Fundamenta Informaticae 125 (2013) 285–312
DOI 10.3233/FI-2013-865
IOS Press

Critical Parameter Values and Reconstruction Properties of Discrete Tomography: Application to Experimental Fluid Dynamics

Stefania Petra[*†]
Image and Pattern Analysis Group, University of Heidelberg
Speyerer Str. 6, 69115 Heidelberg, Germany
petra@math.uni-heidelberg.de

Christoph Schnörr
Image and Pattern Analysis Group, University of Heidelberg
schnoerr@math.uni-heidelberg.de

Andreas Schröder
Institute of Aerodynamics and Flow Technology, German Aerospace Center
Bunsenstr. 10, 37073 Göttingen, Germany
andreas.schroeder@dlr.de

Abstract. We analyze representative ill-posed scenarios of tomographic PIV (particle image velocimetry) with a focus on conditions for unique volume reconstruction. Based on sparse random seedings of a region of interest with small particles, the corresponding systems of linear projection equations are probabilistically analyzed in order to determine: (i) the ability of unique reconstruction in terms of the imaging geometry and the critical sparsity parameter, and (ii) sharpness of the transition to non-unique reconstruction with ghost particles when choosing the sparsity parameter improperly. The sparsity parameter directly relates to the seeding density used for PIV in experimental fluids dynamics that is chosen empirically to date. Our results provide a basic mathematical characterization of the PIV volume reconstruction problem that is an essential prerequisite for any algorithm used to actually compute the reconstruction. Moreover, we connect the sparse volume function reconstruction problem from few tomographic projections to major developments in compressed sensing.

Keywords: compressed sensing, underdetermined nonnegative linear systems, sparsity, large deviation, tail bound, limited angle tomography, TomoPIV

[*]Support by the German Research Foundation (DFG) is gratefully acknowledged, grant SCHN457/11.
[†]Address for correspondence: Image and Pattern Anal. Gr., Univ. of Heidelberg, Speyerer Str. 6, 69115 Heidelberg, Germany

1. Introduction

1.1. Overview, Motivation

In our previous recent work [13] we pointed out the close connection between a 3D tomographic measurement technique that is driving current research work in experimental fluid dynamics, and the abstract problem class concerned with the reconstruction of sparse solutions of underdetermined systems in the field of compressed sensing. This connection is relevant because in the former field knowing how to avoid spurious reconstructions is crucial for applications. Yet, a theoretical underpinning of current practice is lacking so far. On the other hand, research in compressed sensing addresses the problem to devise performance guarantees for ill-posed reconstruction problems under suitable mathematical assumptions.

In [13] we mainly worked out the gap between practice in the former field and theoretical results in the latter, and additionally provided numerical evidence that the actual gap might be considerably smaller, which calls for appropriately modifying the mathematical requirements. In fact, the reconstruction of a random sparse solution will be based on a reduced linear system with (on average) better reconstruction properties. In the present paper we elaborate the latter point and report results of a thorough corresponding numerical study.

More specifically, motivated by a key application in experimental fluid dynamics [9], we investigate conditions for a highly underdetermined nonnegative system of linear equations to have a *unique* nonnegative solution, provided it is sufficiently sparse. The solution to be reconstructed corresponds to a sparse volume function that represents a 3D image of particles immersed in a fluid, and is only known from projections. These projections are 2D images simultaneously recorded by few cameras from different viewing directions, as shown by Fig. 1.

The reconstruction of the 3D image from the 2D images employs a standard algebraic image reconstruction model, which assumes that the 3D image consists of an array of unknowns (called cells or voxels), and sets up algebraic equations relating the unknowns to the measured projection data. The latter are the pixel entries in the recorded 2D images that represent the integration of the original 3D light intensity distribution along the pixels line of sight. The number of cameras is limited (3 to 6, typically 4). As a consequence, the reconstruction problem becomes severely ill-posed.

Thus, we consider very large and severely underdetermined linear systems

$$Ax = b, \qquad A \in \mathbb{R}^{m\times n}, \qquad m \ll n, \tag{1.1}$$

with the following properties: a *very sparse* nonnegative measurement matrix A with *constant small support* of length ℓ of all column vectors,

$$A \geq 0,\ x \geq 0, \qquad \mathrm{supp}(A_{\bullet,j}) = \ell \ll m, \qquad \forall j = 1,\ldots,n, \tag{1.2}$$

and a nonnegative k-sparse solution vector x: $\mathrm{supp}(x) \leq k$. While ℓ equals the number of cameras, k relates to the particle density (equal in the present work, proportional in practice). We also consider the discretization (or resolution) parameter d and relate it to the number of discretization cells and number of measurements:

$$m = \ell \cdot O(d), \quad n = O(d^2), \quad \text{in 2D}, \tag{1.3}$$

$$m = \ell \cdot O(d^2), \quad n = O(d^3), \quad \text{in 3D}. \tag{1.4}$$

We will answer the following question:

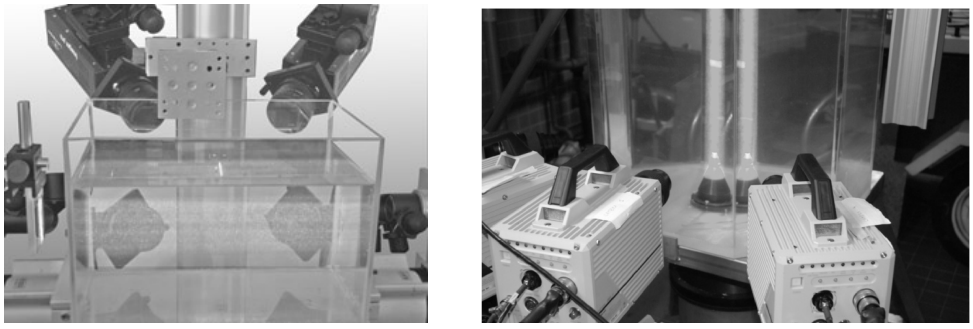

Figure 1. Typical camera arrangements: in circular configuration (right) or all in line (left).

What is the maximal number of particles, depending on the image resolution parameter d, that can be reconstructed uniquely?

Formally, we want to relate the *exact* recovery of x from it's noiseless measurements b to the sparsity k and to the dimensions of m, n of the projection matrix A. Moreover, we investigate critical values of the sparsity parameter k such that *most* k-sparse nonnegative vectors x are the *unique* nonnegative solution of (1.1) for a given b with *high probability*.

Our results provide an answer to both the question above and to the problem of devising suitable relaxations of the mathematical requirements as discussed in the beginning of this section. Moreover, our assumptions on the sensor matrix A *conform* to the imaging set-ups used in practice. Altogether, this closes the gap worked out in [13].

1.2. Related Work: Compressed Sensing

Research on compressed sensing [5, 3] focuses on properties of underdetermined linear systems that guarantee exact recovery of sparse or compressible signals x from measurements b. Donoho and Tanner [7, 8] have computed sharp reconstruction thresholds for random measurement matrices such that for a given signal length n and numbers of measurements m, the maximal sparsity value k which guarantees perfect reconstruction can be determined explicitly. The authors derived their results by connecting it to a problem from geometric probability that n points in general position in $\mathbb{R}^m$ can be linearly separated [18]. This holds with probability $\Pr(n, m) = 1$ for $n/m \leq 1$, and with $\Pr(n, m) \to 1$ if $m \to \infty$ and $1 \leq n/m < 2$, where

$$\Pr(n, m) = \frac{1}{2^{n-1}} \sum_{i=0}^{m-1} \binom{n-1}{i} . \tag{1.5}$$

The authors show in [7, Thm. 1.10] that the probability of uniqueness of a k-sparse nonnegative vector equals $\Pr(n - m, n - k)$, provided A satisfies certain conditions which do not hold in our considered application. Likewise, by exploiting again Wendel's theorem Mangasarian and Recht showed [12] that a binary solution is most likely unique if $m/n > 1/2$, provided that A comes from a centrosymmetric distribution. Unfortunately, the underlying distribution A lacks symmetry with respect to the origin. However, we recently showed in [14] that for a three camera scenario there are thresholds on sparsity (i.e. density of the particles), below which exact recovery will succeed and above which it fails with high

probability. These explicit thresholds depend on the number of measurements (recording pixels in the camera arrays). The current work investigates further geometries and focuses on an *average case analysis* of conditions under which *uniqueness* of x can be expected with *high probability*. A corresponding tail bound implies a weak threshold effect and a criterion for adequately choosing the (maximal) value of the sparsity parameter k.

1.3. Connection to Tomographic PIV and Discrete Tomography

The measurement technique *Particle Image Velocimetry (PIV)* is concerned with the inference of the instantaneous motion of fluids from image measurements. In this context, *Tomographic PIV (TomoPIV)* [9] addressed the problem to reconstruct the time-varying 3D image function in a preprocessing step. A key property of this measurement technique is that at the subsequent processing stage motion estimation techniques can invoke physical prior knowledge (e.g. incompressibility) due to working directly in 3D space. We refer to [15, 16] for further details.

A crucial requirement in this connection concerns the accuracy of the reconstructed 3D image function as input data for motion estimation. This critically depends on the particle density or – mathematically speaking – on the sparsity of the solution. Engineers aim at using high particle densities (less sparse solutions) to enable motion estimates with high spatial resolution at the subsequent processing stage. Using too high densities will inevitably lead to spurious reconstructions, and in turn to erroneous motion estimates. On the other hand, conservatively chosen low particle densities will merely lead to motion estimates at some coarse spatial resolution and hence compromise TomoPIV as an advanced measurement technique.

The present paper studies mathematically substantiated choices of maximal particle densities that provide accurate reconstructions.

Finally, we point out the connection of our work to *discrete tomography* [10]. In this research area discrete representations are used *both* for indexing the underlying domain and for the range of values of the function to be reconstructed from measurements. By contrast, in our setting a continuous bounded spatial domain is discretized by means of expanding the function of interest with compactly supported basis functions (e.g. just indicator functions of voxels), and rather than focusing on discrete-valued functions we focus on the discrete range of the number of non-vanishing coefficients. In our opinion this difference is not a significant one because computational approaches to *large-scale* discrete tomography typically involve some relaxation, for instance by replacing an integer constraint $x_i \in \{0, 1\}$ by $x_i \in [0, 1]$. This blurs the boundary the term "discrete" seems to suggest.

1.4. Contribution and Organization

Our contribution concerns the probabilistic analysis of the dimension of reduced systems (Definition 2.3) depending on the sparsity parameter k, in terms of its expected value together with a deviation bound. Sharp concentration in terms of the latter bound assures that uniqueness holds with high probability, that is informally speaking in most applications.

Such an analysis was recently introduced in [14] for the specific academical scenario considered in our previous work [13]. The present paper extends this analysis to a broader range of scenarios that represent realistic imaging set-ups. Major examples include the set-up depicted by Figure 2, that covers

2D scenarios, which can be easily extended to 3D scenarios by enhancing both camera and volume by one dimension, and a "full" 3D set-up with 4 cameras illustrated by Fig. 3.

Section 2 establishes the connection between realistic sensing matrices used in TomoPIV and recent work on sparse expander graphs in the field of compressed sensing. This enables to formulate criteria for the unique reconstruction of sparse solutions in terms of the dimension of the corresponding reduced system. Moreover, our theoretical analysis suggests that a similar procedure can be applied to different geometries varying the volume size, number of discretization cells and number of projections.

The expected dimension of various imaging set-ups are estimated in Sections 3 and 4. These estimates accurately agree with results from a series of numerical experiments that illustrate in Section 5 the separation of non-recovery and recovery by phase transitions.

We conclude and indicate directions for further research in Section 6.

1.5. Notation

$|X|$ denotes the cardinality of a finite set X and $[n] = \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$. We will denote by $\|x\|_0 = |\{i \colon x_i \neq 0\}|$ and $\mathbb{R}^n_k = \{x \in \mathbb{R}^n \colon \|x\|_0 \leq k\}$ the set of k-sparse vectors. The corresponding sets of non-negative vectors are denoted by $\mathbb{R}^n_+$ and $\mathbb{R}^n_{k,+}$ respectively. The support of a vector $x \in \mathbb{R}^n$, $\operatorname{supp}(x) \subseteq [n]$, is the set of indices of non-vanishing components of x.

For a finite set S, the set $\mathcal{N}(S)$ denotes the union of all neighbors of elements of S where the corresponding relation (graph) will be clear from the context.

$\mathbb{1} = (1, \ldots, 1)^\top$ denotes the one-vector of appropriate dimension.

$A_{\bullet,i}$ denotes the i-th column vector of a matrix A. For given index sets I and J, matrix A_{IJ} denotes the submatrix of A with rows and columns indexed by I and J respectively. I^c, J^c denote the respective complement sets. Similarly, b_I denotes a subvector of b.

$\mathbb{E}[\cdot]$ denotes the expectation operation applied to a random variable and $\Pr(A)$ the probability to observe an event A.

2. Graph Related Properties of Tomographic Projection Matrices

Recent trends in compressed sensing [2, 19] tend to replace random dense matrices by adjacency matrices of "high quality" expander graphs. Explicit constructions of such expanders exist, but are quite involved. However, random $m \times n$ binary matrices with nonreplicative columns that have $\lfloor \ell n \rfloor$ entries equal to 1, perform numerically quite well, even if ℓ is small, as shown in [2]. In [11, 13] it is shown that perturbing the elements of adjacency matrices of expander graphs with low expansion can also improve performance.

2.1. Preliminaries

For simplicity, we will confine ourselves to situations where the intersection lengths of projection rays corresponding to each camera with each discretization cell are all equal. Thus, we assume that the entries of A are binary. It will be useful to denote the set of cells by $C = [n]$ and the set of rays by $R = [m]$.

The incidence relation between cells and rays is then given by

$$(A)_{ij} = \begin{cases} 1, & \text{if } j\text{-th ray intersects } i\text{-th cell,} \\ 0, & \text{otherwise,} \end{cases} \tag{2.1}$$

for all $i \in R$, $j \in C$. Thus, cells and rays correspond to columns and rows of A.

This gives the equivalent representation in terms of a *bipartite graph* $G = (C, R; E)$ with left and right vertices C and R, and edges $(c, r) \in E$ iff $(A)_{rc} = 1$. G has *constant left-degree* ℓ equal to the number of projecting directions.

For any non-negative measurement matrix A and the corresponding graph, the set

$$\mathcal{N}(S) = \{i \in [m] \colon A_{ij} > 0,\, j \in S\}$$

contains all neighbors of S. The same notation applies to neighbors of subsets $S \subset R$ of right nodes. Further, we will call any non-negative matrix *adjacency matrix*, based on the incidence relation of its non-zero entries.

If A is the non-negative adjacency matrix of a bipartite graph with constant left degree ℓ, the **perturbed matrix** $\tilde{A}$ is computed by uniformly perturbing the non-zero entries $A_{ij} > 0$ to obtain $\tilde{A}_{ij} \in [A_{ij} - \varepsilon, A_{ij} + \varepsilon]$, and by normalizing subsequently all column vectors of $\tilde{A}$. In practice, such perturbation can be implemented by discretizing the image by radial basis functions of unequal size or by choosing their locations on an irregular grid.

The following class of graphs plays a key role in the present context and in the field of compressed sensing in general.

Definition 2.1. A (ν, δ)-unbalanced expander is a bipartite simple graph $G = (L, R; E)$ with constant left-degree ℓ such that for any $X \subset L$ with $|X| \leq \nu$, the set of neighbors $\mathcal{N}(X) \subset R$ of X has at least size $|\mathcal{N}(X)| \geq \delta\ell|X|$.

Recovery of a k-sparse nonnegative solution via an (ν, δ)-unbalanced expander was derived in [17]. It employs the smallest expansion constant δ with respect to other similar results in the literature.

Theorem 2.2. Let A be the adjacency matrix of a (ν, δ)-unbalanced expander and $1 \geq \delta > \frac{\sqrt{5}-1}{2}$. Then for any k-sparse vector x^* with $k \leq \frac{\nu}{(1+\delta)}$, the solution set $\{x \colon Ax = Ax^*, x \geq 0\}$ is a singleton.

Now let A denote the tomographic projection matrix, and consider a subset $X \subset C$ of $|X| = k$ columns and a corresponding k-sparse vector x supported on X. Then $b = Ax$ has support $\mathcal{N}(X)$, and we may remove the subset of $\mathcal{N}(X)^c = (\mathcal{N}(X))^c$ rows from the linear system $Ax = b$ corresponding to $b_r = 0$, $\forall r \in R$. Moreover, based on the observation $\mathcal{N}(X)$, we know that

$$X \subseteq \mathcal{N}(\mathcal{N}(X)) \qquad \text{and} \qquad \mathcal{N}(\mathcal{N}(X)^c) \cap X = \emptyset. \tag{2.2}$$

We continue by formalizing the system reduction just described.

Definition 2.3. The *reduced system* corresponding to a given non-negative vector b,

$$A_{red}x = b_{red}, \qquad A_{red} \in \mathbb{R}_+^{m_{red} \times n_{red}}, \tag{2.3}$$

results from A, b by choosing the subsets of rows and columns

$$R_b := \mathrm{supp}(b), \qquad C_b := \mathcal{N}(R_b) \setminus \mathcal{N}(R_b^c) \tag{2.4}$$

with

$$m_{red} := |R_b|, \qquad n_{red} := |C_b|. \tag{2.5}$$

Note that for a vector x and the bipartite graph induced by the measurement matrix A, we have the correspondence (cf. (2.2))

$$X = \mathrm{supp}(x), \qquad R_b = \mathcal{N}(X), \qquad C_b = \mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c).$$

We further define

$$\mathcal{S}^+ := \{x \colon Ax = b, x \geq 0\} \tag{2.6}$$

and

$$\mathcal{S}^+_{red} := \{x \colon A_{R_b C_b} x = b_{R_b}, x \geq 0\}\ . \tag{2.7}$$

The following proposition asserts that solving the reduced system (2.3) will always recover the support of the solution to the original system $Ax = b$

Proposition 2.4. [14, Prop. 5.1] Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ have nonnegative entries only, and let $\mathcal{S}^+$ and $\mathcal{S}^+_{red}$ be defined by (2.6) and (2.7) respectively. Then

$$\mathcal{S}^+ = \{x \in \mathbb{R}^n \colon x_{(C_b)^c} = 0 \ \text{ and } \ x_{C_b} \in \mathcal{S}^+_{red}\}. \tag{2.8}$$

Consequently, we can restrict the linear system $Ax = b$ to the subset of columns $\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c) \subset C$, and only consider properties of this reduced systems.

2.2. Guaranteed Uniqueness

Uniqueness of $x \in \mathbb{R}^n_{\delta k,+}$ is guaranteed if all k or less-sparse supported on $\mathrm{supp}(x)$ induce overdetermined reduced systems with $m_{red}/n_{red} > \delta\ell \geq \frac{\sqrt{5}-1}{2}\ell$.

Proposition 2.5. [14, Th. 3.4] Let A be the adjacency matrix of a bipartite graph such that for all random subsets $X \subset C$ of $|X| \leq k$ left nodes, the set of neighbors $\mathcal{N}(X)$ of X satisfies

$$|\mathcal{N}(X)| \geq \delta\ell|\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c)| \qquad \text{with} \qquad \delta > \frac{\sqrt{5}-1}{2}. \tag{2.9}$$

Then, for any δk-sparse nonnegative vector x^*, the solution set $\{x \colon Ax = Ax^*, x \geq 0\}$ is a singleton.

For perturbed matrices uniqueness is guaranteed for square reduced systems, and thus less sparse solutions.

Proposition 2.6. [14, Th. 3.4] Let A be the adjacency matrix of a bipartite graph such that for all subsets $X \subset C$ of $|X| \leq k$ left nodes, the set of neighbors $\mathcal{N}(X)$ of X satisfies

$$|\mathcal{N}(X)| \geq \delta\ell|\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c)| \qquad \text{with} \qquad \delta > \frac{1}{\ell}. \tag{2.10}$$

Then, for any k-sparse vector x^*, there exists a perturbation $\tilde{A}$ of A such that the solution set $\{x \colon \tilde{A}x = \tilde{A}x^*, x \geq 0\}$ is a singleton.

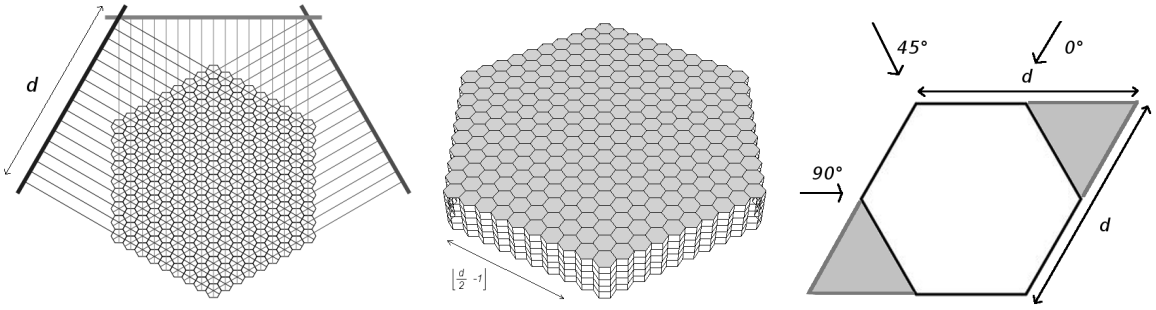

Figure 2. Sketch of a 3-cameras setup in 2D. Left: The hexagonal area discretized in $3 \cdot \frac{d^2+1}{4}$ equally sized cells is projected on three 1D cameras. The resulting projection matrix $A \in \{0,1\}^{m \times n}$ is underdetermined, with $m = 3d$ and $n = 3 \cdot \frac{d^2+1}{4}$, where $\frac{d-1}{2} + 1$ is the number of cells on each hexagon edge. Middle: This geometry can be easily extended to 3D by enhancing both cameras and volume by one dimension, thus representing scenarios of practical relevance when cameras are aligned on a line. Right: When considering a square area along with three projection directions (two orthogonal, one diagonal) one obtains a projection matrix with analogous reconstruction properties. The projection matrix *equals* up to scaling the previous projection matrix corresponding to the hexagonal area, if we remove the $2 \cdot \frac{d^2-1}{8}$ cells in the marked corners along with incident rays.

Recovery via perturbed underdetermined reduced systems is possible and our numerical results from Section 5 suggest the following.

Conjecture 2.7. Let A be the adjacency matrix of a bipartite graph such that for all subsets $X \subset C$ of $|X| \leq k$ left nodes, the set of neighbors $\mathcal{N}(X)$ of X satisfies

$$|\mathcal{N}(X)| \geq \frac{1+\delta}{\ell} |\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c)| \qquad \text{with} \qquad \delta > \frac{\sqrt{5}-1}{2}. \tag{2.11}$$

Then, for any $\frac{k}{\ell}$-sparse vector x^*, there exists a perturbation $\tilde{A}$ of A such that the solution set $\{x \colon \tilde{A}x = \tilde{A}x^*, x \geq 0\}$ is a singleton.

The consequences of Propositions 2.5, 2.6 and Conjecture 2.7 are investigated in the following sections 3.2 and 4.2 by working out critical values of the sparsity parameter k for which the respective conditions are satisfied with high probability.

3. 3 Cameras - Left Degree equals 3

In this section, we analyze the imaging set-up depicted in Figure 2, left panel, which also represents typical 3D scenarios encountered in practice, as shown in Figure 2, center panel.

3.1. Imaging Geometry

Cell centers x_c of hexagonal cells $c \in C$, that partition a region of interest, are given by lattice points corresponding to integer linear combinations of two vectors d^i, $i = \{1, 2\}$,

$$x_c = i_1 d^1 + i_2 d^2, \qquad d^1 = \frac{1}{2}\begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix}, \quad d^2 = \begin{pmatrix}0\\ 1\end{pmatrix}, \quad (i_1, i_2) \in \mathcal{I}, \tag{3.1}$$

for the index set

$$\mathcal{I} = \big\{(i,j)\colon\ -(d-1)/2 \le i, j \le (d-1)/2,\ |i+j| \le (d-1)/2\big\}, \tag{3.2}$$

with problem size $d \in \mathbb{N}$ that we assume (in this section) to be an odd number for simplicity. The number $|R|$ of projections $r \in R = R_1 \cup R_2 \cup R_3$, where R_i, $i = 1, 2, 3$ corresponds to the rays of direction i, is

$$|R| = 3|R_i| = 3d. \tag{3.3}$$

The number of cells incident with projection rays ranges over the interval

$$\{(d+1)/2, (d+1)/2 + 1, \ldots, d\} \tag{3.4}$$

from the periphery towards the center. Thus, indexing with r each projection ray along any particular direction R_i, $i = 1, 2, 3$, from one side of the hexagon across the center towards the opposite side, the numbers of cells incident with ray r is

$$|r| \in \{(d+1)/2, \ldots, d, \ldots, (d+1)/2\}, \qquad r \in R_i,\ i = 1, 2, 3. \tag{3.5}$$

The total number of cells is

$$|C| = \sum_{r \in R_i} |r| = 2 \sum_{j=(d+1)/2}^{d-1} j + d = \frac{1}{4}(3d^2 + 1), \quad i \in \{1, 2, 3\}, \tag{3.6}$$

and

$$\sum_{r \in R} |r| = 3|C|. \tag{3.7}$$

Accordingly, the system of equations representing the imaging geometry depicted by Figure 2, left panel, has dimensions

$$Ax = b, \qquad A \in \{0, 1\}^{|R| \times |C|}, \qquad b \in \mathbb{R}^{|R|}. \tag{3.8}$$

Note that $|R| \ll |C|$. For further reference, we define the quantities

$$q_r = \frac{|r|}{|C|}, \qquad p_r = 1 - q_r, \tag{3.9a}$$

$$\underline{q}_d = \min_{r \in R} q_r, \qquad \overline{q}_d = \max_{r \in R} q_r, \tag{3.9b}$$

$$\underline{p}_d = 1 - \overline{q}_d, \qquad \overline{p}_d = 1 - \underline{q}_d, \tag{3.9c}$$

and list some further relations and approximations for large d,

$$\underline{q}_d = \frac{2(d+1)}{3d^2+1} \approx \frac{2}{3d}, \qquad \overline{q}_d = \frac{1}{3}\frac{|R|}{|C|} \approx \frac{4}{3d}, \tag{3.10a}$$

$$|R|\underline{q}_d \approx 2, \qquad |R|\overline{q}_d \approx 4. \tag{3.10b}$$

3.2. Dimensions of Reduced Systems

We estimate the *expected* dimensions (2.5) of the reduced system (2.3) based on uniformly selecting k cells at random locations.

To each projection ray $r \in R$, we associate a binary random variable X_r taking the value $X_r = 1$ if *not any* of the k cells is incident with ray r, and $X_r = 0$ otherwise. We call the event $X_r = 1$ *zero-measurement*.

We are interested in the random variable

$$X = \sum_{r \in R} X_r \tag{3.11}$$

that determines the number of projection rays not incident with any of the k cells, that is the number of zero measurements. We set

$$N_R^0 := \mathbb{E}[X], \qquad N_R := |R| - N_R^0. \tag{3.12}$$

Hence, N_R is the *expected* size of the support $m_{red} = |\operatorname{supp}(b)|$ of the measurement vector b.

Remark 3.1. Note that random variables X_r are *not* independent because different projection rays may intersect. This dependency does not affect the expected value of X, but it does affect the deviation of observed values of X from its expected value – cf. Section 3.3.

Remark 3.2. We do *not* assume in the derivation below that k *different* cells are selected. In fact, a single cell may be occupied by more than a single particle in practice, because real particles are very small relative to the discretization cells c. The imaging optics enlarges the appearance of particles, and the action of physical projection rays is adequately represented by linear superposition.

Definition 3.3. (Sparsity Parameter)
We refer to the number k introduced above as sparsity parameter. Thus, highly sparse scenarios correspond to low values k.

Lemma 3.4. The expected number N_R^0 of zero measurements is

$$N_R^0 = N_R^0(k) = \mathbb{E}[X] = \sum_{r \in R} p_r^k. \tag{3.13}$$

Proof:
For $k = 1$, X_r has a Bernoulli distribution with

$$\mathbb{E}[X_r] = \Pr[X_r = 1] = 1 - \frac{|r|}{|C|} = 1 - q_r = p_r. \tag{3.14}$$

For k independent trials, we have (cf. Remark 3.2)

$$\mathbb{E}[X_r] = \Pr[X_r = 1] = p_r^k. \tag{3.15}$$

By the linearity of expectations and (3.15), we obtain (3.13),

$$N_R^0 = \mathbb{E}[X] = \sum_{r \in R} \mathbb{E}[X_r] = \sum_{r \in R} p_r^k. \tag{3.16}$$

□

We discuss few specific scenarios depending on the sparsity parameter k.

No particles For $k = 0$, we obviously have

$$N_R^0 = \sum_{r \in R} 1 = |R|.$$

High sparsity By (3.10), we have $\underline{q}_d \leq q_r \leq \overline{q}_d$, hence $q_r = \mathcal{O}(d^{-1})$. Thus, for large problem sizes d and small values of k,

$$N_R^0 \approx \sum_{r \in R} \left(\binom{k}{0} 1^k q_r^0 - \binom{k}{1} 1^{k-1} q_r^1 \right) = \sum_{r \in R} (1 - k q_r).$$

By (3.7), $\sum_{r \in R} q_r = 3$, hence

$$N_R^0 \approx |R| - 3k. \tag{3.17}$$

This approximation says that for sufficiently small values of k each randomly selected cell can be expected to create 3 independent measurements, which just reflects the fact that each cell is met by three projection rays.

Less high sparsity For increasing values of k higher-order terms cannot longer be ignored, due to the increasing number of projection rays meeting *several* particles. Taking the second-order term into account, we obtain in an analogous way

$$\begin{aligned} N_R^0 &\approx \sum_{r \in R} (1 - k q_r + \frac{k(k-1)}{2} q_r^2) \\ &\leq \sum_{r \in R} (1 - k q_r + \frac{k(k-1)}{2} q_r \overline{q}_d) = |R| - 3k + \frac{3}{2} k(k-1) \overline{q}_d, \end{aligned} \tag{3.18}$$

which is a fairly tight upper bound for values of k and N that are relevant to applications.

We consider next the *expected* number of cells $n_{red} = |C_b|$ of cells supporting the set R_b according to (2.4). We denote this expected number by

$$N_C := \mathbb{E}[|C_b|], \qquad N_C^0 := |C| - N_C, \tag{3.19}$$

and by N_C^0 the expected size of the complement.

Let $R = R_1 \cup R_2 \cup R_3$ denote the partition of all projection rays by the three directions. For each cell c, there are three unique rays $r_i(c) \in R_i$, $i = 1, 2, 3$, incident with c. Furthermore, for $i \neq j$ and some ray $r_i \in R_i$, let $R_j(r_i)$ denote the set of rays that intersect with r_i. As before, $|r|$ denotes the number of cells covered by projection ray $r \in R$.

Proposition 3.5. For a given sparsity parameter k, the expected number of cells that can be recognized as empty based on the observations of random variables $\{X_r\}_{r\in R}$ is

$$N_C^0 = N_C^0(k) = 3N_C^1 - 3N_C^2 + N_C^3, \tag{3.20a}$$

$$N_C^1 = \sum_{r\in R_i} |r| \left(1 - \frac{|r|}{|C|}\right)^k, \qquad \text{for any} \quad i \in \{1,2,3\}, \tag{3.20b}$$

$$N_C^2 = \sum_{r_i\in R_i} \sum_{r_j\in R_j(r_i)} \left(1 - \frac{|r_i| + |r_j| - 1}{|C|}\right)^k, \text{ for any } i,j \in \{1,2,3\},\ i \neq j, \tag{3.20c}$$

$$N_C^3 = \sum_{c\in C} \left(1 - \frac{\sum_{i=1}^3 |r_i(c)| - 2}{|C|}\right)^k. \tag{3.20d}$$

Proof:
Each cell intersects with three projection rays $r_i(c)$, $i = 1,2,3$. Hence, given the rays corresponding to zero measurements, each cell that can be recognized as empty if either one, two or three rays from the set $\{r_i(c)\}_{i=1,2,3}$ belong to this set.

We therefore determine separately the expected number of removable cells (i) due to individual rays corresponding to zero measurements, (ii) due to all pairs of rays that intersect and correspond to zero measurements, and (iii) due to all triples of rays that intersect and correspond to zero measurements. The estimate (3.20a) corresponding to the union of these events results from the inclusion-exclusion principle that combines these numbers so as to avoid overcounting.

Consider each projection ray $r \in R_i$ for any fixed direction $i = 1,2,3$. Because these rays do not intersect, the expected number of cells that can be removed based on the observation $\{X_r\}_{r\in R}$, is

$$N_C^1 = \mathbb{E}\Big[\sum_{r\in R_i} X_r|r|\Big] = \sum_{r\in R_i} p_r^k|r| \tag{3.21}$$

by the linearity of expectations and (3.15). Due to the symmetry of the setup, this number is the same for each direction $i = 1,2,3$. Hence we multiply N_C^1 by 3 in (3.20a).

Consider next pairs of directions $i,j \in \{1,2,3\}$, $i \neq j$. For i fixed, the expected number of empty cells based on a zero measurement corresponding to some ray $r_i \in R_i$ *and all* rays $r_j \in R_j(r_i)$ intersecting with r_i, is

$$N_C^2 = \mathbb{E}\Big[\sum_{r_i\in R_i} \sum_{r_j\in R_j(r_i)} X_{r_i}X_{r_j}\Big]. \tag{3.22}$$

The linearity of expectations and $\mathbb{E}[X_{r_i}X_{r_j}] = \Pr\big[(X_{r_i} = 1) \wedge (X_{r_j} = 1)\big]$ gives (3.20c). Due to symmetry, we have to multiply N_C^2 by 3 in (3.20a).

Finally, the expected number of empty cells that correspond to observed zero measurements along *all three* projection directions is

$$N_C^3 = \mathbb{E}\Big[\sum_{c\in C} \prod_{i=1}^3 X_{r_i(c)}\Big], \tag{3.23}$$

which equals (3.20d). □

An immediate consequence of Lemma 3.4 and Prop. 3.5 is

Corollary 3.6. For a given value of the sparsity parameter k, the expected dimensions of the reduced system (2.3) are

$$m_{red} = N_R - N_R^0, \quad n_{red} = N_C - N_C^0, \tag{3.24}$$

with N_R^0, N_C^0 given by (3.13) and (3.20).

3.3. A Tail Bound

We are interested in how sharply the random number X of zero measurements peaks around its expected value $N_R^0 = \mathbb{E}[X]$ given by (3.13).

Because the random variables X_r, $r \in R$, are *not* independent, due to the intersection of projection rays, we apply the following classical inequality for bounding the deviation of a random variable from its expected value based on martingales, that is on sequences of random variables (Y_i) defined on a finite probability space $(\Omega, \mathcal{F}, \mu)$ satisfying

$$\mathbb{E}[Y_{i+1}|\mathcal{F}_i] = X_i, \qquad \text{for all} \quad i \geq 1, \tag{3.25}$$

where $\mathcal{F}_i$ denotes an increasing sequence of σ-fields in $\mathcal{F}$ with Y_i being $\mathcal{F}_i$-measurable.

Theorem 3.7. (Azuma's Inequality [1, 4])
Let $(Y_i)_{i=0,1,2,\ldots}$ be a martingale such that for each i,

$$|Y_i - Y_{i-1}| \leq c_i. \tag{3.26}$$

Then, for all $j \geq 0$ and any $\delta > 0$,

$$\Pr\big(|Y_j - Y_0| \geq \delta\big) \leq 2\exp\Big(-\frac{\delta^2}{2\sum_{i=1}^{j} c_i^2}\Big). \tag{3.27}$$

Let $\mathcal{F}_i \subset 2^R$, $i = 0, 1, 2, \ldots$, denote the σ-field generated by the collection of subsets of R that correspond to all possible events after having observed i randomly selected cells. We set $\mathcal{F}_0 = \{\emptyset, R\}$. Observing cell $i+1$ just further partitions the current state based on the previously observed i cells by possibly removing some ray (or rays) from the set of zero measurements. Thus, we have a nested sequence (filtration) $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_k$ of the set 2^R of all subsets of R.

Based on this, for a fixed value of the sparsity parameter k, we define the sequence of random variables

$$Y_i = \mathbb{E}[X|\mathcal{F}_i], \quad i = 0, 1, \ldots, k, \tag{3.28}$$

where Y_i, $i = 0, 1, \ldots, k-1$, are the random variables specifying the expected number of zero measurements after having observed k randomly selected cells, conditioned on the subset of events $\mathcal{F}_i$ determined by the observation of i randomly selected cells. Consequently, $Y_0 = \mathbb{E}[X] = N_R^0$ due to the absence of any information, and $Y_k = X$ is just the observed number of zero measurements. The sequence $(Y_i)_{i=0,\ldots,k}$ is a martingale by construction satisfying $\mathbb{E}[Y_{i+1}|\mathcal{F}_i] = Y_i$, that is condition (3.25).

Proposition 3.8. Let $N_R^0 = \mathbb{E}[X]$ be the expected number of zero measurements for a given sparsity parameter k, given by (3.13). Then, for any $\delta > 0$,

$$\Pr\left[|X - N_R^0| \geq \delta\right] \;\leq\; 2\exp\left(-\frac{1-\overline{p}_d^2}{18(1-\overline{p}_d^{2k})}\,\delta^2\right). \tag{3.29}$$

Proof:
Let $R_{i-1}^0 \subset R$ denote the subset of rays with zero measurements after the random selection of $i-1 < k$ cells. For the remaining $k-(i-1)$ trials, the probability that not any cell incident with some ray $r \in R_{i-1}^0$ will be selected is

$$p_r^{k-(i-1)} = \mathbb{E}[X_r|\mathcal{F}_{i-1}], \tag{3.30}$$

with p_r given by (3.14). Consequently, by the linearity of expectations, the expectation Y_{i-1} of zero measurements, given the number $|R_{i-1}^0|$ of zero measurements after the selection of $i-1$ cells, is

$$Y_{i-1} = \mathbb{E}[X|\mathcal{F}_{i-1}] = \sum_{r\in R_{i-1}^0} p_r^{k-(i-1)}. \tag{3.31}$$

Now suppose we observe the random selection of the i-th cell. We distinguish two possible cases.

1. Cell i is not incident with any ray $r \in R_{i-1}^0$. Then the number of zero measurements remains the same, and

$$Y_i = \sum_{r\in R_{i-1}^0} p_r^{k-i}. \tag{3.32}$$

Furthermore,

$$\begin{aligned} Y_i - Y_{i-1} &= \sum_{r\in R_{i-1}^0} \left(p_r^{k-i} - p_r^{k-(i-1)}\right) = \sum_{r\in R_{i-1}^0} p_r^{k-i}(1-p_r) \\ &\leq \overline{p}_d^{k-i} \sum_{r\in R} q_r = 3\overline{p}_d^{k-i}. \end{aligned} \tag{3.33}$$

2. Cell i is incident with one, two or three rays contained in R_{i-1}^0. Let R_i^0 denote the set R_{i-1}^0 after removing these rays. Then

$$Y_i = \sum_{r\in R_i^0} p_r^{k-i}.$$

Furthermore, since $R_i^0 \subset R_{i-1}^0$ and $|R_{i-1}^0 \setminus R_i^0| \leq 3$,

$$\begin{aligned} Y_{i-1} - Y_i &= \sum_{r\in R_{i-1}^0\setminus R_i^0} p_r^{k-(i-1)} - \sum_{r\in R_i^0} \left(p_r^{k-i} - p_r^{k-(i-1)}\right) \\ &\leq 3\overline{p}_d^{k-i+1} - \sum_{r\in R_i^0} \overline{p}_d^{k-i} \underline{q}_d. \end{aligned}$$

Further upper-bounding by dropping the second sum shows that the resulting first term is still smaller than the bound (3.33).

As a result, we consider the larger bound (3.33) of these two cases and compute

$$\sum_{i=1}^{k} \left(3\overline{p}_d^{k-i}\right)^2 = 9\frac{1-\overline{p}_d^{2k}}{1-\overline{p}_d^2}.$$

Applying Theorem 3.7 completes the proof. □

Remark 3.9. Expanding the r.h.s. of (3.29) around 0 in terms of the variable d^{-1} shows

$$\Pr\left[|X - N_R^0| \geq \delta\right] \;\leq\; 2\exp\Big(-\frac{\delta^2}{18k}\Big) \qquad \text{for} \qquad d \to \infty. \tag{3.34}$$

This indicates appropriate choices $k = k(d)$ for large but finite problem sizes d occurring in applications, so as to bound the deviation of N_R^0 from its expected value. As a result, for such choices of k our analysis, based on expected values of the key system parameters, will hold in applications with high probability.

3.4. Critical Sparsity Values and Recovery

We derived the expected number $N_R(k)$ of nonzero measurements m_{red} (2.5) induced by random k-sparse vectors $x \in \mathbb{R}^n_{k,+}$ and the corresponding expected number $N_R(k)$ of non-redundant cells n_{red}. The tail bound, Prop. 3.8, guarantees that the dimensions of reduced systems concentrate around the derived expected values, explaining the threshold effects of unique recovery from few tomographic measurements.

We now introduce some further notations and discuss the implication of Section 2.2 on exact recovery of $x \in \mathbb{R}^n_{k,+}$. Let $N_R(k)$ and $N_C(k)$ be the expected dimensions of the reduced system induced by a random k-sparse nonnegative vector as detailed in Corollary 3.6. Let $\delta = \frac{\sqrt{5}-1}{2}$ and denote by k_δ, k_{crit} and $k_{1/\delta}$ the d-dependent sparsity values which solve the equations

$$N_R(k_\delta) = \delta\ell N_C(k_\delta), \tag{3.35}$$
$$N_R(k_{crit}) = N_C(k_{crit}), \tag{3.36}$$
$$N_R(k_{opt}) = \frac{1}{2}N_C(k_{opt}), \tag{3.37}$$
$$N_R(k_{1/\delta}) = \frac{1+\delta}{\ell}N_C(k_{1/\delta}). \tag{3.38}$$

In what follows, the phrase *with high probability* refers to values of the sparsity parameter k for which random supports $|\operatorname{supp}(b)|$ concentrate around the crucial expected value N_R according to Prop. 3.8, thus yielding a desired threshold effect.

Proposition 3.10. The system $Ax = b$, with measurement matrix A, admits unique recovery of k-sparse non-negative vectors x with high probability, if

$$k \leq \frac{N_C(k_\delta)}{1+\delta} =: \tilde{k}_\delta\,. \tag{3.39}$$

For perturbed systems we have.

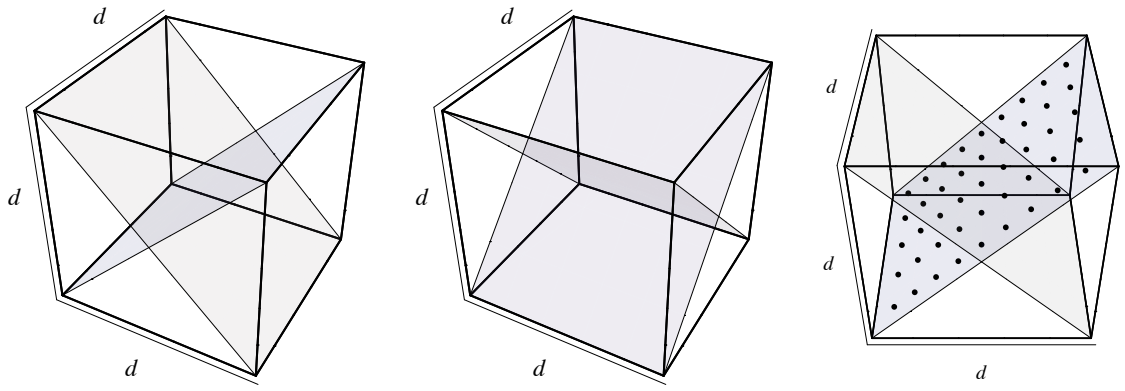

Figure 3. Imaging setup with four cameras corresponding to the image planes shown as two pairs in the left and center panel respectively. Right panel: Cell centers projected onto the first image plane are shown as dots for the case $d = 5$. The cube $\Omega = [0, d]^3$ is discretized into d^3 cells and projected along $4 \cdot d(2d - 1)$ rays.

Proposition 3.11. The system $\tilde{A}x = b$, with perturbed measurement matrix $\tilde{A}$, admits unique recovery of k-sparse non-negative vectors x with high probability, if k satisfies condition $k \leq k_{crit}$.

In case Conjecture 2.7 holds, uniqueness of $x \in \mathbb{R}^n_{k,+}$ is guaranteed if $k \leq k_{1/\delta}$. Finally, the maximal sparsity threshold is given by k_{opt}, in case reduced systems would follow a symmetric distribution with respect to the origin and columns would be in general position.

4. 4 Cameras - Left Degree equals 4

We consider the imaging set-up depicted by Figure 3 and conduct a probabilistic analysis of its recovery properties, analogous to 3.2. This scenario is straightforward to realize and should also be particularly relevant to practical applications.

4.1. Imaging Geometry

Each coordinate of the unit cube $\Omega = [0, d]^3$ is discretized into the intervals $\{0, 1, 2, \ldots, d\}$, resulting in d^3 voxels with coordinates

$$C = \big\{c = (i, j, l) - \frac{1}{2}(1, 1, 1) \colon i, j, l \in [d]\big\}. \tag{4.1}$$

There are 4 sets of parallel projection rays corresponding to the normals of the image planes depicted in Fig. 3,

$$n^1 = \frac{1}{\sqrt{2}}(-1, 0, 1), \quad n^2 = \frac{1}{\sqrt{2}}(1, 0, 1), \quad n^3 = \frac{1}{\sqrt{2}}(0, -1, 1), \quad n^4 = \frac{1}{\sqrt{2}}(0, 1, 1). \tag{4.2}$$

We denote the set of projection rays and its partition corresponding to the four directions

$$R = \cup_{l=1}^{4} R_l. \tag{4.3}$$

Each set R_i contains $(2d-1)\cdot d$ projection rays whose measurements yields a projection image with $(2d-1)\times d$ pixels. We index and denote the pixels by (s,t), and the projection rays through these pixels by

$$r_{s,t}^{i} \in R_i, \quad i \in \{1,2,3,4\}.$$

For each cell $c \in C$ indexed by $i,j,l \in [d]$ according to (4.1), we represent the corresponding pixels after a suitable transformation by

$$(s_1,t_1) = (i+l-1-d, j), \qquad s_1 \in [1-d, d-1],\ t_1 \in [d], \tag{4.4a}$$
$$(s_2,t_2) = (i-l, j), \qquad s_2 \in [1-d, d-1],\ t_2 \in [d], \tag{4.4b}$$
$$(s_3,t_3) = (i, j+l-1-d), \qquad s_3 \in [d],\ t_3 \in [1-d, d-1], \tag{4.4c}$$
$$(s_4,t_4) = (i, j-l), \qquad s_4 \in [d],\ t_4 \in [1-d, d-1]. \tag{4.4d}$$

The cardinalities of the projection rays, i.e. the number of cells covered by each projection ray, are

$$a \in \{1,2\}: \qquad |r_{s,t}^{a}| = d - |s|, \qquad s \in [1-d, d-1], \quad t \in [d], \tag{4.5a}$$
$$b \in \{3,4\}: \qquad |r_{s,t}^{b}| = d - |t|, \qquad s \in [d], \quad t \in [1-d, d-1]. \tag{4.5b}$$

We observe the symmetries

$$|r_{-s,t}^{a}| = |r_{s,t}^{a}|, \qquad |r_{t,s}^{a}| = |r_{s,t}^{b}| \tag{4.6}$$

and define

$$|r_s| := |r_{s,1}^{a}| \tag{4.7}$$

because $|r_{s,t}^{a}|$ does not vary with t. Summing up the cells covered by all rays along the first direction, for example, we obtain by (4.5a), (4.6) and (4.7),

$$\sum_{r^1 \in R_1} |r^1| = \sum_{t\in[d]} \sum_{s=1-d}^{d-1} |r_{s,t}^{1}| = d \sum_{s=1-d}^{d-1} |r_s| = d\big(d + 2\sum_{s=1}^{d-1}(d-s)\big) = d^3 = |C|.$$

We set

$$\begin{aligned} R(k_1,k_2) &:= \Big(1 - \frac{|r_{k_1}| + |r_{k_2}| - 1}{d^3}\Big)^k, \\ R(k_1,k_2,k_3) &:= \Big(1 - \frac{|r_{k_1}| + |r_{k_2}| + |r_{k_3}| - 2}{d^3}\Big)^k, \\ R(k_1,k_2,k_3,k_4) &:= \Big(1 - \frac{|r_{k_1}| + |r_{k_2}| + |r_{k_3}| + |r_{k_4}| - 3}{d^3}\Big)^k. \end{aligned} \tag{4.8}$$

In order to compute for this setup the expected size of the reduced system (2.5) for random k-sparse vectors x, we conduct an analysis analogous to Section 3.2

4.2. Dimensions of Reduced Systems

We first compute the expected number of measurements m_{red} (2.5) as a function of the sparsity parameter k.

Lemma 4.1. The expected number $m_{red} = N_R$ of non-zero measurements is

$$N_R = N_R(k) = \mathbb{E}[|\operatorname{supp}(b)|] = |R| - N_R^0 = 4d(2d-1) - N_R^0, \tag{4.9a}$$

$$N_R^0 = 4d\bigg(\Big(1 - \frac{1}{d^2}\Big)^k + 2\sum_{s=1}^{d-1}\Big(1 - \frac{s}{d^3}\Big)^k\bigg). \tag{4.9b}$$

Proof:
Taking into account symmetry, we have

$$N_R^0 = \mathbb{E}\Big[\sum_{r\in R} X_r\Big] = \sum_{r\in R} p_r^k = 4\sum_{r^1\in R_1}\Big(1 - \frac{|r^1|}{|C|}\Big)^k.$$

Applying (4.5a) yields the assertion. □

Proposition 4.2. The expected size $n_{red} = N_C$ of subset of cells that support random subsets $R_b \subset \mathbb{R}$ of observed non-zero measurements is

$$N_C = N_C(k) = d^3 - N_C^1 + N_C^2 - N_C^3 + N_C^4 \tag{4.10}$$

where

$$\begin{aligned}
N_C^1 &= 4d\bigg(d\Big(1 - \frac{1}{d^2}\Big)^k + 2\sum_{s=1}^{d-1} s\Big(1 - \frac{s}{d^3}\Big)^k\bigg),\\
N_C^2 &= 2d\sum_{i,l\in[d]} R(l+i-1-d, i-l) + 4\sum_{i,j,l\in[d]} R(l-i, l-j),\\
N_C^3 &= 2\sum_{i,j,l\in[d]}\big(R(l+i-1-d, l-i, l-j) + R(l-i, l-j, l+j-1-d)\big),\\
N_C^4 &= \sum_{i,j,l\in[n]} R(l+i-1-d, l-i, l+j-1-d, l-j),
\end{aligned} \tag{4.11}$$

and the functions R are given by (4.8).

Proof:
We consider for each cell $c \in C$ the quadruple of projection rays $(r_C^1, r_C^2, r_C^3, r_C^4)$ meeting in this cell, and the corresponding partition (4.3) of projection rays. Cell c is contained in the set C_b (2.4) supporting

R_b if not any ray of the corresponding quadruple returns a zero measurement. Thus,

$$\begin{aligned}
N_C &= \mathbb{E}\Big[\sum_{c\in C}(1-X_{r_c^1})(1-X_{r_c^2})(1-X_{r_c^3})(1-X_{r_c^4})\Big] \\
&= \sum_{c\in C}\Big(1-\sum_{i=1}^{4}\mathbb{E}[X_{r_c^i}]+\sum_{1\le i\le j\le 4}\mathbb{E}[X_{r_c^i}X_{r_c^j}]-\sum_{1\le i\le j\le l\le 4}\mathbb{E}[X_{r_c^i}X_{r_c^j}X_{r_c^l}]+\mathbb{E}[X_{r_c^1}X_{r_c^2}X_{r_c^3}X_{r_c^4}]\Big) \\
&= \sum_{c\in C}\Bigg(1-\sum_{i=1}^{4}\Big(1-\frac{|r_c^i|}{d^3}\Big)^k+\sum_{1\le i\le j\le 4}\Big(1-\frac{|r_c^i\cup r_c^j|}{d^3}\Big)^k \\
&\qquad -\sum_{1\le i\le j\le l\le 4}\Big(1-\frac{|r_c^i\cup r_c^j\cup r_c^l|}{d^3}\Big)^k+\Big(1-\frac{|\cup_{i=1}^4 r_c^i|}{d^3}\Big)^k\Bigg)
\end{aligned} \tag{4.12}$$

We consider each term in turn.

(i) As for the first term, we obviously have $|C| = d^3$.

(ii) Concerning the second term, taking symmetry into account we compute,

$$\sum_{c\in C}\sum_{i=1}^{4}\Big(1-\frac{|r_c^i|}{d^3}\Big)^k = 4\sum_{c\in C}\Big(1-\frac{|r_c^1|}{d^3}\Big)^k = 4\sum_{r^1\in R_1}\sum_{c\in r^1}\Big(1-\frac{|r_c^1|}{d^3}\Big)^k.$$

Since $r_c^1 = r^1$ for all $c \in r^1$, we obtain using (4.5),

$$\begin{aligned}
\sum_{c\in C}\sum_{i=1}^{4}\mathbb{E}[X_{r_c^i}] &= 4\sum_{r^1\in R_1}|r^1|\Big(1-\frac{|r^1|}{d^3}\Big)^k = 4d\sum_{s_1=1-d}^{d-1}|r^1_{s_1,t_1}|\Big(1-\frac{|r^1_{s_1,t_1}|}{d^3}\Big)^k \\
&= 4d\Bigg(d\Big(1-\frac{1}{d^2}\Big)^k+2\sum_{s=1}^{d-1}s\Big(1-\frac{s}{d^3}\Big)^k\Bigg).
\end{aligned}$$

(iii) Concerning the third term of (4.12), we consider first the contribution of the pair of directions $(i, j) = (1, 2)$. Replacing cell-indices of projection rays by pixel-indices according to (4.4) and using (4.5) and (4.7) we have,

$$\begin{aligned}
\sum_{c\in C}\Big(1-\frac{|r_c^1\cup r_c^2|}{d^3}\Big)^k &= \sum_{i,j,l\in[d]}\Big(1-\frac{|r^1_{(i+l-1-d,j)}|+|r^2_{(i-l,j)}|-1}{d^3}\Big)^k \\
&= d\sum_{i,l\in[d]}\Big(1-\frac{|r_{i+l-1-d}|+|r_{i-l}|-1}{d^3}\Big)^k \\
&= d\sum_{i,l\in[d]}R(i+l-1-d,i-l),
\end{aligned} \tag{4.13}$$

where the factor d appears because the summand does not depend on j by (4.5a), and R is defined by (4.8). For the pair of directions $(3, 4)$, we get

$$\sum_{c\in C}\Big(1-\frac{|r_c^3\cup r_c^4|}{d^3}\Big)^k = \sum_{i,j,l\in[d]}\Big(1-\frac{|r^3_{(i,j+l-1-d)}|+|r^4_{(i,j-l)}|-1}{d^3}\Big)^k, \tag{4.14}$$

which equals (4.13) due to the symmetry (4.6).

Next, we consider the pair of directions $(1,3)$. Taking into account the symmetry (4.6) and using (4.8), we obtain

$$\begin{aligned}\sum_{c\in C}\Big(1-\frac{|r_c^1\cup r_c^3|}{d^3}\Big)^k &= \sum_{i,j,l\in[d]}\Big(1-\frac{|r^1_{(i+l-1-d,j)}|+|r^3_{(i,j+l-1-d)}|-1}{d^3}\Big)^k\\ &= \sum_{i,j,l\in[d]}\Big(1-\frac{|r_{i+l-1-d}|+|r_{j+l-1-d}|-1}{d^3}\Big)^k = \sum_{i,j,l\in[d]}\Big(1-\frac{|r_{l-i}|+|r_{l-j}|-1}{d^3}\Big)^k\end{aligned} \tag{4.15}$$

In the same way, it can be shown that the remaining pairs of directions $(1,4),(2,3),(2,4)$ each contributes to the last expression.

(iv) Concerning the fourth term of (4.10), we get for the triple of directions $(1,2,3)$ the contribution

$$\begin{aligned}\sum_{c\in C}\Big(1-\frac{|r_c^1\cup r_c^2\cup r_c^3|}{d^3}\Big)^k &= \sum_{i,j,l\in[n]}\Big(1-\frac{|r_{i+l-1-d}|+|r_{i-l}|+|r_{j+l-1-d}|-2}{d^3}\Big)^k\\ &= \sum_{i,j,l\in[n]} R(i+l-1-d,l-i,j+l-1-d)\big),\end{aligned} \tag{4.16}$$

and likewise for the remaining triples

$$\begin{aligned}(1,2,4)\colon &\quad \sum_{i,j,l\in[n]} R(i+l-1-d,l-i,l-j),\\ (1,3,4)\colon &\quad \sum_{i,j,l\in[n]} R(i+l-1-d,j+l-1-d,l-j),\\ (2,3,4)\colon &\quad \sum_{i,j,l\in[n]} R(l-i,j+l-1-d,l-j).\end{aligned} \tag{4.17}$$

Evidently, the first and last pair of expressions are equal, respectively.

(v) Finally, the expression for the last term of (4.12) is immediate.

□

We conclude this section by stressing that critical sparsity values k_δ (3.35), k_{crit} (3.36), k_{opt} (3.37), $k_{1/\delta}$ (3.38), can be worked out based on the just derived values N_R and N_C. A tail bound may be derived analogously to Prop. 3.5. We omit this redundant detail due to space constraints.

5. Numerical Experiments and Discussion

In this section we relate the previously derived bounds on the required sparsity that guarantee unique nonnegative or binary k-sparse solutions to numerical experiments. In analogy to [6] we assess the so called *phase transition* ρ as a function of d, which is reciprocally proportional to the undersampling ratio $\frac{m}{n}\in(0,1)$. We vary d, build a specific matrix projection matrix A along with its perturbed version $\tilde{A}$ and consider the sparsity as a fraction of d in 2D or d^2 in 3D. respectively. Thus, $k=\rho d^{D-1}$, with $\rho\in(0,4)$ and $D\in\{2,3\}$. This phase transition $\rho(d)$ indicates the necessary relative sparsity

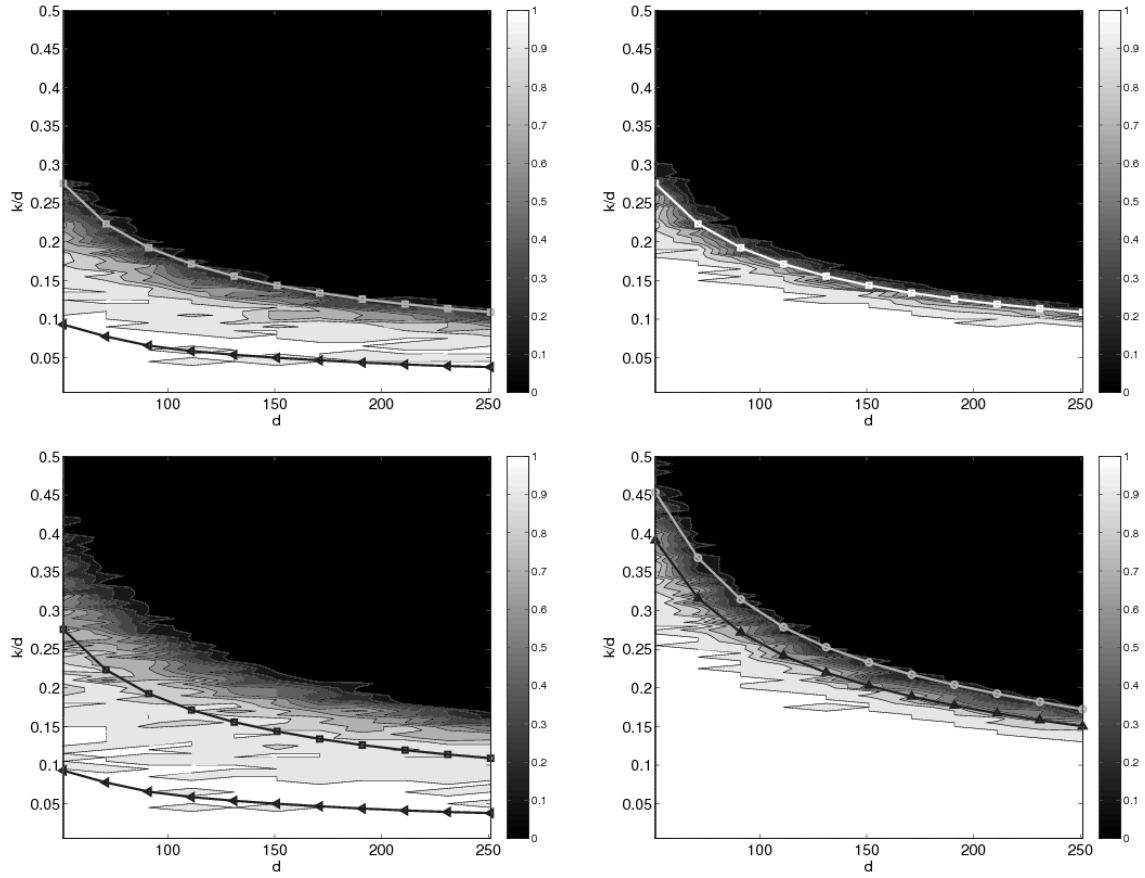

Figure 4. Success and failure empirical phase-transitions for the case of 2D, 3-cameras and hexagonal area discussed in Section 3, Fig. 2, left. Reduced unperturbed (top left) and perturbed (top right) matrices are overdetermined and of full rank with high probability, if the corresponding sparsity level is below k_δ (unperturbed case) or k_{crit} (perturbed case). The $\triangleleft$-marked curve depicts $\tilde{k}_\delta/d^2$ (3.35), and the $\square$-marked curve k_{crit}/d^2 in (3.35). Probability of uniqueness in $[0,1]^n$ of a $k = \rho d^2$ sparse binary vector for unperturbed (bottom left) and perturbed matrices (bottom right). This probability is high, below $\tilde{k}_\delta/d^2$, and decreases slowly, for the unperturbed case (bottom left). In the perturbed case the empirical probability of uniqueness exhibits a sharp transition accurately described by the $\triangleright$-marked curve $k_{1/\delta}/d^2$ from (3.38), which for the 3 camera case lies below the $\circ$-marked curve k_{opt} from (3.37).

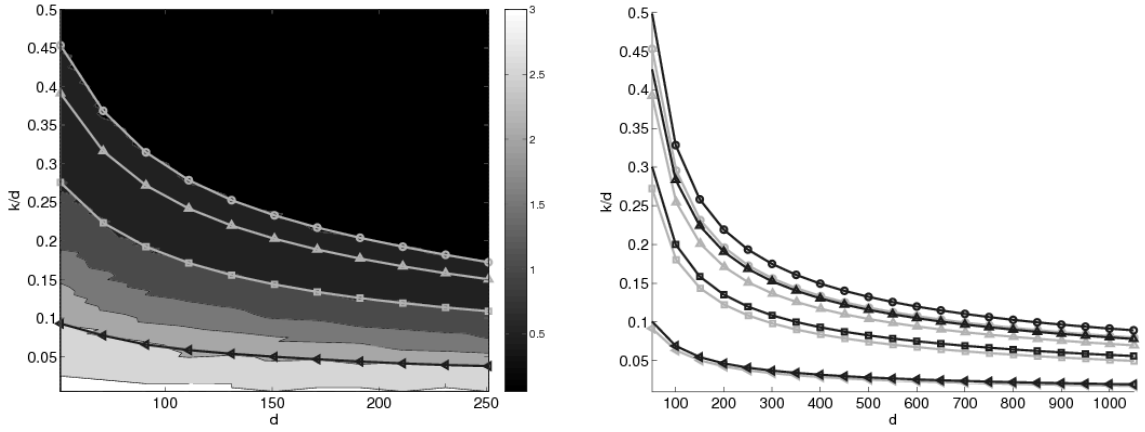

Figure 5. Left: The analytically derived curves from Section 3 correctly follow the contour lines of the average fraction of reduced systems determined empirically as a function of d and relative sparsity k/d. From bottom to top: $\tilde{k}_\delta/d$ ($\triangleleft$-marked curve), k_{crit}/d ($\square$-marked curve), k_{opt}/d ($\circ$-marked curve) and $k_{1/\delta}/d$ ($\triangleleft$-marked curve). Right: These curves are plotted again (light gray) for a wider range and compared to the analogous curves for the geometry in Fig. 2, right, square area, 2 orthogonal, one diagonal projecting direction.

to recover a k-sparse solution with high probability. More precisely, if $\|x\|_0 \leq \rho(d) \cdot d^{D-1}$, then with high probability a random k-sparse nonnegative (or binary) vector x^* is the unique solution in $\mathcal{F}_+ := \{x \colon Ax = Ax^*, x \geq 0\}$ or $\mathcal{F}_{0,1} := \{x \colon Ax = Ax^*, x \in [0,1]^n\}$ respectively. Uniqueness can be "verified" by minimizing and maximizing the same objective $f^\top x$ over $\mathcal{F}_+$ or $\mathcal{F}_{0,1}$, respectively. If the minimizers coincide for several random vectors f we claim uniqueness. The resulting linear programs we solved by a standard LP solver [1]. As shown e.g. in Fig. 9 and confirmed by all our numerical experiments, the threshold for a unique nonnegative solution and a unique $0/1$-bounded solution are quite close, especially for high values of d.

We note that $\tilde{A}$ has the same sparsity structure as A, but random entries drawn from the standard uniform distribution on the open interval $(0.9, 1.1)$.

Further, with $\rho \in [0,1]$ a ρd^{D-1}-sparse nonnegative or binary vector was generated to compute the right hand side measurement vector. For each (d,ρ)-point 50 random problem instances were generated. A threshold-effect is clearly visible in all figures exhibiting parameter regions where the probability of exact reconstruction is close to one and it is much stronger for the perturbed systems. The results are in excellent agreement with the analytically derived thresholds. We refer to the figure captions for detailed explanations and stress that a threshold-effect is clearly visible in all figures exhibiting parameter regions where the probability of exact reconstruction is close to one.

5.1. 2D: Hexagonal Volume and 3 Cameras

In this 2D, 3-cameras case, we generated A according to the geometry described in Section 3, Fig. 2, left. We considered $d \in \{51, 71, 91, \ldots, 251\}$ and varied the sparsity $k = \rho d$, by varying ρ in

[1] MOSEK, `http://www.mosek.com/`

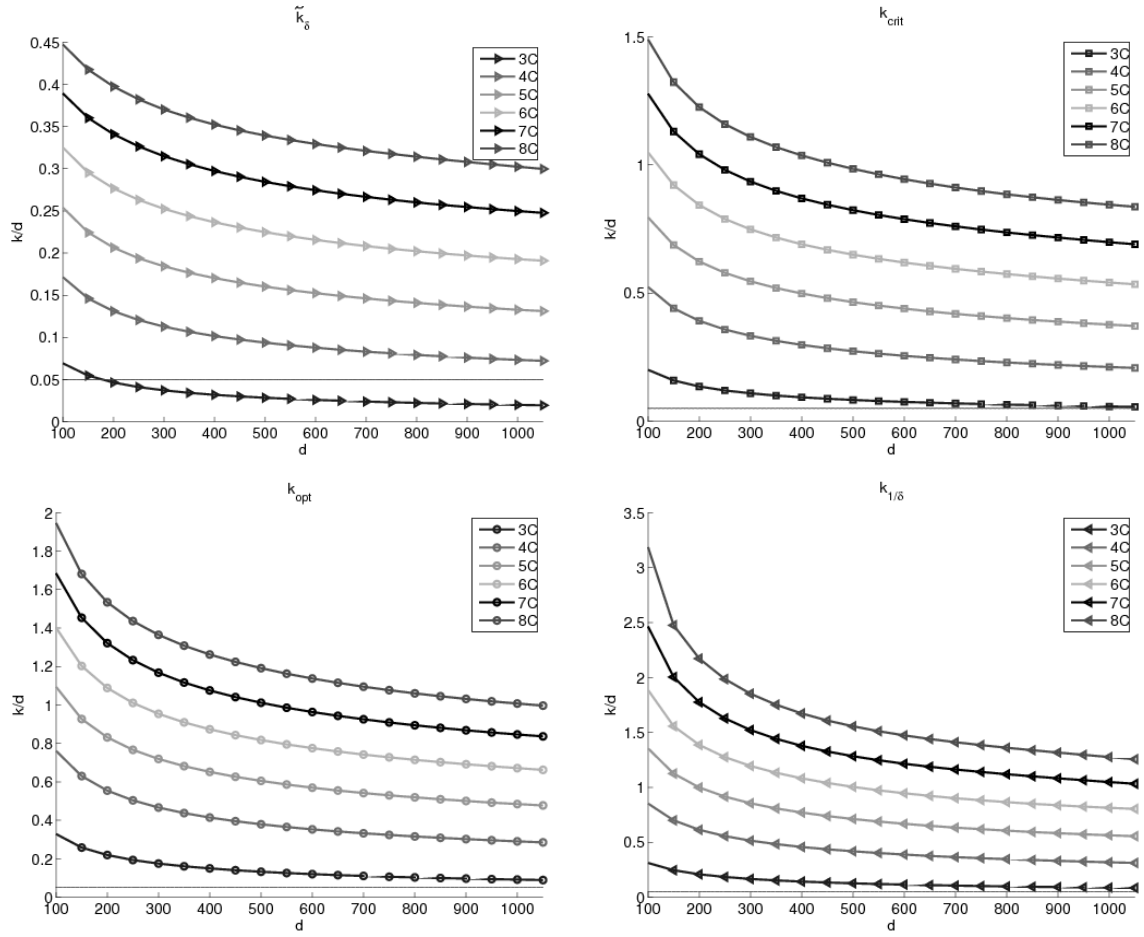

Figure 6. Empirical relative critical curves in for 3 to 8 cameras in 2D. Top left: $\tilde{k}_\delta/d$ ($\triangleright$-marked curve). Top right: k_{crit}/d ($\square$-marked curve). Bottom left: k_{opt}/d ($\circ$-marked curve). Bottom right: $k_{1/\delta}/d$ ($\triangleleft$-marked curve). Perfect recovery is possible for sparsity levels below $k_{1/\delta}$ for perturbed systems. We note that the thin horizontal line at 0.05 represents the most common seeding density currently used in practical TomoPIV systems.

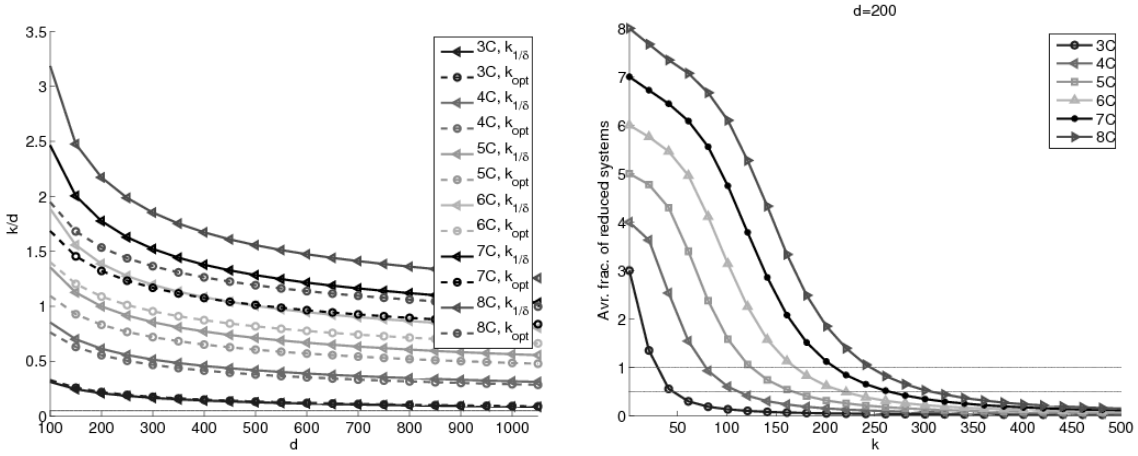

Figure 7. Left: k_{opt}/d ($\circ$-marked curve) along with $k_{1/\delta}/d$ ($\triangleleft$-marked curve). For 3 cameras $k_{1/\delta}/d$ lies below k_{opt}/d, but starting with 5 cameras $k_{1/\delta}$ significantly outperforms k_{opt}. This shows the fundamental difference between the considered 0/1-matrices and random matrices underlying a symmetrical distribution with respect to the origin. For random matrices recovery of k-sparse positive or binary vectors with sparsity levels beyond k_{opt} would be *impossible*. Right: For $d = 200$ the 6 curves depict the average ratio of $m_{red}(k)/n_{red}(k)$ as a function of sparsity k for 3 to 8 cameras from bottom to top.

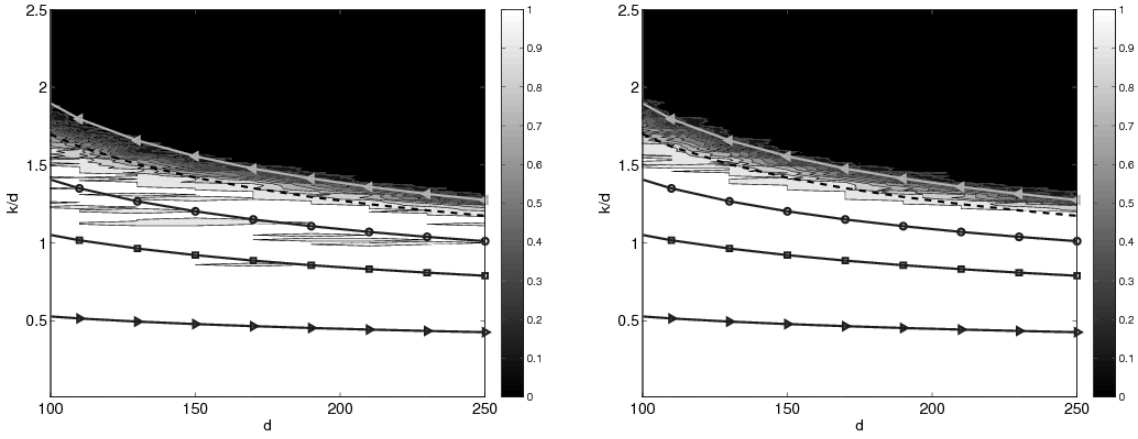

Figure 8. Empirical recovery and non-recovery phase transitions for the 2D, 6 cameras case. Probability of uniqueness in $[0,1]^n$ of a $k = \rho d^2$ sparse binary vector for unperturbed (left) and perturbed matrices (right), along with $\tilde{k}_\delta/d^2$ ($\triangleright$-marked curve), k_{crit}/d^2 ($\square$-marked curve), k_{opt}/d^2 ($\circ$-marked curve), and $k_{1/\delta}/d^2$ ($\triangleleft$-marked curve) from bottom to top. The dashed line depicts k/d^2, with k solving $m_{red}(k) = \frac{2}{\ell} n_{red}(k)$, which accurately follows the border of the highly success area *for all* considered number of cameras, $3, 4 \ldots 8$. Recovery is possible *beyond* k_{opt}, accurately described by $k_{1/\delta}$. In the 6 cameras case there is no evident performance boost for perturbed systems, since the columns of reduced systems are most likely to be in general position for both perturbed and unperturbed systems. However, in the perturbed case recovery is more stable.

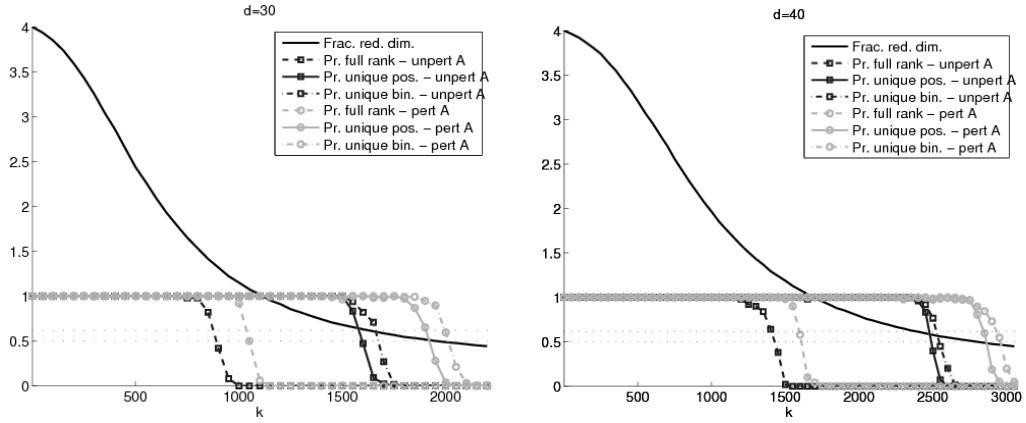

Figure 9. Recovery via the unperturbed matrix from Section 4, Fig. 3 (dark gray □-marked curves), for $d = 30$ (left) and $d = 40$ (right) versus the perturbed counterpart (light gray ∘-marked curves). The dashed lines depict the empirical probability (500 trials) that reduced systems are overdetermined and of full rank. The solid lines show the probability that a k sparse nonnegative vector is unique. The dash-dot curves shows the probability that a k sparse binary solution is the unique solution of in $[0, 1]^n$. Additional information like binarity gives only a slight performance boost, as d increases. The curve k_δ (3.35) correctly predicts that 500 ($d = 30$) and 787 ($d = 40$) particles are reconstructed with high probability via the unperturbed systems. Up to the sparsity level 1136 ($d = 30$) and 1714 ($d = 40$) perturbed systems are overdetermined and of full rank according to k_{crit} (3.36). Moreover, perturbed systems have a unique k sparse solution if k lies in between k_{opt} - 2028 ($d = 30$) and 2856 ($d = 40$) and $k_{1/\delta}$ - 2136 ($d = 30$) and 3128 ($d = 40$).

$(0, 0.5)$ with constant stepsize 0.01. The obtained empirical phase transitions along with the analytically determined ones are depicted in Fig. 4. Fig. 5 additionally shows how these analytically determined critical curves compare to the critical curves obtained empirically for the geometry in Fig. 2, right. The $(4d-1) \times d^2$ projection matrix corresponding to a square area and three projecting directions has similar reconstruction properties and the critical curves slightly change by factor $\frac{10}{9}$.

5.2. 2D: Square Volume and 3 to 8 Cameras

In our theoretical analysis in the previous sections we derived the expected number of nonzero rows $N_R(k)$ induced by the k-sparse vector along with the number $N_C(k)$ of "active" cells which cannot be empty. This can be done also empirically, see e.g. Fig. 5, left. We did this in 2D up to 8 projecting directions and obtained *empirically* the critical curves k_δ, k_{crit}, k_{opt} and $k_{1/\delta}$. To generate the curves we varied $k/d \in (0, 4)$ by a constant stepsize 0.01 and $d \in \{50, 100, \cdots, 1050\}$, and generated for each point $(k/d, d)$ 500 problem instances. Further, we determined the contour lines of $N_R(k/d, d)/N_C(k/d, d)$ corresponding to the levels $\{\delta\ell, 1, 0.5, \frac{1+\delta}{\ell}\}$. The relative sparsity curves k_δ/d, k_{crit}/d, k_{opt}/d and $k_{1/\delta}/d$ are plotted in Fig. 6 and accurately follow the empirical recovery thresholds, as shown e.g. in Fig. 8 for 6 cameras. Fig.7 shows $N_R(k, 200)/N_C(k, 200)$ for varying number of cameras. The projection angles we chosen such that the intersection with all cells is constant, yielding binary projection matrices after scaling. Each camera resolution differs with different angle. We summarize the used parameters in Table 1.

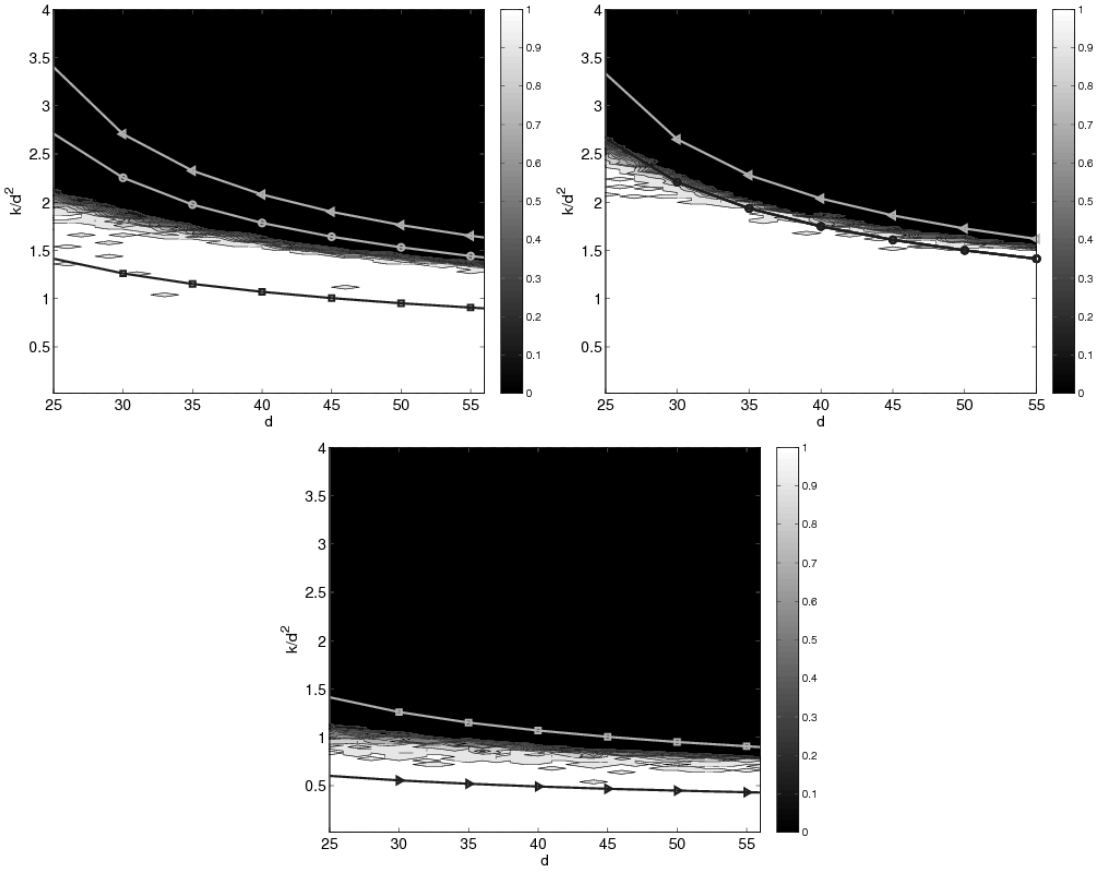

Figure 10. Success and failure empirical phase-transitions for the "full" 3D, 4 cameras case, from Section 4, Fig. 3. Probability of uniqueness in $[0,1]^n$ of a $k = \rho d^2$ sparse binary vector for unperturbed (top left) and perturbed matrices (top right). The $\triangleright$-marked curve depicts $\tilde{k}_\delta/d^2$ (3.35), the $\square$-marked curve k_{crit}/d^2 (3.36), the $\circ$-marked curve k_{opt} (3.37) and the $\triangleleft$-marked curve $k_{1/\delta}/d^2$ (3.38). In case of the perturbed matrix $\tilde{A}$ exact recovery is possible *beyond* k_{opt}/d^2. Moreover $k_{1/\delta}/d^2$ follows most accurately the empirical phase transition for perturbed systems for high values of d. The empirical probability that the reduced unperturbed matrices are overdetermined and of full rank (bottom figure), exhibits a threshold in between the estimated relative critical sparsity level k_δ and k_{crit}. This explains the performance boost for perturbed systems.

#	m	n	projection angles
3rd camera	$4d-1$	d^2	$0°, 90°, 45°$
4th camera	$6d-2$	d^2	$0°, 90°, \mp 45°$
5th camera	$7d+\lfloor\frac{d}{2}\rfloor-2$	d^2	$0°, 90°, \mp 45°, \arctan(2)$
6th camera	$8d+2\lfloor\frac{d}{2}\rfloor-2$	d^2	$0°, 90°, \mp 45°, \mp \arctan(2)$
7th camera	$9d+3\lfloor\frac{d}{2}\rfloor-2$	d^2	$0°, 90°, \mp 45°, \mp \arctan(2), \arctan(0.5)$
8th camera	$10d+4\lfloor\frac{d}{2}\rfloor-2$	d^2	$0°, 90°, \mp 45°, \mp \arctan(2), \mp \arctan(0.5)$

Table 1. Dimensions of full projection matrices.

5.3. 3D: 4 Cameras

For the 3D case, we consider the matrix from Section 4, Fig. 3, vary $d \in \{25, 26, \ldots, 55\}$ and $k = \rho d^2$ by varying $\rho \in (0, 4)$ with stepsize 0.01. For larger values of d, the empirical thresholds follow accurately the estimated curves as shown in Fig. 10. Note that for $d = 55$, A is a 48180×166375 matrix.

6. Conclusion

The measurement paradigm of compressed sensing seeks to capture the "essential" aspects of a high-dimensional but sparse object, using as few measurements as possible, by randomization. Provided that the measurements satisfy certain properties nonnegative, sparse signals can be reconstructed exactly from a surprisingly small number of samples. Moreover, there exist precise thresholds on sparsity such that any nonnegative solution that is sparser than the threshold, is also the unique nonnegative solution of the underlying linear system.

Tomographic projection matrices do not satisfy the conditions that allow to apply directly these results on the image reconstruction problem from few tomographic projections, even if the solution is very sparse. We show in the present work that there exist thresholds on sparsity depending on the numbers of measurements below which uniqueness is guaranteed and recovery will succeed, and above which it will fail *with high probability*. Successful recoveries yield perfect reconstructions without any ghost-particles.

Concerning applied practical work in experimental fluid dynamics, our results provide for the first time a mathematically substantiated basis for deciding what imaging setups will work. On the theoretical side, our work highlights a striking scenario where most recent results in the field of compressed sensing have an immediate impact on advanced measurement techniques in some application area.

Moreover, our analysis shows how similar sparsity thresholds can be derived for different geometries by varying the number of discretization cells in an arbitrary volume, the number of projections and projecting directions.

Our future work will focus on condition (2.11) that is slightly superior to (2.10) and supported by our numerical experiments. However, the constants of (2.11) have not been proven yet. Another line of research concerns the design problem of inserting a new camera given the position of already positioned cameras. Noise sensitivity and stability of solutions will be also addressed in future work.

References

[1] Azuma, K.: Weighted Sums of Certain Dependent Random Variables, *Tohoku Math. J.*, **19**(3), 1967, 357–367.

[2] Berinde, R., Indyk, P.: Sparse Recovery Using Sparse Random Matrices, 2008, MIT-CSAIL Technical Report.

[3] Candes, E. J., Romberg, J., Tao, T.: Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information, *IEEE Trans. Inf. Theor.*, **52**(2), 2006, 489–509.

[4] DasGupta, A.: *Asymptotic Theory of Statistics and Probability*, Springer, 2008.

[5] Donoho, D.: Compressed Sensing, *IEEE Trans. Inf. Theor.I*, **52**(4), 2006, 1289–1306.

[6] Donoho, D., Tanner, J.: Sparse Nonnegative Solution of Underdetermined Linear Equations by Linear Programming, *Proc. National Academy of Sciences*, **102**(27), 2005, 9446–9451.

[7] Donoho, D., Tanner, J.: Counting the Faces of Randomly-Projected Hypercubes and Orthants, with Applications, *Discrete Comput. Geom.*, **43**(3), 2010, 522–541.

[8] Donoho, D., Tanner, J.: Precise Undersampling Theorems, *Proceedings of the IEEE*, **98**(6), 2010, 913–924, ISSN 0018-9219.

[9] Elsinga, G., Scarano, F., Wieneke, B., van Oudheusden, B.: Tomographic Particle Image Velocimetry, *Exp. Fluids*, **41**, 2007, 933–947.

[10] G.T. Herman, G., Kuba, A.: *Discrete Tomography: Foundations, Algorithms and Applications*, Birkhäuser, 1999.

[11] Khajehnejad, M., Dimakis, A., Xu, W., Hassibi, B.: Sparse Recovery of Positive Signals with Minimal Expansion, *IEEE Trans. Sig. Proc.*, **59**, 2011, 196–208.

[12] Mangasarian, O., Recht, B.: Probability of Unique Integer Solution to a System of Linear Equations, *Eur. J Oper. Res.*, **214**(1), 2011, 27–30.

[13] Petra, S., Schnörr, C.: TomoPIV meets Compressed Sensing, *Pure Math. Appl.*, **20**(1-2), 2009, 49–76.

[14] Petra, S., Schnörr, C.: Average Case Recovery Analysis of Tomographic Compressive Sensing, arXiv:1208.5894v2 [math.NA], August 30 2012.

[15] Petra, S., Schröder, A., Schnörr, C.: 3D Tomography from Few Projections in Experimental Fluid Mechanics, in: *Imaging Measurement Methods for Flow Analysis* (W. Nitsche, C. Dobriloff, Eds.), vol. 106 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, Springer, 2009, 63–72.

[16] Vlasenko, A., Schnörr, C.: Variational Approaches for Model-Based PIV and Visual Fluid Analysis, in: *Imaging Measurement Methods for Flow Analysis* (W. Nitsche, C. Dobriloff, Eds.), vol. 106 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, Springer, 2009, 247–256.

[17] Wang, M., Xu, W., Tang, A.: A Unique "Nonnegative" Solution to an Underdetermined System: From Vectors to Matrices, *IEEE Trans. Sig. Proc.*, **59**(3), 2011, 1007–1016.

[18] Wendel, J.: A Problem in Geometric Probability, *Math. Scand.*, **11**, 1962, 109–111.

[19] Xu, W., Hassibi, B.: Efficient Compressive Sensing with Deterministic Guarantees Using Expander Graphs, *Information Theory Workshop, 2007. ITW '07. IEEE*, 2007.