AVERAGE CASE RECOVERY ANALYSIS
OF TOMOGRAPHIC COMPRESSIVE SENSING

STEFANIA PETRA, CHRISTOPH SCHNÖRR

ABSTRACT. The reconstruction of three-dimensional sparse volume functions from few tomographic projections constitutes a challenging problem in image reconstruction and turns out to be a particular instance problem of compressive sensing. The tomographic measurement matrix encodes the incidence relation of the imaging process, and therefore is not subject to design up to small perturbations of non-zero entries. We present an average case analysis of the recovery properties and a corresponding tail bound to establish weak thresholds, in excellent agreement with numerical experiments. Our result improve the state-of-the-art of tomographic imaging in experimental fluid dynamics by a factor of three.

1. INTRODUCTION

Research on compressive sensing [8, 3] focuses on properties of underdetermined linear systems

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

that ensure the accurate recovery of sparse solutions x from observed measurements b. Strong assertions are based on random ensembles of measurement matrices A and measure concentration in high dimensions that enable to prove good recovery properties with high probability [9, 4].

A common obstacle in various application fields are the limited options for *designing* a measurement matrix so as to exhibit desirable mathematical properties, are very limited. Accordingly, recent research has also been concerned with more restricted scenarios, spurred by their relevancy to applications (cf. Section 2.3).

Consequently, we consider a representative scenario, motivated by applications in experimental fluid dynamics (Fig. 1). A suitable mathematical abstraction of this setup gives rise to a huge and severely underdetermined linear system (1.1) that has additional properties: a *very sparse* nonnegative measurement matrix A with *constant small support* of all column vectors, and a nonnegative sparse solution vector x:

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

Our objective is the usual one: relating accurate recovery of x from given measurements b to the sparsity $k = \mathrm{supp}(x)$ of the solution x and to the dimensions m, n of the measurement matrix A. The sparsity parameter k has an immediate physical interpretation (Fig. 1). Engineers require high values of k, but are well aware that too high values lead to spurious solutions. The current practice is based on a rule of thumb leading to conservative low values of k.

In this paper, we are concerned with working out a better compromise along with a mathematical underpinning. The techniques employed are general and only specific to the class of linear systems (1.1), (1.2), rather than to a particular application domain.
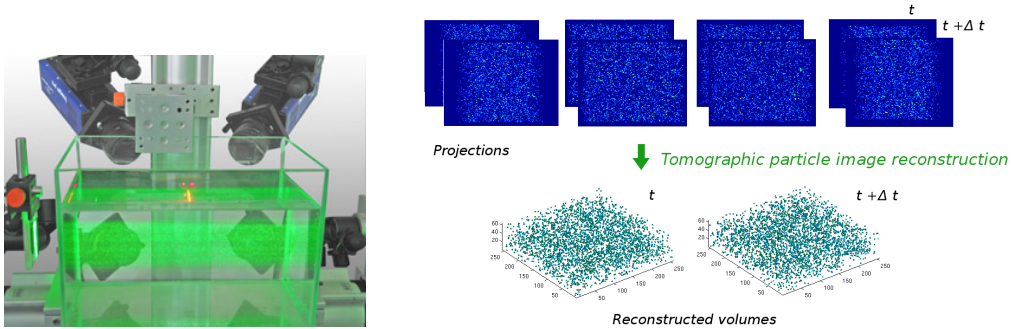
FIGURE 1. Compressive sensing in experimental fluid dynamics: A multi-camera setup gathers few projections from a sparse volume function. This scenario is described by a very large and highly underdetermined sparse linear system (1.1) having the additional properties (1.2). The sparsity parameter k reflects the seeding density of a given fluid with particles. Less sparse scenarios increase the spatial resolution of subsequent studies of turbulent motions, but compromise accuracy of the reconstruction. Research is concerned with working out and mathematically substantiating the best compromise.

We regard the measurement matrix A as *given*. Concerning the design of A, we can only resort to small random perturbations of the non-zero entries of A, thus preserving the sparse structure that encodes the underlying incidence relation of the sensor. Additionally, we exploit the fact that solution vectors x can be regarded as samples from a uniform distribution over k-sparse vectors, which represents with sufficient accuracy the underlying physical situation.

Under these assumptions, we focus on an *average case analysis* of conditions under which *unique recovery* of x can be expected with *high probability*. A corresponding tail bound implies a weak threshold effect and criterion for adequately choosing the value of the sparsity parameter k. Our results are in excellent agreement with numerical experiments and improve the state-of-the-art by a factor of three.

Contribution and Organization. In Section 2, we detail the mathematical abstraction of the imaging process and discuss directly related work. In Section 3, we examine recent results of compressive sensing based on sparse expanders. This sets the stage for an average case analysis conducted in Section 5 and corresponding weak recovery properties, that are in sharp contrast to poor strong recovery properties presented in Section 4. We conclude with a discussion of quantitative results and their agreement with numerical experiments in Section 6.

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 $x \in \mathbb{R}^n$, $\mathrm{supp}(x) \subseteq [n]$, is the set of indices of non-vanishing components of x. With $I^+(x) = \{i\colon x_i > 0\}$, $I^0(x) = \{i\colon x_i = 0\}$ and $I^-(x) = \{i\colon x_i < 0\}$, we have $\mathrm{supp}(x) = I^+(x) \cup I^-(x)$ and $\|x\|_0 = |\mathrm{supp}(x)|$.

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, 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. PRELIMINARIES

2.1. **Imaging Setup and Representation.** We refer to Figure 2 for an illustration of the mathematical abstraction of the scenario depicted by Figure 1. In order to handle in parallel the 2D and 3D cases, we will use the variable

$$D \in \{2, 3\}. \tag{2.1}$$

We measure the **problem size** in terms of $d \in \mathbb{N}$ and consider $n := d^D$ **cells** in a square ($D = 2$) or cube ($D = 3$) and $m := Dd^{D-1}$ **rays**, compare Fig. 2, left and right. 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 given by a $m \times n$ **measurement matrix** A_d^D

$$(A_d^D)_{ij} = \begin{cases} 1, & \text{if } j\text{-th ray intersects } i\text{-th cell,} \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}$$

for all $i \in [m]$, $j \in [n]$. Thus, cells and rays correspond to columns and rows of A_d^D.

The incidence relation encoded by A_d^D gives rise to the equivalent representation in terms of a **bipartite graph** $G = (C, R; E)$ with left and right vertices C and R, and edges $cr \in E$ iff $(A_d^D)_{rc} = 1$. Figure 2 illustrates that G has **constant left-degree** $\ell = D$. It will be convenient to use a separate symbol ℓ.

For a fixed vertex i, any adjacent vertex $j \sim i$ is called **neighbor** of i. For any non-negative measurement matrix A and the corresponding graph, the set

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

contains all neighbors of S. The same notation applies to neighbors of subsets $S \subset [m]$ of right nodes.

With slight abuse, we call the matrix A_d^D that encodes the adjacency $r \sim c$ of vertices $r \in R$ and $c \in C$ **adjacency matrix** of the induced bipartite graph G, deviating from the usual definition of the adjacency matrix of a graph that encodes the adjacency of *all* nodes $v_i \sim v_j$, $V = C \cup R$. Moreover, in this sense, we will call any non-negative matrix adjacency matrix, based on its non-zero entries.

Let A be 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 and choose their locations on an irregular grid, see [14].

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 $|\mathcal{N}(X)| \geq \delta\ell|X|$.

2.2. **Deviation Bound.** We will apply the following inequalities for bounding the deviation of a random variable from its expected value based on martingales, that is on sequences of random variables (X_i) defined on a finite probability space $(\Omega, \mathcal{F}, \mu)$ satisfying

$$\mathbb{E}[X_{i+1}|\mathcal{F}_i] = X_i, \qquad \text{for all} \quad i \geq 1, \tag{2.3}$$

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

This setting applies to random variables associated to measurements that are statistically *dependent* due to the intersection of projection rays (cf. Fig. 2).
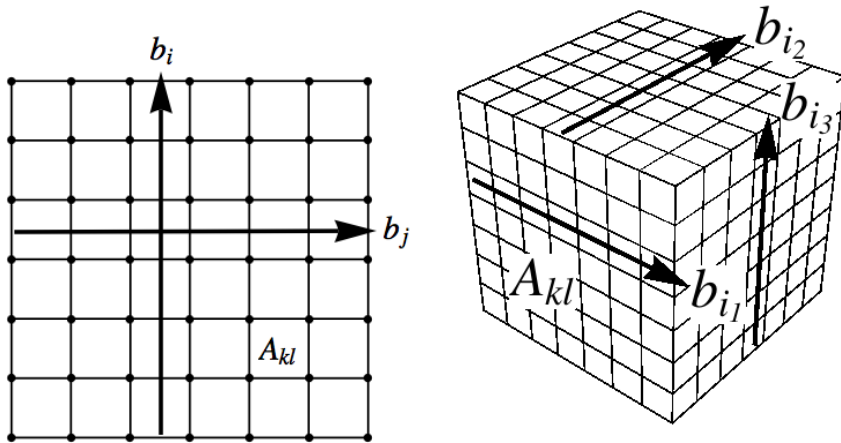
FIGURE 2. **Left:** 2D imaging geometry with d^2 cells and $2d$ projection rays (here: $d = 6$). The incidence relation is given by the measurement matrix $A = A_d^2$ (cf. Eqn. (2.2)) which is the adjacency of a bipartite graph with constant left degree $\ell = 2$. **Right:** 3D imaging geometry with d^3 cells and $3d^2$ rays (here: $d = 7$). The incidence relation given by the measurement matrix $A = A_d^3$ is the adjacency of a bipartite graph with constant left degree $\ell = 3$.

Theorem 2.1 (Azuma's Inequality [1, 6])**.** *Let* $(X_i)_{i=0,1,2,...}$ *be a sequence of random variables such that for each* i,

$$|X_i - X_{i-1}| \leq c_i. \tag{2.4}$$

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

$$\Pr\left(|X_j - X_0| \geq \delta\right) \leq 2\exp\left(-\frac{\delta^2}{2\sum_{i=1}^{j} c_i^2}\right). \tag{2.5}$$

2.3. **Related Work.** Although it was shown [3] that random measurement matrices are optimal for Compressive Sensing, in the sense that they require a minimal number of samples to recover efficiently a k-sparse vector, recent trends [2, 20] 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 extremely well, even if ℓ is small, as shown in [2]. In [12] it is shown that perturbing the elements of adjacency matrices of expander graphs with low expansion, can also improve performance. This findings complement our prior work in [14], where we observed that by slightly perturbing the entries of a tomographic projection matrix its reconstruction performance can be improved significantly.

We wish to inspect the bounds on the required sparsity that guarantee exact reconstruction of *most* sparse signals, and corresponding critical parameter values similar to weak thresholds in [10, 11]. The authors have computed sharp reconstruction thresholds for Gaussian measurements, such that for given a signal length n and numbers of measurements m, the maximal sparsity value k which guarantees perfect reconstruction can be determined precisely.

For a matrix $A \in \mathbb{R}^{m \times n}$, Donoho and Tanner define the undersampling ratio $\delta = \frac{m}{n} \in (0, 1)$ and the sparsity as a fraction of m, $k = \rho m$, for $\rho \in (0, 1)$. The so called *strong phase transition* $\rho_S(\delta)$ indicates the necessary undersampling ratio δ to recover *all* k-sparse solutions, while the *weak phase transition* $\rho_W(\delta)$ indicates when x^* with $\|x^*\|_0 \leq \rho_W(\delta) \cdot m$ can be recovered with overwhelming probability by linear programming.

Relevant for TomoPIV is the setting as $\delta \to 0$ and $n \to \infty$, that is severe undersampling, since the number of measurements is of order $O(10^4)$ and discretization of the volume can

be made accordingly fine. For Gaussian ensembles a strong asymptotic threshold $\rho_S(\delta) \approx (2e \log(1/\delta)^{-1}$ and weak asymptotic threshold $\rho_W(\delta) \approx (2 \log(1/\delta))^{-1}$ holds, see e.g. [10]. In this highly undersampled regime, the asymptotic thresholds are the same for nonnegative and unsigned signals. Exact sparse recovery of nonnegative vectors has been also studied in a series of recent papers [12, 18], while [15, 16] additionally assumes that all nonzero elements are equal to each other. As expected, additional information, improves the recoverable sparsity thresholds.

2.3.1. *Strong Recovery.* The maximal sparsity k depending on m and n, such that *all* sparse signals are *unique* and coincide with the *unique positive* solution of $Ax = b$, is investigated in [10, 11] from the perspective of convex geometry by studying the face lattice of the convex polytope $\mathrm{conv}\{A_{\bullet,1}, \ldots, A_{\bullet,n}, 0\}$. It is related to the nullspace property for nonnegative signals in what follows.

Theorem 2.2 ([10, 12, 18, 14]). *Let $A \in \mathbb{R}^{m \times n}$ be an arbitrary matrix. Then the following statements are equivalent:*

(a) *Every k-sparse nonnegative vector x^* is the unique positive solution of $Ax = Ax^*$.*
(b) *The convex polytope defined as the convex hull of the columns in A and the zero vector, i.e. $\mathrm{conv}\{A_{\bullet,1}, \ldots, A_{\bullet,n}, 0\}$ is outwardly k-neighborly.*
(c) *Every nonzero null space vector has at least $k + 1$ negative (and positive) entries.*

2.3.2. *Weak Recovery.* Thm. 2 in [10] shows the equivalence between (k, ϵ)-weakly (outwardly) neighborliness and weak recovery, i.e. uniqueness of all except a fraction ϵ of k-sparse nonnegative vectors. Weak neighborliness is the same thing as saying that $A\Delta_0^{n-1}$ has at least $(1 - \epsilon)$-times as many $(k - 1)$-faces as the simplex Δ_0^{n-1}. A different form of weak recovery is to determine the probability that a random k-sparse positive vector by probabilistic nullspace analysis. This concepts are related for an arbitrary sparse vector with exactly k nonnegative entries in the next theorem.

Theorem 2.3. *Let $A \in \mathbb{R}^{m \times n}$ be an arbitrary matrix. Then the following statements are equivalent:*

(a) *The k-sparse nonnegative vector x^* supported on S, $|S| = k$, is the unique positive solution of $Ax = Ax^*$.*
(b) *Every nonzero null space vector cannot have all its negative components in S.*
(c) *$A_S \mathbb{R}_+^k$ is a k-face of $A\mathbb{R}_+^n$, i.e. there exists a hyperplane separating the cone generated by the linearly independent columns $\{A_{\bullet,j}\}_{j \in S}$ from the cone generated by the columns of the off-support $\{A_{\bullet,j}\}_{j \in S^c}$.*

Proof. Statement (a) holds if and only if there is no $v \neq 0$ such that $Av = 0$ and $v_{S^c} \geq 0$, compare for e.g. [13, Thm. 1]. Thus (a) $\Leftrightarrow$ (b). By [11, Lem. 5.1], (a) $\Leftrightarrow$ (c) holds as well. $\square$

If, in addition, all k nonzero entries are equal to each other, then a stronger characterization holds.

Theorem 2.4 ([13, Prop. 2]). *Let $A \in \mathbb{R}^{m \times n}$ be an arbitrary matrix. Then the following statements are equivalent:*

(a) *The k-sparse binary vector $x^* \in \{0, 1\}^n$ supported on S, $|S| = k$, is the unique solution of $Ax = Ax^*$ with $x \in [0, 1]^n$.*
(b) *Every nonzero null space vector cannot have all its negative components in S and the positive ones in S^c.*
(c) *There exists a vector r such that $\mathrm{Diag}(z^*)A^\top r > 0$, with $z^* := e - 2x^*$.*
(d) *$0 \in \mathbb{R}^m$ is not contained in the convex hull of the columns of $A\mathrm{Diag}(z^*)$, i.e. $0 \notin \mathrm{conv}\{z_1^* A_{\bullet,1}, \ldots, z_n^* A_{\bullet,n}, 0\}$, with $z^* := e - 2x^*$.*

Proof. If x^* is unique in $\{0,1\}^n$, it is unique in $[0,1]^n$ as well. Uniqueness in $[0,1]^n$ holds, for e.g. by [13, Thm. 1], if there is no $v \neq 0$ such that $Av = 0$, $v_{S^c} \geq 0$ and $v_S \leq 0$, which shows equivalence to (b). With $D := \text{Diag}(e - 2x^*)$ and $DD = I$, (b) can be rewritten as follows: there is no $v \neq 0$ such that $ADDv = 0$, $Dv \geq 0$, $Dv \neq 0$. With $u := Dv$, the above condition becomes:

$$ADu = 0, u \geq 0, u \neq 0 \text{ , has no solution ,}$$

which by Gordon's theorem of alternative gives the equivalent certificate (c):

$$\exists r \quad \text{such that } DA^\top r > 0 . \tag{2.6}$$

In other words, a small k-subset of the columns of A, are "flipped" by multiplication with -1, and these modified columns together with all remaining ones can be separated from the origin, which shows equivalence to (d), i.e. 0 is not contained in the convex hull of these points.   □

Note that statement (d) is related to the necessary condition for uniqueness in [18, Thm. 1]. We further comment on Thm. (c) from a probabilistic viewpoint. Condition (c) says that all points defined by the columns of $A \, \text{Diag}(e - 2x^*)$ are located in a single half space defined by a hyperplane through the origin with normal r. Conditions under which this is likely to hold were studied by Wendel [19]. This problem is also directly related to the basic pattern recognition problem concerning the linear classification[1] of any dichotomy of a finite point set [5].

Assuming n points in $\mathbb{R}^m$ to be in general position, that is any subset of m vectors is linearly independent, and that the distribution from which the given point set is regarded as an i.i.d. sample set is symmetric with respect to the origin, then condition (2.6) holds with probability

$$\Pr(n, m) = \frac{1}{2^{n-1}} \sum_{i=0}^{m-1} \binom{n-1}{i}. \tag{2.7}$$

As Figure 3 illustrates, $\Pr(n, m) = 1$ if $n/m \leq 1$, due to the well known fact that any dichotomy of $m + 1$ points in $\mathbb{R}^m$ can be separated by a hyper-plane [17, 7]. For increasing dimension $m \to \infty$, this also holds almost surely if $n/m < 2$, which can be easily deduced by applying a binomial tail bound. Accordingly, assuming that the measurement matrix A conforms to the assumptions, the authors of [13] conclude that an existing binary solution to (1.1) is unique with probability (2.7) for *underdetermined* systems with ratio $m/n > 1/2$.

We adopt this viewpoint in Section 5.4 and develop a criterion for unique recovery with high probability using the *given* measurement matrix (2.2), based on a probabilistic average case analysis of condition (3.9) (Section 5.1). This criterion currently characterizes best the design of tomographic scenarios (Fig. 2), with recovery performance guaranteed with high probability. We conclude this section by mentioning that *exact nonasymptotic* recovery results for a k-sparse nonnegative vector are obtained in [11, Thm. 1.10] by exploiting Wendel's theorem. Donoho and Tanner show 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.

3. EXPANDERS, PERTURBATION, AND WEAK RECOVERY

This section collects recent results of recovery properties based on expanders associated with sparse measurement matrices, possibly after a random perturbation of the non-zero matrix entries. Section 3.3 applies these results to our specific setting in a form suitable for a probabilistic analysis of recovery performance presented in Section 5.

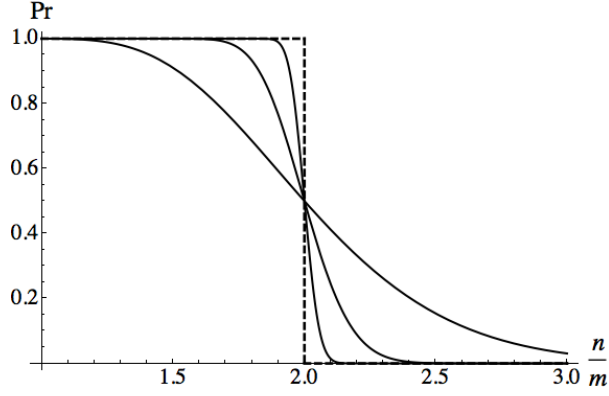[1]In this context, "linear" means affine decision functions.

FIGURE 3. The probability $\Pr(n,m)$ given by (2.7) that n points in general position in $\mathbb{R}^m$ can be linearly separated [19]. 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 $1 \leq n/m < 2$.

3.1. **Expanders and Recovery.** The following theorem is a slight variation of Theorem 4 in [18] tailored to our specific setting.

Theorem 3.1. *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.*

Proof. We will show that every nonzero null space vector has *at least* $\frac{\nu}{(1+\delta)} + 1$ negative and positive entries. Then Theorem 2.2 will provide the desired assertion.

Suppose without loss of generality that there is a vector $v \in \ker(A) \setminus \{0\}$ with

$$s := |I^-(v)| \leq \frac{\nu}{(1+\delta)} . \tag{3.1}$$

Then

$$\ell|I^-(v)| \geq |\mathcal{N}(I^-(v))| \geq \delta\ell s, \tag{3.2}$$

where the second inequality follows by assumption due to the expansion property.

Denoting by S the support of v, $S = I^-(v) \cup I^+(v)$, we have

$$\mathcal{N}(I^-(v)) = \mathcal{N}(I^+(v)) = \mathcal{N}(S) , \tag{3.3}$$

since otherwise $Av \neq 0$ because A is non-negative.

From $\ell|I^+(v)| \geq |\mathcal{N}(I^+(v))|$, (3.3) and (3.2), we obtain

$$|I^+(v)| \geq \delta s . \tag{3.4}$$

Thus,

$$|S| = |I^-(v)| + |I^+(v)| \geq 2\delta s \geq (1+\delta)s. \tag{3.5}$$

Let $\tilde{S} \subseteq S$ such that $|\tilde{S}| = \lfloor (\delta+1)s \rfloor$. Thus $|\tilde{S}| \leq \nu$ and

$$|\mathcal{N}(\tilde{S})| \geq \delta\ell|\tilde{S}| \geq \delta\ell(\delta+1)s > s\ell \tag{3.6}$$

provided $\delta(1+\delta) > 1 \Leftrightarrow \delta > (\sqrt{5}-1)/2$. Summarizing, we get $s\ell < |\mathcal{N}(\tilde{S})| \leq |\mathcal{N}(S)| = |\mathcal{N}(I^-(v))| \leq s\ell$, hence a contradiction. □

The assertion of Theorem 3.1 solely relies on the expansion property of the measurement matrix A. Theorem 3.4 below will be based on it and in turn the results of Section 5.2.

3.2. **Perturbed Expanders and Recovery.** We describe next an alternative route based on the **complete (Kruskal) rank** $r_0 = r_0(A)$ of a measurement matrix A. This is the maximal integer r_0 such that every subset of r_0 columns of A is linearly independent.

While this number is combinatorially difficult to compute in practice, both the number and the corresponding recovery performance can be enhanced by relating it to a particular expansion property of the bipartite graph associated to a *perturbed* measurement matrix $\tilde{A}$. The latter can be easily computed in practice while preserving its sparsity, i.e. the constant left-degree ℓ.

Theorem 3.2 ([14, Thm. 6.2], [12, Thm. 4.1]). *Let A be a non-negative matrix with ℓ non-zero entries in each column and complete rank $r_0 = r_0(A)$. Then $|I^-(v)| \geq r_0/\ell$ for all nullspace vectors $v \in \ker(A)$.*

Remark 3.1. In view of Theorem 2.2, (c), Theorem 3.2 says that all k-sparse non-negative vectors x can be uniquely recovered if $k \leq \lceil r_0/\ell - 1 \rceil$.

The following Lemma asserts that by a perturbation of the measurement matrix the complete rank, and hence the recovery property, may be enhanced provided all subsets of columns, up to a related cardinality, entail an expansion that is *less* however than the one required by Theorem 3.1.

Lemma 3.3 ([12, Lemma 4.2]). *Let A be a non-negative matrix with ℓ non-zero entries in each column. Suppose that for a submatrix formed by $\tilde{r}_0$ columns of A it holds that $|\mathcal{N}(X)| \geq |X|$, for each subset $X \subset C$ of columns of cardinality $|C| \leq \tilde{r}_0$, and with respect to the bipartite graph induced by A. Then there exists a perturbed matrix $\tilde{A}$ that has the same structure as A such that its complete rank satisfies $r_0(\tilde{A}) \geq \tilde{r}_0$.*

Theorem 3.5 below and Section 5.4 will be based on Theorem 3.2 and Lemma 3.3.

3.3. **Weak Reconstruction Guarantees.** We introduce some further notions used subsequently to state our results. Let A denote the matrix A_d^D defined by (2.2), and consider a subset $X \subset C$ of $|X| = k$ columns and a corresponding k-sparse vector 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

$$X \subseteq \mathcal{N}(\mathcal{N}(X)) \qquad \text{and} \qquad \mathcal{N}(\mathcal{N}(X)^c) \cap X = \emptyset. \tag{3.7}$$

Consequently, we can restrict the linear system $Ax = b$ to the subset of columns $\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c) \subset C$. This will be detailed below by Proposition 5.1.

In practical applications, the reconstruction of a random k-sparse vector x will be based on a reduced linear system with the above dimensions. These dimensions will be the same for *all* random sets $X = \text{supp}(x)$ contained in $\mathcal{N}(\mathcal{N}(X))$. Consequently, in view of a probabilistic average case analysis conducted in Section 5, it suffices to measure the expansion with respect to these sets.

Taking this into account, the following theorem tailors Theorem 3.1 to our specific setting.

Theorem 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*

$$|\mathcal{N}(X)| \geq \delta\ell|\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c)| \qquad \text{with} \qquad \delta > \frac{\sqrt{5}-1}{2}. \tag{3.8}$$

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

Likewise, the following theorem applies the statements of Section 3.2 to our specific setting.

Theorem 3.5. *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*

$$|\mathcal{N}(X)| \geq \delta\ell|\mathcal{N}(\mathcal{N}(X)) \setminus \mathcal{N}(\mathcal{N}(X)^c)| \qquad with \qquad \delta > \frac{1}{\ell}. \tag{3.9}$$

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.

The consequences of Theorems 3.4 and 3.5 are investigated in Section 5 by working out critical values of the sparsity parameter k for which the respective conditions are satisfied with high probability.

4. STRONG EQUIVALENCE

In [14] we tested the properties of the discrete tomography matrix in focus against various conditions, like the null space property, the restricted isometry property, etc., and predicted an extremely poor worst case performance of such a measurement system. In the 3D case we showed that the strong threshold on sparsity, that is the maximal sparsity level k_0 for which recovery of *all* k-sparse (positive) vectors, $k \leq k_0$, is guaranteed, is a constant, not depending on the undersampling ratio d.

4.1. **Unperturbed Systems.** Given an indexing of cells and rays, we can rewrite the projection matrix $A_d^D \in \mathbb{R}^{Dd^{D-1} \times d^D}$ from (2.2) in closed form as

$$A_d^D := \begin{cases} \begin{pmatrix} I_d \otimes \mathbb{1}_d^T \\ \mathbb{1}_d^T \otimes I_d \end{pmatrix}, & \text{if } D = 2 , \\ \begin{pmatrix} \mathbb{1}_d^\top \otimes I_d \otimes I_d \\ I_d \otimes \mathbb{1}_d^\top \otimes I_d \\ I_d \otimes I_d \otimes \mathbb{1}_d^\top \end{pmatrix}, & \text{if } D = 3 . \end{cases} \tag{4.1}$$

Since for this matrices a sparse nullspace basis can be computed, we can derive the maximal sparsity via the nullspace property, as shown next.

Proposition 4.1. [14, Prop. 2.2, Prop. 3.2] *Let $D \in \{2,3\}$, $d \in \mathbb{N}$, $d \geq 3$ and A_d^D from (4.1). Define $B_d^D \in \mathbb{R}^{d^D \times (d-1)^D}$ as*

$$B_d^D := \begin{cases} \begin{pmatrix} -\mathbb{1}_{d-1}^T \\ I_{d-1} \end{pmatrix} \otimes \begin{pmatrix} -\mathbb{1}_{d-1}^T \\ I_{d-1} \end{pmatrix}, & \text{if } D = 2, \\ \begin{pmatrix} -\mathbb{1}_{d-1}^\top \\ I_{d-1} \end{pmatrix} \otimes \begin{pmatrix} -\mathbb{1}_{d-1}^\top \\ I_{d-1} \end{pmatrix} \otimes \begin{pmatrix} -\mathbb{1}_{d-1}^\top \\ I_{d-1} \end{pmatrix}, & \text{if } D = 3 . \end{cases} \tag{4.2}$$

Then the following statements hold

(a) $A_d^D B_d^D = 0$.
(b) *Every column in B_d^D has exactly 2^D nonzero (2^{D-1} positive, 2^{D-1} negative) elements.*
(c) *B_d^D is a full rank matrix and $\mathrm{rank}(B_d^D) = (d-1)^D$.*
(d) *$\ker(A_d^D) = span\{B_d^D\}$, i.e. the columns of B_d^D provide a basis for the null space of A_d^D.*
(e) $\mathrm{rank}(A_d^D) = d^D - (d-1)^D$.
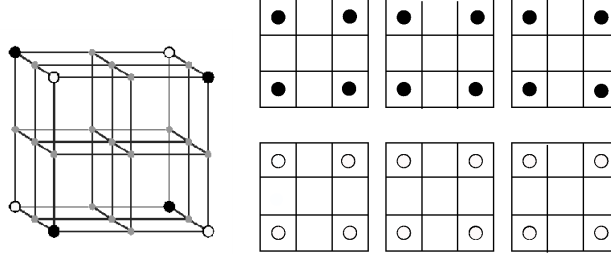(f) $\sum_{i=1}^{n} v_i = 0$ *holds for all $v \in \ker(A_d^D)$.*

FIGURE 4. Two different *non unique* 4-sparse "particle" distributions in a $3 \times 3 \times 3$ volume. Both configurations (represented by black and white dots) yield identical projections in all three directions. Such nonunique configurations correspond to positive or negative entries in an 8-sparse nullspace vector of A_d^3, compare to Prop. 4.1.

(g) *The Kruskal rang of A_d^D is $2^D - 1$, i.e.*

$$\min_{\substack{v \in \ker(A_d^D) \\ v \neq 0}} \|v\|_0 = 2^D .$$

(h) *Every nonzero nullspace vector has at least 2^{D-1} negative entries. i.e.*

$$\min_{\substack{v \in \ker(A_d^D) \\ v \neq 0}} |I^-(v)| = 2^{D-1} .$$

Thus, (g) and (h) imply

Corollary 4.2. *For all $d \in \mathbb{N}$, $d \geq 3$, every $\left(2^{D-1} - 1\right)$-sparse vector x^* is the unique sparsest solution of $A_d^D x = A_d^D x^*$. Moreover, for every $\left(2^{D-1} - 1\right)$-sparse positive vector x^* $\{x \colon A_d^D x = A_d^D x^*\}$ is a singleton.*

This bound is tight, since we can construct two 2^{D-1}-sparse solutions x^1 and x^2 such that $A_d^D x^1 = A_d^D x^2$, compare Fig. 4 for the 3D case. However, when D=3, not every 8-column combination, or more, in A_d^3 is linearly dependent. In fact, only a limited number of k-column combinations can be dependent without violating $\text{rank}(A_d^3) = 3d^2 - 3d + 1$. It turns out that this number is tiny for smaller k when compared to $\binom{n}{k}$. As k increases this number also grows and equals 1 only when $k > \text{rank}(A_d^3)$. Likewise, not every 4-sparse binary vector is nonunique. Due to the simple geometry of the problem it is not difficult to count the "bad" 4-sparse configurations in 3D. Since they are always located in 4 out of 8 corners of a cuboid in the d^3 cube, compare Fig. 4 left, and there are only two possibilities to choose them, the probability that a 4-sparse binary vector is unique, equals

$$1 - \frac{2\binom{d}{2}^3}{\binom{d^3}{4}} = 1 - \frac{6(d-1)^2}{(d^2 + d + 1)(d^3 - 2)(d^3 - 3)} = 1 - \mathcal{O}(d^{-6}) \xrightarrow{d \to \infty} 1 .$$

4.2. **Perturbed Systems.** The weak performance of A_d^D rests upon its small Kruskal rank. In order to increase the maximal number k of columns such that all k (or less) column combinations are linearly independent we perturb the nonzero entries of the original matrix A_d^D. Figure 5, right, indicates that perturbation leads to less sparse nullspace vectors. If we could estimate the Kruskal rank $\tilde{r}_0$ of the perturbed system we could apply Thm. 3.2 and obtain a lower bound on the sparsity yielding strong recovery *for all $\lceil \tilde{r}_0/\ell - 1 \rceil$-sparse vectors*. However, determining $\tilde{r}_0$ for the perturbed matrix seems impossible. We believe however that it increases with d, in contrast to the constant $2^D - 1$ in case of unperturbed systems. Luckily, it will turn out in Section 5.2 that the weak recovery threshold for unperturbed systems will give a *lower bound*
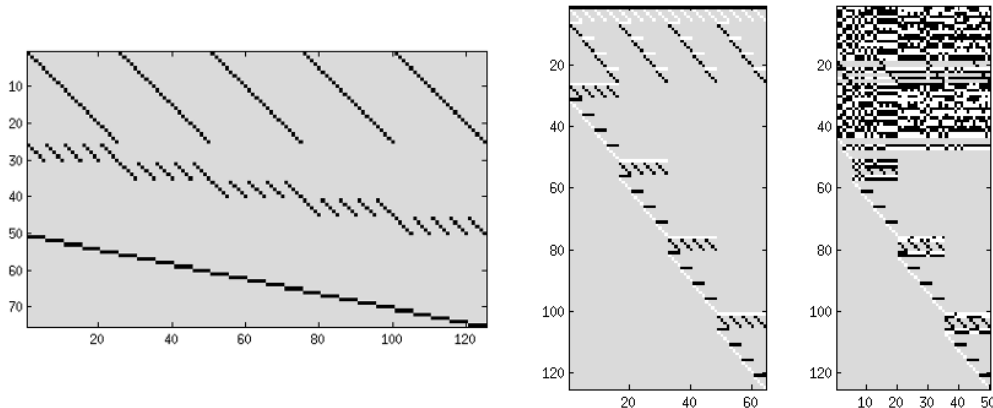
FIGURE 5. **Left:** The 3D projection matrix A_5^3 and, **middle,** a sparse basis which spans its nullspace. **Right:** If we allow a small perturbation of the nonzero entries of A_d^D, all corresponding nullspace vectors of the perturbed matrix will be less sparse and lie in a $d^{D-1}(d-D)$-dimensional subspace, compared to $(d-1)^D$ in the unperturbed case.

on the strong recovery threshold for perturbed matrices, since reduced systems will be strictly overdetermined and guaranteed to have full rank.

5. WEAK RECOVERY

In this section, we consider the recovery properties of the 3D setup depicted in Fig. 2 and establish conditions for weak recovery, that is conditions for unique recovery that holds *on average with high probability*. We clearly point out that our conditions do *not* guarantee unique recovery in *each* concrete problem instance.

Remark 5.1. 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. 5.3, thus yielding a desired threshold effect.

We first inspect in Section 5.1 the effect of sparsity on the expected dimensions of a reduced system of linear equations, along with its equivalence to the original system. Subsequently, we establish the aforementioned conditions based on Theorems 3.4 and 3.5, and on the *expected* quantities involved in the corresponding conditions.

In particular, we establish such uniqueness conditions for reduced underdetermined systems of dimension $m/n > (\sqrt{5} - 1)/2 \approx 0.618$. Our results are in excellent agreement with numerical experiments discussed in Section 6.

5.1. **Reduced System.** We formalize the system reduction described in connection with Eqn. (3.7). Besides checking its equivalence to the unreduced system, we compute the expected reduced dimensions together with a deviation bound. Additionally, we determine critical values of the sparsity parameter k that lead to overdetermined reduced systems.

Recall from Section 2.1 that we regard a given measurement matrix A also as adjacency matrix of a bipartite graph $G = (C, R; E)$.

5.1.1. *Definition and Equivalence.*

Definition 5.1. The **reduced system** corresponding to a given non-negative vector b,

$$A_{red}x = b_{red}, \qquad A_{red} \in \mathbb{R}_+^{m_{red} \times n_{red}}, \tag{5.1}$$

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

$$R_b := \text{supp}(b), \qquad C_b := \mathcal{N}(R_b) \setminus \mathcal{N}(R_b^c) \tag{5.2}$$

with

$$m_{red} := |R_b|, \qquad n_{red} := |C_b|. \tag{5.3}$$

Note that for a vector x and the bipartite graph induced by the measurement matrix A, we have the correspondence (cf. (3.7))

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

We further define

$$\mathcal{S}^+ := \{x \colon Ax = b, x \geq 0\} \tag{5.4}$$

and

$$\mathcal{S}_{red}^+ := \{x \colon A_{R_b C_b} x = b_{R_b}, x \geq 0\}. \tag{5.5}$$

The following proposition asserts that solving the reduced system (5.1) will always recover the support of the solution to the original system $Ax = b$.

Proposition 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 (5.4) and (5.5), respectively. Then*

$$\mathcal{S}^+ = \{x \in \mathbb{R}^n \colon x_{(C_b)^c} = 0 \ \text{and} \ x_{C_b} \in \mathcal{S}_{red}^+\}. \tag{5.6}$$

Proof. Let $S := \{x \in \mathbb{R}^n \colon x_{(C_b)^c} = 0 \text{ and } x_{C_b} \in \mathcal{S}_{red}^+\}$. We first show $S \subseteq \mathcal{S}^+$. Let $x \in S$. From this $x \geq 0$ follows directly. We thus just have to show $\sum_{j=1}^n a_{ij} x_j = b_i, \forall i \in [n]$. Indeed, for

$$i \in R_b: \quad \sum_{j=1}^n a_{ij} x_j = \sum_{j \in C_b} \underbrace{a_{ij} x_j}_{=b_i} + \sum_{j \in (C_b)^c} a_{ij} \underbrace{x_j}_{=0} = b_i \,,$$

whereas for

$$i \in (R_b)^c: \quad \sum_{j=1}^n a_{ij} x_j = \sum_{j \in C_b} \underbrace{a_{ij}}_{=0} x_j + \sum_{j \in (C_b)^c} \underbrace{a_{ij}}_{>a_{ij}} \underbrace{x_j}_{=0} = 0 = b_i \,.$$

Now let $x \in \mathcal{S}^+$ and consider any $i \in (R_b)^c$. Then

$$0 = b_i = \sum_{j=1}^n a_{ij} x_j = \sum_{j \in C_b} \underbrace{a_{ij}}_{=0} x_j + \sum_{j \in (C_b)^c} \underbrace{a_{ij}}_{>a_{ij}} x_j \tag{5.7}$$

holds. Since $x \geq 0$, we obtain from (5.7) that $x_j = 0, \forall j \in (C_b)^c$. To show that $A_{R_b C_b} x_{C_b} = b_{R_b}$, consider

$$i \in R_b: \quad \sum_{j \in C_b} a_{ij} x_j = \sum_{j \in C_b} a_{ij} x_j + \sum_{j \in (C_b)^c} a_{ij} \underbrace{x_j}_{=0} = \sum_{j=1}^n a_{ij} x_j = b_i \,.$$

Hence, $x_{(C_b)^c} = 0$ and $x_{C_b} \in \mathcal{S}_{red}^+$. Thus $x \in S$. $\square$

In the following two sections, we compute the expected values of the reduced system dimension (5.3).

5.1.2. *Expected Number of Non-Zero Measurements.* We consider the uniform random assignment of k particles to the $n = |C|$ cells $c \in C$. A single cell may be occupied by more than a single particle. This corresponds to the physical situation that real particles are very small relative to the discretization depicted by Figure 2. The imaging optics enlarges the appearance of particles, and the action of physical projection rays is adequately represented by linear superposition.

This scenario gives rise to a random vector $x \in \mathbb{R}^n_{k,+}$ with support $|\operatorname{supp}(x)| \le k$. It generates a vector

$$b = A_d^D x \in \mathbb{R}^m_+ \tag{5.8}$$

of measurements. We are interested in the expected size of the support of b,

$$N_R := \mathbb{E}[|\operatorname{supp}(b)|], \qquad N_R^0 := m - N_R, \tag{5.9}$$

that equals the number of projection rays $r \in R$ with non-vanishing measurements $b_r \ne 0$. We denote the event $b_r = 0$ by the binary random variable[2] $X_r = 1$, i.e. $X_r = 0$ corresponds to the event $b_r > 0$ that at least a single particle meets ray r.

The probability that a single c is met by ray r is

$$q_d := \frac{d}{|C|} = \frac{d}{n} = \frac{1}{d^{D-1}}. \tag{5.10}$$

For k particles, the probability that $0 \le i \le k$ particles meet projection ray r is

$$\Pr(b_r = i) = \binom{k}{i} q_d^i p_d^{k-i}, \qquad p_d := 1 - q_d. \tag{5.11}$$

Consequently, we have

$$\Pr[X_r = 1] = \mathbb{E}[X_r] = p_d^k, \tag{5.12a}$$

$$\Pr[X_r = 0] = \sum_{i=1}^{k} \binom{k}{i} q_d^i p_d^{k-i} = 1 - p_d^k. \tag{5.12b}$$

Lemma 5.2. *The expected number of non-zero measurements defined by* (5.9) *is*

$$N_R = N_R(k) = |R|(1 - p_d^k) = Dd^{D-1}\left(1 - \left(1 - \frac{1}{d^{D-1}}\right)^k\right),$$
$$N_R^0 = N_R^0(k) = |R| - N_R = |R|p_d^k = Dd^{D-1}\left(1 - \frac{1}{d^{D-1}}\right)^k. \tag{5.13}$$

Proof. Due to the linearity of expectation, summing over all rays gives

$$N_R = \mathbb{E}\Big[\sum_{r \in R}(1 - X_r)\Big] = |R|(1 - p_d^k).$$

$\square$

Remark 5.2. Note that N_R specifies the expected value of m_{red} in (5.3) induced by random k-sparse vectors $x \in \mathbb{R}^n_{k,+}$. See Figure 6 for an illustration.

---

[2]We economize notation here by re-using the symbol X, a random indicator vector indexed by rays (right nodes) $r \in R$. Due to the context, there should be no danger of confusion with $X = \operatorname{supp}(x)$ denoting random subsets of left nodes used in other sections.
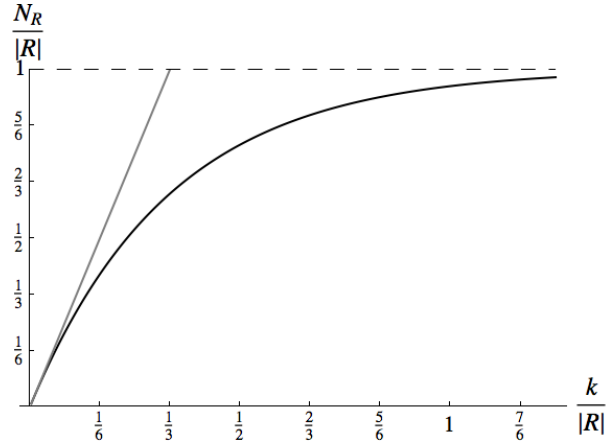
FIGURE 6. The expected number N_R of non-zero measurements (5.12). For highly sparse scenarios (small k), the expected support (5.9) of the measurement vector $|\operatorname{supp}(b)| \approx 3k$. For large values of k, this rate decreases due to the multiple incidence of cells and projection rays.

Bounding the Deviation of N_R^0. We are interested in how sharply the random number $X = \sum_{r \in R} X_r$ of zero measurements peaks around its expected value $N_R^0 = \mathbb{E}[X]$ given by (5.13). We derive next a corresponding tail bound by regarding a sequence of k randomly located cells and by bounding the difference of subsequent conditional expected values of the random variable X. Theorem 2.1 then provides a bound for the deviation $|X - \mathbb{E}[X]|$.

Let the set of rays R represent the elementary events corresponding to the observations $X_r = 1$ or $X_r = 0$ for each ray $r \in R$, i.e. ray r corresponds to a zero measurement or not.

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 $\mathcal{F}_0 = \{\emptyset, R\}$. Because 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, 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
$$Y_i = \mathbb{E}[X|\mathcal{F}_i], \quad i = 0, 1, \ldots, k, \tag{5.14}$$
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 (2.3).

Proposition 5.3. *Let $N_R^0 = \mathbb{E}[X]$ be the expected number of zero measurements for a given sparsity parameter k, given by (5.13). Then, for any $\delta > 0$,*

$$\Pr\left(|X - N_R^0| \geq \delta\right) \leq 2\exp\left(-\frac{1 - p_d^2}{(1 - p_d^{2k})}\frac{\delta^2}{2D^2}\right)$$
$$\nearrow 2\exp\left(-\frac{\delta^2}{2D^2 k}\right) \quad \text{if} \quad d \to \infty. \tag{5.15}$$

This result shows that for large problem sizes d occurring in applications, concentration of observations of N_R^0 primarily depends on the sparsity parameter k. As a consequence, the

bound enables suitable choices of $k = k(d)$ of the sparsity parameter depending on the problem size.

For example, typical values

$$k = \begin{cases} 0.05d & \text{in 2D}, \\ 0.05d^2 & \text{in 3D}, \end{cases} \tag{5.16}$$

chosen by engineers[3] in applications according to a rule of thumb, result in

$$\Pr\left(|X - N_R^0| \geq \delta\right) \leq \begin{cases} 2\exp\left(-\frac{5}{2d}\delta^2\right) & \text{in 2D}, \\ 2\exp\left(-\frac{10}{9d^2}\delta^2\right) & \text{in 3D}. \end{cases} \tag{5.17}$$

For the 3D case (5.16), the probability to observe deviations from N_R^0 larger than 1% drops below 0.01 for problem sizes $d \geq 77$, which is common in practice.

Thus, the bound (5.15) is strong enough to indicate not only that (5.16) is a particular sensible choice, but also leads to more proper choices of k for applications, which still give highly concentrated values of observations of N_R^0. This is the essential prerequisite for threshold effects of unique recovery from sparse measurements.

Proof (Proposition 5.3). 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

$$p_d^{k-(i-1)} = \mathbb{E}[X_r | \mathcal{F}_{i-1}], \tag{5.18}$$

with p_d given by (5.11). Consequently, by linearity, the expectation Y_{i-1} of zero measurements given $|R_{i-1}^0|$ zero measurements after the selection of $i - 1$ cells, is

$$Y_{i-1} = \mathbb{E}[X | \mathcal{F}_{i-1}] = \sum_{r \in R_{i-1}^0} p_d^{k-(i-1)}. \tag{5.19}$$

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

$$Y_i = \sum_{r \in R_{i-1}^0} p_d^{k-i}. \tag{5.20}$$

Furthermore,

$$\begin{aligned} Y_i - Y_{i-1} &= \sum_{r \in R_{i-1}^0} \left(p_d^{k-i} - p_d^{k-(i-1)}\right) = |R_{i-1}^0| p_d^{k-i}(1 - p_d) \\ &\leq (|R| - 1) p_d^{k-i} q_d. \end{aligned} \tag{5.21}$$

(2) Cell i is incident with $1, \ldots, D$ rays contained in R_{i-1}^0. Let R_i^0 denote the set R_{i-1}^0 after removing these rays. Then

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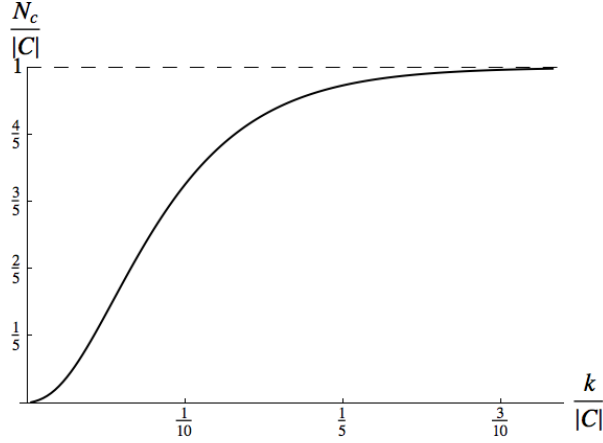
---

[3]Personal communication.

FIGURE 7. The expected number $N_C = \mathbb{E}[|C_b|]$ of cells supporting observed measurement vectors b, given by (5.23). Starting with rate $N_C \propto k$ for very small values of k, it quickly increases and exceeds N_R (Fig. 6), thus leading to underdetermined reduced systems (5.1).

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

$$Y_{i-1} - Y_i = \sum_{r \in R_{i-1}^0 \setminus R_i^0} p_d^{k-(i-1)} - \sum_{r \in R_i^0} \left(p_d^{k-i} - p_d^{k-(i-1)} \right)$$

$$\leq D p_d^{k-i+1} - \sum_{r \in R_i^0} p_d^{k-i}(1 - p_d) \leq D p_d^{k-i+1}. \tag{5.22}$$

Comparing the bounds (5.21) and (5.22), we have with $|R|q_d = D$,

$$(|R| - 1)q_d p_d^{k-i} = (D - q_d)p_d^{k-i}, \qquad D p_d p_d^{k-i} = (D - Dq_d)p_d^{k-i}.$$

Thus, we take the larger bound (5.21), drop the immaterial -1 in the first factor and compute

$$\sum_{i=1}^{k} (D p_d^{(k-i)})^2 = D^2 \frac{1 - p_d^{2k}}{1 - p_d^2}.$$

Inserting p_d from (5.11) and expanding in terms of d^{-1} at 0, we obtain

$$\frac{1 - p_d^{2k}}{1 - p_d^2} = \begin{cases} k + (k - k^2)d^{-1} + \mathcal{O}(d^{-2}), & \text{in 2D} \\ k + (k - k^2)d^{-2} + \mathcal{O}(d^{-4}), & \text{in 3D} \end{cases} \xrightarrow{d \to \infty} k.$$

Applying Theorem 2.1 completes the proof. □

5.1.3. *Expected Number of Cells.* In the previous section, we computed the expected number of measurements $N_R = \mathbb{E}[|\operatorname{supp}(b)|]$ induced by a random unknown k-sparse vector x (Lemma 5.2) along with a tail bound for $N_R^0 = |R| - N_R$ (Prop. 5.3).

In the present section, we determine the expected number of cells corresponding to N_R, denoted by N_C. We confine ourselves to the practically more relevant 3D case.

As in the previous section, $X \in \{0, 1\}^{|R|}$ denotes a random vector indicating subsets of projection rays. $X_r = 1$, $r \in R$, corresponds to a zero observation along ray r. For a subset of rays $R_b \subset R$, we say that the corresponding subset of cells C_b in (5.2) **supports** R_b.

Proposition 5.4. *For a given value of the sparsity parameter k, the expected size of subsets of cells that support random subsets $R_b \subset R$ of observed non-zero measurements, is*

$$N_C = N_C(k) = d^3 \left(1 - 3\left(1 - \frac{1}{d^2}\right)^k + 3\left(1 - \frac{2d-1}{d^3}\right)^k - \left(1 - \frac{3d-2}{d^3}\right)^k \right). \tag{5.23}$$

Proof. We partition the set of rays $R = R_1 \cup R_2 \cup R_3$ according to the three projection images (Fig. 2) and associate with the cells C the corresponding set of triples of projection rays

$$R_{1,2,3} = \big\{(r_1, r_2, r_3)\colon \cap_{i=1}^3 r_i \neq \emptyset, \; r_i \in R_i, \; i = 1, 2, 3\big\},$$

with each triple intersecting in a single cell. Thus, we have $|R_{1,2,3}| = |C| = d^3$, and each cell $c_{ijk} = r_i \cap r_j \cap r_k$ belongs to the set C_b supporting R_b if $R_b \cap (r_i \cup r_j \cup r_k) \neq \emptyset$. In terms of random variables X_r indicating zero-measurements by $X_r = 1$, this means that $c_{ijk} \in C_b$ if not $X_{r_i} = X_{r_j} = X_{r_k} = 1$. Thus,

$$N_C = \mathbb{E}\Big[\sum_{R_{1,2,3}} (1 - X_{r_1})(1 - X_{r_2})(1 - X_{r_3}) \Big]$$

$$= \sum_{R_{1,2,3}} \Big(1 - \big(\mathbb{E}[X_{r_1}] + \mathbb{E}[X_{r_2}] + \mathbb{E}[X_{r_3}]\big) + \sum_{1 \leq i < j \leq 3} \mathbb{E}[X_{r_i} X_{r_j}] - \mathbb{E}[X_{r_1} X_{r_2} X_{r_3}] \Big).$$

This expression takes into account the intersection of projection rays r_i, r_j (inclusion-exclusion principle) in order not to overcount the number of supporting cells.

We have $\mathbb{E}[X_{r_i}] = p_d^k = (1 - d^{-2})^k$ by (5.12) and (5.11). The event $X_{r_i} X_{r_j} = 1$ means that both rays correspond to zero measurements, which happens with probability

$$\Big(1 - \frac{|r_i \cup r_j|}{|C|}\Big)^k = \Big(1 - \frac{2d - 1}{d^3}\Big)^k.$$

We have three pairs of sets of rays from $R = R_1 \cup R_2 \cup R_3$, and each of the d^2 rays $r_i \in R_i$ intersects with d rays $r_j \in R_j$. Finally, three intersecting rays correspond to zero measurements with probability

$$\Big(1 - \frac{|r_1 \cup r_2 \cup r_3|}{|C|}\Big)^k = \Big(1 - \frac{3d - 2}{d^3}\Big)^k,$$

for each of the d^3 cells $c \in C$. $\qquad\square$

Remark 5.3. Note that N_C specifies the expected value of n_{red} in (5.3) induced by random k-sparse vectors $x \in \mathbb{R}^n_{k,+}$. See Figure 7 for an illustration.

5.1.4. *Overdetermined Reduced Systems: Critical Sparsity k.* For small value of k, that is for highly sparse scenarios, the expected value $N_R(k) \approx 3k$ grows faster than $N_C(k) \approx k$. Consequently, the *expected* reduced system due to Definition 5.1 will be overdetermined. This holds up to a critical value $k \leq k_{crit}$ because for increasing values of k, it is more likely that several particles are incident with some projection ray, making N_C increasing faster than N_R.

Proposition 5.5. *For $k \leq k_{crit}$, the reduced system (5.1) will be overdetermined with high probability, where k_{crit} solves*

$$N_R(k_{crit}) = N_C(k_{crit}) \tag{5.24}$$

and $N_R(k_{crit}), N_C(k_{crit})$ are given by (5.13) and (5.23).

Figure 8 shows the dependency $k_{crit} = k_{crit}(d)$ on the problem size d, as defined by (5.24) .

5.2. **Unperturbed Systems.** We consider the recovery properties of the 3D setup depicted in Fig. 2, based on Theorem 3.4 and on the *expected* quantities involved in the corresponding condition (3.8), as worked out in Section 5.1. Concerning the interpretation of the following claims, we refer to Remark 5.1.

Proposition 5.6. *The system $Ax = b$, with measurement matrix A given by (2.2), admits unique recovery of k-sparse non-negative vectors x with high probability, if*

$$k \leq \frac{N_C(k_\delta)}{1 + \delta} = \frac{1}{3\delta(1 + \delta)} N_R(k_\delta), \qquad \delta > \frac{\sqrt{5} - 1}{2}, \tag{5.25a}$$
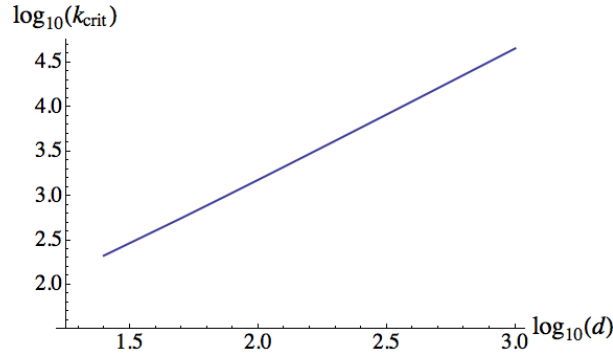
FIGURE 8. Values $k_{crit} = k_{crit}(d)$ of the sparsity parameter such that $k \le k_{crit}$ yield overdetermined reduced systems (5.1). For the depicted and practically relevant range of d, the slope of the log-log curve slightly decreases $1.65 \ldots 1.55$.

where k_δ solves

$$N_R(k_\delta) = 3\delta N_C(k_\delta) \tag{5.25b}$$

and $N_R(k)$, $N_C(k)$ are given by (5.13) and (5.23).

Proof. The assertion follows from replacing the quantities forming condition (3.8) by their expected values, due to Remarks 5.2 and 5.3. □

Remark 5.4. Equation (5.25b) shows that unique recovery of a k-sparse, $k \le \frac{n_{red}}{(1+\delta)}$, non-negative vector can be expected using the unperturbed measurement matrix provided the reduced system (5.1) is by a factor $m_{red} \ge 1.854\, n_{red}$ overdetermined. See Figure 9 for an illustration.

5.3. **Perturbed Systems.** Analogously to the previous section, we evaluate the average recovery performance using perturbed systems based on Theorem 3.5.

Proposition 5.7. *The system $\tilde{A}x = b$, with perturbed measurement matrix $\tilde{A}$ given by (2.2), admits unique recovery of k-sparse non-negative vectors x with high probability, if k satisfies condition $k \le k_{crit}$ from Prop. 5.5, that is, if the reduced system (5.1) is overdetermined.*

Proof. Immediate from Theorem 3.5, replacing the quantities forming condition (3.9) by their expected values, and taking into account $\ell = 3$ for the measurement matrix (2.2) and the case $D = 3$. □

Remark 5.5. In view of this assertion and Remark 5.4, it is remarkable that a significant gain of recovery performance can be obtained by a simple device: structure-preserving perturbation of the measurement matrix. See Figure 9 for an illustration.

5.4. **Underdetermined Perturbed Systems.** Based on (2.7) and the average case analysis of condition (3.9) (Section 5.1), we devise a criterion for determining the maximal sparsity value k (minimal sparse scenario), such that any k-sparse vector x can be uniquely recovered with high probability using the measurement matrix A given by (2.2). Unlike Propositions 5.6 and 5.7, we specifically consider here less sparse scenarios that result in *underdetermined* reduced systems (5.1).

Proposition 5.8. *Let A be a matrix satisfying the assumptions of Lemma 3.3 with $\tilde{r}_0 = N_R(k_{max})$, where k_{max} solves*

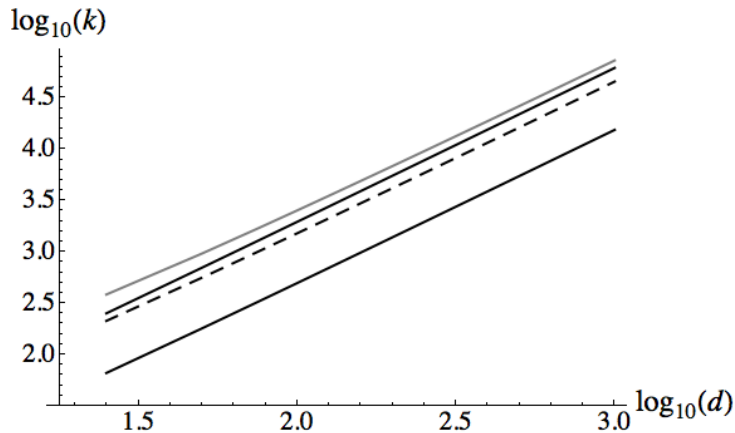$$N_R(\tilde{k}_{max}) = \delta N_C(\tilde{k}_{max}), \qquad \delta > \frac{\sqrt{5}-1}{2}, \tag{5.26}$$

FIGURE 9. Critical upper bound sparsity values $k = k(d)$ that guarantee unique recovery of k-sparse vectors x on average with high probability. From bottom to top: k_δ (5.25a) for unperturbed matrices A, k_{crit} (5.24) resulting in overdetermined reduced systems, k_{max} (5.27) for underdetermined perturbed matrices A, and fully random measurement matrices.

with $N_R(k), N_C(k)$ *given by* (5.13) *and* (5.23). *Then a k-sparse vector x can be uniquely recovered with high probability, if*

$$k \le k_{max} = \frac{N_R(\tilde{k}_{max})}{3}. \tag{5.27}$$

Proof. By assumption and Lemma 3.3, Theorem 3.2 (see also Remark 3.1) implies (5.27), thereby taking into account that Eqn. (5.26) defining k_{max} reflects the expected version of condition (3.8), subdivided by the factor 3 due to (5.27). $\square$

[19, 5, 13] Figure 9 illustrates the value k_{max} (5.27) and compares it to the previous results.

Finally, we comment on the uniqueness condition established in [13] which corresponds to the top $k(d)$ curve in Figure 9. This result does not apply to our setting. The reason is that a basic assumption underlying the application of (2.7) does *not* hold. While after some perturbation the points corresponding to the columns of $\tilde{A}$ and the sparsity value $|I^-(x)| = k$ are in general position, the underlying distribution lacks symmetry with respect to the origin. As a result, we cannot establish the superior performance of "fully" random sensors considered in [13].

5.5. **Two Cameras are Not Enough.** In the present section, we briefly discuss how the previously obtained bounds on sparsity apply in the 2D scenario. To this end, we first compute the expected value of nonempty cells connected to R_b measurements generated by a k sparse nonnegative vector.

Proposition 5.9. *In 2D, the expected size of subsets of cells that support random subsets $R_b \subset R$ of observed non-zero measurements, is*

$$N_C = N_C(k) = d^2 \left(1 - \left(1 - \frac{1}{d}\right)^k\right)^2, \tag{5.28}$$

for a given sparsity parameter k,

Proof. We partition the set of rays $R = R_1 \cup R_2$ according to the two projection images (Fig. 2), left, and associate with the cells C the corresponding set of pairs of projection rays

$$R_{1,2} = \left\{(r_1, r_2) \colon \cap_{i=1}^2 r_i \ne \emptyset, \; r_i \in R_i, \; i = 1, 2\right\},$$

with each pair intersecting in a single cell. Thus, we have $|R_{1,2}| = |C| = d^2$, and each cell $c_{ij} = r_i \cap r_j$ belongs to the set C_b supporting R_b if $R_b \cap (r_i \cup r_j) \neq \emptyset$. In terms of random variables X_r indicating zero-measurements by $X_r = 1$, this means that $c_{ij} \in C_b$ if not $X_{r_i} = X_{r_j} = 1$. Thus,

$$N_C = \mathbb{E}\Big[\sum_{R_{1,2}} (1 - X_{r_1})(1 - X_{r_2}) \Big]$$

$$= \sum_{R_{1,2}} \Big(1 - \big(\mathbb{E}[X_{r_1}] + \mathbb{E}[X_{r_2}] \big) + \sum_{1 \leq i < j \leq 2} \mathbb{E}[X_{r_i} X_{r_j}]] \Big),$$

taking the intersection of projection rays r_i, r_j into account. We obtained $\mathbb{E}[X_{r_i}] = p_d^k = (1 - \frac{1}{d})^k$ in (5.12) and (5.11). The event that both rays correspond to zero measurements $X_{r_i} X_{r_j} = 1$ happens with probability

$$\Big(1 - \frac{|r_i \cup r_j|}{|C|} \Big)^k = \Big(1 - \frac{2d-1}{d^2} \Big)^k = \Big(1 - \frac{1}{d} \Big)^{2k}.$$

$\square$

By Prop. 5.9 and Lemma 5.2 we can now compute the the expected ratio of the dimensions of the reduced system, further denoted by c. We solve the polynomial $N_R(k) = cN_C(k)$ according to and (5.28). Interesting are the values $c \in \{2\delta, 1, \delta, \frac{1}{2}\}$. For example, if $c = 2\delta$, we obtain guaranteed recovery of all 1-sparse vectors, which also equals the strong threshold for the 2D case. If $c = 1$, we obtain, on average, that any k-sparse vector x, with

$$k \leq k_{crit} = \frac{\log\big(\frac{d-2}{d} \big)}{\log\big(\frac{d-1}{d} \big)} \approx 2 \,, \tag{5.29}$$

induces reduced reduced overdetermined systems. Thus two particles can always be reconstructed, after perturbation. If $c = \frac{1}{2}$ the critical sparsity value approximately equals 4 for arbitrary d. This is the best achievable bound, which is obviously useless for application. For $k = 3$ it can be shown that the probability of correct recovery via the perturbed matrix A_d^2 is

$$1 - \frac{2 \cdot 4 \cdot \binom{d}{2}\binom{d}{3} + 4 \cdot \binom{d}{3}^2}{\binom{d^3}{3}} = \frac{d^2 + 6d - 10}{3(d^2 - 2)} \xrightarrow{d \to \infty} 1/3 \,.$$

We mention that the expected relative values of N_R and N_C do not vary much with different two camera arrangements. This highly pessimistic results can be explained by the fact that there is no expander with constant left degree ℓ less than 3.

6. NUMERICAL EXPERIMENTS AND DISCUSSION

In this section we empirically investigate bounds on the required sparsity that guarantee unique nonnegative or binary k-sparse solutions.

6.1. **Reduced Systems versus Analytical Sparsity Thresholds.** The workhorse of the previous theoretical average case performance analysis of the discrete tomography matrix from (4.1) is the derivation of 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 Fig. 10, left, for the 2D case and right, for the 3D case. To generate the figures we varied $k \in \{1, 2, \cdots, 2000\}$ and $d \in \{10, 11, \cdots, 100\}$ in 2D and $k \in \{1, 2, \cdots, 2000\}$ and $d \in \{10, 11, \cdots, 100\}$ in 3D, respectively, and generated for each point (k, d) 500 problem instances. The plots show $N_R(k, d)/N_C(k, d)$ along with the curves: k_δ (5.25a) for unperturbed
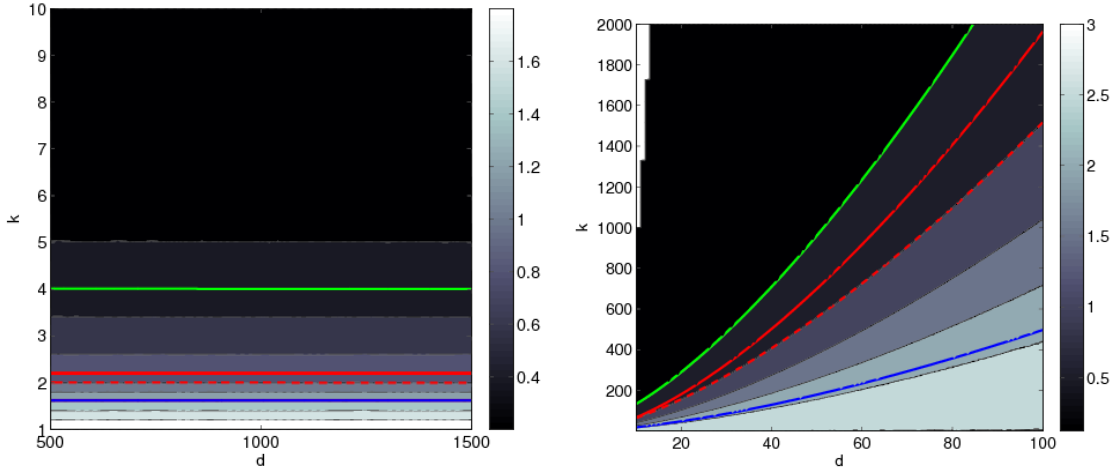
FIGURE 10. The contourplots of the average fraction of the reduced systems as a function of the resolution parameter d and the sparsity parameter k. **Left:** In agreement with the results of Section 5.5, the plots for the 2D case show that level lines of $N_R(k,d)/N_C(k,d)$ are constant with varying d. **Right:** In 3D the situation dramatically changes. Higher sparsity values are allowed for increasing values of d, as the derived threshold curves show. Below the blue curve k_δ (5.25a) reconstruction for unperturbed systems is guaranteed with high probability. Below the dashed red curve k_{crit} (5.24) reduced systems are overdetermined. For points below the solid red curve k_{max} (5.27) reconstruction is guaranteed for perturbed systems. Finally, problem instances under the green curve k_{opt} (6.1) could be recovered if the reduced matrices would follow a symmetrical distribution with respect to the origin.

matrices A, k_{crit} (5.24) resulting in overdetermined reduced systems, k_{max} (5.27) for underdetermined perturbed matrices A, and k_{opt} which solves

$$N_R(k_{opt}) = 0.5 N_C(k_{opt}) \,. \tag{6.1}$$

6.2. **Empirical Phase Transitions.**
We further concentrate on the 3D case. In analogy to [10] 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 consider $d \in \{10, 11, \ldots, 100\}$, the corresponding matrix $A_d^3 \in \mathbb{R}^{3d^2 \times d^3}$ from (4.1) and its perturbed version $\tilde{A}$ and the sparsity as a fraction of d^2, $k = \rho d^2$, for $\rho \in (0,1)$.

This phase transition $\rho(d)$ indicates the necessary relative sparsity to recover a k-sparse solution with overwhelming probability. More precisely, if $\|x\|_0 \leq \rho(d) \cdot d^2$, then with overwhelming 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. As shown in Fig. 12 the threshold for a unique nonnegative solution and a unique 0/1-bounded solution are quite close.

To generate the success and failure transition plots we generated A according to (4.1) and $\tilde{A}$ by slightly perturbing its entries and varying $d \in \{10, 11, \ldots, 100\}$ $\tilde{A}$ has the same sparsity structure as A, but random entries drawn from the standard uniform distribution on the open

interval $(0.9, 1.1)$. We have tried different perturbation levels, all leading to similar results. Thus we adopted this interval for all presented results.

Then for $\rho \in [0, 1]$ a ρd^2-sparse nonnegative or binary vector was generated to compute the right hand side measurement vector and 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 derived analytical thresholds. We refer to the figure captions for detailed explanations. Finally, we refer to the summary in Figure 11 for the computed sharp sparsity thresholds, which are in excellent agreement with our numerical experiments.
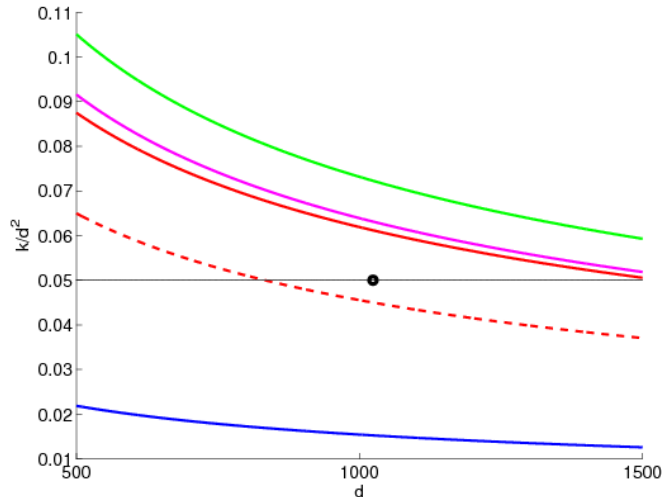

FIGURE 11. Relative critical upper bound sparsity values $k(d)$ in the practical relevant domain $d \in (500, 1500)$ that guarantee unique recovery of k-sparse vectors x on average with high probability. From bottom to top: k_δ (5.25a) for unperturbed matrices A (blue line), k_{crit} (5.24) resulting in overdetermined reduced systems (dashed red line), k_{max} (5.27) and $\tilde{k}_{max}$ (5.26) for underdetermined perturbed matrices A (solid red and pink line), and ideal random measurement matrices k_{opt} (green line). The thin black line depicts the particle density used by engineers in practice, while the black spot corresponds to the typical resolution parameter $d = 1024$. The results demonstrate that specific slight random perturbations of the TomoPIV measurement matrix considerably boost the expected reconstruction performance by at least 150%.

7. CONCLUSIONS

The main contribution of this work is the transfer of recent results on compressive sensing via expander graphs with bad expansion properties to the discrete tomography problem. In particular, we consider a sparse binary measurement matrix, which encodes the incidence relation between projection rays and image discretization cells, along with its slightly perturbed counterpart. While the expected expansion of the underlying graph does not change with perturbation, the recovery performance can be boosted significantly. We investigate the average performance in recovery of exact sparse nonnegative signals by analyzing the properties of reduced systems obtained by eliminating zero measurements and related redundant discretization cells. We compute sharp sparsity thresholds, such that the maximal sparsity can be determined precisely for both perturbed and unperturbed scenarios. Our theoretical analysis suggests that a similar procedure can be applied to different geometries.
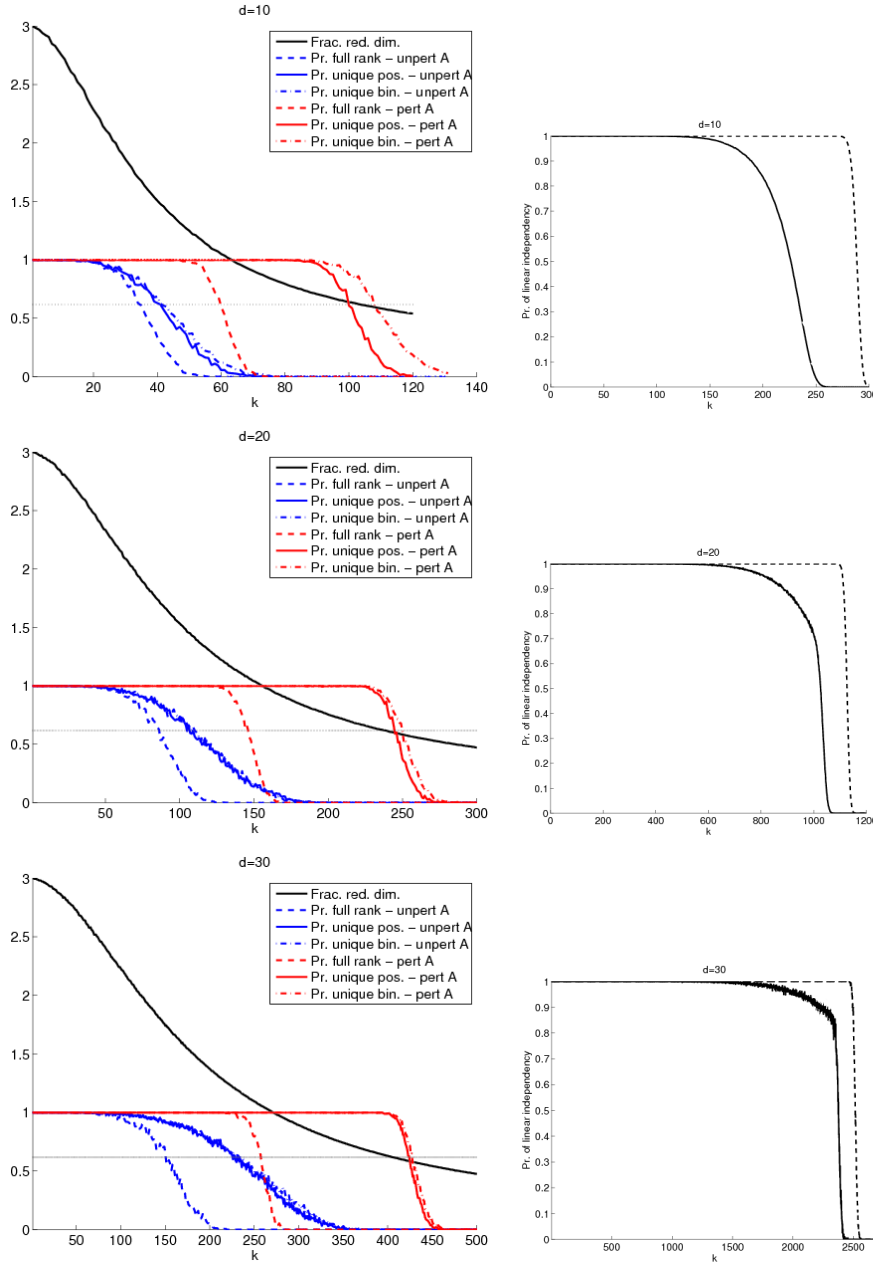
FIGURE 12. **Left:** Recovery via the unperturbed matrix A_d^3 (blue curves), $d \in \{10, 20, 30\}$ (from top to down) versus the perturbed counterpart (red curves). The dash-dot line depicts the empirical probability (500 trials) that reduced systems are overdetermined and have full rank. The solid line (blue: unperturbed, red: perturbed) shows the probability that a k-sparse nonnegative vector is unique. The dashed curve 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. The curve k_δ (5.25a) correctly predicts that 18 ($d = 10$), 48 ($d = 20$), and 85 ($d = 30$) particle are reconstructed with high probability via the unperturbed systems and 66 ($d = 10$), 181 ($d = 20$), 328 ($d = 30$) particles, via the perturbed systems according to k_{max} (5.27). However, 105 ($d = 10$), 241 ($d = 20$), 408 ($d = 30$), by $\tilde{k}_{max}$ from (5.26) are more accurate. Division by three does not seem to be necessary. **Right:** Empirical probability obtained from 10000 trials that k random columns of the unperturbed matrix (solid black line) or of the perturbed matrix (dashed black line) are linearly independent.
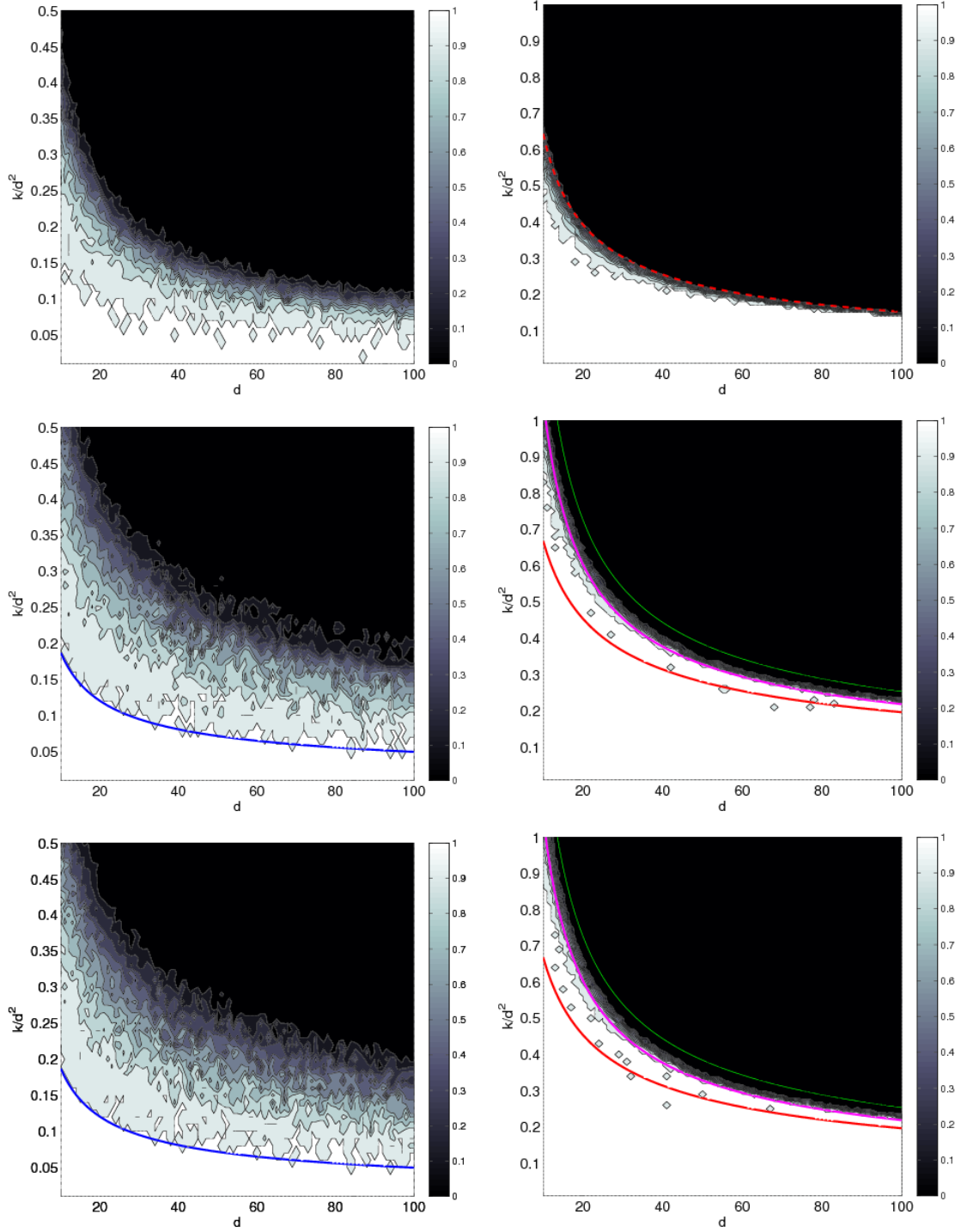
FIGURE 13. **Left:** Success and failure empirical phase transitions for unperturbed and perturbed systems **right**. **Top:** Probability that the reduced matrices are overdetermined and of full rank, along (**right**) with the estimated relative critical sparsity level k_{krit} (dashed red line) which induces overdetermined reduced matrices. **Middle:** Probability of uniqueness of a $k = \rho d^2$ sparse nonnegative vector. **Bottom:** Probability of uniqueness in $[0, 1]^n$ of a $k = \rho d^2$ sparse binary vector. The blue curve depicts again k_δ (5.25a), the dashed red curve k_{crit} (5.24), the solid red curve k_{max} (5.27), $\tilde{k}_{max}$ (5.27) and the green curve k_{opt} (6.1). In case of the perturbed matrix $\tilde{A}$ exact recovery is possible *beyond* overdetermined reduced matrices. Moreover $\tilde{k}_{max}$ follows most accurately the empirical phase transition for perturbed systems.

References

[1] K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J.*, 19(3):357–367, 1967.

[2] R. Berinde and P. Indyk. Sparse recovery using sparse random matrices, 2008. MIT-CSAIL Technical Report.

[3] E. Candès. Compressive sampling. In *Int. Congress of Math.*, volume 3, Madrid, Spain, 2006.

[4] E. Candès, M. Rudelson, T. Tao, and R. Vershynin. Error Correcting via Linear Programming. In *46th Ann. IEEE Symp. Found. Computer Science (FOCS'05)*, pages 295–308, 2005.

[5] T.M. Cover. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Trans. Electr. Comp.*, EC-14(3):326–334, 1965.

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

[7] L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition*. Springer, 1996.

[8] D. Donoho. Compressed Sensing. *IEEE Trans. Information Theory*, 52:1289–1306, 2006.

[9] D.L. Donoho. For most large underdetermined systems of equations, the minimum ℓ_1-norm near-solution approximates the sparsest near-solution. *Comm. Pure Appl. Math.*, 59(7):907–934, 2006.

[10] D.L. Donoho and J. Tanner. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc. National Academy of Sciences*, 102(27):9446–9451, 2005.

[11] D.L. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete & Computational Geometry*, 43(3):522–541, 2010.

[12] M.A. Khajehnejad, A.G. Dimakis, W. Xu, and B. Hassibi. Sparse recovery of positive signals with minimal expansion. *IEEE Trans Signal Processing*, 59:196–208, 2011.

[13] O.L. Mangasarian and B. Recht. Probability of unique integer solution to a system of linear equations. *European Journal of Operational Research*, 214(1):27–30, 2011.

[14] S. Petra and C. Schnörr. TomoPIV meets compressed sensing. *Pure Math. Appl.*, 20(1-2):49–76, 2009.

[15] M. Stojnic. ℓ_1 optimization and its various thresholds in compressed sensing. In *ICASSP*, pages 3910–3913, 2010.

[16] M. Stojnic. Recovery thresholds for ℓ_1 optimization in binary compressed sensing. In *ISIT*, pages 1593–1597, 2010.

[17] V.N. Vapnik and A.Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.

[18] M. Wang, W. Xu, and A. Tang. A unique "nonnegative" solution to an underdetermined system: From vectors to matrices. *IEEE Transactions on Signal Processing*, 59(3):1007–1016, 2011.

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

[20] W. Xu and B. Hassibi. Efficient Compressive Sensing with Deterministic Guarantees Using Expander Graphs. In *Information Theory Workshop, 2007. ITW '07. IEEE*, pages 414–419, 2007.

(Stefania Petra) IMAGE AND PATTERN ANALYSIS GROUP, UNIVERSITY OF HEIDELBERG, SPEYERER STR. 6, 69115 HEIDELBERG, GERMANY

E-mail address: petra@math.uni-heidelberg.de

(Christoph Schnörr) IMAGE AND PATTERN ANALYSIS GROUP, UNIVERSITY OF HEIDELBERG, SPEYERER STR. 6, 69115 HEIDELBERG, GERMANY

E-mail address: schnoerr@math.uni-heidelberg.de