$

$$b^{(k+1)} = b^{(k)} + \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} u^{(k+1)} - \begin{pmatrix} w_1^{(k+1)} \\ \vdots \\ w_m^{(k+1)} \end{pmatrix}. \tag{17}$$

The variables w_i in (16) are not coupled so that we can minimize separately with respect to the w_i, $i = 1, \ldots, m$. As already described in [25], this is very useful if more than one functional in the primal problem is a concatenation of a proper, convex and l.s.c. function with a linear operator. Let us have a brief look at the corresponding Douglas-Rachford splitting algorithm. Since

$$\Psi^*(x) = \langle 0, \cdot \rangle^*(x) = \begin{cases} 0 & \text{if } x = 0, \\ +\infty & \text{if } x \neq 0, \end{cases} \quad \text{and} \quad \Phi^*(x) = \sum_{i=1}^m f_i^*(x_i)$$

the operators A and B in (12) for the corresponding Douglas-Rachford splitting algorithm are $A(p) := \partial(\langle 0, \cdot \rangle^* \circ (-(C_1^* \cdots C_m^*)))(p)$ and $B(p) := (\partial f_1^*(p_1), \ldots, \partial f_m^*(p_m))$. More specifically, since

$$\partial(\langle 0, \cdot \rangle^*)(x) = \begin{cases} \mathbb{R}^N & \text{if } x = 0, \\ \emptyset & \text{if } x \neq 0, \end{cases}$$

we get for A that

$$\begin{aligned} A(p) &= -\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} \partial(\langle 0, \cdot \rangle^*)(-(C_1^* \cdots C_m^*)(p)) \\ &= \begin{cases} \big\{ \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} q, \ q \in \mathbb{R}^N \big\} & \text{if } \sum_{i=1}^m C_i^* p_i = 0, \\ \emptyset & \text{if } \sum_{i=1}^m C_i^* p_i \neq 0. \end{cases} \end{aligned}$$

By the definition of A, the resolvent $J_{\eta A}$ in the Douglas-Rachford splitting algorithm is the orthogonal projection onto the null space of $(C_1^* \cdots C_m^*)$, i.e., for any $p \in \mathbb{R}^{M_1 + \cdots + M_m}$ we have

$$J_{\eta A}(p) = P_{\mathcal{N}(C_1^* \cdots C_m^*)}(p) = p - \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} \operatorname*{argmin}_{y \in \mathbb{R}^N} \frac{1}{2} \| \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} y - p \|_2^2. \tag{18}$$

Observe that formula (18) has already the structure of (15).

By Proposition 2.2, the sequences $(b^{(k)})_{k \in \mathbb{N}}$ and $(w^{(k)})_{k \in \mathbb{N}}$ generated by the alternating split Bregman method (16)-(17) converge, in particular, $\frac{1}{\gamma} b^{(k)}$ converges to a solution of the dual problem. The sequence $(u^{(k)})_{k \in \mathbb{N}}$ converges to the solution of the primal problem (13) if either this primal problem has a unique solution or $\sum_{i=1}^m C_i^* C_i$ is invertible.

3. Deblurring methods for Poissonian images

In [22], problem (3) was considered without the constraint $u \geq 0$. With the notation from the previous section the authors apply the alternating split Bregman algorithm (15)-(17) for $m = 2$ and

$$\operatorname*{argmin}_{u\in\mathbb{R}^N} \big\{ \underbrace{\langle 1, Ku - f\log(Ku)\rangle}_{=:f_1(Ku)} + \underbrace{\lambda\,\phi(Du)}_{=:f_2(u)} \big\},$$

i.e., $(C_1, C_2) := (K, I)$. In particular, we have the constraint $w_2 = u$. Separating the minimization with respect to the variables w_1 and $u = w_2$ in step (16) this leads to the following PIDAL algorithm:

Algorithm (PIDAL)
Initialization: $b_1^{(0)} = b_2^{(0)} = 0$, $w_1^{(0)} = Kf$, $w_2^{(0)} = f$.
For $k = 1, \ldots$ repeat until a stopping criterion is reached

$$\begin{aligned}
u^{(k+1)} &= \operatorname*{argmin}_{u\in\mathbb{R}^N} \big\{ \|b_1^{(k)} + Ku - w_1^{(k)}\|_2^2 + \|b_2^{(k)} + u - w_2^{(k)}\|_2^2 \big\},\\
w_1^{(k+1)} &= \operatorname*{argmin}_{w_1\in\mathbb{R}^N} \big\{ \langle 1, w_1 - f\log w_1\rangle + \frac{1}{2\gamma}\|b_1^{(k)} + Ku^{(k+1)} - w_1\|_2^2 \big\},\\
w_2^{(k+1)} &= \operatorname*{argmin}_{w_2\in\mathbb{R}^N} \big\{ \lambda\phi(Dw_2) + \frac{1}{2\gamma}\|b_2^{(k)} + u^{(k+1)} - w_2\|_2^2 \big\},\\
b_1^{(k+1)} &= b_1^{(k)} + Ku^{(k+1)} - w_1^{(k+1)},\\
b_2^{(k+1)} &= b_2^{(k)} + u^{(k+1)} - w_2^{(k+1)}.
\end{aligned} \tag{19}$$

Output: $\hat{u}_{\text{approx}} := w_2^{(k+1)}$.

To be consistent with [22], we set our final result to be $w_2^{(k+1)}$ instead of $u^{(k+1)}$ which is justified by the constraint $u = w_2$. The first two steps of the PIDAL algorithm can be solved explicitely:

$$\begin{aligned}
u^{(k+1)} &= (I + K^{\mathrm{T}}K)^{-1}\left(K^{\mathrm{T}}(w_1^{(k)} - b_1^{(k)}) + (w_2^{(k)} - b_2^{(k)})\right),\\
w_1^{(k+1)} &= \frac{1}{2}\left(b_1^{(k)} + Ku^{(k+1)} - \gamma + \sqrt{\left(b_1^{(k)} + Ku^{(k+1)} - \gamma\right)^2 + 4\gamma f}\right).
\end{aligned}$$

Note that $I + K^*K$ is positive definite and therefore invertible. Moreover, supposing Neumann boundary conditions for our images, the corresponding linear system of equations can be solved efficiently by the discrete cosine transform (DCT-II) requiring only $\mathcal{O}(N\log N)$ arithmetic operations. In the third step (19), we have to solve an ROF denoising problem via one of the iterative algorithms mentioned in the introduction. The PIDAL algorithm ensures that the sequence $(w_1^{(k)})_{k\in\mathbb{N}}$ is nonnegative so that it converges to a nonnegative image $\hat{w}_1 = K\hat{u}$. Unfortunately, this does not imply that $\hat{u}$ itself is nonnegative even if K has only nonnegative entries summing up columnwise to 1.

The idea of our new method is to use the special structure of the regularization functional $\phi \circ D$ in (3) and to exploit the splitting idea even further. Moreover, we impose a nonnegativity constraint on u. Let $\iota_{u\geq 0}$ denote the indicator function of $\mathbb{R}^N_{\geq 0}$. We apply the alternating split Bregman algorithm (15)-(17) to

$$\operatorname*{argmin}_{u\in\mathbb{R}^N} \big\{ \underbrace{\langle 1, Ku - f\log(Ku)\rangle}_{=:f_1(Ku)} + \underbrace{\lambda\,\phi(Du)}_{=:f_2(Du)} + \underbrace{\iota_{u\geq 0}}_{=:f_3(u)} \big\},$$

i.e., $m = 3$ and $(C_1, C_2, C_3) := (K, D, I)$. In particular, we have the constraint $w_3 = u$. The corresponding alternating split Bregman algorithm, which we call *PIDSplit+* is given below. For

speed comparisons we also apply the algorithm without the positivity constraint, i.e., with $m = 2$ and $(C_1, C_2) = (K, D)$ and call this variant the *PIDSplit* algorithm.

Algorithm (PIDSplit+)
Initialization: $b_1^{(0)} = b_2^{(0)} = b_3^{(0)} = 0$, $w_1^{(0)} = Kf$, $w_2^{(0)} = Df$, $w_3^{(0)} = f$.
For $k = 1, \ldots$ repeat until a stopping criterion is reached

$$\begin{aligned}
u^{(k+1)} &= \operatorname*{argmin}_{u\in\mathbb{R}^N} \big\{ \|b_1^{(k)} + Ku - w_1^{(k)}\|_2^2 + \|b_2^{(k)} + Du - w_2^{(k)}\|_2^2 + \|b_3^{(k)} + u - w_3^{(k)}\|_2^2 \big\},\\
w_1^{(k+1)} &= \operatorname*{argmin}_{w_1\in\mathbb{R}^N} \big\{ \langle 1, w_1 - f\log(w_1)\rangle + \frac{1}{2\gamma}\|b_1^{(k)} + Ku^{(k+1)} - w_1\|_2^2 \big\},\\
w_2^{(k+1)} &= \operatorname*{argmin}_{w_2\in\mathbb{R}^{2N}} \big\{ \lambda\,\phi(w_2) + \frac{1}{2\gamma}\|b_2^{(k)} + Du^{(k+1)} - w_2\|_2^2 \big\},\\
w_3^{(k+1)} &= \operatorname*{argmin}_{w_3\in\mathbb{R}^N} \big\{ \iota_{w_3\geq 0}(w_3) + \frac{1}{2\gamma}\|b_3^{(k)} + u^{(k+1)} - w_3\|_2^2 \big\},\\
b_1^{(k+1)} &= b_1^{(k)} + Ku^{(k+1)} - w_1^{(k+1)},\\
b_2^{(k+1)} &= b_2^{(k)} + Du^{(k+1)} - w_2^{(k+1)},\\
b_3^{(k+1)} &= b_3^{(k)} + u^{(k+1)} - w_3^{(k+1)}.
\end{aligned}$$

Output: $\hat{u}_{\text{approx}} := w_3^{(k+1)}$.

Note that $w_3^{(k)}$ is nonnegative for all k, which is one reason to prefer $w_3^{(k+1)}$ over $u^{(k+1)}$. Thus, in contrast to the PIDAL algorithm we guarantee that $\hat{u}_{\text{approx}} \geq 0$ and we no longer need to solve an ROF denoising problem as inner iteration loop. Instead, $w_1^{(k+1)}, w_2^{(k+1)}, w_3^{(k+1)}$ can be computed by direct formulas and to get $u^{(k+1)}$ we solve a linear system of equations:

$$\begin{aligned}
u^{(k+1)} &= (I + K^{\mathrm{T}}K + D^{\mathrm{T}}D)^{-1}\left(K^{\mathrm{T}}(w_1^{(k)} - b_1^{(k)}) + D^{\mathrm{T}}(w_2^{(k)} - b_2^{(k)}) + (w_3^{(k)} - b_3^{(k)})\right),\\
w_1^{(k+1)} &= \frac{1}{2}\left(b_1^{(k)} + Ku^{(k+1)} - \gamma + \sqrt{\left(b_1^{(k)} + Ku^{(k+1)} - \gamma\right)^2 + 4\gamma f}\right),\\
w_2^{(k+1)} &= shrink_{\gamma\lambda}(b_2^{(k)} + Du^{(k+1)}),\\
w_3^{(k+1)} &= \begin{cases} b_3^{(k)} + u^{(k+1)} & \text{if } b_3^{(k)} + u^{(k+1)} \geq 0,\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

The $shrink_\lambda : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ operator denotes the *coupled soft shrinkage* which is given componentwise for $p := \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$, $p_i := (p_{ij})_{j=1}^N$, $i = 1, 2$, as

$$shrink_\lambda(p_{ij}) := \begin{cases} p_{ij} - \lambda p_{ij}/\sqrt{p_{1j}^2 + p_{2j}^2} & \text{if } \sqrt{p_{1j}^2 + p_{2j}^2} \geq \lambda,\\ 0 & \text{otherwise,}\end{cases}$$

see, e.g., [34]. Supposing Neumann boundary conditions for our images, the matrix $I + K^{\mathrm{T}}K + D^{\mathrm{T}}D$ can be diagonalized by the DCT-II. Thus, also in this case the solution of the corresponding linear system of equations requires only $\mathcal{O}(N \log N)$ arithmetic operations.

4. Numerical Results

Finally, we present numerical examples and compare our new PIDSplit+ algorithm with the EM-TV and PIDAL algorithms in terms of computational speed. To this end, we implemented all algorithms in MATLAB and performed the computations on a dual core desktop with two 2.4 GHz processors and 2.85 GB physical memory. For adding Poisson noise to an image U given

in double precision, we first scaled the image by $U_{\max} I_{\max}/10^{12}$, where $U_{\max}$ is the maximal value of U and $I_{\max}$ is a specified maximal intensity. Then, we applied the MATLAB routine `imnoise(I,'poisson')` and afterwards scaled back again. The values of the images lie all in the interval $[0, 255]$. To plot the results, we used the MATLAB routine `imagesc`, which incorporates an affine gray value scaling.

For solving the ROF-like subproblems (5) and (19) of the EM-TV and PIDAL algorithms in an efficient way, we tried different methods. To gain an additional speed up we applied the following initialization for the inner loops which was suggested by C. Brune [35]:

Inner loop initialization: Use the final values of an inner iteration loop as initial values for the next loop.

For solving (5), we decided to use the gradient descent reprojection algorithm [10], since it outperformed the FISTA algorithm [13] due to the fact that using the inner loop initialization described above the number of iterations needed per loop decreased faster for the gradient descent reprojection algorithm than for FISTA. Similarly, for solving the TV subproblem (19) we compared the gradient descent reprojection algorithm [10] to the alternating split Bregman algorithm in [14]. The second algorithm required to solve in each step a linear system of equations, which we solved again by using the discrete cosine transform (DCT-II). For the same reason as above, the gradient descent reprojection algorithm performed better and we use it for the experiments below. The corresponding step size was set to 0.249. Moreover, as suggested in [22] we used $\gamma = 50/\lambda$ for the PIDAL and also for the PIDSplit/PIDSplit+ algorithms.

Figure 1: Left: Original image showing a part of the 'cameraman' of size 84×84, Middle: Image blurred by a Gaussian kernel (standard deviation $\sigma = 1$) and contaminated by Poisson noise ($I_{\max} = 3000$). Right: Deblurred image by model (3) with $\lambda = 0.008$.

Algorithm	stopping value ε for inner iteration	number of outer iterations	computational time
PIDAL	0.0005	2132	830 sec
EM-TV	0.005	3696	63 sec
PIDSplit+	–	2150	12 sec
(PIDSplit)	–	(2149)	(11 sec)

Table 1: Computational time needed by the different algorithms to compute the result in Fig. 1 up to a maximal pixel difference of 0.1 to a converged reference image.

For our first example in Fig. 1, we took a part of the 'cameraman' image, blurred it by a Gaussian kernel and additionally contaminated it by Poisson noise. The deblurred image is depicted in Fig. 1 (right). Since for this example the restoration result of (3) without the additional

constraint $u \geq 0$ is already positive, all algorithms provide the same result. The computational times of the different algorithms are listed in Table 1. For direct comparison with the PIDAL algorithm, we also included the computation time of the PIDSplit algorithm. As stopping criterion for the inner iteration loops of the PIDAL and EM-TV algorithms, we checked if the maximal change per pixel in the image from one iteration to the next is less than an error bound ε. The results in Table 1 show that the PIDSplit/PIDSplit+ algorithms outperform the PIDAL and EM-TV algorithms significantly. Moreover, all algorithms need a similar number of outer iterations in this example. Only the EM-TV algorithm needed around two times as many.

Figure 2: Left: Original 'cameraman' image of size 256×256. Middle: Image blurred by a Gaussian kernel (standard deviation $\sigma = 1.3$) and contaminated by Poisson noise ($I_{\max} = 1000$). Right: Deblurred image by model (3) with $\lambda = 0.01$.

Algorithm	stopping value ε for inner iteration	number of outer iterations	computational time
EM-TV	0.01	3535	683 sec
PIDSplit+	–	1705	158 sec

Table 2: Computational time needed by the different algorithms to compute the result in Fig. 2 up to a maximal pixel difference of 0.3 to a converged reference image.

A restoration result and a comparison of the speed of the PIDSplit+ and EM-TV algorithms for the whole 'cameraman' image is given in Fig. 2 and Table 2. Note that for these results only a maximal pixel difference of 0.3 to a converged reference image was required.

To apply our PIDSplit+ algorithm also to real confocal data, our last example in Fig. 3 shows at the top left an image of a blurred human cell nucleus of size 401×311 provided by Prof. T. Cremer, LMU Munich. We thank S. Remmele for estimating the blur kernel to be a symmetric Gaussian of standard deviation 3.8. For this example, much more iterations are necessary to obtain equally small errors as for the 'cameraman' image. However, we get already after a few iterations a visually pleasant result as illustrated in Fig. 3. Here, it is also interesting to mention that if we solved model (3) for, e.g., $\lambda = 0.002$ without the additional nonnegativity constraint, the result would contain negative value down to -249 although the blurred image was nonnegative.

5. Summary and Conclusions

We have proposed a simple and very efficient algorithm for the restoration of blurred images corrupted with Poissonian noise by minimizing the TV penalized I-divergence. We use alternating split Bregman techniques (alternating direction method of multipliers) to decouple this problem

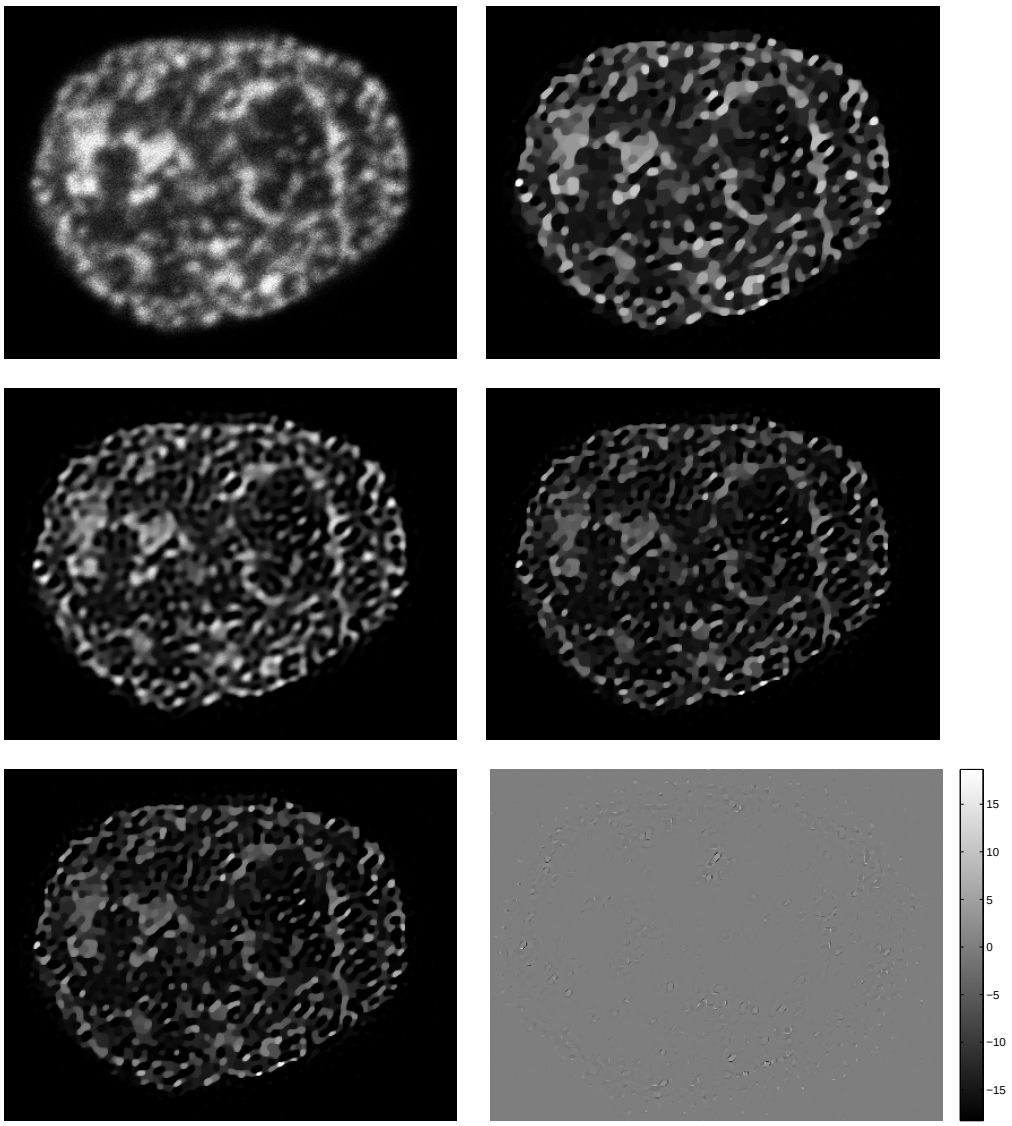

Figure 3: Top left: Blurry cell image of a human nucleus. Top right: Deblurred image by our new algorithm with $\lambda = 0.01$ and $\gamma = 1000$ after 10.000 iterations. Middle to bottom left: Results by our new algorithm with $\lambda = 0.002$ and $\gamma = 1000$ after 30 (middle left), 3000 (middle right) and 50.000 iterations (bottom left). One can hardly see any differences between the result after 3000 and 50.000 iterations, although there are still some differences in single pixels as the difference image at the bottom right shows.

in such a way that inner loops do not occur. Thus, in contrast to other algorithms, we avoid the choice of additional step length parameters and appropriate stopping rules. Further, our algorithm produces nonnegative restored images and the convergence is guaranteed by known convergence results of the alternating split Bregman algorithm.

References

[1] E. Bratsolis, M. Sigelle, A spatial regularization method preserving local photometry for Richardson-Lucy restoration, Astronomy and Astrophysics 375 (3) (2001) 1120–1128.

[2] N. Day, L. Blanc-Feraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, J. Zerubia,

Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution, Microspcopy Research and Technique 69 (4) (2006) 260–266.

[3] A. Kryvanos, J. Hesser, G. Steidl, Nonlinear image restoration methods for marker extraction in 3D fluorescent microscopy, in: C. A. Bouman, E. L. Miller (Eds.), Computational Imaging III, Proc. SPIE, Vol. 5674, 2005, pp. 432–443.

[4] S. Remmele, M. Seeland, J. Hesser, Fluorescence microscopy deconvolution based on Bregman iteration and Richardson-Lucy algorithm with TV regularization, in: T. Tolxdorff, J. Braun, T. M. Deserno, H. Handels, A. Horsch, H.-P. Meinzer (Eds.), Proceedings of the workshop: Bildverarbeitung für die Medizin 2008, Informatik aktuell, Spinger, Berlin, 2008, pp. 72–76.

[5] V. Y. Panin, G. L. Zeng, G. T. Gullberg, Total variation regulated EM algorithm, IEEE Transactions on Nuclear Science 46 (6) (1999) 2202–2210.

[6] L. A. Shepp, Y. Vardi, Maximum likelihood reconstruction for emission tomography, IEEE Transactions on Medical Imaging 1 (2) (1982) 113–122.

[7] I. Csiszár, Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems, The Annals of Statistics 19 (4) (1991) 2032–2066.

[8] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, USSR Computational Mathematics and Mathematical Physics 7 (3) (1967) 200–217.

[9] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992) 259–268.

[10] A. Chambolle, Total variation minimization and a class of binary MRF models, in: A. Rangarajan, B. Vemuri, A. L. Yuille (Eds.), Energy Minimization Methods in Computer Vision and Pattern Recognition, Vol. 3757 of LNCS, Springer, 2005, pp. 136–152.

[11] A. Chambolle, An algorithm for total variation minimization and applications, Journal of Mathematical Imaging and Vision 20 (1-2) (2004) 89–97.

[12] Y. Nesterov, Smooth minimization of non-smooth functions, Mathematical Programming 103 (1) (2005) 127–152.

[13] A. Beck, M. Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, IEEE Transactions on Image Processing. To appear.

[14] T. Goldstein, S. Osher, The Split Bregman method for L1-regularized problems, SIAM Journal on Imaging Sciences 2 (2) (2009) 323–343.

[15] J.-F. Aujol, Some first-order algorithms for total variation based image restoration, Journal of Mathematical Imaging and Vision 34 (3) (2009) 307–327.

[16] I. Loris, M. Bertero, C. De Mol, R. Zanella, L. Zanni, Accelerating gradient projection methods for ℓ_1-constrained signal recovery by steplength selection rules, Applied and Computational Harmonic Analysis 27 (2) (2009) 247–254.

[17] P. Weiss, L. Blanc-Féraud, G. Aubert, Efficient schemes for total variation minimization under constraints in image processing, SIAM Journal on Scientific Computing 31 (3) (2009) 2047–2080.

[18] M. Welk, G. Steidl, J. Weickert, Locally analytic schemes: a link between diffusion filtering and wavelet shrinkage, Applied and Computational Harmonic Analysis 24 (2) (2008) 195–224.

[19] A. Sawatzky, C. Brune, F. Wübbeling, T. Kösters, K. Schäfers, M. Burger, Accurate EM-TV algorithm in PET with low SNR, in: IEEE Nuclear Science Symposium Conference Record, 2008, pp. 5133–5137.

[20] J. A. Fessler, A. O. Hero, Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms, IEEE Transactions on Image Processing 4 (10) (1995) 1417–1429.

[21] C. Brune, A. Sawatzky, M. Burger, Bregman-EM-TV methods with application to optical nanoscopy, in: X.-C. Tai, K. Morken, M. Lysaker, K.-A. Lie (Eds.), Scale Space and Variational Methods in Computer Vision, Vol. 5567 of LNCS, Springer, 2009, pp. 235–246.

[22] M. A. T. Figueiredo, J. M. Bioucas-Dias, Deconvolution of Poissonian images using variable splitting and augmented Lagrangian optimization, in: IEEE Workshop on Statistical Signal Processing, Cardiff, 2009.

[23] C. Chaux, J.-C. Pesquet, N. Pustelnik, Nested iterative algorithms for convex constrained image recovery problems, SIAM Journal on Imaging Sciences 2 (2) (2009) 730–762.

[24] Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM Journal on Imaging Sciences 1 (3) (2008) 248–272.

[25] E. Esser, Applications of Lagrangian-based alternating direction methods and connections to split Bregman, CAM Report 09-31, UCLA (2009).

[26] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, Computers & Mathematics with Applications 2 (1) (1976) 17–40.

[27] R. Glowinski, A. Marrocco, Sur l'approximation par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique 9 (2) (1975) 41–76.

[28] D. Gabay, Applications of the method of multipliers to variational inequalities., in: M. Fortin, R. Glowinski (Eds.), Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary–Value Problems, Vol. 15 of Studies in Mathematics and its Applications, North–Holland, Amsterdam, 1983, Ch. 9, pp. 299–331.

[29] R. T. Rockafellar, R. J.-B. Wets, Variational Analysis, Springer, 2004.

[30] P. L. Combettes, J.-C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE Journal of Selected Topics in Signal Processing 1 (4) (2007) 564–574.

[31] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53 (5–6) (2004) 475–504.

[32] J. Eckstein, D. P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Mathematical Programming 55 (3) (1992) 293–318.

[33] P. L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM Journal on Numerical Analysis 16 (6) (1979) 964–979.

[34] S. Setzer, Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage, in: X.-C. Tai, K. Morken, M. Lysaker, K.-A. Lie (Eds.), Scale Space and Variational Methods in Computer Vision, Vol. 5567 of LNCS, Springer, 2009, pp. 464–476.

[35] C. Brune, Private communication.