A Convex Programming Algorithm for Noisy Discrete Tomography

T. D. Capricelli and P. L. Combettes

Summary. A convex programming approach to discrete tomographic image reconstruction in noisy environments is proposed. Conventional constraints are mixed with noise-based constraints on the sinogram and a binarity-promoting total variation constraint. The noise-based constraints are modeled as confidence regions that are constructed under a Poisson noise assumption. A convex objective is then minimized over the resulting feasibility set via a parallel block-iterative method. Applications to binary tomographic reconstruction are demonstrated.

1 Introduction

The tomographic reconstruction problem is to estimate a multidimensional signal $\overline{x}$ in a Hilbert space $\mathcal{H}$ from lower dimensional measurements of its line integrals. In computerized tomography, the signals are discretized over a bounded domain and $\mathcal{H}$ is the usual Euclidean space $\mathbb{R}^N$ (using standard lexicographic ordering, an N-voxel signal is represented as an N-point vector). A popular approach for solving this problem is to pose it as a convex feasibility problem of the form

$$\text{Find } \overline{x} \in \bigcap_{i=1}^{m} S_i \,, \tag{1}$$

where $(S_i)_{1\leq i\leq m}$ are closed convex sets in $\mathbb{R}^N$ arising from prior knowledge (bounds, support information, spectral information, information about the noise corrupting the measurements, etc) and the discrete line integral (line sum) measurements [4, 6, 21, 31]. This set theoretic approach to tomographic image reconstruction goes back to [18]; further developments can be found in [24, 25, 30, 32, 35]. In some instances it may be justified on physical grounds to seek a feasible image which is optimal in some sense. The problem then assumes the form

$$\text{Find } \overline{x} \in S = \bigcap_{i=1}^{m} S_i \text{ such that } \varphi(\overline{x}) = \inf \varphi(S) \,, \tag{2}$$

where $\varphi\colon \mathbb{R}^N \to]-\infty,+\infty]$ is a convex function. The advantage of this convex programming formulation is to allow for the incorporation of a wide range of prior information in the reconstruction process and, at the same time, to benefit from the availability of powerful algorithms; see [1, 4, 7, 8, 9] and the references therein.

In discrete tomography, the range of the signal $\overline{x}$ to be reconstructed is known to be a finite set, for instance the set $\{0,1\}$ in binary tomography. This additional information is of paramount importance and it has profound consequences on the theoretical and practical aspects of the reconstruction problem [22]. As underlined in [22], classical computer tomography algorithms do not perform well in the presence of few (say 10 or less) views (we do not employ the standard term "projection," as it will be reserved to describe the best metric approximation from a convex set). Consequently, they are not directly applicable in discrete tomography, where such low numbers of views are common. Furthermore, classical algorithms do not exploit nor enforce the discrete nature of the original signal.

In this paper, we propose a convex programming approach to the discrete tomography problem in noisy environments. In recent years, many papers have been devoted to the theoretical and numerical investigations of discrete reconstruction problems; see [15, 17, 19, 22, 23, 29, 34] and the references therein. The novelty of our work is to propose a convex programming formulation of this problem that explicitly takes into account the presence of noise in the measured data, and to provide a numerical method to solve it. Our formulation is of the form (2) and our algorithm is based on the block-iterative methods recently developed in [8, 9]. While our approach is applicable to general discrete problems, we shall focus on the case of binary images for simplicity. Since the set of binary images is nonconvex, our first task will be to find pertinent convex constraints that will promote the binary nature of the image: total variation will be used for this purpose. Thus, the image produced by the algorithm will be relatively close to being binary, which will minimize the number of errors incurred by the final binarization step. Other constraints will exploit standard information (bounds, support) as well as information about the data model and the noise.

The remainder of the paper is organized as follows. In Section 2, we review the parallel block-iterative algorithm that will be employed to solve the convex program (2). In Section 3, we address the construction of constraints for noisy binary tomography. The new constraints are confidence regions that are based on statistical attributes of the noise perturbing the sinogram, and a total variation constraint aiming at promoting the binary nature of the reconstructed image. A Poisson noise model is assumed in our statistical analysis of the confidence regions. In Section 4, we describe several applications of this convex programming framework to binary image reconstruction in the presence of noisy measurements. A few remarks conclude the paper in Section 5.

2 Proposed algorithm

Throughout the paper, the signal space containing the original image $\overline{x}$ is the standard Euclidean space $\mathbb{R}^N$, with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d. The distance to a nonempty set $C \subset \mathbb{R}^N$ is $d_C(x) = \inf \|x - C\|$. If $C \subset \mathbb{R}^N$ is nonempty, closed, and convex then, for every $x \in \mathbb{R}^N$, there is a unique point $P_C x \in C$ such that $\|x - P_C x\| = d_C(x)$; $P_C x$ is called the projection of x onto C.

To solve (2), we use the parallel block iterative algorithm described in [9], where the framework of [8] was adapted to problems with quadratic objective functions. Although more general strictly convex objectives φ can be used [8], we restrict our attention to quadratic functions in this paper, as they lead to particularly simple implementations.

Let

$$\varphi \colon \mathbb{R}^N \to \mathbb{R} \colon x \mapsto \langle R(x - r) \mid x - r \rangle \ , \tag{3}$$

where $r \in \mathbb{R}^N$ and $R \in \mathbb{R}^{N \times N}$ is a positive definite symmetric matrix. In addition, we suppose (without loss of generality) that the closed convex constraint sets $(S_i)_{1 \leq i \leq m}$ in (2) assume the form

$$(\forall i \in \{1, \ldots, m\}) \ \ S_i = \big\{ x \in \mathbb{R}^N \mid f_i(x) \leq \delta_i \big\} \ , \tag{4}$$

where $(f_i)_{1 \leq i \leq m}$ are convex functions from $\mathbb{R}^N$ to $\mathbb{R}$ and $(\delta_i)_{1 \leq i \leq m} \in \mathbb{R}^m$ are such that $S = \bigcap_{i=1}^m S_i \neq \emptyset$. Recall that, under these assumptions, for every $x \in \mathbb{R}^N$, each f_i admits at least one subgradient at x, i.e., a point $g_i \in \mathbb{R}^N$ such that [27]

$$(\forall y \in \mathbb{R}^N) \ \ \langle y - x \mid g_i \rangle + f_i(x) \leq f_i(y) \ . \tag{5}$$

For instance, if C is a nonempty closed convex subset of $\mathbb{R}^N$ and $x \in \mathbb{R}^N \smallsetminus C$, then

$$\partial d_C(x) = (x - P_C x)/d_C(x). \tag{6}$$

The set of all subgradients of f_i at x is the subdifferential of f_i at x and is denoted by $\partial f_i(x)$; if f_i is differentiable at x, then $\partial f_i(x) = \{\nabla f_i(x)\}$. Moreover, the subgradient projection $G_i x$ of x onto S_i is obtained by selecting an arbitrary $g_i \in \partial f_i(x)$ and setting [7]

$$G_i x = \begin{cases} x + \dfrac{\delta_i - f_i(x)}{\|g_i\|^2} g_i \ , & \text{if } \ f_i(x) > \delta_i \ ; \\ x \ , & \text{if } \ f_i(x) \leq \delta_i \ . \end{cases} \tag{7}$$

The algorithm proposed in [9] to solve (2)–(3) constructs a sequence $(x_n)_{n \in \mathbb{N}}$ of approximate solutions as follows.

1. Fix $\varepsilon \in \left]0, 1/m\right[$. Set $x_0 = r$ and $n = 0$.
2. Take a nonempty index set $I_n \subset \{1, \ldots, m\}$.
3. Set $z_n = x_n - \lambda_n R^{-1} u_n$, where:
   - for every $i \in I_n$, $p_{i,n} = G_i x_n$;

- the weights $(\omega_{i,n})_{i\in I_n}$ lie in $[\varepsilon, 1]$ and $\sum_{i\in I_n}\omega_{i,n} = 1$;
- $u_n = x_n - \sum_{i\in I_n}\omega_{i,n}p_{i,n}$;
- $\lambda_n \in [\varepsilon L_n, L_n]$, where

$$L_n = \begin{cases} \dfrac{\displaystyle\sum_{i\in I_n}\omega_{i,n}\|p_{i,n} - x_n\|^2}{\langle R^{-1}u_n \mid u_n\rangle}\ , & \text{if } \max\limits_{i\in I_n}\big(f_i(x_n) - \delta_i\big) > 0\ ; \\ 1/\|R^{-1}\|\ , & \text{otherwise}\ . \end{cases}$$

4. Set $\pi_n = \langle R(x_0 - x_n) \mid x_n - z_n\rangle$, $\mu_n = \langle R(x_0 - x_n) \mid x_0 - x_n\rangle$, $\nu_n = \langle R(x_n - z_n) \mid x_n - z_n\rangle$, and $\rho_n = \mu_n\nu_n - \pi_n^2$.
5. Set

$$x_{n+1} = \begin{cases} z_n\ , & \text{if } \rho_n = 0\ ,\ \pi_n \geq 0\ ; \\ x_0 + \left(1 + \dfrac{\pi_n}{\nu_n}\right)(z_n - x_n)\ , & \text{if } \rho_n > 0,\ \pi_n\nu_n \geq \rho_n\ ; \\ x_n + \dfrac{\nu_n}{\rho_n}\big(\pi_n(x_0 - x_n) + \mu_n(z_n - x_n)\big)\ , & \text{if } \rho_n > 0,\ \pi_n\nu_n < \rho_n\ . \end{cases}$$

6. Set $n = n + 1$ and go to Step 2.

The following convergence result is an application of [9, Theorem 16].

Theorem 1. *Suppose that there exists a strictly positive integer J such that*

$$(\forall n \in \mathbb{N}) \quad \bigcup_{k=n}^{n+J-1} I_k = \{1, \ldots, m\}. \tag{8}$$

Then every sequence $(x_n)_{n\in\mathbb{N}}$ generated by the above algorithm converges to the unique solution to (2)–(3).

Some comments about this result and the algorithm are in order.

- Condition (8) is satisfied in particular when $I_n \equiv \{1, \ldots, m\}$, i.e., when all the sets are activated at each iteration. In general, (8) allows for variable blocks of sets to be used, which provides great flexibility in terms of parallel implementation (see [7] for examples). More details on the importance of block-processing for task scheduling on parallel architectures will be found in [4]. Further flexibility is provided by the fact that the relaxations and the weights can vary at each iteration.
- The algorithm activates the constraints by means of subgradient projections rather than exact projections. The former are significantly easier to implement than the latter, as they require only the computation of subgradients (gradients in the differentiable case). Analytically complex constraints can therefore be incorporated in the recovery algorithm and processed at low cost.

- The parameter L_n is always at least equal to $1/\|R^{-1}\|$ [9, Proposition 12] and it can attain large values. Choosing λ_n large (e.g., equal to L_n) usually yields faster convergence.

In [9], it was shown that, in order to reduce the computational load of the method, an iteration of the algorithm could be implemented as follows (only one application of the matrices R and R^{-1} is required).

1. For every $i \in I_n$, set $a_i = -f_i(x_n)g_i/\|g_i\|^2$, where $g_i \in \partial f_i(x_n)$, if $f_i(x_n) > \delta_i$; $a_i = 0$ otherwise.
2. Choose weights $(\omega_i)_{i\in I_n}$ in $[\varepsilon, 1]$ adding up to 1. Set $v = \sum_{i\in I_n} \omega_i a_i$ and $L = \sum_{i\in I_n} \omega_i \|a_i\|^2$.
3. If $L = 0$, set $x_{n+1} = x_n$ and exit iteration. Otherwise, set $b = x_0 - x_n$, $c = Rb$, $d = R^{-1}v$, and $L = L/\langle d \mid v \rangle$.
4. Choose $\lambda \in [\varepsilon L, L]$ and set $d = \lambda d$.
5. Set $\pi = -\langle c \mid d \rangle$, $\mu = \langle b \mid c \rangle$, $\nu = \lambda \langle d \mid v \rangle$, and $\rho = \mu\nu - \pi^2$.
6. Set $x_{n+1} = \begin{cases} x_n + d\,, & \text{if } \rho = 0 \text{ and } \pi \geq 0\,; \\ x_0 + (1 + \pi/\nu)d\,, & \text{if } \rho > 0 \text{ and } \pi\nu \geq \rho\,; \\ x_n + \dfrac{\nu}{\rho}(\pi b + \mu d)\,, & \text{if } \rho > 0 \text{ and } \pi\nu < \rho\,. \end{cases}$

Remark 1. If the projector P_i onto S_i is easy to implement, one can set $f_i = d_{S_i}$ since S_i can certainly be described as the level set

$$S_i = \big\{x \in \mathbb{R}^N \mid d_{S_i}(x) \leq 0\big\}. \tag{9}$$

In this case, it follows at once from (6) and (7) that $G_i = P_i$ (the subgradient projector reduces to the usual projector) and, moreover, that $a_i = P_i x_n - x_n$ at Step 1 of the algorithm.

3 Construction of closed convex constraint sets

3.1 Data model

The sinogram is the image under the Radon transform of the original image $\overline{x}$. The portion of the sinogram corresponding to a given observation angle θ will be referred to as a *view*. The observed data consists of q noisy views $(z_i)_{1\leq i\leq q}$ at angles $(\theta_i)_{1\leq i\leq q}$. For every $i \in \{1, \ldots, q\}$, we let L_i be the restriction of the Radon transform for a fixed angle θ_i. In other words, the ith measurement is

$$z_i = L_i\overline{x} + w_i\,, \tag{10}$$

where w_i is the noise vector corrupting the observation. Each view z_i is a one-dimensional signal of M points and will be represented by a vector $z_i = [\zeta_{i,k}]^\top_{1\leq k\leq M}$ in $\mathbb{R}^M$; L_i is therefore a matrix in $\mathbb{R}^{M\times N}$. Finally, we denote by $\underline{1}$ the vector $[1, \ldots, 1]^\top$ in $\mathbb{R}^N$ or $\mathbb{R}^M$.

3.2 Standard constraints

Further details on standard constraint sets can be found in [6, 31].

Range

The first standard constraint arises from the fact that pixel values are nonnegative and have known maximal value. After normalization, the corresponding set is

$$[0,1]^N \ . \tag{11}$$

In the context of binary tomography, this is simply the convex hull of the set $\{0,1\}^N$.

Support

A second common piece of a priori information in tomography is the knowledge of the support K of the body under investigation. The associated set is

$$\big\{x \in \mathbb{R}^N \mid x1_K = x\big\} \ , \tag{12}$$

where 1_K is the characteristic function of K and where the product $x1_K$ is taken componentwise. The projector onto this set is $x \mapsto x1_K$ [6].

Pixel sum

Let μ be the sum of the pixel values in the original image, i.e,

$$\mu = \langle \overline{x} \mid \underline{1} \rangle \ . \tag{13}$$

The knowledge of μ leads to the set

$$\big\{x \in \mathbb{R}^N \mid \langle x \mid \underline{1} \rangle = \mu\big\} \ . \tag{14}$$

Since μ is never known exactly, this set should be relaxed into the hyperslab

$$\big\{x \in \mathbb{R}^N \mid \mu^- \leq \langle x \mid \underline{1} \rangle \leq \mu^+\big\} \ , \tag{15}$$

where $[\mu^-, \mu^+]$ is a confidence interval. The projector onto this set is [6]

$$x \mapsto \begin{cases} x + \dfrac{\mu^- - \langle x \mid \underline{1} \rangle}{N} \underline{1} \ , & \text{if } \ \langle x \mid \underline{1} \rangle < \mu^- \ ; \\ x + \dfrac{\mu^+ - \langle x \mid \underline{1} \rangle}{N} \underline{1} \ , & \text{if } \ \langle x \mid \underline{1} \rangle > \mu^+ \ ; \\ x \ , & \text{otherwise} \ . \end{cases} \tag{16}$$

The values of μ^- and μ^+ depend on prior information about the experimental setup and about the noise. An example is provided in Section 3.5.

3.3 Binarity-promoting constraint

The constrained image recovery method presented in Section 2 is limited to problems with convex constraints. As a result, since the set of binary images is nonconvex, the binarity constraint cannot be enforced directly in such a framework. Furthermore, the convex set (11) will not properly enforce binarity and we must find a more effective means to promote binarity through a convex constraint.

In recent years, total variation has emerged as an effective tool to recover piecewise-smooth images in variational methods. This approach was initiated in [28] in the context of denoising problems and has since been used in various image recovery problems. Recently, it has also been applied in certain variational computerized tomography problems [2, 36]. In all of these approaches, total variation appears in the objective of a minimization problem. Such nondifferentiable problems are not easy to solve and they offer very limited potential in terms of incorporating constraints [5, 11].

In [12], it was observed that, in many problems, the total variation $\mathrm{tv}(\overline{x})$ of the original image, which measures the amount of oscillations, does not exceed some known bound τ. This constraint, which is associated with the set

$$\left\{x \in \mathbb{R}^N \mid \mathrm{tv}(x) \leq \tau\right\}, \tag{17}$$

appears to be particularly relevant in binary (and more generally in discrete) tomography, as it attenuates the oscillating components in the image, thereby forcing the creation of flat areas and promoting binarity. An important issue is of course the availability of the parameter τ in (17). In this respect, binary tomography places us on favorable grounds. Indeed, since the total variation of a binary image is simply the length of its contours (i.e., the sum of the perimeters of the elementary shapes), it can be estimated with good accuracy in certain typical problems from prior experiments or by sampling databases [12].

Numerically, the total variation of a discrete image $x = [\xi_{i,j}] \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$ is computed as

$$\begin{aligned}\mathrm{tv}(x) = &\sum_{i=1}^{\sqrt{N}-1}\sum_{j=1}^{\sqrt{N}-1}\sqrt{|\xi_{i+1,j}-\xi_{i,j}|^2+|\xi_{i,j+1}-\xi_{i,j}|^2}\\ &+\sum_{i=1}^{\sqrt{N}-1}|\xi_{i+1,\sqrt{N}}-\xi_{i,\sqrt{N}}|+\sum_{j=1}^{\sqrt{N}-1}|\xi_{\sqrt{N},j+1}-\xi_{\sqrt{N},j}|\,.\end{aligned} \tag{18}$$

The subgradient projector onto the set (17) can be found in [12].

3.4 Constraints on the residual views

Using an approach developed in [13] and [33], a wide range of statistical constraints modeled by closed convex sets can be formed for each of the q views.

Indeed, (10) gives

$$z_i - L_i\overline{x} = w_i \ . \tag{19}$$

Hence, the residual signal $z_i - L_i x$ associated with an estimate x of $\overline{x}$ should be constrained to be statistically consistent with every known attribute (e.g., moment, periodogram, etc) of the noise. Consistency is usually enforced via some statistical confidence bound (see [10] for an analysis of the global confidence level in terms of confidence levels of each set). We now provide examples of such constraint sets.

Amplitude

Let us denote by $(e_k)_{1\leq k\leq M}$ the canonical basis of $\mathbb{R}^M$ (recall that M is the length of each view z_i). If no noise were present, (10) would confine $\overline{x}$ to the intersection of the Mq hyperplanes

$$\big\{x \in \mathbb{R}^N \mid \langle L_i x \mid e_k \rangle = \langle z_i \mid e_k \rangle\big\} \ . \tag{20}$$

These sets were used in the early ART reconstruction technique [18]. In the presence of noise, the hyperplanes must be replaced by hyperslabs of the form [20]

$$\big\{x \in \mathbb{R}^N \mid |\langle L_i x - z_i \mid e_k \rangle| \leq \alpha_{i,k}\big\} \ . \tag{21}$$

The confidence interval $[-\alpha_{i,k}, \alpha_{i,k}]$ can be determined from the distribution of the random variable $\langle w_i \mid e_k \rangle$. Such distributions can be available in certain problems, e.g., [6, 14, 33]. However, in the present problem, the noise is best modeled by a signal-dependent process which makes it impossible to obtain reliable bounds. We shall therefore not use these sets in our experiments.

ℓ^p norms

Let $p \in [1, +\infty[$ and $i \in \{1, \ldots, q\}$. Suppose that an estimate $\delta_{i,p}^{1/p}$ of the ℓ^p norm of the noise vector w_i is available from physical considerations or past experience with tomographic reconstructions of similar objects. Then one can construct the set [13]

$$\big\{x \in \mathbb{R}^N \mid \|L_i x - z_i\|_p^p \leq \delta_{i,p}\big\} \ . \tag{22}$$

The projector onto this set has no simple closed form, but the subgradient projector can be obtained as follows. Let us denote by $(L_{i,k})_{1\leq k\leq M}$ the M rows of L_i and by $L_{i,k,l}$ an entry of the matrix L_i, i.e.,

$$L_i = \Big[L_{i,k,l}\Big]_{\substack{1\leq k\leq M\\ 1\leq l\leq N}} \ . \tag{23}$$

Then, by elementary subdifferential calculus [27],

$$\partial \|L_i x - z_i\|_p^p = \sum_{k=1}^{M} \partial |\langle L_{i,k} \mid x \rangle - \zeta_{i,k}|^p \; . \tag{24}$$

Hence, if $p > 1$, the unique subgradient of $x \mapsto \|L_i x - z_i\|_p^p$ is

$$g = p \left(\sum_{k=1}^{M} L_{i,k,l} |\langle L_{i,k} \mid x \rangle - \zeta_{i,k}|^{p-1} \operatorname{sign}(\langle L_{i,k} \mid x \rangle - \zeta_{i,k}) \right)_{1 \leq l \leq N} \; . \tag{25}$$

Now suppose that $p = 1$. Recall that

$$\operatorname{sign} \colon \xi \mapsto \begin{cases} -1 \; , & \text{if } \xi < 0 \; ; \\ 0 \; , & \text{if } \xi = 0 \; ; \\ +1 \; , & \text{if } \xi > 0 \; , \end{cases} \tag{26}$$

is a selection of the subdifferential of $\xi \mapsto |\xi|$, which is given by

$$(\forall \, \xi \in \mathbb{R}) \quad \partial |\xi| = \begin{cases} \{-1\} \; , & \text{if } \xi < 0 \; ; \\ [-1, 1] \; , & \text{if } \xi = 0 \; ; \\ \{+1\} \; , & \text{if } \xi > 0 \; . \end{cases} \tag{27}$$

Therefore, a subgradient of $x \mapsto \|L_i x - z_i\|_1$ is

$$g = \left(\sum_{k=1}^{M} L_{i,k,l} \operatorname{sign}(\langle L_{i,k} \mid x \rangle - \zeta_{i,k}) \right)_{1 \leq l \leq N} \; . \tag{28}$$

One can then compute the subgradient projection (7) via (25) and (28).

Energy

The case $p = 2$ corresponds to the set

$$\left\{ x \in \mathbb{R}^N \mid \|L_i x - z_i\|^2 \leq \delta_{i,2} \right\} . \tag{29}$$

In this case, (25) reduces to $g = 2L_i^\top (L_i x - z_i)$. As discussed in [7, 33], the projection of an image x onto this set requires an iterative procedure, while the subgradient projector is given explicitly by (7) as

$$x \mapsto \begin{cases} x + \dfrac{\delta_{i,2} - \|L_i x - z_i\|^2}{2 \|L_i^\top (L_i x - z_i)\|^2} L_i^\top (L_i x - z_i) \; , & \text{if } \|L_i x - z_i\|^2 > \delta_{i,2} \; ; \\ x \; , & \text{if } \|L_i x - z_i\|^2 \leq \delta_{i,2} \; . \end{cases} \tag{30}$$

3.5 Bound estimation in the case of Poisson noise

In this section, we address the problem of computing the parameters μ^- and μ^+ in (15), and $\delta_{i,2}$ in (29).

Noise modeling in computerized tomography is a research topic in its own right, and it is beyond the scope of the present paper to attempt to provide a precise model for the various complex underlying physical phenomena. In [3], a simple additive Gaussian noise model was considered. Since many data collection processes in discrete tomography are counting processes (e.g., counting the number of atoms in a structure along a certain direction), we adopt here a Poisson noise model. More specifically, we assume that the observation z_i in (10) is a realization of a random vector

$$Z_i = \big[Z_{i,k}\big]_{1\leq k\leq M}^{\top} , \tag{31}$$

the components of which are independent Poisson variables with means $(\lambda_{i,k})_{1\leq k\leq M}$. It is also assumed that the random vectors $(Z_i)_{1\leq i\leq q}$ are independent. Now, let Λ_i be the mean of Z_i. Then

$$\Lambda_i = \mathsf{E}Z_i = [\lambda_{i,k}]_{1\leq k\leq M}^{\top} = L_i\overline{x} . \tag{32}$$

Bound on the view-sums

The purpose in this section is to determine the parameters μ^- and μ^+ in (15). A property of the discrete Radon transform is that it preserves pixel sums in the sense that

$$(\forall i \in \{1,\dots,q\})\ \ \mu = \langle \overline{x} \mid \underline{1}\rangle = \langle L_i\overline{x} \mid \underline{1}\rangle \ . \tag{33}$$

Since $\langle Z_i \mid \underline{1}\rangle$ is the sum of M independent Poisson variables, it is also a Poisson variable, with mean $\langle \Lambda_i \mid \underline{1}\rangle = \mu$ and variance $\mathsf{var}\,\langle Z_i \mid \underline{1}\rangle = \mu$. The parameter μ can be approximated by the sample mean of the q views, i.e.,

$$\gamma = \frac{1}{q}\sum_{i=1}^{q} \langle z_i \mid \underline{1}\rangle \ . \tag{34}$$

The associated statistical estimator is

$$\Gamma = \frac{1}{q}\sum_{i=1}^{q} \langle Z_i \mid \underline{1}\rangle \ , \tag{35}$$

with mean

$$\mathsf{E}\Gamma = \frac{1}{q}\sum_{i=1}^{q} \mathsf{E}\,\langle Z_i \mid \underline{1}\rangle = \mu \tag{36}$$

and variance

$$\mathsf{var}\Gamma = \frac{1}{q^2}\sum_{i=1}^{q}\mathsf{var}\,\langle Z_i \mid \underline{1}\rangle = \frac{\mu}{q}\ . \tag{37}$$

We look for a confidence interval of the form

$$\left[\ \mathsf{E}\Gamma - \beta_1\sqrt{\mathsf{var}\Gamma},\ \mathsf{E}\Gamma + \beta_1\sqrt{\mathsf{var}\Gamma}\ \right] = \left[\ \mu - \beta_1\sqrt{\mu/q},\ \mu + \beta_1\sqrt{\mu/q}\ \right]\ , \tag{38}$$

for some $\beta_1 > 0$. Upon approximating μ by the observed sample mean γ, the confidence interval becomes $\left[\ \gamma - \beta_1\sqrt{\gamma/q},\ \gamma + \beta_1\sqrt{\gamma/q}\ \right]$. Monte Carlo experiments show that $\beta_1 = 2.3$ gives 98% of the realizations within this interval. Thus, with a confidence level of $\mathsf{c}_1 = 98\%$, we can use the values

$$\mu^- = \gamma - 2.3\sqrt{\gamma/q} \quad \text{and} \quad \mu^+ = \gamma + 2.3\sqrt{\gamma/q} \tag{39}$$

in (15).

Bound on the residual energy

Let $i \in \{1,\dots,q\}$. Our purpose is to determine the parameter $\delta_{i,2}$ in (29). The mean square residual error is

$$\mathsf{E}\|Z_i - L_i\overline{x}\|^2 = \mathsf{E}\|Z_i - \Lambda_i\|^2 = \sum_{k=1}^{M}\mathsf{var}Z_{i,k} = \sum_{k=1}^{M}\lambda_{i,k} = \|\Lambda_i\|_1\ . \tag{40}$$

To compute $\mathsf{E}\|Z_i - L_i\overline{x}\|^4$, we need the moments of $Z_{i,k}$ up to order 4. We have

$$\mathsf{E}|Z_{i,k}|^3 = \lambda_{i,k}^3 + 3\lambda_{i,k}^2 + \lambda_{i,k} \tag{41}$$

and

$$\mathsf{E}|Z_{i,k}|^4 = \lambda_{i,k}^4 + 6\lambda_{i,k}^3 + 7\lambda_{i,k}^2 + \lambda_{i,k}\ . \tag{42}$$

The fourth central moment is therefore

$$\begin{aligned}\mathsf{E}|Z_{i,k} - \lambda_{i,k}|^4 &= \mathsf{E}|Z_{i,k}|^4 - 4\lambda_{i,k}\mathsf{E}|Z_{i,k}|^3 + 6\lambda_{i,k}^2\mathsf{E}|Z_{i,k}|^2 - 4\lambda_{i,k}^3\mathsf{E}Z_{i,k} + \lambda_{i,k}^4\\ &= 3\lambda_{i,k}^2 + \lambda_{i,k}\ .\end{aligned} \tag{43}$$

We can now compute the second order moment of the residual energy as

$$\begin{aligned}\mathsf{E}\|Z_i - L_i\overline{x}\|^4 &= \mathsf{E}\|Z_i - \Lambda_i\|^4\\ &= \mathsf{E}\left|\sum_{k=1}^{M}|Z_{i,k} - \lambda_{i,k}|^2\right|^2\\ &= \sum_{k=1}^{M}\mathsf{E}|Z_{i,k} - \lambda_{i,k}|^4 + 2\sum_{1\le k<j\le M}\mathsf{E}|Z_{i,j} - \lambda_{i,j}|^2\mathsf{E}|Z_{i,k} - \lambda_{i,k}|^2\\ &= \sum_{k=1}^{M}(3\lambda_{i,k}^2 + \lambda_{i,k}) + 2\sum_{1\le k<j\le M}\lambda_{i,j}\lambda_{i,k}\\ &= 2\|\Lambda_i\|^2 + \|\Lambda_i\|_1 + \|\Lambda_i\|_1^2\ ,\end{aligned} \tag{44}$$

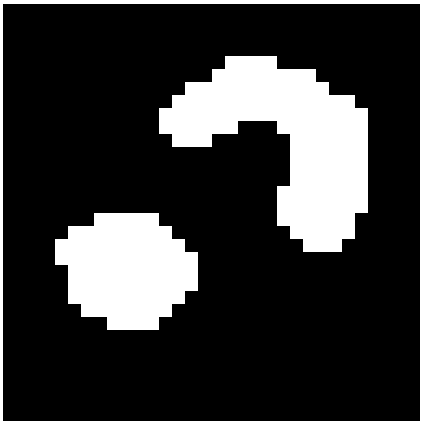


Fig. 1. Original image in last experiment.

and its variance as

$$\mathsf{var}\|Z_i - L_i\overline{x}\|^2 = \mathsf{E}\|Z_i - \Lambda_i\|^4 - \mathsf{E}^2\|Z_i - \Lambda_i\|^2 = 2\|\Lambda_i\|^2 + \|\Lambda_i\|_1 \,. \tag{45}$$

An upper bound on its standard deviation is

$$\sigma = \sqrt{2\|\Lambda_i\|_1^2 + \|\Lambda_i\|_1} \geq \sqrt{2\|\Lambda_i\|^2 + \|\Lambda_i\|_1} = \sqrt{\mathsf{var}\|Z_i - L_i\overline{x}\|^2} \,. \tag{46}$$

We look for a confidence interval of the form $[0, \mu + \beta_2\sigma]$ for some $\beta_2 > 0$. As seen in the previous section, $\mu = \|\Lambda_i\|_1$ can be approximated by the sample mean γ of (34) and, therefore, σ can be approximated by $\sqrt{2\gamma^2 + \gamma}$. In turn, the confidence interval becomes $[0, \gamma + \beta_2\sqrt{2\gamma^2 + \gamma}]$. Monte Carlo experiments show that $\beta_2 = 0.51$ gives 98% of the realizations within this interval. Thus, with a confidence level of $\mathsf{c}_2 = 98\%$, we can use the bound

$$\delta_{i,2} = \gamma + 0.51\sqrt{2\gamma^2 + \gamma} \tag{47}$$

in (29).

Global confidence analysis

There are $q+1$ sets which are confidence regions. In the previous sections, the bounds on the sets (15) and (29) were determined so as to obtain individual confidence levels of $\mathsf{c}_1 = \mathsf{c}_2 = 98\%$. Using the analysis of [10], the global confidence level c on the feasibility set $S = \bigcap_{i=1}^m S_i$ satisfies

$$\mathsf{c} \geq 1 - (1 - \mathsf{c}_1) - q(1 - \mathsf{c}_2) \,. \tag{48}$$

In our experiments, which will involve $q = 4$ or $q = 3$ views, we shall thus obtain global confidence levels of $\mathsf{c} \geq 90\%$ or $\mathsf{c} \geq 92\%$, respectively.

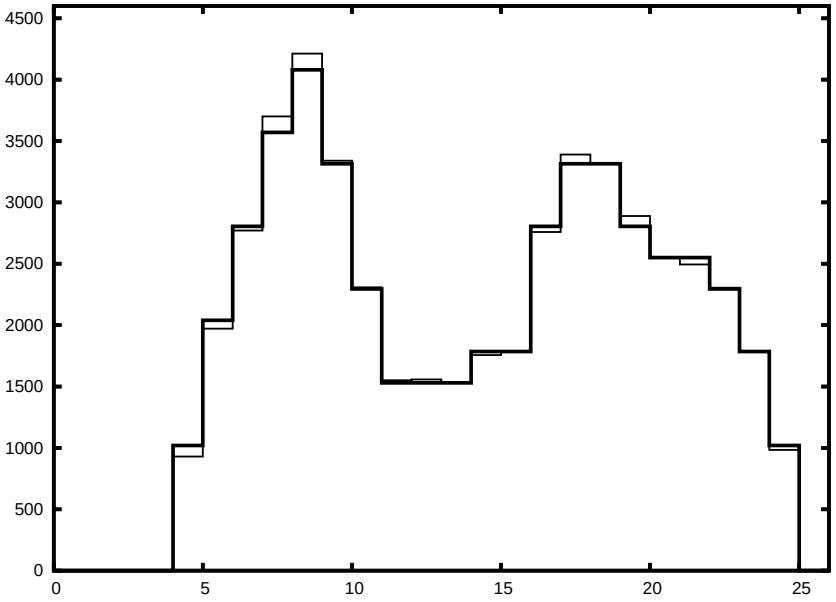


Fig. 2. Example of a sinogram view: noiseless view (bold lines) and view with Poisson noise (thin lines).

4 Numerical simulations

Most theoretical results in discrete tomography impose conditions on the shape of the objects to be reconstructed. Thus, experiments are usually performed on connected, convex, or even hv-convex objects [22]. Our algorithm does not require such stringent assumptions and we shall use the original 32×32 binary image $\overline{x}$ displayed in Fig. 1. This image features two disconnected components, one of which is nonconvex. As in most of the experiments presented in [22], few views will be used, namely $q = 4$ or $q = 3$ views.

4.1 Noise simulation

As discussed above, the noise corrupting the views in (10) is assumed to be Poisson-distributed. The pointwise variance of such a noise is directly related to the amplitude of the signal. On the other hand, the data acquisition process typically induces a multiplicative factor between the actual line sums and the measured views, due for instance to exposure time. In order to obtain a reasonable noise level, we set this multiplicative factor to 255.

We now describe the methodology used for creating the noisy views. First, for each $i \in \{1, \ldots, q\}$, the exact line sum is computed and then multiplied by the proportionality factor 255 in order to generate the noiseless view $L_i\overline{x}$.

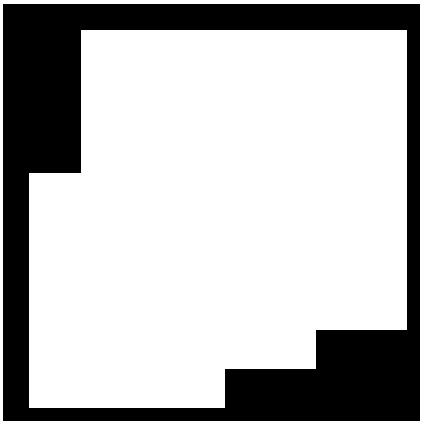


Fig. 3. Support K used in (12) in the experiments.

Then, each point $\langle L_i\overline{x} \mid e_k \rangle$ of the ith view is replaced by a realization of a Poisson variable with mean $\langle L_i\overline{x} \mid e_k \rangle$ (see [16, 26] for the numerical simulation of Poisson noise).

A typical noisy sinogram view is shown in Fig. 2. The SNR (signal-to-noise ratio) on the views varies from 31 dB to 33 dB.

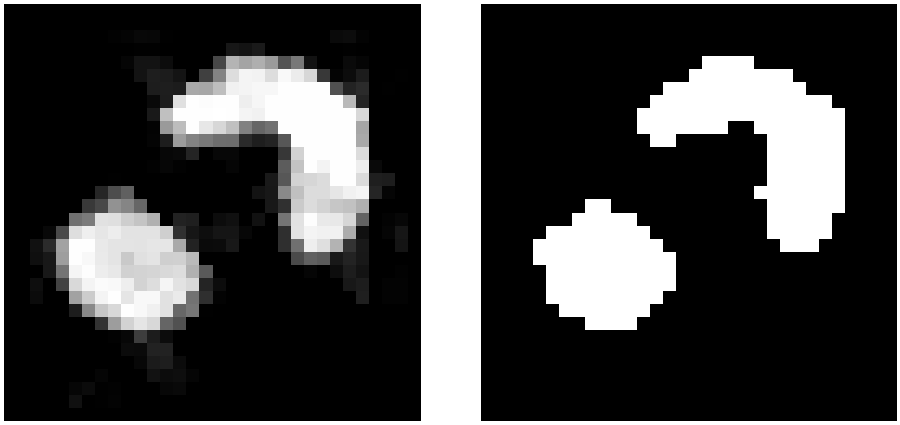
4.2 Binarization

The binarization process consists of mapping an image in $[0,1]^N$ into an image in $\{0,1\}^N$. The scheme we adopt here is straightforward: since the algorithm produces an image in the set (11), each pixel value is rounded to 0 if the original value is less than 0.5, and to 1 otherwise. It is clear that binarization could be performed in a more sophisticated fashion, especially in the light of additional a priori information on the original image.

4.3 Experimental setup

The algorithm described in Section 2 is implemented with $\varphi\colon x \mapsto \|x\|^2$ in (3) (hence R is the identity matrix and r is the zero image). In other words, we seek the image with minimum energy in the feasibility set $S = \bigcap_{i=1}^{m} S_i$. The $m = q + 4$ constraint sets to be used are:

- S_1: pixel range, see (11).
- S_2: image support, see (12) and Fig. 3.
- S_3: sum of pixel values, see (15) and (39).
- $(S_{i+3})_{1\leq i\leq q}$: residual energy of the views, see (29) and (47).
- $S_{q+4} = S_m$: total variation, see (17).



(**a**) (**b**)

Fig. 4. No noise, $q = 4$ views, without the total variation constraint. Before (**a**) and after (**b**) binarization (error: 6 pixels).

To illustrate the benefit of using the total variation set S_m, experiments with the sets $(S_i)_{1\leq i\leq m-1}$ and $(S_i)_{1\leq i\leq m}$ will be performed. Since they are easily computable in closed form, exact projections onto the sets S_1, S_2, and S_3 are used. On the other hand, subgradient projections are used for the sets $(S_i)_{4\leq i\leq m}$.

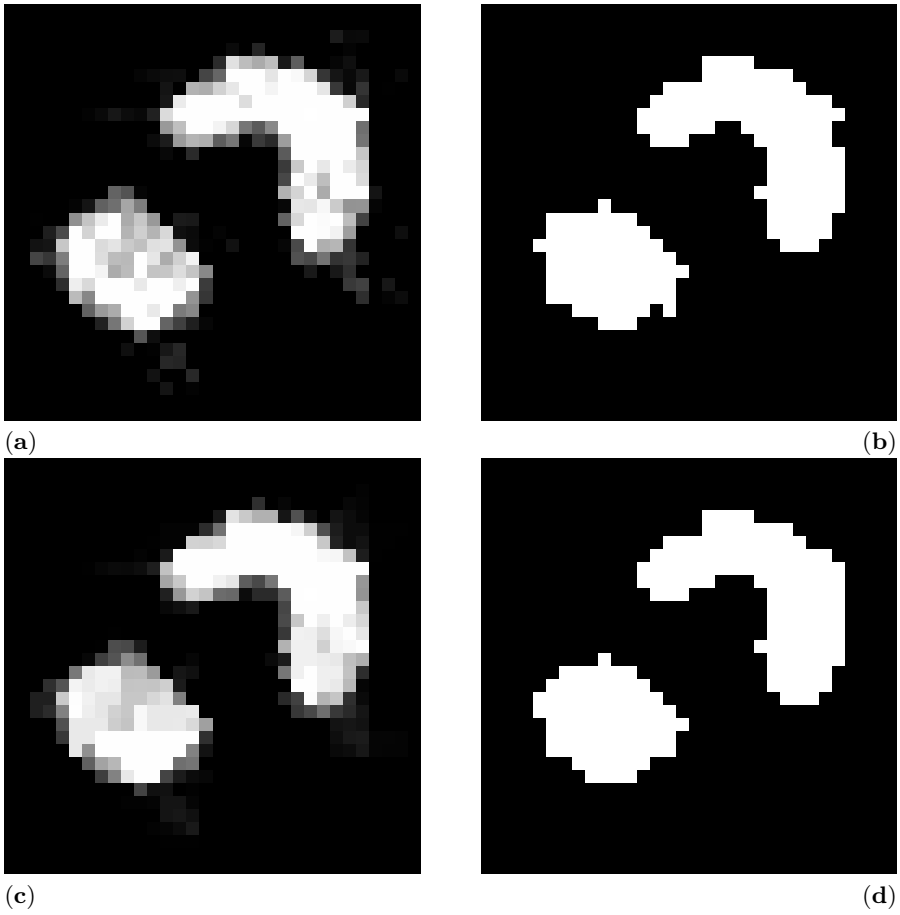
4.4 Numerical results

We first use $q = 4$ views at observation angles 0, $\pi/4$, $\pi/2$, and $3\pi/4$. In the case of noiseless views, the algorithm produces the image shown in Fig. 4.

We now turn to the case of noise-corrupted views. Fig. 5 and 6 show two typical reconstructions produced by the algorithm, for two arbitrary realizations of the noise. To illustrate the impact of the total variation constraint, we display the images obtained with and without this constraint.

In order to test the variability of our results, several hundreds of tests were performed, using different realizations of the noise and various types of images. These experiments reveal that, for a given image, the number of wrong pixels does not vary significantly. This variability is quantified in Fig. 7 in the case of the original image of Fig. 1.

Finally, we show how the algorithm behaves on a standard image from the binary tomography literature. This image, shown in Fig. 8, can be reconstructed uniquely from its exact (noiseless) horizontal and vertical views [22]. This theoretical result is of course no longer true in a noisy environment. However, our algorithm reconstructs the image almost perfectly in the presence of $q = 3$ noise-corrupted views at angles 0, $\pi/4$, and $\pi/2$ (see Fig. 9) with the 6 sets:



(**a**) (**b**)

(**c**) (**d**)

Fig. 5. Poisson noise, $q = 4$ views. Reconstruction without the total variation constraint: before (**a**) and after (**b**) binarization (error: 16 pixels). Reconstruction with the total variation constraint: before (**c**) and after (**d**) binarization (error: 8 pixels).

- S_1: pixel range, see (11).
- S_2: sum of pixel values, see (15) and (39).
- $(S_{i+2})_{1\leq i\leq 3}$: residual energy of the views, see (29) and (47).
- S_6: total variation, see (17).

5 Concluding remarks

We have proposed a convex programming approach to the discrete tomographic image reconstruction problem. Promising results have been obtained

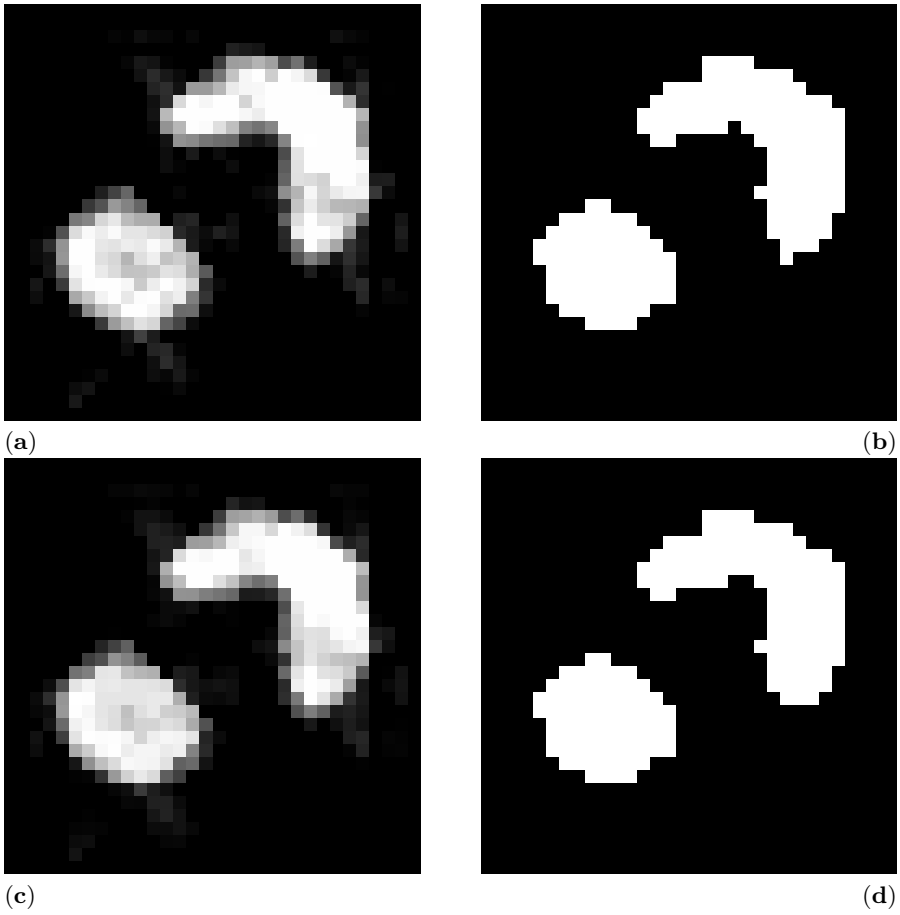


Fig. 6. Same experiments as in Fig. 5 with a different realization of the noise. Reconstruction without the total variation constraint: before (**a**) and after (**b**) binarization (error: 12 pixels). Reconstruction with the total variation constraint: before (**c**) and after (**d**) binarization (error: 9 pixels).

with a limited number of constraints and a quadratic objective function. The proposed algorithm can also handle a wide range of additional constraints as well as more general objective functions. The exploration of such extensions is left for future work. In particular, additional constraints could be derived from a better understanding of the noise process. Likewise, while total variation appears to give good results, soft enforcement of binarity should be possible through alternative convex functionals.

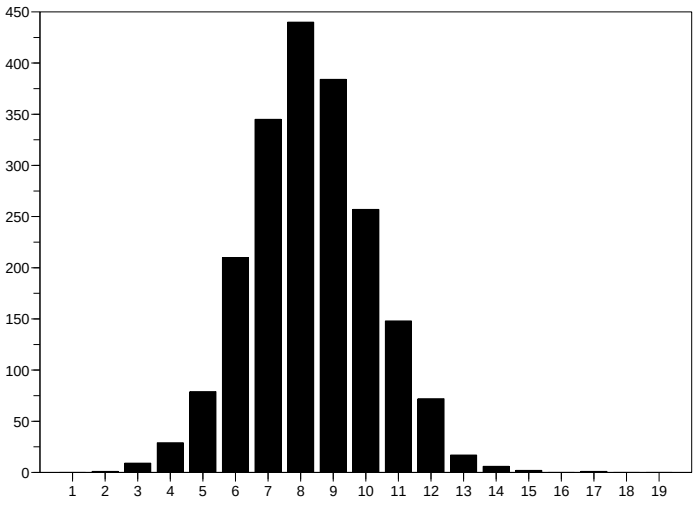


Fig. 7. Experiments of Fig. 5 and 6: histogram of the number of wrong pixels based on 2000 realizations of the noise.

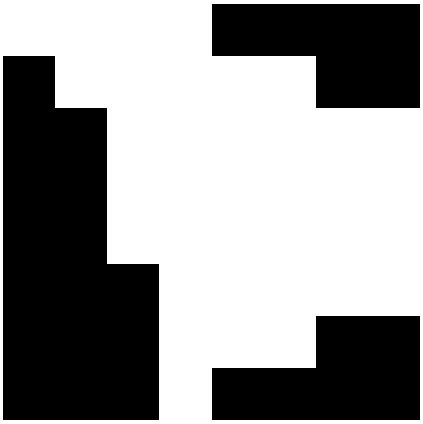
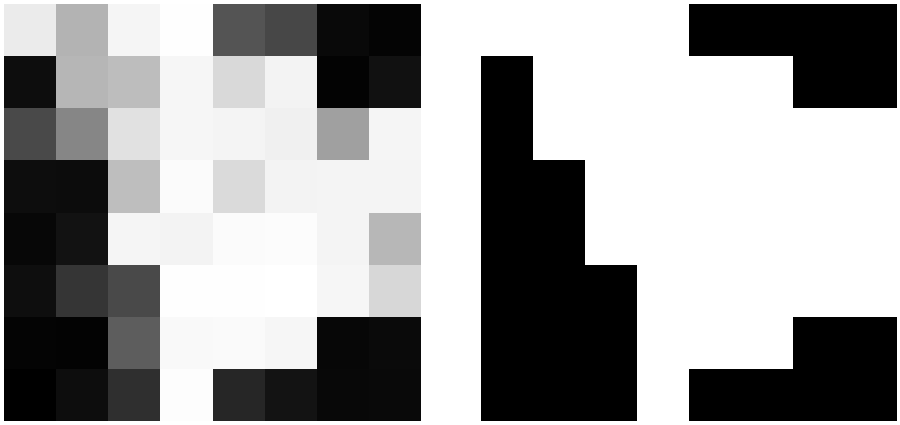


Fig. 8. Original image.

References

1. Bauschke, H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, **38**, 367–426 (1996).
2. Bronstein, M., Bronstein, A.M., Zibulevsky, M., Azhari, H.: Reconstruction in diffraction ultrasound tomography using nonuniform FFT. *IEEE Trans. Med. Imaging*, **21**, 1395–1401 (2002).



(**a**) (**b**)

Fig. 9. Poisson noise, $q = 3$ views. Reconstruction before (**a**) and after (**b**) binarization (error: 1 pixel).

3. Capricelli, T.D., Combettes, P.L.: Parallel block-iterative reconstruction algorithms for binary tomography. In: Herman, G.T., Kuba, A. (eds.), *Electron. Notes Discrete Math., Proceedings of the Workshop on Discrete Tomography and its Applications*, New York, USA, pp. 263–280 (2005).
4. Censor, Y., Zenios, S.A.: *Parallel Optimization – Theory, Algorithms and Applications*. Oxford University Press, New York (1997).
5. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, **20**, 89–97 (2004).
6. Combettes, P.L.: The Convex Feasibility Problem in Image Recovery. *Advances in Imaging and Electron Physics*. Academic Press, New York, **95**, 155–270 (1996).
7. Combettes, P.L.: Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections. *IEEE Trans. Image Process.*, **6**, 493–506 (1997).
8. Combettes, P.L.: Strong convergence of block-iterative outer approximation methods for convex optimization. *SIAM J. Control Optim.*, **38**, 538–565 (2000).
9. Combettes, P.L.: A block-iterative surrogate constraint splitting method for quadratic signal recovery. *IEEE Trans. Signal Process.*, **51**, 1771–1782 (2003).
10. Combettes, P.L., Chaussalet, T.J.: Combining statistical information in set theoretic estimation. *IEEE Signal Process. Lett.*, **3**, 61–62 (1996).
11. Combettes, P.L., Luo, J.: An adaptive level set method for nondifferentiable constrained image recovery. *IEEE Trans. Image Process.*, **11**, 1295–1304 (2002).
12. Combettes, P.L., Pesquet, J.C.: Image restoration subject to a total variation constraint. *IEEE Trans. Image Process.*, **13**, 1213–1222 (2004).
13. Combettes, P.L., Trussell, H.J.: The use of noise properties in set theoretic estimation. *IEEE Trans. Signal Process.*, **39**, 1630–1641 (1991).
14. Combettes, P.L., Trussell, H.J.: Deconvolution with bounded uncertainty. *Int. J. Adapt. Control Signal Process.*, **9**, 3–17 (1995).
15. Del Lungo, A., Gronchi, P., Herman, G.T.: Special issue on discrete tomography. *Linear Algebra Appl.*, **339**, 1–2 (2001).

16. Devroye, L.: *Non-Uniform Random Variate Generation*. Springer-Verlag, New York (1986).
17. Gardner, R.J.: *Geometric Tomography*. Cambridge University Press, New York (1995).
18. Gordon, R., Bender, R., Herman, G.T.: Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *J. Theoret. Biology*, **29**, 471–481 (1970).
19. Gritzmann, P., de Vries, S., Wiegelmann, M.: Approximating binary images from discrete X-rays. *SIAM J. Optim.*, **11**, 522–546 (2000).
20. Herman, G.T.: A relaxation method for reconstructing objects from noisy X-rays. *Math. Program.*, **8**, 1–19 (1975).
21. Herman, G.T.: *Image Reconstruction from Projections, the Fundamentals of Computerized Tomography*. Academic Press, New York (1980).
22. Herman, G.T., Kuba, A. Eds.: *Discrete Tomography: Foundations, Algorithms, and Applications*. Birkhäuser, Boston (1999).
23. Herman, G.T., Kuba, A. Eds.: Proc. Workshop on Discrete Tomography and Its Applications. *Electron. Notes Discrete Math.*, **20**, 1–622 (2005).
24. Kudo, H., Saito, T.: Sinogram recovery with the method of convex projections for limited-data reconstruction in computed tomography. *J. Opt. Soc. Amer. A*, **8**, 1148–1160 (1991).
25. Oskoui-Fard, P., Stark, H.: Tomographic image reconstruction using the theory of convex projections. *IEEE Trans. Med. Imaging*, **7**, 45–58 (1988).
26. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: *Numerical Recipes in C Example Book: The Art of Scientific Computing,* 2nd ed. Cambridge University Press (1992).
27. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton, N.J. (1970).
28. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D*, **60**, 259–268 (1992).
29. Schapira, P.: Tomography of constructible functions. In Cohen, G.D., Giusti, M., Mora, T. (eds.), *Lecture Notes in Comput. Sci., Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 11th International Symposium, AAECC-11*, Paris, France, pp. 427–435 (1995).
30. Sezan, M.I., Stark, H.: Tomographic image reconstruction from incomplete view data by convex projections and direct Fourier inversion. *IEEE Trans. Med. Imaging*, **3**, 91–98 (1984).
31. Stark, H., Ed.: *Image Recovery: Theory and Applications*. Academic Press, San Diego, CA (1987).
32. Stark, H., Yang, Y: *Vector Space Projections: A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics*. Wiley, New York (1998).
33. Trussell, H.J., Civanlar, M.R.: The feasible solution in signal restoration. *IEEE Trans. Acoust., Speech, Signal Process.*, **32**, 201–212 (1984).
34. Weber, S., Schüle, T., Hornegger, J., Schnörr, C.: Binary tomography by iterating linear programs from noisy projections. In: Klette, R, Zunic, J.D. (eds.), *Lecture Notes in Comput. Sci., Combinatorial Image Analysis, 10th International Workshop, IWCIA 2004*, Auckland, New Zealand, pp. 38–51 (2004).
35. Wernick, M.N., Chen, C.T.: Superresolved tomography by convex projections and detector motion. *J. Opt. Soc. Amer. A*, **9**, 1547–1553 (1992).

36. Zhang, X.-Q., Froment, J.: Constrained total variation minimization and application in computerized tomography. In: Rangarajan, A., Vemuri, B.C., Yuille, A.L. (eds.), *Lecture Notes in Comput. Sci., Energy Minimization Methods in Computer Vision and Pattern Recognition, 5th International Workshop, EMMCVPR 2005*, St. Augustine, FL, USA, pp. 456–472 (2005).