

Gradient Flows on a Riemannian Submanifold for Discrete Tomography

Matthias Zisler¹, Fabrizio Savarino¹, Stefania Petra², and Christoph Schnörr¹

¹Image and Pattern Analysis Group, Heidelberg University, Germany

²Mathematical Imaging Group, Heidelberg University, Germany

Abstract. We present a smooth geometric approach to discrete tomography that jointly performs tomographic reconstruction and label assignment. The flow evolves on a submanifold equipped with a Hessian Riemannian metric and properly takes into account given projection constraints. The metric naturally extends the Fisher-Rao metric from labeling problems with directly observed data to the inverse problem of discrete tomography where projection data only is available. The flow simultaneously performs reconstruction and label assignment. We show that it can be numerically integrated by an implicit scheme based on a Bregman proximal point iteration. A numerical evaluation on standard test-datasets in the few angles scenario demonstrates an improvement of the reconstruction quality compared to competitive methods.

1 Introduction

Discrete tomography [9] denotes the problem to reconstruct *piecewise constant* functions from projection data, that are taken from few projection angles only. Such extremely ill-posed inverse problems are motivated by industrial applications, like quality inspection. Regularization of such problems essentially rests upon the fact that the functions to be reconstructed only take values in a finite set of labels $\mathcal{L} := \{c_1, \dots, c_K\} \subset [0, 1]$. This is similar to the common image labeling problem in computer vision, with the essential difference that the function u to be labelled is only *indirectly* observed. Specifically, after a standard problem discretization resulting in the representation $u \in \mathbb{R}^N$, projection data b given by

$$Au = b \quad \text{s.t.} \quad u_i \in \mathcal{L}, \quad \forall i = 1, \dots, N \quad (1)$$

are observed, where the matrix A is underdetermined but known. The task is to reconstruct u subject to the labeling constraints $u_i \in \mathcal{L}$, $\forall i$.

Related Work. A natural class of approaches are based on minimizing *convex sparsifying functionals* of u (e.g. total variation) subject to the affine subject constraints (1), but *without* the labeling constraints [14, 8, 7]. Unless sufficient conditions for unique recovery are met, in terms of the number of projection measurements relative to the complexity of the discontinuity set of u [7], the performance of the necessary rounding post-processing step is difficult to control, however. Likewise, a binary discrete graphical model from labeling was

adopted by [10], and a sequence of s-t graph-cuts was solved to take into account the affine projection constraints. An extension to the non-binary case (multiple labels) seems to be involved. The authors of [15] minimize the ℓ_0 -norm of the gradient directly by a dynamic programming approach, but do not exploit the set \mathcal{L} of feasible labels for regularization.

Approaches that aim to enforce the labeling constraints by *continuous non-convex* optimization include [18, 12, 20, 21]. Unlike our approach proposed below, that limits the degrees of freedom by restricting the feasible set to a Riemannian submanifold, these approach work in the higher-dimensional ambient Euclidean space and hence are more susceptible to poor initializations and local minima. A step towards alleviating these problems was recently done by [19], where a different regularization strategy was proposed based on the Kullback-Leibler (KL) divergence.

Further approaches that define the state of the art include [16, 4]. The authors of [16] proposed a heuristic algorithm that adaptively combines an energy formulation with a non-convex polynomial representation, in order to steer the reconstruction towards the feasible label set. Batenburg et al. [4] proposed the *Discrete Algebraic Reconstruction Technique (DART)* algorithm which starts with a continuous reconstruction by a basic algebraic reconstruction method, followed by a thresholding operation. These steps, interleaved with smoothing, are iteratively repeated to refine the locations of the boundaries. This heuristic approach yields good reconstructions in practice, but cannot be characterized by an objective function which is optimized.

We regard [4, 16, 20] as state-of-the-art approaches in our experimental comparison.

Contribution. We present a novel geometric approach to discrete tomography by optimizing over a Riemannian submanifold of discrete probability measures with full support. Our work is motivated by the recent work [3], where the ordinary labeling problem (with directly observed data) is solved by a Riemannian gradient flow on a manifold of discrete probability measures that represent label assignments. By restricting the feasible set to a submanifold, equipped with a natural extension of the Fisher-Rao metric, we extend this approach to discrete tomography. The resulting gradient flow takes into account the projection constraints and simultaneously performs reconstruction and label assignment. We show that this flow can be numerically integrated by an implicit scheme requires to solve a convex problem at each step. A comprehensive numerical evaluation demonstrates the superior reconstruction performance of our approach compared to related work.

Basic Notation. Functions like log and binary operations (multiplication, subdivision) are applied *component-wise* to vectors and matrices, e.g., $vw = (\dots, v_i w_i, \dots)^T$. The KL-divergence is defined by $\text{KL}(x, y) = \langle x, \log(x/y) \rangle + \langle y - x, \mathbb{1} \rangle$ for both vectors and matrices with non-negative entries, where $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product. We set $\mathbb{1} = (1, 1, \dots, 1)^T$ and $[n] = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$. The linear operator $\text{vec}(\cdot)$ maps matrices to vectors by stacking columns. Finally \otimes denotes the Kronecker product.

2 Approach

We briefly summarize the approach [3]. Then we extend this approach in order to additionally take into account the affine subspace constraint: We construct a smooth Riemannian gradient flow, for any smooth objective function, restricted to the relative interior. Finally, we specify an objective function that is used for the experimental evaluation.

Smooth Geometric Label Assignment. Each label $c_k \in \mathcal{L}$ is represented by a vertex of the probability simplex $e_k \in \mathbb{R}^K$, and the set of feasible label assignments to all pixels corresponds to the set of row-stochastic matrices with full support, denoted by $\mathcal{W} \subset \mathbb{R}_{++}^{N \times K}$. In [3] a smooth geometric approach for labeling is proposed, where \mathcal{W} is turned into a Riemannian manifold using the Fisher-Rao (information) metric [5]. For a given image $u \in \mathbb{R}^N$, the distance between each pixel u_i , $i \in [N]$ and each label $c_k \in \mathcal{L}$, $k \in [K]$ is measured and collected by a distance matrix D_{ik} . Next this matrix is projected onto the tangent space $T^N \simeq T_W \mathcal{W} = \{T \in \mathbb{R}^{N \times K} \mid T \mathbb{1}_K = 0_N\}$ by subtracting pixelwise the mean of D , i.e. $\Pi(D) = D - \frac{1}{K} D \mathbb{1}_K \mathbb{1}_K^T$. The projection $\Pi(D)$ in turn is mapped to the manifold \mathcal{W} by the so-called lifting map

$$\exp: \mathcal{W} \times T^N \rightarrow \mathcal{W}, \quad (W, V) \mapsto \exp_W(V) := \frac{W e^V}{\langle W, e^V \rangle}, \quad (2)$$

to obtain the likelihood matrix $L = \exp_W(\Pi(D))$. Next, spatial regularization is performed by computing Riemannian means of the row vectors L_i within a spatial neighbourhood $\mathcal{N}(i)$ for each pixel $i \in [N]$. It is shown in [3] that these means admit the closed-form solution

$$S(W)_i = \frac{\mathfrak{m}_g(\{L(W)_j\}_{j \in \mathcal{N}(i)})}{\langle \mathfrak{m}_g(\{L(W)_j\}_{j \in \mathcal{N}(i)}), \mathbb{1}_K \rangle}, \quad \mathfrak{m}_g(\{L(W)_j\}_{j \in \mathcal{N}(i)}) := \prod_{j \in \mathcal{N}(i)} L(W)_j^{\frac{1}{|\mathcal{N}(i)|}}. \quad (3)$$

Finally, a labeling in terms of $W \in \mathcal{W}$ is determined by maximizing the correlation $\langle W, S(W) \rangle$. The optimization is carried out on the manifold \mathcal{W} by an explicit Euler scheme for integrating the Riemannian gradient flow (assignment flow).

Tomographic Assignment Flow. We now consider the situation where the image data are only indirectly observed through the projection constraints (1). To this end, we extend the approach [3] using techniques developed by [2], in order to restrict the smooth Riemannian flow to assignments that respect the projection constraints.

Our starting point is the observation that the Riemannian metric used in [3] is induced by the Hessian of the convex Legendre function

$$h(W) := \langle W, \log(W) - \mathbb{1}_N \mathbb{1}_K^T \rangle, \quad (4)$$

with domain restricted to the relative interior of $\overline{\mathcal{W}} = \{W \in \mathbb{R}_+^{N \times K} : W \mathbb{1}_K = \mathbb{1}_N\}$. In order to take into account the projection constraints (1), we introduce

the assignment operator

$$P_{\mathcal{L}}: \mathcal{W} \rightarrow \mathbb{R}^N, \quad W \mapsto P_{\mathcal{L}}(W) = (I_N \otimes c^T) \text{vec}(W) = Wc, \quad (5)$$

that makes explicit the reconstructed function $u = Wc$ in terms of the given labels c and the assignment W . Based on this correspondence and (1), we extend the set $\overline{\mathcal{W}}$ to

$$\overline{\mathcal{F}} = \left\{ W: \mathbb{R}_+^{N \times K}, B \text{vec}(W) = \begin{pmatrix} b \\ \mathbf{1}_N \end{pmatrix} \right\}, \quad B = \begin{pmatrix} A(I_N \otimes c^T) \\ I_N \otimes \mathbf{1}_K^T \end{pmatrix}. \quad (6)$$

The following non-degeneracy property is crucial for the smooth geometric construction below. The proof exploits the structure of B and properties of the Kronecker product. We omit details due to the page limit.

Lemma 1 (rank of B). *The matrix B has full row rank by construction, if the measurement matrix A has full row rank.*

Our next step is to extend the manifold \mathcal{W} to a manifold \mathcal{F} , based on the extension of $\overline{\mathcal{W}}$ to $\overline{\mathcal{F}}$. We adopt the convex Legendre function $h(W)$ from above and take as its domain the linear manifold $\mathcal{M} = \mathbb{R}_{++}^{N \times K}$. Then the Hessian $\nabla^2 h(W) = \frac{1}{W}$ (componentwise inverse) smoothly depends on $W \in \mathcal{M}$ and defines the linear mapping

$$H(W): \mathbb{R}^{N \times K} \rightarrow \mathbb{R}^{N \times K}, \quad U \mapsto H(W)U := (U_{ij}/W_{ij})_{i \in [N], j \in [K]}. \quad (7)$$

Based on the canonical identification of the tangent spaces $T_W \mathcal{M} \simeq \mathbb{R}^{N \times K}$ for linear manifolds, the mapping $H(W)$ defines the Riemannian metric

$$(U, V)_W^H := \langle H(W)U, V \rangle, \quad \forall W \in \mathcal{M}, \quad U, V \in \mathbb{R}^{N \times K}. \quad (8)$$

Given some smooth objective $J(W)$, the corresponding Riemannian gradient field restricted to \mathcal{M} is given by

$$\nabla_H J|_{\mathcal{M}}(W) := H(W)^{-1} \nabla J(W). \quad (9)$$

Next we consider the smooth submanifold $\mathcal{F} := \text{rint}(\overline{\mathcal{F}}) = \mathcal{M} \cap \overline{\mathcal{F}}$ of \mathcal{M} with tangent space $T_W \mathcal{F} \simeq \mathcal{N}(B)$. The metric on \mathcal{M} induces a metric on \mathcal{F} , and the Riemannian gradient field of $J(W)$ restricted to \mathcal{F} is given by

$$\nabla_H J|_{\mathcal{F}}(W) := P_W^{\mathcal{N}(B)}(H(W)^{-1} \nabla J(W)), \quad (10)$$

where $P_W^{\mathcal{N}(B)}$ is the $(\cdot, \cdot)_W^H$ -orthogonal projection onto the nullspace $\mathcal{N}(B)$. Since the matrix B has full rank due to Lemma 1, this projection reads

$$P_W^{\mathcal{N}(B)}(H(W)^{-1} \nabla J(W)) = \text{vec}^{-1} \left[(I - (BD_W^H B^T)^{-1} BD_W^H) (D_W^H)^{-1} \text{vec}[\nabla J(W)] \right], \quad (11)$$

where $D_W^H = \text{Diag}[\text{vec}(H(W))]$. The vector $-\nabla_H J|_{\mathcal{F}}(W)$ for $W \in \mathcal{F}$ is the steepest descent direction in $\mathcal{N}(B)$. Furthermore, minimization of an objective J

on the Riemannian manifold $(\mathcal{F}, (\cdot, \cdot)_{\hat{W}}^H)$ amounts to find the solution trajectory $W(t)$ of the dynamical system

$$\dot{W}(t) + \nabla_H J|_{\mathcal{F}}(W(t)) = 0, \quad W(0) = W^0 \in \mathcal{F}, \quad (12)$$

with initial condition $W^0 \in \mathcal{F}$.

Objective Function. We adopt and modify the approach of [3] sketched in Section 2, for our purpose. Defining the distance matrix $D(\hat{W}) := \frac{1}{\rho} (\|P_{\mathcal{L}}(\hat{W})_i - c_k\|_2^2)_{i \in [N], k \in [K]}$, with the assignment operator $P_{\mathcal{L}}(\hat{W})$ given by (5) and a scaling parameter $\rho > 0$, we compute a similarity matrix $S(\hat{W})$ as described in connection with (3). Based on $S(\hat{W})$, we define the objective function

$$J(W, \hat{W}) = \text{KL}(W, S(\hat{W})^{1+\alpha}), \quad \alpha > 0. \quad (13)$$

Minimizing J with respect to W encodes two aspects. Firstly, the discrete assignment distributions comprising W should be consistent with the spatially regularized similarities $S(\hat{W})$, that correspond to the lifted distances $D(\hat{W})$ between the reconstructed function $P_{\mathcal{L}}(\hat{W})$ and the labels c . Secondly, since W appears as first argument of the KL-distance, W matches the prominent *modes* of the discrete distributions comprising $S(\hat{W})$ (cf. [11]), and hence enforces unique labelings. The damping parameter α enables to control this “rounding property”.

Since the assignment \hat{W} is not given beforehand, we pursue an iterative strategy and set $\hat{W} = W^k$ to the current iterate, in order to compute W^{k+1} by minimizing (13). In the next section, we formulate this process in a more principled way as a fixed point iteration, that properly discretizes and solves the continuous flow (12).

3 Optimization

In this section we want to find a solution trajectory of the initial value problem (12) associated with the steepest Riemannian gradient descent of the convex objective function J in (13) on the smooth manifold \mathcal{F} . Following [2], we reformulate (12) as a differential inclusion for a time interval (T_m, T_M) corresponding to the unique maximal solution of (12) and obtain

$$\frac{d}{dt} \nabla h(W(t)) + \nabla J(W(t)) \in \mathcal{N}(B)^\perp, \quad W(t) \in \mathcal{F}, \quad W(0) = W^0 \in \mathcal{F}, \quad (14)$$

with h given by (4). Since J is convex, an implicit discretization yields the iterative scheme: $\nabla h(W^{k+1}) - \nabla h(W^k) + \mu_k \nabla J(W^{k+1}) \in \mathcal{N}(B)^\perp$, $B \text{vec}(W^{k+1}) = y$ and $W^0 \in \mathcal{F}$, where μ_k is a step-size parameter. These relations are just the optimality conditions of the Bregman proximal point method with the KL-divergence as proximity measure

$$W^{k+1} \in \arg \min_{W \in \mathbb{R}_+^{N \times K}} J(W, \hat{W}) + \frac{1}{\mu_k} \text{KL}(W, W^k) \quad \text{s.t.} \quad B \text{vec}(W) = y. \quad (15)$$

Algorithm 1: Iterated Primal Dual Algorithm**Init:** choose the barycenter for $W^0 \in \mathcal{G}$, dual variable $Q^0 \in \mathbb{R}^m$ and $\tau, \sigma > 0$ **Parameters:** selectivity $\rho > 0$, discretization $\alpha > 0$, trust region $\mu_k > 0$ **while** *not converged* **do** Warmstart for PD: $W^0 = W^k$, $\tilde{W} = T_{\mu_k}(W^k)$, $Q^0 = Q^{last}$, $n = 1$ **while** *not converged* **do**

$$W^{n+1} = \arg \min_{W \in \overline{\mathcal{W}}} \text{KL}(W, \tilde{W}) + \langle W, P_{\mathcal{L}}^T(A^T Q^n) \rangle + \frac{1}{\tau} \text{KL}(W, W^n) \quad (17)$$

$$Q^{n+1} = \arg \min_Q \langle Q, b - AP_{\mathcal{L}}(2W^{n+1} - W^n) \rangle + \frac{1}{2\sigma} \|Q - Q^n\|_2^2 \quad (18)$$

 $n \leftarrow n + 1$ $k \leftarrow k + 1$, $W^k \leftarrow W^n$ **Output:** W^k

We solve (15) for fixed W^k by an iterative algorithm to perform an implicit integration step on the flow (12). In order to update the fixed \hat{W} in $J(W, \hat{W})$ defined by (13), we set $\hat{W} = W^k$. Inserting into (15) and combining the KL-divergences as a multiplicative convex combination with respect to the second argument yields the fixed point iteration

$$W^{k+1} \in \arg \min_{W \in \overline{\mathcal{W}}} \text{KL}(W, \underbrace{(W^k)^{\frac{1}{1+\mu_k}} (S(W^k))^{\frac{\mu_k(1+\alpha)}{1+\mu_k}}}_{:=T_{\mu_k}(W^k)}) \quad \text{s.t.} \quad AP_{\mathcal{L}}(W) = b, \quad (16)$$

where the constraints $W \in \mathbb{R}_+^{N \times K}$ and $B \text{vec}(W) = y$ are rewritten as $W \in \overline{\mathcal{W}}$ and $AP_{\mathcal{L}}(W) = b$. Regarding convergence of the fixed point iteration (16), we use a non-summable diminishing step-size parameter $\mu_k = \frac{1}{0.005 \cdot k \cdot \|AP_{\mathcal{L}}(W^k) - b\|_2}$ with $\lim_{k \rightarrow \infty} \mu_k = 0$. Hence the operator T_{μ_k} becomes $T_{\mu_k} \rightarrow \text{Id}$ for $k \rightarrow \infty$ and the influence of the objective function J vanishes. When the iteration converges, then (16) reduces to the KL-projection onto the fixed feasible set $\overline{\mathcal{F}}$. A rigorous mathematical convergence analysis of the iterations (16) is left for future work.

Solving the Fixed Point Iteration. Algorithm 1 solves equation (16) iteratively using the generalized primal dual algorithm [6]. The primal update step (17) can be evaluated in closed form

$$W^{n+1} = \arg \min_{W \in \overline{\mathcal{W}}} \text{KL}(W, \tilde{W}) + \langle W, P_{\mathcal{L}}^T(A^T Q^n) \rangle + \frac{1}{\tau} \text{KL}(W, W^n) \quad (19a)$$

$$= \frac{(W^n)^{\frac{1}{1+\tau}} (\tilde{W})^{\frac{\tau}{1+\tau}} \exp(-\frac{\tau}{1+\tau} P_{\mathcal{L}}^T(A^T Q^n))}{\langle (W^n)^{\frac{1}{1+\tau}} (\tilde{W})^{\frac{\tau}{1+\tau}}, \exp(-\frac{\tau}{1+\tau} P_{\mathcal{L}}^T(A^T Q^n)) \rangle}. \quad (19b)$$

The dual update step (18) admits a closed form as well,

$$Q^{n+1} = \arg \min_Q \langle Q, b - \text{AP}_{\mathcal{L}}(2W^{n+1} - W^n) \rangle + \frac{1}{2\sigma} \|Q - Q^n\|_2^2 \quad (20a)$$

$$= Q^n + \sigma(\text{AP}_{\mathcal{L}}(2W^{n+1} - W^n) - b). \quad (20b)$$

Parameter Selection. For the step-size parameters τ and σ of the iterated primal-dual algorithm, we adopt the parameter values of [6, Example 7.2] and set $\tau = \sqrt{K}/L_{12}^2$ for the primal update and $\sigma = 1/\sqrt{K}$ for the dual update. This choice implies that the condition $\sigma\tau\|\text{AP}_{\mathcal{L}}(\cdot)\|^2 \leq 1$ for convergence holds, with the operator norm $\|\text{AP}_{\mathcal{L}}(\cdot)\| = \sup_{\|x\|_1 \leq 1} \|A(I_N \otimes c^T)x\|_2 = \max_j \|(A(I_N \otimes c^T))_j\|_2$. This reflects the fact that the negative entropy is 1-strongly convex with respect to the L_1 -norm when restricted to the probability simplex.

4 Numerical Experiments

We compared the proposed approach to state-of-the-art approaches for discrete tomography, including the Discrete Algebraic Reconstruction Technique (*DART*) [4], the energy minimization method of Varga et al. [16] (*Varga*), and the layer-wise total variation approach (*LayerTV*) [20].

Setup. We adapted the binary phantoms from [17] to the multi-label case, shown as phantom 1,2 and 3 in Figure 1. Phantom 4 is the well-known Shepp-Logan phantom [13]. We simulated noisy scenarios by applying Poisson noise to the measurements b with a signal-to-noise ratio of $SNR = 20$ db. The geometrical setup was created by the *ASTRA*-toolbox [1], where we used parallel projections along equidistant angles between 0 and 180 degrees. The width of the sensor-array was set 1.5 times the image size, such that every pixel is intersecting with at least one single projection ray.

Implementation details. The subproblems of Algorithm 1 were approximately solved by the generalized PD algorithm [6]. For the multiplicative updates (19b), we adopted the renormalization strategy from [3] to avoid numerical issues close to the boundary of the manifold, that correspond to unambiguous label assignments. The outer iteration was terminated when $\|\text{AP}_{\mathcal{L}}(W^k) - b\|_2 < 0.1$. For the geometric averaging (cf. (3) and (13)), we used a 3×3 neighborhood for the smaller phantom 1 and 5×5 for all others. In order to reconstruct from *noisy* measurements, we modified the proposed approach by using the squared L_2 -reprojection error as relaxed dataterm, so that the objective (13) reads

$$J(W, \hat{W}) = \text{KL}(W, S(\hat{W})^{1+\alpha}) + \frac{1}{2} \|\text{AP}_{\mathcal{L}}(W) - b\|_2^2, \quad \alpha > 0, \quad (21)$$

which is smooth and convex in W as well. In this case, the fixed point iteration (16) is applied to the modified objective (21) and the dual update step (18) of algorithm 1 is additionally rescaled, i.e. $Q^{n+1} = (Q^n + \sigma(\text{AP}_{\mathcal{L}}(2W^{n+1} - W^n) - b))/(1 + \sigma)$ compared to (20b).

Regarding DART we used the public available implementation included in the ASTRA-toolbox [1], whereas for Varga [16] and LayerTV [20] we used our own implementations in MATLAB. We used the default parameters of the competing approaches as proposed by the respective authors. However, since the test-datasets differ in size, we slightly adjusted the parameters in order to get best results for every algorithm and problem instance.

Results. Figure 1 summarizes the numerical evaluation of the approaches for increasing (but small) numbers of projections, in the noiseless case (filled markers) and in the noisy case (non-filled markers), with Poisson noise $SNR = 20$ dB. Each test-dataset is depicted in the leftmost column, followed by the relative pixel error and runtime. The proposed approach achieved perfect reconstructions with a small number of projection angles in the noiseless case. Only LayerTV needed one projection less at phantom 3 and 4. LayerTV however tends to return non-integral solutions when the regularization parameter is large and then requires a special rounding strategy to obtain a meaningful reconstruction. In noisy scenarios, LayerTV performs better due to use of inequality projection constraints, followed by the proposed method that outperforms both DART and Varga. Figure 2 shows the poor “implicit data terms” generated by the tomographic constraints in case of phantom 2, to illustrate the severe ill-posedness of these inverse problems (see the caption for more details).

Considering the runtime (right plots from figure 1), DART is the fastest approach followed by Varga. The proposed approach and LayerTV are clearly consuming more runtime to return more accurate solutions. In the noiseless and with a sufficient projection angles, the proposed approach is faster. We point out that the proposed approach could be easily parallelized using graphics cards. In figure 3 and 4 the visual results are displayed for the phantoms 2 and 3,

5 Conclusion and Future Work

We presented a novel smooth geometric approach for jointly solving tomographic reconstruction and assignment. We derived a suitable Riemannian structure on the feasible set in order to optimize a smooth objective function on a manifold that respects the projection constraints. The Riemannian gradient flow combines tomographic reconstruction and labeling in a smooth and mathematically sound way.

Our future work will include a rigorous mathematical convergence analysis of the fixed-point iteration (16) and of the stability of the corresponding Riemannian gradient descent flow (12), that entails iterative updates $\hat{W} = W^k$ of the objective function $J(W, \hat{W})$. Such issues are not covered by standard convex programming. A promising extension of the proposed approach concerns the ability to handle inequality constraints, in order to further improve the performance in scenarios with high noise levels.

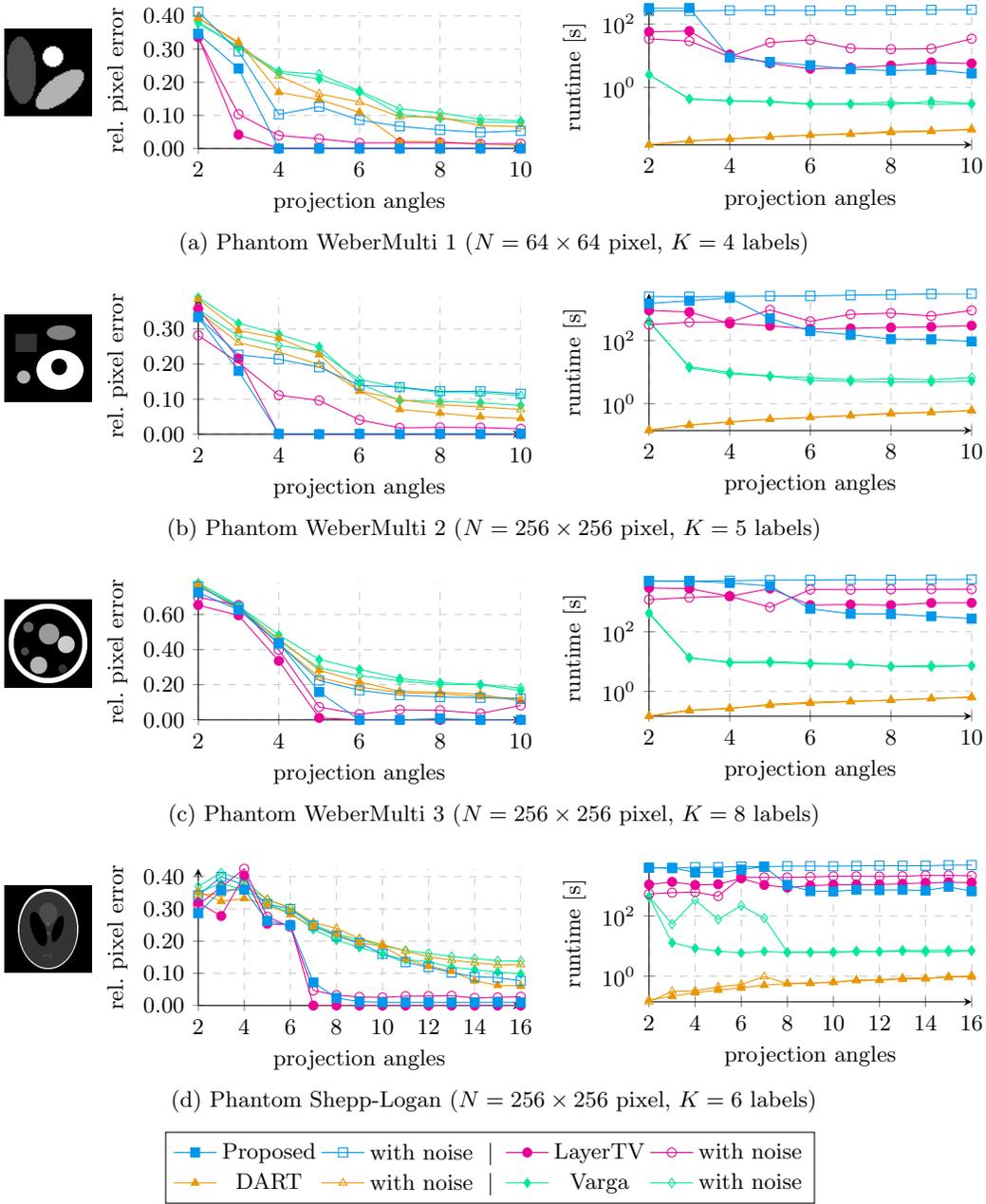


Fig. 1. Evaluation of the approaches for the different test-datasets and increasing (but small) numbers of projections angles, in the noiseless case (**filled markers**) and in the noisy case (**non-filled markers**), noise level $SNR = 20$ dB. The relative pixel error and runtime is displayed. The proposed approach gives perfect reconstructions with a small number of projection angles in the noiseless case and also returns good reconstructions in the presence of noise, compared to the other approaches. The single competing approach, LayerTV, uses a special rounding strategy to obtain meaningful solutions (phantom 3 and 4) and a dedicated data term to cope with Poisson noise.

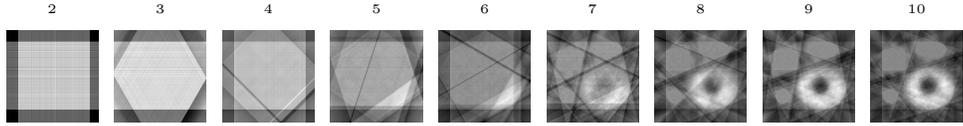


Fig. 2. “Implicit data terms” generated by the tomographic constraints, in terms of the reprojected dual variable $A^T Q$ (scaled to $[0, 1]$ and inverted) after convergence, for WeberMulti 2 and an increasing number of projection angles. The proposed approach achieves a perfect reconstruction from 4 projection angles only. The missing information is effectively compensated by geometric label assignment and spatial coherence due to geometric averaging.

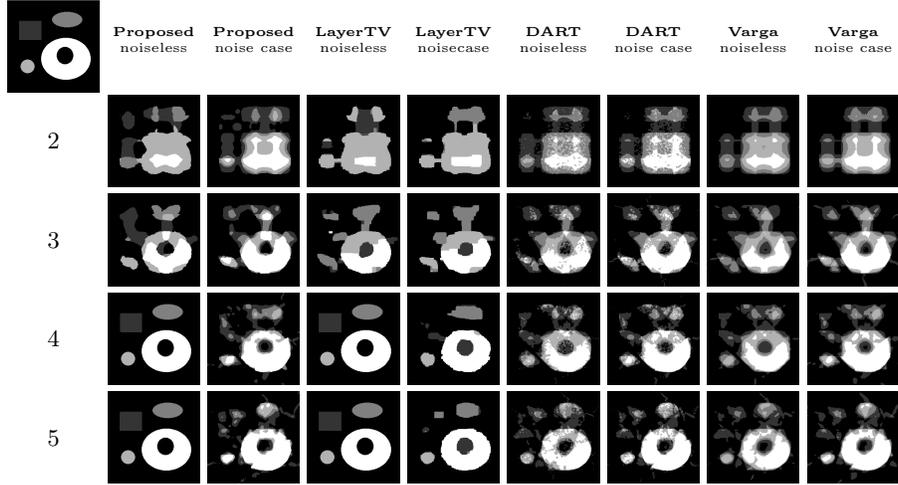


Fig. 3. Experimental results for phantom WeberMulti 2.

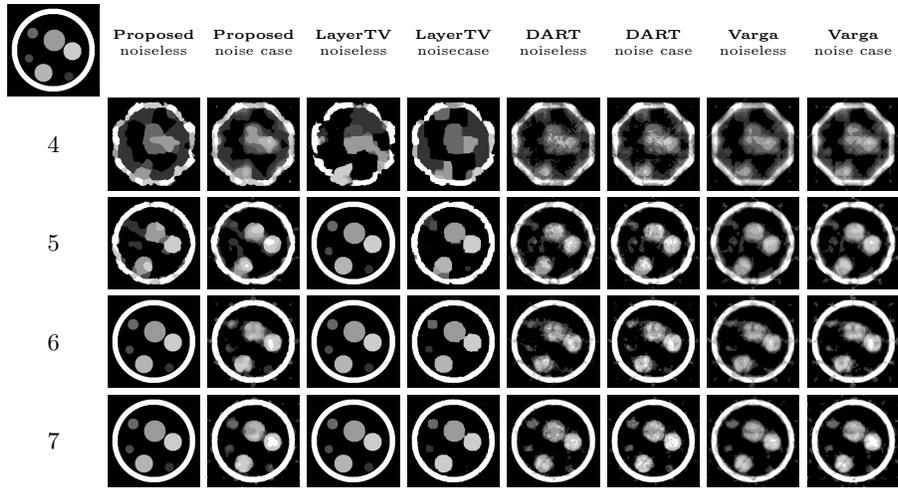


Fig. 4. Experimental results for phantom WeberMulti 3.

References

1. Aarle, W., Palenstijn, W., Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K., Sijbers, J.: The {ASTRA} Toolbox: A Platform for Advanced Algorithm Development in Electron Tomography. *Ultramicroscopy* 157, 35 – 47 (2015)
2. Alvarez, F., Bolte, J., Brahic, O.: Hessian Riemannian Gradient Flows in Convex Programming. *SIAM journal on control and optimization* 43(2), 477–501 (2004)
3. Åström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image Labeling by Assignment. *J. Math. Imaging Vis.* pp. 1–28 (2017), <http://dx.doi.org/10.1007/s10851-016-0702-4>
4. Batenburg, K., Sijbers, J.: DART: A Practical Reconstruction Algorithm for Discrete Tomography. *Image Processing, IEEE Transactions on* 20(9), 2542–2553 (Sept 2011)
5. Burbea, J., Rao, C.R.: Entropy Differential Metric, Distance and Divergence Measures in Probability Spaces: A Unified Approach. *Journal of Multivariate Analysis* 12(4), 575–596 (1982)
6. Chambolle, A., Pock, T.: On the Ergodic Convergence Rates of a First-Order Primal-Dual Algorithm. *Mathematical Programming* 159(1), 253–287 (2016)
7. Denitui, A., Petra, S., Schnörr, C., Schnörr, C.: Phase Transitions and Cospase Tomographic Recovery of Compound Solid Bodies from Few Projections. *Fundamenta Informaticae* 135, 73–102 (2014)
8. Goris, B., Broek, W., Batenburg, K., Mezerji, H., Bals, S.: Electron Tomography Based on a Total Variation Minimization Reconstruction Technique. *Ultramicroscopy* 113, 120 – 130 (2012)
9. Herman, G., Kuba, A.: *Discrete Tomography: Foundations, Algorithms and Applications*. Birkhäuser (1999)
10. Kappes, J.H., Petra, S., Schnörr, C., Zisler, M.: TomoGC: Binary Tomography by Constrained GraphCuts. *Proc. GCPR* 30
11. Minka, T.: Divergence measures and message passing. *Tech. Rep. MSR-TR-2005-173*, Microsoft Research Ltd., Cambridge, UK (2005)
12. Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete Tomography by Convex-Concave Regularization and D.C. Programming. *Discrete Applied Mathematics* 151(13), 229 – 243 (2005)
13. Shepp, L., Logan, B.: The Fourier reconstruction of a head section. *Nuclear Science, IEEE Transactions on* 21(3), 21–43 (June 1974)
14. Sidky, E.Y., Pan, X.: Image Reconstruction in Circular Cone-Beam Computed Tomography by Constrained, Total-Variation Minimization. *Physics in Medicine and Biology* 53(17), 4777 (2008)
15. Storath, M., Weinmann, A., Friel, J., Unser, M.: Joint Image Reconstruction and Segmentation Using the Potts Model. *Inverse Problems* 31(2), 025003 (2015)
16. Varga, L., Balázs, P., Nagy, A.: An Energy Minimization Reconstruction Algorithm for Multivalued Discrete Tomography. In: *3rd International Symposium on Computational Modeling of Objects Represented in Images*, Rome, Italy, Proceedings (Taylor & Francis). pp. 179–185 (2012)
17. Weber, S., Nagy, A., Schüle, T., Schnörr, C., Kuba, A.: A Benchmark Evaluation of Large-Scale Optimization Approaches to Binary Tomography. In: *Discrete Geometry for Computer Imagery (DGCI 2006)*. LNCS, vol. 4245, pp. 146–156. Springer (2006)
18. Weber, S., Schnörr, C., Hornegger, J.: A Linear Programming Relaxation for Binary Tomography with Smoothness Priors. *Electronic Notes in Discrete Mathematics* 12, 243–254 (2003)

19. Zisler, M., Aström, F., Petra, S., Schnörr, C.: Image Reconstruction by Multilabel Propagation. In: Proc. SSVM (in press). LNCS, Springer (2017)
20. Zisler, M., Petra, S., Schnörr, C., Schnörr, C.: Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints. In: Proc. GCPR (2016)
21. Zisler, M., Kappes, J.H., Schnörr, C., Petra, S., Schnörr, C.: Non-Binary Discrete Tomography by Continuous Non-Convex Optimization. IEEE Transactions on Computational Imaging 2(3), 335–347 (2016)