

Non-Binary Discrete Tomography by Continuous Non-Convex Optimization

Matthias Zisler, Jörg Hendrik Kappes, Claudius Schnörr, Stefania Petra, and Christoph Schnörr

Abstract—We study an energy formulation for non-binary discrete tomography and introduce a non-convex coupling term in order to combine discrete constraints with a continuous reconstruction method based on total variation regularization. The optimization is carried out by a generalized forward-backward splitting algorithm for non-convex functions, which exploits the problem structure and is guaranteed to globally converge to a local optimum. A detailed numerical evaluation on standard test-datasets demonstrates that the proposed algorithm returns more accurate reconstructions from a few number of projection angles than competing methods.

Index Terms—Discrete tomography, limited-angle tomography, non-binary, non-convex optimization, reconstruction, relaxation, total variation regularization.

I. INTRODUCTION

A. Overview, Motivation

TOMOGRAPHY is a key imaging method [1]. Applications are widespread from medical imaging [2] to natural sciences, engineering [3] and industry, like non-destructive material testing [4] for quality inspection. In computer tomography (CT), the central task is to reconstruct an object in terms of a volume function that corresponds to the absorption of rays passing through the object. The input data to the reconstruction algorithm are projection images of the object where each pixel corresponds to a line integral along a corresponding ray.

We are particularly interested in scenarios where the number of projections is very small in comparison to the sampling rates established for standard CT reconstruction methods [1]. Such scenarios are motivated by applications that, for instance, enable only limited-angle imaging set-ups or require synchronous recording by several cameras to minimize acquisition time and to eliminate inaccuracies that otherwise would be caused by a moving sensor.

In order to cope with such severely ill-posed reconstruction problems [1], we make the crucial assumption that the range of the volume function is *finite and known*. This assumption effectively restricts the admissible set to piecewise constant func-

tions, which is reasonable in industrial quality inspection where the function values are related to the piecewise homogeneous material properties of the imaged object. Thus, the problem class addressed in this paper belongs to the field of *discrete tomography* [5]. The challenge is to design algorithms that incorporate corresponding constraints into the reconstruction process while avoiding combinatorially complex operations. To this end, we devise and study a continuous non-convex variational approach. Our focus in this paper is rather on the design of the approach and the competitive numerical evaluation using established academic datasets, than on working out a specific application.

B. Related Work

In view of the huge literature on tomographic reconstruction, we limit ourselves to the discussion of few recent, directly related works. We distinguish *binary tomography* and *non-binary discrete tomography*.

Regarding binary tomography, Weber *et al.* [6], [7] suggested a quadratic program together with a non-convex penalty term for gradually enforcing binary functions. These functions were reconstructed by difference of convex functions (DC) programming resulting in a sequence of convex quadratic programs to be solved. The increasing sequence of weights for gradually enforcing the penalty term has to be chosen heuristically, however. The specific case of binary tomography has spurred a lot of research on dedicated algorithms that essentially exploit the fact that each pixel (or voxel) only takes either of two values. Batenburg [8], for instance, used polynomial time network flow solvers for subproblems based on pairwise orthogonal projections directions, yet without proving convergence of the overall scheme. A clever and fast multiscale heuristic is proposed in [9], where a stochastic level-set method is explored in [10]. Recently, Kappes *et al.* [11] proposed a recovery approach based on a discrete graphical model constrained by the projection data. Recovery is performed by a sequence of s-t graph-cuts which gives fast and accurate binary reconstructions from a small number of projections.

The general case of non-binary (multivalued) discrete tomography is considerably more involved because the allowed function values, encoded by unit vectors, span a simplex rather than just a one-dimensional (1D) interval. In view of the piecewise constancy of admissible functions, a natural approach is to consider convex sparsity promoting priors like the ℓ_1 -norm or the total variation (TV) functionals [12], [13] and to round the continuous solution to a piecewise constant one in a postprocessing step. While this approach connects the field of discrete tomography to the fast evolving field of *compressive sensing* and corresponding recovery guarantees [14], the prior

Manuscript received January 08, 2016; revised March 23, 2016; accepted April 20, 2016. Date of publication May 04, 2016; date of current version August 05, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jong Chul Ye.

M. Zisler, J. H. Kappes, S. Petra, and C. Schnörr are with the Heidelberg University, Heidelberg 69117, Germany (e-mail: zisler@math.uni-heidelberg.de; kappes@math.uni-heidelberg.de; petra@math.uni-heidelberg.de; schnoer@math.uni-heidelberg.de).

C. Schnörr is with the Hochschule München, München 80335, Germany (e-mail: schnoerr@cs.hm.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCL.2016.2563321

information in terms of the finite set of function values is not directly used for reconstruction. Regarding non-convex approaches, the authors of [15] proposed a dynamic programming approach for using the ℓ_0 -norm in order to directly enforce sparsity of the gradient. While this non-convex component concerns regularization and necessitates a strategy for avoiding poor local minima, we present a non-convex data term that utilizes prior knowledge about the finites range set of the tomographic reconstruction, without danger of getting trapped in poor local minima as the experiments clearly demonstrate. Using this type of prior knowledge is not possible with the approach [15].

Batenburg and Sijbers [16] proposed a very fast heuristic, the *Discrete Algebraic Reconstruction Technique (DART)*. Starting with a continuous reconstruction by a basic algebraic reconstruction method, a threshold operation is progressively applied to obtain a piecewise constant function, followed by a smoothing operation and performing again a continuous reconstruction, to refine the uncertain boundaries, and further repeating these steps. This approach leads to good results in practice but cannot be characterized by an objective function that is optimized. Moreover, the approach requires considerable parameter tuning, as the authors describe in follow-up work [17]. The latter recent work alleviates these issues by incorporating a TV regularization term along with a scheme for steering the reconstructed function towards discrete values. This combination bears some similarity with the scope of our approach introduced in the present paper. Yet, the objective functions as well as the numerical algorithm are fairly different.

Tuysuzoglu *et al.* [18] suggested a local approximation of a ℓ_2 -fidelity term around an iteratively updated working point in order to cast the non-binary discrete reconstruction problem into a series of submodular binary problems within an α -expansion approach (cf. [19]). This local approximation discards a lot of information, however. As a consequence, a significantly larger number of projections is needed in practice to get reasonable reconstructions.

Maeda *et al.* [20] proposed an alternating optimization approach for maximizing the *a posteriori* probability of a probabilistic model formulation which couples a continuous reconstruction with the Potts model. However, there is no convergence guarantee for the algorithm that alternates a continuous and discrete block coordinate descent.

Varga *et al.* [21] proposed a heuristic algorithm to adaptively combine an energy formulation with a non-convex polynomial in order to steer the solution towards the allowed values. The idea of enforcing integer constraints with a non-convex polynomial penalty is closely related to the DC programming approach [6], [7] in the binary case. Unlike [17], however, the DC programming approach comes along with convergence guarantees.

We regard [16], [21] as state-of-the-art approaches and additionally considered the purely continuous approach [12] for a comparison. We ignored the approach [18] which did not work well unless considerably more projection data was added as input data.

C. Contribution and Organization

We introduce a new coupling term as part of a continuous non-convex variational approach to discrete tomography. This

term resembles the non-convex polynomial proposed by [21] which was motivated as extension of the non-convex approach [6], [7] to non-binary discrete tomography. Our term, however, is derived in a more principled way as relaxation of the minimum distance to the set of admissible values. We explain in detail the rationale behind the design of this coupling term. Furthermore, by introducing an auxiliary variable representing the probability of each admissible value at every pixel, we obtain a procedure for computing the most likely discrete solution based on a local minimum of the continuous approach. The latter is optimized by a non-convex algorithm which exploits the problem structure and is guaranteed to converge to a local optimum. A numerical evaluation on standard test-datasets demonstrates that the proposed approach returns a piecewise constant function which renders any rounding procedure unnecessary. Additionally, these solutions obtained from a small number of projections are more accurate reconstructions in comparison to those obtained by related approaches.

The manuscript is structured as follows. In Section II the non-binary discrete tomography problem is formally defined and the continuous TV reconstruction is briefly reviewed. Next, the coupling term is introduced and discussed, followed by our overall variational approach. We show in Section III how to tackle the resulting non-convex optimization problem by an appropriate algorithm. In Section IV, we report results of numerical experiments and compare to other methods from literature. We conclude in Section V and indicate directions of future research.

II. MODEL FOR JOINT RECONSTRUCTION AND DISCRETE LABELING

We first sketch the combinatorial problem formulation of discrete tomography that motivates our tractable variational approximation introduced and discussed in the subsequent sections.

A. Problem

Discrete tomography is the reconstruction of an image from its projections, which is defined on a bounded region tessellated into pixels (or voxels in 3D) and only takes a few possible intensities as values. Every single projection is a line integral summing up all intensities (absorptions) hit by the projection ray. The *discrete reconstruction problem* amounts to solving a system of linear equations

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

where

- 1) $\mathcal{L} := \{c_1, \dots, c_K\} \subset [0, 1]$ is the set of possible intensities describing the expected CT values for a priori known materials, and $K = |\mathcal{L}|$ is the number of materials;
- 2) u is the vector representing the intensities of the N pixels / voxels of the image to be reconstructed, and $N := n_1 \cdot \dots \cdot n_d$ for images defined on a regular grid graph embedded in \mathbb{R}^d , $d \in \{2, 3\}$;
- 3) $b \in \mathbb{R}^m$ are the measurements or projection values with m denoting the total number of projection rays;

- 4) $A \in \mathbb{R}^{m \times N}$ is a sparse projection matrix, where each entry a_{ij} corresponds to the length of the line segment of the i th projection ray passing through the j th pixel / voxel.

Solving an equation system restricted to few allowed values only is a hard combinatorial problem, however. Even in the binary case, $\mathcal{L} = \{0, 1\}$, the discrete reconstruction problem is NP-hard if the number of projection angles is larger than two [22]. As a consequence, we are interested in a suitable energy formulation which approximates the non-binary discrete tomography reconstruction problem. Our strategy is to introduce an auxiliary variable z which represents the discretization of the reconstructed image with respect to \mathcal{L} . Then the reconstruction u has no longer to be restricted to \mathcal{L} . It suffices to enforce similarity between u and z through a novel coupling term.

B. Variational Reconstruction Approach

Solving solely the projection constraints $Au = b$ does not make sense even without restricting the solution to the set of possible intensities \mathcal{L} . For, the equation system is underdetermined ($m \ll N$) when using only a few projection angles, and noisy measurements take b outside of the range of A , if the projection matrix is rank deficient. Hence, even the ‘‘continuous’’ reconstruction is a severely ill-posed inverse problem.

A basic energy minimization approach is to solve the projection constraints in the least-square sense with additional regularization to ensure spatial coherence of the reconstructed image. Since the image u only consists of a few known materials, u is assumed to be piecewise constant and the gradient ∇u is assumed to be sparse (see [14] for a sufficient uniqueness condition from the compressive sensing viewpoint), where ∇ denotes the discrete gradient with finite differences with Neumann boundary conditions, see [23] for details. The most straightforward sparse gradient regularization is TV , which can be seen as a discrete version of the Rudin-Osher-Fatemi model [24]. This results in the convex energy functional

$$E_{TVL_2}(u) := \frac{1}{2} \|Au - b\|_{l_2}^2 + \lambda \sum_{i=1}^N |(\nabla u)_i| + \delta_{[0,1]^N}(u) \quad (2)$$

where $\lambda > 0$ is the regularization parameter which controls the trade-off between taking into account the measured data and the prior. The norm $|\cdot|$ can be either Euclidean to have isotropic TV or the sum of absolute values for anisotropic TV. Additionally, it has been empirically proven that a non-negativity box constraint $u \in [0, 1]^N$ increases the reconstruction performance. This is accomplished by the indicator function $\delta_{[0,1]^N}(u)$ of the functional (2), where $\delta_C(u)$ for a closed set C is defined as

$$\delta_C(u) := \begin{cases} 0 & \text{if } u \in C \\ \infty & \text{if } u \notin C. \end{cases} \quad (3)$$

C. Discrete Labeling

The constraint $u \in \mathcal{L}^N$, i.e. the restriction of the image to the finite set of possible values, renders the reconstruction problem intractable. To alleviate this problem, we next introduce and discuss a reformulation based on additional auxiliary variables in order to gradually steer the solution towards the feasible set \mathcal{L}^N of discrete labelings.

Our starting point is the minimum squared distance to measure closeness of u to \mathcal{L}^N

$$d(u, \mathcal{L}) := \min_{c \in \mathcal{L}^N} \|u - c\|_{l_2}^2 = \sum_{i=1}^N \min_{c_k \in \mathcal{L}} (u_i - c_k)^2. \quad (4)$$

We introduce an auxiliary variable z supposed to represent the discrete version of u with respect to \mathcal{L}^N . Each value of \mathcal{L} is encoded by a corresponding unit vector in terms of z . Thus, the admissible set for z is given by

$$\mathcal{S}_0 := \left\{ z \in \{0, 1\}^{N \times K} : \sum_{k=1}^K z_{ik} = 1, \forall i = 1, \dots, N \right\}. \quad (5)$$

We rewrite the distance measure (4) accordingly

$$d_{\mathcal{S}_0}^p(u, \mathcal{L}) = \min_{z \in \mathcal{S}_0} D^p(u, z) \quad (6a)$$

$$D^p(u, z) := \sum_{k=1}^K \sum_{i=1}^N z_{ik}^p (u_i - c_k)^2 \quad \text{with } p \in \{1, 2\} \quad (6b)$$

where the minimizer z of the coupling term $D^p(u, z)$, for a given fixed u , can be interpreted as a projection of u onto the set \mathcal{L}^N .

An important aspect of the coupling term that will be discussed in more detail below, is a *tight and smooth* interplay between u and z . To enable this, we avoid hard decisions in terms of the integer-valued variable z and therefore relax the condition $z \in \{0, 1\}^{N \times K}$ to $z \in [0, 1]^{N \times K}$. Accordingly, the feasible set \mathcal{S}_0 of eq. (5) is replaced by its convex hull that we denote by \mathcal{S} . Now, for every pixel i the vector $(z_{ik})_{k=1}^K$ is a point of the standard simplex and hence can be interpreted as the probability that u_i takes the value $c_k \in \mathcal{L}$.

Fig. 1(a) shows the effect of the relaxation of \mathcal{S}_0 to \mathcal{S} on the distance given by the left-hand side of (6a), for both $p = 1$ and $p = 2$ and the label set $\mathcal{L} = \{0.0, 0.4, 1.0\}$. We first observe that the distance values do not change whenever z ranges over the original feasible set \mathcal{S}_0 : $d_{\mathcal{S}_0}^1(u, \mathcal{L}) = d_{\mathcal{S}_0}^2(u, \mathcal{L})$. In the relaxed case, we still have

$$d_{\mathcal{S}_0}^p(u, \mathcal{L}) = d_{\mathcal{S}}^1(u, \mathcal{L}) \quad (7)$$

for $p = 1, 2$ (blue curve in Fig. 1(a)), whereas $d_{\mathcal{S}}^2(u, \mathcal{L})$ is a smoothed version thereof (red curve in Fig. 1(a)).

Fig. 1(b) and (c) show how the variables $z = (z_1, \dots, z_K)^\top$, $K = |\mathcal{L}| = 3$, affect these distances. The plots display the relaxed distances $d_{\mathcal{S}}^p(u(z), \mathcal{L})$, $p = 1, 2$, as z ranges over \mathcal{S} (rather than \mathcal{S}_0) on the right-hand side of (6a). The three vertices correspond to the unique decisions for three labels $c_k \in \mathcal{L}$, $k = 1, 2, 3$. Close to these vertices, the smoothing effect from panel (a) is visible in (c), in comparison to (b). And both versions (b) ($p = 1$) and (c) ($p = 2$) create high costs for switching from the decision c_1 to c_3 and vice versa, that are visible as red peaks.

In summary, the relaxed distance $d_{\mathcal{S}}^2(u, \mathcal{L})$ exhibits two favourable properties: It is exact at values of u determined by the vertices of \mathcal{S} , and it is smooth in contrast to the piecewise quadratic distance $d_{\mathcal{S}_0}^2(u, \mathcal{L})$ and the non-relaxed distances $d_{\mathcal{S}_0}^p(u, \mathcal{L})$, $p = 1, 2$. We will exploit this smooth coupling of u and z for numerical optimization in order to compute high-quality tomographic restorations u .

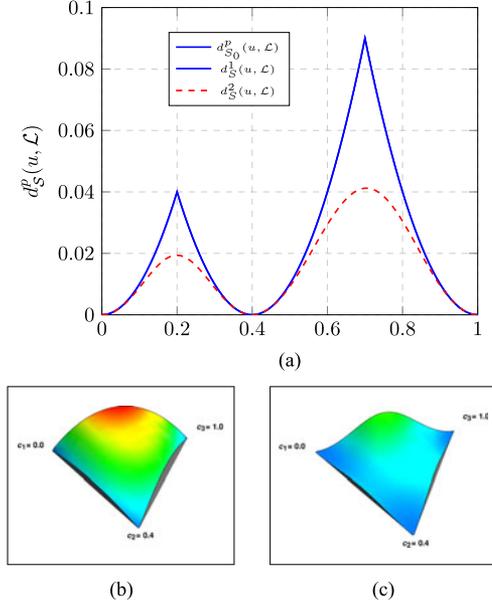


Fig. 1. The distance functions $d_{S_0}^p(u, \mathcal{L})$ and $d_S^p(u, \mathcal{L})$ for the label set $\mathcal{L} = \{0.0, 0.4, 1.0\}$, before and after relaxation, respectively. Panel (a) illustrates that the relaxation d_S^1 is tight whereas d_S^2 is a smooth version thereof. Panels (b) and (c) show the dependency of the relaxed distance functions on z , as z ranges over \mathcal{S} on the right-hand side of (6a). We exploit in this paper the relaxed distance function d_S^2 that is exact if $u \in \mathcal{L}$ and smooth in the entire region. (a) The blue curve shows the distances $d_{S_0}^1(u, \mathcal{L}) = d_{S_0}^2(u, \mathcal{L}) = d_S^1(u, \mathcal{L})$ that all coincide, whereas the relaxation $d_S^2(u, \mathcal{L})$ with $p = 2$ results in the smoothed distance function displayed as red curve. (b) The relaxed distance function $d_S^2(u(z), \mathcal{L})$ as z ranges over \mathcal{S} , with $uz = (z, c)$. (c) The relaxed distance function $d_S^2(u(z), \mathcal{L})$ as z ranges over \mathcal{S} , with $uz = \frac{1}{\sum_{j=1}^K z_j^2} \sum_{k=1}^K z_k^2 c_k$.

Accordingly, based on (6b) we define the mapping

$$z = z(u), \quad z : [0, 1]^N \rightarrow [0, 1]^{N \times K}, \quad u \mapsto \arg \min_{z \in \mathcal{S}} D^2(u, z). \quad (8)$$

This function can be computed in closed form

$$z(u)_{il} := \begin{cases} 0, & \text{if } u_i = c_k \wedge l \neq k, \\ 1, & \text{if } u_i = c_k \wedge l = k, \\ \left(\sum_{j=1}^K \frac{1}{(u_i - c_j)^2} (u_i - c_l)^2 \right)^{-1} & \text{otherwise,} \end{cases} \quad (9)$$

which results from the simplex constraint $z \in \mathcal{S}$ and the Fermat's (first order) optimality condition. Inserting this function into $D^2(u, z)$ and fixing the second argument at some point u_0 yields locally an upper quadratic envelope $D^2(u, z(u_0))$ of the function $(u, z) \mapsto D^2(u, z)$ – see Fig. 2 for an illustration. Minimizing this local second-order approximation steers u towards the value in \mathcal{L} that is closest to u_0 .

In order to further elucidate the coupling of u and z through $D^2(u, z)$, we also visualize the functions $z \mapsto D^2(u_0, z)$ for fixed u_0 over the probability simplex $z \in \mathcal{S}$. This is shown in Fig. 3 for $u_0 \in \{0.4, 0.5, 0.6\}$, where the black dot in each plot shows the minimizer $z(u_0)$ of $D^2(u_0, z)$. Analogous to the local quadratic upper envelopes $u \mapsto D^2(u, z(u_0))$ parametrized by u_0 , the functions $z \mapsto D^2(u_0, z)$ also constitute local quadratic

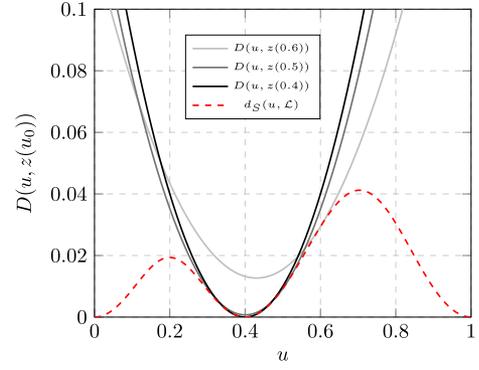


Fig. 2. Quadratic upper envelopes $D^2(u, z(u_0))$ of the distance $D^2(u, z)$ parametrized by $u_0, \mathcal{L} = \{0.0, 0.4, 1.0\}$ and $u_0 \in \{0.4, 0.5, 0.6\}$. Minimizing these envelopes steers u towards the value in \mathcal{L} that is closest to u_0 .

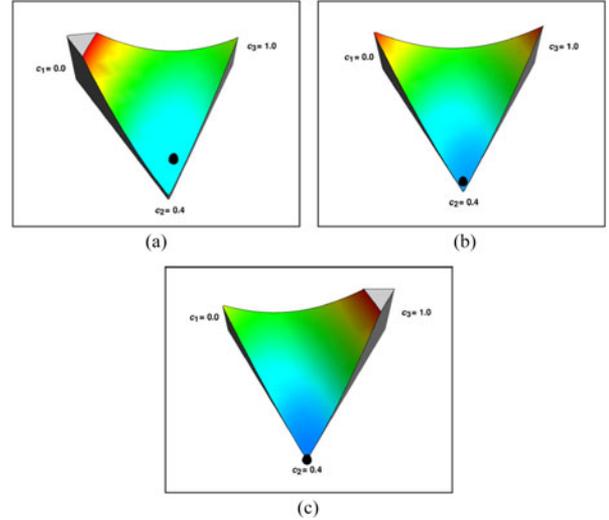


Fig. 3. Visualization of the coupling term $D^2(u_0, z)$ for fixed u_0 over the probability simplex $z \in \mathcal{S}$, where the vertices correspond to the values $\mathcal{L} = \{0.0, 0.4, 1.0\}$. The black dot in each plot marks the minimizer $z(u_0)$ of $D^2(u_0, z)$. The functions $z \mapsto D^2(u_0, z)$ parametrized by u_0 constitute local upper quadratic envelopes of the function $d_S^2(u, \mathcal{L})$ shown in Fig. 1(c). Minimizing these envelopes locally steers z towards the vertices of the simplex. (a) $u_0 = 0.6$. (b) $u_0 = 0.5$. (c) $u_0 = 0.4$.

upper envelopes parametrized by u_0 of the function $d_S^2(u, \mathcal{L})$ shown in Fig. 1(c), which is steering z towards the vertices of the simplex. We observe that if and only if $u_0 \in \mathcal{L}$, then the optimal z is a unit vector.

Finally we combine the coupling $D^2(u, z)$, given by Eq. (6b), and the continuous reconstruction energy $E_{\text{TVL}_2}(u)$, given by Eq. (2). This results in the joint minimization problem with respect to (u, z)

$$E_{\text{Joint}}(u, z) := \frac{1}{2} \|Au - b\|_{l_2}^2 + \lambda \sum_{i=1}^N |(\nabla u)_i| + \delta_{[0,1]^N}(u) + \frac{\alpha}{2} \sum_{k=1}^K \sum_{i=1}^N z_{ik}^2 (u_i - c_k)^2 + \delta_{\mathcal{S}}(z) \quad (10)$$

TABLE I
OVERVIEW OF APPROACHES USED FOR A COMPETITIVE EVALUATION

Shortcut	Reference	Regularization	Implementation
TV- L_2	Section II-B, [12]	anisotropic TV	ours, with PD [23]
Joint	Section II-C	anisotropic TV	ours, with PALM [25]
DART	[16]	none	ASTRA-toolbox [26]
Varga	[21]	Sobolev semi-norm	ours

TABLE II
PARAMETER SETTINGS FOR THE APPROACHES USED IN THE NUMERICAL
EVALUATION OF THE THREE DIFFERENT TEST-DATASETS

Approach	TV- L_2	Joint	Varga [21]			DART [16]			
Parameters	λ	λ	α	α	μ	σ	P_{fix}	r	it_{SART}
Phantom 1: Maple Leaf	0.003	0.003	0.32	0.1	20	1.0	0.85	0.1	3
Phantom 2: Shepp-Logan	0.1	0.1	0.8	0.5	20	1.0	0.85	0.1	3
Phantom 3: Batenburg	0.2	0.2	0.8	1.0	20	1.0	0.85	1.0	3

For a varying number of projection angles, we always used the same parameter values.

with weighting parameter α and the indicator function $\delta_S(z)$ constrains the vectors $(z_{ik})_{k=1}^K$ to the standard simplex at every pixel i .

Of course, our relaxed model $E_{\text{Joint}}(u, z)$ will not deliver a discrete solution with respect to the set of feasible intensities \mathcal{L} . The most straightforward remedy is to round the continuous variable u to the closest value in \mathcal{L} at every pixel. Likewise, the 1-of- K coding variables $(z_{ik})_{k=1}^K$ could be rounded so as to choose the most likely value at every pixel. Our numerical evaluation reported in Section IV however revealed no rounding of z was necessary as long as α was chosen large enough. That is, due to the non-convex component of our approach discussed above, our approach returned in each experiment a unique discrete decision at each pixel.

Although each mapping $u \mapsto E_{\text{Joint}}(u, z)$ and $z \mapsto E_{\text{Joint}}(u, z)$ is convex, the joint mapping $(u, z) \mapsto E_{\text{Joint}}(u, z)$ is non-convex because of the term $D^2(u, z)$. While further convex relaxation of this non-convex term could be done, following McCormick's [McCormick1976] relaxation of products of two functions, for instance, or some recent extensions [21] could be applied to relax the term $z_{ij}^2(u_i - c_k)^2$, the coupling effect then would become too weak.

Therefore, in the next section a non-convex optimization algorithm is adopted which exploits the structure of the proposed objective $E_{\text{Joint}}(u, z)$.

III. NON-CONVEX OPTIMIZATION

A. Proximal Alternating Linearized Minimization (PALM)

The authors of [25] propose the PALM algorithm for non-convex, non-smooth problems. Their algorithm is based on a proximal regularization of alternating minimizations combined with a forward-backward splitting. Hence the objective function is assumed to have the structure

$$\psi(u, z) := f(u) + g(z) + H(u, z) \quad (11)$$

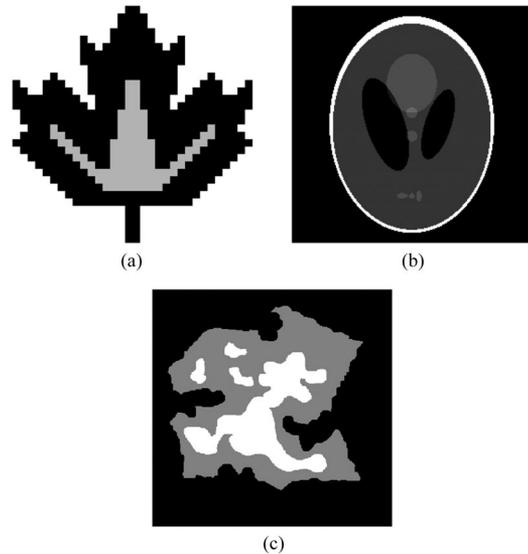


Fig. 4. The different phantom images which are used for the experiments. (a) Phantom 1: Maple Leaf $N = 32 \times 32$, $K = 3$. (b) Phantom 2: Shepp-Logan $N = 256 \times 256$, $K = 6$. (c) Phantom 3: Batenburg $N = 512 \times 512$, $K = 3$.

Algorithm 1: PALM: Proximal Alternating Linearized Minimization.

1. Initialization: start with any $(u^0, z^0) \in \mathbb{R}^n \times \mathbb{R}^m$.
2. For each $k = 0, 1, \dots$ generate a sequence $\{(u^k, z^k)\}_{k \in \mathbb{N}}$ as follows:

- 2.1 Take $\gamma_1 \geq 1$, set $\tau_k = \gamma_1 \cdot L_1(z^k)$ and compute

$$u^{k+1} \in \text{prox}_{\frac{f}{\tau_k}} \left(u^k - \frac{1}{\tau_k} \nabla_u H(u^k, z^k) \right). \quad (13)$$

- 2.2 Take $\gamma_2 \geq 1$, set $\sigma_k = \gamma_2 \cdot L_2(u^{k+1})$ and compute

$$z^{k+1} \in \text{prox}_{\frac{g}{\sigma_k}} \left(z^k - \frac{1}{\sigma_k} \nabla_z H(u^{k+1}, z^k) \right). \quad (14)$$

that suggests to decompose the problem into two block-coordinates u and z and into smooth and non-smooth components given by H and f, g functions, respectively. The analysis of the PALM algorithm relies on the Kurdyka-Łojasiewicz (KL) property of the objective function ψ and ensures global convergence to a critical point. This means, the algorithm converges to some critical point regardless of the initialization, but the critical point depends on the initialization. Since the abstract concept of KL functions is only needed for the proof of convergence of PALM, it is not reviewed herein – see [25] for the exact definition and the proof. In addition, the functions have to satisfy the following properties:

- 1) $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ and $g: \mathbb{R}^m \rightarrow (-\infty, +\infty]$ are proper and lower semicontinuous;
- 2) f, g and ψ are bounded from below;
- 3) the coupling function $H: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is twice continuous differentiable, $H \in C^2(\mathbb{R}^n \times \mathbb{R}^m)$;
- 4) the gradient in each block-coordinate of $H(u, z)$ is globally Lipschitz continuous, $u \mapsto H(u, z) \in C_{L_1(z)}^{1,1}(\mathbb{R}^n)$

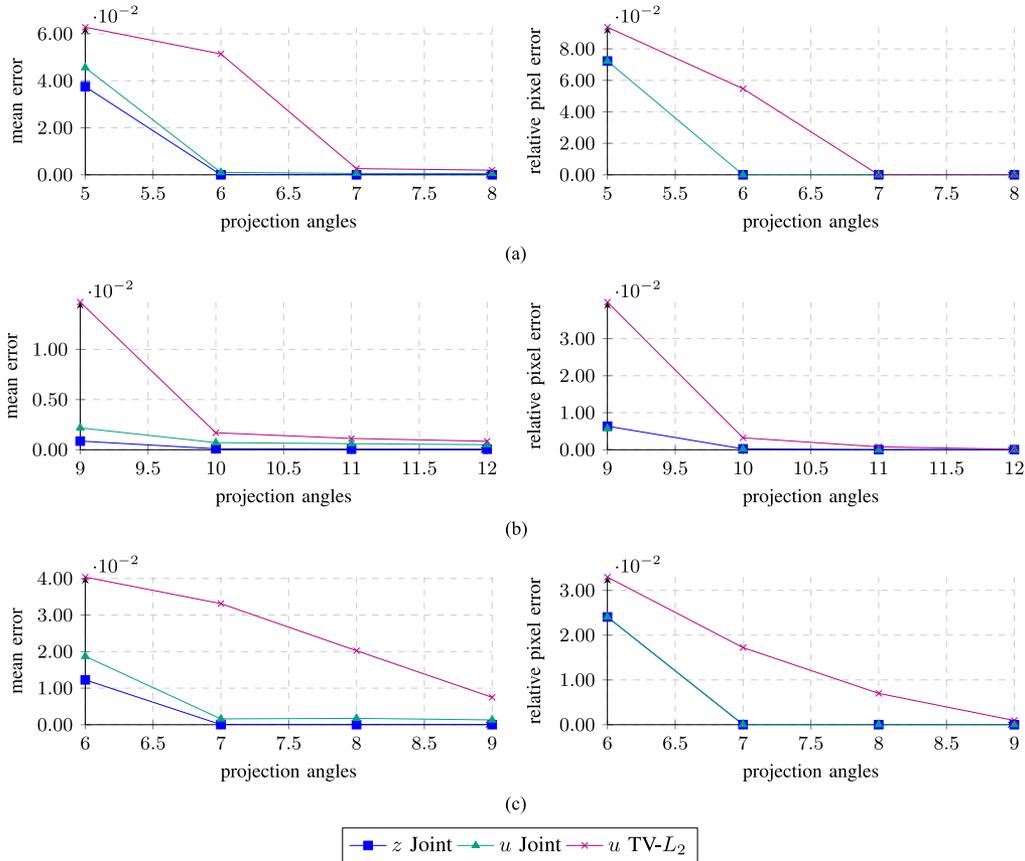


Fig. 5. Numerical evaluation of the TV- L_2 and the Joint approach for the different test-datasets (noiseless case) with a small number of projections in the noiseless case. For both algorithms, the mean error measure was computed for non-rounded reconstructions and the relative pixel error after rounding. The proposed Joint method needs at least one projection angle less to achieve exact reconstruction even in this severely underdetermined scenarios. This clearly demonstrates the beneficial effect of coupling the variables (u, z) by the non-convex coupling term. (a) Phantom 1: Maple Leaf. (b) Phantom 2: Shepp-Logan. (c) Phantom 3: Batenburg.

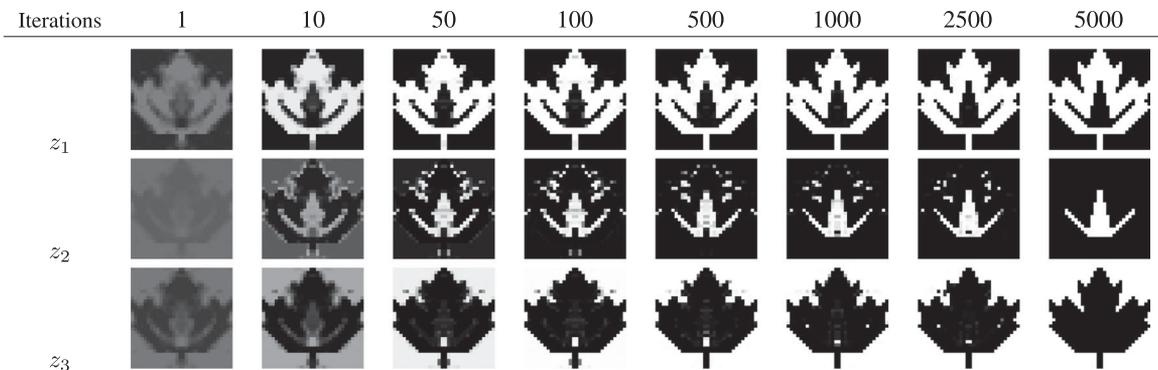


Fig. 6. Evolution of the relaxed 1-of-K coded z_k variables on the test-dataset phantom 1 for 6 projection angles. Each row shows one of the layers z_1, z_2 and z_3 representing the probability for setting a pixel to the values $\mathcal{L} = \{0.0, 0.69, 1.0\}$. Black corresponds to 0 and white to the value 1. We observe that the Joint approach converges without rounding to an exact and binary solution for each z_k .

and $z \mapsto H(u, z) \in C_{L_2(u)}^{1,1}(\mathbb{R}^m)$, where $L_1(z)$ and $L_2(u)$ denote the Lipschitz constants depending on z and u respectively;

5) ψ is a KL function.

The proposed Algorithm 1 alternately updates the two block-coordinates (u, z) where $\text{prox}_\tau^f(p)$ denotes the proximal

operator defined by

$$\text{prox}_\tau^f(p) := \underset{u}{\text{argmin}} \frac{1}{2\tau} \|u - p\|_{l_2}^2 + f(u) \quad (12)$$

and ∇_u, ∇_z denote the partial gradient with respect to block-coordinates u and z .

We note that the conditions $\gamma_i \geq 1$, $i = 1, 2$ differ from the conditions $\gamma_i > 1$, $i = 1, 2$ from the original PALM algorithm [25]. This is due to the convexity of the functions f, g in our case, in which the choices $\gamma_i = 1$, $i = 1, 2$ are also admissible [25, Remark 4, (iv)].

The structure of the objective function (11) which the PALM algorithm minimizes, is well suited for the joint CT reconstruction and discretization energy formulation $E_{\text{Joint}}(u, z)$ given by (10). More precisely, we can identify the corresponding components (11) of the objective function by setting

$$f(u) := \frac{1}{2} \|Au - b\|_{l_2}^2 + \lambda \sum_{i=1}^N |(\nabla u)_i| + \delta_{[0,1]^N}(u), \quad (15)$$

$$g(z) := \delta_{\mathcal{S}}(z), \quad (16)$$

$$H(u, z) := \frac{\alpha}{2} \sum_{k=1}^K \sum_{i=1}^N z_{ik}^2 (u_i - c_k)^2. \quad (17)$$

According to Examples 2 and 4 of [25], the proposed model $E_{\text{Joint}}(u, z)$ (10) enjoys the KL property because it is a sum and composition of norms, real polynomials and indicator functions of semi-algebraic sets. One easily checks that the other requirements (1)–(4) stated above are also fulfilled.

We next compute the partial gradients of the coupling term $H(u, z)$ Eq. (17) and estimate the Lipschitz constants $L_1(z)$ and $L_2(u)$ with trivial assumptions $z \in \mathcal{S}$ and $u_i, c_k \in [0, 1]$

$$(\nabla_u H(u, z))_i = \alpha \sum_{k=1}^K z_{ik}^2 (u_i - c_k) \quad (18)$$

$$\Rightarrow L_1(z) = \alpha \max_i \sum_{k=1}^K z_{ik}^2 \leq \alpha \cdot 1 \quad (19)$$

$$(\nabla_z H(u, z))_{ik} = \alpha z_{ik} (u_i - c_k)^2 \quad (20)$$

$$\Rightarrow L_2(u) = \alpha \max_{i,k} (u_i - c_k)^2 \leq \alpha \cdot 1. \quad (21)$$

Here we can see that if the linear version $p = 1$ of the coupling term $D^p(u, z)$ (6b) would be used with z_{ik} instead of z_{ik}^2 , then the gradient $\nabla_z H(u, z)$ would be constant with respect to z , and thus the Lipschitz constant would be zero and we would not get a proper estimate on the stepsize for minimizing with respect to z . Therefore we consider the quadratic version, $p = 2$, of the coupling term $D^p(u, z)$ (6b).

Finally, the PALM Algorithm 1 requires to attain a minimizer in each subproblem. Thus, we have to evaluate the following two proximal operators

$$\begin{aligned} \text{prox}_{\frac{f}{\tau_k}}^f(p) &= \underset{u}{\text{argmin}} \frac{\tau_k}{2} \|u - p\|_{l_2}^2 + \frac{1}{2} \|Au - b\|_{l_2}^2 \\ &\quad + \lambda \sum_{i=1}^N |(\nabla u)_i| + \delta_{[0,1]^N}(u), \end{aligned} \quad (22)$$

$$\text{prox}_{\frac{g}{\sigma_k}}^g(q) = \underset{z}{\text{argmin}} \frac{\sigma_k}{2} \|z - q\|_{l_2}^2 + \delta_{\mathcal{S}}(z). \quad (23)$$

Exact evaluation is only possible for simple proximal mappings whereas more complicated proximal mappings can only be solved approximately by an iterative algorithm. The second

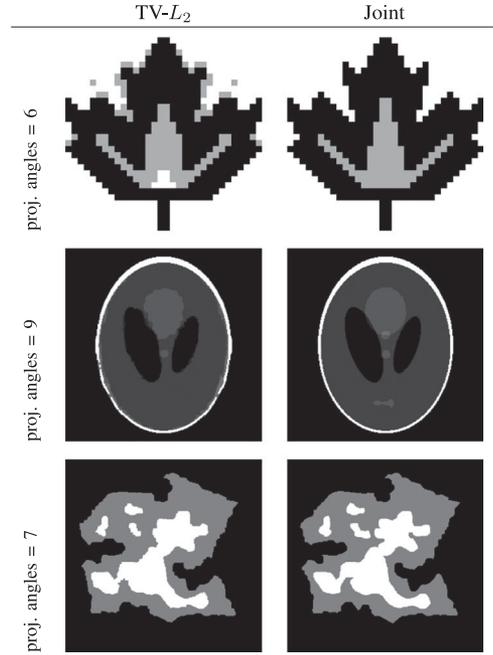


Fig. 7. Demonstrating the improvement of reconstruction performance by the Joint reconstruction and discretization approach in comparison to $\text{TV-}L_2$. From left to right the three different phantoms are shown. The (left) column shows the rounded output u of $\text{TV-}L_2$ and the (right) column the output of z from our proposed approach. The increased accuracy of reconstructions computed using the Joint approach is clearly visible.

proximal mapping of $g(z)$, Eq. (23), can be evaluated exactly: it is the projection onto the standard simplex \mathcal{S} . However, the first proximal mapping of $f(z)$, Eq. (22), cannot be evaluated in closed form. We suggest an approximative evaluation using a standard primal dual (PD) algorithm [23], since the subproblem is convex.

IV. NUMERICAL EXPERIMENTS

In this section, we demonstrate our approach for joint CT reconstruction and discretization to the set of feasible intensities $\mathcal{L} = \{c_1, \dots, c_K\}$ by numerical experiments and compare it to other approaches.

A. Evaluation

1) *Procedure*: We focused on two different aspects in order to compare our new approach.

The first one concerns the evaluation of the plain $\text{TV-}L_2$ model, see Eq. (2), and the proposed joint approach as defined by $E_{\text{Joint}}(u, z)$ in Eq. (10), henceforth short: *Joint*. Both models are based on TV regularization where we used the anisotropic version of TV. One may expect that the joint reconstruction shows reconstruction performance better than $\text{TV-}L_2$, as it enhances the $\text{TV-}L_2$ model by the discretization term which thus incorporates further prior knowledge. For the $\text{TV-}L_2$ model we used our own MATLAB implementation based on the PD algorithm, which returns a continuous output for u . In order to get

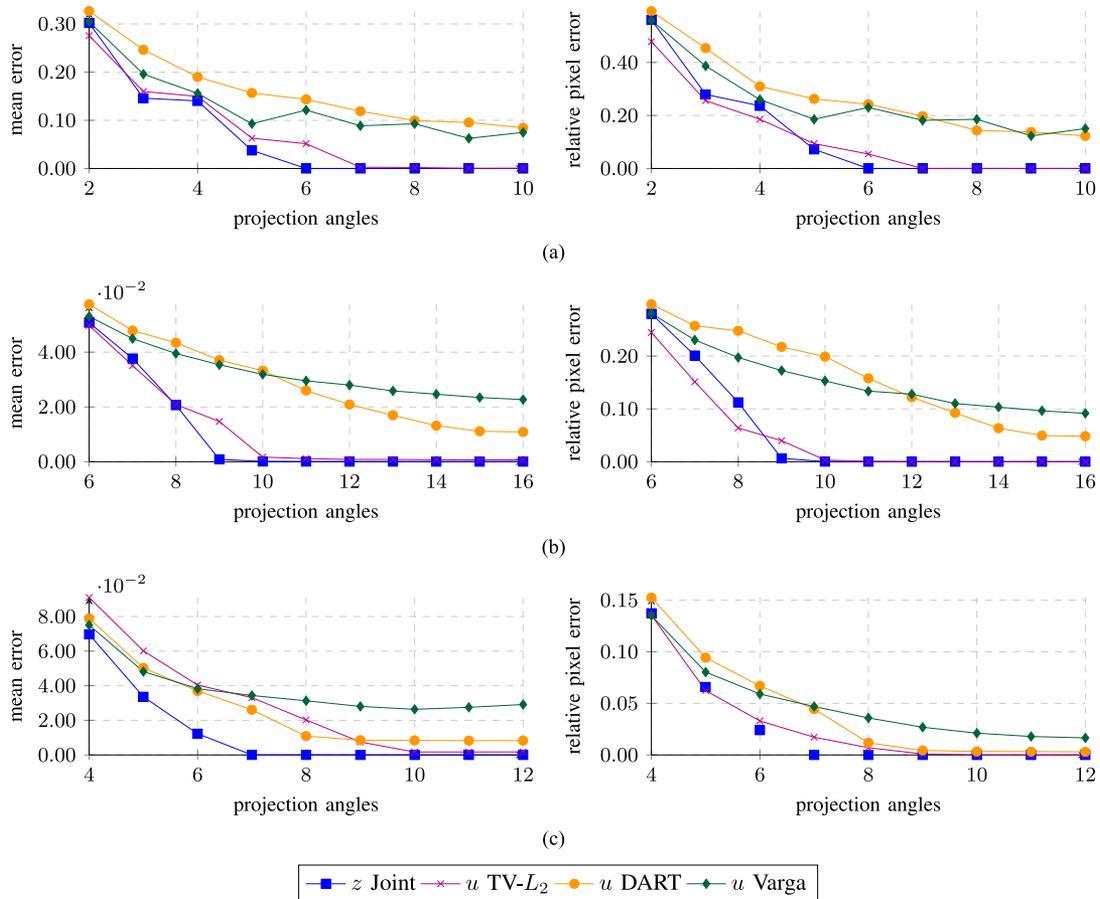


Fig. 8. The plots are showing the numerical evaluation of the approaches for the different test-datasets with a small number of projections in the noise free case. The mean error measure was applied on non-rounded and the relative pixel error on rounded outputs of the algorithms. Our proposed Joint approach has the most accurate reconstruction quality even for fewer projection angles compared to the other methods. (a) Phantom 1: Maple Leaf. (b) Phantom 2: Shepp-Logan. (c) Phantom 3: Batenburg.

a discrete solution, we rounded the continuous solution at each pixel to nearest value of the set of allowed intensities.

We furthermore implemented in MATLAB the proposed Joint approach with the PALM algorithm, as described in Section III. Since the Joint approach is a relaxation, the output of the algorithm may not necessarily be discrete and we also applied rounding. This was done on both coupled variables (u, z) , where for the continuous variable u we rounded to the nearest value of the set \mathcal{L} . The 1-of- K coded variables $(z_{ik})_{k=1}^K$ were rounded to the most likely value at every pixel i , i.e. to the unit vector corresponding to the maximal z_i -component. It turned out, however, that in all experiments we obtained a binary solution (up to numerical precision) for the variables $(z_{ik})_{k=1}^K$, so rounding was not necessary for the proposed joint approach.

The second aspect concerns a comparison of our Joint approach to methods proposed in literature which can handle non-binary tomography for the few-projection limited-angle scenario. Table I lists all considered approaches. We compared to the DART [16] with the publicly available implementation included in the ASTRA-toolbox [26], and we used the SART algorithm for continuous iterative algebraic reconstruction. Regarding the method of Varga [21], we used our own imple-

mentation in MATLAB, because no public implementation was available.

We tried to use the default parameters of the different algorithms as proposed by their authors. However, since the test-datasets sizes differ, we adjusted the parameters to get best results for every algorithm. See Table II for a list of all used parameters. They were only adjusted for different test-datasets but were kept constant for varying numbers of projection angles.

The DART has the parameters P_{fix} as fixed probability, r as radius for the smoothing mask, and it_{SART} which is the number of SART iterations per DART iteration. As termination criterion for DART we stopped after 200 iterations as the suggested in [16].

The approach of Varga [21] involves the parameters α as regularization factor, and μ and σ are controlling the discretization strength. We set as stopping criterion the maximum number of iterations to 5000 and the tolerance between two iterations to $\epsilon = 0.001$.

Finally, for the algorithms TV- L_2 and the proposed Joint approach, we set the maximum number of iterations to 10 000 or terminated if the mean error distance based on (25) between two iterates dropped below $\epsilon = 10^{-6}$. We always used the same

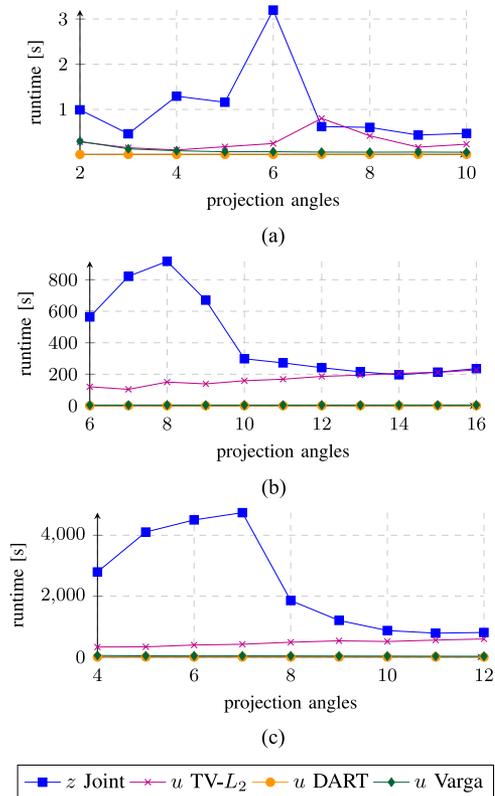


Fig. 9. The plots are showing runtime in seconds needed for the different test-datasets in the noiseless case. DART and the approach of Varga are much faster than the TV- L_2 and the Joint approach. However, the proposed Joint approach returns a fully discretized and more accurate solution, whereas the other approaches stop earlier and have to apply rounding. (a) Phantom 1: Maple Leaf. (b) Phantom 2: Shepp-Logan. (c) Phantom 3: Batenburg.

regularization parameter value λ for the proposed Joint method which was optimal for the alternative TV- L_2 approach. This enabled to directly assess the improvement by the discretization term of the Joint approach. Additionally, the Joint approach involves the parameter α as weight for the coupling term, which enforces discretization. We initialized throughout the variable $u = 0$ and the variable z constant to $1/K$ in order to remove the dependency of the result of the Joint approach from the initialization.

2) *Data Set-Up*: We used the parallel beam setup for our experimental evaluation. The projection matrices were generated using the ASTRA-toolbox [26] with equidistant projection angles between 0 and 180° . The resolution of the sensor-array was chosen 1.5 times larger than the image size, so that every pixel intersects with a projection ray and each sensor has the corresponding pixel size. Each entry of the projection matrix a_{ij} corresponds to the length of the line segment of the i th projection ray passing through the j th pixel. The measurements (projection data) are the line integrals along the projection rays.

Fig. 4 shows the non-binary test-datasets used for the numerical evaluation. The first image 4(a) was taken from [27]. Image 4(b) is the well-known Shepp-Logan phantom, and image 4(c) was taken from the test-datasets of Batenburg and Sijbers [16].

3) *Evaluation Measures*: We used two measures for the evaluation process. The first one is the relative pixel error, that is the relative number of erroneously reconstructed pixels as compared to groundtruth u^*

$$Err_{\text{pxl}}(u) := \frac{1}{N} \sum_i d(u_i, u_i^*); \quad d(x, y) := \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases} \quad (24)$$

This measure was only used for discrete solutions of the algorithms.

As a second error measure we use the mean error, defined as average of the absolute error at every pixel

$$Err_{\text{mean}}(u) := \frac{1}{N} \sum_i |u_i - u_i^*|. \quad (25)$$

This measure was applied to the non-discrete solutions of the algorithms before rounding. Using these two error measures, we quantified the effect of the final rounding step.

B. Numerical Results

We next report experimental results pertaining to the two major aspects of our evaluation, as discussed in Section IV-A1.

1) *Joint Reconstruction versus TV- L_2 Approach*: Fig. 5 shows the evaluation of the two error measures for the different test-datasets with increasing number of projection angles in the noiseless case. On the left are the plots of the mean error, which was computed for the non-rounded outputs, whereas on the right the relative pixel error is displayed, computed for the rounded outputs.

The first row 5(a) shows that the proposed approach needs six projection angles, that is one angle less than TV- L_2 , to return a perfect reconstruction of Phantom 1. The more complex Phantom 2, second row 5(b), can be reconstructed exactly with ten projection angles using the Joint approach, whereas TV- L_2 needs 12 angles. Phantom 3, last row 5(c), is fully reconstructed already with seven angles by the proposed approach, whereas TV- L_2 needs at least nine projection angles.

Furthermore, we noticed that the Joint approach returns already binary output for the 1-of- K coded variables $(z_{ik})_{k=1}^K$ (up to numerical precision) in these examples, thus making obsolete rounding on z . This can also be seen in Fig. 6, where the evolution of the non-rounded z_k is shown on Phantom 1 with six projection angles for different iterations.

Finally, Fig. 7 shows the visual improvement of the reconstruction performance by the Joint approach in comparison to TV- L_2 . The first column shows the rounded results u of TV- L_2 , and the second column shows the output z of the proposed method, for a fixed number of projection angles of the three different phantoms. The results clearly demonstrated the increased accuracy achieved by the Joint approach. In summary, the proposed approach needs at least one projection angle less than the TV- L_2 approach. This is due to the additional prior knowledge in terms of the set of allowed intensities \mathcal{L} , that the Joint approach is able to exploit during the reconstruction process.

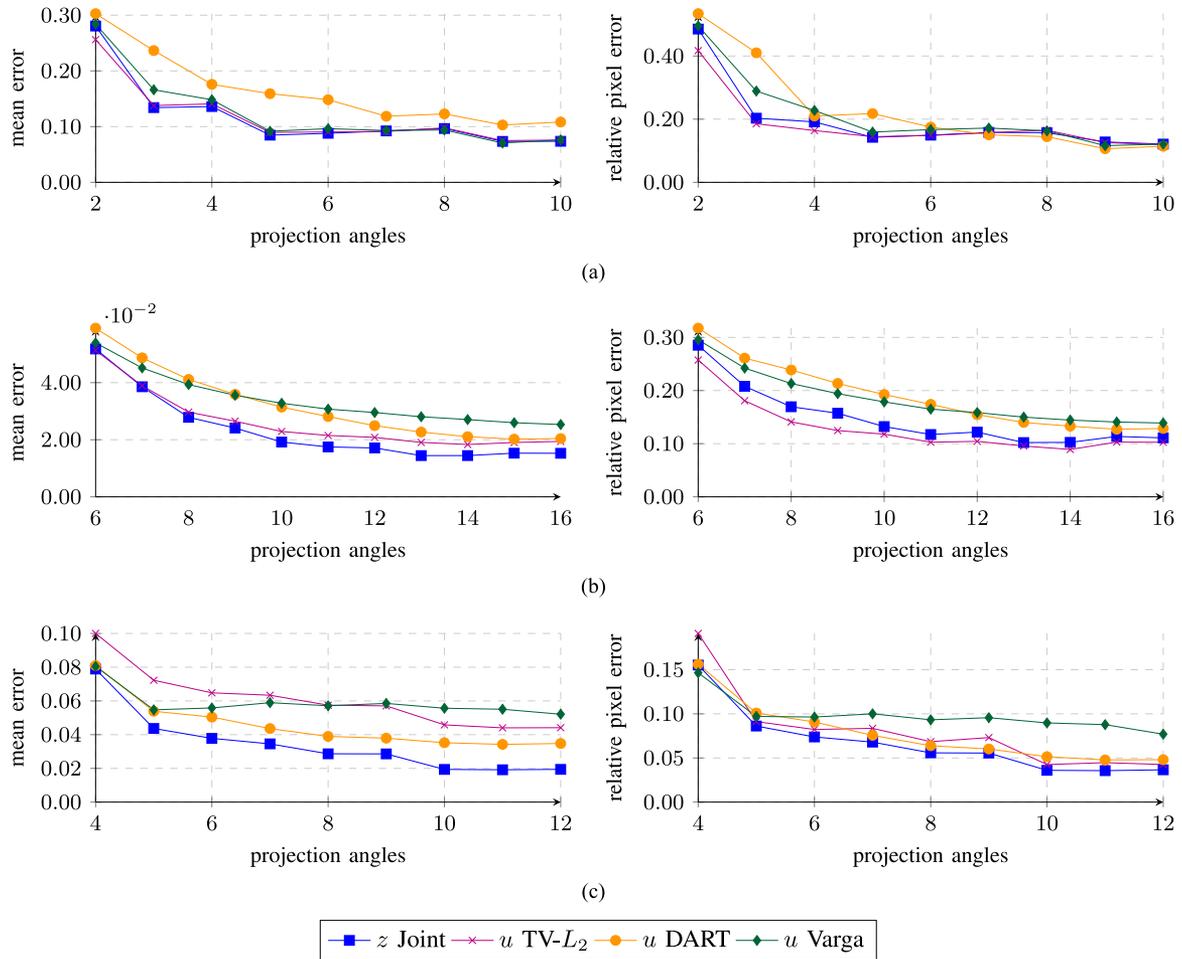


Fig. 10. Numerical evaluation of the approaches for the different test-datasets and increasing small numbers of projections in the noisy case (noise level: $SNR_{db} = 20$). The mean error measure was computed for non-rounded reconstructions and the relative pixel error after rounding. The results demonstrate that the proposed (Joint) approach is robust in the presence of noise. (a) Phantom 1: Maple Leaf. (b) Phantom 2: Shepp-Logan. (c) Phantom 3: Batenburg.

2) *Comparison to Other Methods*: Fig. 8 summarizes the numerical evaluation of the approaches listed in Table I, using the test-datasets in the noiseless case. DART gives better results than the approach of Varga on Phantoms 2 and 3, yet performs worse than the proposed approach. In case of the very small test-dataset Phantom 1, the approach of Varga is better than DART, because DART is applying a smoothing filter mask which limits the reconstruction accuracy for such small image structures.

TV- L_2 also performs better than DART. We attribute this effect to the anisotropic TV term of the TV- L_2 approach. Only for the non-rounded results in case of Phantom 3, see Fig. 8(c), DART is better than the TV- L_2 approach, but no longer so after rounding the results. The gap in reconstruction performance between the approaches based on TV regularization and the other two approaches can be explained by the used regularization method. It would therefore be interesting to extend DART and the approach of Varga to TV regularization.

Fig. 9 shows the runtimes for each algorithm. It is interesting to see how the runtime curves of TV- L_2 and the Joint approach are related. The peak of the Joint approach indicates the number of projections up to which the additional prior knowledge from the set of allowed intensities is effectively used for reconstruc-

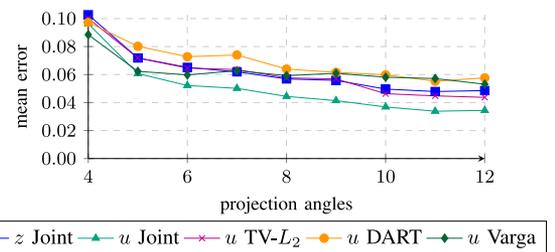


Fig. 11. Numerical evaluation of the approaches for Phantom 3: Batenburg and increasing small numbers of projections in the noisy case (noise level: $SNR_{db} = 20$) and with corrupted intensities $\hat{\mathcal{L}} = \{0.0220, 0.5588, 0.9628\}$. The mean error measure was computed for non-rounded reconstructions against the original groundtruth with $\mathcal{L} = \{0.000, 0.5020, 1.000\}$. Note, that the continuous variable u of the proposed approach is plotted which gives the lowest mean error since it is affected only indirectly by the perturbed intensity values. Interestingly, the TV- L_2 approach has a higher mean error although the noisy intensities are not involved in the corresponding objective function.

tion. Further increasing the number of projection angles yields similar curves for TV- L_2 and Joint, because then TV- L_2 already solves the reconstruction problem.

Clearly, DART and the approach of Varga are much faster than the TV- L_2 and the Joint approach. However, our proposed

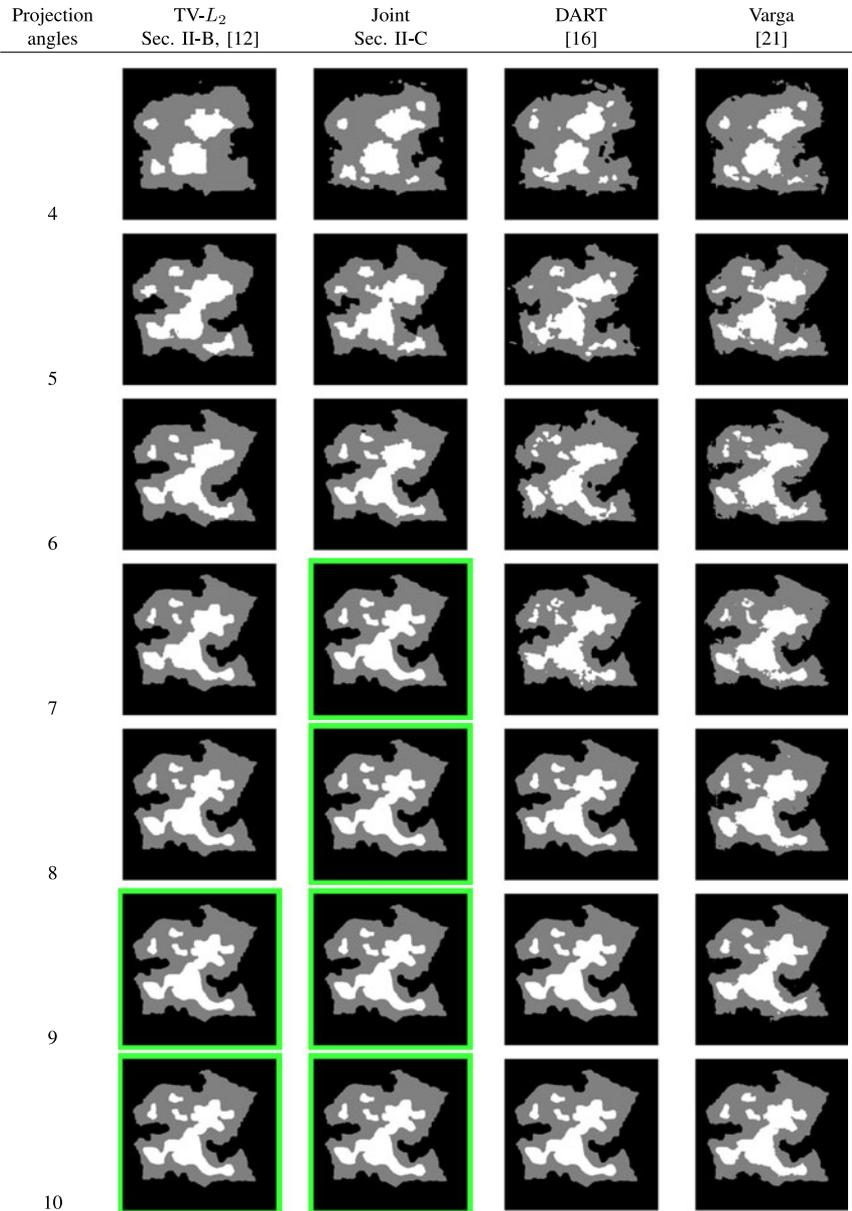


Fig. 12. Exemplary reconstructions (rounded results) from Phantom 3 for different numbers of projection angles to illustrate the visual differences of the compared methods. The proposed Joint approach needs less projections for an exact reconstruction compared to the other approaches. Perfect reconstructions are marked by a green frame.

Joint approach returns a fully discretized and more accurate solution, whereas the other approaches stop earlier and have to apply rounding. Obviously, the proposed approach can be parallelized and implemented e.g. in CUDA to run on a modern graphics card, if computational performance is important.

In Fig. 12 the reconstructions from projections of the test-dataset Phantom 3 are shown for all considered approaches. We see that our proposed Joint approach achieves better results even for a smaller number of projections compared to the other approaches. Exact reconstructions are marked by a green frame.

We also tested the approaches with noise imposed on all the datasets. Poisson noise was applied to the observations with a signal to noise ratio of $SNR = 20$ db. Fig. 10 shows for an increasing number of projections the numerical results using

the two error measures. Plots on the left display the mean errors which were evaluated on the non-rounded outputs, whereas on the right the relative pixel errors are shown, evaluated on the rounded outputs. The results demonstrate that the proposed (Joint) approach is robust in the presence of noise.

These results merely demonstrate that despite a significant amount of noise, the ranking of the performance of the approaches does not change, i.e. neither approach is susceptible to perturbations of the projection data. In realistic applications with such small numbers of projections, however, efforts will be made to obtain accurate measurements and a smaller level of noise. Then the Joint approach will be the method of choice.

Finally, the approaches were tested on noisy data *and* a corrupted set of feasible intensity values, which reflects

imperfect knowledge in practice. The results depicted by Fig. 11 reveal again the robustness of the proposed approach. Interestingly, the TV- L_2 approach has a higher mean error although the noisy intensities are not involved in the corresponding objective function.

V. CONCLUSION

We proposed an energy minimization approach to the non-binary discrete tomography problem. A non-convex coupling term was derived which incorporates additional prior knowledge so as to restrict the reconstruction to a set of feasible intensities. For optimization, we applied a generalized forward-backward splitting algorithm to our non-convex energy formulation. The numerical evaluation demonstrated the superior reconstruction quality of the proposed method. It returns an exact reconstruction with the least number of projection angles as compared to other state-of-the-art approaches from literature. Additionally, the approach converges to a discrete solution without a rounding post processing step, which is required for all other approaches.

There are a range of interesting questions which should be addressed in future work. The running time of the proposed approach, which is long for our current research code, can be improved. Furthermore, the existence of a sufficiently tight convex relaxation for the proposed coupling term is open. In the affirmative case, the approach could be casted into a convex optimization framework. Finally, adopting a norm different from Euclidean for the first term of (10) is conceivable, in order to better fit to the non-Gaussian noise characteristic.

REFERENCES

- [1] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*. Philadelphia, PA, USA: SIAM, 2001.
- [2] J. Bushberg, J. Seibert, E. Leidholdt, and J. Boone, *The Essential Physics of Medical Imaging*. 3rd ed. Alphen aan den Rijn, The Netherlands: Wolters Kluwer, 2011.
- [3] J. Banhart, *Advanced Tomographic Methods in Materials Research and Engineering*. London, U.K.: Oxford Univ. Press, 2008.
- [4] R. Hanke, T. Fuchs, and N. Uhlmann, "X-ray based methods for non-destructive testing and material characterization," *Nucl. Instrum. Methods. Phys. Res. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 591, no. 1, pp. 14–18, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0168900208003914>
- [5] G. Herman and A. Kuba, *Discrete Tomography: Foundations, Algorithms and Applications*. Basel, Switzerland: Birkhäuser, 1999.
- [6] S. Weber, C. Schnörr, and J. Hornegger, "A linear programming relaxation for binary tomography with smoothness priors," *Electron. Notes Discrete Math.*, vol. 12, pp. 243–254, 2003.
- [7] T. Schüle, C. Schnörr, S. Weber, and J. Hornegger, "Discrete tomography by convexconcave regularization and D.C. programming," *Discrete Appl. Math.*, vol. 151, no. 1-3, pp. 229–243, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0166218X05001320>
- [8] K. J. Batenburg, "a network flow algorithm for reconstructing binary images from continuous X-rays," *J. Math. Imag. Vis.*, vol. 30, no. 3, pp. 231–248, 2008.
- [9] S. Roux, H. Leclerc, and F. Hild, "Efficient binary tomographic reconstruction," *J. Math. Imag. Vision*, vol. 49, no. 2, pp. 335–351, 2014.
- [10] L. Wang, B. Sixou, and F. Peyrin, "Binary tomography reconstructions with stochastic level-set methods," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 920–924, Jul. 2015.
- [11] J. H. Kappes, S. Petra, C. Schnörr, and M. Zisler, "TomoGC: Binary tomography by constrained graphcuts," in *Proc. 37th German GCPR Conf.*, 2015, vol. 30, pp. 262–273.
- [12] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, no. 17, 2008, Art. no. 4777. [Online]. Available: <http://stacks.iop.org/0031-9155/53/i=17/a=021>
- [13] B. Goris, W. V. den Broek, K. Batenburg, H. H. Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0304399111002658>
- [14] A. Denitui, S. Petra, C. Schnörr, and C. Schnörr, "Phase transitions and cospare tomographic recovery of compound solid bodies from few projections," *Fundamenta Informaticae*, vol. 135, pp. 73–102, 2014.
- [15] M. Storath, A. Weinmann, J. Friel, and M. Unser, "Joint image reconstruction and segmentation using the Potts model," *Inverse Problems*, vol. 31, no. 2, 2015, Art. no. 025003.
- [16] K. Batenburg and J. Sijbers, "DART: A practical reconstruction algorithm for discrete tomography," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2542–2553, Sep. 2011.
- [17] X. Zhuge, W. Palenstijn, and K. Batenburg, "TVR-DART: A more robust algorithm for discrete tomography from limited projection data with automated gray value estimation," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 455–468, 2016.
- [18] A. Tuysuzoglu, W. Karl, I. Stojanovic, D. Castanon, and M. Unlu, "Graphcut based discrete-valued image reconstruction," *IEEE Trans. Image Process.*, vol. 24, no. 5, pp. 1614–1627, May 2015.
- [19] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
- [20] S. Maeda, W. Fukuda, A. Kanemura, and S. Ishii, "Maximum a posteriori X-ray computed tomography using graph cuts," *J. Phys., Conf. Series*, vol. 233, 2010, Art. no. 012023.
- [21] L. Varga, P. Balázs, and A. Nagy, "An energy minimization reconstruction algorithm for multivalued discrete tomography," in *Proc. 3rd Int. Symp. Comput. Modeling Objects Represented Images, Rome, Italy*, 2012, pp. 179–185.
- [22] R. Gardner and P. Gritzmann, "Discrete tomography: Determination of finite sets by X-rays," *Trans. Amer. Math. Soc.*, vol. 349, no. 6, pp. 2271–2295, 1997.
- [23] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, May 2011. [Online]. Available: <http://dx.doi.org/10.1007/s10851-010-0251-1>
- [24] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
- [25] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Progr.*, vol. 146, pp. 459–494, 2014.
- [26] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0304399115001060>
- [27] C. Pommrenke, "Constrained Graphical Model and its Application in Computer Vision," Master's thesis, Dept. Math. Comput. Sci., Heidelberg University, Heidelberg, Germany, 2015.



Matthias Zisler received the Diploma (FH) degree in mathematics from the University of Applied Sciences, Regensburg, Germany, and the Master's degree in computer science from the University of Applied Sciences, Munich, Germany. He is currently working towards the Ph.D. degree at the Image and Pattern Analysis Group, Institute of Applied Mathematics, Heidelberg University, Heidelberg, Germany.

His research interests include variational methods in imaging and efficient algorithms for determining their minimizers in particular with application to discrete tomography.



Jörg Hendrik Kappes received the Diploma degree in computer science from the University of Mannheim, Mannheim, Germany, in 2005 and the Ph.D. degree from Heidelberg University, Heidelberg, Germany, in 2011. Since 2011, he had been a Postdoctoral researcher at Heidelberg University. He is one of the major developers of OpenGM2, a free library for inference in discrete graphical models.

His research interests include discrete optimization, computer vision and probabilistic graphical

models.



Claudius Schnörr received the Dipl.Ing. degree in electrical engineering from the University of Karlsruhe, Karlsruhe, Germany, in 1989, and the Dr. Ing. degree (with Hons.) in electrical engineering from the University of Stuttgart, Stuttgart Germany, in 1998. Since 2004, he has been a Full Professor at the University of Applied Sciences, Munich, Germany, and a Founding Member of the research group “Computer Vision, Remote Sensing, and Navigation (CORSNAV).”



Stefania Petra received the B.Sc. degree in mathematics and computer science and the M.Sc. degree in mathematics from the Babeş-Bolyai University, Cluj-Napoca, Romania, in 2001 and 2003, respectively. She received the Ph.D. degree in the field of numerical optimization from the University of Würzburg, Würzburg, Germany, in 2006.

She continued as a Research Fellow in the University of Mannheim, Mannheim, Germany, and Heidelberg, Germany in the field of mathematical image processing. During 2013–2015, she was a Margarete von Wrangel-Fellow within the Postdoctoral Lecture Qualification Program of the Ministry of Science, Research and Arts of the state of Baden-Württemberg in Germany. Since 2015, she has been an Assistant Professor at the Heidelberg University, where she is leading the Mathematical Imaging Group at the Institute of Applied Mathematics.

Her research interests include compressed sensing, mathematical models of image analysis with an emphasis on tomography and numerical optimization.



Christoph Schnörr received the degrees from the Technical University of Karlsruhe and the University of Hamburg, Hamburg, Germany, in 1991 and 1998, respectively. He became a Full Professor at the University of Mannheim, Mannheim, Germany, in 1998. In 2008, he joined Heidelberg University, where he is currently heading the Image and Pattern Analysis Group, Institute of Applied Mathematics. Together with colleagues, he has set up and is Codirecting the Heidelberg Collaboratory for Image Processing (funded by the DFG and industrial partners). He also

serves as a Coordinator of a Research Training Group on Probabilistic Graphical Models with Applications to Image Analysis, funded by the German Research Foundation (DFG), that involves 12 Ph.D. positions and 10 colleagues as PIs.

His research interests include mathematical models of image analysis and numerical optimization.