Epigraphical Projection for Solving Least Squares Anscombe Transformed Constrained Optimization Problems

Harizanov Stanislav, Jean-Christophe Pesquet, Gabriele Steidl

Epigraphical Projection for Solving Least Square Anscombe Transformed Constrained Optimization Problems

Stanislav Harizanov[2], Jean-Christophe Pesquet[1], and Gabriele Steidl[2]

[1] Faculty of Mathematics, Université Paris-Est, France
[2] Department of Mathematics, University of Kaiserslautern, Germany

Abstract. This papers deals with the restoration of images corrupted by a non-invertible or ill-conditioned, linear transform and Poisson noise. Poisson data typically occurs in imaging processes where the images are obtained by counting particles, e.g., photons, that hit the image domain. Using the Anscombe transform Poisson noise can be approximated by additive Gaussian noise with zero mean and unit variance. Then the least squares difference between the Anscombe transformed corrupted image and the original image can be estimated by the number of image pixels. We use this information by considering an Anscombe transformed constrained model to restore the image. The advantage to corresponding penalized approaches consists in the existence of a good model parameter estimation. We solve the constrained minimization problem by applying primal-dual algorithms together with a projection onto the epigraph of a convex function related to the Anscombe transform. We show that this epigraphical projection can be computed by Newton's methods with an appropriate initial value. Numerical examples show the performance of our algorithm, in particular, the close relation to the I-divergence constrained model.

1 Introduction

The Poisson distribution exhibits a mean/variance relationship. This mean/variance dependence can be reduced by using variance-stabilizing transformations (VST), one of which is the *Anscombe transform* [1] defined as

$$T\colon [0,+\infty)^n \to (0,+\infty)^n\colon v=(v_i)_{1\le i\le n} \mapsto 2\left(\sqrt{v_i+\frac{3}{8}}\right)_{1\le i\le n}.$$

It transforms Poisson noise to approximately Gaussian noise with zero-mean and unit variance. The Anscombe transform has been employed in order to solve inverse problems where one wants to recover an original signal $\overline{u}\in[0,+\infty)^m$ from observations

$$f=\mathcal{P}(H\overline{u}),$$

where $\mathcal{P}$ denotes an independent Poisson noise corruption process and $H \in [0,+\infty)^{n\times m}$ is a linear degradation operator, e.g. a blur. Note that we consider images of size $M \times N$ columnwise reshaped as vectors of length $m = MN$.
In this context, one of the possible uses of the Anscombe transform is *i*) to transform the degraded observations f, *ii*) to apply a data recovery technique which is valid for an additive white zero-mean Gaussian model and *iii*) to apply an inverse transform to the so-recovered signal [7] (see also [17] for more recent developments). Note that this method appears mainly to be well-founded for denoising problems. When a linear degradation operator H is present, a better approach consists of adopting a variational framework [8, 12] where one minimizes a data fidelity term

$$u \mapsto \|T(Hu) - T(f)\|_2^2 \tag{1}$$

penalized by a (sum of) regularization term(s) serving to incorporate prior information about the sought signal $\overline{u}$. The approach is also closely related to a Maximum A Posteriori (MAP) estimate, where the function in (1) is substituted for the minus-log-likelihood of the Poisson noise, i.e., by the I-divergence (generalized Kullback-Leibler divergence)

$$u \mapsto D(b, Hu) := \begin{cases} \langle \mathbf{1}_n, f \log \frac{f}{Hu} - f + Hu \rangle & \text{if } Hu > 0, \\ +\infty & \text{otherwise,} \end{cases}$$

see [13, 20]. One of the drawbacks of these penalized methods is that multiplicative constants weighting the regularization terms (the so-called regularization parameters) need to be set carefully, which may be a difficult task.
A way of circumventing this problem consists of adopting a constrained approach instead of a regularized one, by imposing that

$$\|T(Hu) - T(f)\|_2^2 \leq \tau \tag{2}$$

where $\tau \in [0,+\infty)$. Based on the statistical properties of the Anscombe transform and the law of large number, a consistent choice for the above bound is $\tau = n$, when the number of pixels n is large. In this work, we will investigate such an approach by solving the following problem:

$$\underset{u\in C}{\text{minimize}} \;\; \Phi(Lu) \quad \text{subject to} \quad \|T(Hu) - T(f)\|_2^2 \leq \tau, \tag{3}$$

where C is a nonempty closed convex subset of $[0,+\infty)^m$, $L \in \mathbb{R}^{q\times m}$, and $\Phi\colon \mathbb{R}^q \to (-\infty,+\infty]$ is a proper, lower-semicontinuous, convex functions. A typical choice for C is the nonnegative orthant of $\mathbb{R}^m$. The classical Total Variation objective function [19] is obtained, as a special case, when Φ is an $\ell_{2,1}$ norm and L corresponds to a discrete gradient operator. Constrained models with the I-divergence have been considered in [5, 21], where in the second paper special attention was paid to the relation between the parameters of the constrained and the penalized problem via discrepancy principles. Note that recently penalized versus constrained problems in a rather general form were handled in [2]. In

[9], the I-divergence constraint was replaced through a polyhedral approximation technique and an epigraphical projection method was applied to solve the problem. In this work, we will also take advantage of an epigraphical projection approach to solve the Anscombe constrained model (1) and we will show that the required epigraphical projections can be easily determined in this context.
The structure of this paper is as follows: Section 2 recalls the notation. In Section 3 we determine the epigraphical projection onto a function related to the constraint (1) which plays the central role in the primal dual algorithms established in Section 4. In particular, we provide a good starting point for the involved Newton method. Numerical examples are presented in Section 5 underlining the good approximation of the I-divergence constrained approach by our Anscombe constrained model. Finally, a summary and conclusions are given in Section 6.

2 Notation

Let $\Gamma_0(\mathbb{R}^n)$ denote the set of proper, lower-semicontinuous, convex functions mapping from $\mathbb{R}^n$ to $(-\infty, +\infty]$. The *epigraph* of $\varphi \in \Gamma_0(\mathbb{R}^n)$ is the nonempty, closed, convex subset of $\mathbb{R}^{n+1}$ defined as

$$\operatorname{epi} \varphi := \{(u, \zeta) \in \mathbb{R}^n \times \mathbb{R} : \varphi(u) \leq \zeta\}.$$

For a nonempty, closed, convex set $C \subset \mathbb{R}^n$ we denote by $\iota_C \in \Gamma_0(\mathbb{R}^n)$ its *indicator function*

$$\iota_C(u) := \begin{cases} 0 & \text{if } u \in C, \\ +\infty & \text{otherwise,} \end{cases}$$

and by P_C the *orthogonal projector* onto C. Beyond epigraphs of functions from $\Gamma_0(\mathbb{R}^n)$ we will consider the half-spaces $V_\tau := \{u \in \mathbb{R}^n : \langle \mathbf{1}_n, u \rangle \leq \tau\}$, where $\mathbf{1}_n$ denotes the vector consisting of n entries 1. For a norm $\|\cdot\|$ on $\mathbb{R}^n$ we denote by $B_{\|\cdot\|} := \{u \in R^n : \|u\| \leq 1\}$ the unit ball with respect to this norm and by $\|\cdot\|_*$ the dual norm. Using this notation and defining, for every $i \in \{1, \ldots, n\}$,

$$\varphi_i \colon [0, +\infty) \to [0, +\infty) \colon s \mapsto \left(2\sqrt{s + \frac{3}{8}} - (T(f))_i\right)^2,$$

problem (1) can be rewritten as

$$\underset{(u,\zeta)}{\text{minimize}} \ \iota_C(u) + \Phi(Lu) + \sum_{i=1}^{n} \iota_{\operatorname{epi} \varphi_i}((Hu)_i + \frac{3}{8}, \zeta_i) + \iota_{V_\tau}(\zeta). \tag{4}$$

Now one can choose a primal-dual splitting algorithm as those proposed in [4, 6, 11, 22] to solve this problem. One step in all these algorithms consists in the orthogonal projections onto the epigraphs of φ_i for all $i = 1, \ldots, n$ which is the topic of the next section.

3 Epigraphical Projection

In this section we deal with the projection onto the epigraph of the function $\varphi \in \Gamma_0(\mathbb{R})$ defined as

$$\varphi(s) = \begin{cases} (2\sqrt{s} - z)^2 & \text{if } s \geq 0, \\ +\infty & \text{otherwise,} \end{cases} \tag{5}$$

where $z > 0$, see Fig. 1.

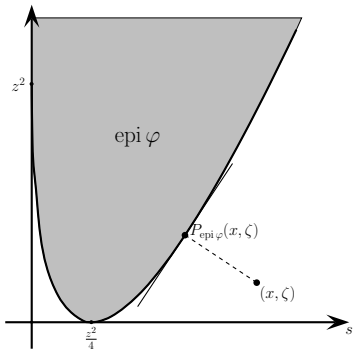

Fig. 1. The epigraph of φ for $z = 3$ and the epigraphical projection $P_{\text{epi}\,\varphi}(x,\zeta)$ of some point (x,ζ).

Proposition 1. *Let φ be defined by (3) with $z > 0$. Then the epigraphical projection of $(x,\zeta) \in \mathbb{R}^2$ is given by*

$$P_{\text{epi}\,\varphi}(x,\zeta) = \begin{cases} (\max\{x,0\},\zeta) & \text{if } (\max\{x,0\},\zeta) \in \text{epi}\,\varphi, \\ \left(\left(\frac{t_+ + z}{2}\right)^2, t_+^2\right) & \text{if } 4x \geq z^2, \\ \left(\left(\frac{t_- + z}{2}\right)^2, t_-^2\right) & \text{if } 4x < z^2, \end{cases}$$

where t_+, resp. t_- is the unique root in $[0,+\infty)$, resp. in $[-z,0)$ of the cubic polynomial

$$p(t) := 17t^3 + 3zt^2 + (3z^2 - 16\zeta - 4x)t + z(z^2 - 4x). \tag{6}$$

Proof. The function φ fulfills $\varphi(0) = z^2$ and

$$\varphi'(s) = 4 - \frac{2z}{\sqrt{s}} \begin{cases} < 0 \text{ if } 0 < s < \frac{z^2}{4}, \\ = 0 \text{ if } s = \frac{z^2}{4}, \\ > 0 \text{ if } s > \frac{z^2}{4} \end{cases}$$

and therefore $\lim_{\substack{s\to 0\\ s>0}} \varphi'(s) = -\infty$. Thus, if $x \le 0$ and $\zeta \ge z^2$, then $P_{\operatorname{epi}\varphi}(x,\zeta) = (0,\zeta)$. If $(x,\zeta) \in \operatorname{epi}\varphi$ then $P_{\operatorname{epi}\varphi}(x,\zeta) = (x,\zeta)$.
We consider the remaining cases $(\max\{x,0\},\zeta) \notin \operatorname{epi}\varphi$. The tangent vector of the curve associated with the graph of φ reads $(1,\varphi'(s))$, $s>0$. The uniquely determined orthogonal projection $(\hat{x},\hat{\zeta}) := P_{\operatorname{epi}\varphi}(x,\zeta)$ has to satisfy

$$\begin{pmatrix} x \\ \zeta \end{pmatrix} - \begin{pmatrix} \hat{x} \\ \hat{\zeta} \end{pmatrix} \perp \begin{pmatrix} 1 \\ \varphi'(\hat{x}) \end{pmatrix} \quad \text{and} \quad \hat{\zeta} = \varphi(\hat{x}), \quad \hat{x} \ge 0 \tag{7}$$

which leads to

$$0 = (x-\hat{x})\sqrt{\hat{x}} + 2\left(\zeta - (2\sqrt{\hat{x}} - z)^2\right)(2\sqrt{\hat{x}} - z), \quad \hat{x} \ge 0.$$

Substituting $\hat{t} := 2\sqrt{\hat{x}} - z \ge -z$ this can be rewritten as

$$0 = 17\hat{t}^3 + 3z\hat{t}^2 + (3z^2 - 16\zeta - 4x)\hat{t} + z(z^2 - 4x) = p(\hat{t}), \quad \hat{t} \ge -z.$$

Conversely, if $\hat{t} \ge -z$ is a root of the polynomial p in (1), then $\hat{x} = \left(\frac{\hat{t}+z}{2}\right)^2$ fulfills (3). When $x \ge z^2/4$ then also $\hat{x} \ge z^2/4$ (see Fig. 1), thus we are interested in the restriction of φ to $[z^2/4, +\infty)$ i.e., the nonnegative roots of p. The restriction of φ to $[z^2/4, +\infty)$ is convex, monotonically increasing, and $(x,\zeta) \notin \operatorname{epi}\varphi$. Hence, there is a unique point $(\hat{x},\hat{\zeta})$ on its graph that satisfies (3), i.e., p has a unique root in $[0,+\infty)$. Analogously, when $x < z^2/4$ then also $\hat{x} < z^2/4$, thus we are interested in the restriction of φ to $[0, z^2/4)$ i.e., the roots of p in the interval $[-z, 0)$. The restriction of φ to $[0, z^2/4)$ is convex and monotonically decreasing, and the uniqueness of the root follows by the same arguments. Finally, it can be noticed that $\hat{\zeta} = \varphi\left(\left(\frac{\hat{t}+z}{2}\right)^2\right) = \hat{t}^2$ since $\hat{t} \ge -z$, which completes the proof. □

The next proposition states that the root t_+, resp. t_-, of polynomial p can be computed efficiently by Newton's method with initial value $t_0 := 2\sqrt{\max\{x,0\}} - z$. Indeed we have seen in our numerical examples that this t_0 is a very good starting point.

Proposition 2. *Let* $(\max\{x,0\},\zeta) \notin \operatorname{epi}\varphi$ *and* $t_0 := 2\sqrt{\max\{x,0\}} - z$. *Let the polynomial* p *be defined by* (1). *Then the Newton method for solving* $p(t) = 0$ *with initial value* t_0 *converges (after a finite number of steps) monotonically to the root* t_+ *if* $4x - z^2 \ge 0$, *resp.,* t_- *if* $4x - z^2 < 0$.

Proof. 1. First we show that

i) $p(t_0)p(0) \le 0$,
ii) $p'(t_0) > 0$,

where equality in i) holds true iff $4x = z^2$.
If $x < 0$, then $t_0 = -z$ and consequently, since $(0,\zeta) \notin \operatorname{epi}\varphi$, i.e., $z^2 > \zeta$, we obtain

$$p(0) = z(z^2 - 4x) > 0, \quad p(t_0) = p(-z) = 16z(\zeta - z^2) < 0$$

which proves i). Further, since $p'(t) = 51t^2 + 6zt + 3z^2 - 4x - 16\zeta$ we obtain

$$p'(t_0) = p'(-z) = 48z^2 - 16\zeta - 4x \geq 16(4z^2 - \zeta) > 0.$$

If $x \geq 0$ then $t_0 = 2\sqrt{x} - z$. Consequently, we have $p(0) = -z(z + 2\sqrt{x})t_0$ and

$$p(t_0) = 17t_0^3 + 2zt_0^2 + (3z^2 - 16\zeta - 4x)t_0 + (t_0^2 + z^2 - 4x)z = 16t_0(t_0^2 - \zeta).$$

Since $(x, \zeta) \notin \operatorname{epi} \varphi$, i.e., $t_0^2 > \zeta$ we conclude that i) holds true. Finally ii) follows by

$$p'(t_0) = 51t_0^2 + 6zt_0 + 3z^2 - 16\zeta - 4x = 8(6t_0^2 - 2\zeta + x) = 16(t_0^2 - \zeta) + 8(4t_0^2 + x) > 0.$$

Since, in both cases, $p(t_0) \neq 0$, equality arises in i) iff $p(0) = 0$ i.e. $4x = z^2$.
2. The following result is well-known, see, e.g., [14, Theorem 18.3]: Newton's method for finding the unique root of a differentiable, convex, strictly increasing function on an interval converges monotonically if we start at the right endpoint of the interval. There is an analogue result for concave functions.
3. Since $p''(t) = 6(17t + z)$, p is convex for $t \geq -\frac{z}{17}$ and concave otherwise.
3.1 Let $t_0 > 0$ which implies $4x - z^2 > 0$. Then $p(0) < 0$ and, according to Part 1i), $p(t_0) > 0$. Hence, by Proposition 1, t_+ is the unique root of p in $(0, t_0)$. Since p is continuous, $p(t_+) = 0$ and $p(t_0) < 0$, we necessarily have $p'(t_+) \geq 0$ (otherwise there would exist another root of p on (t_+, t_0)). Thus, since p is strictly convex, it is strictly monotone increasing on $[t_+, t_0]$ and we can invoke the argument in Part 2 of the proof.
3.2 Let $t_0 < 0$ which implies $4x - z^2 < 0$. Then, $p(0) > 0$ and $p(t_0) < 0$ and by Proposition 1 we know that $t_- \in (t_0, 0)$ is the unique root of p in $[-z, 0)$. If $t_- \leq -\frac{z}{17}$, then we are done by similar arguments as in 3.1 for concave functions. It remains to consider the case when $t_- > -\frac{z}{17}$.
If $t_0 > -\frac{z}{17}$, then we know by the strict convexity of p on $[t_0, +\infty)$ that p' is strictly increasing on this interval and by Part 1ii) we further have $p'(t) > p'(t_0) > 0$ for every $t > t_0$. Thus, p itself is strictly increasing on $[t_0, +\infty)$. Consequently, one Newton step with initialization t_0 generates $t_1 = t_0 - p(t_0)/p'(t_0)$ and the convexity inequality $p(t_1) \geq p(t_0) + p'(t_0)(t_1 - t_0)$ shows that $p(t_1) \geq 0$. We thus are in the setting of Part 2 and the method converges monotonically.
Finally, let $t_0 \leq -\frac{z}{17}$ so that

$$4x \leq \left(\frac{16}{17}\right)^2 z^2. \tag{8}$$

Since $p(-\frac{z}{17}) < 0$ must be valid, a straightforward calculation yields

$$p\left(-\frac{z}{17}\right) = \frac{2z}{17}\left(\frac{z^2}{17} + 7z^2 + 8\zeta - 32x\right) < 0 \Leftrightarrow 8\zeta < 32x - \frac{z^2}{17} - 7z^2. \tag{9}$$

We shall now prove that $p'(t) > 0$ for all $t \in \mathbb{R}$, so that p is strictly increasing. Since p' is quadratic polynomial with minimimum at $-\frac{z}{17}$ we have only to show

that $p'\left(-\frac{z}{17}\right) = 2\left(\frac{24}{17}z^2 - 8\zeta - 2x\right) > 0$. Plugging in (3) and (3) we indeed obtain

$$\frac{1}{2}p'\left(-\frac{z}{17}\right) > \frac{24}{17}z^2 - 34x + \frac{z^2}{17} + 7z^2 > \frac{25}{17}z^2 - \frac{128}{17}z^2 + 7z^2 = \frac{16}{17}z^2 > 0.$$

Thus, after a finite number of steps the Newton algorithm arrives at some $t_k \geq t_-$ and we can apply Part 2 of the proof.
3.3 Let $t_0 = 0$ which implies $4z - t^2 = 0$. Then we get $t_+ = t_0$. □

4 Primal-Dual Algorithms

We can apply the projection onto the epigraph of φ in combination with any primal-dual algorithm proposed in [4, 6, 11, 22] or an alternating direction method of multipliers. Here we use exemplary the **p**rimal-**d**ual **h**ybrid **g**radient algorithm from [6, 18] with an extrapolation (**m**odification) of the dual variable which we will call PDHGMp. Based on the following reformulation of (2),

$$\begin{aligned} &\underset{(u,\zeta),(v,\eta)}{\text{minimize}} \ \iota_C(u) + \|v_2\| + \sum_{i=1}^{n} \iota_{\operatorname{epi}\varphi_i}(v_{1,i}, \eta_i) + \iota_{V_\tau}(\zeta) \\ &\text{subject to} \quad \left(\begin{array}{c|c} H & 0 \\ L & 0 \\ \hline 0 & I \end{array}\right) \begin{pmatrix} u \\ \zeta \end{pmatrix} + \begin{pmatrix} 3/8 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \eta \end{pmatrix}. \end{aligned} \tag{10}$$

where we have used $\Phi := \|\cdot\|$ and the nonnegative orthant C of $\mathbb{R}^m$, this algorithm reads:

Algorithm 1 (PDHGMp for solving the Anscombe constrained problem)
Initialization: $u^{(0)}, \zeta^{(0)}, p^{(0)} = p^{(-1)}$ and $\bar{p}^{(0)} = p^{(0)} + \theta(p^{(0)} - p^{(-1)})$, $\theta \in (0, 1]$, $\rho, \sigma > 0$ with $\rho\sigma < 1/\max\{1, \|H^*H + L^*L\|_2\}$
For $k = 0, 1, \ldots$ repeat until a stopping criteria is reached

1. $u^{(k+1)} = P_C\left(u^{(k)} - \sigma\rho\Big(H^*\bar{p}_1^{(k)} + L^*\bar{p}_2^{(k)}\Big)\right)$
2. $\zeta^{(k+1)} = P_{V_\tau}\left(\zeta^{(k)} - \sigma\rho\bar{p}_3^{(k)}\right)$
3. $(v_{1,i}^{(k+1)}, \eta_i^{(k+1)}) = P_{\operatorname{epi}\varphi_i}\left(p_{1,i}^{(k)} + (Hu^{(k+1)})_i + 3/8\,,\, p_{3,i}^{(k)} + \zeta_i^{(k+1)}\right), \quad i = 1, \ldots, n$
4. $v_2^{(k+1)} = (I - P_{\sigma^{-1}B_{\|\cdot\|_*}}(p_2^{(k)} + Lu^{(k+1)}) = \operatorname{prox}_{\sigma^{-1}\|\cdot\|}(p_2^{(k)} + Lu^{(k+1)})$
5. $p_1^{(k+1)} = p_1^{(k)} + Hu^{(k+1)} + 3/8 - v_1^{(k+1)}$
6. $p_2^{(k+1)} = p_2^{(k)} + Lu^{(k+1)} - v_2^{(k+1)}$
7. $p_3^{(k+1)} = p_3^{(k)} + \zeta^{(k+1)} - \eta^{(k+1)}$
8. $\bar{p}^{(k+1)} = p^{(k+1)} + \theta(p^{(k+1)} - p^{(k)})$.

Step 3 requires the epigraphical projections discussed in the previous section, Step 4 can be performed by coupled soft shrinkage with threshold σ^{-1} if we use the $\ell_{2,1}$-norm. The other steps can be computed in a straightforward way.

We will compare this algorithm with the PDHGMp applied to the I-divergence constrained problem

$$\underset{u\in C}{\text{minimize}} \ \|Lu\| \quad \text{subject to} \quad D(f,Hu) \le \tau_I \tag{11}$$

using a similar splitting as in (4) but without the variables ζ and η:

Algorithm 2 (PDHGMp for solving the I-divergence constrained problem)
Initialization: $u^{(0)}, p^{(0)} = p^{(-1)}$ and $\bar{p}^{(0)} = p^{(0)} + \theta(p^{(0)} - p^{(-1)})$, $\theta \in (0,1]$, $\rho, \sigma > 0$ with $\rho\sigma < 1/\|H^*H + L^*L\|_2$
For $k = 0, 1, \ldots$ repeat until a stopping criteria is reached

1. Step 1 of Algorithm 1
2. $\underset{v_1}{\text{minimize}} \ \|v_1 - (Hu^{(k+1)} + p_1^{(k)})\|_2^2 \quad \text{subject to} \quad D(f,v_1) \le \tau_I$ as in [21].

3. – 7. Steps 4. - 6. and 8. of Algorithm 1

In [23], see also [3], statistical arguments were used to show that $\tau_I = \frac{1}{2}n$ is a good estimate in case of moderate Poisson noise. In [5] this estimate was improved in case f has many zero components.

5 Numerical Examples

In this section we demonstrate the performance of our algorithm by numerical examples implemented in MATLAB (Intel Core i7-870 Processor with 8M Cache, 2.93 GHz, 8 GB physical memory). We have tested the two original images $\overline{u}$, namely 'cameraman' (256×256) and 'brain' (184×140) depicted in Fig. 2 and denoted by B1, resp. B2 in the following. The images were blurred by a Gaussian

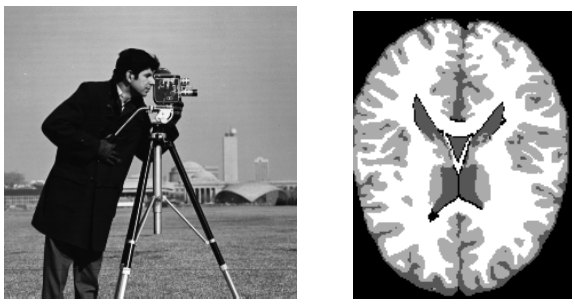

Fig. 2. Original images 'cameraman' (left) and phantom of a brain image (right).

matrix H with standard deviation 1.3 and mirrored boundary. Their gray values are interpreted as photon counts in the range $[0,\ \nu]$, where ν is the intensity of the image. We tested $\nu = 600, 1200, 2000, 3000$ and denoted the blurred, noisy images

by $\mathrm{B1}_\nu$ and $\mathrm{B2}_\nu$. For synthetically adding Poisson noise to the noise-free image we applied the MATLAB routine `imnoise(X,'poisson')`. For a quantitative comparison of the images we computed the peak signal to noise ratio (PSNR) and the (MAE) defined by $\mathrm{PSNR} = 10\log_{10}\frac{|\max\overline{u}-\min\overline{u}|^2}{\frac{1}{n}\|u-\overline{u}\|_2^2}$, and $\mathrm{MAE} = \frac{1}{n}\|\overline{u}-u\|_1$. The 'true' constraints between the blurred, noisy image f and the original image $\overline{u}$ are given in Table 1. As can be seen, the estimates $\tau_I = n/2$ and $\tau = n$ are good approximations of the true constraints $D(f, H\overline{u})$ and $\|T(H\overline{u}) - T(f)\|_2^2$.

	$D(f,H\overline{u})/n$	$\|T(H\overline{u})-T(f)\|_2^2/n$	PSNR	MAE
$\mathrm{B1}_{600}$	0.5020	1.0034	23.38	24.95
$\mathrm{B1}_{1200}$	0.5018	1.0039	23.79	44.60
$\mathrm{B1}_{2000}$	0.4979	0.9960	23.97	69.75
$\mathrm{B1}_{3000}$	0.4994	0.9989	24.06	100.73
$\mathrm{B2}_{600}$	0.5131	1.0004	20.17	36.61
$\mathrm{B2}_{1200}$	0.5122	1.0178	20.37	69.47
$\mathrm{B2}_{2000}$	0.5063	1.0085	20.49	112.05
$\mathrm{B2}_{3000}$	0.4956	0.9899	20.52	165.47

Table 1. The original values $D(f, H\overline{u})/n$, $\|T(H\overline{u}) - T(f)\|_2^2/n$ and PSNR, MAE of f.

We computed the minimizer of our functional (1) with the $l_{2,1}$-norm Φ and the discrete gradient operator L by Algorithm 1 with $\tau = n$. We compared the result with the I-divergence constrained approach (4) and Algorithm 2 with $\tau_I = n/2$. The parameters σ and τ appearing in the PDHGMp are fitted such that the algorithms give (up to two digits after the comma) the same PSNR, MAE and TV-norm (times 10^5 or 10^6) after 1000 iterations as after 100000 iterations. In the I-divergence constrained approach with the brain data we stopped after 5000 iterations. Further we have chosen $\theta = 1$. Fig. 3 shows the restoration results for $\mathrm{B1}_{1200}$ and Fig. 4 for $\mathrm{B2}_{1200}$. Table 2 summarizes the results for the different

Fig. 3. Result for the 'cameraman' image B1$_1$200 corresponding to Table 3. Corrupted image (left), restoration result by the Anscombe constrained model and Alg. 1 (middle), difference image between the middle image and those denoised by I-divergence constrained model and Alg. 2 (right). The gray values in the difference image are between -10 and 10, while the image values were scaled up to 1200.

Fig. 4. Result for the 'brain' image $B2_1200$ corresponding to Table 3. Corrupted image (left), restoration result by the Anscombe constrained model and Alg. 1 (middle), difference image between the middle image and those denoised by the I-divergence constrained model and Alg. 2 (right). The gray values in the difference image are between -15 and 15, while the image values were scaled up to 1200.

intensities. As expected, we observe that the outcomes of the two algorithms are very similar. More precisely, if u_A, resp. u_I denote the output of the restoration procedure with Anscombe, resp. I-divergence constraints, then we get for image B1 that $\|u_A - u_I\|_2/(\nu n)$ ranges from 2.24e-6 to 6.76e-7 and $\max|u_A - u_I|/\nu$ from 0.0082 to 0.0026 for the different noise levels.
Finally, Table 3 compares Algorithms 1 and 2 for different constraints τ and τ_I and a central part of the cameraman of size 130×130 with $\nu = 3000$ after 1000 iterations.

image	σ	ρ	$\|T(H\overline{u}) - T(f)\|_2^2 - n$	$D(f, Hu) - n/2$	TV-norm	PSNR	MAE
$B1_{600}$	0.0599	2.8	-1.5323	-	8.9230e+5	25.58	15.4126
			-	0.0049	8.9402e+5	25.59	15.40
$B1_{1200}$	0.042	3	9.0207	-	1.9190e+6	26.08	28.99
			-	0.045	1.9198e+6	26.08	28.98
$B1_{2000}$	0.021	4.108	0.7789	-	2.6780e+6	22.72	67.06
			-	0.7310	2.7331e+6	22.73	67.44
$B1_{3000}$	0.0329	4.001	-0.3559	-	5.1667e+6	26.64	67.49
			-	0.2767	5.1673e+6	26.64	67.48
$B2_{600}$	0.040	3.04	-2.7434	-	7.4158e+5	21.81	24.04
			-	9.0129	7.5693e+5	21.91	23.58
$B2_{1200}$	0.034	3.97	0.0079	-	1.5598e+6	22.33	43.52
			-	4.2683	1.5673e+6	22.36	43.34
$B2_{2000}$	0.021	4.108	0.7789	-	2.6780e+6	22.72	67.06
			-	0.8585	2.6885e+6	22.73	66.97
$B2_{3000}$	0.0182	5.413	-0.5284	-	4.0565e+6	22.93	96.23
			-	1.5536	4.0603e+6	22.94	96.09

Table 2. Results of Algorithms 1 and 2 with $\tau = n$ and $\tau_I = \frac{n}{2}$.

6 Summary and Conclusions

We have considered a constrained restoration model for images corrupted by a linear transform and Poisson noise which involves the Anscombe transform. An advantage of this approach in contrast to penalized ones is the existence of estimates for the model parameter. We have provided algorithms to find a minimizer

scale	σ	ρ	$\|T(H\overline{u}) - T(f)\|_2^2 - n$	$D(f, Hu) - n/2$	TV-norm	PSNR	MAE
0.8	0.2	0.975	-0.0813	-	2.2400e+6	26.51	66.10
			-	-0.0585	2.2416e+6	26.5176	66.0681
1.0	0.0269	4	-2.4868	-	1.7071e+6	25.60	63.60
			-	-0.0425	1.7073e+6	25.60	63.57
1.0095 (optimal)	0.0239	4.002	-0.0616	-	1.6964e+6	25.56	63.87
			-	-1e-4	1.6914e+6	25.54	63.98
1.2	0.011	8	1.1405	-	1.5563e+6	24.93	68.89
			-	0.0611	1.5565e+6	24.93	68.87
2	0.004	30.065	-1.7501	-	1.3388e+6	23.70	81.05
			-	-0.0257	1.3391e+6	23.71	81.06

Table 3. Results of Algotithms 1 and 2 on a part of $\mathrm{B1}_{3000}$ for different constraining parameters $\tau = \text{scale} \cdot n$ and $\tau_I = \text{scale} \cdot \frac{n}{2}$. The optimal scale is computed as in Table 1.

of the model based on epigraphical projections and have shown that the performance is similar to a recently introduced I-divergence constrained model. Future research comprises the following: *i)* replacing or combining the discrete gradient operator L with other ones (discrete higher order operators, nonlocal means, wavelet-like transforms) and handling other problems than deblurring ones, *ii)* considering convex optimization problems involving multiple constraints

$$\underset{u\in\mathbb{R}^n}{\text{minimize}} \sum_{r=1}^{R} \Phi_r(L_r u) \quad \text{subject to} \quad H_s u \in C_s, \quad s = 1, \ldots, S$$

with linear operators $L_r : \mathbb{R}^n \to \mathbb{R}^{N_r}$, $H_s : \mathbb{R}^n \to \mathbb{R}^{M_s}$, $\Phi_r \in \Gamma_0(\mathbb{R}^{N_r})$ and nonempty, closed, convex subsets $C_s \subset \mathbb{R}^{M_s}$, where the epigraphical projection approach may be quite efficient, see [10], *iii)* restoring images with Poissonian+Gaussian noise, see [15, 16], and *iv)* finding numerically efficient methods to map the constraining parameter to the parameters of the corresponding penalizing functional.

References

1. F. J. Anscombe. The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, 35:246–254, 1948.
2. A. Y. Aravkin, J. V. Burkey, and M. P. Friedlander. Variational properties of value functions. *Preprint Univ. British Columbia*, 2012.
3. J. M. Bardsley and J. Goldes. Regularization parameter selection methods for ill-posed Poisson maximum likelihood estimation. *Inverse Problems*, 25(9):095005, 2009.
4. R. I. Bot and C. Hendrich. Convergence analysis for a primal-dual monotone + skew splitting algorithm with application to total variation minimization. *Preprint Univ. Chemnitz*, 2012.
5. M. Carlavan and L. Blanc-Feraud. Sparse Poisson noisy image deblurring. *IEEE Transactions on Image Processing*, 21(4):1834–1846, 2012.
6. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

7. C. Chaux, L. Blanc-Féraud, and J. Zerubia. Wavelet-based restoration methods: Application in 3d confocal microscopy images. In *Proc. SPIE Conf. Wavelets*, page 67010E, San Diego, 2007.
8. C. Chaux, J.-C. Pesquet, and N. Pustelnik. Nested iterative algorithms for convex constrained image recovery problems. *SIAM Journal on Imaging Science*, 2(2):730–762, 2009.
9. G. Cherchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu. A proximal approach for constrained cosparse modelling. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, 2012.
10. G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu. Epigraphical projection and proximal tools for solving constrained convex optimization problems - part I. *Preprint*, 2012.
11. P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering, (H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Editors)*, pages 185–212, New York, 2011. Springer-Verlag.
12. F.-X. Dupé, J. Fadili, and J.-L. Starck. A proximal iteration for deconvolving Poisson noisy images using sparse representations. *IEEE Transactions on Image Processing*, 18(2):310–321, 2009.
13. M. A. T. Figueiredo and J. M. Bioucas-Dias. Restoration of Poissonian images using alternating direction optimization. *IEEE Transactions on Image Processing*, 19(12):3133–3145, 2010.
14. M. Hanke–Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Teubner, Stuttgart, 2002.
15. A. Jezierska, E. Chouzenoux, J.-C. Pesquet, and H. Talbot. A primal-dual proximal splitting approach for restoring data corrupted with Poisson-Gaussian noise. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012)*, Kyoto, Japan, 2012.
16. J. Li, Z. Shen, R. Jin, and X. Zhang. A reweighted ℓ_2 method for image restoration with Poisson and mixed Poisson-Gaussian noise. *UCLA Preprint*, 2012.
17. M. Mikkitalo and A. Foi. Optimal inversion of the Anscombe transformation in low-count Poisson image denoising. *IEEE Transactions on Image Processing*, 20(1):99–109, 2011.
18. T. Pock, A. Chambolle, D. Cremers, and H. Bischof. A convex relaxation approach for computing minimal partitions. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 810–817, 2009.
19. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
20. S. Setzer, G. Steidl, and T. Teuber. Deblurring Poissonian images by split Bregman techniques. *Journal of Visual Communication and Image Representation*, 21(3):193–199, 2010.
21. T. Teuber, G. Steidl, and R.-H. Chan. Minimization and parameter estimation for seminorm regularization models with I-divergence constraints. *Preprint Univ. Kaiserslautern*, 2012.
22. B. C. Vu. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, 2012. Accepted.
23. R. Zanella, P. Boccacci, L. Zanni, and M. Bertero. Efficient gradient projection methods for edge-preserving removal of Poisson noise. *Inverse Problems*, 25(4):045010, 2009.