Analysis of the Statistical Restricted Isometry Property for Deterministic Sensing Matrices Using Stein's Method

Lu Gan*, Cong Ling†, Thong T. Do‡ and Trac D. Tran‡
*School of Engineering and Design
Brunel University, UK
Email: lu.gan@brunel.ac.uk
† Department of Electrical and Electronic Engineering
Imperial College London, UK
Email: c.ling@imperial.ac.uk
‡ Department of Electrical and Computer Engineering
The Johns Hopkins University, USA
Email: thongdo@jhu.edu; trac@jhu.edu

***Abstract*— Statistical restricted isometry property (STRIP) was recently formulated by Calderbank *et al.* to analyze the performance of deterministic sampling matrices for compressed sensing. In this paper, we study the STRIP by taking advantage of concentration inequalities using Stein's method. In particular, we derive the STRIP performance bound in terms of the mutual coherence of the sampling matrix and the sparsity level of the input signal. Based on such connections, we show that a large class of deterministic matrices can satisfy the STRIP with high probability provided that they can nearly meet the Welch bound. Such matrices include many classical spreading codes in code-division multiple access (CDMA) such as the Kasami code, the Gold code, the Frank-Zadoff-Chu code, and the more recent partial FFT matrices based on difference sets etc. Simulation results show that these deterministic sensing matrices can offer reconstruction performance similar to that of random operators.**

I. Introduction

Over the past few years, there have been increased interests in the study of compressed sensing (CS) [1]–[4], a new framework for simultaneous sampling and compression of signals. The CS theory is based on the assumption that a signal is compressible or sparse. Consider a discrete-time, length-N signal $\mathbf{x}$ that can be represented (or approximated) by only K ($K \ll N$) coefficients. CS is accomplished by computing a measurement vector $\mathbf{y}$ through the following linear transformation [1], [2]:

$$\mathbf{y} = \Phi\mathbf{x}, \tag{1}$$

where $\mathbf{y}$ represents an $M \times 1$ sampled vector and Φ is an $M \times N$ measurement matrix. It was proved in [1], [2] that under certain conditions, $\mathbf{x}$ can be well recovered from only $M = \mathcal{O}(K \log(N/K))$ measurements through *non-linear* optimization.

One fundamental question in CS is how to construct Φ so that the number of measurements is (near) minimal. Candès and Tao [5] formulated a sufficient condition called the *restricted isometry property* (RIP). Simply speaking, the RIP implies that the l_2 norms of *all* K-sparse vectors are preserved under the linear transform of Φ. However, it is often computationally infeasible to verify the RIP for a given Φ. Besides, most known families of sampling operators are random matrices, including the random i.i.d Gaussian or Bernoulli matrices and partial Fourier or Walsh-Hadamard matrices [6]. But these operators usually require huge memory for storage and some of them are even difficult to implement in the hardware. Deterministic matrices with the RIP have been proposed in [7], [8]. However, their guaranteed RIP performance is not comparable to that of random operators.

Recently, the statistical versions of the RIP have been developed in [9], [10], where a deterministic Φ is required to preserve the l_2 norm for *any* K-sparse input vector $\mathbf{x}$ within a small fraction. Using elementary proof, the performance bound of statistical TRIP (STRIP) has been developed and the construction of a large class of deterministic matrices was also proposed [10]. But it was noted in [10] that the empirical performance of some deterministic matrices are much better than the derived theoretical bounds.

In this paper, we investigate the STRIP of deterministic matrices by exploiting Stein's method. In particular, we show that a normalized $M \times N$ deterministic sensing operator has the STRIP with high probability if its mutual coherence is on the order of $\mathcal{O}(1/\sqrt{M})$. Examples of these deterministic matrices include the Kasami and Gold codes [11], the Frank-Zadoff-Chu (FZC) code [12] and the deterministic partial FFT matrices based on the difference sets [13]. Empirical results show that these deterministic matrices can provide similar reconstruction performance to that of the random ones despite their low storage requirement.

The rest of this paper is organized as follows. In Section II, we briefly review the concepts of the RIP and the STRIP and introduce Stein's method for concentration inequalities. Section III presents our theoretical analysis of the STRIP performance bound. Several examples deterministic operators were also provided. Simulation results are given in Section IV, followed by conclusions in Section V.

II. Review

A. The RIP and the STRIP

The RIP [5] given below is an important sufficient condition to provide theoretical guarantee for exact sparse signal recovery.

Definition 1 (RIP [5]): Let Ω denote the set of all length-N vectors with K non-zero coefficients. An $M \times N$ measurement matrix Φ has the restricted isometry property (RIP) with parameters (K, δ) for $\delta \in (0,1)$ if it satisfies [5]

$$(1-\delta)\|\mathbf{x}\|^2 \leq \|\Phi\mathbf{x}\|^2 \leq (1+\delta)\|\mathbf{x}\|^2, \ \textbf{for all } \mathbf{x} \in \Omega.$$

Note that the RIP implies that for *all* $N \times K$ sub-matrices of Φ, the eigenvalues of their Gram matrices lie in the interval of $[1-\delta, 1+\delta]$. This is a very restrictive condition and the currently known measurement matrices satisfying the RIP with (near) optimal number of measurements fall into two categories [6]: (*i*) Random matrices with i.i.d. sub-Gaussian variables, e.g., normalized i.i.d. Gaussian or Bernoulli matrices; (*ii*) Random partial bounded orthogonal matrices in which the sensing operators are obtained by choosing M rows uniformly at random from a normalized $N \times N$ Fourier or Walsh-Hadamard transform matrices.

In some applications, deterministic sensing operators are highly desirable due to storage limitations. However, the construction of deterministic RIP matrices still remains as a challenging task. As an alternative, statistical versions of the restricted isometry property were proposed for deterministic sensing operators [9], [10]. In this paper, we follow the STRIP formulation given by Calderbank *et al*. in [10].

Let Φ be a normalized $M \times N$ deterministic matrix and denote ϕ_i $(1 \leq i \leq N)$ as its i-th column. In [10], the input signal $\mathbf{x}$ is modelled as a K-sparse random vector with non-zero coefficients $x_1, x_2, \cdots, x_K$ and the positions of these coefficients are chosen uniformly at random. Under such a model, the measurement vector can be expressed as [10]

$$\mathbf{y}(\pi) = \Phi\mathbf{x} = \sum_{i=1}^{N} \phi_{\pi(i)} x_i, \tag{2}$$

where π is drawn from the uniform distribution over the set of all permutations of $\{1, \cdots, N\}$. Φ is said to have the STRIP provided that the following inequality holds with high probability:

$$\left| \|\mathbf{y}(\pi)\|^2 - \|\mathbf{x}\|^2 \right| \leq \delta \|\mathbf{x}\|^2. \tag{3}$$

The STRIP defined above is a much weaker condition than the RIP. Through elementary proof, the STRIP performance bound has been derived in [10] for a large class of deterministic matrices, as summarized by the following theorem:

Theorem 1 (STRIP [10]): Let Φ be a deterministic $M \times N$ sensing matrix satisfying the following properties:

- The columns of Φ form a group under point-wise multiplication;
- The rows of Φ are orthogonal and all row sums are equal to zero, i.e., $\sum_{i=1}^{N} \phi_i = 0$.

Let $\mathbf{x}$ be a K-sparse signal where the positions of the K non-zero entries are equiprobable. Then, for $\frac{K-1}{N-1} < \delta < 1$, the following inequality holds:

$$\mathbf{P}\left(\left| \|\Phi\mathbf{x}\|^2 - \|\mathbf{x}\|^2 \right| < \delta\|\mathbf{x}\|^2 \right) \geq 1 - \frac{\left(\frac{2K}{M} + \frac{2K+7}{N-3}\right)}{\left(\delta - \frac{K-1}{N-1}\right)^2}. \tag{4}$$

Note that the restrictions in Theorem 1 are rather weak. For example, almost all linear codes [11] and partial FFT matrices (excluding the first row) meet the conditions in Theorem 1. However, not all of them have good performance for compressed sensing. Besides, (4) implies that

$$\mathbf{P}\left(\left| \|\Phi\mathbf{x}\|^2 - \|\mathbf{x}\|^2 \right| > \delta\|\mathbf{x}\|^2 \right) < \mathcal{O}\left(\frac{K}{M\delta^2}\right). \tag{5}$$

As pointed out by Remark 2 in [10], empirical results suggested that the actual bound of some deterministic matrices decays much faster with M than that predicted by (5).

The above analysis suggests that in order to construct deterministic operators satisfying the STRIP with high probability, stronger restrictions need to be imposed on Φ. It is also desirable to derive a better performance bound than (5). We will address these issues in Section III by exploiting Stein's method for concentration inequality.

B. Concentration inequality based on Stein's method

Stein's method is a powerful tool to obtain bounds on the distance between two probability distributions with respect to a probability metric. One significant idea in Stein's method is the *exchangeable pair*. Two random variables Z and Z' are said to form an exchangeable pair if the probability density function of (Z, Z') is equal to that of (Z', Z). In [14], Stein's method was used to derive measure concentration, as stated in the following theorem:

Theorem 2 ([14]): Suppose that (Z, Z') is an exchangeable pair of random variables. Let $F(Z, Z')$ be an antisymmetric function, i.e., $F(Z, Z') = -F(Z', Z)$ with $E(F(Z, Z')|Z) = f(Z)$. Define $\Delta(Z)$ as

$$\Delta(Z) = \frac{1}{2}\mathbb{E}\left(|(f(Z) - f(Z'))F(Z, Z')| \, \middle| Z \right).$$

Then, $\mathbb{E}(f(Z)) = 0$ and if there exist non-negative constants a_0 and a_1 such that $\Delta(Z) \leq a_0 + a_1 f(Z)$ almost surely, then for any $t \geq 0$, we have

$$\mathbf{P}(f(Z) \geq t) \leq \exp\left(-\frac{t^2}{2a_0 + 2a_1 t}\right) \tag{6}$$

and

$$\mathbf{P}(f(Z) \leq -t) \leq \exp\left(-\frac{t^2}{2a_0}\right). \tag{7}$$

As the STRIP is based on the permutation operator π, the next section will derive new probability performance bound of (3) by applying Theorem 2 to concentration inequality of random permutation (i.e., when $Z = \pi$).

III. Analysis of the STRIP of Deterministic Matrices

For an $M \times N$ deterministic sensing matrix Φ whose row sums are equal to zero, we derived the performance bound of its STRIP in terms of its mutual coherence, as presented in Theorem 3 below:

Theorem 3: Let $\mathbf{x}$ be a length-N, K-sparse signal with non-zero coefficients $x_1, x_2, x_3, \cdots x_K$. Assume that $\mathbf{x}$ has zero-mean (i.e., $\sum_{i=1}^{K} x_i = 0$) and the positions of the K non-zero entries are equiprobable. Let $\Phi = \left[\phi_1, \phi_2, \cdots, \phi_N\right]$ be an $M \times N$ deterministic normalized sensing matrix, where each column has unit norm, i.e., $\|\phi_i\| = 1$ $(1 \leq i \leq N)$. Assume further that all row sums of Φ are equal to zero (i.e. $\sum_{i=1}^{N} \phi_i = \mathbf{0}$), and let μ denote its mutual coherence, i.e.,

$$\mu = \max_{i \neq j} | < \phi_i, \phi_j > |. \tag{8}$$

Then,

$$\mathbb{E}\left(\|\Phi\mathbf{x}\|^2\right) = \left(1 + \frac{1}{N-1}\right)|\mathbf{x}\|^2. \tag{9}$$

Besides, the following inequality holds for $\delta > \frac{1}{N-1}$:

$$\begin{aligned} &\mathbf{P}\left(\left|\|\Phi\mathbf{x}\|^2 - \|\mathbf{x}\|^2\right| < \delta\|x\|^2\right) \\ \geq &1 - 2\exp\left(-\frac{\left(\delta - \frac{1}{N-1}\right)^2}{16\mu^2 K}\right). \end{aligned} \tag{10}$$

Remark: The restriction that $\mathbf{x}$ has zero mean is needed only to simplify the derivation so that $\mathbb{E}(\|\Phi\mathbf{x}\|^2)$ is fixed. In practice, for $\mathbf{x}$ with non-zero mean, we can add an all-ones row vector $\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}$ to measure the DC component of the signal.

The proof of the above theorem is outlined in the Appendix. Note that for μ in (8), its lower-bound is given by [15]

$$\mu \geq \sqrt{\frac{N-M}{(N-1)M}}. \tag{11}$$

When the equality holds, Φ is said to meet the *Welch bound* [15]. Next, if we further requires

$$\mu \leq \frac{\alpha_1}{\sqrt{M}}, \quad \text{for some constant } \alpha_1, \tag{12}$$

we can derive from (10) that

$$\mathbf{P}\left(\left|\|\Phi\mathbf{x}\|^2 - \|\mathbf{x}\|^2\right| > \delta\|x\|^2\right) < 2\exp\left(-\mathcal{O}\left(\frac{M\delta^2}{K}\right)\right). \tag{13}$$

One can see that (13) decays *exponentially* with $\frac{M}{K}$, while the bound in (5) only decays linearly with $\frac{M}{K}$. Thus, (13) approaches 0 much faster with the increase of $\frac{M}{K}$. The improvement mainly comes from two facts (*i*) the restriction of mutual coherence in (12) and (*ii*) the use of exponential concentration inequalities in (6) and (7).

From (13), we can easily arrive at the following proposition:

Proposition 1: Suppose that $\mathbf{x}$ and Φ follow the same definitions in Theorem 3 with Φ's mutual coherence μ satisfying (12). Then, we have

$$\mathbf{P}\left(\left|\|\Phi\mathbf{x}\|^2 - \|\mathbf{x}\|^2\right| \leq \delta\|x\|^2\right) > 1 - \frac{1}{N},$$

TABLE I

Examples of Deterministic Matrices satisfying the STRIP with High Probability

$M \times N$ Sampling Operator Φ	Relations of M and N	Mutual coherence μ
Kasami Codes [11]	$N = (M+1)^{\frac{3}{2}}$ $M = 2^{2p} - 1, p \in \mathbb{Z}^+$	$\frac{1+\sqrt{M+1}}{M}$
Gold/Gold-like Codes [11]	$M = 2^p - 1, p \in \mathbb{Z}^+$ $N = (M+1)^2$	$\frac{1+2\sqrt{M+1}}{M}$ or $\frac{1+\sqrt{2(M+1)}}{M}$
FZC Codes [12]	$N = M^2$, M prime	$\frac{1}{\sqrt{M}}$
Partial FFT Codes based on the difference sets [13]	$N = 2^p M + 1$ $p \in \mathbb{Z}^+$ and N prime	$\sqrt{\frac{N-M}{(N-1)M}}$

if the sparsity level K satisfies

$$K \leq c_0(\delta)\frac{M}{\log N}, \tag{14}$$

in which $c_0(\delta) = \frac{\delta^2}{8\alpha_0}$.

Recall that for i.i.d Gaussian and Bernoulli matrices, the RIP holds with high probability when

$$K \leq c_1(\delta)\frac{M}{\log(N/M)}, \tag{15}$$

where $c_1(\delta)$ is a function depending only on δ. One can observe that (14) takes a similar form to (15). Note that there are no existing solutions of deterministic $M \times N$ sensing operators that could achieve the RIP bound of (15). But Proposition 1 suggests that when K is on the similar order, the STRIP can be satisfied with high probability.

In fact, there are many families of deterministic matrices that can achieve the bound in (14). Table I lists some of these operators along with their mutual coherence and size limitations. All these operators have zero row sums. Note that the construction of Φ nearly meeting the Welch bound has been extensively studied in code-division multiple access (CDMA), such as the Kasami, Gold/Gold-like codes [11] (these binary operators have hardware friendly implementation), and the FZC codes [12]. It is worth mentioning that many codes in [10] also satisfy the condition (12). For example, the dual BCH codes are closely related to Gold/Gold-like codes, while chirp codes have the same mutual coherence as FZC codes. The last class in Table I is the deterministic partial FFTs (PFFTs) that can exactly meet the Welch bound [13]. It includes the family of equiangular tight frames [16] as a special case. The row indexes of these PFFT are selected according to the difference set in combinatorial design theory.

IV. Simulation Results

Extensive simulations have been carried out to compare the reconstruction performance of different random and deterministic sampling operators. For illustration purposes, we only present here some results for 63×512 and 169×677 sampling operators as follows:

1) 63×512 *binary-coefficient matrices*: These sampling operators include the random i.i.d. Bernoulli $\{+1, -1\}$

matrix, the random partial Walsh-Hadamard matrix and the deterministic Kasami code;

2) 169×677 *complex-coefficient sampling operators*: Sampling operators under comparison are the random complex i.i.d. Gaussian matrix, the random PFFT matrix and the deterministic PFFT based on the quartic difference sets (PFFT-DS) [13].

The reconstruction algorithms are based on the orthogonal matching pursuit (OMP) [18] and the subspace pursuit (SP) [19], respectively. In particular, for the OMP, the non-zero coefficients x_i $(i = 1, \cdots K)$ of the input signal have the Gaussian distribution; while for the SP, x_i $(i = 1, \cdots K)$ follow the Bernoulli distributions.

Fig. 1 and Fig. 2 depict the empirical frequencies of exact reconstruction for 63×512 binary-coefficient matrices and 169×677 complex-coefficient matrices, respectively. In these simulations, 1000 trials were run for each sparsity level K and the positions of non-zero elements are selected uniformly at random. Besides, we assume that the exact reconstruction is achieved if the signal to noise ratio is greater than 50 dB.

From these figures, one can observe that the performance of deterministic matrices are quite similar to those of the random matrices. These results imply the promising application of deterministic compressed sensing. They also suggest that construction of deterministic Φ satisfying the STRIP with high probability may be an interesting research direction.

V. Conclusions and Future Work

This paper investigates the STRIP for deterministic sampling operators in compressed sensing. Specifically, Theorem 3 derives the concentration inequalities of the STRIP performance bound by exploiting Stein's method. Based on Theorem 3, we show that there exists a large class of deterministic matrices satisfying the STRIP with high probability. Examples of such matrices include the FZC codes, Gold codes, Kasami codes, partial deterministic FFT matrices based on the difference sets etc. Experimental results show that these deterministic operators compare favorably with existing random operators in reconstruction performance.

There are many intriguing questions that future work should consider. First, the noise resilience of these deterministic operators needs to be analyzed and evaluated. Secondly, as indicated in Table I, existing deterministic matrices nearly meeting the Welch bound have certain restrictions on N and M. The construction of deterministic sensing operators with arbitrary sizes will be an interesting question. Thirdly, new reconstruction algorithms for deterministic compressed sensing need to be developed. Furthermore, it is interesting to consider the statistical model of the input signal in STRIP formulation.

Appendix
Outline of Proof of Theorem 3

Here, we only provide an outline for the proof of Theorem 3. Details will be presented in the journal version of this paper. For the random permutation operator π in (2), we can define its exchangeable pair as $\pi' = \pi \circ (I, J)$ [14], where I and J are chosen uniformly and independently at random from $\{1, \cdots, N\}$. (I, J) denotes the transposition of I and J, i.e., $\pi'(I) = \pi(J)$, $\pi'(J) = \pi(I)$ and $\pi'(i) = \pi(i)$ for $i \neq I, J$.

Define the antisymmetric function $\mathbf{F}(\pi, \pi')$ as follows:

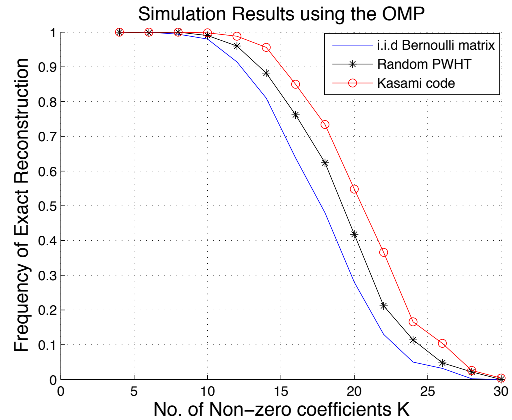
$$\begin{aligned} \mathbf{F}(\pi, \pi') &= \frac{N^2}{4(N-1)} \left(\|\mathbf{y}(\pi)\|^2 - \|\mathbf{y}(\pi')\|^2 \right) \\ &= \frac{N^2}{4(N-1)} \left(2\mathbf{y}^H(\pi)\mathbf{r}(\pi) - \|\mathbf{r}(\pi)\|^2 \right) \end{aligned} \quad (16)$$

where $\mathbf{y}(\pi)$ is given by (2) and $\mathbf{r}(\pi)$ is the difference between $\mathbf{y}(\pi)$ and $\mathbf{y}(\pi')$ that can be expressed as

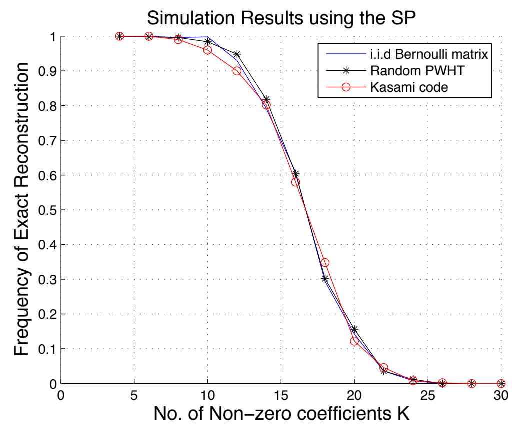
$$\mathbf{r}(\pi) = \phi_{\pi(I)} x_I + \phi_{\pi(J)} x_J - \phi_{\pi(I)} x_J - \phi_{\pi(J)} x_I. \quad (17)$$

Next, according to Theorem 2, we need to find

$$\begin{aligned} f(\pi) &= \mathbb{E}\left(F(\pi, \pi') | \pi \right) \\ &= \frac{N^2}{4(N-1)} \mathbb{E}_{I,J} \left(2\mathbf{y}^H(\pi)\mathbf{r}(\pi) - \|\mathbf{r}(\pi)\|^2 \right). \end{aligned}$$

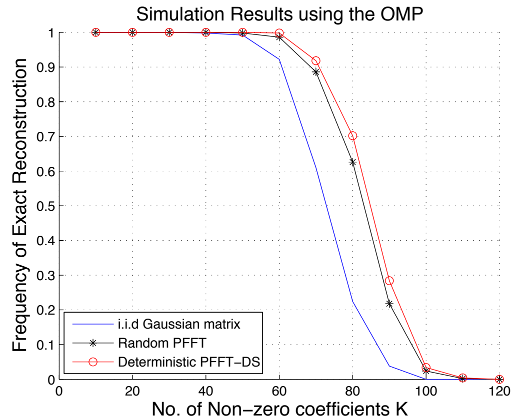


(a)

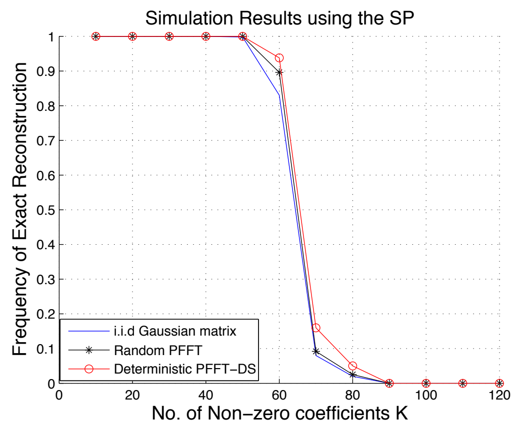


(b)

Fig. 1. Simulation results for different 63×512 binary sampling operators, including the i.i.d Bernoulli matrix, the random partial Walsh-Hadamard transform (PWHT) and the Kasami code [11]. (a) Results using the orthogonal matching pursuit (OMP) [18] algorithm with Gaussian sparse signal. (b) Results using the subspace pursuit (SP) [19] algorithm with $(1, -1)$ binary sparse signal.



(a)



(b)

Fig. 2. Simulation results for different 169×677 complex sampling operators, including the complex i.i.d Gaussian matrix, the random partial FFT (PFFT) matrix and the deterministic PFFT based on the difference sets (PFFT-DS) [13]. (a) Results using the orthogonal matching pursuit (OMP) [18] algorithm with Gaussian sparse signal. (b) Results using the subspace pursuit (SP) [19] algorithm with $(1, -1)$ binary sparse signal.

From the expression of $\mathbf{r}(\pi)$ in (17), by summing over all choices of $1 \le I, J \le N$, we have

$$\mathbb{E}_{I,J}\left(2\mathbf{y}^H(\pi)\mathbf{r}(\pi)\right) = \frac{4\|\mathbf{y}(\pi)\|^2}{N}, \tag{18}$$

$$\mathbb{E}_{I,J}\left(\|\mathbf{r}(\pi)\|^2\right) = \frac{4\|\mathbf{x}\|^2}{N} + \frac{4\|\mathbf{y}(\pi)\|^2}{N^2}, \tag{19}$$

where we have used the assumptions that $\sum_{i=1}^N \phi_i = \mathbf{0}$ and $\sum_{i=1}^N x_i = 0$. Substituting the above results into $f(\pi)$ yields

$$f(\pi) = \|\mathbf{y}(\pi)\|^2 - \|\mathbf{x}\|^2 - \frac{\|\mathbf{x}\|^2}{N-1}. \tag{20}$$

By Theorem 2, we know that $\mathbb{E}(f(\pi)) = 0$, which implies that (9) holds.

To get the concentration inequality in (10), we thus need to bound

$$\begin{aligned}\Delta(\pi) &= \frac{1}{2}\mathbb{E}\left(\left|\left(f(\pi) - f(\pi')\right)F(\pi,\pi')\right| \,|\pi\right) \\ &= \frac{N^2}{8(N-1)}\mathbb{E}\left(\left(2\mathbf{y}^H(\pi)\mathbf{r}(\pi) - \|\mathbf{r}(\pi)\|^2\right)^2 |\pi\right).\end{aligned} \tag{21}$$

Using Cauchy-Schwartz inequality and (8), we can derive that

$$\left(2\mathbf{y}^H(\pi)\mathbf{r}(\pi) - \|\mathbf{r}(\pi)\|^2\right)^2 \le 16\mu^2(K-2)\|\mathbf{x}\|^2\left(x_I - x_J\right)^2.$$

As a result, $\Delta(\pi)$ can be bounded as follows

$$\begin{aligned}\Delta(\pi) &= \frac{N^2}{8(N-1)}\mathbb{E}_{I,J}\left(\left(2\mathbf{y}^H\mathbf{r} - \|\mathbf{r}\|^2\right)^2 |\pi\right) \\ &\le \frac{N^2}{8(N-1)}16\mu^2\|\mathbf{x}\|^2(K-2)E\left[\left(x_I - x_J\right)^2 |\pi\right] \\ &\le \frac{4\mu^2 NK}{N-1}\|\mathbf{x}\|^4 \\ &\le 8\mu^2 K\|\mathbf{x}\|^4.\end{aligned}$$

Finally, based on the above result, (10) can be easily obtained from (6) and (7) in Theorem 2.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, July 2006.

[2] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[3] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, 2006.

[4] Compressive sensing resources. [Online]. Available: http://www-dsp.rice.edu/cs

[5] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, 2005.

[6] M. Rudelson and R. Vershynin, "On sparse reconstruction from fourier and gaussian measurements," *Communications on Pure and Applied Mathematics*, vol. 61, pp. 1025–1045, 2008.

[7] R. A. Devore, "Deterministic construction of compressed sensing matrices," *Journal of Complexity*, vol. 23, pp. 918–925, 2007.

[8] M. A. Iwen, "Simple deterministically constructible rip matrices with sublinear fourier sampling requirements," in *in Proc. of Proceedings of CISS 2008*, 2008.

[9] S. Guverich and R. Hadani, "The statistical restricted isometry property and the wigner semicircle distribution of incoherent dictionaries," *Annals of Applied Probability*, 2009, submitted.

[10] R. Calderbank, S. Howard, and S. Jafarpour. (2009) Construction of a large of deterministic sensing matrices that satisfy a statistical isometry property. [Online]. Available: http://www.dsp.ece.rice.edu/files/cs/strip-final.pdf

[11] D. V. Sarwate, "Meeting the Welch bound with equality," in *Sequences and Their Applications*. Springer, 1998, pp. 79–102.

[12] ——, "Bounds on crosscorrelation and autocorrelation of sequences," *IEEE Trans. Inform. Theory*, vol. 25, pp. 720–724, Nov. 1979.

[13] P. Xia, S. Zhou, and G. B. Giannakis, "Achieving the welch bound with difference sets," *IEEE Trans. Inform. Theory*, vol. 138, pp. 1900–1907, 2006.

[14] S. Chatterjee, "Stein's method for concentration inequalities," *Probability Theory Related Fields*, vol. 138, pp. 305–321, 2007.

[15] L. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. 20, pp. 397–399, May 1974.

[16] H. Konig, "Cubature formulas on spheres," in *Advances in Multivariate Approximation*. Wiley, New York, 1979, pp. 201–211.

[17] L. Applebaum, S. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," in *in Applied and Computational Harmonic Analysis*, August 2009.

[18] J. A. Tropp, "Greedy is good: Algorithm results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2231–2242, 2004.

[19] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing: Closing the gap between performance and complexity. [Online]. Available: http://arxiv.org/abs/0803.0811