Ark. Mat., **56** (2018), 45–52
DOI: 10.4310/ARKIV.2018.v56.n1.a4

Salem sets in vector spaces over finite fields

Changhao Chen

Abstract. We prove that almost all random subsets of a finite vector space are weak Salem sets (small Fourier coefficient), which extend a result of Hayes to a different probability model.

1. Introduction

Let F_p denote the finite field with p elements where p is prime, and $\mathbb{F}_p^d$ be the d-dimensional vector space over this field. Let $E\subset\mathbb{F}_p^d$. We use the same notation as in Babai [3], Hayes [4] to define that

$$\Phi(E)=\max_{\xi\neq 0}|\widehat{E}(\xi)|. \tag{1}$$

Here and in what follows, we simply write $E(x)$ for the characteristic function of E, $\widehat{E}$ its discrete Fourier transform which we will define in Section 2. For $\xi\neq 0$, we mean that ξ is a non-zero vector of $\mathbb{F}_p^d$. Applying the Plancherel identity, we have that for any $E\subset\mathbb{F}_p^d$ with $\#E\leq p^d/2$,

$$\sqrt{\#E/2}\leq\Phi(E)\leq\#E. \tag{2}$$

See Babai [3, Proposition 2.6] for more details. The notation $\#E$ stands for the cardinality of a set E. Observe that the optimal decay of $\widehat{E}(\xi)$ for all $\xi\neq 0$ are controlled by $O(\sqrt{\#E})$. We write $X=O(Y)$, which means that there is a positive constant C such that $X\leq CY$, and $X=\Theta(Y)$ if $X=O(Y)$ and $Y=O(X)$. Iosevich and Rudnev [6] called these sets Salem sets. To be precise, we show the definition here.

Definition 1.1. ([6]) A subset $E\subset\mathbb{F}_p^d$ is called a Salem set if for all non-zero ξ of $\mathbb{F}_p^d$,

$$|\widehat{E}(\xi)|=O(\sqrt{\#E}). \tag{3}$$

Key words and phrases: finite fields, Salem sets.
2010 Mathematics Subject Classification: 05B25, 52C99.

Note that this is a finite fields version of Salem sets in Euclidean spaces. Roughly speaking, a set in Euclidean space is called a Salem set if measures exist on this set, and the Fourier transform of these measures have optimal decay; see [2] and [9, Chapter 3] for more details on Salem sets in Euclidean spaces.

It is well known that the sets with small Fourier coefficient play an important role in many topics, e.g., see [3], [9] and [12]. For some applications of Salem sets in vector spaces over finite fields, see [5], [6] and [7].

In [4, Theorem 1.13] Hayes proved that almost all m-subset of $\mathbb{F}_p^d$ are (weak) Salem sets which answers a question of Babai. To be precise, let $E=E^\omega$ be selected uniformly at random from the collection of all subsets of $\mathbb{F}_p^d$ which have m vectors. Let $\Omega(\mathbb{F}_p^d, m)$ denotes the probability space.

Theorem 1.2. (Hayes) *Let $\varepsilon>0$. Let $m\leq p^d/2$. For all but an $O(p^{-d\varepsilon})$ probability $E\in\Omega(\mathbb{F}_p^d, m)$,*

$$\Phi(E) < 2\sqrt{2(1+\varepsilon)m\log p^d} = O\left(\sqrt{m\log p^d}\right). \tag{4}$$

For convenience we call this kind of subset of $\mathbb{F}_p^d$ weak Salem set.

1.1. Percolation on $\mathbb{F}_p^d$

There is an another random model which is closely related to the random model $\Omega(\mathbb{F}_p^d, m)$. First, we show this random model in the following. Let $0<\delta<1$. We choose each point of $\mathbb{F}_p^d$ with probability δ and remove it with probability $1-\delta$, all choices being independent of each other. Let $E=E^\omega$ be the collection of these chosen points, and $\Omega=\Omega(\mathbb{F}_p^d, \delta)$ be the probability space. Note that both random models $\Omega(\mathbb{F}_p^d, m)$ and $\Omega(\mathbb{F}_p^d, \delta)$ are related to the well known Erdös-Rényi-Gilbert random graph models.

Note that the random model $\Omega(\mathbb{F}_p^d, \delta)$ has more independence than the random model $\Omega(\mathbb{F}_p^d, m)$. The independence of different vectors in the model $\Omega(\mathbb{F}_p^d, \delta)$ often make the analysis easier. For example, let $F\subset\mathbb{F}_p^d$ then (under the model $\Omega(\mathbb{F}_p^d, \delta)$)

$$\mathbb{P}(F\subset E) = \delta^{\#F}.$$

On the other hand for the random model $\Omega(\mathbb{F}_p^d, m), m\geq\#F$, we have

$$\mathbb{P}(F\subset E) = \frac{m(m-1)\dots(m-\#F+1)}{p^d(p^d-1)\dots(p^d-\#F+1)}. \tag{5}$$

We also show roughly that the model $\Omega(\mathbb{F}_p^d, \delta)$ is closely related to the model $\Omega(\mathbb{F}_p^d, m)$ with $m=p^d\delta$. Observe that if $\#F$ is uniformly bounded, i.e. $\#F\leq C$

where C is a positive constant and $m\to\infty$, then the identity (5) becomes (the dependence become weaker)

$$\mathbb{P}(F\subset E)\longrightarrow\left(\frac{m}{p^d}\right)^{\#F}.$$

Meanwhile, the law of large numbers implies that with high probability each element of $\Omega(\mathbb{F}_p^d,\delta)$ has roughly $p^d\delta=m$ amount of vectors.

We note that Hayes [4] proved a similar result to Theorem 1.2 for the random model $\Omega(\mathbb{F}_p^d,1/2)$. However, the martingale argument for $\Omega(\mathbb{F}_p^d,1/2)$ and $\Omega(\mathbb{F}_p^d,m)$ of [4] do not apply easily to the random model $\Omega(\mathbb{F}_p^d,\delta)$ for other values of $\delta\neq 1/2$. Babai [3, Theorem 5.2] used the Chernoff bounds for the model $\Omega(\mathbb{F}_p^d,1/2)$, but it seems that the method also can not be easily extended to general δ. We note that Babai [3], Hayes [4] proved their results in general finite Abelian group, see [3] and [4] for more details. For the finite vector space $\mathbb{F}_p^d$ (special Abelian group) we extend their result to general δ.

Theorem 1.3. *Let $\varepsilon>0$. Let $\delta\in(p^{\varepsilon_0-d},1)$ with fixed small $\varepsilon_0>0$. For all but an $O(p^{-d\varepsilon})$ probability $E\in\Omega(\mathbb{F}_p^d,\delta)$,*

$$\Phi(E)<2\sqrt{(1+\varepsilon)\delta p^d\log p^d}=O\left(\sqrt{\delta p^d\log p^d}\right). \tag{6}$$

We know that almost all set $E\in\Omega(\mathbb{F}_p^d,\delta)$ has size roughly δp^d. This follows by Chebyshev's inequality,

$$\mathbb{P}(|\#E-p^d\delta|\geq\frac{1}{2}p^d\delta)\leq\frac{4p^d\delta(1-\delta)}{(p^d\delta)^2}=O\left(\frac{1}{\delta p^d}\right). \tag{7}$$

We immediately have the following corollary, which says that almost all $E\in\Omega(\mathbb{F}_p^d,\delta)$ is a weak Salem set.

Corollary 1.4. *Let $\varepsilon>0$. Let $\delta\in(p^{\varepsilon_0-d},1)$ with fixed small $\varepsilon_0>0$. For all but an $O(\max\{p^{-d\varepsilon},\frac{1}{\delta p^d}\})$ probability $E\in\Omega(\mathbb{F}_p^d,\delta)$,*

$$|\widehat{E}(\xi)|=O\left(\sqrt{\#E\log p^d}\right). \tag{8}$$

In $\mathbb{F}_p^d$, it seems that the only known examples of Salem sets are discrete paraboloid and discrete sphere. We note that both the size of the discrete paraboloid and the discrete sphere are roughly p^{d-1}, see [6] for more details. It is natural to ask that does there exist Salem set with any given size $m\leq p^n$. The above results and [8, Problem 20] suggest the following conjecture.

Conjecture 1.5. Let $s\in(0,d)$ be a non-integer number and C be a positive constant. Then

$$\min_{E}\frac{\Phi(E)}{\sqrt{\#E}}\longrightarrow\infty\quad\text{as } p\longrightarrow\infty,$$

where the minimal taking over all subsets $E\subset\mathbb{F}_p^d$ with $p^s/C\leq\#E\leq Cp^s$.

2. Preliminaries

In this section we show the definition of the finite field Fourier transform, and some easy facts about the random model $\Omega(\mathbb{F}_p^d,\delta)$. Let $f:\mathbb{F}_p^d\longrightarrow\mathbb{C}$ be a complex value function. Then for $\xi\in\mathbb{F}_p^d$ we define the Fourier transform

$$\hat{f}(\xi)=\sum_{x\in\mathbb{F}_p^d}f(x)e^{-\frac{2\pi i x\cdot\xi}{p}},\tag{9}$$

where the dot product $x\cdot\xi$ is defined as $x_1\xi_1+...+x_p\xi_p$. Recall the following Plancherel identity,

$$\sum_{\xi\in\mathbb{F}_p^d}|\hat{f}(\xi)|^2=p^d\sum_{x\in\mathbb{F}_p^d}|f(x)|^2.$$

Specially for the subset of $E\subset\mathbb{F}_p^d$, we have

$$\sum_{\xi\in\mathbb{F}_p^d}|\widehat{E}(\xi)|^2=p^d\#E.\tag{10}$$

For more details on discrete Fourier analysis, see Stein and Shakarchi [11].

We show some easy facts about the random model $\Omega(\mathbb{F}_p^d,\delta)$ in the following. Let $\xi\neq0$, then the expectation of $\widehat{E}(\xi)$ is

$$\mathbb{E}(\widehat{E}(\xi))=\delta\sum_{x\in\mathbb{F}_p^d}e^{-\frac{2\pi i x\cdot\xi}{p}}=0.$$

Since

$$\begin{aligned}|\widehat{E}(\xi)|^2&=\sum_{x,y\in\mathbb{F}_p^d}E(x)E(y)e^{-\frac{2\pi i(x-y)\cdot\xi}{p}}\\&=\sum_{x\in\mathbb{F}_p^d}E(x)+\sum_{x\neq y\in\mathbb{F}_p^d}E(x)E(y)e^{-\frac{2\pi i(x-y)\cdot\xi}{p}},\end{aligned}$$

we have

$$\begin{aligned}\mathbb{E}\left(|\widehat{E}(\xi)|^2\right)&=\delta p^d+\delta^2\sum_{x\neq y\in\mathbb{F}_p^d}e^{-\frac{2\pi i(x-y)\cdot\xi}{p}}\\&=p^d\delta\left(1-\delta\right).\end{aligned}$$

We may read this identity as (for small δ)

$$|\widehat{E}(\xi)| = \Theta\left(\sqrt{p^d\delta}\right) = \Theta\left(\sqrt{\#E}\right).$$

3. Proof of Theorem 1.3

For the convenience of use, we formulate a special large deviations estimate in the following. For more background and details on large deviations estimate, see Alon and Spencer [1, Appendix A].

Lemma 3.1. *Let $\{X_j\}_{j=1}^N$ be a sequence independent random variables with $|X_j|\le 1$, $\mu_1 := \sum_{j=1}^N \mathbb{E}(X_i)$, and $\mu_2 := \sum_{j=1}^N \mathbb{E}(X_j^2)$. Then for any $\alpha>0$, $0<\lambda<1$,*

$$\mathbb{P}(\big|\sum_{j=1}^N X_j\big| \geq \alpha) \leq e^{-\lambda\alpha+\lambda^2\mu_2}(e^{\lambda\mu_1}+e^{-\lambda\mu_1}). \tag{11}$$

Proof. Applying Markov's inequality to the random variable $e^{\lambda\sum_{j=1}^N X_j}$. This gives

$$\begin{aligned}\mathbb{P}(\sum_{j=1}^N X_j \geq \alpha) &= \mathbb{P}(e^{\lambda\sum_{j=1}^N X_j} > e^{\lambda\alpha}) \\ &\leq e^{-\lambda\alpha}\mathbb{E}(e^{\lambda\sum_{j=1}^N X_j}) \\ &= e^{\lambda\alpha}\prod_{j=1}^N \mathbb{E}(e^{\lambda X_j}),\end{aligned} \tag{12}$$

the last equality holds since $\{X_j\}_j$ is a sequence independent random variables.

For any $|x|\le 1$ we have

$$e^x \leq 1+x+x^2.$$

Since $|\lambda X_j|\le 1$, we have

$$e^{\lambda X_j} \leq 1+\lambda X_j+\lambda^2 X_j^2,$$

and hence

$$\begin{aligned}\mathbb{E}(e^{\lambda X_j}) &\leq 1+\mathbb{E}(\lambda X_i)+\mathbb{E}(\lambda^2 X_j^2) \\ &\leq e^{\mathbb{E}(\lambda X_i)+\mathbb{E}(\lambda^2 X_j^2)}.\end{aligned}$$

Combining this with (12), we have

$$\mathbb{P}(\sum_{j=1}^N X_j \geq \alpha) \leq e^{-\lambda\alpha+\lambda\mu_1+\lambda^2\mu_2}.$$

Applying the similar way to the above for $\mathbb{P}(\sum_{j=1}^N X_j \geq -\alpha)$, we obtain

$$\mathbb{P}(-\sum_{j=1}^N X_j \geq \alpha) \leq e^{-\lambda\alpha - \lambda\mu_1 + \lambda^2\mu_2}.$$

Thus we finish the proof. □

The following two easy identities are also useful for us.

$$\begin{aligned} \sum_{x\in\mathbb{F}_p^d} \cos\frac{2\pi x\cdot\xi}{p} &= \mathrm{Re}\left(\sum_{x\in\mathbb{F}_p^d} e^{-\frac{2\pi i x\cdot\xi}{p}}\right) = 0 \\ \sum_{x\in\mathbb{F}_p^d} \cos^2\frac{2\pi x\cdot\xi}{p} &= \sum_{x\in\mathbb{F}_p^d} \frac{1+\cos\frac{4\pi x\cdot\xi}{p}}{2} = \frac{1}{2}p^d \end{aligned} \tag{13}$$

Proof of Theorem 1.3. Let $\xi\neq 0$ and $E\in\Omega(\mathbb{F}_p^d,\delta)$. Let

$$\widehat{E}(\xi) = \sum_{x\in\mathbb{F}_p^d} E(x) e^{-\frac{2\pi i x\cdot\xi}{p}} = \mathcal{R} + i\mathcal{I}$$

where $\mathcal{R}$ and is the real part of $\widehat{E}(\xi)$, and $\mathcal{I}$ is the imagine part of $\widehat{E}(\xi)$. First we provide the estimate to the real part $\mathcal{R}$. By the Euler identity, we have

$$\mathcal{R} = \sum_{x\in\mathbb{F}_p^d} E(x)\cos(\frac{2\pi x\cdot\xi}{p}).$$

Note that

$$E(x)\cos\left(\frac{2\pi x\cdot\xi}{p}\right), \quad x\in\mathbb{F}_p^d$$

is a sequence of independent random variables. Furthermore, applying the identities (13), we have

$$\mu_1 = 0 \quad \text{and} \quad \mu_2 = \frac{1}{2}p^d\delta. \tag{14}$$

Here μ_1, μ_2 are defined as the same way as in Lemma 3.1. Let

$$\alpha := \sqrt{2(1+\varepsilon)p^d\delta\log p^d}, \quad \text{and} \quad \lambda := \frac{\alpha}{p^d\delta}. \tag{15}$$

Note that $0<\lambda<1$ for large p. Applying Lemma 3.1, we have

$$\begin{aligned} \mathbb{P}(|\mathcal{R}|\geq\alpha) &\leq 2e^{-\lambda\alpha+\lambda^2\mu_2} \\ &= 2e^{-\frac{\alpha^2}{2p^d\delta}} = \frac{2}{p^{d(1+\varepsilon)}}. \end{aligned} \tag{16}$$

Now we turn to the imagine part $\mathcal{I}$. Applying the similar argument to the real part $\mathcal{R}$, note that the identities (13) also hold if we take sin instead of cos, we obtain

$$\mathbb{P}(|\mathcal{I}|\geq\alpha)\leq\frac{2}{p^{d(1+\varepsilon)}}.$$

Combining this with the estimate (16), we obtain

$$\mathbb{P}(|\widehat{E}(\xi)|\geq\sqrt{2}\alpha)\leq\mathbb{P}(|\mathcal{R}|\geq\alpha)+\mathbb{P}(|\mathcal{I}|\geq\alpha)\leq\frac{4}{p^{d(1+\varepsilon)}} \tag{17}$$

Observe that the above argument works to any non-zero vector ξ. Therefore, we obtain

$$\mathbb{P}(\exists\,\xi\neq 0,\ \text{s.t}\ |\widehat{E}(\xi)|\geq\sqrt{2}\alpha)\leq\frac{4}{p^{d\varepsilon}}. \tag{18}$$

Recall the value of α in (15),

$$\alpha=\sqrt{2(1+\varepsilon)p^d\delta\log p^d},$$

this completes the proof. □

Remark 3.2. Let $E\subset\mathbb{F}_p^d$ with $\#E=p^s$. By Mockenhaupt and Tao [10, p.47], we define the Fourier transform of E at ξ as

$$\widehat{E}(\xi):=\frac{1}{\#E}\sum_{x\in\mathbb{F}_p^d}E(x)e^{-\frac{2\pi i x\cdot\xi}{p}}.$$

Then the estimate (3) in the definition of Salem sets becomes (for $\xi\neq 0$)

$$|\widehat{E}(\xi)|=O(p^{-\frac{s}{2}}).$$

We note that this form is the 'same' as the definition of Salem sets in Euclidean spaces, see [9, Chapter 3].

Acknowledgements. I am grateful for being supported by the Vilho, Yrjö, and Kalle Väisälä foundation.

References

1. Alon, N. and Spencer, J., *The probabilistic method.* New York: WileyInterscience, 2000.
2. Bluhm, C., Random recursive construction of Salem sets. *Ark. Mat.* **34** (1996), 51–63;
3. Babai, L., Fourier Transforms and Equations over Finite Abelian Groups, An introduction to the method of trigonometric sums. http://people.cs.uchicago.edu/~laci/reu02/fourier.pdf

4. Hayes, T., A Large-Deviation Inequality for Vector-valued Martingales. (see https://www.cs.unm.edu/~hayes/papers/VectorAzuma/VectorAzuma20050726.pdf)
5. Iosevich, A., Morgan, H., and Pakianathan, J., On directions determined by subsets of vector spaces over finite fields, *Integers* **11** (2011), 815–825.
6. Iosevich, A. and Rudnev, M., Erdös distance problem in vector spaces over finite fields, *Trans. Am. Math. Soc.* **359** (2007), 6127–6142.
7. Koh, D. and Shen, Chun-Yen, Additive energy and the Falconer distance problem in finite fields, *Integers* **13** (2013), 1–10.
8. Mattila, P., Hausdorff dimension, projections, and the Fourier transform, *Publ. Mat.* **48** (2004), 3–48.
9. Mattila, P., *Fourier analysis and Hausdorff dimension*, Cambridge Studies in Advanced Mathematics, vol. **150**, Cambridge University Press, 2015.
10. Mockenhaupt, G. and Tao, T., Restriction and Kakeya phenomena for finite fields, *Duke Math. J.* **121** (2004), 35–74.
11. Stein, E. and Shakarchi, R., *Fourier Analysis: An Introduction.* Princeton and Oxford: Princeton UP, 2003. Print. Princeton Lectures in Analysis.
12. Tao, T. and Vu, V., Additive Combinatorics, Cambridge University Press.

Changhao Chen
School of Mathematics and Statistics
The University of New South Wales
Sydney
NSW AU-2052
Australia
changhao.chenm@gmail.com

Received January 29, 2017
in revised form May 10, 2017