An Entropic Perturbation Approach to TV-Minimization for Limited-Data Tomography

Andreea Deniţiu[1,2], Stefania Petra, Claudius Schnörr[2], and Christoph Schnörr[1]

[1] Image and Pattern Analysis Group, University of Heidelberg, Germany
[2] Hochschule München, Fakultät für Informatik und Mathematik, München, Germany
{denitiu,petra,schnoerr}@math.uni-heidelberg.de
denitiu@hm.edu, schnoerr@cs.hm.edu

Abstract. The reconstruction problem of discrete tomography is studied using novel techniques from compressive sensing. Recent theoretical results of the authors enable to predict the number of measurements required for the unique reconstruction of a class of co-sparse dense 2D and 3D signals in severely undersampled scenarios by convex programming. These results extend established ℓ_1-related theory based on sparsity of the signal itself to novel scenarios not covered so far, including tomographic projections of 3D solid bodies composed of few different materials. As a consequence, the large-scale optimization task based on total-variation minimization subject to tomographic projection constraints is considerably more complex than basic ℓ_1-programming for sparse regularization. We propose an entropic perturbation of the objective that enables to apply efficient methodologies from unconstrained optimization to the perturbed dual program. Numerical results validate the theory for large-scale recovery problems of integer-values functions that exceed the capacity of the commercial MOSEK software.

Keywords: discrete tomography, compressed sensing, underdetermined systems of linear equations, cosparsity, phase transitions, total variation, entropic perturbation, convex duality, convex programming.

1 Introduction

This paper addresses the problem of reconstructing compound solid bodies u from few tomographic projections: $Au = b$. Theoretical guarantees of reconstruction performance relate to current research in the field of compressive sensing, concerned with *real* sensor matrices A that do *not* satisfy commonly made assumptions like the restricted isometry property.

By adopting the cosparse signal model [1] that conforms to our tomographic scenario, the authors [2] recently established the existence of "phase transitions" that relate the required limited(!) number of measurements to the cosparsity level of u in order to guarantuee unique recovery. This result was extensively validated numerically using the commercial MOSEK software, to avoid that numerical optimization issues affect the validation.

For our 3D problems a dedicated numerical optimization algorithm is necessary, however, because MOSEK cannot handle medium and large problem sizes. We present in this paper such an approach by utilizing the fact that adequate perturbations of the optimization problem may lead to a simpler dual problem [3, 4]. We work out a corresponding approach to our specific reconstruction problem

$$\min_u \mathrm{TV}(u) \quad \text{subject to} \quad Au = b, \tag{1}$$

that minimizes the total variation $\mathrm{TV}(u)$ subject to the projection constraints.

Organization. Section 2 briefly reports the above-mentioned phase transitions, followed by presenting a reformulation in Section 3 together with uniqueness results. A suitable perturbation is worked out in Section 4 that favourably compares to MOSEK for relevant problem sizes (Section 5).

Basic Notation. For $n \in \mathbb{N}$, we denote $[n] = \{1, 2, \ldots, n\}$. The complement of a subset $J \subset [n]$ is denoted by $J^c = [n] \setminus J$. For some matrix A and a vector z, A_J denotes the submatrix of rows indexed by J, and z_J the corresponding subvector. A_i will denote the ith row of A. $\mathcal{N}(A)$ and $\mathcal{R}(A)$ denote the nullspace and the range of A, respectively. Vectors are columns vectors and indexed by superscripts. Sometimes we write e.g. $(u, v)^\top$ instead correctly $(u^\top, v^\top)^\top$. $A^\top$ denotes the transposed of A. $\mathbb{1} = (1, 1, \ldots, 1)^\top$ denotes the one-vector whose dimension will always be clear from the context. The dimension of a vector z is $\dim(z)$. $\langle x, z\rangle$ denotes the standard scalar product in $\mathbb{R}^n$ and we write $x \perp z$ if $\langle x, z\rangle = 0$. The indicator function of a set C is denoted by $\delta_C(x) := \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{if } x \notin C. \end{cases}$ $\sigma_C(x) := \sup_{y\in C}\langle y, x\rangle$ denotes the support function of nonempty set C. $\partial f(x)$ is the subdifferential of f at x and $\operatorname{int} C$ and $\operatorname{rint} C$ denote the interior and the relative interior of a set C. f^* conjugate function of f. We refer to [5] for related properties.

In what follows, we will work with an anisotropic discretization of the TV operator in (1) given by

$$\mathrm{TV}_d(u) := \|Bu\|_1, \qquad B := \begin{pmatrix} \partial_1 \otimes I \otimes I \\ I \otimes \partial_2 \otimes I \\ I \otimes I \otimes \partial_3 \end{pmatrix} \in \mathbb{R}^{p\times n} \quad , \tag{2}$$

where $\otimes$ denotes the Kronecker product and ∂_i, $i = 1, 2, 3$, are derivative matrices forming the forward first-order differences of u wrt. the respective coordinates. Implicitly, it is understood that u on the r.h.s. of (2) is a vector representing all voxel values in appropriate order.

2 Weak Phase Transitions for TV-Based Reconstruction

We summarize in this section an essential result from [2] concerning the unique recovery from tomographic projection by solving problem (1), depending on the cosparsity level ℓ of u (Definition 1) and the number m of measurements (projections) of u. This result motivates the mathematical programming approach discussed in Section 3 and the corresponding numerical optimization approach presented in Section 4.

Definition 1 (cosparsity, cosupport). *The* cosparsity *of* $u \in \mathbb{R}^n$ with respect to $B \in \mathbb{R}^{p \times n}$ *is*

$$\ell := p - \|Bu\|_0 \ , \tag{3}$$

and the cosupport *of* u with respect to B *is*

$$\Lambda := \{r \in [p] \colon (Bu)_r = 0\}, \qquad |\Lambda| = \ell \ . \tag{4}$$

Thus, $B_\Lambda u = 0$. In view of the specific operator B given by (2), ℓ measures the homogeneity of the volume function u. A large value of ℓ can be expected for u corresponding to solid bodies composed of few homogeneous components.

Proposition 1 ([2, Cor. 4.6]). *For given* $A \in \mathbb{R}^{m \times n}$ *and* $B \in \mathbb{R}^{p \times n}$*, suppose the rows of* $\left(\begin{smallmatrix} A \\ B \end{smallmatrix}\right)$ *are linearly independent. Then a* ℓ*-cosparse solution* u *to the measurement equations* $Au = b$ *will be unique if the number of measurements satisfies*

$$m \geq 2n - (\ell + \sqrt{2\ell + 1} - 1) \qquad (d = 2) \ , \tag{5a}$$

$$m \geq 2n - \frac{2}{3}\Big(\ell + \sqrt[3]{3\ell^2} + 2\sqrt[3]{\frac{\ell}{3}} - 2\Big) \qquad (d = 3) \ . \tag{5b}$$

The above lower bounds on the number of measurements m required to recover a ℓ-cosparse vector u imply that recovery can be carried out by solving $\min_u \|Bu\|_0$ subject to $Au = b$. Replacing this objective by the convex relaxation (2) yields an excellent agreement of empirical results with the prediction (5), as shown in [2], although the independency assumption made in Prop. (1) does not strictly hold for the sensor matrices A and the operator B (2) used in practice.

This motivates to focus on efficient and sparse numerical optimization techniques that scale up to large problem sizes.

3 TV-Recovery by Linear Programming

We consider the discretized TV-term (2), an additional nonnegative constraint on image u and express $Bu = z$. Thus, (1) becomes

$$\min_{u,z} \|z\|_1 \qquad \text{s.t.} \quad Bu = z, \quad Au = b, \quad u \geq 0 \ . \tag{6}$$

3.1 Primal Linear Program and its Dual

By splitting the variable z in its positive $v^1 := \max\{0, z\}$ and negative part $v^2 := -\min\{0, z\}$ we convert problem (6) into a linear program in normal form. With

$$M := \begin{pmatrix} B & -I & I \\ A & 0 & 0 \end{pmatrix}, \quad q := \begin{pmatrix} 0 \\ b \end{pmatrix} \ , \tag{7}$$

and the polyhedral set

$$\mathcal{P} := \{y \in \mathbb{R}^{n+2p} \colon Mx = q,\ x \geq 0\}, \quad x := (u, v^1, v^2)^\top \ , \tag{8}$$

problem (6) becomes the linear program (P)

$$(P) \qquad \min_{x \in \mathcal{P}} \langle c, x\rangle = \min_{(u,v^1,v^2) \in \mathcal{P}} \langle \mathbb{1}, v^1 + v^2\rangle, \quad c = (0, \mathbb{1}, \mathbb{1})^\top \ . \tag{9}$$

We further assume that $\mathcal{P} \neq \emptyset$ and a feasible solution always exists. Due to $c \geq 0$, the linear objective in P is bounded on $\mathcal{P}$. Thus (P) always has a solution under the feasibility assumption. In view of basic linear programing theory, compare [5, 11.43], the dual program also has a solution. The dual program (D) reads

$$(D) \qquad \min_{y} -\langle q, y\rangle, \qquad M^\top y \leq c \ .$$

With

$$y = \begin{pmatrix} y_0 \\ y_b \end{pmatrix}, \quad M^\top y = \begin{pmatrix} B^\top & A^\top \\ -I & 0 \\ I & 0 \end{pmatrix}, \quad y = \begin{pmatrix} B^\top y_0 + A^\top y_b \\ -y_0 \\ y_0 \end{pmatrix} \ , \tag{10}$$

this reads

$$\min_{y_0, y_b} -\langle b, y_b\rangle \quad \text{s.t.} \quad B^\top y_0 + A^\top y_b \leq 0, \quad -1 \leq y_0 \leq 1 \ . \tag{11}$$

Moreover, both primal and dual solutions $(\overline{x}, \overline{y})$ will satisfy the following optimality conditions

$$0 \leq c - M^\top y \perp x \geq 0 \ , \tag{12}$$

$$Mx = q \ . \tag{13}$$

3.2 Uniqueness of Primal LP

A classical argument for replacing $\|\cdot\|_0$ by $\|\cdot\|_1$ and solving for (6) is uniqueness of the LP solution. Let $\overline{x} = (\overline{u}, \overline{v}) = (\overline{u}, \overline{v}^1, \overline{v}^2)$ be ℓ-cosparse and solve (9). We assume throughout

$$\overline{u}_i > 0,\ i \in [n] \ . \tag{14}$$

Based on $\overline{x}$, we define the corresponding support set

$$J := \{i \in [\dim(x)] \colon \overline{x}_i \neq 0\} = \operatorname{supp}(\overline{x}), \quad \overline{J} := J^c = [\dim(x)] \setminus J \ . \tag{15}$$

Denoting $k := p - \ell$ the cardinality of the index sets J and $\overline{J}$ is

$$|\overline{J}| = 2\ell + k = p + \ell, \qquad |J| = n + 2p - |\overline{J}| = n + k \ , \tag{16}$$

compare [2, Lem. 5.3]. This shows that $x \in \mathbb{R}^{n+2p}$ is a $n+k$-sparse vector.

Theorem 1 ([6, Thm. 2(iii)]). *Let $\overline{x}$ be a solution of the linear program* (9)*. The following statements are equivalent:*

(i) $\overline{x}$ is unique.
(ii) There exists no x satisfying

$$Mx = 0, \quad x_{\overline{J}} \geq 0, \quad \langle c, x \rangle \leq 0, \quad x \neq 0 \ . \tag{17}$$

Theorem (1) can be turned into a *nullspace condition* w.r.t. the sensor matrix A, for the unique solvability of problems (9) and (6).

Proposition 2 ([2, Cor. 5.3]). *Let $\overline{x} = (\overline{u}, \overline{v}^1, \overline{v}^2)$ be a solution of the linear program* (9) *with component $\overline{u}$ that has cosupport Λ with respect to B. Then $\overline{x}$, resp. $\overline{u}$, are unique if and only if*

$$\forall x = \begin{pmatrix} u \\ v \end{pmatrix}, \ v = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \quad \text{s.t.} \quad u \in \mathcal{N}(A) \setminus \{0\} \quad \textit{and} \quad Bu = v^1 - v^2 \tag{18}$$

the condition

$$\|(Bu)_\Lambda\|_1 > \big\langle (Bu)_{\Lambda^c}, \operatorname{sign}(B\overline{u})_{\Lambda^c} \big\rangle \tag{19}$$

holds. Furthermore, any unknown ℓ-cosparse vector u^*, with $Au^* = b$, can be uniquely recovered as solution $\overline{u} = u^*$ to (6) *if and only if, for all vectors u conforming to* (18)*, the condition*

$$\|(Bu)_\Lambda\|_1 > \sup_{\Lambda \subset [p] \colon |\Lambda| = \ell} \ \sup_{\overline{u} \in \mathcal{N}(B_\Lambda)} \big\langle (Bu)_{\Lambda^c}, \operatorname{sign}(B\overline{u})_{\Lambda^c} \big\rangle \tag{20}$$

holds.

Remark 1. Conditions (19) and (20) clearly indicate the direct influence of cosparsity on the recovery performance: if $\ell = |\Lambda|$ increases, then these conditions will more likely hold. On the other hand, these results are mainly theoretical since numerically checking (20) is infeasible. However we will assume that uniqueness of (6) is given, provided that the cosparsity ℓ of the unique solution $\overline{u}$ satisfies the conditions in (5a) and (5b). This assumption is motivated by the comprehensive experimental assessment of recovery properties reported in [2].

Remark 2. We note that, besides the condition for uniqueness from Thm. (1), uniqueness of a LP solution is provided in case of a unique feasible point. For high cosparsity levels ℓ, this seems to be often the case.

Let $\overline{x}$ be a (possibly unique) primal solution of (P) and $\overline{y}$ a dual solution. In view of (15) and (12) we have

$$(c - M^\top \overline{y})_i = 0, \quad \forall i \in J \ . \tag{21}$$

We note that non-degeneracy of the primal-dual pair $(\overline{x}, \overline{y})$ implies uniqueness of the dual variable $\overline{y}$.

4 Recovery by Perturbed Linear Programming

Preliminaries: Fenchel Duality Scheme. We will use the following result.

Theorem 2 ([5]). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ and $A \in \mathbb{R}^{m\times n}$. Consider the two problems*

$$\inf_{x\in\mathbb{R}^n} \varphi(x), \qquad \varphi(x) = \langle c, x\rangle + f(x) + g(b - Ax), \tag{22a}$$

$$\sup_{y\in\mathbb{R}^m} \psi(y), \qquad \psi(y) = \langle b, y\rangle - g^*(y) - f^*(A^\top y - c) \ . \tag{22b}$$

where the functions f and g are proper, lower-semicontinuous (lsc) and convex. Suppose that

$$b \in \operatorname{int}(A \operatorname{dom} f + \operatorname{dom} g), \tag{23a}$$

$$c \in \operatorname{int}(A^\top \operatorname{dom} g^* - \operatorname{dom} f^*) \ . \tag{23b}$$

Then the optimal solutions $\overline{x}, \overline{y}$ are determined by

$$0 \in c + \partial f(\overline{x}) - A^\top \partial g(b - A\overline{x}), \qquad 0 \in b - \partial g^*(\overline{y}) - A\partial f^*(A^\top \overline{y} - c) \tag{24a}$$

and connected through

$$\overline{y} \in \partial g(b - A\overline{x}), \qquad \overline{x} \in \partial f^*(A^\top \overline{y} - c), \tag{25a}$$

$$A^\top \overline{y} - c \in \partial f(\overline{x}), \qquad b - A\overline{x} \in \partial g^*(\overline{y}) \ . \tag{25b}$$

Furthermore, the duality gap vanishes: $\varphi(\overline{x}) = \psi(\overline{y})$.

Entropic Perturbation and Exponential Penalty. In various approaches to solving large-scale linear programs, one regularizes the problem by adding to the linear cost function a separable nonlinear function multiplied by a small positive parameter. Popular choices of this nonlinear function include the quadratic function, the logarithm function, and the $\langle x, \log(x)\rangle$-entropy function. Our main motivation in following this trend is that by adding a strictly convex and separable perturbation function, the dual problem will become unconstrained and differentiable. Consider

$$(P_\varepsilon) \qquad \min \langle c, x\rangle + \varepsilon\langle x, \log x - \mathbb{1}\rangle \quad \text{s.t.} \quad Mx = q, x \geq 0 \ . \tag{26}$$

The perturbation approach by the entropy function was studied by Fang et al. [4, 7] and, from a dual exponential penalty view, by Cominetti et al. [8].

The Unconstrained Dual. We write (P_ε) (26) in the form (22a)

$$\min \varphi(x), \quad \varphi(x) := \langle c, x\rangle + \underbrace{\varepsilon\langle x, \log x - \mathbb{1}\rangle + \delta_{\mathbb{R}^n_+}(x)}_{:=f(x)} + \delta_0(q - Mx) \ . \tag{27}$$

With $g := \delta_0$, we get $g^* \equiv 0$, since $\delta_C^* \equiv \sigma_C$ and thus

$$g^*(y) = \delta_0^*(y) = \sigma_0(y) = \sup_{z=0} \langle y, z\rangle = 0, \quad \forall y \in \mathbb{R}^n$$

holds. On the other hand, we have $f^*(y) = \varepsilon\langle \mathbb{1}, e^{\frac{y}{\varepsilon}}\rangle$. Now (22b) gives immediately the dual problem

$$\sup \psi(y), \quad \psi(y) := \langle q, y\rangle - \varepsilon\langle \mathbb{1}, e^{\frac{M^\top y - c}{\varepsilon}}\rangle \ . \tag{28}$$

We note that ψ is unconstrained and twice differentiable with

$$\nabla\psi(y) = q - Me^{\frac{M^\top y - c}{\varepsilon}} \qquad \text{and} \tag{29a}$$

$$\nabla^2\psi(y) = -\frac{1}{\varepsilon} M \operatorname{diag} e^{\frac{M^\top y - c}{\varepsilon}} M^\top \ . \tag{29b}$$

Moreover, $-\nabla^2\psi \succ 0$ for all y, with $e^{\frac{M^\top y - c}{\varepsilon}} \in \mathcal{R}(M) = \mathcal{N}(M)^\perp$, in view of (29b). Note that if ψ has a solution then it is unique and the strictly feasible set must be nonempty, see (29a), thus $\operatorname{rint}\mathcal{P} = \{x \colon Mx = q, x > 0\} \neq \emptyset \Leftrightarrow q \in M(\mathbb{R}^n_{++})$. Further we can rewrite (28) in more detailed form in view of (10)

$$(D_\varepsilon) \quad \min_{y_0, y_b} -\langle b, y_b\rangle + \varepsilon\langle \mathbb{1}_n, e^{\frac{B^\top y_0 + A^\top y_b}{\varepsilon}}\rangle + \varepsilon\langle \mathbb{1}_p, e^{\frac{-y_0 - \mathbb{1}_p}{\varepsilon}}\rangle + \varepsilon\langle \mathbb{1}_p, e^{\frac{y_0 - \mathbb{1}_p}{\varepsilon}}\rangle \ . \tag{30}$$

Connecting Primal and Dual Variables. With $\operatorname{dom} g = 0$, $\operatorname{dom} g^* = \mathbb{R}^n$, $\operatorname{dom} f^* = \mathbb{R}^n$ and $\operatorname{dom} f = \mathbb{R}^n_+$, the assumptions (23) become $q \in \operatorname{int} M(\mathbb{R}^n_+) = M(\operatorname{int}\mathbb{R}^n_+) = M(\mathbb{R}^n_{++})$, compare [5, Prop. 2.44], and $c \in \operatorname{int}\mathbb{R}^n = \mathbb{R}^n$. Thus, under the assumption of a strictly feasible set, we have no duality gap. Moreover both problems (27) and (28) have a solution.

Theorem 3. *Denote by x_ε and y_ε a solution of (P_ε) and (D_ε) respectively. Then the following statements are equivalent:*

(a) $q \in M(\mathbb{R}^n_{++})$, thus the strictly feasible set is nonempty.
(b) The duality gap is zero $\psi(y_\varepsilon) = \varphi(x_\varepsilon)$.
(c) Solutions x_ε and y_ε of (P_ε) and (D_ε) exist and are connected through

$$x_\varepsilon = e^{\frac{M^\top y_\varepsilon - c}{\varepsilon}} \ . \tag{31}$$

Proof. (a) $\Rightarrow$ (b): holds due to Thm. 2. On the other hand, (b) implies solvability of ψ and thus (a), as noted after Eq. (29b). (a) $\Rightarrow$ (c): The assumptions of Thm. 2 hold. Now $\partial f^*(y) = \{\nabla f^*(y)\} = \{e^{\frac{y}{\varepsilon}}\}$ and the r.h.s. of (25a) gives (c). Now, (c) implies $Mx_\varepsilon = q$ and thus (a).

The following result shows that for $\varepsilon \to 0$ and under the nonempty feasible set assumption, x_ε given by (31) approaches the least-entropy solution of (P), if y_ε is a solution of (D_ε). The proof follows along the lines of [9, Prop. 1].

Theorem 4. *Denote the solution set of (9) by $\mathcal{S}$. Assume $\mathcal{S} \neq \emptyset$. Then, for any sequence of positive scalars (ε_k) tending to zero and any sequence of vectors (x_{ε_k}), converging to some x^* , we have $x^* \in \operatorname{argmin}_{x\in S}\langle x, \log x - \mathbb{1}\rangle$. If $\mathcal{S}$ is a singleton, denoted by $\overline{x}$, then $x_{\varepsilon_k} \to \overline{x}$.*

Partial Perturbation. In the case of a unique and sparse feasible point $\overline{x}$ the assumption $q \in M(\mathbb{R}^n_{++})$ does not hold. With $J = \mathrm{supp}(\overline{x})$ the primal reads

$$\min \langle c, x\rangle + \varepsilon \langle x_J, \log x_J - \mathbb{1}_J\rangle \quad \text{s.t.} \quad Mx = q, x_{J^c} = 0, x \geq 0,$$

and the dual becomes

$$\max_y \langle q, y\rangle - \varepsilon \langle \mathbb{1}, e^{\frac{(M^\top)_J y - c_J}{\varepsilon}}\rangle \ .$$

However, the solution support J is unknown. Using (21), one can show that an approximative solution y_ε of (D_ε), i.e. $\|\nabla\psi(y_\varepsilon)\| \leq \tau_\varepsilon$, with $\tau_\varepsilon > 0$ small, can be used to construct x_ε according to (31), such that $x_\varepsilon \to \overline{x}$.

Exponential Penalty Method. We discussed above how problem (P_ε) tends to (P) as $\varepsilon \to 0$. Likewise, (D_ε) tends to (D). This was shown by Cominetti et al. [8, Prop. 3.1]. The authors noticed that the problem (D_ε) is a exponential penalty formulation of (D), compare (10) and (30).

They also investigated the asymptotic behavior of the trajectory y_ε and its relation with the solution set of (D). They proved the trajectory y_ε is approximatively a straight line directed towards the center of the optimal face of (D), namely $y_\varepsilon = y^* + \varepsilon d^* + \eta(\varepsilon)$, where y^* is a particular solution of (D). Moreover, the error $\eta(\varepsilon)$ goes to zero exponentially fast, i.e. at the speed of $e^{\frac{-\mu}{\varepsilon}}$ for some $\mu > 0$. See the proof of [8, Prop. 3.2].

5 Numerical Experiments

In this section, we illustrate the performance of our perturbation approach compared to the LP solver MOSEK, in noisy and non-noisy environments, for 2D and 3D cases. We implemented the entropic quadratic approach and solved the perturbed dual formulations by a conventional unconstrained optimization approach, the *Limited Memory BFGS* algorithm, see [10], which scales to large problem sizes. In all experiments the perturbation parameters were kept fixed to $\varepsilon = 1/50$ and $\alpha = 1$, see Fig. 3 for a justification. We allowed at maximum 1500 iterations and stopped when the norm of the gradient of the perturbed dual function satisfies $\|\nabla\psi(y^k)\| \leq 10^{-4}$.

The first performance test was done on 2D $d \times d$ images of randomly located ellipsoids with random radii along the coordinate axes. See Fig. 2 (right) for two sample images. The relative cosparsity is denoted by $\rho := \frac{\ell}{n}$. Parameters p and n vary for two- and three-dimensional images as

$$n = \begin{cases} d^2 & \text{in 2D} \\ d^3 & \text{in 3D} \end{cases} \quad , \quad p = \begin{cases} 2d(d-1) & \text{in 2D} \\ 3d^2(d-1) & \text{in 3D} \end{cases} \ . \tag{32}$$

Our parametrization relates to the design of the projection matrices $A \in \mathbb{R}^{m \times n}$, see [2] for details.

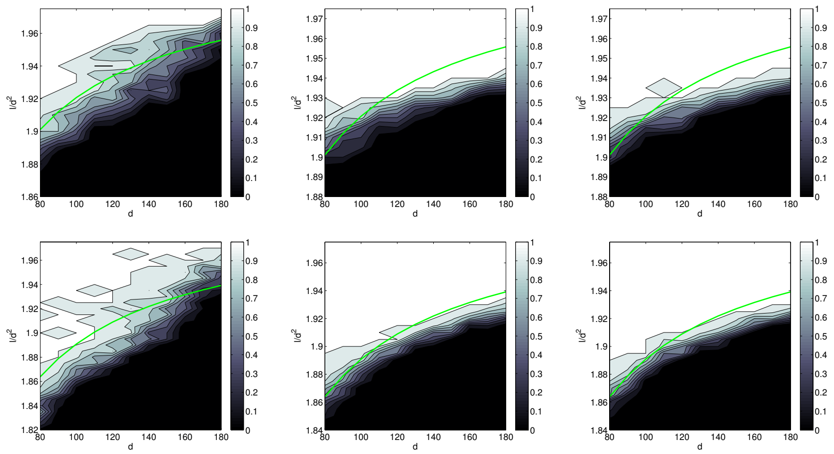

Fig. 1. Phase transitions for the 2D case, 4 cameras (top row) and 6 cameras (bottom row), computed for the noiseless case with MOSEK (left column), our approach (middle column) and our approach for the noisy case (right column). The green solid line corresponds to the theoretical curve (5a).

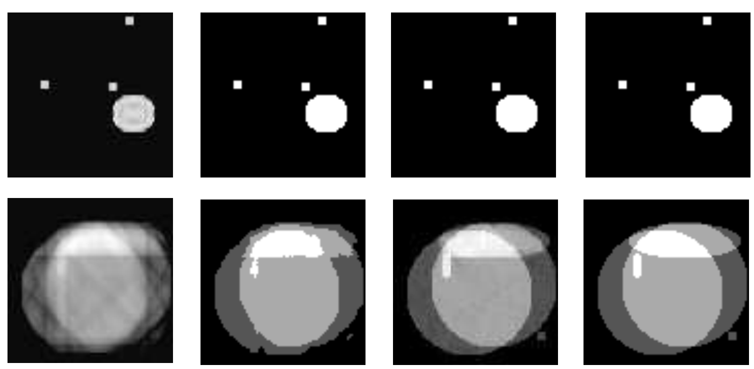

Fig. 2. Comparison between the quadratic perturbation approach (left two columns) and entropic perturbation approach (right two columns) for two relative cosparsity levels. Two 80×80 images, are projected along 6 directions. For both $\rho = \ell/d^2 = 1.7786$ (top row) and $\rho = \ell/d^2 = 1.8586$ (bottom row) reconstruction should in theory be *exact*. Result (left column) and rounded result (second left column) of the quadratic perturbation approach for $\alpha = 1$. Results for the entropic perturbation approach (right two columns) with $\varepsilon = 1/50$. Here the rounded result *exactly equals* the original image (right column).

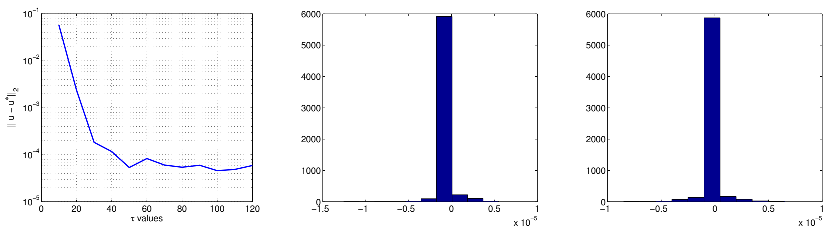

Fig. 3. Experimental finite perturbation property of the entropic approach. Here $\varepsilon = 1/50$ is a reasonable value since the reconstruction error varies insignificantly (left). The histograms of $(u - u^*)$ for $\varepsilon = 1/50$ (middle) and $\varepsilon = 1/120$ (right) are highly similar.

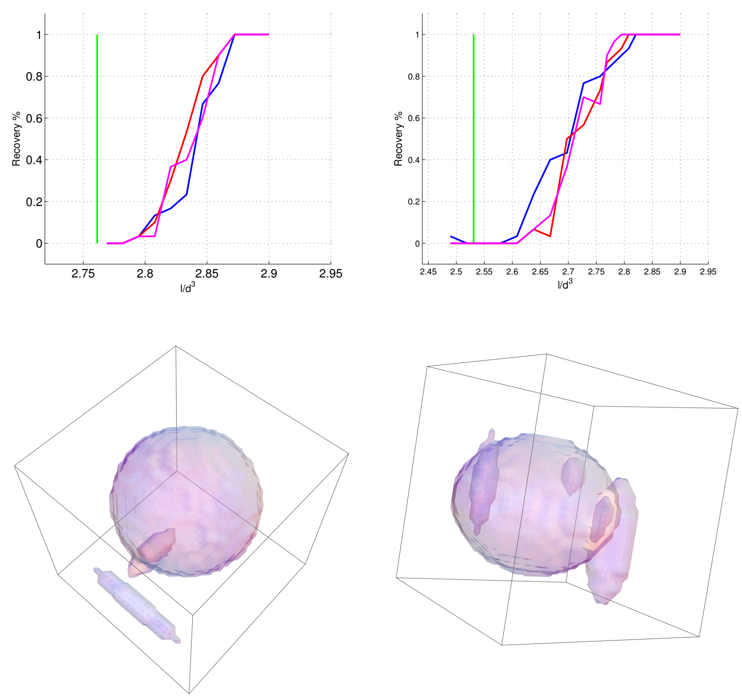

Fig. 4. Phase transitions for the 3D case, 3 cameras (top left) and 4 cameras (top right) and random example of perfectly reconstructed images $d = 31$ (bottom). The average performance of MOSEK (blue line) for the noiseless case, and the entropic approach in the noiseless (red line) and noisy (magenta line) case for $\varepsilon = 1/50$ as a variation of relative cosparsity. The green solid line corresponds to the theoretical curve (5b). Measurements were corrupted by Poisson noise of SNR = 50db.

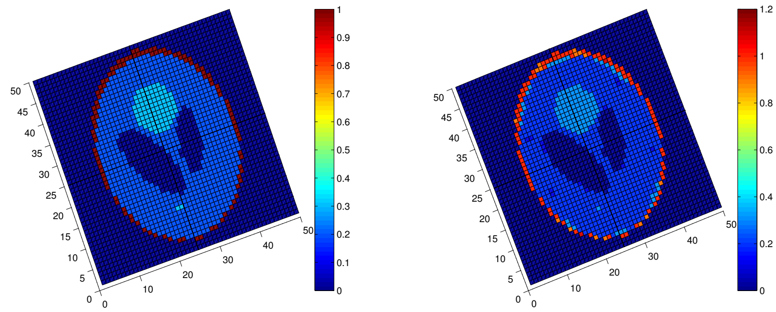

Fig. 5. Slices through the 3D volume of an original Shepp-Logan image (left) and the reconstructed image from 7 *noisy* projecting directions via the entropic perturbation approach, satisfying $\|u - u^*\|_\infty < 0.5$ (right). This shows that the approach is also stable for low noise levels as opposed to MOSEK. Measurements were corrupted by Poisson noise of SNR = 50db.

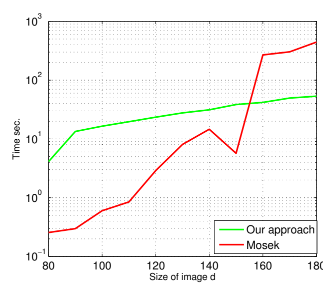

Fig. 6. Comparison between computation times of the proposed approach and MOSEK.

The phase transitions in Fig. 1 display the empirical probability of exact recovery over the space of parameters that characterize the problem. Here we performed 90 tests for each (ρ, δ) parameter combination.

We analyzed the influence of the image cosparsity, also for 3D images, see Fig. 4. In 3D for each problem instance defined by a (ρ, d)-point we generated 60 random images. In both 2D and 3D we declared a random test as successful if $\|u - u^*\|_\infty < 0.5$, which leads to perfect reconstruction after rounding. Fig. 1 and 4 display a phase transition and exhibit regions where exact image reconstruction has probability equal or close to one. The solid green line in the plots, stands for the theoretical curve (5a). In the noisy case, projection data was corrupted by Poisson noise of $SNR = 50$db. The perturbation parameter has been set as in the noiseless case, i.e. $\varepsilon = 1/50$ and $\alpha = 1$. MOSEK was unable to solve the given problem, stating that either the primal or the dual might be infeasible. The algorithm proposed in this paper scales much better with the problem size and is significantly more efficient for large problem sizes that are relevant to applications. In particular, problems sizes can be handled where MOSEK stalls, see Fig. 6. We also note that our perturbation approach is also stable to low noise levels as opposed to MOSEK. Finally, we underline that the entropic perturbation approach performs significantly better then quadratic perturbation as shown in Fig. 2.

6 Conclusion

We presented a mathematical programming approach based on perturbation that copes with large tomographic reconstruction problems of the form (1). While the perturbation enables to apply efficient sparse numerics, it does not compromise reconstruction accuracy. This is a significant step in view of the big data volumes of industrial scenarios.

Our further work will examine the relation between the geometry induced by perturbations on the u-space and the geometry of Newton-like minimizing paths, and the potential for parallel implementations.

Acknowledgement. SP gratefully acknowledges financial support from the Ministry of Science, Research and Arts, Baden-Württemberg, within the Margarete von Wrangell postdoctoral lecture qualification program. AD and the remaining authors appreciate financial support of this project by the Bayerische Forschungsstiftung.

References

1. Nam, S., Davies, M., Elad, M., Gribonval, R.: The Cosparse Analysis Model and Algorithms. Applied and Computational Harmonic Analysis **34**(1) (2013) 30–56
2. Deniţiu, A., Petra, S., Schnörr, C., Schnörr, C.: Phase Transitions and Cosparse Tomographic Recovery of Compound Solid Bodies from Few Projections. ArXiv e-prints (November 2013)
3. Ferris, M.C., Mangasarian, O.L.: Finite perturbation of convex programs. Appl. Math. Optim **23** (1991) 263–273
4. Fang, S.C., Tsao, H.S.J.: Linear programming with entropic perturbation. Math. Meth. of OR **37**(2) (1993) 171–186
5. Rockafellar, R., Wets, R.J.B.: Variational Analysis. 2nd edn. Springer (2009)
6. Mangasarian, O.L.: Uniqueness of Solution in Linear Programming. Linear Algebra and its Applications **25**(0) (1979) 151–162
7. Fang, S.C., Tsao, H.S.J.: On the entropic perturbation and exponential penalty methods for linear programming. J. Optim. Theory Appl. **89** (1996) 461–466
8. Cominetti, R., San Martin, J.: Asymptotic analysis of the exponential penalty trajectory in linear programming. Math. Progr. **67** (1994) 169–187
9. Tseng, P.: Convergence and Error Bound for Perturbation of Linear Programs. Computational Optimization and Applications **13**(1-3) (1999) 221–230
10. Bonnans, J.F., Gilbert, Lemaréchal, C., Sagastizábal, C.: Numerical Optimization – Theoretical and Practical Aspects. Springer Verlag, Berlin (2006)