NONUNIFORM SPARSE RECOVERY WITH FUSION FRAMES*

ULAŞ AYAZ† AND HOLGER RAUHUT†

Abstract. Fusion frames are generalizations of classical frames that provide a richer description of signal spaces where subspaces are used in the place of vectors as signal building blocks. The main idea of this work is to extend ideas from Compressed Sensing (CS) to a fusion frame setup. We use a sparsity model for fusion frames and then show that sparse signals under this model can be compressively sampled and reconstructed in ways similar to standard CS. In particular we invoke a mixed ℓ_1/ℓ_2 norm minimization in order to reconstruct sparse signals. The novelty of our research is to exploit an incoherence property of the fusion frame which allows us to reduce the number of measurements needed for sparse recovery.

Key words. Compressed Sensing, fusion frames, sparse recovery, random matrices, $\ell_{2,1}$ minimization

AMS subject classifications. 94A20, 60B20

1. Introduction.

The problem of recovering sparse signals in $\mathbb{R}^N$ from $m < N$ measurements has drawn a lot of attention in recent years [2, 3, 6]. Compressed Sensing (CS) achieves such performance by using classical signal representations and imposing a sparsity model on the signal of interest. The sparsity model, combined with randomized linear acquisition, guarantees that non-linear reconstruction can be used to efficiently and accurately recover the signal.

Fusion frames are recently emerged mathematical structures than can better capture the richness of natural and man-made signals compared to classically used representations [5]. Fusion frames generalize frame theory by using subspaces in the place of vectors as signal building blocks. For further information on motivations and applications of fusion frames, we refer to [1].

In this paper, we extend the concepts of CS to fusion frames. We demonstrate that a sparse signal in a fusion frame can be sampled using very few random projections and exactly constructed using a convex optimization program. Our sparsity model assumes that signals lie only in very few subspaces of the fusion frame. It is not required that the signals are sparse within the subspace. For the reconstruction, a mixed ℓ_1/ℓ_2 minimization is invoked.

2. Background on Fusion Frames.

A *fusion frame* for $\mathbb{R}^d$ is a collection of N subspaces $W_j \subset \mathbb{R}^d$ and associated weights v_j that satisfies

$$A\|x\|_2^2 \leq \sum_{j=1}^{N} v_j^2 \|P_j x\|_2^2 \leq B\|x\|_2^2$$

for all $x \in \mathbb{R}^d$ and for some universal fusion frame bounds $0 < A \leq B < \infty$, where $P_j \in \mathbb{R}^{d\times d}$ denotes the orthogonal projection onto the subspace W_j. For simplicity we assume that the dimensions of the W_j are equal, $\dim(W_j) = k$.

For a fusion frame $(W_j)_{j=1}^N$, let us define the Hilbert space $\mathcal{H}$ as

$$\mathcal{H} = \{(x_j)_{j=1}^N : x_j \in W_j,\ \forall j \in [N]\} \subset \mathbb{R}^{d\times N},$$

*Date produced 14.02.2013.

†Hausdorff Center for Mathematics, University of Bonn, Endenicher Allee 60, 53115 Bonn, Germany ({ulas.ayaz,rauhut}@hcm.uni-bonn.de).

where we denote $[N] = \{1, \ldots, N\}$. The *mixed $\ell_{2,1}$-norm* of a vector $\mathbf{x} \equiv (x_j)_{j=1}^N \in \mathcal{H}$ is defined as

$$\|(x_j)_{j=1}^N\|_{2,1} \equiv \sum_{j=1}^N \|x_j\|_2.$$

We note that the $\ell_{2,2}$-norm coincides with the usual ℓ_2-norm of a block vector. Furthermore, the 'ℓ_0-norm' (which is actually not even a quasi-norm) is defined as

$$\|\mathbf{x}\|_0 = \sharp\{j \in [N] : x_j \neq 0\}.$$

We call a vector $\mathbf{x} \in \mathcal{H}$ s-sparse, if $\|\mathbf{x}\|_0 \leq s$.

2.1. Sparse Recovery Problem. We take m linear combinations of an s-sparse vector $\mathbf{x}^0 = (x_j^0)_{j=1}^N \in \mathcal{H}$

$$\mathbf{y} = (y_i)_{i=1}^m = \left(\sum_{j=1}^N a_{ij} x_j^0 \right)_{i=1}^m, \quad y_i \in \mathbb{R}^d.$$

Let us denote the block matrices $\mathbf{A_I} = (a_{ij} I_d)_{i\in[m],j\in[N]}$ and $\mathbf{A_P} = (a_{ij} P_j)_{i\in[m],j\in[N]}$ that consist of the blocks $a_{ij} I_d$ and $a_{ij} P_j$ respectively. Here I_d is the identity matrix of size $d \times d$. Then we can formulate this measurement scheme as

$$\mathbf{y} = \mathbf{A_I} \mathbf{x}^0 = \mathbf{A_P} \mathbf{x}^0.$$

We can replace $\mathbf{A_I}$ by $\mathbf{A_P}$ since the relation $P_j x_j = x_j$ holds for all $j \in [N]$. We wish to recover $\mathbf{x}^0$ from those measurements. This problem reduces to the following optimization problem

$$(L0) \quad \hat{\mathbf{x}} = \mathrm{argmin}_{x \in \mathcal{H}} \|\mathbf{x}\|_0 \quad s.t. \quad \mathbf{A_P} \mathbf{x} = \mathbf{y}.$$

The optimization problem $(L0)$ is NP-hard. Instead we propose the following program

$$(L1) \quad \hat{\mathbf{x}} = \mathrm{argmin}_{x \in \mathcal{H}} \|\mathbf{x}\|_{2,1} \quad s.t. \quad \mathbf{A_P} \mathbf{x} = \mathbf{y}.$$

2.2. Relation with Previous Work. A special case of the sparse recovery problem above appears when all subspaces coincide with the ambient space $W_j = \mathbb{R}^d$ for all j. Then the problem reduces to the well studied *joint sparsity setup* [8] in which all the vectors have the same sparsity structure.

Furthermore, our problem is itself a special case of *block sparsity setup* [7], with significant additional structure that allows us to enhance existing results.

Finally in the case $d = 1$, the projections equal 1, and hence the problem reduces to the *classical recovery problem* $Ax = y$ with $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^m$.

2.3. Incoherence Parameter. We define the parameter λ as a measure of the coherence of the fusion frame subspaces as

$$\lambda = \max_{i \neq j} \|P_i P_j\|_{2\to 2}, \quad i, j \in [N].$$

Note that $\|P_i P_j\|_{2\to 2}$ equals the largest absolute value of the cosines of the principle angles between W_i and W_j. Observe that if the subspaces are all orthogonal to each other, i.e. $\lambda = 0$, then only one measurement suffices to recover $\mathbf{x}^0$. This observation suggests that fewer measurements are necessary when λ gets smaller. In this work our goal is to provide a solid theoretical understanding of this observation.

3. Nonuniform Sparse Recovery with Random Matrices. In this section we study the nonuniform recovery from fusion frame measurements. Such a result states that a fixed sparse signal can be recovered with high probability using a random draw of the measurement matrix $A \in \mathbb{R}^{m\times N}$ that has independent Rademacher entries (± 1 with equal probability) or Gaussian entries.

3.1. Notation. For matrices we use $\|\cdot\|$ to denote the spectral norm. We also denote the ℓ-th block column of the matrix $\mathbf{A_P}$ by $(\mathbf{A_P})_\ell = (a_{i\ell}P_\ell)_{i\in[m]}$ and the column submatrix restricted to $S \subset [N]$ by $(\mathbf{A_P})_S = (a_{ij}P_j)_{i\in[m],j\in S}$. Boldface notation refers to block vectors and matrices throughout this paper. For a set $S \subset [N]$ with card$(S) = s$, let $\mathbf{P}_S$ denote the $s\times s$ block diagonal matrix with entries P_i, $i \in S$, where P_i is the projection onto the subspace W_i. Given a vector $\mathbf{z} \in \mathcal{H}$, we define $\operatorname{sgn}(\mathbf{z}) \in \mathbb{R}^{dN}$ as

$$\operatorname{sgn}(\mathbf{z})_i = \begin{cases} \frac{z_i}{\|z_i\|_2} & \text{if} \quad \|z_i\|_2 \neq 0, \\ 0 & \text{if} \quad \|z_i\|_2 = 0. \end{cases}$$

3.2. Recovery Lemma. The following lemma gives a sufficient condition for recovery based on an "inexact dual vector". It is a modified version (as in [4]) of the standard method for establishing recovery via a "dual certificate". Below, $A_{|\mathcal{H}}$ denotes the restriction of a matrix A to $\mathcal{H}$.

LEMMA 3.1 (Inexact duality). *Let $A \in \mathbb{R}^{m\times N}$ and $(W_j)_{j=1}^N$ be a fusion frame for $\mathbb{R}^d$ and $x \in \mathcal{H}$ be with support $S \subset [N]$. Assume that*

$$\|[(\mathbf{A_P})_S^*(\mathbf{A_P})_S]^{-1}_{|\mathcal{H}}\| \leq 2 \quad \textit{and} \quad \max_{\ell\in[N]\setminus S} \|(\mathbf{A_P})_S^*(\mathbf{A_P})_\ell\| \leq 1. \tag{3.1}$$

Suppose there exist a block vector $\mathbf{u} \in \mathbb{R}^{Nd}$ of the form $\mathbf{u} = \mathbf{A_P}^*\mathbf{h}$ with block vector $\mathbf{h} \in \mathbb{R}^{md}$ such that

$$\|\mathbf{u}_S - \operatorname{sgn}(\mathbf{x}_S)\|_2 \leq 1/4 \quad \textit{and} \quad \max_{i\in[N]\setminus S} \|u_i\|_2 \leq 1/4. \tag{3.2}$$

Then x is the unique minimizer of $\|\mathbf{z}\|_{2,1}$ subject to $\mathbf{A_P z} = \mathbf{A_P x}$.

3.3. Main Result. We state our theoretical result pertaining to the sparse recovery from fusion frame measurements.

THEOREM 3.2. *Let $\mathbf{x} \in \mathcal{H}$ be s-sparse. Let $A \in \mathbb{R}^{m\times N}$ be Rademacher or Gaussian matrix and $(W_j)_{j=1}^N$ be given with parameter $\lambda \in [0,1]$. Assume that*

$$m \geq C(1+\lambda s)\ln^\alpha(\max\{N, sd\})\ln(\varepsilon^{-1}), \tag{3.3}$$

where $C > 0$ is a universal constant. Then with probability at least $1-\varepsilon$, (L1) recovers $\mathbf{x}$ from $\mathbf{y} = \mathbf{A_P x}$. (Here $\alpha = 1$ in the Rademacher case and $\alpha = 2$ in the Gaussian case.)

3.4. Remarks. In summary, Theorem 3.2 states that if (3.3) is satisfied, one can exactly recover an s sparse vector via $(L1)$ program from its random measurements. An additional log-factor appears in (3.3) for the Gaussian case, which we believe can be removed. This result asserts that as the subspaces become closer to being orthogonal, i.e., as λ decreases, the number of measurements decreases.

We are also able to prove stability of reconstruction with respect to noise on the measurements and under passing to approximately sparse signals.

3.5. Proof Ideas. Due to lack of space, we only present the main ideas of the proof of Theorem 3.2. We note that complete results and proofs will appear in a future paper of the authors which is in preparation. The proof relies on constructing an inexact dual vector satisfying Condition (3.2) in Lemma 3.1. To this end, we use the so-called *golfing scheme* due to Gross [9]. This method partitions the matrix $\mathbf{A_P}$ into submatrices and then define a candidate $\mathbf{u}$ in a recursive manner. We prove a series of lemmas which invoke tools from the non-asymptotic random matrix theory in order to verify that $\mathbf{u}$ satisfies the sufficient conditions.

The analysis of Condition (3.1) requires to obtain estimates on the conditioning of the submatrix $(\mathbf{A_P})_S$ associated to the support of the vector to be recovered. In particular, we will give the full proof of this result which is a main ingredient of the proof. It also shows how the parameter λ comes into play. First we introduce the rescaled matrix $\tilde{\mathbf{A}}_\mathbf{P} = \frac{1}{\sqrt{m}}\mathbf{A_P}$.

Proposition 3.3 (Conditioning of submatrix). *Let $A \in \mathbb{R}^{m\times N}$ be a measurement matrix whose entries are i.i.d. Rademacher random variables ϵ_{ij} and $(W_j)_{j=1}^N$ be a fusion frame with parameter $\lambda \in [0,1]$. Let $S \subset [N]$ with* $\mathrm{card}(S) = s$. *If, for $\delta \in (0,1)$,*

$$m \geq \frac{8}{3}\delta^{-2}(1+\lambda s)\ln(2sd/\varepsilon),$$

then the block matrix $\tilde{\mathbf{A}}_\mathbf{P}$ defined above satisfies $\|(\tilde{\mathbf{A}}_\mathbf{P})_S^*(\tilde{\mathbf{A}}_\mathbf{P})_S - \mathbf{P}_S\| \leq \delta$ with probability at least $1-\varepsilon$.

Proof. We can assume that $S = [s]$ without loss of generality. Denote $\mathbf{Y}_\ell = \frac{1}{\sqrt{m}}(\epsilon_{\ell j}P_j)_{j\in S}$ for $\ell \in [m]$ as the ℓ-th block column vector of $(\tilde{\mathbf{A}}_\mathbf{P})_S^*$. Observing that $\mathbb{E}(\mathbf{Y}_\ell\mathbf{Y}_\ell^*)_{j,k} = \mathbb{E}\frac{1}{m}(\epsilon_{\ell j}P_j\epsilon_{\ell k}P_k) = \frac{1}{m}\delta_{jk}P_jP_k$, we have $\mathbb{E}\mathbf{Y}_\ell\mathbf{Y}_\ell^* = \frac{1}{m}\mathbf{P}_S$. Then we can write

$$(\tilde{\mathbf{A}}_\mathbf{P})_S^*(\tilde{\mathbf{A}}_\mathbf{P})_S - \mathbf{P}_S = \sum_{\ell=1}^m (\mathbf{Y}_\ell{\mathbf{Y}_\ell}^* - \mathbb{E}\mathbf{Y}_\ell{\mathbf{Y}_\ell}^*).$$

This is a sum of independent self-adjoint random matrices. To this end, we will use the noncommutative Bernstein inequality due to Tropp [10, Theorem 1.4]. The block matrices $\mathbf{X}_\ell := \mathbf{Y}_\ell{\mathbf{Y}_\ell}^* - \mathbb{E}\mathbf{Y}_\ell{\mathbf{Y}_\ell}^*$ have mean zero. Moreover,

$$\begin{aligned}\|\mathbf{X}_\ell\| &= \max_{\|\mathbf{x}\|_2=1}\left|\langle \mathbf{Y}_\ell\mathbf{Y}_\ell^*\mathbf{x},\mathbf{x}\rangle - \frac{1}{m}\langle \mathbf{P}_S\mathbf{x},\mathbf{x}\rangle\right| \\ &\leq \max\left\{\max_{\|\mathbf{x}\|_2=1}\|\mathbf{Y}_\ell^*\mathbf{x}\|_2^2 - \min_{\|\mathbf{x}\|_2=1}\frac{1}{m}\|\mathbf{P}_S\mathbf{x}\|_2^2, \max_{\|\mathbf{x}\|_2=1}\frac{1}{m}\|\mathbf{P}_S\mathbf{x}\|_2^2 - \min_{\|\mathbf{x}\|_2=1}\|\mathbf{Y}_\ell^*\mathbf{x}\|_2^2\right\} \\ &\leq \max\left\{\|\mathbf{Y}_\ell^*\|^2, \frac{1}{m}\|\mathbf{P}_S\|^2\right\} = \max\left\{\|\mathbf{Y}_\ell^*\|^2, \frac{1}{m}\right\}.\end{aligned}$$

Here we used $\|\mathbf{P}_S\| = 1$. Let us bound the spectral norm of the block matrix $\mathbf{Y}_\ell^*$. We separate a vector $\mathbf{x} \in \mathbb{R}^{sd}$ into s blocks of length d and denote it as $\mathbf{x} = (x_i)_{i=1}^s$. Then

$$\begin{aligned}\|\mathbf{Y}_\ell^*\|^2 &= \frac{1}{m}\max_{\|\mathbf{x}\|_2=1}\left\|\sum_{i=1}^s \epsilon_i P_i x_i\right\|^2 = \frac{1}{m}\max_{\|\mathbf{x}\|_2=1}\sum_{i,j=1}^s \epsilon_i\epsilon_j\langle P_i x_i, P_j x_j\rangle \\ &\leq \frac{1}{m}\max_{\|\mathbf{x}\|_2=1}\sum_{i,j=1}^s |\langle P_iP_jx_j, x_i\rangle| \leq \frac{1}{m}\max_{\|\mathbf{x}\|_2=1}\sum_{i,j=1}^s \|P_iP_j\|\|x_i\|_2\|x_j\|_2\end{aligned}$$

$$\leq \frac{1}{m} \max_{\|\mathbf{x}\|_2=1} \sum_{j=1}^{s} \|x_j\|_2^2 + \sum_{i\neq j} \|P_i P_j\| \left[\frac{1}{2}(\|x_i\|_2^2 + \|x_j\|_2^2)\right]$$
$$\leq \frac{1}{m}(1+\lambda s).$$

In the third inequality above, we used that $ab \leq \frac{1}{2}(a^2+b^2)$ for $a, b \in \mathbb{R}$. This implies $\|\mathbf{X}_\ell\| \leq \frac{1}{m}(1+\lambda s)$. We denote $\mathbf{E}_{ii}(A) \in \mathbb{R}^{sd\times sd}$ as the block matrix (consisting of $s \times s$ blocks) with a single entry $A \in \mathbb{R}^{d\times d}$ at the intersection of i-th row and i-th column, and 0-matrix everywhere else. Furthermore,

$$\mathbb{E}\mathbf{X}_\ell^2 = \mathbb{E}\left(\mathbf{Y}_\ell\mathbf{Y}_\ell^*\mathbf{Y}_\ell\mathbf{Y}_\ell^* + \frac{1}{m^2}\mathbf{P}_S - \mathbf{Y}_\ell\mathbf{Y}_\ell^*\frac{1}{m}\mathbf{P}_S - \frac{1}{m}\mathbf{P}_S\mathbf{Y}_\ell\mathbf{Y}_\ell^*\right)$$
$$= \mathbb{E}\frac{1}{m}\mathbf{Y}_\ell\left(\sum_{j=1}^{s} P_j\right)\mathbf{Y}_\ell^* + \frac{1}{m^2}\mathbf{P}_S - \mathbb{E}(\mathbf{Y}_\ell\mathbf{Y}_\ell^*)\frac{1}{m}\mathbf{P}_S - \frac{1}{m}\mathbf{P}_S\mathbb{E}(\mathbf{Y}_\ell\mathbf{Y}_\ell^*)$$
$$= \frac{1}{m^2}\sum_{i=1}^{s}\mathbf{E}_{ii}\left(P_i\left(\sum_{j=1}^{s}P_j\right)P_i\right) - \frac{1}{m^2}\mathbf{P}_S.$$

In the first equality above, we used the independence of $\epsilon_{\ell j}$ for $j \in S$ and the fact that $\epsilon_{\ell j}^2 = 1$. The variance parameter appearing in the noncommutative Bernstein inequality is estimated as

$$\sigma^2 := \left\|\sum_{\ell=1}^{m}\mathbb{E}(\mathbf{X}_\ell^2)\right\| = \frac{1}{m}\left\|\sum_{i=1}^{s}\mathbf{E}_{ii}\left(P_i\left(\sum_{j=1}^{s}P_j\right)P_i\right) - \mathbf{P}_S\right\|$$
$$\leq \frac{1}{m}\max\left\{\left\|\sum_{i=1}^{s}\mathbf{E}_{ii}\left(P_i\left(\sum_{j=1}^{s}P_j\right)P_i\right)\right\|, \|\mathbf{P}_S\|\right\}$$
$$= \frac{1}{m}\max\left\{\max_{i\in[s]}\left\|P_i\left(\sum_{j=1}^{s}P_j\right)P_i\right\|, 1\right\}.$$

The first inequality above is the triangle inequality. We further estimate, for any $i \in [s]$,

$$\left\|P_i\left(\sum_{j=1}^{s}P_j\right)P_i\right\| = \left\|\sum_{j=1}^{s}P_iP_jP_i\right\| \leq \sum_{j=1}^{s}\|P_iP_j\|\|P_jP_i\| \leq 1 + \lambda^2(s-1).$$

Finally, we have

$$\sigma^2 \leq \frac{1}{m}(1+\lambda^2(s-1)).$$

Then it holds, for $\delta \in (0,1)$,

$$\mathbb{P}\left(\|(\tilde{\mathbf{A}}_\mathbf{P})_S^*(\tilde{\mathbf{A}}_\mathbf{P})_S - \mathbf{P}_S\| > \delta\right) = \mathbb{P}\left(\|\sum_{\ell=1}^{m}\mathbf{X}_\ell\| > \delta\right)$$
$$\leq 2sd\exp\left(-\frac{\delta^2 m/2}{1+\lambda^2(s-1)+(1+\lambda s)\delta/3}\right) \leq 2sd\exp\left(-\frac{3}{8}\frac{\delta^2 m}{(1+\lambda s)}\right).$$

Bounding the right hand side by ε completes the proof. □

4. A Preliminary Uniform Recovery Result. In this section, we present uniform recovery conditions for random matrices, i.e., once the random matrix is chosen, then with high probability all sparse signals can be recovered using the same matrix. One common way to study uniform recovery conditions is via the restricted isometry property. The following definition for the restricted isometry property for fusion frames was given in [1].

DEFINITION 4.1 (Fusion RIP). *Let $A \in \mathbb{R}^{m \times N}$ and $(W_j)_{j=1}^N$ be a fusion frame for $\mathbb{R}^d$. The fusion restricted isometry constant δ_s is the smallest constant such that*

$$(1 - \delta_s)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A_P x}\|_2^2 \leq (1 + \delta_s)\|\mathbf{x}\|_2^2$$

for all $\mathbf{x} \in \mathcal{H}$ of sparsity $\|\mathbf{x}\|_0 \leq s$.

Boufounos et al. [1] also show that the condition $\delta_{2s} < 1/3$ is sufficient to guarantee uniform recovery. Our analysis of the fusion RIP constant yields the following recovery result.

THEOREM 4.2. *Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with independent subgaussian entries and $(W_j)_{j=1}^N$ be given with $\dim(W_j) = k$ and parameter $\lambda \in [0, 1]$. Let $\delta \in (0, 1)$. Assume that*

$$m \geq C\delta^{-2} k \sqrt{\lambda s^2 + s} \ln^4(\max\{N, d\}) \ln(2\varepsilon^{-1}) \tag{4.1}$$

where $C > 0$ is a universal constant. Then with probability at least $1 - \varepsilon$, the fusion RIP constant δ_s of $\frac{1}{\sqrt{m}}\mathbf{A_P}$ satisfies $\delta_s \leq \delta$.

In [1], it is proved that if the underlying random matrix A satisfies the classical RIP, then uniform recovery is implied for fusion frames. This means that $m \gtrsim s \ln(N/s)$ is a sufficient condition for recovery. However this bound does not take the parameter λ into account. Presently the uniform result (4.1) suffers from additional log-terms and behave slightly worse than the nonuniform one (3.3).

Acknowledgments. The authors would like to thank the Hausdorff Center for Mathematics for support, and acknowledge funding through the WWTF project SPORTS (MA07-004) and the ERC Starting Grant StG 258926.

REFERENCES

[1] P. Boufounos, G. Kutyniok, and H. Rauhut. Sparse recovery from combined fusion frame measurements. *IEEE Trans. Inform. Theory*, 57(6):3864–3876, 2011.

[2] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.

[3] E. Candès and T. Tao. Near optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inf. Theory*, 52(12):5406–5425, 2006.

[4] E. J. Candès and Y. Plan. A probabilistic and ripless theory of compressed sensing. *CoRR*, abs/1011.3854, 2010.

[5] P. G. Casazza, G. Kutyniok, and S. Li. Fusion frames and distributed processing. *Appl. Comput. Harmon. Anal.*, 254(1):114–132, 2008.

[6] D. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[7] Y. C. Eldar and H. Bölcskei. Block-sparsity: Coherence and efficient recovery. *CoRR*, abs/0812.0329, 2008.

[8] M. Fornasier and H. Rauhut. Recovery algorithms for vector valued data with joint sparsity constraints. *SIAM J. Numer. Anal.*, 46(2):577–613, 2008.

[9] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *CoRR*, abs/0910.1879, 2009.

[10] J. A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.