



# Epigraphical Projection for Solving Least Squares Anscombe Transformed Constrained Optimization Problems

Harizanov Stanislav, Jean-Christophe Pesquet, Gabriele Steidl

## ► To cite this version:

Harizanov Stanislav, Jean-Christophe Pesquet, Gabriele Steidl. Epigraphical Projection for Solving Least Squares Anscombe Transformed Constrained Optimization Problems. 2013. hal-00773246v2

**HAL Id: hal-00773246**

**<https://hal.science/hal-00773246v2>**

Submitted on 14 Feb 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Epigraphical Projection for Solving Least Squares Anscombe Transformed Constrained Optimization Problems

Stanislav Harizanov<sup>1</sup>, Jean-Christophe Pesquet<sup>2</sup>, and Gabriele Steidl<sup>1</sup>

<sup>1</sup> Department of Mathematics, University of Kaiserslautern, Germany

<sup>2</sup> Laboratoire d'Informatique Gaspard Monge, Université Paris-Est, France

**Abstract.** This paper deals with the restoration of images corrupted by a non-invertible or ill-conditioned linear transform and Poisson noise. Poisson data typically occur in imaging processes where the images are obtained by counting particles, e.g., photons, that hit the image support. By using the Anscombe transform, the Poisson noise can be approximated by an 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 observations. We use this information by considering an Anscombe transformed constrained model to restore the image. The advantage with respect to corresponding penalized approaches lies in the existence of a simple model for parameter estimation. We solve the constrained minimization problem by applying a primal-dual algorithm together with a projection onto the epigraph of a convex function related to the Anscombe transform. We show that this epigraphical projection can be efficiently computed by Newton's methods with an appropriate initialization. Numerical examples demonstrate the good performance of our approach, in particular, its close behaviour with respect 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

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

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  $\bar{u} \in [0, +\infty)^m$  from observations

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

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 [18] 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, 13] where one minimizes a data fidelity term

$$u \mapsto \|T(Hu) - T(f)\|_2^2 \quad (1)$$

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

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

where  $\langle \cdot, \cdot \rangle$  denotes the standard Euclidean inner product and  $\mathbf{1}_n$  denotes the vector consisting of  $n$  entries equal to 1 (see [14, 21]). 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

$$\|T(Hu) - T(f)\|_2^2 \leq \tau \quad (2)$$

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

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

where  $C$  is a nonempty closed convex subset of  $[0, +\infty)^m$ ,  $L \in \mathbb{R}^{q \times m}$ , and  $\Phi: \mathbb{R}^q \rightarrow (-\infty, +\infty]$  is a proper, lower-semicontinuous, convex function. A typical choice for  $C$  is the nonnegative orthant of  $\mathbb{R}^m$ . The classical Total Variation objective function [20] is obtained, as a special case, when  $\Phi$  is an  $\ell_{2,1}$  norm and  $L$  corresponds to a discrete gradient operator. Constrained models based on the  $I$ -divergence have been considered in [5, 22], 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 (3) 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 for a function related to constraint (2) which plays a 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 emphasizing the good approximation of the  $I$ -divergence constrained approach achieved by our Anscombe constrained model. Finally, a summary of our contribution and some 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

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

For a nonempty, closed, convex set  $C \subset \mathbb{R}^m$  we denote by  $\iota_C \in \Gamma_0(\mathbb{R}^m)$  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-space  $V_\tau := \{\zeta \in \mathbb{R}^n : \langle \mathbf{1}_n, \zeta \rangle \leq \tau\}$ . Using this notation and defining, for every  $i \in \{1, \dots, n\}$ ,

$$\varphi_i : [0, +\infty) \rightarrow [0, +\infty) : s \mapsto (2\sqrt{s} - (T(f))_i)^2,$$

problem (3) can be rewritten as

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

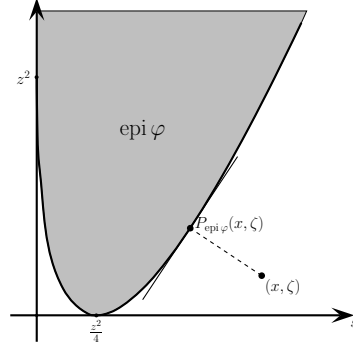
Now one can choose a primal-dual splitting algorithm as those proposed in [4, 6, 11, 12, 23] to solve this problem. One step in all these algorithms consists of the orthogonal projections onto the epigraphs of  $\varphi_i$  for all  $i \in \{1, \dots, n\}$  which is the topic of the next section.

### 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

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

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



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

**Proposition 1.** Let  $\varphi$  be defined by (5) 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

$$p: t \mapsto 17t^3 + 3zt^2 + (3z^2 - 16\zeta - 4x)t + z(z^2 - 4x). \quad (6)$$

**Proof.** The function  $\varphi$  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  $\lim_{s \rightarrow 0} \varphi'(s) = -\infty$ . Thus, if  $x \leq 0$  and  $\zeta \geq z^2$ , then  $P_{\text{epi } \varphi}(x, \zeta) = (0, \zeta)$ . In addition, if  $(x, \zeta) \in \text{epi } \varphi$ , then  $P_{\text{epi } \varphi}(x, \zeta) = (x, \zeta)$ .

We consider the remaining cases when  $(\max\{x, 0\}, \zeta) \notin \text{epi } \varphi$ . The tangent vector of the curve associated with the graph of  $\varphi$  reads  $(1, \varphi'(s))$ ,  $s > 0$ . The uniquely determined orthogonal projection  $(\hat{x}, \hat{\zeta}) := P_{\text{epi } \varphi}(x, \zeta)$  has to satisfy

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

which leads to

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

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

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

Conversely, if  $\hat{t} > -z$  is a root of the polynomial  $p$  in (6), then  $\hat{x} = \left(\frac{\hat{t}+z}{2}\right)^2$  fulfills (7). When  $x \geq z^2/4$  then also  $\hat{x} \geq z^2/4$  (see Fig. 1), thus we are interested in the restriction of  $\varphi$  to  $[z^2/4, +\infty)$  i.e., the nonnegative roots of  $p$ . The restriction of  $\varphi$  to  $[z^2/4, +\infty)$  is convex, monotonically increasing, and  $(x, \zeta) \notin \text{epi } \varphi$ . Hence, there is a unique point  $(\hat{x}, \hat{\zeta})$  on its graph that satisfies (7), 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  $\varphi$  to  $[0, z^2/4)$  i.e., the roots of  $p$  in the interval  $(-z, 0)$ . The restriction of  $\varphi$  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  $\hat{t} > -z$ , which completes the proof.  $\square$

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  $t_0$  is a very good starting point.

**Proposition 2.** *Let  $(\max\{x, 0\}, \zeta) \notin \text{epi } \varphi$  and  $t_0 := 2\sqrt{\max\{x, 0\}} - z$ . Let the polynomial  $p$  be defined by (6). Then the Newton method for finding a zero of  $p$  with initial value  $t_0$  converges (after a finite number of steps) monotonically to the root  $t_+$  if  $4x \geq z^2$ , resp.,  $t_-$  if  $4x < z^2$ .*

**Proof.** 1. First we show that

- i)  $p(t_0)p(0) \leq 0$ ,
- 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  $(0, \zeta) \notin \text{epi } \varphi$ , i.e.,  $z^2 > \zeta$ , we obtain:  $p(0) = z(z^2 - 4x) > 0$  and  $p(t_0) = p(-z) = 16z(\zeta - z^2) < 0$ , which proves i). Further, since  $p': t \mapsto 51t^2 + 6zt + 3z^2 - 4x - 16\zeta$ , 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  $(x, \zeta) \notin \text{epi } \varphi$ , 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., [15, 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  $p'' : t \mapsto 6(17t + z)$ ,  $p$  is convex on  $[-\frac{z}{17}, +\infty)$  and concave on  $(-\infty, -\frac{z}{17}]$ .

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 study 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

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

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

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

We shall now prove that  $p'(t) > 0$  for all  $t \in \mathbb{R}$ , so that  $p$  is strictly increasing. Since  $p'$  is a quadratic polynomial with minimum at  $-\frac{z}{17}$ , we have only to show that  $p'(-\frac{z}{17}) = 2(\frac{24}{17}z^2 - 8\zeta - 2x) > 0$ . Plugging in (8) and (9), we indeed obtain

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

The sequence  $(t_k)_{k \in \mathbb{N}}$  generated by Newton's algorithm is such that there exists  $k_0 \in \mathbb{N} \setminus \{0\}$  such that  $t_{k_0} \geq t_-$ . Otherwise,  $(t_k)_{k \in \mathbb{N}}$  would be an increasing sequence which would necessarily converge to  $t_-$  and there would exist  $k_1 \in \mathbb{N}$  such that  $p$  is convex over  $[t_{k_1}, +\infty[$ . Then, we would have  $p(t_{k_1+1}) \geq 0$ . Thus, after a finite number of steps, the algorithm arrives at  $t_{k_0} \geq t_-$  and we can apply Part 2 of the proof.

3.3 Let  $t_0 = 0$  which implies  $4z - t^2 = 0$ . Then, we get  $t_+ = t_0$ .  $\square$

## 4 Primal-Dual Algorithms

We can apply the projection onto the epigraph of  $\varphi$  in combination with any primal-dual algorithm proposed in [4, 6, 11, 12, 23] or an alternating direction method of multipliers. For example, we use here the **primal-dual hybrid gradient** algorithm from [6, 19] with an extrapolation (**m**odification) of the dual variable which will be designated by PDHGMp. Based on the following reformulation of (4),

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

this algorithm reads:

**Algorithm 1** (PDHGMp for solving the Anscombe constrained problem)

Initialization:  $u^{(0)}, \zeta^{(0)}, (p_j^{(0)})_{1 \leq j \leq 3} = (\bar{p}_j^{(0)})_{1 \leq j \leq 3}$ ,  $\theta \in (0, 1]$ ,  $(\rho, \sigma) \in (0, +\infty)^2$  with  $\rho\sigma < 1/\max\{1, \|H^*H + L^*L\|_2\}$

For  $k = 0, 1, \dots$  repeat until a stopping criterion is reached

1.  $u^{(k+1)} = P_C \left( u^{(k)} - \sigma\rho \left( H^* \bar{p}_1^{(k)} + L^* \bar{p}_2^{(k)} \right) \right)$
2.  $\zeta^{(k+1)} = P_{V_\tau} \left( \zeta^{(k)} - \sigma\rho \bar{p}_3^{(k)} \right)$
3.  $(v_{1,i}^{(k+1)}, \eta_i^{(k+1)}) = P_{\text{epi } \varphi_i} \left( p_{1,i}^{(k)} + (Hu^{(k+1)})_i + 3/8, p_{3,i}^{(k)} + \zeta_i^{(k+1)} \right)$ ,  $i = 1, \dots, n$
4.  $v_2^{(k+1)} = \text{prox}_{\sigma^{-1}\Phi}(p_2^{(k)} + Lu^{(k+1)})$
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.  $\bar{p}_j^{(k+1)} = p_j^{(k+1)} + \theta(p_j^{(k+1)} - p_j^{(k)})$ ,  $j = 1, 2, 3$ .

The projection in step 1 is quite simple if  $C$  is the nonnegative orthant of  $\mathbb{R}^m$ . Step 3 requires the epigraphical projections discussed in the previous section, Step 4 can be performed by coupled soft shrinkage with threshold  $\sigma^{-1}$  if we use the  $\ell_{2,1}$ -norm. The other steps can be computed in a straightforward way.



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

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

using a similar splitting to (10) but without the extra-variables  $\zeta$  and  $\eta$ :

**Algorithm 2** (PDHGMp for solving the  $I$ -divergence constrained problem)

Initialization:  $u^{(0)}, (p_j^{(0)})_{1 \leq j \leq 2} = (\bar{p}_j^{(0)})_{1 \leq j \leq 2}$ ,  $\theta \in (0, 1]$ ,  $(\rho, \sigma) \in (0, +\infty)^2$  with  $\rho\sigma < 1/\|H^*H + L^*L\|_2$

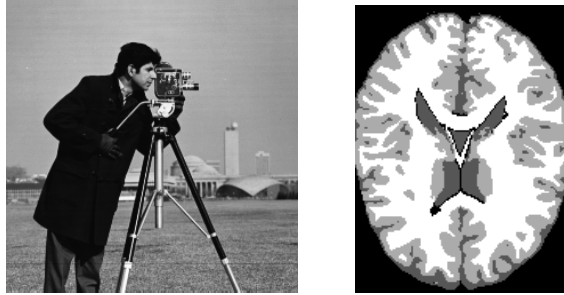
For  $k = 0, 1, \dots$  repeat until a stopping criteria is reached

1. Step 1 of Algorithm 1
2.  $\underset{v_1}{\text{minimize}} \quad \|v_1 - (p_1^{(k)} + Hu^{(k+1)})\|_2^2 \quad \text{subject to} \quad D(f, v_1) \leq \tau_I$  as in [22].
3. – 5. Steps 4. - 6. of Algorithm 1
6.  $\bar{p}_j^{(k+1)} = p_j^{(k+1)} + \theta(p_j^{(k+1)} - p_j^{(k)})$ ,  $j = 1, 2$ .

In [24] (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  $\bar{u}$ , namely 'cameraman' ( $256 \times 256$ ) and 'brain' ( $184 \times 140$ ), depicted in Fig. 2 and denoted by B1, resp. B2 in the following.



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

The images were blurred by a matrix  $H$  corresponding to a Gaussian kernel with standard deviation 1.3 and mirrored boundary (we have then  $m = n$ ). Their gray values are interpreted as photon counts in the range  $[0, \nu]$ , where

	$D(f, H\bar{u})/n$	$\ T(H\bar{u}) - T(f)\ _2^2/n$	PSNR	MAE
B1 <sub>100</sub>	0.5075	1.0086	20.58	66.41e-3
B1 <sub>600</sub>	0.5020	1.0034	23.38	41.59e-3
B1 <sub>1200</sub>	0.5018	1.0039	23.79	37.17e-3
B1 <sub>2000</sub>	0.4979	0.9960	23.97	34.87e-3
B1 <sub>3000</sub>	0.4994	0.9989	24.06	33.58e-3
B2 <sub>100</sub>	0.4954	0.8866	18.34	82.45e-3
B2 <sub>600</sub>	0.5131	1.0004	20.17	61.01e-3
B2 <sub>1200</sub>	0.5122	1.0178	20.37	57.90e-3
B2 <sub>2000</sub>	0.5063	1.0085	20.49	56.02e-3
B2 <sub>3000</sub>	0.4956	0.9899	20.52	55.16e-3

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

$\nu$  is the intensity of the image. We tested  $\nu = 100, 600, 1200, 2000, 3000$  and denoted the blurred, noisy images by B1 $_{\nu}$  and B2 $_{\nu}$ . In order to synthetically add 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  $\text{PSNR} = 10 \log_{10} \frac{|\max \bar{u} - \min \bar{u}|^2}{\frac{1}{n} \|u - \bar{u}\|_2^2}$ , and  $\text{MAE} = \frac{1}{n\nu} \|\bar{u} - u\|_1$ . The 'true' constraints between the blurred, noisy image  $f$  and the original image  $\bar{u}$  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  $D(f, H\bar{u})$  and  $\|T(H\bar{u}) - T(f)\|_2^2$ .

We computed a minimizer of our functional (3) with  $C = [0, +\infty)^n$ , the  $\ell_{2,1}$ -norm for  $\Phi$ , and the discrete gradient operator  $L$  ( $q = 2n$ ) by using Algorithm 1 with  $\tau = n$ . We compared the result with the  $I$ -divergence constrained approach (11) and Algorithm 2 with  $\tau_I = n/2$ . The parameters  $\sigma$  and  $\rho$  appearing in PDHGMp (in this setting convergence is theoretically guaranteed for  $\sigma\rho < 1/9$ ) 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. Furthermore, we have chosen  $\theta = 1$ . Fig. 3 shows the restoration results for B1<sub>1200</sub> and Fig. 4 for B2<sub>1200</sub>.



**Fig. 3.** Result for the 'cameraman' image B1<sub>1200</sub> 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 the one recovered 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.

image	$\sigma$	$\rho$	$\ T(H\bar{u}) - T(f)\ _2^2 - n$	$D(f, Hu) - n/2$	TV-norm	PSNR	MAE
B1 <sub>100</sub>	0.09	1.225	2.2837	-	1.2217e+5	24.28	30.76e-3
	0.12		-	2.5643	1.2646e+5	24.39	30.36e-3
B1 <sub>600</sub>	0.0599	2.8	-1.5323	-	8.9230e+5	25.58	25.69e-3
			-	0.0049	8.9402e+5	25.59	25.67e-3
B1 <sub>1200</sub>	0.042	3	9.0207	-	1.9190e+6	26.08	24.16e-3
			-	0.045	1.9198e+6	26.08	24.15e-3
B1 <sub>2000</sub>	0.027	3.03	0.2634	-	3.3007e+6	26.35	23.30e-3
			-	-0.1911	3.3017e+6	26.36	23.30e-3
B1 <sub>3000</sub>	0.0329	4.001	-0.3559	-	5.1667e+6	26.64	22.50e-3
			-	0.2767	5.1673e+6	26.64	22.49e-3
B2 <sub>100</sub>	0.55	0.25	0.9730	-	0.9349e+5	19.91	59.78e-3
			-	0.001	1.0459e+5	20.39	53.35e-3
B2 <sub>600</sub>	0.040 0.050	3.04	-2.7434	-	7.4158e+5	21.81	40.05e-3
			-	9.0129	7.5693e+5	21.91	39.28e-3
B2 <sub>1200</sub>	0.034 0.042	3.97	0.0079	-	1.5598e+6	22.33	36.26e-3
			-	4.2683	1.5673e+6	22.36	36.11e-3
B2 <sub>2000</sub>	0.021 0.041	4.108	0.7789	-	2.6780e+6	22.72	33.53e-3
			-	0.8585	2.6885e+6	22.73	33.48e-3
B2 <sub>3000</sub>	0.0182 0.0282	5.413	-0.5284	-	4.0565e+6	22.93	32.08e-3
			-	1.5536	4.0603e+6	22.94	32.03e-3

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



**Fig. 4.** Result for the 'brain' image B2<sub>1200</sub> 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 the one recovered 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.

Table 2 summarizes the results for the different intensities. As expected, we observe that the outcomes of the two algorithms are very similar. More precisely, if  $u_A$ , resp.  $u_I$  denotes the output of the restoration procedure with Anscombe, resp.  $I$ -divergence constraints, then we get for image B1 that  $\|u_A - u_I\|_2/(\nu\sqrt{n})$  ranges from 0.004 to 1.74e-4 and  $\max |u_A - u_I|/\nu$  from 0.0612 to 0.0031 for the different noise levels.

Finally, Table 3 compares Algorithms 1 and 2 for different constraints  $\tau$  and  $\tau_I$  and a central part of the cameraman of size  $130 \times 130$  with  $\nu = 3000$  after 1000 iterations.

## 6 Summary and Conclusions

We have considered a constrained restoration model for images corrupted by a linear transform and Poisson noise by making use of the Anscombe transform.

scale	$\sigma$	$\rho$	$\ T(H\bar{u}) - T(f)\ _2^2 - n$	$D(f, Hu) - n/2$	TV-norm	PSNR	MAE
0.8	0.2	0.975	-0.0813	-	2.2400e+6	26.51	22.03e-3
			-	-0.0585	2.2416e+6	26.52	22.02e-3
1.0	0.0269	4	-2.4868	-	1.7071e+6	25.60	21.20e-3
			-	-0.0425	1.7073e+6	25.60	21.19e-3
1.0095 (optimal)	0.0239	4.002	-0.0616	-	1.6964e+6	25.56	21.29e-3
			-	-1e-4	1.6914e+6	25.54	21.33e-3
1.2	0.011	8	1.1405	-	1.5563e+6	24.93	22.96e-3
			-	0.0611	1.5565e+6	24.93	22.96e-3
2	0.004	30.065	-1.7501	-	1.3388e+6	23.70	27.02e-3
			-	-0.0257	1.3391e+6	23.71	27.02e-3

**Table 3.** Results of Algorithms 1 and 2 on a part of B1<sub>3000</sub> 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.

In contrast with penalized approaches, a main advantage of the proposed one is that it makes it possible to employ a simple estimate for the model parameter. We have provided proximal algorithms to find a minimizer of the model, which are based on epigraphical projections, and we have shown that the performance is similar to a recently introduced  $I$ -divergence constrained model. Future research directions include 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 for which the epigraphical projection approach may be quite efficient, see [10], *iii*) restoring images with Poisson+Gauss noise, see [16, 17], and *iv*) finding numerically efficient methods to map the constraint bound to the parameter of the corresponding penalized 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. Carlván and L. Blanc-Féraud. 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. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu. A proximal approach for constrained cospase 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. P. L. Combettes and J.-C. Pesquet. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *newblock Set-Valued and Variational Analysis*, 20(2):307–330, 2012.
13. 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.
14. 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.
15. M. Hanke-Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Teubner, Stuttgart, 2002.
16. 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.
17. J. Li, Z. Shen, R. Jin, and X. Zhang. A reweighted  $\ell_2$  method for image restoration with Poisson and mixed Poisson-Gaussian noise. *UCLA Preprint*, 2012.
18. 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.
19. 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.
20. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
21. 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.
22. T. Teuber, G. Steidl, and R.-H. Chan. Minimization and parameter estimation for seminorm regularization models with  $I$ -divergence constraints. *Preprint Univ. Kaiserslautern*, 2012.
23. B. C. Vu. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, 2012. Accepted.
24. 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.