TomoGC: Binary Tomography by Constrained GraphCuts

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

Heidelberg University

Abstract. We present an iterative reconstruction algorithm for binary tomography, called TomoGC, that solves the reconstruction problem based on a constrained graphical model by a sequence of graphcuts. TomoGC reconstructs objects even if a low number of measurements are only given, which enables shorter observation periods and lower radiation doses in industrial and medical applications. We additionally suggest some modifications of established methods that improve state-of-the-art methods. A comprehensive numerical evaluation demonstrates that the proposed method can reconstruct objects from a small number of projections more accurate and also faster than competitive methods.

1 Introduction

Limited-data tomography deals with the problem of reconstructing 3D-volumes or 2D-images denoted by $x \in \mathbb{R}^N$, from a small number of (noisy) projections given by $b = Ax + \nu \in \mathbb{R}^M$. The range of applications for tomography includes industrial [21] and medical [29] applications. In many situations it is desirable to reduce the number of required measurements M that are represented by the rows of the matrix $A \in \mathbb{R}^{M\times N}$. If M is much smaller than N, then the reconstruction problem is ill-posed and regularization is required.

The tomography reconstruction problem can be formulated as a regularized least squares (1) or a constrained minimization of the regularizer (2).

$$x^* \in \operatorname*{arg\,min}_{x\in\mathbb{R}^N} R(x) + \|Ax - b\|_2^2 \tag{1}$$

$$x^* \in \operatorname*{arg\,min}_{x\in\mathbb{R}^N} R(x), \qquad \text{s.t.} \quad \underline{b} \leq Ax \leq \bar{b} \tag{2}$$

While problem (1) is searching for a solution that has a low score of the regularizer and good data-fidelity, problem (2) is searching in the feasible set (given by the data-constraints) for the solution with the lowest score of the regularizer.

Early approaches such as filtered back projection (FBP) [7], deal with the tomography problem by analytical reconstruction methods, which provides reasonably accurate reconstructions in very short times, but usually require many projection angles. The algebraic reconstruction methods (ARMs) such as ART, SIRT or SART solve problem (1) without any regularization term $R(x)$. They fall into the category of row-action methods [10,11] also known as iterated projection methods for systems of linear (in)equalities. ARMs give better results than FBP, but due to the lack of regularization usually the number of required projections is still large.

For a further reduction of required observations, several regularization techniques have been proposed depending on the prior knowledge at hand. Convex sparsity promoting priors like ℓ_1- or total variation minimization [17], smoothness priors [36] or box constraints conserve the convexity of the overall problem. Such non-smooth, constrained, distributed and large-scale optimisation problems can be addressed by proximal algorithms in an unified fixed point theoretical framework [14,9] as finding solutions to monotone inclusion problems or more specifically by projections on convex sets [5,8]. In this context the alternating direction method of multipliers [28] and in particular the Chambolle Pock Algorithm [12], which is a decomposable method for minimizing the sum of two convex functions subject to linear constraints, can be considered for tomograhic inversion [31]. Interestingly, the ADMM framework can be adopted also when considering a non-convex regularization term like the ℓ_0-prior as done in [32]. However several questions concerning convergence remain open. For a sufficient uniqueness condition for the ℓ_0-regularized tomographic reconstruction problem in terms of the image gradient sparsity and the number of tomographic measurements, we refer to [15].

A further reduction of required measurements can be obtained if the range of x is a finite set. The tomography problems (1) and (2) with the additional constraint that $x \in \{v_1, \dots, v_K\}^N$ is known as *discrete tomography* problem, the subject of the present paper. A special case of this problem is *binary tomography* where the set is restricted to two possible values ($K = 2$) for each x_i, which in practice occurs e.g. when air pockets in work pieces need to be detected without destroying the object.

We underline that several heuristics have been designed to intervene between consecutive steps of a non-binary iterative image reconstruction algorithm in order to gradually steer the iterates towards a binary solution. Batenburg *et al.* suggested a (Soft) Discrete Algebraic Reconstruction Technique known as (S)DART [4,6], which is a very fast heuristic that starts from a continuous reconstruction, applies a segmentation step to restrict the reconstruction to the allowed values, and then restarts the continuous reconstruction on boundary regions of the segmentation, iteratively. While this leads to good results quite fast, it does not optimize an objective function. In another line of research Batenburg and Sijbers [2,3] presented an algorithm for the binary tomography problem that is based on a sequence of minimum cost flow problems. For two projection directions (with non-overlapping rays for each direction) this method is exact. In the general case, it is a greedy approximation.

An alternative ansatz is to reformulate problem (1) into a discrete graphical model. For the binary tomography problem (1) this leads to a fully connected second-order binary model [30]. The multi-label case can be reduced to a sequence of such binary problems in a α-expansion framework [30]. As this is in general not sub-modular, Raj *et al.* [30] have suggested to use QPBO [25] to solve a relaxation of the problem which give additional persistence certificates.

The main limitation of this approach is that the number of pairwise terms grows quadratically with the number of pixels and the complexity of QPBO roughly grows cubically with the number of pairwise terms in the worst case, which caused e.g. [30] to consider only restricted projection matrices A. To overcome this problem, Tuysuzoglu *et al.* [34] consider local approximations of the non-sub-modular terms around a working point which is iteratively improved. Similar methods have been also studied for more general graphical models [16,33].

Weber *et al.* [36,35] suggested to solve problem (1) by a quadratic program. The binary constraints are enforced by iteratively increasing a non-convex balloon-term that pushes the labels to zeros and ones. The subproblems are solved by the difference-of-convex-function programming technique that iteratively and locally approximates the non-convex part of the objective by an affine upper bound. While there is no guarantee that this method finds the global optimum, it generally returns good results.

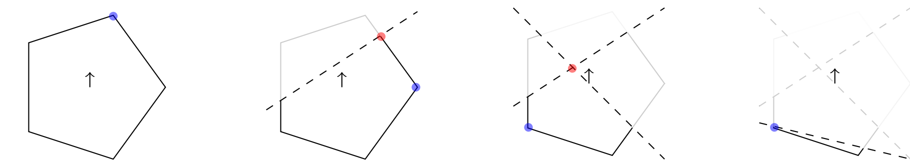
Gouillart *et al.* [18] have proposed a belief propagation algorithm for the discrete tomography problem (1). In order to handle the higher order interactions induced by the projection constraints, they include Lagrangian multipliers that enforce that these constraints are fulfilled on average. However, this algorithm only estimates the marginal distributions, which then are rounded to obtain a discrete reconstruction.

Outside the application area of tomography, Lagrangian relaxation has been used amongst others for multicommodity max flow [37], graphical models [26], and graphical models with a few constraints [27]. While in [26], contrary to our work, variable duplication is used to relax the problem, [37] and [27] use the same mathematical idea as we do in the present context of discrete tomography.

Contributions. We present a novel method for solving the binary tomography problem, which solves the dual of a relaxation of problem (2) by a sequence of graphcut problems. The size of these problems scales linearly with the number of primal variables and, besides the graphcut computation, only a few simple matrix-vector operations are required. Consequently, the proposed method is very efficient and scales up well to large data. On the other side, it is mathematically sound and is the only currently available method that provides a lower bound on the optimal objective value. Furthermore, we provide a comprehensive experimental comparison of state-of-the-art methods, which was lacking so far in the literature. We also suggest some modifications of standard methods which improve their performance or are even necessary to make these methods applicable in all considered scenarios.

2 Constrained GraphCuts for Binary Tomography

We consider problem (2) for $K = 2$. To ease the presentation, we temporarily consider the noise-free case where $\underline{b} = \overline{b} = b$ and generalize it later on. Without loss of generality, we assume that $v_1 = 0$ and $v_2 = 1$. We define a grid graph $G = (V, E)$, with V corresponding to image pixels and $E \subset V \times V$ defining the neighborhood system. As regularization term, we use $R(x) := \sum_{uv \in E} \beta \cdot |x_u - x_v|$,



(a) unconstrainted (b) 1 constraint (c) 2 constraints (d) 3 constraints

Fig. 1: Polyhedral illustration of the constrained linear program. In the unconstrained case **(a)**, the optimal solution is integral. With one additional constrained **(b)**, the LP-solution (•) and optimal integer solution in the constrained set (•) are not identical. When adding another constrained **(c)**, the LP-solution moves in the interior of the original polytope. By adding more constraints **(d)**, the feasible set gets smaller and finally the LP solution gets integral.

so the problem at hand is given by

$$x^* \in \operatorname*{arg\,min}_{x \in \{0,1\}^{|V|}} \sum_{uv \in E} \beta \cdot |x_u - x_v|, \qquad \text{s.t.} \quad Ax = b. \tag{3}$$

Without the additional constraints $Ax = b$ and $\beta \geq 0$, this problem can be solved as a linear program by relaxing the $\{0,1\}$ constraints to $[0,1]$ constraints and by representing $|x_u - x_v|$ linearly by means of additional auxiliary variables. This would be even the case if additional unary terms are added [23,13]. However, in the presence of projection constraints as part of the problem, this is no longer true, as illustrated in Fig. 1. The relaxed linear program can then have non-binary solutions.

In order to find efficiently a solution of the relaxed problem (3), we consider its Lagrangian dual

$$\max_{\lambda} \min_{x \in [0,1]^{|V|}} \sum_{uv \in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax - b \rangle \tag{4}$$

$$= \max_{\lambda} \underbrace{\min_{x \in [0,1]^{|V|}} \sum_{uv \in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax \rangle - \langle \lambda, b \rangle}_{=: g(\lambda)}. \tag{5}$$

By weak duality, we know that for every feasible primal x and feasible dual value λ, the inequality (6) holds.

$$\sum_{uv \in E} \beta \cdot |x_u - x_v| \geq g(\lambda) \tag{6}$$

If the optima x^* and λ^* exist, equality in (6) holds (strong duality). If a feasible finite primal solution exists, then also the dual has a feasible finite solution. In the case that no feasible primal value exists, the dual problem is unbounded.

As a consequence, if a feasible primal solution exists, then we may solve the dual problem instead of the primal, followed by recovering a primal solution from the dual solution.

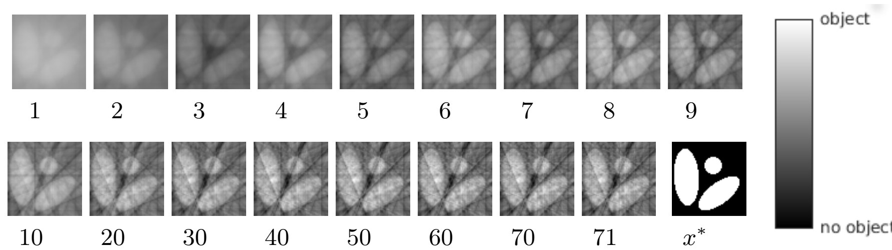


Fig. 2: Shows the evolution of the unary data-term $A^\top\lambda$ during the iterations. After 71 iterations, the data term leads to a duality gap of zero. This illustrates that after a few iterations, the data term does not change so much any more.

The most simple algorithm to optimize the dual problem (5) is iterative subgradient ascent with a proper stepsize sequence γ_i. For any λ and

$$x^\lambda \in \arg\min_{x\in[0,1]^{|V|}} \sum_{uv\in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax \rangle,$$

a lower bound on the optimal value is given by $\sum_{uv\in E} \beta \cdot |x_u^\lambda - x_v^\lambda| + \langle \lambda, Ax^\lambda - b\rangle$. For the dual objective $g(\lambda)$ and its subdifferential $\partial g(\lambda)$, we compute a subgradient by

$$\partial g(\lambda) \ni Ax^\lambda - b, \qquad x^\lambda \in \arg\min_{x\in[0,1]^{|V|}} \sum_{uv\in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax \rangle. \tag{7}$$

The calculation of x^λ can be further simplified by making use of the relation to graphcuts [23,13], which guarantees that a binary solution exists, that is globally optimal. This can be efficiently calculated by a graphcut (max-flow) algorithm. As long as $\beta > 0$, the optimal solution x^* does not depend on the value of β. Only the optimal dual variable λ^* will scale according to β.

An interesting observation is that by optimizing the dual objective, we iteratively build up a unary data term $A^\top\lambda$, as illustrated in Fig. 2. Due to regularization, the unary terms do not have to be perfect. While a reasonable data term is found after a few iterations, most of the iterations are required to close the primal-dual-gap without changing the dual variables much.

The construction of a feasible primal solution is non-trivial. While general primal construction rules exists [20], these produce an optimal and feasible solution only in the limit. More advanced methods for solving the dual, for example bundle methods [22], have a faster convergence and also provide primal estimates. However, a study of those methods is beyond the present work.

As we are interested in binary solutions anyway, we have come up with the following framework to generate primal solutions. Each subgradient yields a primal solution x^λ. If this solution is feasible and strong duality holds, *i.e.* $\langle \lambda, Ax^\lambda - b\rangle = 0$, this is an optimal primal solution. If the optimal primal solution is non-binary, the sub-gradients will oscillate around the non-binary solu-

Algorithm 1 TomoGC (noise free case)

Require: $A \in \mathbb{R}^{M\times N},\ b \in \mathbb{R}^{M\times 1},\ \beta > 0,\ E \subset [N]^2$
Ensure: $v \leq \min_{x\in[0,1]^N,\ Ax=b} \sum_{uv\in E} \beta \cdot |x_u - x_v|$ if feasible
initialize: $i = 0, \quad \lambda = [0]^{1\times M}, \quad \bar{x} = [0]^{N\times 1}$
$x^\lambda \in \arg\min_{x\in\{0,1\}^{|V|}} \sum_{uv\in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax\rangle$
while ($\|Ax^\lambda - b\| > 0$ and $\langle \lambda, Ax^\lambda - b\rangle \neq 0$) and $i < i_{\max}$ **do**
  $\lambda = \lambda + \gamma_i(x^\lambda) \cdot [Ax^\lambda - b]$
  $x^\lambda \in \arg\min_{x\in\{0,1\}^{|V|}} \sum_{uv\in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax\rangle$
  if $\|A\bar{x} - b\| > \|Ax^\lambda - b\|$ **then**
    $\bar{x} = x$
  end if
  $i = i + 1$
end while
$x = \bar{x}$
$v = \sum_{uv\in E} \beta \cdot |x_u - x_v| + \langle \lambda, Ax - b\rangle$

tion. But if the solution is binary and unique, the dual objective will have the optimal primal solution as subgradient at the optimal dual point.

The pseudocode of our method is given in Alg. 1. In each iteration, we update the dual variable in the direction of the subgradient. The non-summable diminishing step length that ensures convergence, is defined by $\gamma_i(x) = \frac{20}{(0.1\cdot i+1)\cdot\|Ax-b\|_2}$, $i \in \mathbb{N}$.

Noisy data case. In the case where we have to deal with noise and $\underline{b} < \overline{b}$, we have to replace $Ax - b$ in eq. 4 and Alg. 1 by $\max\{0, Ax - \underline{b}\} + \min\{Ax - \overline{b}, 0\}$. The values $\underline{b}$ and $\overline{b}$ have to be selected with respect to the noisy measurements b and the assumed noise level such that a feasible solution exists.

3 Experiments

For our experimental evaluation, we used the binary test-datasets of Weber *et al.* [35] and Batenburg and Sijbers [4]. We generated the projection matrices with the ASTRA-toolbox [1] and simulated parallel projections within the range of 0 and 180 degrees. The width of the sensor-array is 1.5 times the image size and each sensor has the same size as a pixel. The entries of the projection matrix A are given by the length of the intersection of the pixels and the rays. We restricted our evaluation to algorithms that can deal with arbitrary projection matrices and excluded methods that make additional assumptions such as $A \in \{0,1\}^{M\times N}$. Table 1 lists all methods that we evaluated.

As a baseline for continuous methods we considered Filtered Back Projection (**FBP**) [7], Simultaneous Iterative Reconstruction Technique (**SIRT**) [19], and a total variation regularized reconstruction with hard projection constraints (**tomoTV**) [15]. For the former two, we used the implementation available in the ASTRA-toolbox, the latter was kindly provided by Denitiu *et al.*.

shortcut	reference	label	regularization	implementation	objective
FBP	[7]	cont.	no	ASTRA-toolbox	-
SIRT	[19]	cont.	no	ASTRA-toolbox	eq. (1)
tomoTV	[15]	cont.	TV	Denitiu et al.	eq. (2)
tomoDC	[36]	binary	Potts	ours	eq. (1)
tomoFTR*	[34,16]	binary	Potts	ours	eq. (1)
tomoPB*	[34,33]	binary	Potts	ours	eq. (1)
tomoGC*	Sec. 2	binary	Potts	ours	eq. (2)
DART	[4]	discrete	-	ASTRA-toolbox	-
DART-S*	[4]	discrete	Potts	ASTRA-toolbox + TRWS	-

Table 1: Compared Methods. Methods marked with * are either novel methods or extensions of existing methods proposed in the present work.

We furthermore compared to the Discrete Algebraic Reconstruction Technique (**DART**) [4]. We used the publicly available implementation of the ASTRA-toolbox. For the continuous iterative reconstructions we used SIRT. We set the smoothing intensity and percentage of random points to 0.1, which are the suggested default values, and run DART for 20 iterations. Additionally, we suggest a variation of the DART method by replacing the elementary nearest neighbor segmentation of the DART-method by a structured segmentation that also includes a smoothness-term. In order to be able to deal with multi-label problems, we used TRWS [24] to solve the segmentation problems. To the best of our knowledge, this combination of DART and structured segmentation (**DART-S**) has not been considered before.

For the binary case, we implemented the difference-of-convex-functions approach (**tomoDC**) from Weber *et al.* [36] which is known to give good results even with a low number of projections. We used the same parameter setting as described in [35] and the implementation of the spectral projected gradient (SPG) method of Mark Schmidt[1] for solving the subproblems. When running tomoDC on the large instances from [4], we observed that the method got stucked in non-binary equilibriums due to numerical reasons. Because adding some additional noise as suggested by Weber *et al.* did not solve the problem, we initialized tomoDC with the solution of FBP. This resolved all numerical problems for all our problem instances.

In recent work Tuysuzoglu *et al.* [34] solve binary tomography problems by a set of surrogate problems that approximate the original function around the current solution. The surrogate problems are designed to be solvable by graph-cut (max-flow) methods. If the best solution of all surrogate problems improves the original energy, then the current solution is updated accordingly and the procedure continues, otherwise it stops. The downside of this approach is that the selection of the surrogate problems in [34] is rather greedy and inefficient.

[1] `http://www.cs.ubc.ca/~schmidtm/Software/minConf.html`

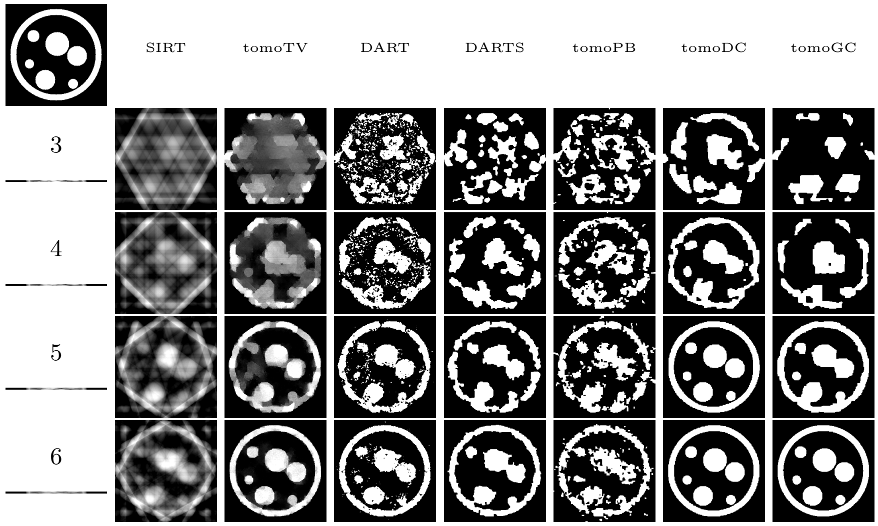
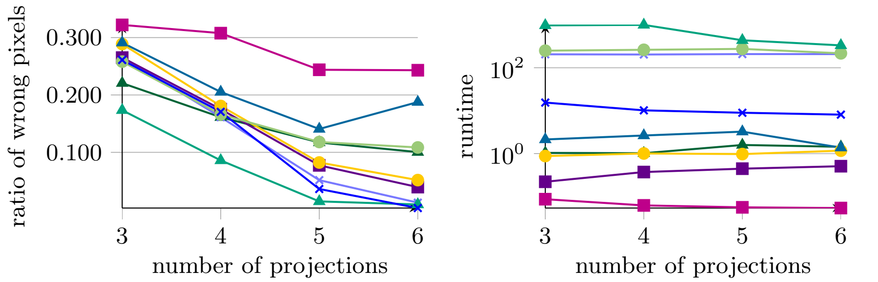


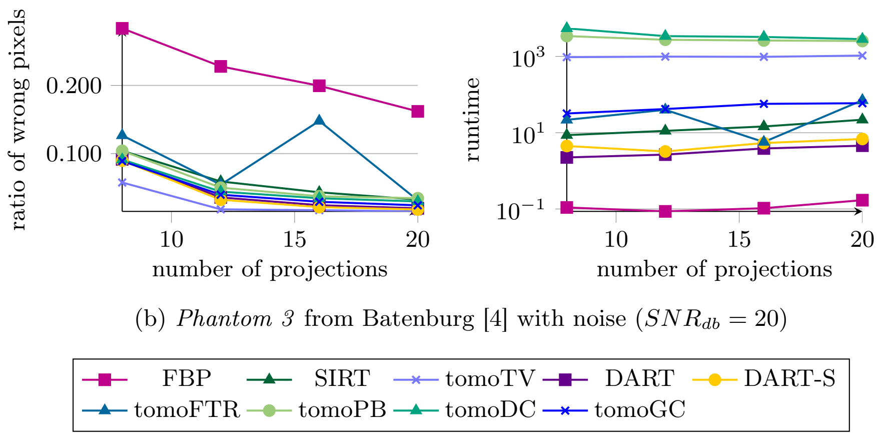
Fig. 3: *Phantom 3* from Weber [35] with no noise

Inspired by this work, we recognize some relations to recent works in the area of discrete optimization [16,33] which better indicate how to choose these surrogate problems. Tang *et al.* [33] consider all possible surrogate problems (pseudo bounds) with respect to the free parameter, and find all possible solutions by parametric max-flow. By using parametric max-flow a greedy selection is only required if the number of possible solutions is too large - this can be the case if the current solution is bad. In such a case, we simply greedy-like suppress nearby solutions. Typically, after a few iterations, the solution is good enough such that the number of possible solutions is small. We call this method tomography with pseudo bounds (**tomoPB**). A similar approach was suggested by Gorelick *et al.* [16] - originally also not applied to tomography problems. They use also a first-order Taylor expansion as an upper bound of the original function around the working point. An additional trust region term, based on the Euclidean distance, enforces solutions in the local region where the objective function is approximated well. We call this method tomography with fast trust region (**tomoFTR**).

A full evaluation of all test-instances is reported in the supplementary material. Due to lack of space, we can only show here two examples and some reconstructions. Fig. 3 and 5 show in the first row the original data and the sinograms (b) which are measured with k projection angles. Fig. 4 shows the ratio of wrongly reconstructed pixels and runtime for a different number of projection angles for two examples. In the noise free case tomoDC and tomoGC give the best results, but tomoGC is typically one magnitude faster. As shown in Fig. 3 and 4a, those are able to obtain nearly optimal reconstructions with only 5 projections. DART-S gives a reasonable result with 5 projections, which



(a) *Phantom 3* from Weber [35] with no noise



(b) *Phantom 3* from Batenburg [4] with noise ($SNR_{db} = 20$)

Fig. 4: Exemplary plots for the runtime and pixel accuracy in the noise free and noisy case for small number of projections. In the noise free case, tomoDC, tomoTV and tomoGC give the best results, but tomoGC is typically one magnitude faster. In the presence of noise tomoTV, DART and DART-S give the best results, since their greediness/rounding make them robust against noise.

is much better than the original DART method with only slightly increased runtime. FBP, SIRT, tomoFTR and tomoPB have problems with this small number of projections and require more projections for reasonable reconstructions.

We also simulated noisy observations by adding Poisson noise to the sinograms (b). The reconstruction results shown in Fig. 4b and 5 are obtained with a signal to noise ratio (SNR) of 20db. None of the problem formulations are designed to deal with Poisson noise, which is the most realistic approximation of noise in tomography. DART, DART-S and tomoTV include a rounding procedure, which removes noise in a greedy way. This seemed to work better than more sophisticated approaches, like tomoDC or tomoGC, which use a "wrong" noise model and added some artefacts to fulfill the projection constraints. The best results are obtained by tomoTV after rounding and DART-S, which again gives better results compared to the original DART method. In the presence

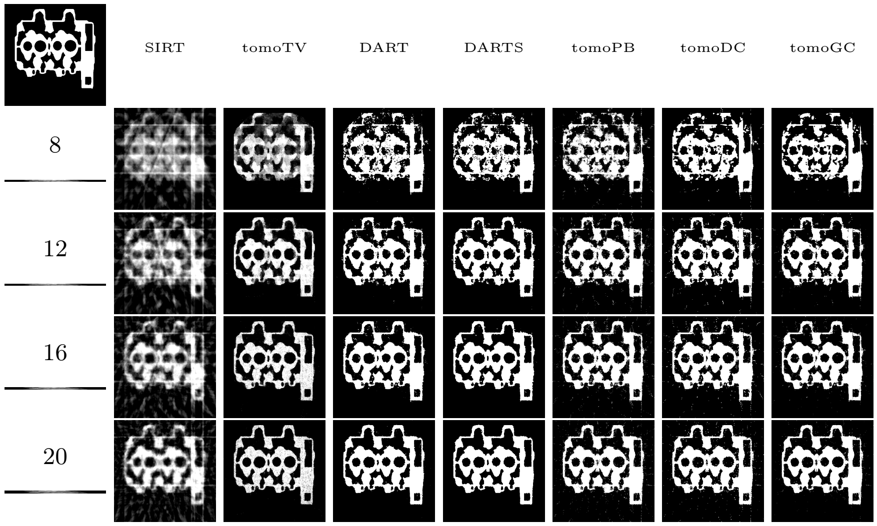


Fig. 5: *Phantom 3* from Batenburg [4] with noise ($SNR_{db} = 20$)

of noise, tomoFTR got sometimes stucked in local fixed points, and tomoPB performs better than tomoDC, but worse than tomoGC.

4 Conclusion and Future Work

We presented a new method for efficient binary reconstruction problems. In each iteration, our method only has to perform simple matrix vector operations and a graphcut problem of the size of the image/volume. For large-scale problems, solving the graphcut problem becomes the limiting factor, but efficient parallel implementations for this problem have been suggested in the recent literature. Even without this specialized implementations, our method is by more than one magnitude faster than competitive methods and provides additional theoretical guarantees, which makes it appealing to be used as a sub-solver within a α-expansion like algorithm, as suggested in [34].

For the generalization to multi-label tomography, we obtained some first promising results by replacing graph cuts with graphical models, which is equivalent in the binary case. However, in the multi-label case two additional problems have to be considered. Firstly, the discrete inference problem is no longer tractable in polynomial time and secondly, the allowed values span a simplex and no longer live on an one-dimensional space.

In future work, we also plan to replace naive subgradient ascent by the more advanced bundle method with automatic stepsize choice [22]. This should give a further speedup and non-binary primal estimates which can be used to suppress noise similar to tomoTV.

Acknowledgements: Financial support of our research work by the DFG, grant GRK 1653. is gratefully acknowledged.

References

1. van Aarle, W., Palenstijn, W.J., Beenhouwer, J.D., Altantzis, T., Bals, S., Batenburg, K.J., Sijbers, J.: The ASTRA-toolbox: A platform for advanced algorithm development in electron tomography. Ultramicroscopy 0(0), – (2015)
2. Batenburg, K.J.: A network flow algorithm for reconstructing binary images from continuous x-rays. Journal of Math. Imaging and Vision 30(3), 231–248 (2008)
3. Batenburg, K.J., Sijbers, J.: Generic iterative subset algorithms for discrete tomography. Discrete Applied Mathematics 157(3), 438–451 (2009)
4. Batenburg, K.J., Sijbers, J.: DART: A practical reconstruction algorithm for discrete tomography. IEEE Trans. Image Processing 20(9), 2542–2553 (2011)
5. Bauschke, H.H., Borwein, J.M.: On Projection Algorithms for Solving Convex Feasibility Problems. SIAM Rev. 38(3), 367–426 (Sep 1996)
6. Bleichrodt, F., Tabak, F., Batenburg, K.J.: SDART: an algorithm for discrete tomography from noisy projections. Computer Vision and Image Understanding 129, 63–74 (2014)
7. Bracewell, R.N., Riddle, A.C.: Inversion of fan-beam scans in radio astronomy. Astronomical Journal 150(2), 427–434 (1967)
8. Capricelli, T., Combettes, P.: A convex programming algorithm for noisy discrete tomography. In: Advances in Discrete Tomography and Its Applications. Birkhäuser, Boston (2007)
9. Cegielski, A.: Iterative Methods for Fixed Point Problems in Hilbert Spaces, Lect. Not. Math., vol. 2057. Springer (2013)
10. Censor, Y.: Row-action methods for huge and sparse systems and their applications. SIAM Review 23(4), 444–466 (1981)
11. Censor, Y., Zenios, S.: Parallel Optimization: Theory, Algorithms, and Applications. Oxford Univ. Press, New York (1997)
12. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Math. Imaging and Vision 40(1), 120–145 (2011)
13. Chambolle, A.: Total variation minimization and a class of binary mrf models. In: Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science, vol. 3757, pp. 136–152. Springer (2005)
14. Combettes, P.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization 53(5-6), 475–504 (2004)
15. Denitiu, A., Petra, S., Schnörr, C., Schnörr, C.: Phase transitions and cosparse tomographic recovery of compound solid bodies from few projections. Fundamenta Informaticae 135, 73–102 (2014)
16. Gorelick, L., Schmidt, F.R., Boykov, Y.: Fast trust region for segmentation. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013. pp. 1714–1721 (2013)
17. Goris, B., Van den Broek, W., Batenburg, K., Mezerji, H., Bals, S.: Electron tomography based on a total variation minimization reconstruction techniques. Ultramicroscopy 113, 120–130 (2012)
18. Gouillart, E., Krzakala, F., Mezard, M., Zdeborova, L.: Belief-propagation reconstruction for discrete tomography. Inverse Problems 29(3), 035003 (2013)
19. Gregor, J., Benson, T.: Computational analysis and improvement of SIRT. IEEE Trans. Med. Imaging 27(7), 918–924 (2008)
20. Gustavsson, E., Patriksson, M., Strömberg, A.B.: Primal convergence from dual subgradient methods for convex optimization. Mathematical Programming 150(2), 365–390 (2015)

21. Hanke, R., Fuchs, T., Uhlmann, N.: X-ray based methods for non-destructive testing and material characterization. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 591(1), 14 – 18 (2008), radiation Imaging Detectors 2007 Proceedings of the 9th International Workshop on Radiation Imaging Detectors
22. Kiwiel, K.C.: Proximity control in bundle methods for convex nondifferentiable minimization. Math. Program. 46, 105–122 (1990)
23. Kolmogorov, V., Zabin, R.: What energy functions can be minimized via graph cuts? IEEE Trans. Pattern Anal. Mach. Intell. 26(2), 147–159 (Feb 2004)
24. Kolmogorov, V.: Convergent tree-reweighted message passing for energy minimization. IEEE Trans. Pattern Anal. Mach. Intell. 28(10), 1568–1583 (Oct 2006)
25. Kolmogorov, V., Rother, C.: Minimizing nonsubmodular functions with graph cuts-a review. IEEE Trans. Pattern Anal. Mach. Intell. 29(7), 1274–1279 (2007)
26. Komodakis, N., Paragios, N., Tziritas, G.: MRF energy minimization and beyond via dual decomposition. IEEE Trans. Pattern Anal. Mach. Intell. 33(3), 531–552 (2011)
27. Lim, Y., Jung, K., Kohli, P.: Efficient energy minimization for enforcing label statistics. IEEE Trans. Pattern Anal. Mach. Intell. 36(9), 1893–1899 (2014)
28. Parikh, N., Boyd, S.: Proximal algorithms. Found. Trends Optimization 1(3), 1–108 (2013)
29. R, S.B., J, L., R, M., et al: Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Archives of Internal Medicine 169(22), 2078–2086 (2009)
30. Raj, A., Singh, G., Zabih, R.: MRF's for MRI's: Bayesian reconstruction of MR images via graph cuts. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006), 17-22 June 2006, New York, NY, USA. pp. 1061–1068 (2006)
31. Sidky, E.Y., Jakob H Jörgensen, J.H., Pan, X.: Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. Physics in Medicine and Biology 57(10), 3065 (2012)
32. Storath, M., Weinmann, A., Frikel, J., Unser, M.: Joint image reconstruction and segmentation using the Potts model. Inverse Problems 31(2), 025003 (2015)
33. Tang, M., Ayed, I.B., Boykov, Y.: Pseudo-bound optimization for binary energies. In: Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V. pp. 691–707 (2014)
34. Tuysuzoglu, A., Karl, W., Stojanovic, I., Castanon, D., Unlu, M.: Graph-cut based discrete-valued image reconstruction. IEEE Trans. Image Processing 24(5), 1614–1627 (May 2015)
35. 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, 13th International Conference, DGCI 2006, Szeged, Hungary, October 25-27, 2006, Proceedings. Lecture Notes in Computer Science, vol. 4245, pp. 146–156. Springer (2006)
36. 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)
37. Xiao, L., Johansson, M., Boyd, S.: Simultaneous routing and resource allocation via dual decomposition. IEEE Trans. on Communications 52(7), 1136–1144 (July 2004)