The Restricted Isometry Property for Measurements from Group Orbits

Hartmut Führ
Chair of Geometry and Analysis
RWTH Aachen University
Aachen, Germany
fuehr@mathga.rwth-aachen.de

Timm Gilles
Chair of Geometry and Analysis
RWTH Aachen University
Aachen, Germany
gilles@mathga.rwth-aachen.de

***Abstract*—It is known that sparse recovery by measurements from random circulant matrices provides good recovery bounds. We generalize this to measurements that arise as a random orbit of a group representation for some finite group G. We derive estimates for the number of measurements required to guarantee the restricted isometry property with high probability. Following this, we present several examples highlighting the role of appropriate representation-theoretic assumptions.**

***Index Terms*—compressed sensing, group representations.**

I. Introduction

We are interested in recovering an s-sparse signal $x \in \mathbb{C}^n$, i.e. $\|x\|_0 := |\{j \in \{1,\dots,n\} \mid x_j \neq 0\}| \leq s$, which is measured by the linear measurement process

$$y = \Phi x$$

from as few measurements as possible. Here, $\Phi \in \mathbb{C}^{m\times n}$ is called *measurement matrix* and $y \in \mathbb{C}^m$ is the vector of observed measurements. For a matrix $\Phi \in \mathbb{C}^{m\times n}$ and $1 \leq s \leq n$, the restricted isometry constant δ_s is the smallest $\delta \geq 0$ such that

$$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2$$

for all s-sparse vectors $x \in \mathbb{C}^n$. It is well known that if Φ satisfies the above condition, also called restricted isometry property, for a small enough constant δ_s, there exist efficient algorithms that recover any s-sparse signal x, see [1], [2], [3]. It has been shown in [4] that if Φ is a Gaussian random matrix, i.e its entries are independent and follow a normal distribution with mean 0 and variance 1, and

$$m \geq C\delta^{-2} s \ln\left(\frac{n}{s}\right) \tag{1}$$

holds, then $\delta_s \leq \delta$ with high probability. Moreover, all so far known measurement matrices that provide optimal bounds, i.e. m scales as in (1) [5], are random. However, in practice, one usually has more structured measurement processes. Specific choices of structured measurement matrices have been studied, such as Random Partial Fourier Matrices in [6], or Partial Random Circulant Matrices and Time-Frequency Structured Random Matrices in [7].

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1481 – 442047500

In this work however, we aim for a more general setup of structured measurement matrices that arise from inner products of the signal x with elements of some group orbit. To precisely state the measurement matrix Φ of interest, we first need some notation: Let G be a finite group. We define a projective representation $\pi\colon G \to \mathrm{GL}(\mathbb{C}^n)$ to be a mapping such that for all $g,h \in G$ there exists $\lambda(g,h) \in \mathbb{C}$ with $|\lambda(g,h)| = 1$ and

$$\pi(g)\pi(h) = \lambda(g,h)\pi(gh).$$

π is called unitary if $\pi(g)$ is unitary for every $g \in G$. Lastly, π is a representation if $\lambda(g,h) = 1$ for all $g,h \in G$. For a more detailed introduction to group representations see [8]. Lastly, we call a random vector X on $\mathbb{C}^n$ L-subgaussian if $\mathbb{E}X = 0$, $\mathbb{E}|\langle X,\theta\rangle|^2 = 1$ for every $\theta \in S^{n-1}$ and

$$\mathbb{P}(|\langle X,\theta\rangle| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

for every $\theta \in S^{n-1}$ and any $t > 0$. With this notation at hand we are ready to define the measurement matrix Φ.

Definition 1. *Let $\pi\colon G \to GL(\mathbb{C}^n)$ be a unitary (projective) representation. Fix some sampling set $\Omega = \{\omega_1,\dots,\omega_m\} \subset G$ of size $|\Omega| = m$. Let ξ be a random vector with independent mean 0, variance 1, and L-subgaussian entries. Set*

$$\Phi = \frac{1}{\sqrt{m}} R_\Omega\big(\pi(g)\xi\big)^*_{g\in G} \in \mathbb{C}^{m\times n}$$

where $(\pi(g)\xi)_{g\in G}$ denotes the matrix in $\mathbb{C}^{n\times n}$ which columns are the vectors $\pi(g)\xi$ and $R_\Omega\colon \mathbb{C}^G \to \mathbb{C}^m$ restricts a vector to its entries in Ω, i.e. $(R_\Omega y)_l = y(\omega_l)$.

A special case of Definition 1 are the Partial Random Circulant Matrices mentioned before. Therefore, our main result, Theorem 2, generalizes [7, Theorem 4.1]. We will make this precise in Section III.

II. The restricted isometry property

Intuitively, different representations should give different sparse recovery results. For example, it can be shown that sparse recovery is not possible when considering the trivial representation $\pi(g) = \mathrm{Id}$. To measure this performance difference between representations, we have added the restriction (2) in Theorem 2. Here, the size of the constant $C_{\Omega,\pi}$ plays

a crucial role. It essentially characterizes how well a given representation performs in sparse recovery. In Section III, we present examples for the choice of $C_{\Omega,\pi}$.

The proof idea of Theorem 2 was first seen in [7].

Theorem 2. *Let Φ be as in Definition 1. Further, let $C_{\Omega,\pi} > 0$ be a constant, which can depend on all of Ω's characteristics except its size m, such that*

$$\sup_{1\le j\le n, y\in\mathbb{C}^n} \Big\| \frac{1}{\sqrt{m}} R_\Omega\big(\pi(g)y\big)^*_{g\in G} e_j \Big\|_2 \le \sqrt{\frac{C_{\Omega,\pi}}{m}} \|y\|_2 \quad (2)$$

holds. If, for $s \le n$ and $\delta, \eta \in (0,1)$,

$$m \ge c\delta^{-2} s\, C_{\Omega,\pi} \max\left\{(\log(s\, C_{\Omega,\pi})\log n)^2,\, \log(\eta^{-1})\right\}, \quad (3)$$

then with probability at least $1-2\eta$, the restricted isometry constant of Φ satisfies $\delta_s \le \delta$. Here, $c > 0$ only depends on L.

Proof: We start by explaining the main idea of the proof: The restricted isometry constant of Φ can be rewritten as supremum of a chaos process for some set of matrices $\mathcal{A}$. For this type of process, we can then use the concentration inequality of the following theorem.

Theorem 3. *[7, Theorem 3.1] Let $\mathcal{A}$ be a set matrices, and let ξ be a random vector whose entries ξ_j are independent, mean 0, variance 1 and L-subgaussian random variables. Define $d_{2\to2}(\mathcal{A}) = \sup_{A\in\mathcal{A}} \|A\|_{2\to2}$ and $d_F(\mathcal{A}) = \sup_{A\in\mathcal{A}} \|A\|_F$. By $\gamma_2(\mathcal{A}, \|\cdot\|_{2\to2})$ we denote Talagrand's functional (see [7, Definition 2.1]). Set*

$$\begin{aligned} E &= \gamma_2(\mathcal{A}, \|\cdot\|_{2\to2})(\gamma_2(\mathcal{A}, \|\cdot\|_{2\to2}) + d_F(\mathcal{A})) \\ &\quad + d_F(\mathcal{A}) d_{2\to2}(\mathcal{A}), \\ V &= d_{2\to2}(\mathcal{A})(\gamma_2(\mathcal{A}, \|\cdot\|_{2\to2}) + d_F(\mathcal{A})), \\ U &= d^2_{2\to2}(\mathcal{A}). \end{aligned}$$

Then, for $t > 0$,

$$\begin{aligned} &\mathbb{P}\left(\sup_{A\in\mathcal{A}} \left|\|A\xi\|_2^2 - \mathbb{E}\|A\xi\|_2^2\right| \ge c_1 E + t\right) \\ &\le 2\exp\left(-c_2 \min\left\{\frac{t^2}{V^2}, \frac{t}{U}\right\}\right). \end{aligned}$$

The constants $c_1, c_2 > 0$ depend only on L.

Then, it will be sufficient to find suitable bounds for $d_{2\to2}(\mathcal{A})$, $d_F(\mathcal{A})$ and $\gamma_2(\mathcal{A}, \|\cdot\|_{2\to2})$. Let us start with some notation. Let

$$D_{s,n} = \{x \in \mathbb{C}^n \mid \|x\|_2 \le 1, \|x\|_0 \le s\}$$

and for every $x \in \mathbb{C}^n$ define the matrix

$$A_x = \frac{1}{\sqrt{m}} R_\Omega\left(\pi(g)^* x\right)^*_{g\in G} \in \mathbb{C}^{m\times n}.$$

Further, set $\mathcal{A} = \{A_x \mid x \in D_{s,n}\}$. Since the entries of ξ are independent and have mean 0 and variance 1, it holds

$$\mathbb{E}|\langle \xi, y\rangle|^2 = \mathbb{E}\sum_{i,j=1}^n y_i \overline{y_j}\, \overline{\xi_i}\xi_j = \sum_{i,j=1}^n y_i\overline{y_j}\, \underbrace{\mathbb{E}\overline{\xi_i}\xi_j}_{=\delta_{ij}} = \|y\|_2^2$$

for all $y \in \mathbb{C}^n$. Together with the fact that the representation π is unitary we get

$$\begin{aligned} \mathbb{E}\|A_x\xi\|_2^2 &= \mathbb{E}\sum_{g\in\Omega} \frac{1}{m} |\langle \xi, \pi(g)^* x\rangle|^2 = \frac{1}{m}\sum_{g\in\Omega} \mathbb{E}|\langle \xi, \pi(g)^* x\rangle|^2 \\ &= \frac{1}{m}\sum_{g\in\Omega} \|\pi(g)^* x\|_2^2 = \|x\|_2^2 \end{aligned}$$

for all $x \in \mathbb{C}^n$. Now we are able to rewrite the restricted isometry constant as supremum of a chaos process, i.e

$$\delta_s = \sup_{x\in D_{s,n}} \left|\|\Phi x\|_2^2 - \|x\|_2^2\right| = \sup_{x\in D_{s,n}} \left|\|A_x\xi\|_2^2 - \mathbb{E}\|A_x\xi\|_2^2\right|.$$

For better readability, we prove bounds for the above quantities separately in the following lemmata.

Lemma 4. *It holds*

$$\|A_x\|_{2\to2} \le \sqrt{\frac{s\, C_{\Omega,\pi}}{m}} \|x\|_2$$

for all $x \in \mathbb{C}^n$ with $\|x\|_0 \le s$. Thus, $d_{2\to2}(\mathcal{A}) \le \sqrt{\frac{s\, C_{\Omega,\pi}}{m}}$.

Proof: Let $x \in \mathbb{C}^n$ with $\|x\|_0 \le s$. We can rewrite the s-sparse vector x as $\sum_{j\in J} \alpha_j e_j$ with $|J| \le s$. Then, the triangle inequality and Cauchy-Schwarz inequality give

$$\begin{aligned} \Big\|\frac{1}{\sqrt{m}} R_\Omega(\pi(g)y)^*_{g\in G} x\Big\|_2 &\le \sum_{j\in J} |\alpha_j| \Big\|\frac{1}{\sqrt{m}} R_\Omega(\pi(g)y)^*_{g\in G} e_j\Big\|_2 \\ &\le \sum_{j\in J} |\alpha_j| \sqrt{\frac{C_{\Omega,\pi}}{m}} \|y\|_2 \\ &\le \sqrt{\frac{C_{\Omega,\pi}}{m}} \|y\|_2 \sqrt{|J|} \Big(\sum_{j\in J} |\alpha_j|^2\Big)^{\frac{1}{2}} \\ &\le \sqrt{\frac{s\, C_{\Omega,\pi}}{m}} \|y\|_2 \|x\|_2 \end{aligned}$$

for all $y \in \mathbb{C}^n$ and $g \in \Omega$. Since $|\langle y, \pi(g)^* x\rangle| = |\langle x, \pi(g) y\rangle|$ holds for all $y \in \mathbb{C}^n$ and $g \in \Omega$, it follows that

$$\|A_x\|_{2\to2} = \sup_{\|y|_2=1} \Big\|\frac{1}{\sqrt{m}} R_\Omega\big(\pi(g)y\big)^*_{g\in G} x\Big\|_2 \le \sqrt{\frac{s\, C_{\Omega,\pi}}{m}} \|x\|_2.$$

Hence, $d_{2\to2}(\mathcal{A}) \le \sqrt{\frac{s\, C_{\Omega,\pi}}{m}}$. ∎

Lemma 5. *It holds $d_F(\mathcal{A}) = 1$.*

Proof: Let $x \in \mathbb{C}^n$ with $\|x\|_0 \le s$. Then, π being unitary implies

$$\begin{aligned} \|A_x\|_F^2 &= \frac{1}{m}\Big\|R_\Omega\left(\pi(g)^* x\right)^*_{g\in G}\Big\|_F^2 = \frac{1}{m}\Big\|\left(\pi(g)^* x\right)_{g\in\Omega}\Big\|_F^2 \\ &= \frac{1}{m}\Big\|\big(\pi(g_1)^* x \mid \cdots \mid \pi(g_m)^* x\big)\Big\|_F^2 \\ &= \frac{1}{m}\sum_{i=1}^m \|\pi(g_i)^* x\|_2^2 = \frac{1}{m}\sum_{i=1}^m \|x\|_2^2 = \|x\|_2^2. \end{aligned}$$

Thus, $d_F(\mathcal{A}) = 1$. ∎

Lastly, we have to bound the term $\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2})$. In [9], a bound in form of a Dudley integral (first considered in [10])

$$\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2}) \lesssim \int_0^{d_{2\to 2}(\mathcal{A})} \sqrt{\log \mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t)}\, \mathrm{d}t$$

was shown. Here, $A \lesssim B$ means that there exists a universal constant $c_1 > 0$ such that $A \leq c_1 B$. Usually, one then proves two bounds for the covering number $\mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t)$, which is the minimal number of open balls in $(\mathcal{A}, \|\cdot\|_{2\to 2})$ of radius t that is needed to cover $\mathcal{A}$. The first bound uses the empirical method of Maurey [11] and the second one a volumetric argument. This approach has also been used in several similar settings, see [12] or [13].

Lemma 6. *Let $t > 0$. We have the following inequalities:*

$$\log \mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t) \lesssim \frac{s\, C_{\Omega,\pi}}{m} \frac{1}{t^2} \log(n) \log(4n)$$

and

$$\log \mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t) \lesssim s\Big(\log\Big(1 + 2\frac{\sqrt{s\, C_{\Omega,\pi}}}{\sqrt{m}\, t}\Big) + \log\Big(\frac{en}{s}\Big)\Big).$$

Proof: Let $t > 0$. Define a set of matrices by

$$B = \{A_{\pm\sqrt{2}e_1}, \dots, A_{\pm\sqrt{2}e_n}, A_{\pm\sqrt{2}ie_1}, \dots, A_{\pm\sqrt{2}ie_n}\}.$$

It is easy to check that

$$D_{s,n} \subseteq \sqrt{s}\, \mathrm{conv}(\pm\sqrt{2}e_1, \dots, \pm\sqrt{2}e_n, \pm\sqrt{2}ie_1, \dots, \pm\sqrt{2}ie_n)$$

holds, where conv() denotes the convex hull. Combining this with the fact that

$$A_{\sum_{j=1}^N \alpha_j z^j} = \sum_{j=1}^N \overline{\alpha_j}\, A_{z^j}$$

for all $\alpha_1, \dots, \alpha_N \in \mathbb{C}$ and $z^1, \dots, z^N \in \mathbb{C}^n$, gives the relation $\mathcal{A} \subset \sqrt{s}\, \mathrm{conv}(B)$. Thus,

$$\mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t) \leq \mathcal{N}\left(\mathrm{conv}(B), \|\cdot\|_{2\to 2}, \frac{1}{\sqrt{s}} t\right). \quad (4)$$

To bound the right side we use [7, Lemma 4.2] which is based on the empirical method of Maurey. Let $N \in \mathbb{N}$, $(A_1, \dots, A_N) \in B^L$ and let $(e_j)_{j=1}^N$ be a random vector with independent Rademacher distributed entries. It holds

$$\begin{aligned}\mathbb{E}_\epsilon \Big\| \sum_{j=1}^N \epsilon_j A_j \Big\|_{2\to 2} &\lesssim \sqrt{\log(n)} \max\Big\{ \Big\| \sum_{j=1}^N A_j^* A_j \Big\|_{2\to 2}, \\ &\qquad \Big\| \sum_{j=1}^N A_j A_j^* \Big\|_{2\to 2} \Big\}^{\frac{1}{2}} \\ &\leq \sqrt{\log(n)} \Big(\sum_{j=1}^N \|A_j\|_{2\to 2}^2 \Big)^{\frac{1}{2}}\end{aligned}$$

where we used the non-commutative Khintchine inequality due to Lust-Piquard [14], [15]. The definition of B and our assumption on π in (2) gives

$$\sqrt{\log(n)} \Big(\sum_{j=1}^N \|A_j\|_{2\to 2}^2 \Big)^{\frac{1}{2}} \leq \sqrt{\log(n)}\, \frac{\sqrt{2\, C_{\Omega,\pi}}}{\sqrt{m}} \sqrt{N}.$$

Now it follows from [7, Lemma 4.2] that

$$\log \mathcal{N}\Big(\mathrm{conv}(B), \|\cdot\|_{2\to 2}, \frac{t}{\sqrt{s}}\Big) \lesssim \frac{s\, C_{\Omega,\pi}}{m} \frac{1}{t^2} \log(n) \log(4n). \quad (5)$$

Putting (4) and (5) together yields the first inequality.

To prove the second inequality, first note by inspecting the proof of Lemma 4 we have

$$\|A_x - A_y\|_{2\to 2} = \|A_{x-y}\|_{2\to 2} \leq \sqrt{\frac{C_{\Omega,\pi}}{m}} \|x - y\|_1.$$

Therefore,

$$\mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t) \leq \mathcal{N}\left(D_{s,n}, \sqrt{\frac{C_{\Omega,\pi}}{m}} \|\cdot\|_1, t \right).$$

Now following the arguments presented in Section 8.4 of [13] shows that

$$\mathcal{N}\left(D_{s,n}, \sqrt{\frac{C_{\Omega,\pi}}{m}} \|\cdot\|_1, t \right) \leq \left(1 + 2\sqrt{\frac{s\, C_{\Omega,\pi}}{m}} \frac{1}{t} \right)^{2s} \Big(\frac{en}{s}\Big)^s.$$

The second inequality follows immediately. ■

Now we are able to prove the bound for $\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2})$.

Lemma 7. *We have*

$$\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2}) \lesssim \sqrt{C_{\Omega,\pi}} \sqrt{\frac{s}{m}} \sqrt{\log(n) \log(4n)} \log(s\, C_{\Omega,\pi}).$$

Proof: Using the first inequality that was shown in Lemma 6 together with Lemma 4 yields

$$\begin{aligned}&\int_{\frac{1}{\sqrt{m}}}^{d_{2\to 2}(\mathcal{A})} \sqrt{\log \mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t)}\, \mathrm{d}t \\ &\lesssim \int_{\frac{1}{\sqrt{m}}}^{\sqrt{\frac{s\, C_{\Omega,\pi}}{m}}} \sqrt{C_{\Omega,\pi}} \sqrt{\frac{s}{m}} \sqrt{\log(n)} \sqrt{\log(4n)}\, \frac{1}{t}\, \mathrm{d}t \\ &= \frac{1}{2} \sqrt{C_{\Omega,\pi}} \sqrt{\frac{s}{m}} \sqrt{\log(n)} \sqrt{\log(4n)} \log(s\, C_{\Omega,\pi}).\end{aligned}$$

Performing a similar calculation as presented in Section 8.4 in [13] together with the second bound in Lemma 6 gives

$$\begin{aligned}&\int_0^{\frac{1}{\sqrt{m}}} \sqrt{\log(\mathcal{N}(\mathcal{A}, \|\cdot\|_{2\to 2}, t))}\, \mathrm{d}t \\ &\lesssim \sqrt{s} \int_0^{\frac{1}{\sqrt{m}}} \sqrt{\log\Big(\frac{en}{s}\Big) + \log\Big(1 + 2\frac{\sqrt{s\, C_{\Omega,\pi}}}{\sqrt{m}\, t}\Big)}\, \mathrm{d}t \\ &\leq \sqrt{\frac{s}{m}} \Big(\sqrt{\log\Big(\frac{en}{s}\Big)} + \sqrt{\log\Big(e\big(1 + 2\sqrt{s\, C_{\Omega,\pi}}\big)\Big)} \Big).\end{aligned}$$

The statement follows from putting both inequalities together. Note that if we were to just use the second bound we would get worse scaling in the exponent of s. ■

With these three bounds at hand, we return to proving Theorem 2. Let $\delta \in (0,1)$. Using the above lemmata as well as the inequality (3) yields

$$E \leq \frac{\delta^2}{c} + \frac{\delta}{\sqrt{c}} + \frac{\delta}{\sqrt{c}} \leq \frac{3\delta}{\sqrt{c}},$$

where we assumed that $c \geq 1$. Now by choosing the constant c large enough, we get the bound $E \lesssim \frac{\delta}{2c_1}$ where c_1 is the absolute constant from Theorem 3. Using that same Theorem then gives

$$\begin{aligned}\mathbb{P}(\delta_s \geq \delta) &\leq \mathbb{P}\Big(\delta_s \geq c_1 E + \frac{\delta}{2}\Big) \\ &\leq 2\exp\Big(-c_2 \min\Big\{\frac{0.25\delta^2}{V^2}, \frac{0.5\delta}{U}\Big\}\Big).\end{aligned}$$

Using the Lemmata 4, 5 and 7 as well as inequality (3) gives the bound

$$\begin{aligned}&2\exp\Big(-c_2 \min\Big\{\frac{0.25\delta^2}{V^2}, \frac{0.5\delta}{U}\Big\}\Big) \\ &\leq 2\exp\left(-c_2 c \frac{0.25}{\big(c^{-\frac{1}{2}}+1\big)^2}\log(\eta^{-1})\right).\end{aligned}$$

Choose the constant c again large enough and deduce that

$$2\exp\left(-c_2 c \frac{0.25\log(\eta^{-1})}{\big(c^{-\frac{1}{2}}+1\big)^2}\right) \leq 2\exp\big(-\log(\eta^{-1})\big) = 2\eta.$$

This establishes the theorem. ■

III. The constant $C_{\Omega,\pi}$

The dependence of the estimate in Theorem 2 on the choice of representation and sampling set is encapsulated in the constant $C_{\Omega,\pi}$, on which the number of measurements m depends linearly. The determination of this constant (preferably close to 1) remains a challenge. In this section, we present some examples that illustrate this fact.

The next proposition establishes that the optimal value 1 is obtained by the left regular representation $L\colon G \to \mathrm{GL}(\mathbb{C}^G)$, defined by $(L(g)f)(h) = f(g^{-1}h)$.

Proposition 8. *Let Φ be constructed by choosing the left regular representation L and as in Definition 1. Then, it holds*

$$\Big\|\frac{1}{\sqrt{m}} R_\Omega\big(L(g)y\big)^*_{g\in G} e_h\Big\|_2 \leq \sqrt{\frac{1}{m}}\|y\|_2$$

for all $y \in \mathbb{C}^G$ and all canonical vectors $e_h \in \mathbb{C}^G$.

Proof: For $y \in \mathbb{C}^G$ and a canonical vector e_h we have

$$\begin{aligned}\Big\|\frac{1}{\sqrt{m}} R_\Omega\big(L(g)y\big)^*_{g\in G} e_h\Big\|_2^2 &\leq \frac{1}{m}\sum_{g\in G}|\langle e_h, L(g)y\rangle|^2 \\ &= \frac{1}{m}\sum_{g\in G}|y(g^{-1}h)|^2 = \frac{1}{m}\|y\|_2^2.\end{aligned}$$

■

It is straightforward to verify that choosing $G = \mathbb{Z}_n$ and $\pi = L$ results in Φ being a Partial Random Circulant Matrix as examined in [7]. So, combining Theorem 2 and Proposition 8 generalizes [7, Theorem 4.1] from $\mathbb{Z}_n$ to arbitrary, even non-abelian, finite groups.

Remark 9. *It is possible to generalize Proposition 8 to subrepresentations of the left regular representation L to a certain degree by using more knowledge about group representations than we have presented here.*

To illustrate another advantage of stating Theorem 2 depending on the sampling set as well as the representation, we add one more example: Let $p \in \mathbb{N}$ be prime. We call the set $G = \mathbb{Z}_p \times \mathbb{Z}_p^*$ together with the operation

$$(k,l)(k',l') := (k + lk' \bmod p, ll' \bmod p)$$

the affine group. We define a unitary representation of G by

$$\rho\colon G \to \mathrm{GL}(\mathbb{C}^p),\ (\rho(k,l)f)(j) = f(l^{-1}(j-k))$$

for all $(k,l) \in G$.

Proposition 10. *Let Φ be constructed by choosing the representation ρ and as in Definition 1. Then, it holds*

$$\Big\|\frac{1}{\sqrt{m}} R_\Omega\big(\rho(k,l)y\big)^*_{(k,l)\in G} e_j\Big\|_2 \leq \sqrt{\frac{|\Omega_2|}{m}}\|y\|_2$$

for all $y \in \mathbb{C}^p$ and all canonical vectors $e_j \in \mathbb{C}^p$ with $\Omega_2 := \{l \in \mathbb{Z}_p^* \mid \exists k \in \mathbb{Z}_p\colon (k,l) \in \Omega\}$.

Proof: For $y \in \mathbb{C}^p$ and a canonical vector e_j we have

$$\begin{aligned}&\Big\|\frac{1}{\sqrt{m}} R_\Omega\big(\rho(k,l)y\big)^*_{(k,l)\in G} e_j\Big\|_2^2 \leq \frac{1}{m}\sum_{l\in\Omega_2}\sum_{k=1}^{p}|\langle e_j, \rho(k,l)y\rangle|^2 \\ &= \frac{1}{m}\sum_{l\in\Omega_2}\sum_{k=1}^{p}|y(lj+k)|^2 = \frac{|\Omega_2|}{m}\|y\|_2^2.\end{aligned}$$

■

Since the measurement matrix is known, we can use that knowledge by a priori restricting the sampling set Ω to be a subset of e.g. $G = \{(k,1) \in G \mid k \in \mathbb{Z}_p\}$. Thus, we get $|\Omega_2| = 1$ in Proposition 10 and $C_{\Omega,\rho} = 1$.

Remark 11. *This idea of a priori restricting the sampling set can actually be generalized to a broader class of representations, namely induced representations of split group extensions. The details will be provided elsewhere.*

We want to finish this section with the following important observation: The availability of sampling sets Ω with $C_{\Omega,\pi}$ close to 1 can depend on the realization of the representation. For example, the realization of the left regular representation of $\mathbb{Z}_n$ in the Fourier domain gives a measurement matrix Φ that does not do sparse recovery for any Ω. However, randomizing the sampling set Ω and slightly adapting Φ allows for sparse recovery independently of the chosen realization at the cost of slightly larger sampling sets than prescribed in Theorem 2.

Understanding this phenomenon, as well as further questions that are linked to the analysis of the constant $C_{\Omega,\pi}$, is still ongoing work and will be presented elsewhere.

ACKNOWLEDGMENT

We thank H. Rauhut (LMU Munich) for fruitful discussion and advice.

REFERENCES

[1] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE transactions on information theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[2] Q. Mo and S. Li, "New bounds on the restricted isometry constant δ_{2k}," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 460–468, 2011.

[3] S. Foucart, "Hard thresholding pursuit: an algorithm for compressive sensing," *SIAM Journal on numerical analysis*, vol. 49, no. 6, pp. 2543–2563, 2011.

[4] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE transactions on information theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[5] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Birkhäuser New York, NY, 2013.

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[7] F. Krahmer, S. Mendelson, and H. Rauhut, "Suprema of chaos processes and the restricted isometry property," *Communications on Pure and Applied Mathematics*, vol. 67, no. 11, pp. 1877–1904, 2014.

[8] A. Terras, *Fourier analysis on finite groups and applications*. Cambridge University Press, 1999, no. 43.

[9] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*. Cambridge university press, 2018, vol. 47.

[10] R. M. Dudley, "The sizes of compact subsets of Hilbert space and continuity of Gaussian processes," *Journal of Functional Analysis*, vol. 1, no. 3, pp. 290–330, 1967.

[11] B. Carl, "Inequalities of Bernstein-Jackson-type and the degree of compactness of operators in Banach spaces," in *Annales de l'institut Fourier*, vol. 35, no. 3, 1985, pp. 79–118.

[12] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 61, no. 8, pp. 1025–1045, 2008.

[13] H. Rauhut, "Compressive sensing and structured random matrices," *Theoretical foundations and numerical methods for sparse recovery*, vol. 9, no. 1, p. 92, 2010.

[14] F. Lust-Piquard, "Inégalités de Khintchine dans $\mathcal{C}_p$ $(1 < p < \infty)$," *CR Acad. Sci. Paris*, vol. 303, pp. 289–292, 1986.

[15] M. Rudelson, "Random vectors in the isotropic position," *Journal of Functional Analysis*, vol. 164, no. 1, pp. 60–72, 1999.