On the Conditioning of Random Block Subdictionaries

Waheed U. Bajwa,* Robert Calderbank, and Marco F. Duarte*

September 22, 2010

1   Introduction

The use of sparsity as a prior for efficient signal acquisition and processing has become prevalent in the last decade. In particular, compressed sensing (CS) [3, 4] uses sparsity to reduce the number of samples/measurements needed to acquire a signal. The performance of CS hinges on certain properties of the measurement matrix that maps the input signal to the obtained measurements. Early results in CS focused on deterministic conditions on the measurement matrix for the case of arbitrary (deterministic) signals. However, the conditions are either difficult to verify or provide highly pessimistic bounds on recovery performance.

Recent results by Tropp [11] have focused on the performance of measurement matrices for sparse signals under a probabilistic signal model where the support is drawn at random from a uniform distribution over all sparse supports of a given size. The result exploits the fact that the product of a matrix with a sparse signal is essentially controlled by the subdictionary corresponding to the signal support. In this case, recovery performance is tied to the coherence and spectral norm of the matrix through the calculation of the expected spectral norm of randomly selected subdictionaries of the matrix.

In this report, we adapt these results to applications where the signal exhibits block (or group) sparsity, rather than simple sparsity. That is, the block-sparse signal has support that is drawn uniformly at random from all block-sparse supports containing a given number of blocks. In block sparsity, the nonzero coefficients of the sparse signal are grouped together as intervals in the signal coefficient vector. The block sparsity model is popular for multiband signals, fusion frames, and signal ensemble applications [1, 2, 5, 6, 8]. Thus, our results focus on random block subdictionaries, where the columns selected from the measurement matrix are grouped together into intervals. The results in this paper can also have implications on the performance of block-sparse signal recovery under the random block-sparse support model.

This report is organized as follows. Section 2 sets up notation and presents our main theorem. Section 3 provides necessary lemmas, and Section 4 gives the proof of our main theorem.

2 Notation and Main Result

The measurement matrix (or dictionary) is $\Phi = [\Phi_1 \ \Phi_2 \ \ldots \ \Phi_p]$, where each block $\Phi_i = [\phi_{i,1} \ \ldots \ \phi_{i,m}]$ is an $n \times m$ matrix. The dictionary has coherence

$$\mu := \max_{(i,j) \neq (k,l)} |\langle \phi_{(i,j)}, \phi_{(k,l)} \rangle|$$

and each of its columns $\phi_{i,j}$ has unit norm. We define random variables $\delta_1, \ldots, \delta_p$ that are independent and identically distributed (i.i.d.) Bernoulli with parameter $\delta := k/p$, and form a block subdictionary $X = [\Phi_i : \delta_i = 1]$. The hollow Gram matrix for X is $A := X^H X - I$, where I denotes the identity matrix of appropriate size. To calculate the spectral norm of A, we use a masking matrix $R = S \otimes I_m$, where $S = \mathrm{diag}(\delta_1, \ldots, \delta_p)$ is a random matrix and $\otimes$ denotes the standard Kronecker product. We then have $\|A\|_2 = \|RGR\|_2$ with $G := \Phi^H \Phi - I$. We denote $\mathbb{E}_q X := [\mathbb{E}|X|^q]^{1/q}$. We also define the block coherence[1]

$$\mu_B = \max_{1 \leq i,j \leq p} \|\Phi_i^H \Phi_j - 1_{\{i=j\}} I\|_2.$$

Finally, we denote $T \subseteq \{1, \ \ldots \ , p\}$ to be a subset of the block index set, with $T^C := \{1, \ \ldots \ , p\} \setminus T$. Using this notation, the main result of this report can be stated as follows.

Theorem 1. *For $q = 2 \log(pm)$, we have the bound*

$$\mathbb{E}_q \|RGR\|_2 \leq 20 \mu_B \log(pm) + 9\sqrt{\delta \log(pm)(1 + (m-1)\mu)} \|\Phi\|_2 + \delta \|\Phi\|_2^2.$$

---

[1]We recently became aware of [5], which also introduces the term "block coherence" in the context of recovery of block-sparse signals, defined as $\max_{i \neq j} \|\Phi_i^H \Phi_j\|_2$.

The proof of the theorem uses many lemmas and tools, following the ideas of the proof in [11]. To begin, we denote the matrix G in block-diagonal fashion:

$$G = [G_1 \; G_2 \; \ldots \; G_p] = \begin{bmatrix} G_{1,1} & G_{1,2} & \ldots & G_{1,p} \\ G_{2,1} & G_{2,2} & \ldots & G_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ G_{p,1} & G_{p,2} & \ldots & G_{p,p} \end{bmatrix}$$

where $G_{i,j} = \Phi_i^H \Phi_j - 1_{\{i=j\}} I$ for $1 \leq i, j \leq p$. We then split $G = H + D$, where D contains the diagonal blocks $G_{i,i}$, and H contains only the non-diagonal blocks. We also define the following "norms" for block matrices:

- when we group only the columns, we define $\|G\|_{B,1} := \max_{1 \leq i \leq p} \|G_i\|_2$;

- when we group both columns and rows, we define $\|G\|_{B,2} := \max_{1 \leq i,j \leq p} \|G_{i,j}\|_2$.

Finally, we will make use of the following standard inequalities:

- Cauchy-Schwarz Inequality: $|\mathbb{E}(XY)|^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

- Hölder's Inequality: $\|fg\|_1 \leq \|f\|_p \|g\|_q$, $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$.

- Jensen's Inequality for a convex function f: $f(\mathbb{E}X) \leq \mathbb{E}f(X)$.

- Scalar Khintchine Inequality: Let $\{a_i\}$ be a finite sequence of complex numbers and $\{\epsilon_i\}$ be a Rademacher (random ± 1 binary) sequence. For each $q \geq 0$, we have

$$\mathbb{E}_q \left| \sum_i \epsilon_i a_i \right| \leq C_q \left(\sum_i |a_i|^2 \right)^{1/2},$$

where $C_q \leq 2^{1/4} \sqrt{q/e}$.

- Noncommutative Khintchine Inequality [12]: Let $\{X_i\}$ be a finite sequence of matrices of the same size and $\{\epsilon_i\}$ be a Rademacher sequence. For each $q \geq 2$,

$$\mathbb{E}_q \left\| \sum_j \epsilon_j X_j \right\|_{S_q} \leq W_q \max \left\{ \left\| \left(\sum_j X_j X_j^H \right)^{1/2} \right\|_{S_q}, \left\| \left(\sum_j X_j^H X_j \right)^{1/2} \right\|_{S_q} \right\},$$

where $\|X\|_{S_q} = \|\sigma(X)\|_q$ denotes the Schatten q-norm for a matrix (equal to the ℓ_q-norm of the vector containing its singular values) and $W_q \leq 2^{-1/4} \sqrt{\pi q/e}$.

3   Lemmata

We use the following lemmas in our proof of Theorem 1. The first two lemmas are used to prove later ones.

Lemma 1. *Let $X = [X_1 \; \ldots \; X_p]$ be a block matrix and D_X be its block diagonalization, i.e., a block-diagonal matrix $D_X = \text{diag}(X_i)_{i=1}^p$ containing the square matrices X_i in its diagonal, with all other elements being equal to zero. Then, we have*

$$\|D_X\|_2 \leq \|X\|_{B,1}.$$

Proof. For a vector a of appropriate length, we evaluate the ratio $\frac{\|D_X a\|_2^2}{\|a\|_2^2}$. We partition $a = [a_1 \; a_2 \; \ldots \; a_p]$ into pieces matching the number of columns of the blocks X_i, $1 \leq i \leq p$. Then, we will have

$$\frac{\|D_X a\|_2^2}{\|a\|_2^2} = \frac{\sum_{i=1}^p \|X_i a_i\|_2^2}{\sum_{i=1}^p \|a_i\|_2^2} \leq \frac{\sum_{i=1}^p \|X_i\|_2^2 \|a_i\|_2^2}{\sum_{i=1}^p \|a_i\|_2^2}$$

$$\leq \frac{\max_{1 \leq i \leq p} \|X_i\|_2^2 \sum_{i=1}^p \|a_i\|_2^2}{\sum_{i=1}^p \|a_i\|_2^2} \leq \max_{1 \leq i \leq p} \|X_i\|_2^2.$$

Thus, the spectral norm obeys

$$\|D_X\|_2 = \max_a \frac{\|D_X a\|_2}{\|a\|_2} \leq \max_{1 \leq i \leq p} \|X_i\|_2 = \|X\|_{B,1}.$$

$\square$

The next lemma is adapted from [9].

Lemma 2. *Let $X = [X_1 \; X_2 \; \ldots \; X_p]$ be a block matrix where each block X_i has m columns. For any $q \geq 2\log(pm)$, we have*

$$\mathbb{E}_q \left\| \sum_{i=1}^p \epsilon_i X_i X_i^H \right\|_2 \leq 1.5\sqrt{q}\|X\|_{B,1}\|X\|_2,$$

where $\{\epsilon_i\}$ is a Rademacher sequence.

Proof. We start by bounding the spectral norm by the Schatten q-norm:

$$E := \mathbb{E}_q \left\| \sum_{i=1}^p \epsilon_i X_i X_i^H \right\|_2 \leq \mathbb{E}_q \left\| \sum_{i=1}^p \epsilon_i X_i X_i^H \right\|_{S_q}.$$

Now use the noncommutative Khintchine inequality (noting that the two terms in the inequality's max are equal) to get

$$E \leq W_q \left\| \left(\sum_{i=1}^p X_i X_i^H X_i X_i^H \right)^{1/2} \right\|_{S_q}.$$

4

We can bound the Schatten q-norm by the spectral norm by paying a multiplicative penalty of $(pm)^{1/q}$, where pm is the maximum rank of the matrix sum. By the hypothesis, this penalty does not exceed $\sqrt{e}$.

$$E \le W_q \sqrt{e} \left\| \left(\sum_{i=1}^{p} X_i X_i^H X_i X_i^H \right)^{1/2} \right\|_2$$

$$\le W_q \sqrt{e} \left\| \sum_{i=1}^{p} X_i X_i^H X_i X_i^H \right\|_2^{1/2}$$

We note that the sum term is a quadratic form that can be expressed in terms of X and its block diagonalization, as follows:

$$E \le W_q \sqrt{e} \| X D_X^H D_X X^H \|_2^{1/2} \le W_q \sqrt{e} \| D_X X^H \|_2$$
$$\le W_q \sqrt{e} \| D_X \|_2 \| X \|_2$$
$$\le W_q \sqrt{e} \| X \|_{B,1} \| X \|_2,$$

where the last step used Lemma 1. Now replace $W_q \le 2^{-1/4} \sqrt{\pi q / e}$ to complete the proof. $\square$

The next lemma is adapted from [11].

Lemma 3. *Let H be a a Hermitian matrix with zero block diagonal. Then $\mathbb{E}_q \| RHR \|_2 \le 2 \mathbb{E}_q \| RHR' \|_2$, where $R' = S' \otimes I_m$ with S' denoting an independent realization of the random matrix S.*

Proof. We establish the result for $q = 1$ for simplicity. Denote $H_{i,j} = \Phi_i^H \Phi_j - 1_{\{i=j\}} I$ for $1 \le i, j \le p$. Further, denote $\widetilde{H}_{i,j}$ as the masking of the matrix H that preserves only the subblock $H_{i,j}$.

$$\mathbb{E} \| RHR \|_2 = \mathbb{E} \left\| \sum_{1 \le i < j \le p} \delta_i \delta_j (\widetilde{H}_{i,j} + \widetilde{H}_{j,i}) \right\|_2 .$$

Let η_i be i.i.d. Bernoulli random variables with parameter $1/2$. Here, we use Jensen's inequality on the new random variable $\eta = \{\eta_i\}_{1 \le i \le p}$: we define $M_{i,j}(\eta) = \eta_i(1 - \eta_j) + \eta_j(1 - \eta_i)$, and note that $\mathbb{E}_\eta M_{i,j}(\eta) = 1/2$ for all i, j. We also define the function

$$f(M_{i,j}(\eta)) = \mathbb{E}_\delta \left\| \sum_{1 \le i < j \le p} 2 \delta_i \delta_j M_{i,j}(\eta) (\widetilde{H}_{i,j} + \widetilde{H}_{j,i}) \right\|_2 .$$

Thus, by applying Jensen's inequality to $f(\cdot)$, we obtain

$$\mathbb{E} \| RHR \|_2 \le 2 \mathbb{E}_\eta \mathbb{E}_\delta \left\| \sum_{1 \le i < j \le p} [\eta_i(1 - \eta_j) + \eta_j(1 - \eta_i)] \delta_i \delta_j (\widetilde{H}_{i,j} + \widetilde{H}_{j,i}) \right\|_2 .$$

There is a 0-1 vector η^* for which the expression exceeds its expectation over η. Let $T = \{i : \eta_i^* = 1\}$.

$$\mathbb{E}\|RHR\|_2 \leq 2\mathbb{E}_\delta \left\|\sum_{i \in T, j \in T^C} \delta_i \delta_j (\widetilde{H}_{i,j} + \widetilde{H}_{j,i})\right\|_2 = 2\mathbb{E}_\delta \left\|\sum_{i \in T, j \in T^C} \delta_i \delta_j \widetilde{H}_{i,j}\right\|_2$$

$$\leq 2\mathbb{E}_\delta \left\|\sum_{i \in T, j \in T^C} \delta_i \delta_j' \widetilde{H}_{i,j}\right\|_2,$$

where $\{\delta_j'\}$ is an independent realization of the sequence $\{\delta_i\}$. The first equality follows from a standard identity for block counter-diagonal Hermitian matrices. Now the norm of a submatrix does not exceed the norm of the matrix, so we re-introduce the missing blocks to complete the argument:

$$\mathbb{E}\|RHR\|_2 \leq 2\mathbb{E}_\delta \left\|\sum_{1 \leq i,j \leq p, i \neq j} \delta_i \delta_j' \widetilde{H}_{i,j}\right\|_2 = 2\mathbb{E}_\delta \|RHR'\|_2.$$

$\square$

The next lemma is adapted from $[10, 12]$.

Lemma 4. *Let $B = [B_1 \ldots B_p]$ be a matrix with p column blocks, and suppose $q \geq 2\log(pm) \geq 2$. Then*

$$\mathbb{E}_q\|BR\|_2 \leq 3\sqrt{q}\mathbb{E}_q\|BR\|_{B,1} + \sqrt{\delta}\|B\|_2.$$

Proof. We denote $E := \mathbb{E}_q\|BR\|_2$, and have

$$E^2 = \mathbb{E}_{q/2}\|BRB^H\|_2 = \mathbb{E}_{q/2} \left\|\sum_{1 \leq i \leq p} \delta_i B_i B_i^H\right\|_2$$

$$\leq \mathbb{E}_{q/2} \left\|\sum_{1 \leq i \leq p} (\delta_i - \delta) B_i B_i^H\right\|_2 + \delta \left\|\sum_{1 \leq i \leq p} B_i B_i^H\right\|_2$$

Here we replace $\delta = \mathbb{E}\delta_i'$, with $\{\delta_i'\}$ denoting an independent copy of the sequence $\{\delta_i\}$. We then take the expectation out of the norm with Jensen's inequality to get

$$E^2 \leq \mathbb{E}_{q/2} \left\|\sum_{1 \leq i \leq p} (\delta_i - \delta_i') B_i B_i^H\right\|_2 + \delta\|BB^H\|_2.$$

We symmetrize the distribution by introducing a Rademacher sequence $\{\epsilon_i\}$, noticing that the expectation does not change due to the symmetry of the random variables $\delta_i - \delta'_i$:

$$E^2 \le \mathbb{E}_{q/2} \left\| \sum_{1 \le i \le p} \epsilon_i (\delta_i - \delta'_i) B_i B_i^H \right\|_2 + \delta \|B\|_2^2.$$

We apply the triangle inequality to separate δ_i and δ'_i, and by noticing that they have the same distribution, we obtain

$$E^2 \le 2\mathbb{E}_{q/2} \left\| \sum_{1 \le i \le p} \epsilon_i \delta_i B_i B_i^H \right\|_2 + \delta \|B\|_2^2.$$

Writing $\Omega = \{i : \delta_i = 1\}$, we see that

$$E^2 \le 2\mathbb{E}_{q/2,\Omega} \left(\mathbb{E}_{q/2,\epsilon} \left\| \sum_{i \in \Omega} \epsilon_i B_i B_i^H \right\|_2 \right) + \delta \|B\|_2^2.$$

Here we have split the expectation on the random variables Ω and $\{\epsilon_j\}$. Now we use Lemma 2 on the term in parentheses to get

$$E^2 \le 3\sqrt{q} \mathbb{E}_{q/2} (\|BR\|_{B,1} \|BR\|_2) + \delta \|B\|_2^2.$$

Using the Cauchy-Schwarz inequality, we get

$$E^2 \le 3\sqrt{q} \mathbb{E}_q \|BR\|_{B,1} \mathbb{E}_q \|BR\|_2 + \delta \|B\|_2^2.$$

This inequality takes the form $E^2 \le bE + c$. We bound E by the largest solution of this quadratic form:

$$E \le \frac{b + \sqrt{b^2 + 4c}}{2} \le b + \sqrt{c},$$

proving the lemma. $\qquad\square$

The last lemma is adapted from [12].

Lemma 5. *Let*

$$B = [B_1 \ B_2 \ \dots \ B_p] = \begin{bmatrix} B_{1,1} & B_{1,2} & \dots & B_{1.p} \\ B_{2,1} & B_{2,2} & \dots & B_{2.p} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p,1} & B_{p,2} & \dots & B_{p.p} \end{bmatrix}$$

be a block matrix, where each block $B_{i,j}$ has size $m \times m$. Assume $q \ge 2 \log p$. Then we have

$$\mathbb{E}_q \|RB\|_{B,1} \le 2^{1.5} \sqrt{q} \|B\|_{B,2} + \sqrt{\delta} \|B\|_{B,1}.$$

Proof. We begin by seeing that

$$E^2 := (\mathbb{E}_q \|RB\|_{B,1})^2 = \mathbb{E}_q \left(\max_{1 \le j \le p} \|RB_j\|_2^2 \right) = \mathbb{E}_{q/2} \left(\max_{1 \le j \le p} \sum_{i=1}^{p} \delta_i \|B_{i,j}\|_2^2 \right)$$

In the sequel, we abbreviate $t = q/2$ and $y_{i,j} = \|B_{i,j}\|_2^2$. We continue by using the same technique as in the proof of Lemma 4: we split a term for the mean value of the sequence $\{\delta_i\}$, then replace the term by $\mathbb{E}\delta_i'$ — an independent copy of the sequence, then exploit symmetrization by introducing a Rademacher sequence $\{\epsilon_i\}$, and then finish by merging the two terms due to their identical distributions:

$$E^2 \le \mathbb{E}_t \left(\max_{1 \le j \le p} \sum_{i=1}^{p} (\delta_i - \delta) y_{i,j} \right) + \delta \max_{1 \le j \le p} \sum_{i=1}^{p} \|B_{i,j}\|_2^2$$

$$\le \mathbb{E}_t \left(\max_{1 \le j \le p} \sum_{i=1}^{p} (\delta_i - \delta_i') y_{i,j} \right) + \delta \|B\|_{B,1}^2$$

$$= \mathbb{E}_t \left(\max_{1 \le j \le p} \sum_{i=1}^{p} \epsilon_i (\delta_i - \delta_i') y_{i,j} \right) + \delta \|B\|_{B,1}^2$$

$$\le 2\mathbb{E}_t \left(\max_{1 \le j \le p} \sum_{i=1}^{p} \epsilon_i \delta_i y_{i,j} \right) + \delta \|B\|_{B,1}^2.$$

Now we bound the maximum by the sum and separate the expectations on the two sequences:

$$E^2 \le 2\mathbb{E}_{t,\delta} \left(\sum_{j=1}^{p} \left(\mathbb{E}_{t,\epsilon} \sum_{i=1}^{p} \epsilon_i \delta_i y_{i,j} \right)^t \right)^{1/t} + \delta \|B\|_{B,1}^2.$$

For the inner term, we can use the scalar Khintchine inequality to obtain

$$E^2 \le 2C_t \mathbb{E}_{t,\delta} \left(\sum_{j=1}^{p} \left(\sum_{i=1}^{p} \delta_i y_{i,j}^2 \right)^{t/2} \right)^{1/t} + \delta \|B\|_{B,1}^2.$$

We continue by bounding the outer sum by the maximum term:

$$E^2 \le 2C_t p^{1/t} \mathbb{E}_{t,\delta} \left(\max_{1 \le j \le p} \left(\sum_{i=1}^{p} \delta_i y_{i,j}^2 \right)^{t/2} \right)^{1/t} + \delta \|B\|_{B,1}^2.$$

Since $t \ge \log p$, it holds that $p^{1/t} \le e$, which implies that $2C_t p^{1/t} \le 4\sqrt{t}$. We now use Hölder's inequality inside the sum term $\delta_i y_{i,j}^2 = y_{i,j} \cdot \delta_i y_{i,j}$ with $p = \infty$,

$q = 1$:

$$E^2 \leq 4\sqrt{t} \left(\max_{1 \leq i,j \leq p} y_{i,j} \right)^{1/2} \mathbb{E}_{t,\delta} \left(\max_{1 \leq j \leq p} \left(\sum_{i=1}^{p} \delta_i y_{i,j} \right)^{t/2} \right)^{1/t} + \delta \|B\|_{B,1}^2$$

$$\leq 4\sqrt{t} \left(\max_{1 \leq i,j \leq p} y_{i,j} \right)^{1/2} \mathbb{E}_{t,\delta} \left(\max_{1 \leq j \leq p} \left(\sum_{i=1}^{p} \delta_i y_{i,j} \right)^{t} \right)^{1/2t} + \delta \|B\|_{B,1}^2.$$

Now we recall that $t = q/2$ and $y_{i,j} = \|B_{i,j}\|_2^2$, to get

$$E^2 \leq 2^{1.5}\sqrt{q} \max_{1 \leq i,j \leq p} \|B_{i,j}\|_2 \mathbb{E}_{q/2} \left(\max_{1 \leq j \leq p} \left(\sum_{i=1}^{p} \delta_i \|B_{i,j}\|_2^2 \right)^{q/2} \right)^{1/q} + \delta \|B\|_{B,1}^2$$

$$\leq 2^{1.5}\sqrt{q} \max_{1 \leq i,j \leq p} \|B_{i,j}\|_2 \mathbb{E}_{q/2} \left(\max_{1 \leq j \leq p} \sum_{i=1}^{p} \delta_i \|B_{i,j}\|_2^2 \right)^{1/2} + \delta \|B\|_{B,1}^2$$

$$\leq 2^{1.5}\sqrt{q} \|B_{i,j}\|_{B,2} \mathbb{E}_q \max_{1 \leq j \leq p} \|RB_j\|_2 + \delta \|B\|_{B,1}^2$$

$$\leq 2^{1.5}\sqrt{q} \|B_{i,j}\|_{B,2} \mathbb{E}_q \|RB\|_{B,1} + \delta \|B\|_{B,1}^2$$

and notice that E has appeared on the right hand side. By following the same argument that ends the proof of Lemma 4, we comlpete the proof. $\qquad\square$

4 Proof of Theorem 1

We can now prove the main theorem. Split G into its diagonal blocks D (containing $X_i^H X_i - I$, $1 \leq i \leq p$) and off-diagonal blocks H (containing $X_i^H X_j$, $1 \leq i \neq j \leq p$) and apply Lemma 3:

$$\mathbb{E}_q \|RGR\|_2 \leq 2\mathbb{E}_q \|RHR'\|_2 + \mathbb{E}_q \|RDR\|_2.$$

To estimate the first term, we apply Lemma 4 twice: once for R, and once for R':

$$\mathbb{E}_q \|RHR'\|_2 \leq 3\sqrt{q}\mathbb{E}_q \|RHR'\|_{B,1} + \sqrt{\delta}\mathbb{E}_q \|R'H\|_2$$

$$\leq 3\sqrt{q}\mathbb{E}_q \|RHR'\|_{B,1} + 3\sqrt{\delta q}\mathbb{E}_q \|HR'\|_{B,1} + \delta \|H\|_2.$$

By applying Lemma 5 on the first term, we obtain

$$\mathbb{E}_q \|RHR'\|_2 \leq 3\sqrt{q} \left[2^{1.5}\sqrt{q}\mathbb{E}_q \|HR'\|_{B,2} + \sqrt{\delta}\mathbb{E}_q \|HR'\|_{B,1} \right] + 3\sqrt{\delta q}\mathbb{E}_q \|HR'\|_{B,1}$$

$$+ \delta \|H\|_2.$$

Since R and R' have the same distribution, we can collect terms to get

$$\mathbb{E}_q \|RHR\|_2 \leq 9q\mathbb{E}_q \|HR\|_{B,2} + 6\sqrt{\delta q}\mathbb{E}_q \|HR\|_{B,1} + \delta \|H\|_2 + \mathbb{E}_q \|RDR\|_2.$$

To bound $\|HR\|_{B,1}$, we denote $\Phi_{\{i\}^C} = [\Phi_1^H \ \ldots \ \Phi_{i-1}^H \ \Phi_{i+1}^H \ \ldots \ \Phi_p^H]^H$; we then have

$$\|HR\|_{B,1} \le \|H\|_{B,1} = \max_{1 \le i \le p} \|\Phi_i^H \Phi_{\{i\}^C}\|_2 \le \max_{1 \le i \le p} \|\Phi_i^H \Phi\|_2 \le \max_{1 \le i \le p} \|\Phi_i\|_2 \|\Phi\|_2.$$

Using the Geršgorin circle theorem [7], we can show that for each $1 \le i \le p$ $\|\Phi_i\|_2 \le \sqrt{1 + (m-1)\mu}$, so that $\|HR\|_{B,1} \le \sqrt{1 + (m-1)\mu}\|\Phi\|_2$. Now we use the facts $\|HR\|_{B,2} \le \mu_B$, $\|H\|_2 \le \|G\|_2 + \|D\|_2 = \|\Phi\|_2^2 + \|D\|_2$ and

$$\mathbb{E}_q \|RDR\|_2 \le \|D\|_2 = \max_{1 \le i \le p} \|\Phi_i^H \Phi_i - I\|_2 \le \mu_B$$

(from Lemma 1) to prove the theorem:

$$\begin{aligned}
\mathbb{E}_q \|RGR\|_2 &\le 9q\mu_B + 6\sqrt{\delta q(1 + (m-1)\mu)}\|\Phi\|_2 + \delta(\|\Phi\|_2^2 + \mu_B) + \mu_B \\
&\le 20\mu_B \log(pm) + 9\sqrt{\delta \log(pm)(1 + (m-1)\mu)}\|\Phi\|_2 + \delta\|\Phi\|_2^2.
\end{aligned}$$

5 Acknowledgement

References

[1] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk. Distributed compressed sensing. 2005. Available at http://www.dsp.rice.edu/cs.

[2] P. T. Boufounos, G. Kutyniok, and H. Rauhut. Sparse recovery from combined fusion frame measurements. Dec. 2009. Preprint.

[3] E. J. Candès. Compressive sampling. In *Proc. International Congress of Mathematicians*, volume 3, pages 1433–1452, Madrid, Spain, 2006.

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

[5] Y. C. Eldar, P. Kuppinger, and H. Bölcksei. Block-sparse signals: Uncertainty relations and efficient recovery. *IEEE Trans. Signal Processing*, 58(6):3042–3054, June 2010.

[6] M. Fornassier and H. Rauhut. Recovery algorithms for vector valued data with joint sparsity constrains. *SIAM Journal on Numerical Analysis*, 46(2):577–613, Feb. 2008.

[7] S. A. Geršgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk SSSR Ser. Fiz.-Mat.*, 6:749–754, 1931.

[8] M. Mishali and Y. Eldar. Blind multi-band signal reconstruction: Compressed sensing for analog signals. *IEEE Trans. Signal Proc.*, 57(3):993–1009, Mar. 2009.

[9] M. Rudelson. Random vectors in the isotropic position. *J. Functional Anal.*, 164(1):60–72, 1999.

[10] M. Rudelson and R. Vershynin. Sampling from large matrices: An approach through geometric functional analysis. *J. Assoc. Comput. Mach.*, 54(4):1–19, 2007.

[11] J. A. Tropp. Norms of random submatrices and sparse approximation. *C. R. Acad. Sci. Paris, Ser. I*, 346(23–24):1271–1274, 2008.

[12] J. A. Tropp. The random paving property for uniformly bounded matrices. *Studia Math.*, 185(1):67–82, 2008.