THE STATISTICAL RESTRICTED ISOMETRY PROPERTY FOR GABOR SYSTEMS

Alihan Kaplan, Volker Pohl

Institute of Theoretical Inform. Technology
Technische Universität München
80333 München, Germany

Dae Gwan Lee

Mathematisch-Geographische Fakultät
Katholische Universität Eichstätt-Ingolstadt
85071 Eichstätt, Germany

ABSTRACT

Gabor matrices are important in many different areas of time-frequency analysis like radar or communications. For applications with sparse data, the question arises whether these matrices satisfy some recovery guarantees for compressive sampling, and which generating windows yield a matrix with restricted isometric property. This paper proves the uniqueness-guaranteed statistical restricted isometric property for Gabor matrices generated by an Alltop window. In that way, we present a recovery guarantee for this deterministic measurement matrix where the number of measurements scales linearly with the sparsity of the signals.

Index Terms— deterministic compressive sampling, Gabor matrices, statistical RIP

1. INTRODUCTION

The idea of compressive sampling (CS) is to reconstruct k-sparse vectors $f \in \mathbb{C}^N$ from only a few number $m << N$ of measurements. If one focuses on linear measurements, the measurement data is a vector $g = Af \in \mathbb{C}^m$ wherein $A \in \mathbb{C}^{m\times N} = [a_1, \dots, a_N]$ is the *sensing matrix* with m rows and N columns.

A major challenge in CS is to find good sensing matrices A. The *coherence* $\mu(A)$ is a simple (easy to compute) tool for measuring the quality of measurement matrices A. However, due to the *Welch bound*, the best recovery guarantee one can possibly get from coherence methods is $m \geq c\,k^2$, which is known as the *quadratic bottleneck*. This obviously limits the performance analysis of recovery algorithms to rather small sparsity levels [5]. For this reason, the concept of the *restricted isometric property (RIP)* for measurement matrices was developed [1]. Based on RIP and for probabilistic constructions of the measurement matrix [2–4], recovery guarantees (with high probability) were proven for $m \geq c\,k\,\log(...)$ measurements. This is essentially the backbone of compressed sensing [5,6].

However, random constructions have some major drawback in terms of reconstruction complexity and efficiency (see, e.g., [7]). Moreover, in applications, the measurement matrix is often not random but has a certain structure, predefined by the measurement setup. For these reasons, there are several attempts to find deterministic constructions for sensing matrices [7, 8] satisfying RIP. However, it is generally difficult to verify RIP for a given (deterministic) matrix. Therefore [7] introduced the notion of *statistical RIP (StRIP)* to investigate deterministic sensing matrices systematically. In this framework, a deterministic matrix is a good sensing matrix if a sparse vector f can be recovered with high probability (with respect to random vectors f).

This paper considers the construction of Gabor matrices which have StRIP. Gabor matrices G_ϕ are obtained by translations and modulations of a fixed generating vector ϕ. These matrices are of fundamental importance in time-frequency analysis and they appear in many applications such as radar and communications [9–13]. Matrices G_ϕ generated by a random vector ϕ where investigated in [14, 15] and it was shown that such matrices satisfy with high probability (with respect to choice of ϕ) a k-th order RIP if the number of measurements is of the order $m \geq c\,k\,(\log k)^2(\log m)^2$. Moreover, it is known that there exist vectors ϕ such that the coherence of G_ϕ is close to the optimal Welch bound [5]. Nevertheless, to the best of our knowledge, there are no results verifying StRIP or RIP for deterministic Gabor matrices. So up to now, we do not know any recovery guarantee where the number of measurements m scales linearly with the sparsity level k. This paper will show that Gabor matrices G_α generated by a so-called Alltop window α have StRIP if the number of measurements m satisfy $m \geq c\,k\,\log(N)$ with a certain constant $c > 0$. This recovery guarantee is illustrated by some numerical simulations at the end of the paper.

2. NOTATIONS AND STATISTICAL RIP

General notations We write $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ for the m-element cyclic group and we identify the Hilbert space $\mathcal{H}(\mathbb{Z}_m)$ of all functions over $\mathbb{Z}_m$ with the m-dimensional Euclidean vector space $\mathbb{C}^m$ with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. Throughout this paper, Σ_k^N stand for the set of all k-sparse vectors in $\mathbb{C}^N$. Given a matrix $\Phi \in \mathbb{C}^{m\times N}$, its column vectors are $\phi_1, \phi_2, \dots, \phi_N \in \mathbb{C}^m$ and $\phi_n(x)$ stands for the x-th entry of ϕ_n. If the columns of Φ are normalized as $\|\phi_n\| = 1$ for all $n = 1, \dots, N$, then $\mu(\Phi) = \max_{i\neq j} |\langle\phi_i, \phi_j\rangle|$ is the *coherence* of Φ. If S is a subset of $\{1, 2 \dots, N\}$, then Φ_S is the matrix containing the columns of Φ indexed by S. Finally, $\delta_{x,y}$ is the usual *Kronecker-delta* with $\delta_{x,y} = 1$ if $x = y$ and $\delta_{x,y} = 0$ if $x \neq y$.

StRIP The *statistical restricted isometry property (StRIP)* was introduced in [7]. The main idea is to analyze a statistical version of the RIP for deterministic sensing matrices. Since the matrices are deterministic the probability enters in the signal model. For the sake of completeness we give a short overview of StRIP [7].

Definition 1: *A matrix $A \in \mathbb{C}^{m\times N}$ with ℓ_2-normalized columns is said to have (k, δ, ϵ)-StRIP if*

$$(1-\delta)\|f\|_2^2 \leq \|Af\|_2^2 \leq (1+\delta)\|f\|_2^2$$

holds with probability exceeding $1-\epsilon$ for $f \in \Sigma_k^N$ having the uniform distribution over the set $S^{N-1} = \{f \in \Sigma_k^N \,:\, \|f\|_2 = 1\}$.

This work was supported by the German Research Foundation (DFG) within the priority program *Compressed Sensing in Information Processing (CoSIP)* under Grants PO 1347/3-1 and PF 450/9-1.

*Moreover, we say that A has (k,δ,ϵ)-Uniqueness-guaranteed StRIP (abbr. (k,δ,ϵ)-*UStRIP*) if additionally*

$$\left\{h \in \Sigma_k^N \,:\, Ah = Af\right\} = \{f\}$$

is satisfied with probability exceeding $1-\epsilon$.

By this definition, unique recovery of signal $x \in \Sigma_k^N$ is guaranteed (with high probability) for UStRIP-matrices but not in general for StRIP-matrices. The following conditions were introduced in [7] to characterize matrices which have StRIP or even UStRIP.

Definition 2: *Let $\Phi = [\phi_1,\ldots,\phi_N] \in \mathbb{C}^{m\times N}$ be a matrix with all entries having absolute value 1. Then the matrix $A := \frac{1}{\sqrt{m}}\Phi$ is said to be η-*StRIP-able* if Φ satisfies the following conditions*

- ***(St1)*** *The rows of Φ are mutually orthogonal, and the sum of all entries in each row is zero, i.e.,*

$$\sum_{j=1}^N \phi_j(x)\overline{\phi_j(y)} = 0 \qquad \text{if } x \neq y,$$
$$\sum_{j=1}^N \phi_j(x) = 0 \qquad \text{for all } x.$$

- ***(St2)*** *The columns of Φ form a group under 'pointwise multiplication', i.e., for all $j,j' \in \{1,\ldots,N\}$ there exists $j'' \in 1,\ldots,N$ such that $\phi_j(x)\phi_{j'}(x) = \phi_{j''}(x)$ for all x. In particular, there is an identity column consisting of ones. Without loss of generality, we assume that $\phi_1(x) = 1$ for all x.*
- ***(St3)*** *There exists $\eta > 0$ such that for all $j \in \{2,\ldots,N\}$,*

$$\left|\sum_{x=0}^{m-1} \phi_j(x)\right|^2 \leq m^{2-\eta}.$$

The importance of (ST1)–(ST3) stems from the fact that these three properties are easy to verify, whereas RIP is usually hard to check. Several classes of matrices which are StRIP-able were given in [7]. The next theorem shows that under some mild conditions on η and the number of rows m, any η-StRIP-able matrix has StRIP or even UStRIP.

Theorem 1 ([7]): *Let $A = \frac{1}{\sqrt{m}}\Phi \in \mathbb{C}^{m\times N}$ be an η-StRIP-able matrix with $\eta > 1/2$, and assume that $k < 1 + (N-1)\delta$. Then*

- *(a) A has (k,δ,ϵ)-*StRIP* with $\epsilon = 2\exp\left(-\left(\delta - \frac{k-1}{N-1}\right)^2 \frac{m^\eta}{8k}\right)$.*
- *(b) if additionally $m \geq c\,k\,\delta^{-2}\log N$ for some $c > 0$, then A has $(k,\delta,2\epsilon)$-*UStRIP* with $\epsilon = 2\exp\left(-\left(\delta - \frac{k-1}{N-1}\right)^2 \frac{m^\eta}{8k}\right)$.*

It is important to notice that an η-StRIP-able matrix has UStRIP if the number of measurements satisfy $m \geq c\,k\,\log N$. Thus the necessary number of measurements m grows linearly with the sparsity k. This is basically the same behavior as for random matrices [1, 5].

3. STRIP OF GABOR SYSTEMS WITH ALLTOP WINDOW

Gabor systems generated by an Alltop window On the Hilbert space $\mathcal{H}(\mathbb{Z}_m) = \mathbb{C}^m$, the *translation operator* T and *modulation operator* M are the unitary operators defined by

$$\big(\mathrm{T}f\big)(x) = f(x-1) \qquad \text{and} \qquad \big(\mathrm{M}f\big)(x) = f(x)\,\mathrm{e}^{\mathrm{i}\frac{2\pi}{m}x}\,,$$

respectively. Moreover, for any $\lambda = (\tau,\nu) \in \mathbb{Z}_m \times \mathbb{Z}_m$, the *time-frequency shift* operator $\pi_\lambda : \mathbb{C}^m \to \mathbb{C}^m$ is defined by

$$\big(\pi_\lambda f\big)(x) = \big(\mathrm{M}^\nu \mathrm{T}^\tau f\big)(x) = f(x-\tau)\,\mathrm{e}^{\mathrm{i}\frac{2\pi}{m}\nu x}, \quad x \in \mathbb{Z}_m.$$

For any subset $\Lambda \subseteq \mathbb{Z}_m \times \mathbb{Z}_m$ and $\phi \in \mathbb{C}^m$, the collection $\{\pi_\lambda\phi\}_{\lambda\in\Lambda}$ is said to be a *Gabor system* generated by the window ϕ, and we write $G_\phi \in \mathbb{C}^{m\times m^2}$ for the *Gabor matrix* whose columns form the *full Gabor system* $\{\pi_\lambda\phi\}_{\lambda\in\mathbb{Z}_N\times\mathbb{Z}_N}$.

Here, we consider Gabor systems G_α generated by the *Alltop window* $\alpha \in \mathbb{C}^m$, where $m \geq 5$ is a prime, which is defined [16] by

$$\alpha(x) = \tfrac{1}{\sqrt{m}}\exp\left(\mathrm{i}\tfrac{2\pi}{m}x^3\right)\,, \qquad x \in \mathbb{Z}_m\,. \tag{1}$$

Note that the columns of the corresponding Gabor matrix G_α all have unit length, i.e. $\|\pi_\lambda\alpha\|_2 = 1$ for all $\lambda \in \mathbb{Z}_m \times \mathbb{Z}_m$. Moreover, it is known that the coherence of G_α is $\mu(G_\alpha) = 1/\sqrt{m}$, which is close to the optimal lower (Welch) bound $1/\sqrt{m+1}$ [17].

Properties of Gabor matrices We want to prove that the Gabor matrix G_α has StRIP. Nevertheless, it is not hard to see that a Gabor matrix G_ϕ *cannot* be StRIP-able for any window ϕ. Therefore, the theory developed in [7] (cf. Sec. 2) cannot be applied directly but has to be adapted slightly. The next lemma shows that G_α satisfies some properties which are very similar to (St1)-(St3). These properties are needed later to prove StRIP. The simple proof of this lemma is omitted, because of space constraints.

Lemma 2: *Let $m \geq 5$ be a prime of the form $m = 3n+2$, $n \in \mathbb{N}$, let $\alpha \in \mathbb{C}^m$ be the Alltop window (1) and let G_α be the associated Gabor matrix. With α we associate the $m\times m$ diagonal matrix*

$$S_\alpha = m\,\mathrm{diag}\big(\overline{\alpha(0)}, \overline{\alpha(1)}, \ldots, \overline{\alpha(m-1)}\big)\,.$$

Then $\Phi := S_\alpha G_\alpha \in \mathbb{C}^{m\times m^2}$ has the following properties.

- ***(P1)*** *For any $x,y \in \mathbb{Z}_m$,*

$$\sum_{\lambda\in\mathbb{Z}_m\times\mathbb{Z}_m} \phi_\lambda(x)\,\overline{\phi_\lambda(y)} = m^2\delta_{x,y}.$$

- ***(P2)*** *For any $\lambda_1,\ \lambda_2 \in \mathbb{Z}_m\times\mathbb{Z}_m$, there exist unique $\gamma \in \mathbb{Z}_m$ and $\lambda_3 \in \mathbb{Z}_m\times\mathbb{Z}_m$ such that*

$$\phi_{\lambda_1}(x)\,\overline{\phi_{\lambda_2}(x)} = \omega^\gamma\phi_{\lambda_3}(x) \quad \text{for all } \ x \in \mathbb{Z}_m,$$

where $\omega = \mathrm{e}^{\mathrm{i}2\pi/m}$. For later use, we introduce the notation $\gamma = \gamma[\lambda_1,\lambda_2]$ and $\lambda_3 = \sigma[\lambda_1,\lambda_2]$.

- ***(P3)*** *For any $\lambda \in (\mathbb{Z}_m\times\mathbb{Z}_m)\backslash(0,0)$, we have*

$$\left|\sum_{x=0}^{m-1}\phi_\lambda(x)\right| \leq m\,\mu\Big(\tfrac{1}{\sqrt{m}}\Phi\Big) = \sqrt{m}.$$

- ***(P4)*** *For all $x \in \mathbb{Z}_m$, we have*

$$\sum_{\lambda_1,\lambda_2\in\mathbb{Z}_m\times\mathbb{Z}_m} \phi_{\lambda_1}(x)\,\overline{\phi_{\lambda_2}(x)} = 0, \tag{2}$$

Remark 1: Note that (P3) shows that Φ satisfies (St3) with $\eta = 1$. Moreover, (2) holds for $x \neq 0$ without any restriction on $m \geq 2$. However, it does not necessarily hold for $x = 0$ if m is not of the form $m = 3n+2$, $n \in \mathbb{N}$.

Statistical RIP After these preparations, we are ready to show that G_α, generated by the Alltop window (1) has StRIP.

Theorem 3: *Let $m \geq 5$ be a prime of the form $m = 3n+2$, $n \in \mathbb{N}$, and let G_α be the Gabor matrix generated by the Alltop window $\alpha \in \mathbb{C}^m$. If $k < 1 + (m^2-1)\delta$ then G_α has (k,δ,ϵ)-StRIP, i.e.*

$$(1-\delta)\,\|f\|_2^2 \leq \|G_\alpha f\|_2^2 \leq (1+\delta)\,\|f\|_2^2 \tag{3}$$

holds with probability exceeding $1-\epsilon$ for all random $f \in \Sigma_k^{m^2}$ uniform distributed over the set $S^{N-1} = \{h \in \Sigma_k^{m^2} : \|h\|_2 = 1\}$, and where

$$\epsilon = 2\exp\left(-\Big(\delta - \frac{k-1}{m^2-1}\Big)^2\frac{m}{8k}\right)\,. \tag{4}$$

Remark 2: It is known that there exists infinity many primes of the form $m = 3n+2$ (the so-called Eisenstein primes).

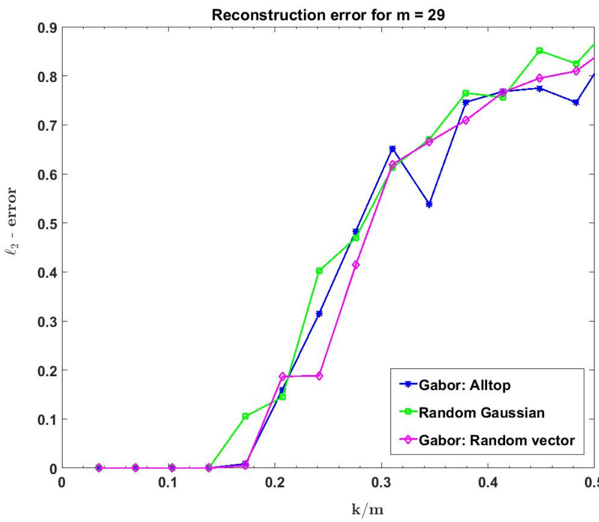

Fig. 1. Simulation results for $m = 29$ measurements.

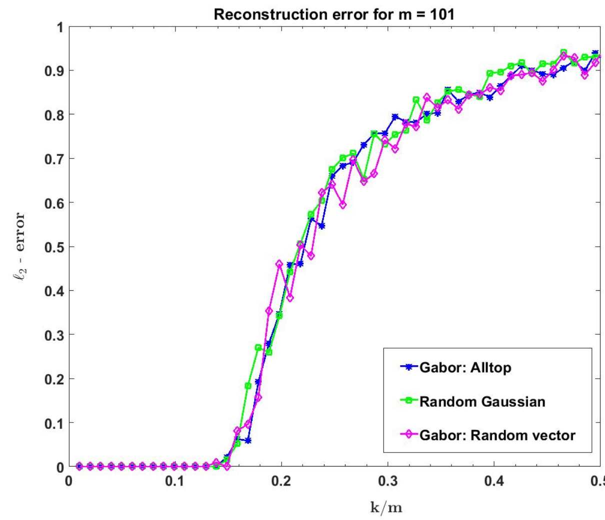

Fig. 2. Simulation results for $m = 101$ measurements.

4. USTRIP OF GABOR SYSTEMS

Theorem 3 shows that a Gabor matrix G_α generated by the Alltop window α has StRIP. However, as mentioned earlier, StRIP does not imply unique recovery of the signals, even not with high probability. In fact, while it is known that a Gabor matrices G_ϕ has full spark for almost every $\phi \in \mathbb{C}^m$ [18, 19], the matrix G_α does not have this property because its $m \times m$ submatrix $[\mathrm{T}^0\alpha, \mathrm{T}^1\alpha, \ldots, \mathrm{T}^{m-1}\alpha]$ is singular, in fact there exists a vector $f \in \mathbb{C}^{m^2}\backslash\{0\}$ with $G_\alpha f = f_{(0,0)}\mathrm{T}^0\alpha + f_{(1,0)}\mathrm{T}^1\alpha + \ldots + f_{(m-1,0)}\mathrm{T}^{m-1}\alpha = 0$. Nevertheless, UStRIP requires the uniqueness only up to high probability with respect to the uniform distribution of f over all k-sparse signals in S^{N-1}. Therefore, even though G_α does not have full spark, it may still have UStRIP. Indeed, this is the case as shown next.

Theorem 4: *Let $m \geq 5$ be a prime of the form $m = 3n+2, n \in \mathbb{N}$, and let G_α be the Gabor matrix generated by the Alltop window $\alpha \in \mathbb{C}^m$. If $k < 1 + (m^2 - 1)\,\delta$ and $m \geq c\,k\,\delta^{-2}\log m$ for some constant $c > 0$, then G_α has $(k, \delta, 2\epsilon)$-UStRIP with ϵ given in* (4).

Remark 3: We would like to stress again the difference between some known results on RIP of Gabor matrices generated by random windows [14, 15]. The main difference between those results and ours is that we consider deterministic matrices and random signals, whereas they consider randomness in both the matrices and signals.

5. NUMERICAL EXPERIMENTS & DISCUSSION

This section presents some numerical experiments which should illustrate our findings from the previous sections. Fig. 1 and 2 show simulation results where we recovered k-sparse vectors f from measurements $g = Af$ using basis pursuit, i.e.

$$\widehat{f} = \arg\min_{f \in \mathbb{C}^{m^2}} \|f\|_1 \quad \text{subject to } Af = g\ .$$

As measurement matrix A, we compared: 1. A Gabor matrix G_α generated by the Alltop window (1). 2. A random Gaussian matrix whose entries are independent, identically distributed normal random variables. 3. A Gabor matrix G_ϕ with a random Gaussian vector ϕ. All matrices are of size $m \times m^2$. In our experiments, we varied the sparsity k of the data vector f and evaluated the normalized quadratic reconstruction error $\|f - \widehat{f}\|^2 / \|f\|^2$, shown on the vertical axis. For each point, shown in the diagrams, the recovery results where averaged over 10 experiments with random data vectors f.

We observe in our simulations that the deterministic Gabor matrix with Alltop window (dimension $m = 29, 101$ were chosen) performs similarly good as random Gaussian matrices and Gabor matrices with random window. The last two classes of random matrices are known to have RIP with high probability for $m \geq c\,k\,\log(...)$. Clearly, our simulations show that the deterministic Gabor matrices G_α (generated with the Alltop window) have the same behavior as those random matrices. In this way, the linear-scale recovery guarantee for G_α proved in Theorem 4 is supported numerically.

6. APPENDIX – PROOF SKETCHES

We shortly sketch the proof of our main results. In general, these proofs follow the same steps as in [7] but working with the properties (P1)–(P4) of Lemma 2 instead of the (ST1)–(ST3) from Def. 2. It is assumed that random vectors $f \in \Sigma_k^{m^2}$ with $\|f\|_2 = 1$ are generated as follows. First, the support set of f is chossen as $S = \{\pi_1, \ldots, \pi_k\}$, where $\{\pi_j\}_{j=1}^{m^2}$ is a random permutation of $\{1, \ldots, m^2\}$. Then one chooses the k entry values of f randomly so that $\|f\|_2 = 1$ (with no distribution specified).

Sketch of proof (Theorem 3): We establish (3) for the matrix $\Phi = S_\alpha\, G_\alpha$ given in Lemma 2, and then use that $\|\Phi f\|_2^2 = m\,\|G_\alpha f\|_2^2$ for all $f \in \mathbb{C}^m$ to obtain (3).

1) Let $f \in \Sigma_k^{m^2}$ be a random vector supported on $S = \{\pi_1, \ldots, \pi_k\}$ with k nonzero entries denoted as $f_1, \ldots, f_k$. Set $g = \frac{1}{\sqrt{m}}\Phi f$. Then

$$\left(1 - \tfrac{k-1}{m^2-1}\right)\|f\|_2^2 \leq \mathbb{E}_\pi\left[\|g\|_2^2\right] \leq \left(1 + \tfrac{1}{m^2-1}\right)\|f\|_2^2\ , \qquad (5)$$

wherein $\mathbb{E}_\pi$ stands for the expectation over all possible choices π. To see this, one writes the k-sparse vector f as $f = \sum_{j=1}^k f_j e_{\pi_j}$. This yields $g = \frac{1}{\sqrt{m}}\sum_{j=1}^k f_j\,\phi_{\pi_j}$ and

$$\|g\|_2^2 = \sum_{x=0}^{m-1} |g(x)|^2 = \|f\|_2^2 + \tfrac{1}{m}\sum_{i \neq j} f_i\,\overline{f_j}\,\langle \phi_{\pi_i}, \phi_{\pi_j}\rangle \quad (6)$$

using that $\|\phi_k\|_2 = 1$ for all k. By Property (P3), we have for $i \neq j$

$$\mathbb{E}_\pi \left[\langle \phi_{\pi_i}, \phi_{\phi_j} \rangle\right] = \frac{1}{m^2(m^2-1)} \sum_{\substack{\lambda_1,\, \lambda_2 \in \mathbb{Z}_m \times \mathbb{Z}_m \\ \lambda_1 \neq \lambda_2}} \langle \phi_{\lambda_1}, \phi_{\lambda_2} \rangle$$
$$= \frac{1}{m^2(m^2-1)} \sum_{x=0}^{m-1} (-m^2) = -\tfrac{m}{m^2-1}$$

and therefore

$$\mathbb{E}_\pi \left[\|g\|_2^2\right] = \|f\|_2^2 - \tfrac{1}{m^2-1} \sum_{i \neq j} f_i \overline{f_j}. \tag{7}$$

Then (5) follows by applying the Cauchy-Schwarz inequality.

2) Next, we are going to show that the random variable $\|g\|^2$ is concentrated around its mean. The main idea is to upper bound $\Pr_\pi \left[\left|\|g\|_2^2 - \|f\|_2^2\right| \geq \delta \|f\|_2^2\right]$, and we already know from (5) that $\mathbb{E}_\pi \left[\|g\|_2^2\right]$ is close to $\|f\|_2^2$. Recalling the dependence of g on $\pi_1, \ldots, \pi_k$ as shown in (6), we write $h(\pi_1, \ldots, \pi_k) = \|g\|_2^2$. For $\beta > 0$ we consider $\Pr_\pi \left(\left|\|g\|_2^2 - \mathbb{E}_\pi \left[\|g\|_2^2\right]\right| \geq \beta \|f\|_2^2\right) = \Pr_\pi \left(|h - \mathbb{E}_\pi [h]| \geq \beta \|f\|_2^2\right)$. Together with Property (P2) and (P3) of Φ, one obtains, after some manipulations, the estimate

$$\left|h(\pi_1, \ldots, \pi_\ell, \ldots, \pi_k) - h(\pi_1, \ldots, \pi'_\ell, \ldots, \pi_k)\right|$$
$$\leq \frac{4}{m} |f_\ell| \sum_{\substack{1 \leq j \leq k \\ j \neq \ell}} |f_j| \sqrt{m} = \frac{4}{\sqrt{m}} |f_\ell| \sum_{\substack{1 \leq j \leq k \\ j \neq \ell}} |f_j| =: c_\ell,$$

Applying the *Self-Avoiding McDiarmid inequality* [7], one obtains

$$\Pr_\pi \left(|h - \mathbb{E}[h]| \geq \beta \|f\|_2^2\right)$$
$$\leq 2 \exp\Big(- \beta^2 m \|f\|_2^4 \,/ \Big[8 \textstyle\sum_{\ell=1}^k |f_\ell|^2 \Big(\sum_{j \neq \ell} |f_j| \Big)^2\Big]\Big) .$$

Together with the estimate

$$\textstyle\sum_{\ell=1}^k |f_\ell|^2 \Big(\sum_{j \neq \ell} |f_j| \Big)^2 \leq k \, \|f\|_2^4 \ ,$$

which follows from Cauchy–Schwarz inequality, it follows that $\Pr_\pi \left(|h - \mathbb{E}[h]| \geq \beta \|f\|_2^2\right) \leq 2 \exp\left(-\frac{\beta^2 m}{8k}\right)$. Substituting $\|g\|_2^2$ for h and using (5), we obtain finally

$$\Pr_\pi \left[\left|\|g\|_2^2 - \|f\|_2^2\right| \geq \left(\beta + \tfrac{k-1}{m^2-1}\right) \|f\|_2^2\right] \leq 2 \exp\left(-\tfrac{\beta^2 m}{8k}\right).$$

3) For $\delta > \frac{k-1}{m^2-1}$, we set $\beta = \delta - \frac{k-1}{m^2-1}$ and ϵ as in (4). Then, using that $\|g\|_2^2 = \frac{1}{m}\|\Phi f\|_2^2 = \|G_\alpha f\|_2^2$, one obtains

$$\Pr_\pi \left(\left|\|G_\alpha f\|_2^2 - \|f\|_2^2\right| \geq \delta \|f\|_2^2\right) \leq \epsilon.$$

This shows that G_α is a near-isometry for k-sparse vectors with probability at least $1 - \epsilon$ and proves the statement of the theorem. ∎

Sketch of proof (Theorem 4): From Theorem 3, G_α has (k, δ, ϵ)-*StRIP* and therefore it has also $(k, \delta, 2\epsilon)$-*StRIP*. To prove that G_α has $(k, \delta, 2\epsilon)$-UStRIP, it suffices to show that the probability of having $G_\alpha f = G_\alpha f'$ with distinct k-sparse vectors $f \neq f'$ is less than ϵ. Note that $G_\alpha f = G_\alpha f'$ is equivalent to $\Phi f = \Phi f'$, where $\Phi = S_\alpha G_\alpha$ as in the proof of Theorem 3. Again, we only sketch the main steps, leaving some longer calculations to the reader.

1) Fix any $r \in \{1, \ldots, m^2\}$ and let $S = \{\pi_1, \ldots, \pi_k\}$ be the set of the first k elements in a random permutation of the $m^2 - 1$ elements $\{1, \ldots, m^2\} \backslash \{r\}$. First we show that

$$\mathbb{E}_S \left[\|\Phi_S^* \phi_r\|^2\right] = \frac{km^2}{m+1}, \tag{8}$$

where the expectation is taken with respect to the choice of S. This equations follows from

$$\mathbb{E}_S \left[\|\Phi_S^* \, \phi_r\|^2\right] = \textstyle\sum_{i=1}^k \mathbb{E}_S \left[|\langle \phi_r, \phi_{\pi_i} \rangle|^2\right] k \, \mathbb{E}_S \left[|\langle \phi_r, \phi_{\pi_1} \rangle|^2\right]$$

and by using Properties (P2) and (P3) of Φ.

2) Next, we show that with probability exceeding $1 - \epsilon$, any random subset $S \subset \{1, \ldots, m^2\}$ of size k and any $r \in S^c$ satisfy

$$\|\Phi_S^* \phi_r\|^2 \leq k\, m + m \sqrt{2k \log (m^2/\epsilon)}. \tag{9}$$

To see this, let $y(t_1, \ldots, t_k) = \sum_{i=1}^k \left|\phi_{t_i}^* \phi_r\right|^2$ where $t_1, \ldots, t_k$ are k distinct elements chosen randomly from $\{1, \ldots, m^2\} \setminus \{r\}$, with r fixed. For $t_i \neq t'_i$, we have

$$\mathbb{E}\left[y(t_1, \ldots, t_i, \ldots, t_k)\right] - \mathbb{E}\left[y(t_1, \ldots, t'_i, \ldots, t_k)\right]$$
$$= \left| |\phi_i^* \phi_r|^2 - |\phi_{t'_i}^* \phi_r|^2 \right|$$
$$= \left| \left|\textstyle\sum_{x=0}^{m-1} \phi_{\gamma(r, t_i)}(x)\right|^2 - \left|\sum_{x=0}^{m-1} \phi_{\gamma(r, t'_i)}(x)\right|^2 \right| \leq 2m,$$

using (P1) and (P2) together with the fact that $\gamma(t_i, r) \neq 0$ and $\gamma(t'_i, r) \neq 0$, since t_i, t'_i are distinct from r. Similarly as in the proof of Theorem 3, we apply the McDiarmid inequality which yields

$$\Pr\left[\|\Phi_S^* \phi_r\|^2 \geq km + \xi\right] \leq \Pr\left[\|\Phi_S^* \phi_r\|^2 \geq \tfrac{km^2}{m+1} + \xi\right]$$
$$\leq \exp\left(-\tfrac{\xi^2}{2km^2}\right).$$

Taking the union bound over all m^2 choices of r yields (9).

3) When $m = \mathrm{O}(k\delta^{-2} \log m)$, (9) becomes

$$\left\| \tfrac{1}{\sqrt{m}} \Phi_S^* \, \tfrac{1}{\sqrt{m}} \phi_w \right\|^2 \leq \frac{k}{m} + \frac{1}{m} \sqrt{2k \log m^2/\epsilon}$$
$$= \mathrm{O}\left(\frac{\delta^2}{\log m}\right) + \mathrm{O}\left(\delta^2 \left(1 + \frac{|\log \delta|}{\log m}\right)^{1/2} (k \log m)^{-1/2}\right),$$

which indicates a small coherence between the columns of Φ_S and the remaining columns.

4) Since G_α has (k, δ, ϵ)-*StRIP* (see Theorem 3), one can deduce the following statement, as in [7, Lemma 19]:

Let $S = \{\pi_1, \ldots, \pi_k\}$ be the set of the first k elements in a random permutation of $\{1, \ldots, m^2\}$. With probability exceeding $1 - \epsilon$, any subset $\lambda \subset \{1, \ldots, m^2\}$ of size k with $\lambda \neq S$ satisfies

$$\dim \left(\mathrm{range}\,(G_\alpha)_\lambda \cap \mathrm{range}\,(G_\alpha)_S\right) < k. \tag{10}$$

5) We consider $f \in \mathbb{C}^{m^2}$ as a random vector supported on $S = \{\pi_1, \ldots, \pi_k\}$. The k nonzero entries are chosen randomly from a k-dimensional vectors space equipped with a measure which is absolutely continuous with respect to the Lebesgue measure.

6) Using Theorems B and D in [20] and the fact that G_α has (k, δ, ϵ)-*StRIP*, it follows that with probability exceeding $1 - \epsilon$, $(G_\alpha)_S$ is injective, i.e, $\dim(\mathrm{range}(\Phi_S)) = k$. This means that with probability exceeding $1 - \epsilon$, no two signals supported on S can have the same measurement.

7) If there exists f' supported on $\lambda \neq S$ with $G_\alpha f' = G_\alpha f$, then (10) implies that with probability exceeding $1 - \epsilon$, the vector f restricted to S lies in at most $(k-1)$-dimensional subspace of $\mathbb{C}^k$, which is of zero measure with respect to any absolutely continuous measure on $\mathbb{C}^k$. That is, with probability larger $1 - \epsilon$, the set of vectors f satisfying $G_\alpha f' = G_\alpha f$ for some f' supported on $\lambda \neq S$ is a measure zero set.

8) Combining the above two statements, we conclude that with probability exceeding $1 - 2\epsilon$ (with respect to random choice of f), f is the only k-sparse vector satisfying the equation $y = G_\alpha f$. ∎

7. REFERENCES

[1] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[2] D. L. Donoho, "Compressive Sensing," vol. 52, no. 4, pp. 1289–1306, 2006.

[3] S. Mendelson and A. Pajor and N. Tomczak-Jaegermann, "Uniform uncertainty principle for Bernoulli and subgaussian ensembles," *Constr. Approx.*, vol. 28, no. 3, pp. 277–289, Dec. 2008.

[4] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[5] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Basel: Birkhäuser, 2013.

[6] M. Elad, *Sparse and Redundant Representations*. New York, USA: Springer, 2010.

[7] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[8] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complexity*, vol. 23, no. 4, pp. 918–925, 2007.

[9] D. Gabor, "Theory of communication," *J. of IEE, Part III: Radio and Commun. Eng.*, vol. 93, no. 26, pp. 429–441, Nov. 1946.

[10] S. D. Howard, A. R. Calderbank, and W. Moran, "The finite Heisenberg–Weyl groups in radar and communications," *EURASIP Journal on Applied Signal Processing*, vol. Articel ID 85685, pp. 1–12, 2006.

[11] G. E. Pfander and D. F. Walnut, "Measurement of time-variant linear channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4808–4820, Nov. 2006.

[12] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, Jun. 2009.

[13] R. Heckel and H. Bölcskei, "Identification of sparse linear operators," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7985–8000, Dec. 2013.

[14] G. E. Pfander, H. Rauhut, and J. A. Tropp, "The restricted isometry property for time–frequency structured random matrices," *Probab. Theory Related Fields*, vol. 156, no. 3-4, pp. 707–737, 2013.

[15] F. Krahmer, S. Mendelson, and H. Rauhut, "Suprema of chaos processes and the restricted isometry property," *Comm. Pure Appl. Math.*, vol. 67, no. 11, pp. 1877–1904, 2014.

[16] W. Alltop, "Complex sequences with low periodic correlations," *IEEE Trans. Inf. Theory*, vol. 26, no. 3, pp. 350–354, May 1980.

[17] L. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.

[18] J. Lawrence, G. E. Pfander, and D. Walnut, "Linear independence of Gabor systems in finite dimensional vector spaces," *J. Fourier Anal. Appl.*, vol. 11, no. 6, pp. 715–726, 2005.

[19] R.-D. Malikiosis, "A note on Gabor frames in finite dimensions," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 2, pp. 318–330, 2015.

[20] J. Tropp, "On the conditioning of random subdictionaries," *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 1–24?, 2008.